

Are Vanikoro languages really Austronesian?

A.L.L.2 ~ UKARG
Oxford, 3rd June 2006

Alexandre FRANÇOIS
LACITO — C.N.R.S — Paris

Good afternoon.

I'd like to talk to you today about the languages of Vanikoro, and more precisely about the delicate issue of their genealogy.

Presentation

- Darrell Tryon (1994: 635) says:
“The status of the six languages of Utupua and Vanikoro is in no doubt. These are clearly Austronesian languages by any standards.”
- Focusing on the languages of Vanikoro, I'd like to show there is room for doubt on that assumption:
 - either the languages of Vanikoro are *not* Austronesian
 - or if they indeed *are* Austronesian, this is anything but 'clear', because most of their AN characteristics have disappeared.
- This paper will not come up with definite conclusions: at least I'd like to **show *there is a problem to be solved by future research*** — a problem overlooked by former accounts.

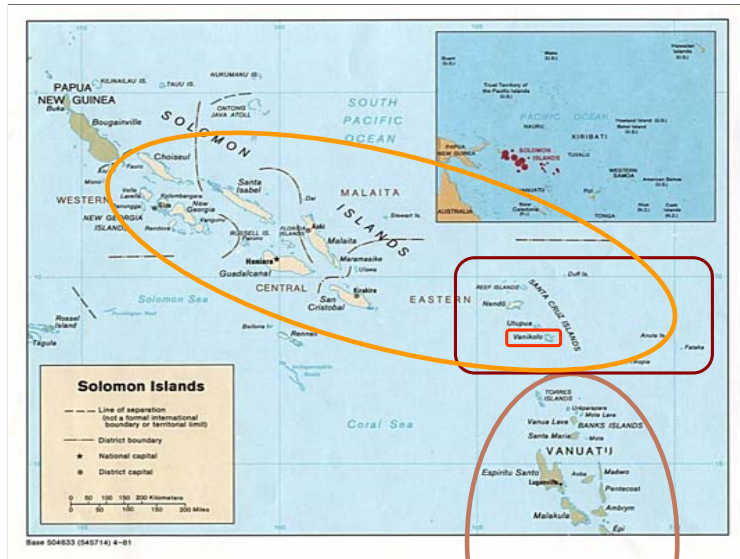
After doing some fieldwork in Vanikoro, Darrell Tryon did not seem to have any problem with the origin of these languages. He said:

“The status of the six languages of Utupua and Vanikoro is in no doubt. These are clearly Austronesian languages by any standards.”

What I propose today is to show that such a conclusion, at least for Vanikoro, is far from obvious.

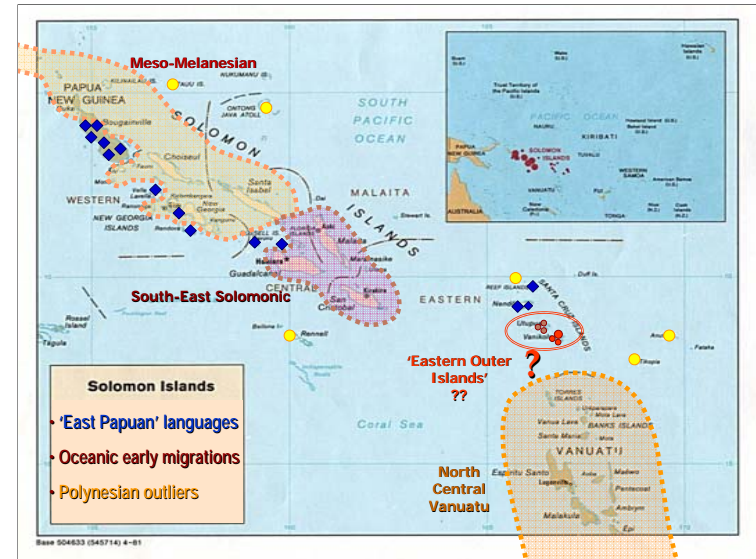
In my opinion, there is some possibility that these languages are in fact *not* Austronesian; and even if they were proven to be indeed Austronesian, this would be anything but 'clear', because most of their Austronesian characteristics have disappeared.

My paper will not necessarily come up with a definite conclusion; at least what I'd like to demonstrate today, is that **there is a problem** that deserves to be addressed by future research — a problem which had been overlooked until now.



This map shows the Solomon Islands and the northern part of Vanuatu, which is the area I am most familiar with.

At the crossroads between these two archipelagoes, in the southeastern tip of the Solomons, lies the small island group known as Santa Cruz islands; and Vanikoro is the second largest island of this group.



First of all, it may be useful to situate the three languages that are spoken on Vanikoro, in their linguistic context.

Indeed, the Santa Cruz area is one of the most complex regions of Island Melanesia, in terms of layers of language families and migrations.

First of all, this is the last place where Papuan languages are found, at least according to current classification (this includes Āiwoo, the focus of Åshild's paper).

Secondly, several dozens of modern languages reflect the early migrations of Oceanic populations: we have the Meso-Melanesian branch, the Southeast Solomonic, and finally the North Central Vanuatu.

Finally, several Polynesian outliers can be found in the area, including a community of Tikopian speakers on the island of Vanikoro itself.

In the Santa Cruz group, the only languages that are classified neither Papuan nor Polynesian, are the 3 of Vanikoro, plus 3 languages of Utupua, on which little is actually known.

These 6 languages have been identified as Oceanic by Darrell Tryon, but appeared sufficiently aberrant to deserve being considered as an autonomous, tentative branch below "Central Eastern Oceanic", under the name 'Eastern Outer Islands' – a putative group which is actually not defined by any shared innovations.

I will not discuss this grouping here as such, and will focus on the 3 languages of Vanikoro.

The three languages of Vanikoro

- The three languages of Vanikoro *my text corpus*
 - *Teanu* (Buma) 800 speakers 14300 words
 - *Lovono* (Vano) 4 speakers 1410 words
 - *Tanema* 3 speakers 1133 words

- Although these 3 languages share the same syntax, they show spectacular formal (lexical, morphological) diversity.

- Example

'We were in our garden, we've been planting crops.'

1exc:pl-*stay in garden* *our* 1exc:pl-*plant food*

TEA Pi-te ne sekele iupa, pi-wowo none.

LVN Nupe-lu ne amenonga iemitore, nupe-ngo nane.

TNM Tei-o ini vasangola akegamuto, ti-oa bauva.

Among these three languages, one is now dominant, Teanu or Buma. It has virtually overwhelmed the two other languages, Lovono and Tanema.

Teanu is the language on which most data are available, both in previous authors' publications, and in my own corpus.

A first important observation about these 3 languages, is that they share the same syntax, but show surprising diversity in their forms, both in the lexicon and most of their morphology.

I'll briefly illustrate this point with an example from my corpus:

If you consider a sentence like 'We were in our garden, we've been planting crops.'

you will get exactly parallel structures in the three languages, but quite diverse forms:

- in Teanu "*Pite ne sekele iupa, piwowo none*",
- in Lovono "*Nupelu ne amenonga iemitore, nupengoa nane*",
- and in Tanema "*Teio ini vasangola akegamuto, tioa bauva*" !

The three languages of Vanikoro

- Before even looking at the relations of Vanikoro with external languages, this spectacular formal diversity *within* Vanikoro is in itself a puzzle that deserves to be solved in the future.
- This situation recalls (at least impressionistically) the diversity found between Papuan languages – much more so than between any two close AN languages.
- For methodological reasons, I will essentially focus on one language in this demonstration: **Teanu**. It can be seen as representative of the 3 languages, at least regarding the difficulty of proving their AN status.

Before we even begin to look at the relationship between Vanikoro and external languages, I think this surprising formal diversity *within* Vanikoro is in itself a puzzle that warrants future consideration.

Incidentally, this diversity is not exactly typical of Austronesian languages, and somehow resembles more the sort of variety that is typically found among Papuan languages – if I dare make this impressionistic observation.

A corollary of this situation, is that it is probably safer in this paper to focus on only one language, **Teanu**.

In a way, it can be seen as representative of the 3 languages, at least with regard to the difficulty of proving their Austronesian status.

Searching for Oceanic cognates

- Observing the lexicon of Teanu:
 - Most lexical items seem disconnected from any known POc etyma.
 - Only 87 items (out of 1100) retained my attention as of possible Oceanic origin.
 - **87** is not negligible
(indeed these items are often cited as evidence that Teanu is Oceanic)
 - but this corresponds to only **7.9** percent of the lexicon, which is close to the margin of error (chance and borrowing)
 - Moreover:
Among these 87 items, phonological correspondences are irregular and often require *ad hoc* hypotheses.

Now let's delve into the data.

In comparison with more classical Oceanic languages, the first striking observation I made in Vanikoro, was that the vast majority of the lexicon seems disconnected from any known POc etyma.

Out of a Teanu lexicon of 1100 entries, I was personally able to recognise no more than **87** lexical items of possible Oceanic origin, the vast majority of them rather doubtful.

[note that I'm not counting here the many recent loanwords from Polynesian]

The first comment I'll make, is that **87** is clearly not negligible. And indeed, these words could easily be cited as evidence that Teanu is Oceanic.

Yet I have two methodological concerns with that figure.

- One is that, 87 items out of 1100 corresponds to only 7.9 percent of the lexicon. This is close to the margin of error, of chance similarities (which some linguists estimate at 5 to 6% for any two languages), or borrowing.
- And my second problem, as we shall see now, is that very few of these items are really secure. For most of them, cognacy judgments are more than problematic, with no way of defining regular phonological correspondences, other than resorting to *ad hoc* hypotheses.

Searching for Oceanic cognates

	POc	Kahua [S.E.Solomonic]	Teanu [Vanikoro]	Lo-Toga [N.C.Vanuatu]
<i>yesterday</i>	*nañorap	nanora	pepane	nənorə
<i>red</i>	*meRaŋ	memera		məmi
<i>snake</i>	*m ^w ata	m ^w a	abilo	ŋ ^w etə
<i>village</i>	*panua	hinua	kulumoe	βənūə
<i>bamboo</i>	*qauR	yau		ɔ
<i>die</i>	*mate	mae	-bu	met
<i>sleep</i>	*maturuR	mauru		mətür
<i>his/her hair</i>	*pulu-ña	huru-na	fiabasa ini	lū-nə
<i>sago palm</i>	*qatop	ao		et
<i>you two</i>	*kamiu rua	yamirua		kəmor

Let me take a random list of ten well-established Proto-Oceanic etyma.

First of all, I propose to take a witness sample of two *external* languages that are clearly Oceanic, and which are geographically closest to Vanikoro: one is Kahua, spoken on Makira to the west; the other is Lo-Toga, to the south.

What is striking in these data, is the relative conservativeness of these two languages. And even when sound change has occurred, this happened in regular patterns.

Now, let's come back to Vanikoro, and have a look at Teanu.

The first thing which strikes the observer is an overwhelming majority of forms that look totally disconnected from POc: compare *nañorap and pepane, *mwata and abilo, *panua and kulumoe, *mate and bu, and so on and so forth.

This situation of non-cognacy corresponds to more than **92** % of my lexical data, which I think is a very high rate for an Austronesian language.

Searching for Oceanic cognates

	POc	Kahua [S.E.Solomonic]	Teanu [Vanikoro]	Lo-Toga [N.C.Vanuatu]
<i>yesterday</i>	*nañorap	nanora	pepane	nənorə
<i>red</i>	*meRaq	memera	<i>moloe</i>	məmi
<i>snake</i>	*mʷata	mʷa	abilo	ŋʷetə
<i>village</i>	*panua	hinua	kulumœ	βənüə
<i>bamboo</i>	*qauR	ɣau	<i>okoro</i>	ɔ
<i>die</i>	*mate	mac	-bu	met
<i>sleep</i>	*maturuR	mauru	<i>-mokoïu</i>	mətür
<i>his/her hair</i>	*pulu-ña	huru-na	fiabasa ini	lü-nə
<i>sago palm</i>	*qatop	ao	<i>otovo</i>	et
<i>you two</i>	*kamiu rua	ɣamirua	<i>kela</i>	kəmør

Sometimes, some phonetic similarity can be suggested, but it is often doubtful.

- Thus 'you two' /kela/ is vaguely reminiscent of the reconstruction *kamiu-rua;
- /mokoïu/ 'sleep' at least shares a couple of phonemes with *maturuR
- /moloe/ 'red' *might* be a reflex of *meRaq
- /okoro/ 'bamboo' is a likely reflex for *qauR
- and likewise, /otovo/ 'sago palm' strongly recalls *qatop.

The trouble is, not all proposals are equally convincing, and it's difficult to come up with any satisfying phonological correspondences.

Searching for Oceanic cognates

	POc	Kahua [S.E.Solomonic]	Teanu [Vanikoro]	Lo-Toga [N.C.Vanuatu]
<i>yesterday</i>	*nañorap	nanora	pepane	nənorə
<i>red</i>	*meRaq	memera	<i>moloe</i>	məmi
<i>snake</i>	*mʷata	mʷa	abilo	ŋʷetə
<i>village</i>	*panua	hinua	kulumœ	βənüə
<i>bamboo</i>	*qauR	ɣau	<i>okoro</i>	ɔ
<i>die</i>	*mate	mac	-bu	met
<i>sleep</i>	*maturuR	mauru	<i>-mokoïu</i>	mətür
<i>his/her hair</i>	*pulu-ña	huru-na	fiabasa ini	lü-nə
<i>sago palm</i>	*qatop	ao	() <i>otovo</i>	et
<i>you two</i>	*kamiu rua	ɣamirua	<i>kela</i>	kəmør

- Thus, *R would be reflected sometimes as /l/ as in moloe, sometimes as /r/ as in okoro;
- the glottal stop would surface sometimes as /k/ and most of the times as zero (as in otovo);
- and so on and so forth...

And I'm saying nothing of vowel correspondences, which are by no means regular.

Searching for Oceanic cognates

menuko	'bird'	<	POc *manuk	
ma	'arm, hand'	<	POc *lima	?
luro	'coconut'	<	POc *niuR	?
aele	'leg, foot'	<	POc *qaqe	?
uo	'yam'	<	POc *qupi	?
basa	'head'	<	POc *b ^(w) atu	?
bwogo	'night'	<	POc *boŋi	?
foŋoro	'Canarium'	<	POc *[ka]ŋaRi	?
dijobe	'Kingfisher'	<	POc *sikon	?
iawo	'fire'	<	POc *api	?
-wene	'lie down'	<	POc *qenop	?
-punuo	'steal'	<	POc *panako	?
-leŋi	'hear, feel'	<	POc *roŋoR	?
-le	'go'	<	POc *lako	?

I won't have time to go into detail here.

But what I'd like to point out with this list, is that apart from a handful of items whose Oceanic origin is beyond doubt (like *menuko* from *manuk), most other lexical candidates are much more problematic, and would require adhoc explanations.

Searching for Oceanic morphosyntax

- The typological features of Vanikoro languages are generally reminiscent of other Oceanic languages
 - S-V-O order
 - Possessed-Possessor
 - Prepositions
 - [Noun-Adj-Dem]
 - Verbs take subject prefixes (Realis vs Irrealis)
 - Serial verb constructions
 - Verb 'say' grammaticalised as Complementiser
 - Three numbers (sing, dual, plural); no gender, no case
 - Contrast of (in)alienability + Food/drink classifiers
 - Geocentric system of space reference

Now let's have a look at the morphosyntax.

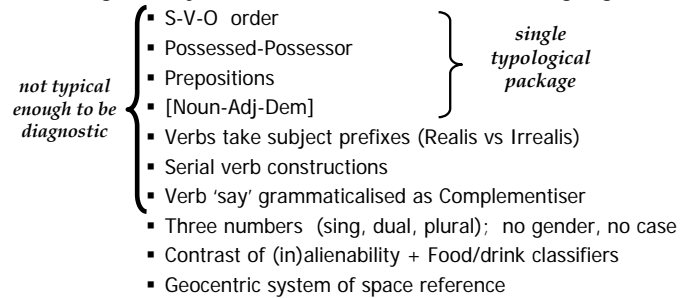
Generally speaking, it is true that the typological features of Vanikoro languages are essentially reminiscent of other Oceanic patterns.

- The order of constituents is S.V.O.; Possessed-Possessor; Prepositions; the order in the NP is *Noun-Adjective-Demonstrative*;
- Verbs take subject prefixes that get inflected for mood;
- We have extensive Serial verb constructions;
- The verb 'say' has grammaticalised as a Complementiser;
- there is a dual, but no gender and no case;
- the grammar of possession shows contrast of inalienability, as well as the existence of food and drink classifiers
- and finally Spatial reference makes use of the same geocentric system as is found elsewhere in Oceania.

This is a bit disturbing indeed. However, it is unclear to me whether these features constitute solid enough diagnostic evidence for genetic status.

Searching for Oceanic morphosyntax

- The typological features of Vanikoro languages are generally reminiscent of other Oceanic languages



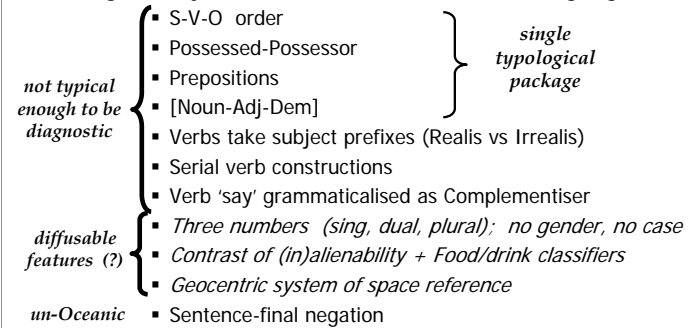
Indeed, several of these properties are in fact quite common typologically.

... Moreover, the first four properties are known to come generally bundled together as a single typological package.

And as far as I know, they are indeed attested in some of the East Papuan languages of the area.

Searching for Oceanic morphosyntax

- The typological features of Vanikoro languages are generally reminiscent of other Oceanic languages



Certain properties seem more promising.

However, these can also be seen as areal features that may be diffused and calqued from one language to another...

And finally, there is at least one typological feature that is rather un-Oceanic and more typically found in East Papuan languages: that is, the sentence-final position of the negation.

In summary, typological properties, in the case of Vanikoro languages, do not constitute clearcut evidence for our discussion.

This is not that surprising, since we know that typological features can easily change, be borrowed or be lost.

Syntactic structures generally constitute dubious evidence when it comes to genetic matters.

Searching for Oceanic morphology

■ Subject prefixes of Teanu

	SING	DUAL	PLUR
1incl		la-	li- le-
1excl	ni- ne-	ba-]	pi- pe-]
2	a- u-		
3	i-	la-	li- le-

■ Subject prefixes of POC (Lynch, Ross, Crowley 2002: 67)

	SING	DUAL	PLUR
1incl		*ta-	*ta-
1excl	*ku- ~ *au-	—	—
2	*mu- ~ *ko-	—	—
3	*a- ~ *i-	*ra-	*ra-

Much more useful should be the morphology.

Most surprisingly, I was unable to find any solid trace of the morphology reconstructed for Proto-Oceanic.

Let's have a look at personal markers first, more precisely subject prefixes.

No correspondence here between Teanu and POC is obvious.

What's more, the three Vanikoro languages show an unusual pattern of merger of 1excl with 2nd person, and 1incl with 3rd person; a pattern which is not attested in other Oceanic languages.

On the contrary, this sort of mergers between pronouns appears quite commonly among East Papuan languages – with the only problem that the combinations attested are different.

Searching for Oceanic morphology

■ Independent pronouns of Teanu

	SING	DUAL	PLUR
1incl		kia	kiapa
1excl	ene	keba	kupa
2	eo	kela	kaipa
3	ini	da	dapa

■ Independent pronouns of POC (Lynch, Ross, Crowley 2002: 67)

	SING	DUAL	PLUR
1incl		*kita rua	*kita
1excl	*[i]au	*ka[m]i rua	*ka[m]i
2	*[i]ko[e]	*kamiu rua	*kamiu
3	*ia	*ra rua	*[k]ira

The evidence is not much more convincing for independent pronouns (*ene, eo, ini...*).

Some similarities could be suggested, but once again they would rest upon no clear phonological correspondences.

Searching for Oceanic morphology

■ Possessive markers of Teanu

	SING	DUAL	PLUR
1incl		i-akia	i-akapa
1excl	enone	i-aba	i-upa
2	i-ono	i-amela	i-aipa
3	i-ape	i-ada	i-adapa

■ Possessive suffixes of POC (Lynch, Ross, Crowley 2002: 67)

	SING	DUAL	PLUR
1incl		(* <i>-da rua</i>)	* <i>-da</i>
1excl	* <i>-gu</i>	—	* <i>-ma[m]i</i>
2	* <i>-mu</i>	—	* <i>-m[i]u</i>
3	* <i>-na</i>	(* <i>-ra rua</i>)	* <i>-ra</i>

The situation is even worse for Possessive markers.

In fact, Teanu is the only Oceanic language I know of (if it is Oceanic!), which has lost absolutely all traces of POC possessive suffixes.

Even those languages which have changed their system, at least show traces of these suffixes in other parts of their morphology. Teanu does not.

Thus compare POC **gu, mu, na* with Teanu *enone, iono, iape*.

Searching for Oceanic morphology

■ Noun morphology

- **no** trace of article **na*
- **no** affixation on inalienable nouns (personal suffixes, **-qi, ...*)
- Food/Drink classifier ($\approx \emptyset$ -) does not reflect **ka-*
 - Drink classifier *me* (<POC **ma-*) in little use [borrowed?]

■ Verb morphology

- **no** trace of POC object suffixes
- **no** trace of transitivity **-i*
- **no** trace of applicative **-aki(n)*
- **no** valency-changing affixes (Causative **paka-*, Reciprocal **paRi-...*)
- **no** derivation { Verbs \leftrightarrow Nouns } (!)
- **no** use of reduplication (!)

- The evidence that we are dealing with an AN language is **slim**.

In fact, the same observations hold for the rest of Teanu morphology:

In the domain of nouns, we find

- no trace of article **na*
- no affixation on inalienable nouns...
- a Food and Drink classifier basically with the form /zero/, that does not reflect POC **ka*
- there is a second Drink classifier of the form /me/ (which recalls POC **ma* of course), but which appears to be in little use, and may well be borrowed from an Oceanic language.

Likewise, in the domain of verbs, we find

- no trace of POC object suffixes... [see slide]

In sum, there's very little evidence that we are dealing with an Austronesian language.

Possible explanations

- We have *roughly*
{ Oceanic structures, non-Oceanic forms }

- Possible explanations
 - language contact *or* language shift
between AN and Non-AN (=Papuan) language

 - language-internal change

Roughly speaking, what we have for Teanu is {Oceanic structures, but massively non-Oceanic forms}.

How can we account for this situation?

I can see two types of possible explanations:

- one would involve language contact, or language shift; that is, a kind of blend between Austronesian and non-Austronesian (in other words, Papuan) languages
- the second sort of hypothesis would entail *language-internal change*.

I'll review first the hypotheses based on language contact.

The language contact hypothesis

1. **Papuan language** influenced by Austronesian input?
 - would account for the massively non-AN/OC forms
 - requires less than **7.9 %** of lexicon (+ **1 %** of morphology) being borrowed from AN source
→ these are likely figures

 - *but* requires almost all syntactic patterns to be borrowed (calqued?) from AN source.
→ is this a likely scenario??

 - Would be a case of *metatypy* (Ross 1996):
a NAN language having its syntactic structures reshaped through contact with AN...

One possibility would suggest that Teanu is originally a Papuan language that was later influenced by Austronesian.

[incidentally, this scenario would be very similar to the one proposed by Stephen Wurm in the case of Äiwoo or Reef-Santa Cruz]

- this hypothesis would account for the massively non-AN/OC forms
- and it would require less than **8 %** of the lexicon (plus about 1 % of the morphology) being borrowed from an Austronesian source
... note that these are likely figures.
- But at the same time, this same hypothesis would require almost all syntactic patterns to be borrowed (or calqued?) from an Austronesian source.

Now, how likely would such a scenario be?

In fact, this kind of structural reshaping of a morphosyntax through language contact has already been described for other parts of Melanesia by Malcolm Ross, who calls the phenomenon *metatypy*.

The language contact hypothesis

2. Austronesian language influenced by Papuan input?

- would account for the AN/OC features
- *but* would require **92 %** of lexical borrowing (or change)
+ **99 %** of morphological borrowing (or change)
- **Problem** of the 'Papuan' side of the coin: the surrounding 'East Papuan' languages do not seem to qualify as likely donors.
[...but more information would be welcome]

The alternative possibility, would be that we are dealing with an Austronesian language that was heavily influenced by a Papuan input.

[interestingly, this scenario would parallel the way Peter Lincoln proposed to see the Reef-Santa Cruz languages, as opposed to Stephen Wurm]

- this would account for the presence of Austronesian features
- but it would require up to 92 % of lexical borrowing (or change)
as well as 99 % of morphological borrowing (or change)
... These are indeed high percentages.

• A further problem:

Ideally, such a hypothesis would require the identification of a Papuan language possibly at the source of that influence.

But the trouble is, none of the surrounding languages identified as Papuan really qualify as likely donors ... at least to the current extent of my knowledge.

A possible answer to this problem here would be to say that there is actually very little shared vocabulary, or shared morphosyntactic patterns, among the East Papuan languages *anyway*; so to find little commonalities with other Papuan languages is not really a strong argument against its Papuan status. As we know, Papuan languages have a much more ancient history in the region than Austronesian, which is one of the reasons why they are so heterogeneous.

The esoterogeny hypothesis

3. A 'purely' Austronesian language?

- We may be dealing with an essentially AN / OC language, which for some reason would have drastically renewed its lexicon and morphology, 'language internally'.
 - There would be no Papuan donor to look for.
 - Structures kept intact, 92% of forms renewed.
- One possible sociolinguistic explanation:
the phenomenon of *esoterogeny* (Thurston 1989, Ross 1996).
= The tendency for speakers to keep their language increasingly distinct from neighbouring languages, especially through lexical innovations.

Now, there may still be one last possibility:

Namely, that Teanu is in fact no more than a 'purely' Oceanic language (so to speak), which for some reason would have drastically renewed its lexicon and morphology, on a language-internal basis.

In this case, the principal source of lexical change would not have to be sought in an external, Papuan language, but essentially within the proper resources of the language itself.

According to this hypothesis, we would have a language whose grammatical structures have been essentially kept intact through time, but in which more than 92% of the forms would have undergone innovation.

Such radical language-internal evolution would be spectacular, but still possible.

I can think of one phenomenon that is known to trigger language-internal lexical replacement: this is what William Thurston called "*esoterogeny*".

– that is, *the tendency for speakers to keep their language increasingly distinct from neighbouring languages, especially through lexical innovations.*

There must be other possible explanations I'm not aware of, and I will welcome your suggestions.

Of course there is still a lot to say, but at least I think I have exposed the main elements of my Vanikoro puzzle.

Conclusion

- Something spectacular has happened...
 1. NAN influenced by AN:
spectacular case of metatypy (calqued morphosyntax)
 2. AN influenced by NAN:
spectacular case of extensive lexical borrowing
 3. AN language with mainly 'internal' evolution:
spectacular case of esoterogeny and lexical replacement
- No hypothesis really prevails.
- But the *esoterogeny hypothesis* seems favoured by:
 - existence of about 80 words of probable/possible Oceanic origin
 - seeming absence of any likely Papuan lexifier (?)
 - strong tendency for esoterogeny and lexical replacement exists already *between* the three languages of Vanikoro

So, to conclude, I would say that, whatever solution should ultimately be retained, obviously *something spectacular has happened* in the history of Teanu:

- If this is a Papuan language influenced by Austronesian, then we are faced with an extreme case of metatypy, or *morphosyntactic borrowing*.
- If it is an Austronesian language influenced by Papuan, then this illustrates extensive *lexical borrowing*.
- Finally, if Teanu is simply an Austronesian lgg having gone through language-internal evolution, then we are confronted with an impressive case of *esoterogeny* and *lexical replacement*.

None of these three hypotheses really prevails.

Yet if I really had to choose right now, I believe that the third scenario would be the most likely, as is suggested by the following points:

- that is, the presence of a few dozen words of possible Oceanic origin
- the seeming absence of any likely Papuan lexifier
- finally, the strong tendency for esoterogeny and lexical innovation that can be observed among the 3 languages of Vanikoro.

Conclusion

**The languages of Vanikoro
may be of Austronesian origin**

(?)

So my final word would be

that the languages of Vanikoro
may be of Austronesian origin ...

... or maybe *not* !

Thank you.