Social ecology and language history in the northern Vanuatu linkage

A tale of divergence and convergence*

Alexandre François
Langues et Civilisations à Tradition Orale (Lacito-CNRS) & Australian National University

This study describes and explains the paradox of related languages in contact that show signs of both linguistic divergence and convergence. Seventeen distinct languages are spoken in the northernmost islands of Vanuatu. These closely related Oceanic languages have evolved from an earlier dialect network, by progressive diversification. Innovations affecting word forms — mostly sound change and lexical replacement — have usually spread only short distances across the network; their accumulation over time has resulted in linguistic fragmentation, as each spatially-anchored community developed its own distinctive vocabulary. However, while languages follow a strong tendency to diverge in the form of their words, they also exhibit a high degree of isomorphism in their linguistic structures, and in the organization of their grammars and lexicons. This structural homogeneity, typically manifested by the perfect translatability of constructions across languages, reflects the traditions of mutual contact and multilingualism which these small communities have followed throughout their history. While word forms are perceived as emblematic of place and diffuse to smaller social circles, linguistic structures are left free to diffuse across much broader networks. Ultimately, the effects of divergence and convergence are the end result, over time, of these two distinct forms of horizontal diffusion.

Keywords: language change, language contact, areal diffusion, structural borrowing, lexical replacement, linguistic diversification, historical dialectology, Melanesia

1. Sister languages in contact: Divergence or convergence?

The present study examines the processes of divergence and convergence which characterize the historical evolution of a group of closely related languages spoken

Journal of Historical Linguistics 1:2 (2011), **175–246**. DOI 10.1075/jhl.1.2.03fra
ISSN 2210–2116 / E-ISSN 2210-2124

in northern Vanuatu. I will argue that these two conflicting tendencies operate on different areas of these languages' systems: divergence typically affects word FORMS, whereas convergence mostly takes place between linguistic STRUCTURES. I will propose to associate this twofold evolution with two sociolinguistic factors playing a major role in the region: social group emblematicity, and language contact.

1.1 Divergence and convergence between languages

The reconstruction of historical relationships between languages potentially involves two types of processes: processes of divergence, whereby languages become less similar over time; and processes of convergence, whereby languages become more and more similar.

Linguistic divergence and convergence are typically observed in different contexts, involving different sets of languages. For example, within a given language family, the process whereby a proto-language split into a number of descendant languages is essentially conceived of as a process of divergence (Bossong 2009). If genetically related languages have long been separated from one another, their comparison will highlight the various processes which have made them different, since each descendant went through its own linguistic innovations.

Conversely, research on language contact demonstrates how multilingual speakers tend to make their languages more and more similar over time. The story then told is one of convergence. The majority of studies in language contact involve languages which are genetically diverse; this genetic diversity is often emphasized as a typical ingredient in the observation of contact phenomena, e.g., in this definition of linguistic areas: "A linguistic area is generally taken to be a geographically delimited area including languages *from two or more language families*, sharing significant traits" (Aikhenvald & Dixon 2001: 11, my emphasis).

Admittedly, this perspective offers the most extreme view of convergence phenomena, as languages can be seen narrowing the gap that used to separate them. However, contact phenomena are not limited to that particular configuration; it is equally legitimate to observe contact between languages which are closely related. This configuration underlies a number of studies dealing with various parts of the world: see, *inter alia*, the papers in Braunmüller & House (2009), and also Trudgill (1986) and Labov (1991) for English dialects, Chappell (2001) for Sinitic languages, Enfield (2001) for Eastern Mon Khmer, Toulmin (2009) for a subgroup of Indo-Aryan, and Donohue (2002) for Skou languages; within the Oceanic subgroup, see in particular Geraghty (1983) for Fijian dialects, and Ross (1988, in prep.) for Western Oceanic.

The reason why this case — related languages in contact — is special lies therefore not in its rarity, but rather in the potential contradiction it entails. For a group of languages to be genetically related means that they have diverged from a common ancestor, in order to become distinct languages over time. But processes of language contact between them typically entail the diffusion of linguistic traits, and would be expected to result in some form of linguistic convergence. This raises the question of how these two contradictory tendencies interact, and what outcome results from their interference.

1.2 Divergence of forms, convergence of structures

This study proposes to observe a group of 17 languages spoken by the 9,400 inhabitants of the Banks and Torres Islands, in the northernmost parts of Vanuatu. These languages are all related: first, they all belong to the Oceanic subgroup of the Austronesian family; second, they form a "linkage" (Ross 1988), insofar as they descend from an earlier dialect network (see Section 3 below). Yet despite their common origin, three millennia of linguistic change have made these languages mutually unintelligible.

While both processes of divergence and convergence are at play in the history of these languages, they have affected different dimensions of their systems. In a nutshell, processes of divergence have been particularly intense with respect to linguistic FORMS, whether this concerns sound change or lexical replacement. By contrast, languages show much more areal homogeneity as far as their STRUCTURES are concerned — whether in the syntactic and combinatorial properties of words, the semantic organisation of the lexicon, or phraseological strategies. (The contrast I draw between the forms of words and their structures, or structural properties, will be presented in more detail in Section 5.1.)

This particular alchemy has brought about an intricate situation, where languages show relative diversity in the forms of their words, but also a strong isomorphism in the structural organization of meaning. Table 1 illustrates the configuration that typically obtains, by showing the 17 translations of a sentence chosen randomly (meaning 'They don't know our language very well yet').

The typical situation is that a single line of interlinear glossing corresponds to a variety of different word forms. These northern Vanuatu languages illustrate in a somewhat extreme way a tendency which has already been observed in other parts of Melanesia (Haudricourt 1961, Thurston 1987, 1989, Dutton 1995, Ross 2001, François 2009b). Pawley (2006: 216) describes in which respects the most innovative (or "aberrant", Grace 1990) of the Oceanic languages typically differ from their more conservative sisters: "In these areas, most of the aberrant languages (…) are

Table 1. Divergence of forms, convergence of structures in the 17 languages of northern Vanuatu: an example sentence

Hiw	sisə	tati	jəjməᵍʟen	wuᵍʟɔɣ	kʷe	i	nə	məŋa	ta
Ltg	nihə	tat	lolmərɛn	ʉrβɛ	kʷɛ	e	nə	βəɣəβaɣə	məʈə
Lhi	kɛj	tɛtnɛ	ɣlal	ɣalsɛ	kʷɒ		n-	βap	munɣɛn
Lyp	k͡iɛ̂j	tɛ	ɣilal	ʧøjmat	ʧɛk͡pʷɛ		n-	βaβap	ŋ͡mʷɔn͡iɛ̂n
Vlw	ⁿgɪj	ɛt	ɣilal	ɣalsi	tɛⁿg͡bʷɛ		n-	ɣatɣat	njɔnɣɪn
Mtp	kɪj	ɛt	ɪɣlal	ɣalsi	k͡pʷɛtɛ		nɔ-	hɔhɔlɛ	nɔnɔnɣɪn
Lmg	tær	ɪ	ɣɒlɒl	ʔørmaʔ	ʔæ.kiʔis		n	tɛktɛk	mʊɣʊt
Vra	ⁿdir	ɪʔ	lamai	ɛntɛɣ	ʔɪn		ɪn	tɪktɪk	muⁿdɪ
Vrs	nɪr	ɣɪtɪ	ɣilal	warɛɣ	tɛn		ɔ	k͡pʷak͡pʷ	namøɣynɪn
Msn	nɪr	ɛtɛ	lɪlɪ	maŋtɛ	βɪs		ɔ	ɣatlɛ	mɔɣɔnin
Mta	nira	ɣate	ɣlala	mantaɣ	tuk͡pʷe		o	βaβae	naŋ͡mʷunina
Num	nir	βitis	ɣil	liŋliŋi	mi		u	luwluw	namɣin
Drg	nɪr	sɔwsɛ	βrɪɣɪl	taβul	tɛ		na	lŋa	ɣɪn
Kro	nɪr	tɪ	rɔŋ	taβul	wʊs.mɛlɛ		ɔ	βalβalaw	namɪɣɪn
Olr	nɪj	tɪ	rɔŋ	βɪlɪː	wʊs.mɛlɛ			ususraː	mʊʧ
Lkn	ɣɪː	atɪ	rɔŋ	kɛrɛ	aβʊh.malɛ			ɛlŋa	nɣɪʧ
Mrl	kɛr	ti	βalɣɛ̂ar	mɪnmɪn	tɪkʷɪtɛ̂a		nɵ	liŋɪ	ɣɛ̂an
	3pl	NOT.YET$_1$	know	properly	NOT.YET$_2$	[OBL]	ART	speech	POSS:1INCL.PL
	'They don't know our language very well yet.'								

atypical in lexicon and phonology, but are not markedly atypical in respect of (…) morphology and syntax."

Other parts of the world show similar configurations, that is, cases where languages in contact show almost perfect intertranslatability, in spite of the difference between their forms. Consider, for example, the way Gumperz & Wilson (1971) describe the relation between Kannada, Marathi and Urdu as spoken in the Indian village of Kupwar:

> [S]entence-by-sentence comparison of natural conversation texts in all three main local varieties reveals an extraordinary degree of translatability from one local utterance to the other. (…) The sentences (…) are lexically distinct in almost every respect, yet they have identical grammatical categories and identical constituent structures.
> (Gumperz & Wilson 1971: 154)

However, the Indian situation described there involves languages which are genetically diverse (Dravidian vs. Indo-Aryan), in such a way that lexical forms were

historically different to begin with; the spectacular process in the languages of Kupwar is the amount of structural CONVERGENCE which they have gone through. But in the northern Vanuatu case which I propose to describe here, the languages share a common ancestor, and have maintained contact ever since their geographical dispersal. In other words, what needs to be explained here is a paradox, whereby a cluster of languages have simultaneously gone through processes of DIVERGENCE and of CONVERGENCE.

1.3 The present study

Until recently, northern Vanuatu languages had been only known through grammar sketches by early missionaries (Codrington 1885), and wordlists (Tryon 1976). The present study, based on data collected by the author since 1997, proposes to describe this region in some detail. The complex situation that prevails here, and which was illustrated in Table 1, will be explained in light of the social and anthropological conditions of these languages' historical developmen, that is, their LANGUAGE ECOLOGY (Haugen 1972; Mühlhäusler 1996).

I will argue that the syntactic and semantic isomorphism observable across the northern Vanuatu area results from sustained relations of language contact and multilingualism. On the other hand, the heterogeneity of word forms will be shown to correlate with social perceptions of what normally constitutes a distinct community. I will suggest that this interplay of centrifugal and centripetal forces in the social ecology of these groups — distinct village communities in constant interaction — provides the key to the complex linguistic situation observable today, in which languages show both divergence and convergence.

The present article will unfold as follows. Section 2 will present the 17 language communities of northern Vanuatu in their social setting, showing how it involves both some degree of separation and various forms of contact. Section 3 will situate these languages genetically, and establish that they share a single ancestor.

Sections 4 and 5 will present in some detail the facts which constitute the paradox under discussion. Section 4 will discuss the various processes leading to linguistic diversification in northern Vanuatu, especially sound change and lexical replacement. Section 5, in turn, will cite a number of cases where languages show parallel syntactic or semantic structures despite the diversity of their word forms.

The final discussion, in Section 6, will attempt to explain the social underpinnings that may account for this paradox, whereby languages can simultaneously show signs of divergence and convergence. Finally, I will suggest that these two apparently contradictory tendencies observed on the macro scale are really the superficial result of deeper processes that combine social biases and cognitive factors. Both divergence and convergence are rooted in a single, universal process of

micro-diffusion across individuals — or "linguistic epidemiology" (Enfield 2003) — as linguistic innovations emerge locally and spread, by imitation, across individuals in contact. However, at least in this part of the world, this process of micro-diffusion does not target the same social circles, depending on the nature of the innovations that spread. Those changes affecting the FORM of words (sound change, lexical replacement), take on an emblematic role as markers of place, and tend to diffuse through smaller social circles, typically targeting the village or local community to which each speaker relates more or less consciously. Conversely, those innovations that affect the STRUCTURAL ORGANISATION of meaning (morphosyntax, semantic categorisations, phraseology) escape this biased anchoring to any particular social group, and are left free to diffuse across entire social networks. Eventually, as the two types of change accumulate over time, the end result is an epiphenomenon whereby these related languages in contact appear to have simultaneously "diverged" in their word forms and "converged" in their structures. The article concludes with a summary in Section 7.

2. The social ecology of northern Vanuatu languages: Distinct communities in constant interaction

2.1 A fragmented linguistic landscape

2.1.1 *The situation of the Banks and Torres languages*

The modern country of Vanuatu, known as the New Hebrides until its independence in 1980, hosts the world's highest density of languages per capita (Crowley 2000). A total of 106 vernacular languages have been recorded there (Tryon 1976, Lynch & Crowley 2001) for a current population of 234,000 inhabitants, which makes for an average number of about 2200 speakers per language. All the languages of Vanuatu belong to the Oceanic subgroup of Austronesian (see Section 3 below).

The present study will focus on the northernmost area of Vanuatu, namely the groups of islands known as the Torres Islands and the Banks Islands. Figure 1 situates northern Vanuatu within Island Melanesia and the Pacific.

Throughout this paper, the label "northern Vanuatu", or the word "archipelago", will consistently refer to the Torres and Banks groups. This relatively small archipelago includes ten inhabited islands, and has a total land surface of 882 km², most of which is uninhabited. Its population of about 9,400, distributed across approximately 50 villages, today speaks 17 different languages.[1] Figure 2 provides a map of the area, together with the approximate number of speakers for each language.

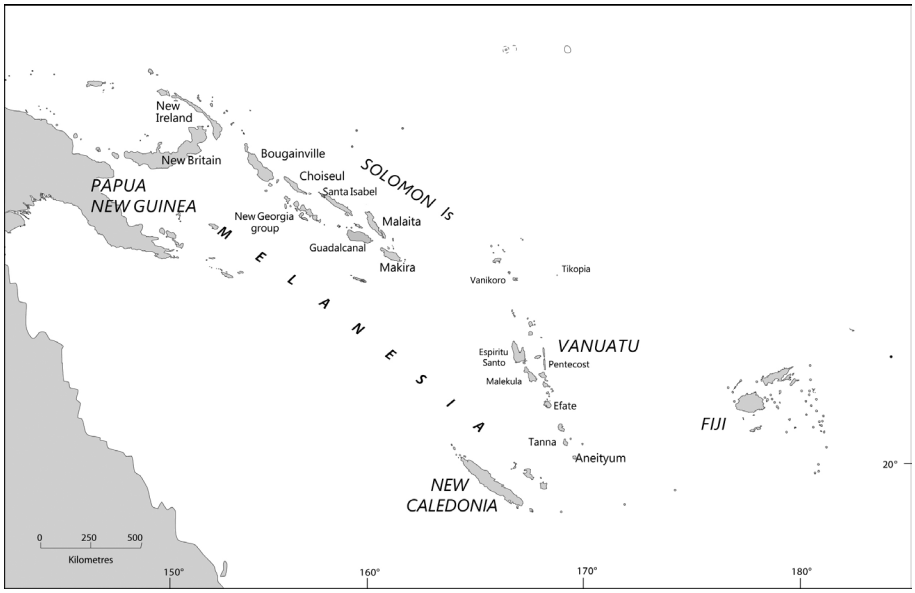

Figure 1. The situation of northern Vanuatu in the Pacific

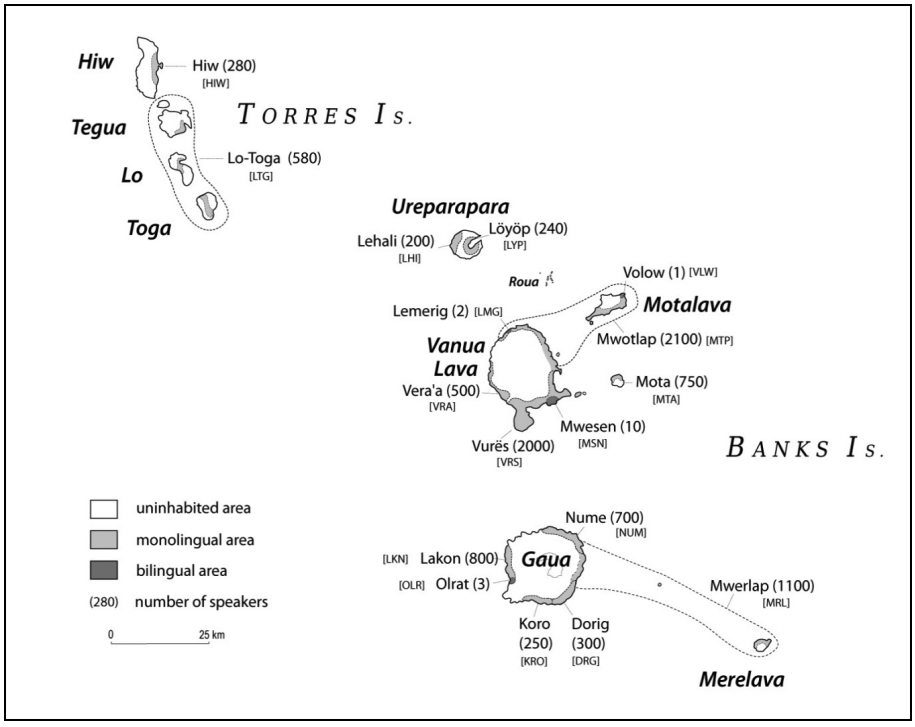

Figure 2. The 17 languages of northern Vanuatu (Torres and Banks islands)

Among these 17 speech traditions or "communalects" (Pawley & Sayaba 1971, Geraghty 1983), some show enough similarity that they might be considered, under some definitions, dialects of a single language (e.g., Mwesen and Vurës, Mwotlap and Volow, or Koro and Olrat). However, they exhibit enough differences, in their phonologies and/or their lexicons, for speakers to consider them different languages, with limited mutual intelligibility. Because the present study focuses precisely on linguistic diversification, I will follow the social representation of these speech traditions, and treat them as separate languages. Conversely, when the difference between two varieties is tenuous and of little (linguistic and social) significance, I group them as dialects of a single language. Thus Lo and Toga are subsumed under "Lo-Toga", Vurē and Qätärew under "Lakon", Viar and Olrat under "Olrat", and Merig and Mwerlap under "Mwerlap".

2.1.2 *Demography and vitality*

As the figures on the map suggest, these 17 languages today enjoy varying degrees of vitality. Four of them are clearly moribund, no longer transmitted to younger generations, and only remembered by a handful of elderly individuals. Mwesen, with about 10 speakers, is giving way to the locally dominant language Vurës. The three speakers of Olrat are shifting to Lakon, the major language on the west coast of Gaua. Lemerig had five speakers in 2003, and has now gone down to just two individuals, who live in an area settled by Mwotlap speakers.

Volow can now be considered extinct, since its last fluent speaker (Wanhan) died in 1986. The reason why I still include it here among the 17 languages of the area is because it is still remembered today, at least passively, by its "last hearer" (Evans 2010: 209), Wanhan's son Wolman. A valuable recording of the late Wanhan, which the anthropologist Bernard Vienne carried out in 1969 and later handed over to me, has been extremely useful in the task of reconstructing, with Wolman's help, what spoken Volow used to be like.

These four moribund tongues (Mwesen, Olrat, Lemerig, and Volow) are remnants of languages which seem to have always been spoken by small inland communities, and probably never more than one or two hundred speakers each in their heyday.

The other Torres and Banks languages tend to number in the hundreds, from 200 for Lehali to 2,000 for Vurës. At the upper end of the spectrum, Mwotlap is currently thriving, with as many as 2,100 speakers of all ages, of whom 1,650 live on the island of Motalava. All these languages, except the four moribund ones, are still transmitted to children. In this regard, despite figures which may seem low by world standards, they are safe from immediate endangerment (see Crowley 1995, 2000).

In this part of island Melanesia, it appears that just a few hundred speakers may form a viable speech community. The average number of speakers per language is 550, or 720 if one removes the four moribund languages from the count. These figures are much lower than Vanuatu's national average of 2,200 speakers per language, which is already a world record in terms of language density per capita (see Section 2.1.1 above). This alone shows the extreme linguistic fragmentation which is characteristic of the northern Vanuatu area.

While 17 languages may seem a high number, historical evidence shows that it was even higher a few generations ago. François (forthcoming), compiling accounts from both missionary documents (Codrington 1885, 1891) and from the oral tradition, lists a total of 35 distinct communalects (languages or dialects) for the Banks and Torres islands, which were alive around the middle of the 19th century. Various social processes took place in the decades around 1900, including epidemics, labour trade, and island-internal migrations, which resulted in a drastic demographic downturn (Vienne 1984: 400), as well as in the reshaping of communities and social groups. In just a couple of generations, the consequence was the dramatic erosion of linguistic diversity in the area (François forthcoming), which went from 35 communalects to 17 today, and down to 13 in the near future.

In terms of social setting, the languages which have survived until today are often spoken in only one or two villages. The maximum for one language is six villages, and the average is three (50 villages for 17 languages). The only way to reach another village on the same island is by walking. These villages and hamlets are scattered across the archipelago. Although some hamlets were occasionally built on higher parts of mountainous islands, especially on Vanua Lava or Gaua, many of them have now been abandoned in favour of coastal villages, where fishing and inland resources are more easily combined. Today, island-internal migrations and the accretion of smaller hamlets have occasionally led to the formation of larger and more densely-populated villages, with up to 1500 people living together in one contiguous area, in the west of Motalava.

Some places in the archipelago show particularly high linguistic density. For example, the village of Jōlap (500 inhabitants) on the west coast of Gaua counts as many as three distinct languages, which are all heard in public every day: Lakon (itself a composite of two dialects Vurē and Qätärew), Olrat (and its variety Viar) and Dorig. These three languages have been spoken and transmitted by families in this village for more than three generations. To these one may add Koro, Mwerlap and Nume; these are the languages of immigrants (spouses, new settlers) and are spoken within certain households. There is enough multilingualism (see Section 2.2.1 below) in the small population of this village for everybody to understand, and occasionally speak, each others' language(s).

2.2 A homogeneous social network

Despite their relative social separation, the populations of the Torres and Banks Islands have always formed a coherent network. All throughout their history, the communities of this small archipelago have practiced various forms of social interactions and linguistic contact, whether in their traditional way of life (see Section 2.2.1 below) or in modern times (see Section 2.2.2 below). These constant interactions are an essential component of the social ecology of languages in the region.

2.2.1 *Sustained relations of contact*

The cultural and linguistic similarities found today across the Torres and Banks area are partly due to the common origin of its population (see Section 3.1 below), but also to the constant relations of contact which local communities have always entertained with each other. Even though contact took place regularly with islands further south as well, the Torres and Banks Islands together evidently form a particularly cohesive social network.

The distances between the two ends of the archipelago, say, between the Torres Islands and Merelava, were too great to allow for any sustained contact between all pairs of islands of the group. However, it is generally easy for people to travel from their own language community to the closest ones. The distance between two language communities on the same island is, on average, about two or three hours of strenuous walking. It did not take much longer to cross the sea to the next island when sailing canoes were still in use. In sum, while contact was minimal between distant islands, it has always been possible for a given community to keep in contact with the few communities that surrounded them.

Cultural contact between groups was not only possible, it was socially encouraged. Until today, the relationship prevailing between communities has always been one of a mutual balance of power, with little socio-economic dominance of one group over others (Vienne 1972). While the oral traditions of the Torres Islands still recall warfare and conflict between villages, this is hardly the case across the various islands of the Banks group. Even if they also occasionally allude to past conflicts, their stories tend rather to revolve around alliances between villages and chiefs, and also around the trade of valuable items, including women. The deity Qat, who lived on Vanua Lava, sailed northwest as far as Vava, the ancient name of the Torres Islands; in return for one of his pigs, he bought the Night, as well as a rooster, and a piece of obsidian (François 2003b: 2). Another myth tells how a variety of yam was first discovered in the Torres Islands, and disseminated to populations further south. As for the prestigious white yam, it is claimed by Motalava people to have been first cultivated on their island before a child took

it to Mota. Another story tells how the chief of Ureparapara appeased a conflict with the chief of Maewo, to the south of Merelava, by offering him his daughter (François 2003b: 26). Finally, the Merelava people who have colonised the eastern coast of Gaua (see Figure 2) like to relate how their ancestors once taught the native people of this island how to marry and beget offspring.

All these stories, as well as many others, portray an archipelago where communities have constantly exchanged goods and women. A similar picture is drawn by archaeologists and anthropologists. Huffman (1996) shows how the economy of the Torres and Banks communities, as well as much of the Vanuatu archipelago further south, was dependent on the trade of goods. In particular, the grade-taking ceremonies involved the consumption and exchange of vast amounts of shell-money, which used to be essentially produced on the now deserted atoll of Roua (see Figure 2), in northern Banks (Vienne 1972, Bedford & Spriggs 2008: 96). These trade contacts had important implications in terms of traditional multilingualism, especially among men: the men of a given community used to know and practise the languages of the neighboring groups with whom they had contact. Crucially, this form of multilingualism was always EGALITARIAN (Haudricourt 1961: 9), in the sense that no group or language was considered dominant or more prestigious than others. Instead, neighboring groups typically entertained relations of mutual bilingualism. This egalitarian multilingualism is still observed today, with the caveat that much interisland communication today tends to be carried out in Bislama (François forthcoming), the English-lexifier pidgin used as a lingua franca throughout modern Vanuatu (Tryon & Charpentier 2004).

Major public events, such as grade-taking ceremonies for new initiates, regularly took the form of large gatherings during which people of remote communities would meet and communicate together. These intercommunity events were also an opportunity for unions to form.

Vienne (1984: 232–240) documents in detail the patterns of interisland marriages in the northern islands of the Banks group (Vanua Lava, Ureparapara, Motalava, Mota — see Figure 1). What he shows is a strong habit of exogamy, with more than 20 percent of unions being interisland marriages. Figure 3 is borrowed from Vienne (1984: 240). The outer circle identifies the islands; the dots in the inner circle correspond to individual villages. His data, collected in the 1960's and 1970's, documented 455 individual marriages. Each line on the figure symbolizes the marriage of at least one woman from the source village with a man from the target village. These unions resulted in the relocation of the wife in 61 percent of cases, of the husband in 39 percent.

Due to this exogamous tendency, it is very common for someone to have relatives in various neighboring communities of the archipelago, and occasionally visit them. Because residence tends to be patrilocal, every village will count a number

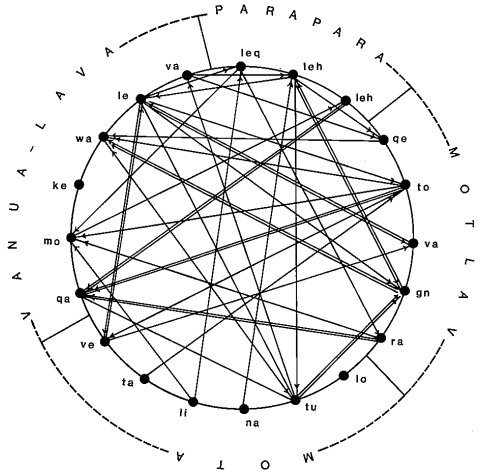

Figure 3. Interisland marriages in the northern Banks Islands (Vienne 1984)

of immigrant women originating from various parts of the archipelago, who have established their residence in their spouses' villages. Incidentally, Vienne's observations regarding the northern Banks Islands can be safely extrapolated to the entire Torres-Banks archipelago. For example, the small population of Lo Island (Torres) currently includes immigrants from Hiw, Tegua, Toga, Motalava, Mota, Vanua Lava and Gaua.

These social conditions have important consequences in terms of language contact. Each of these adult women quickly becomes bilingual, being able to switch between her own mother tongue and the language of her new village. The linguistic environment of her children will be shaped by this form of language contact, whether directly or indirectly. Should each parent speak his or her own language at home, then the children are raised bilingually. Even when the mother speaks the language of the father at home, the variety her children will hear will often be tainted by her own linguistic background. Together with other causes of language contact, interisland marriage plays an important role in the development of linguistic convergence (see Section 6.1 below).

2.2.2 *Language contact in the modern world*

Traditional patterns of social exchange, whether through trade or interisland marriages, have been extended in the modern era. The christianization of the archipelago by Anglican missionaries has put an end to any remaining hostilities that may have existed. Missionaries also encouraged the merger of inland hamlets into larger coastal villages, thereby fostering *de facto* language contact. Major religious gatherings, such as Christmas celebrations, have replaced earlier traditional ceremonies, yet still play a similar role in sustaining inter-island exchanges and marriages.

Akin to the effects of christianization are the consequences of modern education as it is organised in these low-density communities (François forthcoming). Because some small islands lack schools beyond the primary level, parents send their children to boarding schools on another island. Relations of power between the vernaculars are not so much affected by the school language itself (French or English) as by the demographic imbalance between students of various linguistic backgrounds. For example, when young speakers of smaller languages are schooled in the Arep secondary school of east Vanua Lava, they tend to learn Mwotlap; even though this language is not indigenous to this island, it is becoming increasingly dominant in the region. This present organization of schools favors, at best, multilingualism at a young age, and in the long term, endangerment of the weakest language varieties.

Finally, the modern era has brought about some technical progress made in the domains of transportation (engine boats, interisland flights) and communication (satellite phones have quickly spread since the end of 2008). The combined effect of all these modern conditions is to considerably increase the frequency of contact between linguistic communities, even beyond what was already the case under traditional circumstances.

3. Shared ancestry and *in situ* diversification

In many parts of the world, modern linguistic diversity within a given area results from population movements, whereby genetically diverse languages have come in contact during their history. This, however, is not the case in northern Vanuatu. Both archaeological and linguistic evidence show that the local populations share a common ancestry, and have formed a relatively unified network for about three millennia, until today (see Section 2.2).

Before European discovery, there is little evidence that the islands of northern Vanuatu had any sustained contact with any exogenous group. While occasional encounters may have happened with the Temotu people in the north, or with Polynesian voyagers (Clark 1994), these were so scarce that they left virtually no linguistic trace other than a handful of lexical loanwords. The only sustained contact that can be established for northern Vanuatu remained internal to the Torres and Banks Islands. Contact with islands further south, via a Merelava–Maewo connection, also existed (Tryon 1996, François 2011), but still involved closely related languages.

This is an important point for our discussion, since this entails that the present-day linguistic diversity does not result from external causes (historical migrations or contact between genetically diverse languages), but from a historical

process of in situ diversification from what was initially a linguistically unified social network.

3.1 A common ancestor

Unlike the Solomon Islands to the northwest (see Figure 1), the Vanuatu archipelago shows no archaeological trace of any human settlement prior to 3200 BP. Non-Austronesian ("Papuan") populations had settled much earlier (more than 40,000 years ago) in New Guinea and part of the Solomons chain, a region labelled "Near Oceania" by Pawley & Green (1973). As for the rest of the Pacific islands, known as "Remote Oceania", their colonization only began after a population of navigators, speakers of Austronesian languages, migrated from Island Southeast Asia into Melanesia, between 4,000 and 3,000 BP.

This is when populations living off the coast of New Guinea, in the Bismarck archipelago close to New Britain I (see Figure 1), developed a civilization known as "Lapita" (Spriggs 1995, 1997, Kirch 1997, 2010). The attributes of this culture complex included a certain style of pottery artifacts ("Lapita pottery") as well as elaborate seafaring techniques (Pawley & Ross 1995: 48). Between 3200 and 2900 BP, Lapita navigators managed to cross large ocean gaps, and rapidly settled the then unpopulated archipelagoes of Remote Oceania: the Santa Cruz islands (east of the Solomon chain), Vanuatu, New Caledonia, Fiji and West Polynesia. Although the detail of their itinerary and settlement patterns remains to be fully understood, it is a clear fact that the modern populations of northern Vanuatu descend from these early Lapita settlers (Shing et al. 2007, Bedford & Spriggs 2008).

As far as languages are concerned, all indicators suggest that the language spoken by the bearers of the Lapita civilisation who left the Bismarck archipelago essentially corresponded to what is reconstructed as "Proto-Oceanic", the common ancestor of all the indigenous languages of Remote Oceania (Ross 1988, Pawley & Ross 1995, Pawley 2003, 2007b, Lynch, Ross & Crowley 2002). The language spoken by Lapita navigators had little time to evolve significantly during the swift migrations that led them to settle across island Melanesia: "The rapid spread of Lapita from the Bismarcks to West Polynesia between 3200 and 2900 BP had a linguistic correlate. The speech of the Lapita colonists in the different island groups must have been relatively homogeneous, little differentiated from Proto Oceanic" (Pawley 2007a: 11). This observation has important consequences for the linguistic history of the Oceanic family, whose tree must largely be reconstructed as a "rake" rather than a series of neatly nested subgroups (Pawley 1999) — see Figure 4 below. In other words, under the Proto-Oceanic node of the family tree, no significantly different intermediate stage was able to form between the time when the populations left the Bismarck archipelago, and the time when they reached

the islands of Vanuatu. The only common ancestor which is shared by northern Vanuatu languages, strictly speaking, is thus Proto-Oceanic itself.

3.2 Pre-dispersal ancestry vs. post-dispersal linkages

There are in fact two different ways in which the languages of northern Vanuatu can be said to share a common "ancestry". For one thing, they all descend from Proto-Oceanic, the language spoken by the initial settlers prior to their demographic dispersal. But another form of shared development corresponds to the linguistic network in which the populations continued to participate even after their dispersal.

After the initial settlement of northern Vanuatu, archaeological evidence (Pawley 2007a; Bedford & Spriggs 2008) shows that the newly settled populations maintained a unified social network across vast archipelagoes for several generations. Such social links are evidenced, for example, by the early presence of obsidian from the Banks Islands in Fiji (Ambrose 1976; Best 1987, in Bedford & Spriggs 2008: 98), and later on Efate (Bedford & Spriggs 2008: 112). The examination of ceramic evidence allows archaeologists to identify some post-Lapita features which are shared across certain clusters of islands within Vanuatu: "[T]here is regional diversification in ceramic traditions right across Vanuatu soon after Lapita (…) these traditions can be divided into a number of regions: southern, central, northern, and far northern Vanuatu (*Banks and Torres Islands*)" (Bedford & Spriggs 2008: 107; my emphasis).

The linguistic correlates of this ancient social unity are a number of post-dispersal innovations, which historically diffused across a geographically spread-out network of then mutually intelligible dialects. Because some of these innovations did not always spread to all members of the dialect network, they do not define proper SUBGROUPS, but LINKAGES (Ross 1988, 1997, Pawley 1999). A linkage designates "a group of communalects which have arisen by dialect differentiation" (Ross 1988: 8).

Some isoglosses encompass the whole of Vanuatu and New Caledonia, and define the SOUTHERN OCEANIC LINKAGE (Lynch 2000). These include certain morphological innovations (Lynch 2000, Lynch & Ozanne-Rivierre 2001), as well as the irregular loss of the phoneme *R in a vast number of lexical items (Geraghty 1990). The detailed history of *R loss in Vanuatu (François 2011) provides crucial evidence to demonstrate that Vanuatu populations participated in dense communication networks even after the dispersal of the first settlers.

Several innovations also spread across the northern two thirds of Vanuatu, from Hiw to Efate. These formed the basis for what Clark (1985, 2009) has identified as the North and Central Vanuatu linkage, to which he assigned 95 of

Vanuatu's 106 languages, including the 17 Torres-Banks languages of the present study. The putative "proto-language" for this linkage is referred to as PNCV "Proto North Central Vanuatu" (Clark 2009). Finally, those linguistic innovations which spread across the Torres-Banks linkage may be conveniently subsumed under the label PTB "Proto Torres-Banks". The tree-like diagram in Figure 4 summarizes the mutual relations between these linkages; here I adopt the convention in Ross (1988) of representing linkages with a double line, to distinguish them from proper subgroups.

Figure 4. Proto Oceanic has split into various dialect networks, which have given rise to loosely-bound linkages of various sizes

There are thus two ways in which a group of closely related languages may share a "common ancestor", as conceived of in a family-tree model. First, they fully inherit the features of their PRE-DISPERSAL SHARED ANCESTOR, i.e. the language which was spoken by the initial settlers as they dispersed across territory, in this case, Proto-Oceanic. But should the dispersed communities continue to sustain a social and linguistic network for some time, then they will also witness a number of innovations which will diffuse to part or all of the network. Over time, the accumulation of such innovations defines a linkage, i.e., a cluster of adjacent dialects sharing a common post-dispersal history. Such a linkage will in part have the properties of a SUBGROUP (e.g., some innovations shared throughout the network), but also properties that make it necessary to assign it a distinct status (e.g., uneven distribution of innovations that allow for conservative pockets; intersecting isoglosses; one language may belong simultaneously to several overlapping linkages). One property of the evolution of dialect networks is that those innovations which are exclusively shared across several dialects do not reflect vertical inheritance from an earlier (pre-dispersal) ancestor, but rather post-dispersal events of horizontal diffusion across dialects in contact (Geraghty 1983, Pawley 1999, Garrett 2006, Bossong 2009, Heggarty et al. 2010, Babel et al. forthcoming).

To summarize, from the strict point of view of phylogenetics, the only single ancestor uncontroversially shared by the modern languages of northern Vanuatu is Proto-Oceanic. All other lower-level clusterings are linkages, i.e., post-dispersal dialect networks which formed areas of diffusion. This dialect-network model explains how linguistic innovations managed to spread across social groups of varying scope during the historical development of the northern Vanuatu languages.

3.3 The changing scope of linguistics innovations

An important point for our discussion is that the spread of the various linguistic innovations differed in geographical scope. Some covered a broad territory, such as the ones which are shared across the North-Central Vanuatu linkage. By definition, these innovations did not trigger language fragmentation on a local scale, because they were adopted equally by all (or nearly all) adjacent members of the relevant dialect network. Other innovations, however, only spread across very limited territory, for example, a single island. Such local innovations result in greater linguistic diversity.

The varying distribution of innovations partially corresponds to successive stages in the history of the region. What started as a far-flung network of small voyaging communities maintaining some form of social and linguistic unity slowly crumbled into smaller social networks on a reduced scale. As local populations grew and developed economic self-sufficiency, they relied less and less upon their distant relationships. This breakup of social networks has linguistic correlates, in bringing about divergence among once unified languages: "After the first phase of colonisation, the archaeological and linguistic record indicates that in the Southern Melanesian archipelagos, a sequence of demographic and cultural changes occurred which led to weakening or loss of communication between distant sister communities. (…) *Most linguistic innovations spread only short distances*, and the speech traditions of distant communities diverged" (Pawley 2007a: 21; my emphasis).

In sum, the diversity found among the modern languages of northern Vanuatu, as illustrated in Table 1 (Section 1.2), is neither due to any preexisting genetic diversity, nor to external input, but reflects three millennia of *in situ* diversification. Section 4 will now illustrate some of the linguistic forms taken by these processes of fragmentation.

4. Processes of language diversification in northern Vanuatu

4.1 Addressing the issue of language diversification

The 17 languages spoken in the Torres and Banks Islands are thus related in two ways. On the one hand, they are all genetically related, all being descendants of the same ancestor. On the other hand, they have always entertained close relations of contact and multilingualism. These two factors could be expected to result in some form of linguistic homogeneity, or at least preserved mutual intelligibility. But this is not what can be observed today. How did it come about that the languages are so

divergent in their phonologies and lexicons? Did their history follow universals of language change? Or do we need to resort to region-specific principles of language change, which would only apply to this part of the world? Are these languages so chaotic that they make it impossible to apply the comparative method?

In fact, far from challenging the validity of the comparative method, the data presented here will strongly confirm its relevance in attempting to reconstruct the history of these languages. The careful identification of regular sound correspondences and structural changes, whether in the phonology, the morphology or the lexicon, will prove methodologically crucial in identifying the precise history of phonemes, morphemes and words in the area. Regular correspondences (cf. François 2005a) form the background against which it will become possible to detect irregular developments that are limited to a few adjacent languages. This will allow us to compare languages in their historical development, and track the path taken by each innovation separately.

Linguistic diversification, as I will show in this section, took here essentially three forms: REGULAR SOUND CHANGE (Section 4.2), IRREGULAR SOUND CHANGE (Section 4.3), and LEXICAL REPLACEMENT (Section 4.4). These processes raise two different questions, of different natures. One question is whether these changes are typologically unusual, and require explanations specific to this group of languages; my answer to this will be negative: in themselves, these three types of change follow universal trends of language evolution, and do not require any local explanation. This contrasts with another question, discussed in the final section of this paper (Section 6), namely, how can one explain the sheer *density* of these innovations, and their propensity to stabilize to small communities? This density, I will eventually suggest, mirrors the social fabric of the archipelago, and the accepted size of social communities in this particular part of the world.

But before I comment on the issue of their social correlates, I first propose to describe here the innovations themselves, and discuss which linguistic processes can explain them.

4.2 Language diversification through regular sound change

The initial linguistic unity that can be reconstructed for the early times of settlement gave way to a network of dialects which diverged more and more over the three millennia of their coexistence. One of the forces of such evolution was sound change. This cover term encompasses two distinct types of change:

– REGULAR sound change: changes occurring in the phonological system, and affecting the whole lexicon of a given communalect. The unit of change is the phoneme.

– Irregular sound change: changes specific to some individual lexical items, independent of regular sound change in the system as a whole. The unit of change is the word.

These two forms of sound change are independent of each other, and are not ordered with respect to one another. In the course of their history, individual lexical items are affected both by regular (phoneme-based) and by irregular (word-based) sound change. The first of these two types is the object of the present section, Section 4.2; the second one will be described in Section 4.3.

Just like any other innovation, both types of sound change can be described as having emerged locally, in the speech of a few individuals, and diffused gradually to their peers across a network of idiolects, following processes of the kind described by Labov (1963, 1994, 2001). As we will see, sound changes in northern Vanuatu often stabilized to a small area, usually smaller than the whole Torres-Banks archipelago, e.g., just one island, and sometimes just one village.

4.2.1 *Linguistic diversity at the phonological level*

The linguistic fragmentation due to regular sound change is most conspicuous in the domain of vowels. Proto-Oceanic, the common ancestor of these languages, is reconstructed with five vowels: /i e a o u/. This simple inventory is still preserved in Mota, a conservative language of the Banks Islands. However, all other languages of the area have increased their inventory of vowel phonemes via a historical process of metaphony, akin to umlaut, correlated with stress (François 2005a). A former trochaic foot *CV_1CV_2 regularly yielded a closed syllable CV′C, in which V′ is a new vowel quality resulting from the coloring of V_1 (stressed) under the influence of V_2 (unstressed), e.g., Pncv *ⁿdámu 'yam' > Vrs ⁿdœm.

Crucially, the patterns of vowel change (*CV_1CV_2 > CV′C) are highly regular within each language, but they show strong variation from one language to another. Thus in Koro, a stressed *a followed by an unstressed *u is systematically reflected as a diphthong /ɛ͡a/: *ⁿ*damu* > ⁿ*dɛ͡am* 'yam'. The same sequence *a(C)u regularly yielded /a/ in Nume (ⁿ*dam*) or Lakon (*ʧam*), a long /aː/ in Dorig (ⁿ*daːm*); /ɑ̟/ in Lehali (ⁿ*dɑ̟ŋ*), /ɔ/ in Vera'a (ⁿ*dɔm*) or Mwesen (*nɔm*), /ɐ/ in Mwerlap (ⁿ*dɐm*), /œ/ in Lemerig (*tœm*), /ø/ in Löyöp (ⁿ*døm*), /ɛ/ in Mwotlap and Volow, /ɪ/ in Olrat (*ʧɪm*). As this pervasive phenomenon of metaphony applied regularly to all possible combinations of vowels, it resulted in the phonemicization of new vowels, in different ways in each language. Table 2 (adapted from François 2005a: 445) lists the vowel inventories of the 17 modern languages of the Banks and Torres Islands. The rightmost column indicates the total number of vowel phonemes in each language, including short monophthongs, diphthongs and long vowels.

Table 2. The 17 languages of northern Vanuatu and their vowel inventories

lgg	name	vowel inventory	nb V
Hiw	*Hiw*	i ɪ ɛ a ə ɵ ɔ o ʉ	9
Ltg	*Lo-Toga*	i e ɛ a ə ɔ o ʉ + i͡ɛ ɪ͡ɛ i͡a o͡ə o͡ɔ	8+5
Lhi	*Lehali*	i ɪ ɛ æ a ə ɒ̟ ɔ ʊ u	10
Lyp	*Löyöp*	i ɪ ɛ æ a œ ø y ɔ ʊ + ɪ͡ɛ	10+1
Vlw	*Volow*	i ɪ ɛ a ɔ ʊ u	7
Mtp	*Mwotlap*	i ɪ ɛ a ɔ ʊ u	7
Lmg	*Lemerig*	i ɪ ɛ æ a ɒ̟ œ ø ɔ ʊ u	11
Vra	*Vera'a*	i ɪ ɛ a ɔ ʊ u	7
Vrs	*Vurës*	i ɪ ɛ a œ ø y ɔ ʊ + i͡a	9+1
Msn	*Mwesen*	i ɪ ɛ a ɔ ʊ u	7
Mta	*Mota*	i e a o u	5
Num	*Nume*	i ɪ ɛ a ɔ ʊ u	7
Drg	*Dorig*	i ɪ ɛ a ɔ ʊ u + aː	7+1
Kro	*Koro*	i ɪ ɛ a ɔ ʊ u + ɛ͡a	7+1
Olr	*Olrat*	i ɪ ɛ a ɔ ʊ u + iː ɪː ɛː aː ɔː ʊː uː	7×2
Lkn	*Lakon*	i ɪ ɛ æ a ɔ ʊ u + iː ɪː ɛː æː aː ɔː ʊː uː	8×2
Mrl	*Mwerlap*	i ɪ ɛ a ɞ ɵ ʉ ɔ ʊ+ ɛ͡a ɔ͡ɞ ɵ͡ʊ	9+3

Modern vowel inventories are thus diverse, ranging from a conservative 5-vowel system (Mota) to innovative sets of 12 (Mwerlap) or 13 (Lo-Toga). If short and long vowels are counted as distinct phonemes, then the highest numbers are found in Olrat (14) and Lakon (16). These systems also vary in their internal logic, some being perfectly symmetrical (e.g., Mwesen, Lemerig), others less so (see Figure 5).

Only in very few cases can one group two adjacent languages together (e.g., Mwotlap with Volow, or Dorig with Koro), and suggest their vowel systems may

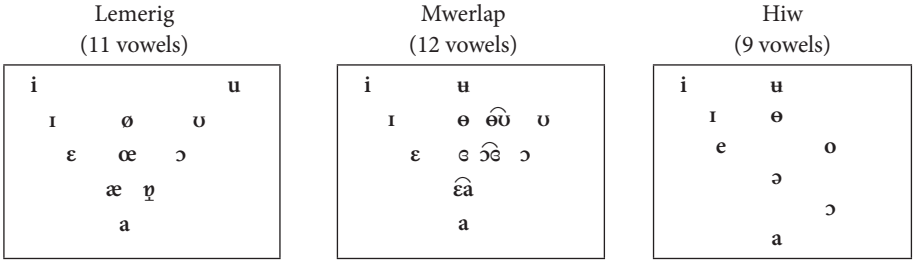

Figure 5. Modern vowel systems are more or less symmetrical

have followed a single history. But everywhere else, it appears that each local community has modified its vowel system in a different way than its immediate neighbors.

The phonological systems of these modern languages also show some differences in their consonant inventories. Thus, some languages have prenasalized stops (/ᵐb/, /ⁿd/, /ᵑg/); others do not. While most languages have rounded labialvelars (/k͡pʷ/ or /ᵑᵐg͡bʷ/, /ŋ͡mʷ/), four languages have rounded velars (/kʷ/, /ŋʷ/) instead. Certain segments, such as /ʧ/, /h/ or /ʔ/, are found in some languages, but not in their immediate neighbors. Some languages have two liquids /r/ and /l/, others just /l/; the only liquid of Hiw is a prestopped velar lateral /ᵍʟ/ [g͡ʟ], unheard elsewhere (François 2010a).

Languages also differ regarding their phonological constraints and phonotactic properties. Some license labialvelar stops syllable-finally (e.g., Vurës /k͡pʷak͡pʷ/ 'speech' in Table 1), while others do not. Some license tautosyllabic consonant clusters, while others systematically avoid them. Even though both Hiw and Dorig share the same syllabic template (CCVC), it can be shown that constraints of sonority are operational in the clusters of Hiw, but not in Dorig (François 2010a).

Even intonation can differ across languages. Whereas most languages express polar questions with a rising pitch, two sets of languages (the two Torres languages, plus those of Gaua) provide these questions with a falling contour.

4.2.2 *Linguistic fragmentation through regular sound change*

These regular phonological changes have sometimes obscured the relations between word forms in neighboring languages, even when they are ultimately cognate.

Although this is not the typical case (see Section 4.4), some lexical items show a high rate of retention, in the sense that (nearly) all the languages of the Torres and Banks Islands reflect the same original etymon. This etymon may be a retention from their common ancestor (Proto-Oceanic), or the result of post-dispersal lexical innovations which spread across dialects during the early times of linguistic unity (Pncv, Ptb).

While all the modern words may be cognate, their surface forms have often become different enough to hamper mutual intelligibility. Table 3 provides a small selection of such cognate sets. Regular phonological correspondences (François 2005a) make it possible to reconstruct a possible protoform for each cognate set (the last row of Table 3), at the level of "Proto Torres-Banks" (Ptb, see Section 3.2).

A proportion of the divergent forms can be explained simply by the regular changes outlined in Section 4.2.1. Thus, the variety of modern forms for 'song', 'stroke', and 'coconut crab' result from the application of regular patterns of sound change to the three reconstructable etyma, Ptb **asi*, **saraβi*, and **ⁿdaeru*

Table 3. Language differentiation through sound change: A few cognate sets from northern Vanuatu[a]

	'song'	'stroke'	'coconut crab'	'crayfish'	'breathe'	'1incl:Du'
Hiw	ɔs	sə^{g}ʟɔβ	*ɣʉtʉt*	^{g}ʟɔɣ	*mənawə*	tɵ^{g}ʟɵ
Lo-Toga	ɛh	hərɛβ	*ɣəhʉh*	rɔɣ	mək^{w}hɛ	ʈor
Lehali	n-æh	hɪjæp	n-^{n}dɪj	n-jʊ	mɔksæ	ɣɪnjɔ
Löyöp	n-ɪs	sɪjɪp	*n-^{n}døɣi͡ɛj*	n-i͡ɛj	ŋ͡m^{w}ɔŋsɪ	i͡ɛ^{n}dʊ
Volow	n-ɪh	hɪjɪɸ	n-^{n}dij	n-ɪj	mɔŋhɪɣ	^{n}dʊ(jʊ)
Mwotlap	n-ɛh	hɛjɛp	na-^{n}dij	n-ɪj	ŋ͡m^{w}ʊkhɛɣ	^{n}dʊ(jʊ)
Lemerig	n-ɛs	sœrœɸ	tɪr	n-ær	mœpsɪ	ɣætru
Vera'a	nɛs	saraβ	^{n}diɪr	nirɪ	mɔmsɛ	ɣi^{n}duʊ
Vurës	œs	sœrœβ	^{n}dɪr	ʊr	mɛmsɛɣ	^{n}dʊrʊk
Mwesen	ɛs	sɛrɛɸ	nɪr	ʊr	mɔpsɛ	ninrʊ
Mota	as	saraɸ	naer	ura	ŋ͡m^{w}aɸsaɣ	nara
Nume	wɛs	saraɸ	^{n}dɪr	wʊr	mamsɛŋ	^{n}duru
Dorig	aːs	sraːβ	ŋ^{n}dɪr	ʊr	maː^{m}bsɪɣ	^{n}daːr
Koro	ɛ͡as	sɛrɛ͡aβ	^{n}dɪr	rɛ͡aŋ	mɛmsɛ͡aɣ	^{n}duru
Olrat	nɪs	ʔ	ʧɪj	nurɪŋ	mɪpsaː	ʧʊrʊ
Lakon	^{ʔ}æh^{ʔ}æh	hæræβ	ʧɪː	uræŋ	mahpæɣ	wʊʧʊ
Mwerlap	n-ɛs	sɛrɛp	*nɵɣɵtɵt͡ɛak*	n-ɵ͡ʊr	mɔm	^{n}dʊrʊ
ProtoTB	*asi	*saraβi	*^{n}daeru	*uraŋ	*ma^{m}busaɣi	*ɣi^{n}darua

[a] The prefix /n/ before nouns is a separable article. Forms in italics are not cognate with the rest of the list.

respectively. Aside from a couple of exceptions, the way each language treats the proto-consonants (*s, *r, *^{n}d) or sequences of proto-vowels (e.g., *áCi) is very regular.

Most of the changes illustrated here are in line with typologically common trends in sound change, whether with respect to individual segments (e.g., *s > /h/; *t > /ʧ/, or *t > /ʔ/), or processes like umlaut (e.g., *áCi > /ɛC/). Other instances of sound change are perhaps less common (e.g., *r > /j/, *r > /^{g}ʟ/; *^{m}b > /ŋ/, *^{m}b > /k/; *^{n}d > /ʧ/, *^{n}d > /ʈ/; *áCi > /ɔC/, etc.), but they are still within the limits of the explicable. Overall, what is more conspicuous here in these data is not so much the sound changes *per se* as their sheer density in such a limited territory.

The impression of fragmentation is even stronger if the data are plotted onto a map. Indeed, even those languages which share the same sound change (e.g., *s > /h/, *r > /j/, or *áCi > /æC/), instead of forming a coherent cluster are geographically scattered. Figure 6 illustrates this with the reflexes of **asi* 'song' (or with the article, **na asi*).

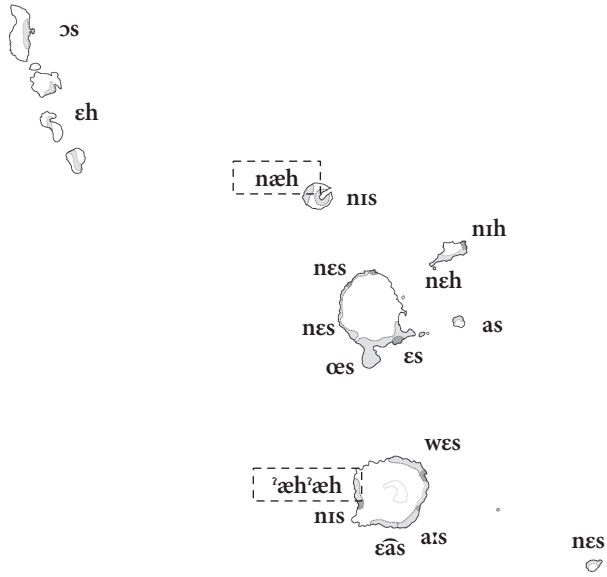

Figure 6. Lexical differentiation through regular sound change: Reflexes of the noun **asi* 'song'

When a change is shared across neighboring languages, it may result from the spread of a single innovation across adjacent dialects on the chain (e.g., *s > /h/ in Mwotlap and Volow). However, it is often the case that a given change is restricted to just a single dialect, and absent from its immediate neighbors (e.g., **asi* > Vrs /œs/). When the same change took place in two geographically separate dialects, it probably reflects parallel development. For example, given what is otherwise known about these two languages, it would not make sense to group the geographically distant Lehali and Lakon together on the basis of their shared changes *s > /h/ and *a(C)i > /æ(C)/, as illustrated by **asi* > Lhi *næh*, Lkn *ˀæhˀæh* (see the boxes in Figure 6). Clearly these are two parallel developments, which occurred independently in both places.

Now, from the perspective of the forthcoming discussion, it may be interesting to underline the coincidence that both in Lehali and Lakon, the resulting string /æh/ turns out to be as different as one can get from the forms of their actual day-to-day neighbors, Löyöp /n-ɪs/ and Olrat /nɪs/ respectively. And indeed, in this area of Vanuatu, it is common to find a higher degree of phonological differentiation between languages which are otherwise, both geographically and

linguistically, very close to each other. It looks as though sound change might have been (sub)consciously exploited by each speech community as a means to distinguish itself from its immediate neighbors (see Section 6.2).

4.3 Language diversification through irregular sound change

4.3.1 *The diffusion of irregular forms*

In addition to regular sound change, some forms have gone through irregular changes, i.e., changes restricted to particular lexical items. The precise dialectological history of each cognate set can often be traced with reasonable accuracy. By observing modern forms in the light of known regular sound correspondences, it is possible to reconstruct intermediate protoforms, some of which show irregular innovations, and observe their distribution in the dialect chain.

For example, in the case of the noun 'crayfish' (Table 3), while all the modern forms are probably ultimately cognate (but see below), it is necessary to reconstruct not one, but four distinct protoforms:

- Mta *ura*, Vrs *ʊr*, Num *wʊr*, Mrl *n-ə͡ʊr*, etc. all point to a pre-modern form *(*na*) *úra*
- Lyp *nɪ̂ɛ̀j*, Vlw/Mtp *nɪj*, Lmg *nær*, and Vra *nirɪ* all point to *(*na*) *íra*
- Kro *rɛ̂ə̀ŋ*, Olr *nurɪŋ*, and Lkn *uræŋ* all presuppose a form *(*na*) *uráŋi*
- Hiw *ᵍlɔɣ*, Ltg *rɔɣ*, and Lhi *njʊ* all point to a protoform *(*na*) *ráɣu*

For reasons of space, I will not detail here the individual changes for each of the 17 forms of this noun, but each of them can be regularly derived from one of these four intermediate protoforms, e.g., Lhi *njʊ* < **nrʊɣ* < **na ráɣu*; Lyp *nɪ̂ɛ̀j* < **nɪ̂ɛ̀r* < **na íra*, etc.[2] The geographical range of these different protoforms is shown in Figure 7.

The four intermediate reconstructions are all ultimately derived from the Proto-Oceanic etymon for 'crayfish', namely **quraŋ* (Ross, Pawley & Osmond 2011: 165). Yet among them, only one, namely **ura*, shows the expected form based on regular sound correspondences. Thus, considering that the loss of POc *q in all positions and the loss of word-final consonants are both common changes in the North-Central Vanuatu linkage (Clark 2009: 10, 17), the protoform **quraŋ* regularly yields a pre-modern form **ura*, of which a word like /ʊr/ in Vurës is a perfectly regular reflex (via metaphony and unstressed vowel deletion, see Section 4.2.1). By contrast, the three other reconstructions involve an extra change which cannot be explained by regular sound change alone. Thus **ira* entails an irregular vowel change *u > *i, found only in this lexeme. The form **uraŋi* illustrates the locally frequent yet unpredictable accretion of a non-etymological *-i, and the subsequent retention of the root-final consonant (*ŋ). As for **raɣu*, it is unclear

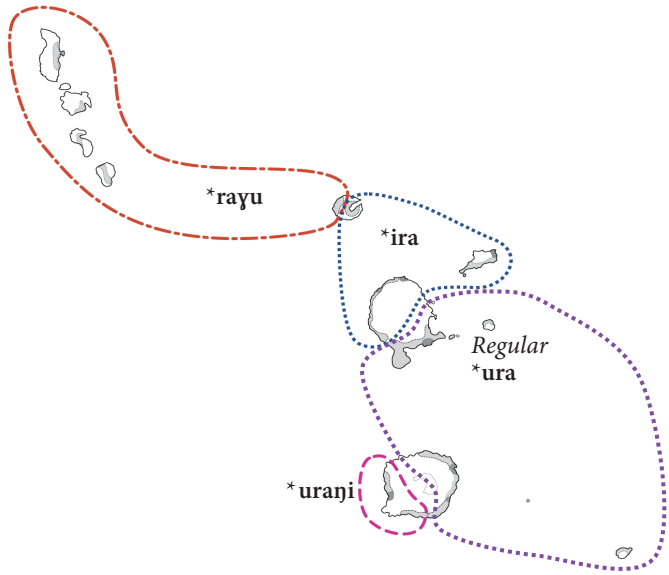

Figure 7. Lexical differentiation through irregular sound change: Intermediate innovations for the noun 'crayfish' (< POc **quraŋ*)

whether it is an irregular reflex of **quraŋ* (via ***ɣura* and metathesis? or ***uraŋu* > ***uraɣu*?), or if it is in fact non-cognate with **quraŋ*, in which case it would be a case of lexical replacement (Section 4.4). This minor question makes little difference for our current discussion, which involves the spread of innovative forms.

Because **ura* is regular, its reflexes are merely cases of shared retention: the processes of sound change affecting them have to do with regular phonological changes that do not target this particular lexeme. But the case is different for the three other forms, which are clearly innovative. Each of these three forms (**ira*, **uraŋi*, **rayu*) is unpredictable, and cannot be assigned any simple phonetic motivation. The fact that they are each found in several adjacent languages (Figure 7) is best explained by a diffusional scenario: each form had to emerge in some dialect, and then spread by imitation, across a portion of the social network.

The complete history of the individual cognate set for 'crayfish' can thus be reconstructed as a layering of multiple post-dispersal innovations, each with a different geographic scope. Some spread across vast dialect networks, such as the (regular) change *q > *ʔ > Ø (**quraŋ* > **ura*(*ŋ*)) witnessed throughout North-Central Vanuatu. Others limited their impact to just a few adjacent dialects, such as the vowel change *(*q*)*ura*(*ŋ*) > *(*q*)*ira*(*ŋ*).

Interestingly, in those cases for which a relative chronology of changes is reconstructable, it can be shown that innovations affecting a smaller area sometimes took place *before* those affecting the larger network. Thus, the accretion of a paragogic vowel *-i (as in *(*q*)*uraŋ* > *(*q*)*uraŋi*), which happened with various lexemes in southwest Gaua (François 2005a: 479), must necessarily have taken place at a time when POc word-final consonants were still present. The regular loss of word-final consonants (e.g. *(*q*)*uraŋ* > *(*q*)*ura*#), a sound change found across the entire NCV linkage, necessarily took place later. At first sight, this scenario seems at odds with the tree-like form of Figure 4 (Section 3.2), since it implies that certain innovations characteristic of higher nodes (large social networks, in this case the NCV linkage) must have taken place *after* other changes in the lower nodes (small social networks, in this case three dialects of Gaua). But this is precisely an important difference between the classic interpretation of phylogenetic trees in terms of successive demographic splits, and events of post-dispersal horizontal diffusion such as the ones I am describing here. Because the "nodes" of Figure 4 do not define subgroups, but linkages, they are not strictly ordered in time; they merely reflect relations of spatial and social inclusion.

In other cases, the reverse ordering of events (large-scope first, small-scope later) must be reconstructed. It must be borne in mind that such processes of diffusion, especially those regarding sound change, were only possible as long as the dialects in contact maintained enough mutual intelligibility to let innovations spread. For example, it is highly likely that the replacement of *(*q*)*ura*(*ŋ*), across the northern Banks area (Figure 7), with an innovative form *(*q*)***i***ra*(*ŋ*), occurred at a time when the dialects involved still had very similar phonological systems, with five vowels and a trill /r/. It is only later, once the form **ira* had settled in that area, that each dialect went through its own idiosyncratic forms of phonological change, e.g., **na íra* > /n-îɛ̂j/ in Löyöp, /n-ɪj/ in Mwotlap, /nær/ in Lemerig, /nirɪ/ in Vera'a.[3] If the order of sound changes had been the opposite (phonological sound change first, lexically-specific change later), this would have meant that the word 'crayfish' would have first evolved into distinct forms such as (**na ura* >) ***n-ʊj* in Löyöp and Mwotlap, ***nʊr* in Lemerig, ***nurʊ* in Vera'a. Later, as speakers of, say, Lemerig shifted irregularly from ***nʊr* to /nær/, their Vera'a neighbors would have "borrowed" the irregular change, and altered their own ***nurʊ* into /nirɪ/. Such a scenario, whereby languages could have borrowed an irregular sound change after they had become phonologically differentiated, would be highly unlikely. In such a case, it is therefore safer to reconstruct a sequence whereby the change with a larger geographic scope, irregular *(*q*)*ura*(*ŋ*) > *(*q*)*ira*(*ŋ*), took place first, whereas sound changes limited to smaller communities (e.g., regular *iCa > /æC/ in Lemerig) are more recent (François 2005a: 456). Occasionally, the recent date of certain phonological changes can be supported by external evidence. For example,

the shift *r > j in the northern Banks area can be shown to have occurred within the last century or so, both on Motalava (François 2001:62) and on Ureparapara; the {*r > j} isogloss one observes today results here from a recent process of diffusion.[4]

Overall, events of diffusion sometimes spread across large chunks of the network, and sometimes across reduced areas, with no principled way to order these in time. Both orders are attested:

- Innovation spread to a smaller area first: [*(*q*)*uraŋ* > *(*q*)*uraŋi* in southwest Gaua], followed by an innovation spread to a larger area: [loss of *C# in all NCV languages]
- Innovation spread to a larger area first: [*(*q*)*ura*(*ŋ*) > *(*q*)*ira*(*ŋ*) in northern Banks], followed by an innovation spread to a smaller area: [e.g., *r > /j/; vowel metaphony in each distinct language]

In sum, irregular processes of (word-based) sound change were the reason why a single inherited etymon **quraŋ* separated, as it were, into four intermediate forms. In turn, these followed various patterns of phonological and morphological change, both regular and irregular, which eventually brought about the various different forms observable today. A similar demonstration could account for the various modern forms shown in Table 3 for 'breathe', which combine layered processes of irregular and regular sound change.

In each case, language fragmentation results from an accumulation of various innovations (in these examples, sound change), each of which diffused to a different portion of the social network. The overlay of multiple intersecting isoglosses ended up providing each local communalect with its own linguistic identity.

4.3.2 *Parallel innovation in lexically-specific sound change*

Geographically speaking, the distribution of etymological variants for 'crayfish' as shown in Figure 7 was orderly enough to group adjacent languages together, and to reconstruct a single irregular sound change that diffused by contact. But the situation can sometimes be less tidy, in ways that make it necessary to consider the possibility of drift or parallel innovations, in a way similar to what was proposed in Section 4.2.2 for phonological change.

Consider, for example, the case of the 1st person dual inclusive pronoun, given in the rightmost column of Table 3. Its full etymological form can be reconstructed as **ɣi*^{n}*da-rua*, itself from Proto-Oceanic **kita* '1incl:Pl' + **rua* 'two'.[5] This pronoun **ɣi*^{n}*darua* evidently acquired three irregular reflexes at some point in its history: **^{n}darua* (via apheresis); **^{n}daru* (apheresis + apocope); **ɣi*^{n}*dua* (via syncope). For reasons of space, I will not detail here the individual changes for each of the 17 modern forms, but each of them can be regularly derived (François 2005a) from one of these intermediate reconstructions, for example:

–	Lhi	*yɪnjɔ*	< *ɣɪ^{n}drɔ < *ɣi^{n}darúa	
–	Hiw	*tə^{g}lə*	< *^{n}dərə < *^{n}darúa	(< *ɣi^{n}darua)
–	Drg	*^{n}daːr*	< *^{n}dáru	(< *ɣi^{n}darua)
–	Lkn	*wʊʧʊ*	< *ɣʊʧʊ < *ɣə^{n}dʊ < *ɣi^{n}dúa	(< *ɣi^{n}darua)

The dialectological distribution of these four intermediate protoforms (Figure 8) is geographically more chaotic than the neat isoglosses we saw in Figure 7 above. It is likely that at least some of the innovations shared by distant communities reflect parallel innovation, or drift. For example, the syncope which yielded the protoform *ɣi^{n}dúa must probably have taken place separately in the history of Löyöp, Vera'a and Lakon; hypotheses resorting to contact or migration, although not excluded, would not be very convincing here.

In sum, just like regular sound change (Section 4.2), we have seen that processes of word-specific sound change sometimes involved diffusion, as in Figure 7, and sometimes occurred independently in separate languages. In both cases, the change was restricted to a relatively small area, and has resulted in greater linguistic fragmentation at the scale of the Torres-Banks archipelago.

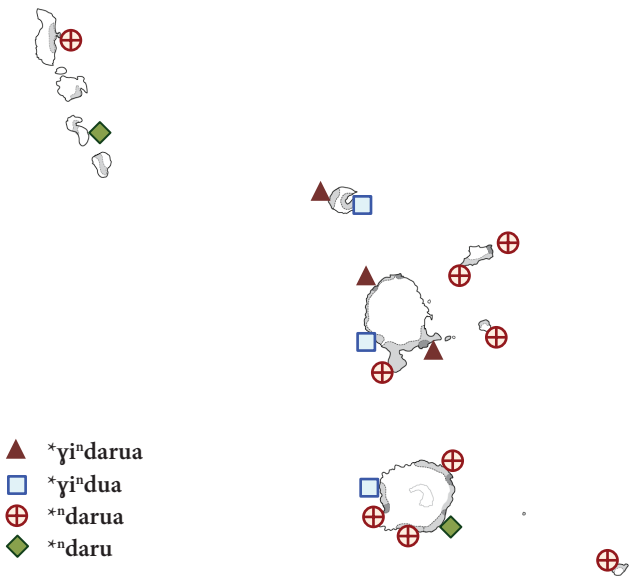

Figure 8. Irregular morphological change and chaotic distribution: Intermediate protoforms for the 1Incl:Dual pronoun

4.4 Language diversification through lexical replacement

In the course of three millennia, the combined effects of regular and irregular sound change thus managed to create enough diversity to make these languages mutually unintelligible, even in those cases where the modern forms are ultimately cognate with each other. But in addition to this widespread tendency towards sound change, another process has proven even more powerful in bringing about linguistic divergence over time. This is LEXICAL REPLACEMENT, when an inherited form is replaced by another word, historically unrelated, for a given meaning.

4.4.1 *Lexicostatistics and the rate of lexical replacement*

In his lexicostatistical study of the New Hebrides, Tryon (1976) calculated percentages of lexical cognacy between pairs of languages, based on a list of about 250 "basic" vocabulary items. In his tables of Torres and Banks Islands (1976: 95), the highest figure of lexical closeness between any two languages, Mwesen and Vurës, is 88.4% of shared basic vocabulary.[6] But more significant is the lowest figure of his chart, which shows that the two lexically most distant languages of these islands, Hiw and Lakon, share no more than 44.5% of their basic lexicon. All other numbers for the Torres and Banks Islands range between 45 and 88% of shared vocabulary. Not surprisingly, lexical diversity becomes even higher if the scope of observation is extended south of the Banks area to other languages of the same North-Central Vanuatu linkage; for example, Lakon shares only 35.2% of basic vocabulary with Apma (Pentecost Island), and 25.7% with Lelepa (Efate Island).

By definition, these figures tell us nothing about linguistic differentiation among cognate forms through sound change (Section 4.2, 4.3). Yet they are quite instructive when it comes to weighing the degree of lexical replacement in the Torres-Banks area. By way of comparison, lexicostatistical figures calculated with the same method for better-known subgroups of Indo-European (Dyen, Kruskal & Black 1992) show that the rates of shared vocabulary range between 91.4 and 61.0% in the Slavic subgroup, between 97.0 and 53.6% in Germanic, and between 89.9 and 51.1% within Romance.[7] None of these three groups approaches the rate of 44.5% which is found within Banks-Torres alone, to say nothing of the North-Central Vanuatu linkage as a whole.

One could debate how statistically significant the difference is between the figures of 44.5% and the lower rates (e.g., 51.1%) found for the Indo-European subgroups. The time depth involved in these different cases of linguistic diversification is essentially similar: the "protolanguage" ancestral to the Torres-Banks area ("Proto-Torres-Banks") began its existence about 3200 years ago, but must have only broken up several centuries later, after a period of shared development during which it formed a homogeneous dialect network (a "single language"). Such

dates, between three and two millennia of age, are thus comparable to those generally accepted for the three subgroups of Indo-European mentioned here. In this perspective, the chronological pace of lexical diversification in the Torres-Banks linkage may be said to be essentially similar to each of the three Indo-European subgroups mentioned here.

While lexical replacement does not seem to be significantly faster than elsewhere, what is extreme here is rather its geographical and human density. Indo-European high-order subgroups such as Romance, Germanic or Slavic all involve immense language communities whose demographies number hundreds of millions of speakers dispersed across vast territories. By contrast, the Torres-Banks archipelago represents a land surface about 10% the size of Corsica, and has a population of just a few thousand individuals (today 9,400); it only forms a subset of the North-Central Vanuatu linkage, itself merely a branch of Oceanic. The fact that this very low-order subgroup of Austronesian, with such a small population, has reached such a degree of lexical diversity is in itself worthy of notice.

4.4.2 *Semantic change, markedness shift, and lexical replacement*

Combined with the various patterns of sound change we saw earlier, lexical replacement largely accounts for the high level of divergence found between the word forms of northern Vanuatu languages. See, for example, the variety of words for 'properly' or 'speech' in Table 1 (Section 1.2).

One question arises from this observation, namely, what motivates lexical replacement. Is this a conscious process? Is it primarily driven by social factors? In the final section of this study, I will indeed argue that sociolinguistic factors in this area of Melanesia play an important role in the diffusion and spread of lexical innovations. However, I will also propose (Section 6.3) that these factors only account for the DIFFUSION of innovations once these have already emerged. As for the question of how lexical innovations themselves ARISE in the first place, I will claim here that this is best explained by universal tendencies of semantic change and lexical replacement, which are not specific to this region.

Languages replace lexical items for a variety of reasons. They may renew items of their vocabulary, for example, to avoid homophony (Gilliéron & Roques 1912, Keller 1989, Campbell 2004a: 322). Another motivation may be related to the social practice of linguistic taboo (Blust 1981). Indeed, Simons (1982: 183) considers linguistic taboo to have played a key role in lexical replacement in the Austronesian family. However, this explanation is not entirely convincing for northern Vanuatu. The only form of taboo reported for this area (Codrington 1891: 43), which is still in use today, forbids individuals from uttering the names of their in-laws; should they need to use a word phonetically similar to a forbidden name, they will have to replace it with a synonym. Obviously, this form of word taboo

differs from person to person depending on their position in the kinship network, in such a way that it never affects an entire community. Such a cultural explanation is therefore not the final word to account for lexical replacement.

Most often, the replacement of one word by another simply results, I believe, from universal processes of semantic shift (Blust 1987, 2010, Wilkins 1996, Blank & Koch 1999b, Vanhove 2008). In typical cases of lexical replacement, the starting point is a language which, at a given point in time, has a number of close synonyms for a given meaning. These may initially differ in their semantic nuances, their connotations, or their stylistic registers. As speakers manage their way through the manifold adjustments of daily communication, some semantic or pragmatic extension of one lexeme may gain ground over its lexical neighbor, and tend to replace it in certain situations. Should the new usage take root and spread through the speech community, eventually the old term becomes obsolete and ends up being replaced by the new term, following a process called MARKEDNESS SHIFT (Dik 1989: 44).

This is how, for example, Late Latin progressively replaced its old term *caput* 'head' with another noun *testa*, originally 'earthen pot' (cf. Blank & Koch 1999a). I propose to illustrate the four stages taken by this case of lexical replacement in Table 4. The brackets symbolize when a term is marked, as opposed to being the default term for a given meaning (in bold). The first innovation (STAGE 1) was to use the noun 'pot' as a figurative, slang word for 'head', in competition with the inherited form *caput*, in a way similar to French slang *carafe* 'jug' for 'head'. Eventually, *testa* lost its jocular connotations, and ended up as the standard term for this body part (STAGE 2), as evidenced by Italian *testa*, and French *tête*. The older term *caput* (> Italian *capo*, O.Fr. *chief*) resisted for some time, but had become the marked term in the pair; in 17th century French, *chef* was still used as an archaic, highly marked synonym of *teste*, until it disappeared in this sense. In modern French, *chef* survives only with another meaning (historically secondary) of *caput*, namely 'leader, chief' (STAGE 3). The whole process can be summarized by stating that Latin/Romance "replaced" its original noun *caput* (STAGE 0) with an innovative form *testa* (STAGE 3), but it must be borne in mind that this formulation is simply a shortcut for what is really a gradual, four-stage sequence involving lexical competition and markedness shift among polysemous terms (see Sweetser 1990: 9; Evans & Wilkins 2000: 549). Rather than a direct shift from STAGE 0 to STAGE 3, the key turning point in this evolution is really the shift from STAGE 1 to STAGE 2, and the inversion of markedness.

Most histories of lexical replacement in northern Vanuatu are likely to be explained by a four-stage sequence similar to this *caput* > *testa* example. Sometimes the source of an innovative form can be identified, giving a clue to the semantic path which was followed. One such example can be found in Table 3 (Section 4.2.2).

Table 4. When markedness shift drives lexical replacement: Words for 'head' in Late Latin/Romance

	'leader, chief'	'head'	'earthen pot'
Stage 0	*caput*	***caput***	*testa*
Stage 1	*caput*	***caput*** ~ (*testa*)	*testa*
Stage 2	*caput*	(*caput*) ~ ***testa***	*testa*
Stage 3	*caput*	***testa***	*testa*

While most languages reflect Ptb *^{n}*daeru* for 'coconut crab (*Birgus latro*)', Löyöp has a non-cognate form *n-^{n}døɣîɛ̂j*. Literally, this noun parses as 'pandanus leaves' (i.e., ^{n}*dø-* < *^{n}*rau* < POc **raun* 'leaf' + *ɣîɛ̂j* < **ɣire* < POc **kiRe* '*Pandanus tectorius*'); this is due to an analogy between the prickles of those leaves and the spines on the crab's carapace. In a way similar to the Latin example above, one can imagine that *n-^{n}døɣîɛ̂j* 'pandanus leaves' started its life as a colourful metaphor for the animal alongside the inherited form, which was presumably a reflex of **^{n}daeru*. In spontaneous speech, the old form was to be replaced more and more often by its marked synonym, which eventually became the unmarked name for the crustacean. Similarly, Mwerlap has replaced **^{n}daeru* with another compound form: *nɵ-ɣɵtɵ-tɛ̂ak*, literally 'the spearing hermitcrab', with reference to the crustacean's dangerous limbs. These are cases of lexical replacement via compounding and markedness shift.

4.4.3 *Lexical competition in progress*

Quite often, the process of lexical replacement has not yet taken any definite shape, but already lurks in the ongoing competition between two semantically very close words. By observing the patterns of variation between speakers, it is possible to actually spot lexical replacement in progress. Thus, to take a last example from Table 3 above, the Hiw noun for 'song' used indeed to be *ɔs* (< **asi*); yet this is now a marked, archaic term which is reserved for 'sacred song' or 'Christian psalm'. The unmarked word for 'song' today, often the only one known to younger speakers, is *təntənɔ*, derived from the verb *tənɔ* 'learn'.

Instead of showing lexical competition as a completed process (Stage 3), what these cases illustrate is precisely the phase of lexical competition between two words (Stage 1–2), as a forerunner of lexical replacement. The modern languages of northern Vanuatu teem with synonym pairs, and it is likely that this was also the case in earlier stages of their history. Some synonyms are simply free variants, and indeed are used freely by the same speakers in the same contexts, e.g., *na-mtɛ* and *nɛ-^{m}bɛɣ*, two perfect synonyms used for the breadfruit tree in Mwotlap. But in many cases, the two words differ either semantically or stylistically, in terms of

register. For example, Mwotlap has a verb *jɔŋtɛɣ* 'hear, listen'. All speakers would agree that this form is a "casual word" (*hɔhɔlɛ βasapsawjɛɣ*), and that in certain social contexts, e.g., in order to adorn one's story-telling with a colorful style or to show respect to one's in-laws by carefully choosing a more delicate wording, it would be more appropriate to use *haltiŋk͡pʷɔj*, a "heavy word" (*hɔhɔlɛ map*) with the same meaning.

Similarly, young Hiw speakers are taught to use "respectful speech" (*məŋa tɵmmaβə*) or "avoidance speech" (*maŋə βisəβisə*) as a token of respect when addressing important interlocutors. In such a situation, rather than using dull and direct everyday words such as in (1), it is more appropriate to resort to some sophisticated paraphrasis such as in (1′):

(1) Hiw nɔkə ɣənɣɔn
1sg eat
'I'm going to eat.'

(1′) Hiw nɔkə ᵍLak-aᵍL-ɣe ti nɔkə
1sg seize-seek-thing dat 1sg
'I shall seize something for myself.' [respectful equivalent of 'I'm going to eat']

Insofar as the morphological sequence /ᵍLak-aᵍL-ɣe/ has already become conventionalized in the respectful register as a synonym for 'eat' (Stage 1), it is not difficult to imagine a potential scenario whereby this expression, after progressively losing its stylistic markedness, might end up becoming the new standard way to say 'eat' in Hiw (Stage 2), and eventually replaces *ɣənɣɔn* in the lexicon (Stage 3).

The dynamic lexicons of the northern Vanuatu languages display all possible degrees on the scale of stylistic markedness, from the familiar variant to the fading-out, literary archaism. Crucially, what appear to be stylistic differences from the synchronic point of view are often correlated with historical processes of lexical change. Whether the new expression reflects a slangish shortcut or an elevated metaphor, it will begin lexical competition with the word previously in use for the position of being the default form for a given meaning. While this arm-wrestling may continue for a long time, it will often end up with the victory of one form over the other.

4.4.4 *The diffusion of lexical innovations*

These processes of lexical replacement have resulted in a high degree of linguistic divergence among the northern Vanuatu languages. Table 5 provides a random sample of words of different syntactic categories, showing various degrees of lexical replacement. Other similar cases can be found in Table 1 (Section 1.2), and in Table 6 (Section 5.2).

Table 5. Language differentiation through lexical replacement: a few lexical sets from northern Vanuatu

	'person'	'woman'	'true'	'Complementizer'	'(be) like'
Hiw	[1] tajə	[1] jəkʷen	[1] βəwjə	[1] tɔm	[1] ŋʷe
Lo-Toga	[1] telə	[1] ləkʷɛβinə	[1] βəwiə	[1] tɛ	[1] wɛ
Lehali	[2] nat	[1] nlɔkβɛn	[2] tŋʷas	[2] ⁿdɛ	[2] nan
Löyöp	[2] nat	[1] nlik͡pʷîɛn	[2] taŋ͡mʷas	[3] sɛ	[3] mɛnɛ
Volow	[2] n-at	[1] nlɛⁿg͡bʷɛβɪn	[3] hɛjwɪ	[4] ᵑgɔ	[4] βɪh
Mwotlap	[2] n-ɛt	[1] nalk͡pʷʊβɪn	[3] hɪjwɪ	[3] sɔ	[5] k͡pʷɛlɛ
Lemerig	[3] nʔʊŋsʊr	[2] nrɛŋɛ	[4] ʔirɣi	[5] wœ	[6] mækæ
Vera'a	[3] ʔaŋsar	[2] rɛŋɛ	[5] irʊr	[3] sɔ	[7] sɛ
Vurës	[3] taŋsar	[3] rɛk͡pʷɛ	[6] ⁿdyn	[6] βita	[8] timîak
Mwesen	[3] taŋ͡mʷsar	[3] rɛk͡pʷɛ	[6] nun	[6] (wɔ)ta	[8] tɛmɛk
Mota	[4] tanun	[4] taβne	[6] nun	[5] was	[9] taŋ͡mʷa
Nume	[4] tuⁿdun	[5] tawa	[6] ⁿdun	[7] si	[9] taŋ
Dorig	[4] tⁿdun	[3] rk͡pʷa	[6] lɛⁿdun	[4] (k)ak	[10] ŋ͡mʷray
Koro	[4] tuⁿdun	[3] rak͡pʷa	[7] βutwɪ	[4] (k)ak	[10] ŋ͡mʷar
Olrat	[4] ʧuʧuŋ	[3] rak͡pʷa	[7] βutwɪ	[4] ka	[10] ŋ͡mʷaj
Lakon	[4] ʧaːʧun	[6] ɣamtʊ	[7] wuswɪ	[3] sa	[10] ŋ͡mʷɪrɪ
Mwerlap	[4] nɛtɛⁿdʉn	[7] naβaβɛ̂an	[3] sʉrwɛ̂a	[7] si	[10] ŋʷɛr
nb etyma	4	7	7	7	10

In some cases, the comparison of lexical isoglosses combined with knowledge of regular sound correspondences makes it possible to group forms together that are cognate. Each column of Table 5 has subscript numbers labelling those forms which derive from the same protoform, i.e., whose modern differences are only due to regular or irregular sound change; the last row shows the number of distinct cognate sets that need to be posited for each meaning.

For example, the 17 modern forms for 'person', in the first column of Table 5, can be reduced to four cognate subsets, reflecting **talua*, **ata*, **tamʷasara* and **taⁿdunu* respectively. Of these, **ata* is PNCV **ʔata* 'spirit, soul; person' (Clark 2009:76), itself from POC **qata* 'shadow, reflection; soul'. The **ta* element in other forms is either a truncated form of the same noun, or a reflexe of POC **tau* 'person'. These longer forms are compounds, literally PTB **ta-ⁿdunu* 'real person' (cf. the column 'true') and **ta-mʷasara* 'ordinary person' (cf. **mʷasara* 'poor, needy'). The etymology of **talua* is unknown.

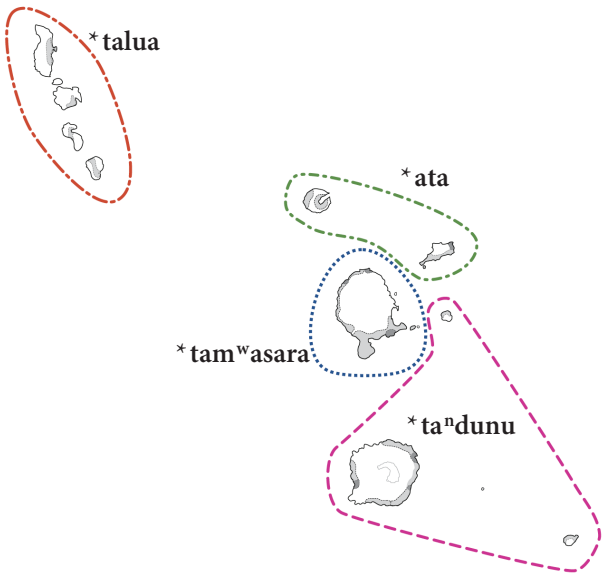

Figure 9. Language differentiation through lexical replacement: Intermediate protoforms for the noun 'person'

The dialectological distribution of these four etyma defines four geographically coherent clusters of languages. In line with maps given earlier for sound change, lexical isoglosses for the noun 'person' are charted in Figure 9. In the absence of compelling evidence, Figure 9 remains agnostic as to which of the four forms are to be considered conservative vs innovative.

Just as we saw earlier for sound change, linguistic divergence here followed a multilayered process. First, the early dialect chain witnessed the emergence of new forms for the same meaning, each of which spread to a portion of the dialect network. At some point, the initial unity had broken up into four isoglosses for this particular word. Later on, each of the four etyma underwent local patterns of sound change, which increased the degree of differentiation within each dialect set: e.g., **tamʷasara* > Msn *taŋ͡mʷsar*, Lmg *nʔvŋsvr*; **taⁿdunu* > Mta *tanun*, Olr *ʧuʧuŋ*, etc.

Notice, incidentally, that the distribution of lexical isoglosses for this word 'person' does not match the one observed earlier for the noun 'crayfish' (Figure 7), let alone the random scattering found with the 1incl:Dual pronoun (Figure 8). The more examples of individual lexical items are brought in, the more inconsistent the picture becomes. This shows how the spreading of innovations differed from one individual item to another. Should a historical linguist attempt to map together various kinds of isoglosses with the hope of identifying coherent clusters of

shared innovations, this would turn out to be a vain endeavor. Even though some dialects may have evolved together with respect to certain innovations (cf. the four languages of Vanua Lava I. in Figure 9), they are always separated by a large number of other innovations (cf. the same four Vanua Lava languages in Figure 7 and Figure 8). This pervasive intersection of isoglosses, which is typical of linkage phenomena (Section 3.2), jeopardizes any hope of ever coming up with a neat phylogenetic tree for the region.

The noun 'person' is the simplest case of Table 5. The other examples show even more intricate histories, which I will not detail here. This is true not only of lexemes, e.g., 'woman' and 'true', but also of grammatical morphemes, like the complementizer. In these cases, the combination of lexical innovations with morphological or phonological change appears to have defined small areas of diffusion, sometimes restricted to just one or two adjacent dialects.

4.5 Summary: On linguistic diversification in northern Vanuatu

Overall, three major kinds of change have contributed to the diversification of northern Vanuatu vocabularies: REGULAR PHONOLOGICAL CHANGE at the system's level, IRREGULAR SOUND CHANGE and morphological innovations in specific words, and LEXICAL REPLACEMENT.

Beyond their differences, these three types of change share a few properties. Each innovation must have arisen in the speech of some individuals in the archipelago, and then diffused from speaker to speaker, by imitation, until it settled in a certain portion of the social network. Sometimes innovations were confined to only a single communalect, or a set of adjacent villages. In other cases, the isogloss covered a larger, contiguous area, involving dialect contact across islands. Some changes found in widely separate places may be due to drift or parallel innovations. As more innovations added up over time, each member of the dialect chain further differentiated itself even from its closest neighbors; this is what brought about the linguistic fragmentation we observe today.

Considered individually, most of the innovations which have occurred in the history of the northern Vanuatu languages are unproblematic, and can be explained by typologically common processes of language change. Two points, however, are characteristic of this group of languages, and are still in need of an explanation.

One point is the high density of these innovations. For example, we saw (Section 4.4.1) that rates of lexical diversity in basic vocabulary in the northern Vanuatu languages are similar to, and even higher than, those found within certain high-order subgroups of Indo-European, for a much smaller population. I will propose, in Section 6.2 below, that this high degree of linguistic diversification can

be explained by the language ecology of the area and social attitudes that prevail among the local population.

The second paradox is that the high degree of linguistic divergence observed among these 17 languages goes along with a reverse trend, one which could be described as a form of regional CONVERGENCE. However, this convergence does not show up in the same linguistic domains as the diverging tendency we just observed. Fundamentally, the domains in which the northern Vanuatu languages show most intense differentiation are those that affect the FORM of words, whether this involves sound change or lexical replacement. By contrast, the same languages show much more homogeneity with respect to the STRUCTURAL COMPONENT of these words, whether in terms of their morphosyntactic properties or their semantic makeup. This structural homogeneity is the subject of the next section.

5. Structural isomorphism in northern Vanuatu

5.1 Structures vs. forms

A sign may be described as the association of a specific phonological string or FORM with properties of STRUCTURE, the latter referring to the organization of meaning into language-specific categories and constructions.

This contrast, *structure* vs. *form* corresponds more or less closely to what Ross (2001: 148) called *lemma* vs. *form*, or to the contrast drawn by Grace (1981: 24) between, respectively, *content form* vs. *lexification* (see the citation in Section 6.4). The reason why I am not using Grace's term *content form*, a term originally from Hjelmslev (1961: 52), is its relative opacity, especially since it is supposed to contrast with what I (following Ross and others) call *form*.

As for the term *lexification*, it is too restrictive if it is taken to apply exclusively to the form of lexemes. My use of the pair of terms *structure* vs. *form* is really orthogonal to the traditional divide between grammar and lexicon. "Form" refers to the phonological substance of the linguistic sign, whether it is a lexeme, a grammatical morpheme or a construction. As for the term "structure", it refers to those relational properties between signs, whether relations of contrast or of the combination of signs. This includes those structural relations on the syntagmatic axis (morphosyntax, combinatorics, phraseology), as well as those on the paradigmatic axis (lexical and grammatical semantics). Under this structuralist model, semantic properties of lexemes and morphemes (their "meaning") are subsumed under their structural properties, insofar as they are defined by their contrasts, and by their various combinations within the language's repertoire of constructions.

Two languages can be said to be ISOMORPHIC in a particular domain of their system if the signs they use in a given speech situation share the same structural (or constructional) properties. For example, the two sentences in (2) are perfectly isomorphic because they express the same functional content by resorting to identical structures:

(2) ENG *You can take your time.*
FR *Tu peux prendre ton temps.*
2SG POT.AUX take POSS:2SG time
'You don't need to hurry.'

The isomorphism between English and French here concerns not only the parallel order of words, but also the inherent properties of each component in the sentence. Thus both languages have parallel constructions in which the abstract notion 'time' is modified by a possessor, and the object of an active verb 'take'. The two languages provide that construction with the same meaning ('act as slowly as necessary or desired'), and with the same array of pragmatic implications ('no need to hurry, you can relax', etc.).

Structural isomorphism can be perfect as in (2), or quasi-perfect as in (3):[8]

(3) Eng *I have caught a cold.*
Fr *J' ai attrapé froid.*
1SG PRF catch:PAST.PCP [ART:INDF] cold
'I am sick due to cold weather.'

Even though the two sentences in (3) show one syntactic difference (presence vs. absence of the indefinite article on the object), they are otherwise structurally parallel. Thus in English and French, *have* and *avoir* are both lexical verbs whose meanings include 'possess', and which can both be used as auxiliaries to encode perfect aspect. *Cold* and *froid* are both temperature adjectives which can be used as nouns. Finally, *catch (a) cold* and *attraper froid* are two parallel constructions that share the same semantic properties, considered as a whole ('become sick due to cold weather') as well as in their internal make-up (identical metaphor involving a 'catching' event).

Crucially, structural isomorphism across languages is independent of the relationship between the phonological forms of words. In example (3), the two pronouns *I* and *je*, while ultimately cognate, have evolved in such a way to lose any form of similarity. All other words in (2) and (3) are not cognate between French and English.

5.2 Diversification of forms, isomorphism of structures in northern Vanuatu

As Table 1 (Section 1.2) already suggested, the languages of northern Vanuatu exhibit countless patterns of structural isomorphism, regardless of the diversity of their forms. This parallelism among languages was already witnessed by the missionary-linguist Codrington at the end of the 19th century: "an examination of their languages discovers a very considerable underlying sameness" (Codrington 1891: 20).

Because these languages are all closely related, it can often be ambiguous whether a particular case of isomorphism simply reflects shared inheritance from their common ancestor (Proto-Oceanic), or a post-dispersal innovation which diffused across a linguistic network (Section 3.2). Only in the latter case is it appropriate to speak of CONVERGENCE strictly speaking. Usually, external evidence will make it possible to identify what is shared among the Oceanic languages and may potentially be reconstructed for their common ancestor. In this section, I will focus on examples which appear to be more specific to northern Vanuatu, since they presumably reflect events of convergence restricted to this area as opposed to shared retentions from their Proto-Oceanic ancestor.[9]

Sometimes, structural homogeneity goes along with a set of forms that happen to all be cognate. For example, the verbs meaning 'breathe' in Table 3 (Section 4.2.2) all colexify the two meanings 'breathe' and 'take rest' (cf. François 2008: 183). Because most of the modern forms are cognate, Occam's razor simply suggests assigning this polysemy to their ancestral etymon, **ma*^{m}*busaɣi*, a shared lexical innovation of the Torres-Banks linkage (cf. François 2005a: 482, Clark 2009: 132), and concludes that the meaning diffused together with the form.[10] In a way, this is also a case of structural diffusion, in the sense that some structural property (in this case, an apparently innovative case of polysemy) spread across the whole area, but this is not a spectacular nor a problematic case because it most likely diffused together with the form.

More revealing are those cases which involve both isomorphism of structures and diversity of forms. Indeed, these cases entail that *the very same words* have gone, so to speak, in two opposite directions. On the one hand, their forms have gone through processes of divergence, either through phonological diversification or lexical replacement; on the other hand, their structural properties have maintained — or created — cross-linguistic homogeneity.

What follows will concentrate on the latter cases, since they clearly make it necessary to discriminate between the divergence of forms on one side and the stability or convergence of structures on the other hand. For reasons of space, I will confine the discussion to five sets of words, listed in Table 6.

Table 6. A selection of word sets combining formal diversification and structural parallelism across languages

	LEXICON			GRAMMAR	
	Adjective 'random'	**Postverb 'find'**	**Restrictive 'just'**	**Negative Existential**	**Time Focus aspect**
Hiw	βitikəje	sɔᵍʟ	ŋʷutʉjə	təɣɔ	takə
Lo-Toga	kʷure	hɛr	wərəŋɔ	tatəɣɛ	akə
Lehali	βælækjæ	sɪɣɪt	ɛwwɛ	tɛtɣɛ	jak
Löyöp	βɔlɔkjɛ	søɣøt	wi͡ɛ	mamʧɛɣɛ	pɛj
Volow	ᵑgɪlᵑgɪl	sas	ɣɛwɪ	tatɪh	ᵑg͡bʷaja
Mwotlap	sɔlɔtɛɣ	sas	ɪwɪ	tatɛh	k͡pʷɔjɔ
Lemerig	βʋlʋkrɛɣ	ʔɛsɣøʔ	k͡pʷɔɣɔr	niβ	mak
Vera'a	βalakra	sɛr	ɣiβa	ɣitaɣ	mak
Vurës	βalakrɛɣ	siœr	ɣɛm	ɔⁿdiaŋ	kara
Mwesen	pɔlɛtɛ	siɛr	ɣɔp	ɛnɛŋ	kara
Mota	popolotaɣ	suar	ɣap	taɣai	k͡pʷara
Nume	kɪlkɪl	wun	am	ᵐbɛk	k͡pʷar
Dorig	kɪlkɪl	wʊn	wɔr	(ɔ)ᵐbɛk	k͡pʷra
Koro	kɪlkɪl	wɛs	wɔr	ᵐbɛk	k͡pʷara
Olrat	kɪlkɪl	wɛn	wɔj	ɣaiβ	tak
Lakon	sælɪn	ɣɪn	wɔː	ɣaiβ	lak
Mwerlap	kɪlkɪl	β͡ɔ͡ɘn	ɣɔm	tɪɣɪ	kʷɛr

Whether they belong to the lexicon or to the grammar, these words are structurally isomorphic, insofar as they display the same properties across languages: the same semantic range (polysemy, polyfunctionality), the same combinatorics, and parallel usage in phraseology. An obvious manifestation of these languages' isomorphism is the fact that most of their sentences are amenable to a single line of interlinear translation, just as in Table 1 above. In the following sections, even when a given construction is attested in most or all of the archipelago, for reasons of space I will keep the examples to just a small sample of up to five languages. The selection will be arbitrary, but will attempt to cover as wide an area as possible, following the same north-to-south order as in previous tables.

5.3 Structural isomorphism in the lexicon

5.3.1 *Adjective 'random'*

All the northern Vanuatu languages have an adjective (often found in verb-modifying position) showing a broad polysemy around a meaning which may be glossed 'random'. Its typical spatial sense describes a motion as random or aimless:

(4)	Mtp	kɪ	ni-	^{n}dɛl^{n}dɛl	*sɔlɔtɛɣ*		lɛ-	pnʊ
	Lmg	ti	n-	tɛltɛl	*βɒlʊkrɛɣ*		lɪ	βunu
	Kro	ni	t-	ɛlɛ͡al	*kɪlkɪl*	ti	l-	βʊnʊ
	Lkn	ni	tɪ	βanβan	*sælɪn*	tɔ	lɪ	βanʊ
		3sg	ipfv	wander~red	random	[ipfv$_2$]	loc	village
		'He was wandering *around* in the village.'						

Outside the spatial domain, a common semantic extension is 'messily, casually, following no specific principle or rule'. This may have negative connotations, as in situations where some social rule should be followed:

(5)	Hiw	titə	tat	βəɣəβaɣə	*βitikəje*		iə
	Mtp	ɣɪn	tit	kaka	*sɔlɔtɛɣ*	βɪstɛ	aɪ
	Lkn	ɣɪʧ	ɣatɪ	suː	*sælɪn*	ɪsaβʊh	ɪhɪ
		1inc:Pl	pot:neg	chat	random	[pot:neg$_2$]	about.it
	Lmg	ɣæt	ɪ	tɛkʔøʔ	*βɒlʊkrɛɣ*	ŋ͡m^{w}æsʔæ	
	Kro	ɣɪn	βtɛ	sarɔr	*kɪlkɪl*	latɛ	
		1inc:Pl	pot:neg	mention	random	pot:neg$_2$	
		[speaking about a taboo issue] 'We can't just talk *casually* about that.'					

(6)	Hiw	ikə	tati	mːɪmːɪ^{g}l	*βitikəje*	
	Mtp	nɪk	tɔɣ	^{m}bɔ^{m}bɔɛl	*sɔlɔslɔtɛɣ*	
		2-sg	proh	be.angry~red	random	
	Lmg	n-	lø-m	ʔɔɣ	sɛsɛʔ	*βɒlʊkrɛɣ*
	Kro	na	llɔ-ŋ	t-	kɔtkɔt	*kɪlkɪl* lɛr
	Lkn	lɔlɔ-ŋ	mɪtɪ	kɔtkɔt	*sælɪn*	lɛː
		[art]	mind2sg	proh	be.angry~red	random [proh$_2$]
		'Stop getting angry *with no reason*!'				

The combination of this adjective with the verb 'do, make' will form a compound verb meaning 'mess up, spoil, damage':

(7)	Hiw	ikə	mik	^{g}Lak	*βitikəje*	nə	tətajwɵ
	Lyp	niŋ	nin	aŋ	*βɔlɔkjɛ*	n-	ɣɪnɣɪn.lɔwɔ
	Mtp	nɪk	tiplɛ	ɣalɛɣ	*sɔlɔtɛɣ*	na-	laβɪt
		2sg	appreh	make	random	art	party
		'You might *mess up* the party.'					

But the same adjective can also have positive connotations:

(8)

Hiw	təkŋʷa	*βitikəje*	
Mtp	iɣɛ	*sɔlɔtɛɣ*	
	hum:pl	random	

'*ordinary* people' (as opposed to rich, powerful or big people)

(9)

Hiw	nɔkə	ne	*βitikəje*	ŋʷutʉjə	
Ltg	nɛkə	na	*kʷure*	wərəŋɔ	
Mtp	nɔ	nɛ-	*slɔtɛɣ*	iwi	
Lmg	nœ	ɣɛ	*βʋlʋkrɛɣ*	k͡pʷɔɣɔr	
	1sg	stat	random	just	

'I'm just a *simple* guy.' (i.e., I don't think I'm special …)
or 'I don't mind!' (i.e., I don't have special demands, I'm easy.)

In all cases, the languages show structural isomorphism, regardless of whether their forms are cognate or not.

5.3.2 *The postverb 'find'*

The northern Vanuatu languages lack any verb 'find'. What they have is a "postverb", or verb modifier, which they use to create compound verbs with the sense 'find'. For want of a better translation, I will gloss this postverb as 'find' even though this is usually not synchronically a proper verb. A rough parallel would be Englis *come across X*, where the sense 'find' results from the combination of *come* and *across*.

The verb V_1 that precedes the postverb describes the action which led to the finding event, i.e., looking (10), listening, groping, searching, walking (11), and so on:

(10)

Lyp	niŋ	mal	ɛt	*søɣøt*?
Mtp	nɪk	m-	ɛt	*sas*?
Num	nik	mɛ	kɪrɪ	*wun*?
Lkn	nɪk	ɪn	ʔætæ	*ɣɪn*?
Mrl	nɛ͡ak	nu-	mətɵ	*βɔ͡ɔn*?
	2sg	prf	see	(find)

'Have you found it?'

(11)

Hiw	nɔkə		to	*sɔᵍʟ*	ti	i	nə	woβɔt
Ltg	nɛkə	na	βen	*hɛr*	t'	e	nə	βɔt
Mtp	nɔ	ma-	βan	*sas*			nɪ-	sɪm
Drg	na	m-	ᵐban	*wʊn*			ɔ	sʊm
	1sg	[prf]	walk	(find)	[prf]	[obl]	art	money

'I've found some money (as I was walking).'

In all these languages, the same postverb is also used in combination with the verb 'think', to construct the meaning 'remember':

(12)	VLW	nɛ	m-	ⁿdɔn	*sas*	ɪɣɪn!
	LMG	nœ	m-	tœt	*ʔɛsɣøʔ*	βaʔanɪ!
	MTA	na	me	nom	*suar*	nakp͡ʷa!
	LKN	na	n-	ʧɪːʧɪm	*ɣɪn*	ɣɪnɛk!
		1SG	PRF	think	(find)	now
		'I remember now!'				

The map in Figure 10 shows two layers of isoglosses for this postverb 'find'. The dotted lines show the distribution of cognate sets across the archipelago, and are thus concerned with the dialectological distribution of word forms, in line with Figure 9 (Section 4.4.4). The solid line, on the other hand, shows that all modern words, despite events of lexical replacement and sound change, share the same structural features.

Recently collected data (Henri 2010) show that the next language further south (Sungwadia, the northernmost language of Maewo Island), despite occasional contact with the Banks Islands, does not have any word which shares all the structural properties of the postverb 'find' found in the Torres-Banks languages. In other words, the solid line in Figure 10 outlines exactly the maximal extent of the structural isogloss for this postverb 'find'.

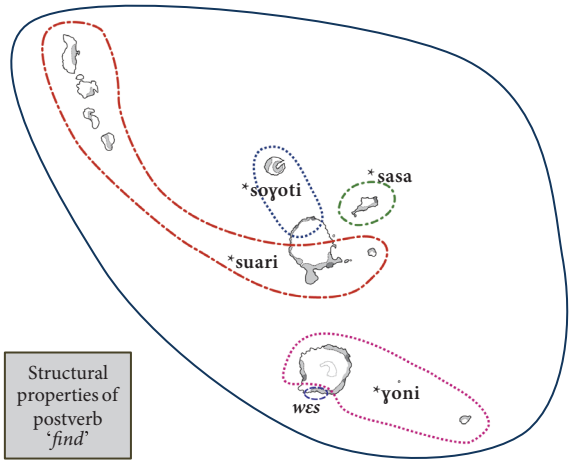

Figure 10. Diversification of forms, isomorphism of structures: The case of the postverb '*find*'

5.3.3 *The restrictive postverb*

All the northern Vanuatu languages have a functionally similar restrictive word which may be glossed as 'just, only'. Everywhere, it fits in the same slot of the postverb, and thus occurs at the end of the predicative phrase. This restrictive postverb indicates that the predicate is restricted to a low value, as in English *only three*, or *just you*. This use was illustrated in (9) above.

Combined with a verb, the restrictive commonly means 'do nothing else than':

(13)	LTG	nikə		βɛrwulimə	*wərəŋɔ*
	MTP	nɪk		ɣalɣal	*ɪwɪ*
	MTA	ka		ɣalɣale	*ɣap*
	LKN	nɪk	tɪ	ɣæl	*wɔː* tɔ
		2SG	[IPFV$_1$]	deceive	just [IPFV$_2$]
		'You're (just) lying!'			

Languages will idiomatically use this restrictive in combination with stative predicates that semantically imply a limited quantity or quality, e.g., 'small', 'light', 'easy', 'close', 'identical', etc.:

(14)	HIW	n′	ɵŋʷə	=kjə	ᵍLIptɔɣ	*ŋʷutʉjə*
	LHI	n-	ɛŋ	jɪkɪs	jiptæ	*′wwɛ*
	LMG	n-	k͡pʷɛk	muk	rɪβʔɪ	*k͡pʷɔɣɔr*
	DRG	n-	ɣaːβru	-k	βriɣtɪɣ	*wɔr*
	MRL	n-	ɛ͡aŋʷɔ	-k	ɣɔmɣɔmtʉ	*ɣɔm*
		ART	house	my	close	just
		'My house is (very) close.'				

(15)	LTG	hor	na	ʈaɣɛtʉwe	*wərəŋɔ*
	LYP	kɪ͡ɛjʊ	nyɛ	ʧːatjɛs	*wɪ͡ɛ*
	LMG	tæru	ɣɛ	ʔæræs	*k͡pʷɔɣɔr*
	VRS	rʊrʊ	ɣa-	sasarɪt	*ɣɛm*
	LKN	wʊrʊ	ɣa	haʧ	*wɔː*
		3DU	STAT	identical	just
		'They're (exactly) the same.'			

Another idiomatic use is to combine the restrictive with the predicate '[be] good'. Because the meaning of the latter is semantically downgraded by the restrictive, what results is a phrase meaning 'it's just fine; no worries' (see (21) below). Finally, example (24) will illustrate how the same restrictive postverb is a typical component of the immediate perfect construction (English 'have just done').

In line with Figure 10 above, Figure 11 superimposes two types of isoglosses. The broadest circle, which encompasses the whole northern Vanuatu archipelago, may be called the "structural isogloss", that is, the isogloss grouping all the languages

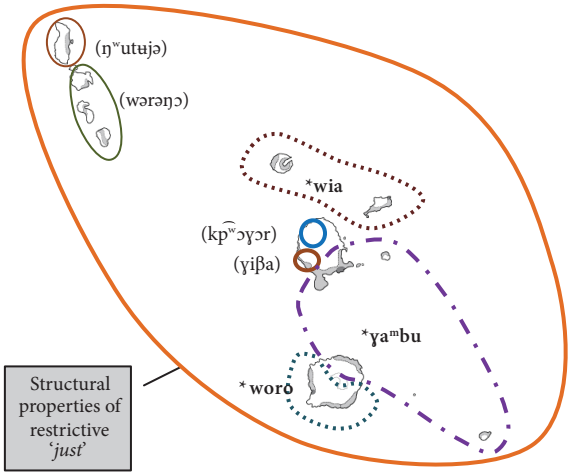

Figure 11. Diversification of forms, isomorphism of structures: The case of the restrictive postverb '*just*'

that share the semantic and phraseological uses of the restrictive postverb, as described in this section. As for the smaller circles, they delineate the maximal area of diffusion of each cognate set. Some of these are identified by reconstructions (e.g., **ɣa^{m}bu*, the regular source of the modern forms *ɣɔp*, *ɣɛm*, *ɣap*, *am*, *ɣɔm*); when a cognate set is restricted to one language, I indicate the form itself.

5.4 Structural isomorphism in the grammar

The isomorphism found in the lexical domain is also present, to a large extent, in the functional properties of most grammatical morphemes. I will take only two examples here: the Negative Existential; and Time Focus aspect.

5.4.1 *The Negative Existential*

All the northern Vanuatu languages possess a word with similar semantics and grammatical properties, which I will gloss as 'Negative Existential' (NEG:EXIST).

First, this word shows the same grammatical features everywhere. It is a member of a closed class of "directly predicative words", which is distinct from verbs. Like other major parts of speech (e.g., verbs, adjectives, and nouns) they can form the head of a predicate phrase, yet unlike them, these predicative words can do so directly, without any inflectional morphology. This small class of uninflected words patterns with other non-verbal predicates (François 2005b,c), such as locatives and demonstratives.

The isomorphism concerns not only the morphosyntactic behaviour of the Negative Existential, but also its semantic and phraseological properties. When it takes an overt noun phrase as its subject, NEG:EXIST states the absence of the referent in a given place (English 'there is no X'):

(16)
HIW	sːe	βɔtwu	*təɣɔ*		ŋʷutə.pənə
VLW	n	ɣasɛl	*tatɪh*		ɣɪs
MTA	o	ɣasal	*taɣai*		iak
LKN		ɣahæl	*ɣaiβ*		βaan.hɛɣ
	[ART]	knife	NEG:EXIST		here

'*There is no* knife here.'

When its subject is a possessed NP, the Negative Existential forms possessive statements with negative polarity, the equivalent of English 'have not':

(17)
HIW	ɣə-ma	pɪta	*təɣɔ*	jə-ŋʷə	=ma
LKN	ɣa-mæ	ʧam	*ɣaiβ*	l-uŋ͡mʷæ	mɔːmæ
	POSS:FOOD-1EXC:PL	yam	NEG:EXIST	LOC-house	POSS:1EXC:PL

'We *don't have any* yam in our house.' [lit. '*Our yam is lacking* in our house …']

With a human referent as its subject, the same predicate word will often translate as 'not be there, be absent':

(18)
LYP	iβɛp	*mamʧɛɣɛ*,	kɪ͡ɛ̀	m-	βan	lɪ	ʧɪk͡pʷɪ
MTP	tita	*tatɛh*,	kɪ	ma-	βan	lɪ-	tk͡pʷɪ
VRS	ⁿdiɛ	*ɔⁿd͡ià*ŋ,	nɪ	ma-	βan	lɛ	ŋɛŋɛrɛ
LKN	apʊ	*ɣaiβ*,	ni	n-	βan	lɪ	ŋ͡mʷɪhŋ͡mʷɪh
	Mum	NEG:EXIST	3SG	PRF-	go	LOC	garden

'Mum's *not here*, she's gone to the garden.'

The same word 'NEG:EXIST' is also often used absolutely (i.e., with no overt subject), with various meanings. For one thing, it forms the heavy equivalent of English 'no' in an answer to a yes/no question. In all the northern Vanuatu languages, a negative reply may consist of two parts:

- the 'light no', consisting of a prosodic gesture of the form [ɭ.ɻ.ʎ] uttered either on a schwa ə́ə̀ə̌ [ɭə.ɻə.ʎə] or on /ɔ/ ɔ́ɔ̀ɔ̌ [ɭɔ.ɻɔ.ʎɔ]
- the 'heavy no', which is the Negative Existential used absolutely.

(19)	Hiw	jəkʷen	ənɔmə	nenə?	–	ə̀ə̀ə̌,	*təɣɔ*!	kʷukʷuj	ŋʷutʉjə!
		woman	your	that		(no)	NEG:EXIST	friend	just
	Mtp		iɣni	nɛn?	–	ɔ̀ɔ̀ɔ̌,	*tatɛh*!	k͡pʷulk͡pʷul	ɪwɪ!
	Msn	ɛ	ɣunum	nɔ?	–	ɔ̀ɔ̀ɔ̌,	*ɛnɛŋ*!	pulsal	ɣɔp!
	Mrl		rɔnatɔŋ	kan?	–	ɔ̀ɔ̀ɔ̌,	*tɪɣɪ*!	ᵐbulsal	ɣɔm!
		[art]	wife.2sg	that		(no)	NEG:EXIST	friend	just

'Is that your girlfriend? –Oh no, *no*! She's just a friend.'

Another common usage of 'NEG:EXIST' in discourse is when one's effort is frustrated by a lack of result. In this case, the predicate would translate 'in vain, to no avail':

(20)	Lhi	lɪ	kʷʊŋ	nʊ	nɛ-	mjɪs	ⁿdɛ	k-	mutuj,	pɔ	*tɛtyɛ*!
	Vra	lɪn	k͡pʷʊŋ	nɔ	ɣa	mʊrʊs	sɔ	k-	miʔir,	ᵐba	*ɣitay*!
	Drg	lɪ	k͡pʷʊŋ	na	βa	mrʊs			ŋɔr,	la	*ᵐbɛk*!
		loc	night	1sg	stat	want	[comp	1sg]	sleep	but	NEG:EXIST

'Last night, I tried to sleep, but *no way*!'

Finally, a widespread formulaic pattern employs this word 'NEG:EXIST' in greetings. The literal meaning must be understood as '[there's] no [problem]', but it is routinely used to mean 'I'm fine':

(21)	Hiw	ŋʷe.na?	–	*təɣɔ*,	ne	məᵍʟawə	ŋʷutʉjə
	Lmg	mækæβæ.nɛ?	–	*niβ*,	ɣɛ	wi	k͡pʷɔɣɔr
	Vrs	timi͡ak.aβɪ?	–	*ɔⁿdi͡aŋ*,	ɣʊ-	wɪ	ɣɛm
	Mta	taŋ͡mʷa.βɛa?	–	*taɣai*,	we	wia	ɣap
	Drg	ⁿdaksa?	–	*ᵐbɛk*,	βa-	wɪ	wɔr
		how	–	NEG:EXIST	stat	good	just

'How's things? –*No* [*worries*]! Everything's alright.'

In sum, even though the form of the Negative Existential differs considerably between languages, its morphosyntactic properties as well as its constructional combinatorics are identical across the entire region.

5.4.2 *The Time Focus aspect*

Tense-Aspect-Mood (TAM) paradigms show a certain degree of variation across the area. Languages will sometimes differ in the way they divide the semantic space of TAM reference, some languages making contrasts where others do not. For example, the TAM category called "Aorist" (encompassing sequential, prospective, subjunctive, hortative, etc.) is found only in the northern part of the archipelago; in the southern Banks languages, it is broken down into several categories (François 2009a). However, even this domain shows some areas of cross-linguistic homogeneity, since certain TAM categories are found to be very similar across the entire northern Vanuatu group. Thus all languages possess a distinct modal

category of Apprehensive (cf. Lichtenberk 1995), they all distinguish Perfect vs. Complete or Future vs. Potential, and so on.

One particularly idiosyncratic aspectual category, which can be found everywhere in the area despite the variety of its forms, is one I suggest calling "Time Focus" (TmFoc). Its underlying semantic mechanism (François 2003a: 199–216) consists in anchoring an event in time by establishing an explicit contrast with earlier possible time anchorings. In other words, a general gloss for this TAM marker would be '[do] then, and not earlier'.[11]

As often happens in these languages which encode aspect but not tense, Time Focus is compatible with any time anchoring, whether past, present or future. This accounts for the polyfunctionality of this morpheme. When narrating past events, Time Focus may be used to anchor an event contrastively in the timeline:

(22)	Hiw	paβen	sə^{g}lə	*takə*	pʉn
	Vra	sʊwlɪ	^{n}duru	*mak*	βɛryʊl
	Drg	ti	raːr	*k͡p^{w}ra*	βlala
	Lkn	tɔ	wʊrʊ	*lak*	rulɛ
	Mrl	lan	karar	*k^{w}ɛr*	rɔs
		then	3du	TmFoc	argue
		'… So *that's when* they (started to) argue.'			

If the clause includes a time adverb referring to the past (e.g., 'yesterday') or the present ('now'), Time Focus will receive an inaugurative interpretation, i.e., 'do sth for the first time':

(23)	Hiw	ninə	*takə*	ni	nə	ɣa	pə	nənɔ^{g}lə
	Lmg	ti	*mak*	sun	n-	ɣa		lɪ.nɔr
	Mtp	kɪ	*k͡p^{w}ɔjɔ*	in	na	ɣa		anɔj
	Olr	ni	*tak*	un		k͡p^{w}ɔɣɔɣɔ		nanɔnɔ
		3sg	TmFoc	drink	[art]	kava	[foc]	yesterday
		'He drank kava yesterday *for the first time*.'						

A construction which is shared by all languages in the region combines the Time Focus aspect marker with a perfective postclitic, plus (optionally) the Restrictive postverb 'just'. This combination forms the semantic equivalent of an immediate perfect:

(24)	Ltg	n′	*akə*	pah	wərəŋɔ	si
	Mtp		*k͡p^{w}ɔjɔ*	^{m}bah	ɪwɪ	tʊ
	Lmg		*mak*	k͡p^{w}ɪʔ	k͡p^{w}ɔɣɔr	ʔi
	Vrs		*kara*	k͡p^{w}ɪt	ɣɛm	ti
	Lkn		*lak*	pah	wɔː	tɔ
		[3sg]	TmFoc	finish	just	prf
		'It *just* finished.'				

This construction combines the aspectual operation of Time Focus ('then and not earlier') with the semantics of the perfect, as well as with the pragmatic orientation of the restrictive towards a low quantity ('only' > 'just a moment ago').

Finally, when used in an irrealis context, TmFoc works as a sort of delayed future, thereby drawing a subtle argumentative contrast with the ordinary Future (François 2003a: 205). While both markers present an event as bound to happen (e.g., 'we *will* play, I promise you'), Time Focus is used as a DILATORY FUTURE. That is, it places the emphasis on the necessary delay between the moment of utterance and the beginning of that event in a way consistent with its core semantic mechanism 'then, and not earlier'.

(25)	Hiw	titə	*takə*	tɔɣəkesə,	pa	tɔkʷɛ
	Lyp	i͡ɛn	*pɛj*	wulwuli͡ɛs,	pɛ	ʧɛk͡pʷɛ
	Vra	ɣiⁿdɪ	*mak*	ɔrɔrɔ,	ᵐba	ɣitaɣ.ʔɪn
	Vlw	ɣɪn	*ᵑg͡bʷaja*	sisɛɣ,	ᵐba	maᵑgʊh
	Lkn	ɣɪʧ	*lak*	ˀɔːˀɔː,	la	tamalɛ
	Mrl	ɣe͡an	*kʷɛr*	ɔrɔr,	nɛn	tɪkʷɪte͡a
		1inc:Pl	TmFoc	play	but	not.yet
		'Yes, we *will* (*eventually*) play, but not yet.'				

As these examples show, the northern Vanuatu languages all share the same idiosyncratic aspect category of "Time Focus", both in the fine-grained detail of its semantic polyfunctionality and in its various combinatorial abilities.

5.5 Summary: "One grammar, 17 lexicons"

The examples cited in the preceding sections illustrate the various forms of structural isomorphism which can be found across the languages of northern Vanuatu: parallel word order, identical polysemies and semantic categorizations, and similar phraseological strategies.

Such a survey could potentially be expanded infinitely, considering that a vast proportion of these languages' structures show isomorphism, whether "perfect" or "quasi-perfect" (Section 5.1). Of course there are always exceptions, and careful observation unveils subtle differences across languages; for example, the fine-grained semantic detail of certain TAM markers, or of possessive classifiers, seldom matches exactly across the archipelago. Also, while some patterns of polyfunctionality encompass all the Torres and Banks Islands, other isoglosses are restricted to smaller areas. Ideally, the structural distances between the northern Vanuatu languages could be measured in a way similar to the percentages of lexical distance in basic vocabulary (Section 4.4.1), but such an endeavour would go beyond the limits of this study. At this stage, I can at least offer an impressionistic

figure based on my observations, and estimate that any two languages of the Torres-Banks area would exhibit 80 to 85% structural isomorphism in their utterances. The closer (geographically or socially) the two language communities, the higher this figure would be.

One finds the same distribution of meaning into functional categories and constructions everywhere, even though the boxes are filled with heterogeneous phonological forms. One could almost propose to summarize the situation using a bold formula, and say that these 17 languages represent "one grammar with 17 lexicons", in a way reminiscent of Friedman's (1997) paper "One grammar, three lexicons", about language contact in the Balkan area. Of course, this would be slightly exaggerated, since each of these languages also has its own idiosyncrasies and innovations; but it would capture quite well the paradox under discussion here. Also, such a formula would have to be taken with caution due to its deceptive terminology, as the contrast is really orthogonal to the traditional divide between 'grammar' and 'lexicon' (Section 5.1). What is observed is rather a strong diversification in the phonological forms of words (regardless of whether they belong to the 'lexicon' or the 'grammar') vs. a strong isomorphism between the structural properties of these same words (whether syntactic or semantic properties).

Now that these empirical facts have been established, it is time to investigate the motivations for such a complex situation. This is the object of the final section.

6. Divergence and convergence: Two forms of diffusion

6.1 Explaining structural isomorphism

Among the two major trends at play in northern Vanuatu, the ISOMORPHISM is probably easier to account for. The main force at work here is language contact, and the mutual influence that neighboring languages exert on each other.

6.1.1 *Structural convergence in multilingual contexts*

We know (Sections 2.2.1 and 3.2) that northern Vanuatu language communities have always been in contact with each other. They have maintained relations of multilingualism through their social relations, especially sustained patterns of exogamy. The fact that many children in the archipelago are raised bilingually, or by a mother who is herself bilingual, increases the rate of interference whereby languages tend to constantly reshape each others' structures. This multilingualism is evidently the key to the structural homogeneity observed in the area. Haudricourt (1961:9), in his discussion of the New Caledonian situation, also proposed

correlating patterns of grammatical isomorphism with the social configuration he calls "egalitarian bilingualism".

Many other cases around the world have been described where distinct languages in contact have developed parallel structures, sometimes including perfect morpheme-by-morpheme correspondences, despite their words' different origins. To take just a few examples, patterns of structural convergence have been observed between various branches of Indo-European in the Balkans (Weinreich 1953, Thomason & Kaufman 1988), between Indo-Aryan and Dravidian languages in southern India (Gumperz & Wilson 1971), between various linguistic families in southeast Asia (Enfield 2001, 2003) or America (Campbell et al. 1986, Bossong 2009), or between the Austronesian and Papuan languages of New Guinea (Thurston 1987). The latter area of contact also gave rise to Ross's (2001) concept of METATYPY, which describes the processes of structural influence which one language can exert upon another in multilingual situations. To these examples one can add the entire field of pidgin and creole studies, which regularly unearth patterns of structural parallelism between pidgins or creoles and their substrate languages (Keesing 1988, 1991, Lefebvre 1998, 2009, Siegel 2000, 2008), or research in second-language acquisition, which consistently shows how the structures of a first language L1 influence the acquisition of a second language L2 (Odlin 1989, 2003, Doughty & Long 2003, Winford 2003). However different they may be in their social circumstances, all these contact situations typically follow a similar tendency, which Sasse (1985, in Ross 2001:149) summarized in an enlightening formula: "With advanced language contact, there arises the tendency to develop a single language with different vocabularies."

These processes of structural convergence are ultimately rooted in cognitive pressures in multilingual situations. Structural convergence arises when multilingual speakers in contact give in to a "trend towards word-for-word translatable codes" (Gumperz & Wilson 1971:165). Achieving cross-linguistic structural isomorphism makes it cognitively easier for speakers to learn and speak other language varieties. Also, by limiting the risk of semantic loss in translation, such parallelism increases the chances for successful communication events. Individuals who learn or practice a language distinct from their own are tempted to "transfer" (Odlin 1989) the structures of the language they are most familiar with to the new system they are attempting to use. In doing so, they tend to "relabel" (Lefebvre 2008) their native categories using forms from the target language. This classic model provides the explanation for the ongoing processes of convergence observed in northern Vanuatu.

6.1.2 *Convergence or retention?*

A particularity of the northern Vanuatu area, however, is that languages share the same ancestor. For this reason, it might be objected that some cases of structural isomorphism, rather than resulting from the areal spread of an innovation, really reflect the RETENTION of an inherited construction or structural configuration; this configuration would have been transmitted over time in spite of lexical replacement. This is clearly a possibility, which can sometimes be confirmed through external comparisons (François 2010b). For example, each Torres-Banks language has a noun, different in each language, which colexifies 'tree stump, tree base' with 'beginning', 'origin', 'source' or 'cause'. The same polysemy is found in most Austronesian languages (Fox 1995: 5); it is reconstructed by Blust (1995) for Proto Malayo-Polynesian as **puna(ŋ)*, and by Evans (2008: 90) for POc as **puqun*. This is a case where languages may prove innovative in their forms, yet conservative in their semantic structures. The same polysemy found across the Torres-Banks languages is thus likely to be an inherited pattern. Strictly speaking, it could hardly be described as a case of areal CONVERGENCE, since it simply reflects shared retention; logically, language contact should be irrelevant in this case.

However, even in such a case where the structural isomorphism is probably inherited rather than innovative, I believe that language contact does play a role if and only if lexical replacement has taken place. Consider, for example, the domain of WATER, whose semantic outline is identical in all the Vanuatu languages: one noun colexifies '(fresh) water' with 'river' or 'lake', and is also a generic term for '(drinkable) liquid', but it cannot refer to 'salty water' or 'sea', which is a separate word. While the meaning '(fresh) water+' is lexified, in most Torres-Banks languages, with reflexes of a root **^{m}bei*, the language Lakon has replaced **^{m}bei* with an innovative form *tuŋ*, etymologically from **tuŋu* 'sap, juice of a plant or fruit'. The most likely "bridging context" (Evans & Wilkins 2000: 549) which can account for the word's semantic shift from 'juice' to 'fresh water' would be a generic meaning 'drinkable liquid'. At some point in history, pre-Lakon ended up having two synonyms for 'fresh water': the inherited form **^{m}bei*, and an innovative synonym **tuŋu*, presumably with some stylistically marked connotation, for example if 'juice' were used as a slang term for 'water'. In Table 7, this corresponds to a shift from STAGE A to STAGE B.

The shift from STAGE A to STAGE B illustrates lexical replacement via semantic shift well, which typically results in a change of semantic organization over time. Thus we saw earlier, in Table 4 (Section 4.4.2), that Latin initially colexified 'head' with 'leader' (*caput*), but eventually, after a process of semantic shift, lost that polysemy and acquired a new one 'head' ~ 'pot' (*testa*). Likewise, the semantic shift affecting **tuŋu* in pre-Lakon logically resulted in a new semantic organization corresponding to STAGE B, in which 'drinkable liquid' (**tuŋu*) and 'river, lake'

Table 7. When lexical replacement preserves semantic structures: Words for 'water' in Lakon

	'(fruit) sap, juice'	**'(drinkable) liquid'**	**'fresh water'**	**'river, lake'**	**'sea water, salt, sea'**
Stage A	*tuŋu	*^{m}*bei*	*^{m}*bei*	*^{m}*bei*	**nawo*
Stage B	*tuŋu	*tuŋu	*tuŋu~*^{m}*bei*	*^{m}*bei*	**nawo*
Stage C	tuŋuhwi:	*tuŋu	*tuŋu	*tuŋu	**nawo*

(*^{m}*bei*) were presumably lexified separately. But this is not what we observe in modern Lakon. Instead, some time after the meaning 'water' had been relexified using the innovative form **tuŋu*, this form ended up replacing the former *^{m}*bei* in each and every part of its initial polysemy. Streams and rivers, and even the large sulphuric lake of Gaua, are lexified as *tuŋ* in Lakon, despite its etymology as 'juice'. Not only has the new label *tuŋ* taken up all the senses formerly associated with *^{m}*bei*, but it has even lost its meaning 'juice', which is now lexified with a derived form (*tuŋuhwi:*); in parallel, Lakon has lost all traces of *^{m}*bei*. Stage C illustrates how Lakon now assigns to **tuŋu* exactly the same semantic outline (shaded area) as that which can be reconstructed for the other root *^{m}*bei*.

In a way, the structural properties of Lakon *tuŋ* constitute a form of "retention" from Stage A to Stage C. But considering how lexical replacement normally results in some form of semantic restructuring, this unexpected preservation of semantic structures needs to be explained. The paradox whereby a language can renew its word forms without renewing their semantic outline would be difficult to explain if Lakon had evolved in isolation from its neighbors: why should a language, after having gone through semantic change, retrieve polysemies of the past? Sometimes, such a result may be triggered by language-internal processes of analogy, provided the right conditions are present (Winters 1997). However, given the language ecology of northern Vanuatu, the most likely explanation here is probably simply language contact. After Lakon had innovated a new word **tuŋu* for the sense 'drinking water', its multilingual speakers felt the cognitive pressure to (re)align its semantic properties with the ones found in neighboring languages. This triggered the shift from Stage B to Stage C, and the generalization of **tuŋu* to all senses (and only those) formerly associated with *^{m}*bei*.

While the process ends up giving the illusion of structural conservatism, it can only be satisfactorily explained by positing a two-step process whereby languages first diverge, and later partially reconverge:

– Stage A > B: a language undergoes a lexical innovation that changes the form associated with a certain meaning, but also the structural organization of its semantic domain; this is a normal process of semantic change.

– Stage B > C: While the new form is adopted, its structural properties (semantic outline, combinatorics, syntactic behaviour) are subsequently realigned with a similar word from neighboring languages due to areal pressure. Sometimes, this realignment with neighboring languages happens to result in language-internal structural conservatism.

The same reasoning could be proposed to account for all other cases in which structural isomorphism comes along with lexical replacement, whether in the lexicon or the grammar (François 2010b: 137). I propose to name this two-step process language-internal relexification; it combines the preservation of a specific pattern of semantic organization with a change only in the phonological material associated with that meaning.

In sum, language contact does not only account for the spread of innovative structures. Even *conservative* structures, whenever they survived processes of semantic shift, are best explained by the pressure of language contact in a multilingual setting.

6.2 Language as an emblem of place

While contact-induced convergence accounts for the structural homogeneity of the northern Vanuatu languages, one still has to define a model capable of integrating it with what seems like a contradictory trend, namely the strong tendency toward language divergence. I will suggest that the attested diversity is the effect of an ideological bias widespread in Melanesia, which *de facto* fosters cultural and linguistic differentiation.

6.2.1 *A social bias for cultural differentiation*

In these parts of Melanesia, constant reference is made to the precise anchoring of things and people within the social and geographic space (Fox 2006). Individuals identify themselves, and each other, as belonging in a vast network of kinship relations, but also as tied to specific locations in the archipelago: to their island, their village, or even to a smaller area within their village. This constant awareness of the spatial anchoring of people and things has its corollary, namely the attention paid to whatever may *vary* between local communities. People are often heard commenting, with conspicuous pleasure and excitement, on particular manners of cooking, mat-weaving, dancing or singing, which may differ, sometimes quite subtly, between two social groups. However they arose in real life, those cultural differences are not so much assigned to contrasts in age or gender, or to individual performers, but tend to be formulated and remembered in terms of specific spatially-bound communities.

These two dimensions of cultural life in northern Vanuatu — the constant reference to places and the constant attention to local differences — design an idealized geography in which the world is conceived of as a mosaic of small-sized communities, as diverse in their cultural forms as their landscape can be rugged and variegated.

This magnetic attraction to cultural diversity has a linguistic manifestation. Should some novel way of speaking emerge in the speech of a few villagers, it can quickly spread to a whole village. And given the social bias towards the differentiation of local groups, there is a high likelihood that the new usage will become a trademark of this community, in contrast with its neighbors.

The *linguistic* principles underlying these changes are essentially universal, and the languages of Vanuatu follow the same processes as other languages. What makes them conspicuous is their high density within such a small human community. The high degree of language divergence observed in this region appears to be indirectly correlated with a prevailing *social* bias that tends to foster cultural and linguistic differentiation between local communities.

6.2.2 *Avoiding teleology*

The sociolinguistic explanation I propose here in order to account for the emergence of linguistic diversity in northern Vanuatu may be reminiscent (at least apparently) of previous studies on language differentialism in the same region or elsewhere.

Thus several scholars, trying to explain the high linguistic fragmentation of Melanesia, have proposed the idea that such a form of language change may be deliberate and conscious. For example, here is how Dutton (1995) summarizes the observations made earlier by Laycock (1982) on "conscious" linguistic change in Papua New Guinea: "Melanesians thus appear to foster linguistic diversity *purposefully* because they see linguistic differences as important badges of group identity. It is, as it were, a Melanesian choice to promote diversity" (Dutton 1995: 219; my emphasis). Likewise, Ross (1997) borrows from Thurston (1987, 1989, 1994) the notion of ESOTEROGENY; he defines it as a process whereby communities increase the complexity of their language, so as to develop "an 'esoteric' lect from which outsiders are *consciously* excluded" (Ross 1997: 239; my emphasis). Finally, Thomason (2007: 52) describes similar processes elsewhere in the world, which she explicitly characterizes as "deliberate change".

Can this model apply to the situation in northern Vanuatu? This depends on the way these ideas are to be interpreted. The general notion that linguistic divergence can be explained, at least partly, by sociolinguistic factors, and in particular by the building process of in-group identities, is a point which can safely be granted (Labov 1963, 2001, Le Page & Tabouret-Keller 1985). However, the use of

such words as "deliberately", "purposefully" or "consciously" might lead to some misunderstanding.

It is true that the difference between languages or dialects in this part of the world is a social phenomenon of which individuals are highly aware, and about which they can even have ideological views. In various part of Vanuatu, I have often heard the same idea that linguistic opacity and fragmentation are valuable assets rather than problems, since they potentially form a protective shield for each social group. This being said, there is a fine line between acknowledging such ideological representations and providing them with the explanatory power to account for the emergence of linguistic change itself (cf. Silverstein 2001).

Indeed, the wording chosen by these authors suggests that the push for divergence may form the primary reason why languages change. This runs the risk of representing linguistic change in a teleological fashion, as though the speakers' conscious craving for a distinctive identity, or for linguistic isolation, were alone able to trigger the sort of phonological, lexical or morphosyntactic innovations whereby languages diverge from each other. Campbell (2004b: 127) rightfully highlights how difficult it would then be to identify exactly, for a given language, which instances of linguistic change would result from such a "separatist motive" as opposed to "changes which just happen with no such motive".

It is thus desirable to design a model of language change which can take into account such factors as speakers' social attitudes, yet manage to avoid the snare of teleology. I'd like to propose that DIFFERENTIALISM is never the primary motivation for language change; the fundamental reason why innovations diffuse is rather due to speakers' desire to *converge* on particular speech habits. The linguistic differentiation between groups is just an indirect and epiphenomenal effect of that fundamental push for in-group homogeneity, and of the socially-accepted size of such groups. This proposal makes it necessary to redefine the contrast between convergence and divergence.

6.3 The social group as a focus of convergence

6.3.1 *Divergence and convergence: A single process*

As a first approximation, the terms DIVERGENCE and CONVERGENCE portrayed the apparent contradiction which emerge from the data well, as though the mutual relations between languages move in two opposite directions at once. But while this formulation is a valid metaphor, it may be insufficient to understand the precise mechanism at stake here. Now that I have discussed the linguistic facts behind both trends, I would like to propose another way of framing the paradox.

We know that languages "converge" with each other due to contact and diffusion (Section 6.1). On the other hand, we saw earlier that the various ways in

which languages appear to "diverge" from each other really result from an accumulation of linguistic innovations, each of which spread to some portion of the dialect chain. In other words, both cases of convergence and of apparent divergence ultimately involve processes of social diffusion. The difference between them is a matter of geographical scope for these patterns of diffusion — large scope in the case of convergence phenomena vs. restricted scope in the case of divergence.

For some innovation to emerge and spread across a certain linguistic territory necessarily means that some new way of speaking (an innovative pronunciation, a new metaphor, or a new lexical collocation) was first used by an individual or a group of individuals, and was then felt worthy of imitation by other people who heard it. As more and more people came to adopt this new speech habit, the innovation diffused to larger social groups, moving across age classes, across villages, or across entire networks. This process has been dubbed "linguistic epidemiology" by Enfield (2003, 2008), and "Propagation" by Toulmin (2009: 23). Once a given innovation has settled down to a certain group, it draws an invisible line between two sets of speakers, one circle within which individuals have converged together in adopting the new form vs. the rest of the population, for whom this innovation means more linguistic divergence. In other words, what may appear to be an effect of divergence on the macro scale (between different dialects) is first and foremost the result of events of convergence on the micro scale (between individual speakers); see Figure 12.

The term "group" used in Figure 12 is intended as a neutral term likely to refer to any social network of any size. It may refer to a family, a village, a group of villages, an island, or an entire archipelago; it may be specific in age, gender, or any other dimension of social significance. Each unit is here simply defined as the

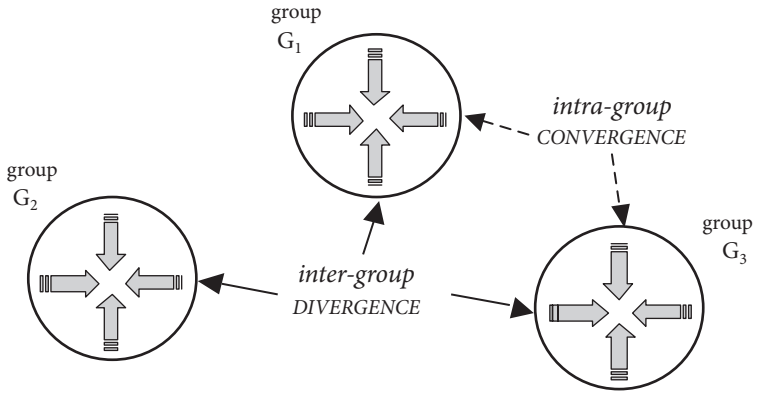

Figure 12. Global divergence results from events of local convergence: Individual speakers adopt local innovations within their social network, thereby increasing the gap with neighboring groups

social group which has adopted a given linguistic innovation; it corresponds to Toulmin's (2009: 27) notion of "Propagation Network", whose outline can differ for each propagation event.

Crucially, the same diagram (Figure 12) fits equally with the cases I described earlier as "divergence" and those which were understood as "convergence". Whether one observes the processes of sound change, or the adoption of certain phraseological patterns, ultimately what we observe is that some social network of some size has converged in adopting some new speech habit, and that this shared innovation has incidentally resulted in the differentiation between this group and the others around it which have not adopted this speech habit.

The way dialects have differentiated themselves from each other was by adopting certain innovations which their neighbors did not adopt. The typical case was that lexical or phonological innovations would encompass only a subset of the archipelago, and stop somewhere between two dialects. Given a chain of adjacent dialects A, B, C, D and E, an innovation in dialect C might spread to B and D, but not go any further. While this particular change did sanction the social bonds between communities B–C–D, it also highlighted the perceived difference between communities A and B, or between D and E. Later, another innovation would arise in dialect B and spread to A but not C, thereby splitting B–C–D into different groups again. Compare, for example, the maps in Figures 7–11. The more innovations accumulate over time, each with its own spatial distribution, the more different each member of the earlier dialect chain would become with respect to its immediate neighbors. Eventually, the distribution of intersecting isoglosses has resulted in language diversity. However, the very process whereby each individual innovation diffused across space was in itself not a process of DIVERGENCE, but of (MICRO-)CONVERGENCE between speakers.

6.3.2 *The critical size of social groups*

The transitional phase of competition between old and new forms is the moment when social attitudes can exert their maximal influence. However, rather than portraying northern Vanuatu, or Melanesia as a whole, as a region of "deliberate" differentiation, it might be more accurate to describe it as a cultural area where each stable social group, recognised with its distinctive — and internally homogeneous — identity, can be extremely reduced in size.

Indeed, a parameter which differs considerably across cultures is the minimal dimension of the social unit which may constitute a stable focus of linguistic convergence (Nettle 1999: 139). In societies with centralised political structures and immense social networks — as in Polynesian kingdoms (Pawley 1981, 2007a), or in modern Europe — an emergent change will only really start to stabilize in mainstream varieties once it has managed to spread to vast groups of speakers. Should

an innovation be restricted to just a small local group, it might survive for a while but then eventually fade out, under the wide-scale pressure towards levelling and uniformity.

Conversely, in the islands of northern Vanuatu, as well as in most parts of island Melanesia and elsewhere, the socially accepted size of an identifiable community can be much smaller (Pawley & Ross 1995: 66, Pawley 2006, Lynch, Ross & Crowley 2002: 92, Evans 2010: 14). The typically pertinent circle within which individual speakers will tend to imitate each others' linguistic behavior will coincide with the perceived limits of a local community — a handful of adjacent villages in daily interaction. In this part of Melanesia, social actors are highly aware of their environment's linguistic geography, and the notion of "language" as a discrete notion is meaningful to them (*pace* Mühlhäusler 1996: 5). The social unit encompassing speakers of "the same language" typically forms a FOCUS OF CONVERGENCE for its members (François 2001: 16, cf. Enfield 2003: 20). It is the primary circle within which individuals, in order to optimize the success of communication with their day-to-day peers, will most likely feel the pressure to imitate each others' ways of speaking. Should one innovation emerge as a new linguistic usage in their community, the chances are great that they might quickly adopt it; it will only take a few hundred speakers adopting it before it characterizes the speech of an entire community. In principle, nothing should prevent local innovations from spreading further to neighboring dialects or languages, especially considering the amount of areal contact and multilingualism traditional in the region. But crucially, this expansion to other communities is not a necessary step for the new usage to become socially stable. Should it remain confined to a single village and be lacking from the speech of its neighbors, then this will not threaten its survival; on the contrary, its emblematic power may precisely facilitate its social stabilization within a local group. The potential drawback in terms of intervillage communication would in any case be offset by people's willingness to learn the tongues of their neighbors (Section 2.2).

Speakers do not deliberately bring about language diversity. In a way reminiscent of Adam Smith's "Invisible Hand" (cf. Keller 1989, 1994), the linguistic kaleidoscope of modern Vanuatu is best understood as the unplanned outcome of multiple micro-events of CONVERGENCE between individual speakers, within the invisible boundaries of what they spontaneously identify as their most salient social group.

6.4 Forms vs. structures: A difference in awareness

The paradox raised by the northern Vanuatu languages, which was initially formulated in terms of DIVERGENCE vs. CONVERGENCE, can be rephrased in terms

of the size of social networks. While each propagation event outlines a different social group, the first sections of this study have shown a correlation between the size of the network and the nature of the linguistic innovation (see Figure 11 as an illustration):

- Those innovations which have to do with structural properties of signs typically diffuse across larger networks
  [effect on the macro scale = convergence]
- Those innovations which affect the form of signs typically diffuse across smaller networks
  [effect on the macro scale = divergence]

These observations entail the necessity of distinguishing between the form of words and their structural properties as two different components of language (Section 5.1), ultimately capable of evolving in opposite directions. Arguably, the reason why linguistic forms tend to differ among small-scale communities is that the phonological form of words is, it seems, more salient for awareness. In speakers' meta-discourse, the distinctive identity of each particular language is typically defined by the form of its words rather than by their structural properties, which appear to be less accessible to spontaneous representations.

The asymmetry between these two components explains why they do not respond to social attitudes in the same way. On the one hand, word forms undergo a strong pressure to be preempted for their emblematic function, since they signal the distinctive identity of a specific language community. In contrast, the structural or typological properties of linguistic systems are not invested with this emblematic function. Consequently, this structural component is left free to comply with constraints on ease-of-processing, which multilingual speakers tend to resolve through the mutual alignment of their systems.

Impressed by the situation of structural convergence described by Gumperz & Wilson (1971) in India, Grace (1981) arrived at a similar conclusion when he proposed distinguishing between the structural organization of languages (his "content form" — see Section 5.1) and the phonological material attached to these structures (his "lexification"):

> [T]he two components of language — content form and lexification — can be distinguished. They 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. […] If the language serves as the emblem of the group, it is the lexification in turn which is the emblem of the language. (Grace 1981: 30–31)

The situation Grace was referring to was a case of contact-induced convergence between Dravidian and Indo-European languages in the village of Kupwar. The contrast in word forms between these genetically unrelated languages was present from the start, and had only later taken up the secondary function of social emblematicity as the two communities historically came into contact. What the Vanuatu case shows is that a similar asymmetry between heterogeneity of forms and homogeneity of structures can also arise spontaneously in communities that once spoke the same language. Where linguistic diversity does not exist, it can be invented.

6.5 Issues for further research

The social and linguistic ecology of northern Vanuatu combines a number of specific ingredients: scattered habitat in a landscape of mountainous islands, stateless social organization, egalitarian multilingualism, an economy based on land ownership and horticultural practices, preference for village exogamy, related languages sharing a single ancestor (proto-language) as well as a shared history of contact (linkage), and ideological emphasis placed on the value of local identities. But while this region may be extreme in some ways, there is no reason to believe that the social and linguistic processes at work there should be unique. Even though the present article chooses to avoid any universal claim, future debate and research should help assess its broader relevance.

Some aspects of the situation described for the Banks and Torres Islands presumably apply to a wider regional area, say Melanesia as a whole. Other parts of the world are also characterized by strong language fragmentation, and it would be useful to know which factors are shared with the ones observed in Vanuatu. More broadly again, some processes described here might have universal relevance, such as the proposed correlation between the culturally accepted size of distinct social groups and the scope of certain forms of linguistic diffusion.

A question for further studies would be to tease out the different contributions of the various parameters that make up the social ecology of northern Vanuatu, in bringing about a specific linguistic configuration.

For example, do diffusional areas evolve differently in island settings compared to continents? Is language diversification stronger in situations of demic separation than in contact? How are linguistic convergence and divergence affected by economic dimensions (self-sufficiency vs. mutual dependency among communities, hunter-gatherers vs. horticulturalists, etc.) or political factors (egalitarian vs. stratified societies, or polycentric social networks vs. centralized states)?

Is the value given to ethnicity or local identities (assuming this can be empirically identified) a key factor in language diversification? What is the exact role played by speakers' awareness of their own linguistic practices (cf. Silverstein

2001), and by conscious ideological discourse on homogeneity vs. heterogeneity? What configuration obtains when the most salient dimensions for group identities are not so much the anchoring in space as the anchoring in social class, gender, or urban communities?

In the various parts of the world exhibiting a similar configuration to the one described here (structural homogeneity, heterogeneity of word forms), what points of variation can be observed? For instance, while the phonological systems of northern Vanuatu rather stand on the diverging side (Section 4.2), other contact areas show the reverse tendency, namely the alignment of phonologies.

Finally, how do processes of post-dispersal diffusion across dialect networks affect the methodology of linguistic subgrouping (cf. Ross 1988, Pawley 1999, Aikhenvald & Dixon (eds) 2001, Bossong 2009, Toulmin 2009, Heggarty et al. 2010)? It is desirable to define a mode of representation that can address the shortcomings of the family-tree model and accurately capture the history of these social networks as it can be reconstructed based on the comparative method.

While many of these issues are already the object of ongoing debates, it is hoped that this study of the northern Vanuatu languages can shed a new light on processes of language change and their social underpinnings.

7. Summary

The paradoxical relations that can be observed among the modern languages of northern Vanuatu reflect the intricacies of their social ecology. In particular, the contradiction between their two evolutionary trends reflects the two major concentric social circles to which all individuals belong.

Under traditional circumstances, the day-to-day social unit which is most relevant for an individual (and also the most salient in people's representations and social practices) is the village, or the group of nearby villages which constitute a stable language community. This level of social organization constitutes a privileged focus of convergence for those linguistic components that are more salient to awareness, namely the form of words. The spatial distribution of these communities is mirrored in the 17 distinct languages which comprise the linguistic mosaic of the archipelago and differ mostly in their phonologies and vocabularies.

But at the same time, whether they realize it or not, individuals are also entangled in a much larger network of social relations that transcends the dimension of their village. This vast web of acquaintances and kinship, tirelessly spun through generations of cross-community contact, intermarriage, and cultural exchange, spreads from village to village, from island to island, and ultimately encompasses the entire Torres and Banks archipelago. While these languages' lexical idiosyn-

crasies reflect people's fundamental attachment to their local communities, ultimately the strong parallelism found among their structures reveals the profound coherence of their peaceful universe.

Abbreviations

Languages

The abbreviations for language names appear in Figure 1 and are repeated below.

Bsl	Bislama	Ltg	Lo-Toga	Olr	Olrat
Drg	Dorig	Lyp	Löyöp	Vlw	Volow
Hiw	Hiw	Mrl	Mwerlap	Vra	Vera'a
Kro	Koro	Mta	Mota	Vrs	Vurës
Lhi	Lehali	Mtp	Mwotlap	POc	Proto-Oceanic
Lkn	Lakon	Msn	Mwesen	Pncv	Proto-North Central Vanuatu
Lmg	Lemerig	Num	Nume	PTB	Proto Torres–Banks

Interlinear glosses

Example sentences are glossed according to the Leipzig rules. More specific abbreviations are listed below.

appell	appellative	poss	possessive classifier or linker
appreh	apprehensive	pot	potential
art	article	prf	perfect
asp	aspect	proh	prohibitive
comp	complementiser	prosp	prospective
cplt	complete aspect	Pssd	article for inalienably possessed nouns
exist	existential predicate	quot	quotative particle
food	possessive classifer for food	red	reduplication
hum	human article	stat	stative
ipfv	imperfective	TmFoc	time focus aspect
irr	irrealis	top	topicaliser
mix	mixed gender	unp	article for non-possessed nouns

Notes

* Various aspects of this research were presented at the international workshops *Ecology and Language Evolution* of Lacito-CNRS (2008, Paris) and *Grammatical Constructions in Time and Space* (2008, Oslo); at the 7th conference of the *Association of Linguistic Typology* (2007, Paris); and the annual meeting of the Australian Linguistics Society (2009, Melbourne). On these occasions, I found of great help the questions and comments from my colleagues. I am also indebted to several people for their useful comments on earlier drafts of this paper: Mark Donohue, Beth

Evans, Rob Mailhammer, Martine Mazaudon, Alexis Michaud, Andy Pawley, Malcolm Ross, Claudia Wegener; I also thank Jóhanna Barðdal as well as two anonymous reviewers of *JHL*. Any remaining flaws and errors remain my own.

The collection of data took place during several trips to the Banks and Torres islands, between 1997 and 2007. This research would not have been possible without the moral support of the Vanuatu Cultural Centre, and the financial support of the Académie Française, of the Centre National de la Recherche Scientifique (LACITO), as well as a scholarship *Jeunes Chercheurs* granted by the French Ministère de la Recherche.

But more than anyone else, I would like to express my gratitude to the many individuals from Vanuatu who so generously shared their time with me, teaching me their beautiful languages and telling me their countless stories. To name but a few, my warmest thanks go to †William Hagēt, Moses Meywēlgen, Richard Woris Lerig, Edgar Howard (MTP); René Wolman (VLW); Fred William Qasvarong (LYP); Stanley Lengson (LHI); †Wolta Robin, †Taitus Sërortelsöm (LMG); Fedrik Qarngi (VRA); Tevēt Mesigteltōk (MSN); Mama Isikiel Wurvegqiat, Alban Mofet (VRS), Eli Field (VRS, VRA, LMG); Mama Febaian Din (MTA); Antonet Rovanle (NUM); Mrānga Sales Maklēn (DRG), John Star (DRG, OLR); Derek Bogo, Banabas Womal (KRO); †Maten Womal (OLR); Moses Stiven Wēting, Nelson Vagēl (LKN); Mata Uli Rowon, Nikson Wevalēs (LKN, MRL); Nora Philip (MRL); Mama Brian, Elton Gëgelikwe, Eron Jefri Muglol (LTG); Sipo Ngwoypitvën, Stanley Elfrik Vëniwyōy, Sekōp Elison and Mama Jimmy Tiwyoy (HIW).

1. The 2009 census carried out by the *Vanuatu National Statistics Office* (VNSO 2009) gives a figure of 9,359 inhabitants for the province "TorBa" (Torres-Banks). This reveals a +20.7% increase from the figure of 7,757 inhabitants observed in 1999.

2. François (2005a) discusses the various regular patterns of sound change attested in the 17 languages. It provides tables of regular correspondences as well as principles for the reconstruction of protoforms in this area, based on the comparative method.

3. Various morphological changes also affected each language in ways different from its neighbors. This is the case, for example, with the accretion of a former NP article, whether *na or *wo (François 2007), to the noun stem, e.g., *na íra > *niri > VRA *niri*; *na uráŋi > *nuriŋ > OLR *nuriŋ*; *wo úra > *wur > NUM *wur*.

4. Here is what Codrington (1885: 323) wrote about the languages of "Bligh Island" (Ureparapara): "The practise of pronouncing *r* as *y* cannot be limited precisely; it is always followed at Retan, but in the Bay, children and some adults do it; *r* would at any rate be written."

5. The prenasalised /^{n}d/ in *ɣi^{n}da is an irregular development shared by all NCV languages (Clark 2009: 114).

6. Note that Tryon's statistics include neither Olrat nor Volow. Seeing as how Olrat can be often close to Koro (or to Lakon) and Volow very close to Mwotlap, it is possible that the inclusion of these two languages might have yielded slightly higher figures for maximal lexical cognacy. Conversely, Tryon distinguishes between "Merig" and "Merlav" (94.5% of cognacy), which I consider to be simply two varieties of Mwerlap (Section 2.1.1).

7. Excluding creoles, the lexically most remote languages in each group are: for Slavic, Slovenian and Byelorussian (61.0%); for Germanic, Faroese and English (53.6%); for Romance as a whole, French and Vlach (51.1%).

8. In the following examples, e.g., (3), (5) or (6), square brackets in the interlinear glossing will indicate morphemes that are found in only some languages of the sample.

9. I will return to the issue of structural convergence vs retention in Section 6.1.2.

10. The polysemy 'breathe; take rest' is not mentioned in Ross's reconstruction for POc *mañawa (Ross 1988: 461), yet it is found in the Hiw reflex *mənawə*. This could reflect areal influence from *ma^{m}busaɣi.

11. Mandarin Chinese has a similar aspect particle *cái* which can be glossed '[then and] only then' (De Francis 1966: 20, Li & Thompson 1989: 333).

References

Aikhenvald, Alexandra & Dixon, R.M.W. 2001. Introduction. In Aikhenvald & Dixon, eds., 126.

Aikhenvald, Alexandra & Dixon, R.M.W., eds. 2001. *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.

Ambrose, Wallace R. 1976. Obsidian and its Distribution in Melanesia. *Ancient Chinese Bronzes and South East Asian Metal and Other Archaeological Artifacts*, ed. by Noel Barnard, 351–378. Melbourne: National Gallery of Victoria.

Babel, Molly, Andrew Garrett, Michael J. Houser, & Maziar Toosarvandani. Forthcoming. Descent and Diffusion in Language Diversification: A Study of Western Numic Dialectology. *International Journal of American Linguistics*.

Bedford, Stuart & Matthew Spriggs. 2008. Northern Vanuatu as a Pacific Crossroads: The Archaeology of Discovery, Interaction, and the Emergence of the "Ethnographic Present". *Asian Perspectives* 47:1.95–120.

Bellwood, Peter S., James J. Fox, & Darrell Tryon, eds. 1995. *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University.

Best, Simon. 1987. Long-Distance Obsidian Travel and Possible Implications for the Settlement of Fiji. *Archaeology in Oceania* 22.31–32.

Blank, Andreas, & Peter Koch. 1999a. Onomasiologie et Étymologie Cognitive: L'exemple de la *tête*. *Actas do 1º Encontro Internacional de Linguística Cognitiva* ed. by Mario Vilela & Fatima Silva, 49–72. Porto: Faculdade de Letras.

Blank, Andreas & Peter Koch, eds. 1999b. *Historical Semantics and Cognition*. Berlin: Mouton de Gruyter.

Blust, Robert. 1981. Linguistic Evidence for Some Early Austronesian Taboos. *American Anthropologist* 83:2.285–319.

Blust, Robert. 1987. Lexical Reconstruction and Semantic Reconstruction: The Case of Austronesian 'House' Words. *Diachronica* 4.79–106.

Blust, Robert. 1995. Austronesian Comparative Dictionary (ACD). Unpublished computer files. University of Hawai'i.

Blust, Robert. 2010. Five Patterns of Semantic Change in Austronesian languages. *A Journey through Austronesian and Papuan Linguistic and Cultural Space: Papers in Honour of Andrew K. Pawley* ed. by John Bowden, Nikolaus P. Himmelmann & Malcolm Ross, 525–546. Canberra: Australian National University.

Bonnemaison, Joël, Kirk Huffman, Christian Kaufmann & Darrell Tryon, eds. 1996. *Arts of Vanuatu*. Bathurst: Crawford House Press.

Bossong, Georg. 2009. Divergence, Convergence, Contact: Challenges for the Genealogical Classification of Languages. In Braunmüller & House, eds., 13–40.

Braunmüller, Kurt, & Juliane House, eds. 2009. *Convergence and Divergence in Language Contact Situations*. *Hamburg Studies on Multilingualism 8*. Amsterdam: John Benjamins.

Camden, William. 1979. Parallels in Structure and Lexicon and Syntax between New Hebrides Bislama and the South Santo Language Spoken at Tangoa. *Papers in Pidgin and Creole Linguistics, No.2*, 51–117. Canberra: Australian National University.

Campbell, Lyle. 2004a. *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Campbell, Lyle. 2004b. Historical Linguistics: The State of the Art. *Linguistics Today: Facing a Greater Challenge* ed. by Piet van Sterkenburg, 109–140. Philadelphia: John Benjamins.

Campbell, Lyle, Terrence Kaufman & Thomas Smith-Stark. 1986. Meso-America as a Linguistic Area. *Language* 62:3.530–570.

Chappell, Hilary. 2001. Language Contact and Areal Diffusion in Sinitic Languages: Problems for Typology and Genetic Affiliation. In Aikhenvald & Dixon, eds., 328–357.

Clark, Ross. 1985. Languages of North and Central Vanuatu: Groups, Chains, Clusters and Waves. *Austronesian Linguistics at the 15th Pacific Science Congress* ed. by Andrew Pawley & Lois Carrington, 199–236. Canberra: Australian National University.

Clark, Ross. 1994. The Polynesian Outliers as a Locus of Language Contact. *Language Contact and Change in the Austronesian World* ed. by Tom Dutton & Darrell Tryon, 109–140. Berlin: Mouton de Gruyter.

Clark, Ross. 2009. **Leo Tuai: A Comparative Lexical Study of North and Central Vanuatu Languages*: *Pacific Linguistics 603*. Canberra: Australian National University.

Codrington, Robert. 1885. *The Melanesian Languages*. Oxford: Clarendon Press.

Codrington, Robert. 1891. *The Melanesians: Studies in Their Anthropology and Folklore*. Oxford: Clarendon Press.

Crowley, Terry. 1995. Melanesian Languages: Do They Have a Future? *Oceanic Linguistics* 34:2.327–344.

Crowley, Terry. 2000. The Language Situation in Vanuatu. *Current Issues in Language Planning* 1:1.47–132.

De Francis, John. 1966. *Advanced Chinese*. New Haven: Yale University Press.

Dik, Simon. 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause* ed. by A. Machtelt Bolkestein. Dordrecht: Foris.

Donohue, Mark. 2002. Which Sounds Change: Descent and Borrowing in the Skou Family. *Oceanic Linguistics* 41:1.171–221.

Doughty, Catherine, & Michael H. Long. 2003. *The Handbook of Second Language Acquisition*. Oxford: Wiley-Blackwell.

Dutton, Thomas Edward. 1995. Language Contact and Change in Melanesia. In Bellwood, Fox & Tryon, eds., 207–228.

Dyen, Isidore, Joseph B. Kruskal & Paul Black. 1992. *An Indoeuropean Classification: A Lexicostatistical Experiment*. Philadelphia: American Philosophical Society.

Enfield, Nicholas. 2001. On Genetic and Areal Linguistics in Mainland Southeast Asia: Parallel Polyfunctionality of 'Acquire'. In Aikhenvald & Dixon, eds., 255–290.

Enfield, Nicholas. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge-Curzon.

Enfield, Nicholas. 2008. Transmission Biases in Linguistic Epidemiology. In Nicolaï & Comrie, eds., 299–310.

Evans, Bethwyn. 2008. Parts of Plants. *The Lexicon of Proto-Oceanic: Plants* ed. by Malcolm Ross, Andrew Pawley & Meredith Osmond, 85–127. Canberra: Australian National University.

Evans, Nicholas. 2010. *Dying Words: Endangered Languages and What They Have to Tell Us*. The Language Library: Wiley-Blackwell.

Evans, Nicholas, & David Wilkins. 2000. In the Mind's Ear: The Semantic Extensions of Perception Verbs in Australian Languages. *Language* 76:3.546–592.

Fox, James J. 1995. Introduction. *Origins, Ancestry and Alliance: Explorations in Austronesian Ethnography* ed. by James J. Fox & Clifford Sather, 1–18. Canberra: Australian National University.

Fox, James J. 2006. *The Poetic Power of Place: Comparative Perspectives on Austronesian Ideas of Locality*. Canberra: Australian National University.

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

François, Alexandre. 2002. *Araki. A Disappearing Language of Vanuatu*. Canberra: Australian National University.

François, Alexandre. 2003a. *La sémantique du prédicat en mwotlap (Vanuatu)*. Louvain: Peeters.

François, Alexandre. 2003b. *Tog tog i van en: Navap tam̃ag non ige to M̆otlap* ('Once upon a time: Traditional stories from Motalava'). Paris.

François, Alexandre. 2005a. Unraveling the History of the Vowels of Seventeen Northern Vanuatu Languages. *Oceanic Linguistics* 44:2.443–504.

François, Alexandre. 2005b. A Typological Overview of Mwotlap, an Oceanic Language of Vanuatu. *Linguistic Typology* 9:1.115–146.

François, Alexandre. 2005c. Diversité des Prédicats Non Verbaux dans quelques Langues Océaniennes. *Les Constituants Prédicatifs et la Diversité des Langues: Actes de la Journée de la Société de Linguistique de Paris* ed. by Jacques François & Irmtraud Behr, 179–197. Louvain: Peeters.

François, Alexandre. 2007. Noun Articles in Torres and Banks Languages: Conservation and Innovation. *Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley* ed. by John Lynch, Jeff Siegel & Diana Eades, 313–326. Amsterdam: John Benjamins.

François, Alexandre. 2008. Semantic Maps and the Typology of Colexification: Intertwining Polysemous Networks across Languages. In Vanhove, ed., 163–215.

François, Alexandre. 2009a. Verbal Aspect and Personal Pronouns: The History of Aorist Markers in North Vanuatu. *Austronesian Historical Linguistics and Culture History: A Festschrift for Bob Blust* ed. by Andrew Pawley & Alexander Adelaar, 179–195. Canberra: Australian National University.

François, Alexandre. 2009b. The Languages of Vanikoro: Three Lexicons and One Grammar. *Discovering History through Language: Papers in Honour of Malcolm Ross* ed. by Bethwyn Evans, 103–126. Canberra: Australian National University.

François, Alexandre. 2010a. Phonotactics and the Prestopped Velar Lateral in Hiw: Resolving the Ambiguity of a Complex Segment. *Phonology* 27:3.393–434.

François, Alexandre. 2010b. Des Valeurs en Héritage: Les Isomorphismes Sémantiques et la Reconstruction des Langues. *Typologie et Comparatisme. Hommages offerts à Alain Lemaréchal* ed. by Injoo Choi-Jonin, Marc Duval & Oivier Soutet, 129–145. Louvain: Peeters.

François, Alexandre. 2011. Where *R They All? The Geography and History of *R-loss in Southern Oceanic Languages. *Oceanic Linguistics* 50:1.140–197.

François, Alexandre. Forthcoming. The Dynamics of Linguistic Diversity: Egalitarian Multilingualism and Power Imbalance among Northern Vanuatu Languages. *Language Use in Melanesia* ed. by Pete Unseth & Lynn Landweer. *International Journal of the Sociology of Language*.

Friedman, Victor. 1997. One Grammar, Three Lexicons: Ideological Overtones and Underpinnings of the Balkan Sprachbund. *CLS 33: Papers from the Panels on Linguistic Ideologies in Contact; Universal Grammar; Parameters and Typology; The Perception of Speech and other Acoustic Signals* ed. by Kora Singer, Randall Eggert & Gregory Anderson, 23–44. Chicago: Chicago Linguistic Society.

Garrett, Andrew. 2006. Convergence in the Formation of Indo-European Subgroups: Phylogeny and Chronology. *Phylogenetic Methods and the Prehistory of Languages* ed. by Peter Forster & Colin Renfrew, 139–151. Cambridge: McDonald Institute for Archaeological Research.

Geraghty, Paul A. 1983. *The History of the Fijian languages*. Honolulu: University of Hawaii Press.

Geraghty, Paul A. 1990. Proto-Eastern Oceanic *R and Its Reflexes. *Pacific Island Languages: Essays in honour of G. B. Milner* ed. by Jeremy H.C.S. Davidson & George Bertram Milner, 51–93. London: University of London, School of Oriental and African Studies.

Gilliéron, Jules & Mario Roques. 1912. *Études de Géographie Linguistique d'après l'Atlas Linguistique de la France*. Paris: Honoré Champion.

Grace, George W. 1981. *An Essay on Language*. Columbia, SC: Hornbeam Press.

Grace, George W. 1990. The "Aberrant" (vs. "Exemplary") Melanesian Languages. *Linguistic Change and Reconstruction Methodology* ed. by Philip Baldi, 155–173. Berlin: Mouton de Gruyter.

Gumperz, John J., & Robert Wilson. 1971. Convergence and Creolization: A Case From the Indo-Aryan/Dravidian Border in India. *Pidginization and Creolization of Languages* ed. by Dell Hymes, 151–168. Cambridge: Cambridge University Press.

Haudricourt, André-Georges. 1961. Richesse en Phonèmes et Richesse en Locuteurs. *L'Homme* 1:1.5–10.

Haugen, Einar I. 1972. The Ecology of Language. *The Ecology of Language: Essays by Einar Haugen*, ed. by Anwar S. Dil, 325–339. Stanford, CA: Stanford University Press.

Heggarty, Paul, Warren Maguire & April McMahon. 2010. Splits or Waves? Trees or Webs? How Divergence Measures and Network Analysis Can Unravel Language Histories. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.3829–3843.

Henri, Agnès. 2010. *Éléments de Description d'une Langue Mélanésienne du Vanuatu, le Suñwadia*. Paris: Université Paris-IV Sorbonne PhD dissertation.

Hjelmslev, Louis. 1961 [1943]. *Prolegomena to a Theory of Language*. 2nd Edition. Translated by Francis J. Whitfield. Madison, WI: University of Wisconsin Press.

Huffman, Kirk. 1996. Trading, Cultural Exchange and Copyright: Important Aspects of Vanuatu Arts. In Bonnemaison, Huffman, Kaufmann & Tryon, eds., 182–194.

Keesing, Roger. 1988. *Melanesian Pidgin and the Oceanic Substrate*. Stanford, CA: Stanford University Press.

Keesing, Roger. 1991. Substrates, Calquing and Grammaticalization in Melanesian Pidgin. *Approaches to Grammaticalization* ed. by Elizabeth C. Traugott & Bernd Heine, 315–342. Amsterdam: John Benjamins.

Keller, Rudi. 1989. Invisible-Hand Theory and Language Evolution. *Lingua* 77:2.113–127.

Keller, Rudi. 1994. *On Language Change: The Invisible Hand in Language*. London: Routledge.
Kirch, Patrick V. 1997. *The Lapita Peoples: Ancestors of the Oceanic World*. Cambridge: Blackwell.
Kirch, Patrick V. 2010. Peopling of the Pacific: A Holistic Anthropological Perspective. *Annual Review of Anthropology* 39.131–148.
Labov, William. 1963. The Social Motivation of Sound Change. *Word* 19.273–309.
Labov, William. 1991. The Three Dialects of English. *Quantitative Analyses of Sound Change* ed. by Penelope Eckert, 1–44. New York: Academic Press.
Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.
Laycock, Donald C. 1982. Linguistic Diversity in Melanesia: A Tentative Explanation. *Gava': Studies in Austronesian Languages and Cultures, Dedicated to Hans Kähler* ed. by Rainer Carle, Martina Henschke, Peter W. Pink, Christel Rost & Karen Stadtlender, 31–37. Berlin: Dietrich Reimer.
Lefebvre, Claire. 1998. *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge: Cambridge University Press.
Lefebvre, Claire. 2008. Relabelling: A Major Process in Language Contact. In Nicolaï & Comrie, eds., 91–111.
Lefebvre, Claire. 2009. The Contribution of Relexification, Grammaticalisation, and Reanalysis to Creole Genesis and Development. *Studies in Language* 33:2.277–311.
Le Page, Robert Brock & Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
Li, Charles & Sandra Thompson. 1989. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA: University of California Press.
Lichtenberk, Frantisek. 1995. Apprehensional Epistemics. *Modality in Grammar and Discourse* ed. by Joan Bybee & Suzanne Fleischmann, 293–327. Amsterdam: John Benjamins.
Lynch, John. 2000. Linguistic Subgrouping in Vanuatu and New Caledonia. *Proceedings of the Second International Conference on Oceanic Linguistics (SICOL), Vol. 2: Historical and Descriptive Studies* ed. by Bill Palmer & Paul A. Geraghty, 155–184. Canberra: University of the South Pacific.
Lynch, John & Terry Crowley. 2001. *Languages of Vanuatu: A New Survey and Bibliography*. Canberra: Australian National University.
Lynch, John & Françoise Ozanne-Rivierre. 2001. Some Shared Developments in Pronouns in Languages of Southern Oceania. *Oceanic Linguistics* 40:1.33–66.
Lynch, John, Malcolm Ross & Terry Crowley. 2002. *The Oceanic Languages*. Richmond: Curzon Press.
Mühlhäusler, Peter. 1996. *Linguistic Ecology: Language Change and Linguistic Imperialism in the Pacific Region*. London: Routledge.
Nettle, Daniel. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.
Nicolaï, Robert & Bernard Comrie, eds. 2008. *Language Contact and the Dynamics of Language: Theory and Implications. Journal of Language Contact*. Thema 2. http://www.jlc-journal.org
Odlin, Terence. 1989. *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
Odlin, Terence. 2003. Cross-Linguistic Influence. In Doughty & Long, eds., 436–486.
Pawley, Andrew. 1981. Melanesian Diversity and Polynesian Homogeneity: A Unified Explanation for Language. *Studies in Pacific Languages and Cultures in Honour of Bruce Biggs* ed. by Jim Hollyman & Andrew Pawley, 259–310. Auckland: Linguistic Society of New Zealand.
Pawley, Andrew. 1999. Chasing Rainbows: Implications of the Rapid Dispersal of Austronesian Languages for Subgrouping and Reconstruction. *Selected Papers from the Eighth Interna-*

tional Conference on Austronesian Linguistics ed. by Elizabeth Zeitoun & Paul Jen-Kuei Li, 95–138. Taipei: Institute of Linguistics, Academia Sinica.

Pawley, Andrew. 2003. Locating Proto Oceanic. *The Lexicon of Proto-Oceanic : The Physical Environment* ed. by Malcolm Ross, Andrew Pawley & Meredith Osmond, 17–34. Canberra: Australian National University.

Pawley, Andrew. 2006. Explaining the Aberrant Austronesian Languages of Southeast Melanesia: 150 Years of Debate. *Journal of the Polynesian Society* 115.213–256.

Pawley, Andrew. 2007a. Why do Polynesian Island Groups Have One Language and Melanesian Island Groups Have Many? Patterns of Interaction and Diversification in the Austronesian Colonization of Remote Oceania. Paper presented at *"Migrations"* conference, Porquerolles (Var, France).

Pawley, Andrew. 2007b. The Origins of Early Lapita Culture: The Testimony of Historical Linguistics. *Oceanic Explorations: Lapita and Western Pacific Settlement* ed. by Stuart Bedford, Christophe Sand & Sean P. Connaughton, 17–50. Canberra: ANU E-Press.

Pawley, Andrew & Roger C. Green. 1973. Dating the Dispersal of the Oceanic Languages. *Oceanic Linguistics* 12.1–67.

Pawley, Andrew & Roger C. Green. 1984. The Proto-Oceanic Language Community. *Journal of Pacific History* 19.123–146.

Pawley, Andrew & Malcolm Ross. 1995. The Prehistory of Oceanic Languages: A Current View. In Bellwood, Fox & Tryon, eds., 39–80.

Pawley, Andrew & Timoci Sayaba. 1971. Fijian Dialect Divisions: Eastern and Western Fijian. *Journal of the Polynesian Society* 80.405–436.

Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian Languages of Western Melanesia*. Canberra: Australian National University.

Ross, Malcolm. 1997. Social Networks and Kinds of Speech-Community Event. *Archaeology and Language 1: Theoretical and Methodological Orientations*, ed. by Roger Blench & Matthew Spriggs, 209–261. London: Routledge.

Ross, Malcolm. 2001. Contact-Induced Change in Oceanic Languages in North-West Melanesia. In Aikhenvald & Dixon, eds., 134–166.

Ross, Malcolm. In Prep. Reconstructing the History of Contact in Northwest New Britain. MS, Australian National University.

Ross, Malcolm, Andrew Pawley, & Meredith Osmond, eds. 2011. *The Lexicon of Proto-Oceanic, Volume 4: The Fauna*. Canberra: Australian National University.

Sasse, Hans-Jürgen. 1985. Sprachkontakt und Sprachwandel: Die Gräzisierung der albanischen Mundarten Griechenlands [Language Contact and Language Change: The Hellenization of the Albanian Dialects of Greece]. *Papiere zur Linguistik* 32:1.37–95.

Shing, Richard, Stuart Bedford, Matthew Spriggs & Ralph Regenvanu. 2007. *Ripot long saed blong stadi long akioloji long not blong Vanuatu (2005–2007)*. Port-Vila: Vanuatu National Library.

Siegel, Jeff. 2008. *The Emergence of Pidgin and Creole languages*. Oxford: Oxford University Press.

Siegel, Jeff, ed. 2000. *Processes of Language Contact: Studies from Australia and the South Pacific*. Montréal: Fides.

Silverstein, Michael. 2001. The Limits of Awareness. *Linguistic Anthropology: A Reader* ed. by Alessandro Duranti, 382–401. Malden: Blackwell.

Simons, Gary. 1982. Word Taboo and Comparative Austronesian Linguistics. *Papers from the Third International Conference on Austronesian linguistics, Vol. 1: Currents in Oceanic* ed.

by Amran Halim, Lois Carrington & Stephen A. Wurm, 157–226. Canberra: Australian National University.
Spriggs, Matthew. 1995. The Lapita Culture and Austronesian Prehistory in Oceania. In Bellwood, Fox & Tryon, eds., 119–142.
Sweetser, Eve E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural aspects of Semantic Structure*. Cambridge: Cambridge University Press.
Thomason, Sarah G. 2007. Language Contact and Deliberate Change. *Journal of Language Contact Thema* 1.41–62.
Thomason, Sarah G. & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
Thurston, William R. 1987. *Processes of Change in the Languages of North-Western New Britain*. Canberra: Pacific Linguistics.
Thurston, William R. 1989. How Exoteric Languages Build a Lexicon: Esoterogeny in Western New Britain. *VICAL1: Oceanic Languages, Papers from the Fifth International Conference on Austronesian Linguistics, Part 2* ed. by Ray Harlow & Robin Hooper, 555–579. Auckland: Linguistic Society of New Zealand.
Thurston, William R. 1994. Renovation and Innovation in the Languages of North-Western New Britain. *Language Contact and Change in the Austronesian World* ed. by Tom Dutton & Darrell Tryon, 573–609. Berlin: Mouton de Gruyter.
Toulmin, Matthew. 2009. *From Linguistic to Sociolinguistic Reconstruction: The Kamta Historical Subgroup of Indo-Aryan*. Pacific Linguistics: Australian National University.
Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.
Tryon, Darrell T. 1976. *New Hebrides Languages: An Internal Classification*. Pacific Linguistics, C50. Canberra: Australian National University.
Tryon, Darrell T. 1996. Dialect Chaining and the Use of Geographical Space. In Bonnemaison, Huffman, Kaufmann & Tryon, eds., 170–181.
Tryon, Darrell & Jean-Michel Charpentier. 2004. *Pacific Pidgins and Creoles: Origins, Growth and Development*. Berlin: Mouton de Gruyter.
Vanhove, Martine, ed. 2008. *From Polysemy to Semantic Change: Towards a Typology of Lexical Semantic Associations*. Amsterdam: John Benjamins.
Vienne, Bernard. 1972. Les Formes du Pouvoir et de l'Autorité dans les Communautés Paysannes des Îles Banks (Nouvelles-Hébrides). Nouméa, Nouvelle-Calédonie: ORSTOM.
Vienne, Bernard. 1984. *Gens de Motlav. Idéologie et Pratique Sociale en Mélanésie*. Paris: Société des Océanistes.
VNSO [Vanuatu National Statistics Office]. 2009. *2009 National Census of Population and Housing: Summary Release*. Port Vila. (released 31 August 2009). 34 pp.
Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York: Linguistic Circle.
Wilkins, David. 1996. Natural Tendencies of Semantic Change and the Search for Cognates. *The Comparative Method Reviewed* ed. by Mark Durie & Malcolm Ross, 264–304. Oxford: Oxford University Press.
Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Wiley-Blackwell.
Winters, Margaret E. 1997. Kuryłowicz, Analogical Change, and Cognitive Grammar. *Cognitive Linguistics* 8:4.359–386.

Author's address:

Alexandre FRANÇOIS
Linguistics
School of Culture, History and Language
Australian National University
ACT 0200
AUSTRALIA

francois@vjf.cnrs.fr