



Structures et dynamiques des systèmes linguistiques: Documentation, description, comparaison des langues océaniennes

Alexandre François

► To cite this version:

Alexandre François. Structures et dynamiques des systèmes linguistiques: Documentation, description, comparaison des langues océaniennes. Linguistique. Institut National des Langues et Civilisations Orientales (INALCO), 2014. tel-02004673

HAL Id: tel-02004673

<https://theses.hal.science/tel-02004673>

Submitted on 27 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Alexandre FRANÇOIS

LACITO-CNRS, Paris
Australian National University

Structures et dynamiques des systèmes linguistiques

**Documentation, description, comparaison
des langues océaniques**

MÉMOIRE DE SYNTHÈSE

Travaux rassemblés en vue de
l'Habilitation à Diriger des Recherches

sous la direction de Stéphane Robert (LLACAN-INALCO)

Rapporteurs :

Martin Haspelmath (Univ. Leipzig),
Nicholas Thieberger (Univ. Melbourne)

Jury de soutenance :

Isabelle Bril (LACITO-CNRS), Annie Montaut (INALCO),
Denis Creissels (DDL-CNRS), Antoine Guillaume (DDL-CNRS).

Soutenance : le 26 juin 2014

Institut National des Langues et Civilisations Orientales, Paris

Table des matières

<i>Avant-propos</i>	7
1 INTRODUCTION	9
1.1 <i>Cadre général de mes travaux</i>	9
1.2 <i>Problématiques de recherche</i>	10
1.2.1 Structures et dynamiques des systèmes linguistiques	11
1.2.2 Diversité des langues et écologie sociale	12
1.2.3 Panorama des productions scientifiques	14
1.3 <i>Participation aux réseaux de la recherche</i>	17
1.3.1 Aperçu de mon parcours académique	17
1.3.2 Contributions à la vie de la recherche	18
1.3.3 Au-delà de la communauté scientifique	19
1.4 <i>Organisation de ce volume</i>	21
2 DOCUMENTATION DES LANGUES : LE TERRAIN	23
2.1 <i>Le contexte géographique</i>	23
2.1.1 Le Vanuatu, une mosaïque linguistique	24
2.1.2 Mes enquêtes au Vanuatu	26
2.1.3 Langues océaniques hors du Vanuatu	28
2.2 <i>Recueillir la langue spontanée : méthodologie</i>	29
2.2.1 Description et documentation des langues	29
2.2.2 Questionnaires vs. énoncés authentiques	29
2.2.3 La constitution du corpus	32
2.2.4 Pour un terrain monolingue	34
2.2.5 Les cahiers de terrain	35
2.3 <i>Parole spontanée et approche constructionniste</i>	37
3 LA DESCRIPTION GRAMMATICALE : CATÉGORIES ET CONSTRUCTIONS	39
3.1 <i>Décrire chaque langue comme un système</i>	40
3.2 <i>Phonologie</i>	42
3.2.1 Phonologie des voyelles du mwotlap	42
3.2.2 Phonologie des groupes consonantiques en hiw et en dorig	44
3.3 <i>Grammaire et catégorisation</i>	45
3.3.1 Les parties du discours	45
3.3.2 Les fonctions des noms	46
3.4 <i>Sémantique du Temps-Aspect-Mode</i>	49
3.4.1 La description des systèmes TAM	49
3.4.2 Le Gabarit standard de procès, et l'Aktionsart	51
3.4.3 Sémantique de la reduplication	54
3.4.4 Marquage aspecto-modal et subordination	55
3.5 <i>Prédicats complexes et relations actancielles</i>	58
3.5.1 Macroverbes et verbes sériels en mwotlap	58

3.5.2	Constructions sérielles et relations actanciennes	61
3.5.3	Syntaxe de la ditransitivité en araki	62
3.6	<i>Les systèmes spatiaux</i>	65
3.7	<i>Prospective</i>	67
4	DESCRIPTION LEXICALE ET CHANGEMENT SÉMANTIQUE	69
4.1	<i>Le travail lexicographique</i>	69
4.1.1	Dictionnaires publiés et en préparation	69
4.1.2	Principes généraux de lexicographie	71
4.2	<i>L'atlas linguistique de Polynésie Française</i>	73
4.3	<i>Études de sémantique lexicale</i>	76
4.3.1	Sémantique lexicale et variations syntaxiques	76
4.3.2	Termes de température	77
4.3.3	Étymologie et changement sémantique	78
4.4	<i>Approche typologique de la colexification</i>	80
4.4.1	Pour une typologie de la colexification	81
4.4.2	Les cartes sémantiques	81
5	LINGUISTIQUE HISTORIQUE : MODÉLISER LA DIVERSIFICATION	83
5.1	<i>Études comparatives au nord du Vanuatu</i>	83
5.1.1	Les systèmes de voyelles	83
5.1.2	Morphologie historique des marques d'aspect	86
5.1.3	Les articles des noms	87
5.1.4	Les systèmes de directionnels géocentriques	89
5.2	<i>Comprendre la diversification des langues</i>	94
5.2.1	Contexte	94
5.2.2	Unité et diversité dans une même île : Vanikoro	95
5.2.3	Divergence et convergence parmi les langues Torres–Banks	98
5.3	<i>Reconstituer les processus historiques de diffusion</i>	104
5.3.1	La disparition de la consonne *R	104
5.3.2	Diffusion lexicale, diffusion dialectale	105
5.4	<i>Modéliser les processus de diversification</i>	107
5.4.1	Problèmes avec le modèle arborescent	107
5.4.2	Réhabiliter la Théorie des Ondes	109
5.4.3	La glottométrie historique	112
5.4.4	Diffusion de ces résultats	115
6	LES LANGUES DANS LEUR CONTEXTE SOCIAL ET CULTUREL	117
6.1	<i>Un regard sociolinguistique</i>	117
6.2	<i>Agir pour la préservation des langues</i>	119
6.2.1	Donner aux langues leurs premières publications	119
6.2.2	Archivage et diffusion de mes enregistrements	121
6.3	<i>Documenter les pratiques esthétiques</i>	123
6.3.1	La littérature orale	123
6.3.2	Musique et poésie	124
6.4	<i>Valorisation, popularisation</i>	127

7	PARTICIPATION À LA VIE SCIENTIFIQUE	129
7.1	<i>Contributions à la recherche</i>	129
7.1.1	Organisation d'événements scientifiques	129
7.1.2	Séminaires	129
7.1.3	Administration de la recherche	130
7.1.4	Comités de sélection	130
7.2	<i>Activités éditoriales</i>	131
7.2.1	Comptes rendus	131
7.2.2	Comités de lecture, consultance	131
7.2.3	Édition scientifique	131
7.3	<i>La formation des étudiants</i>	132
7.3.1	Enseignements	132
7.3.2	Interventions invitées	133
7.3.3	Direction d'étudiants	133
7.3.4	Jury de thèse	134
7.3.5	L'école thématique	134
7.3.6	Autres formes d'encadrement	135
8	PROJETS	137
8.1	<i>Poursuivre la description des langues</i>	137
8.1.1	Grammaires de référence	137
8.1.2	Études spécifiques	138
8.1.3	Un ouvrage de dialectologie aréale	140
8.1.4	Ouvrage collectif sur les langues du Vanuatu	140
8.2	<i>Recherches en sémantique lexicale</i>	141
8.2.1	Typologie de la colexification	141
8.2.2	Polysémies et phénomènes aréaux	142
8.3	<i>Développer la Glottométrie historique</i>	142
8.4	<i>Littérature orale</i>	143
8.5	<i>Vie de la recherche et formation</i>	145
8.5.1	Séminaires	145
8.5.2	Enseignements	145
8.5.3	Supervision d'étudiants	146
8.6	<i>À plus long terme</i>	146
9	ABRÉVIATIONS	147
9.1	<i>Langues</i>	147
9.2	<i>Gloses</i>	147
10	RÉFÉRENCES BIBLIOGRAPHIQUES	147

Toutes les pièces de ce dossier d'Habilitation
(mémoire de synthèse, publications, documents, liens)
sont également accessibles au format numérique :

<http://alex.francois.free.fr/hdr>

Avant-propos

Passage obligé dans le monde universitaire français, l'Habilitation à Diriger des Recherches constitue un exercice original, sans équivalent exact à l'étranger. Mi-travail scientifique, mi-rite de passage, l'Habilitation incite à interrompre un moment l'activité de publication et de recherche, pour se retourner sur les années écoulées. Les questions qu'on se pose alors sont celles qui nous hantent, mais qu'on n'est pas toujours prêt à affronter de face : mes travaux depuis la thèse ont-ils apporté une contribution utile à l'effort collectif en linguistique ? Qu'ai-je accompli, du programme de recherche que je m'étais assigné au début de ma carrière ? Quelles directions prendre à l'avenir ?

L'incessant tourbillon de la vie académique moderne, faite d'échéances et de projets à court terme, n'offre pas souvent l'occasion de prendre ainsi du recul sur son propre travail. Il faut admettre que l'exercice a un côté redondant : il consiste à reprendre et résumer des recherches publiées parfois des années plus tôt, territoires dont on a depuis longtemps quitté les rivages pour en explorer de nouveaux. Mais ce travail réflexif a aussi ses bienfaits : il donne l'occasion d'identifier des liens insoupçonnés entre des publications pourtant éloignées, et dégager des lignes de force dans un parcours scientifique parfois vécu, sur le moment, comme aléatoire et contingent. Considérées individuellement, telle mission de terrain, telle publication, étaient nées d'une rencontre fortuite, d'un projet de collaboration, d'un congrès scientifique cette année-là... – le travail du chercheur est souvent soumis aux hasards de la vie. Pourtant, avec le recul des années, apparaissent des cohérences et des leitmotivs entre ces divers travaux, comme s'ils étaient guidés, malgré tout, par un discret fil d'Ariane.

Mesurer le chemin parcouru est aussi l'occasion de jauger le travail qui reste à accomplir, d'imaginer les voies que l'on suivra au cours des prochaines années, que ce soit seul ou en équipe. On a envie de progresser, d'aller de l'avant, de continuer à mettre en œuvre ce beau métier de chercheur.

Enfin, cette pause est aussi le moment de repenser à ceux qui m'ont accompagné le long de ces années, toutes ces personnes qui ont donné forme, chacune à sa manière, à ma recherche. Je pense à ma famille, mes professeurs, mes mille collègues d'ici ou d'ailleurs – sans oublier, bien entendu, les amis que l'on se fait aux antipodes, lors des séjours sur le terrain, rencontres si précieuses dans la vie d'un linguiste, et d'un homme.

Je dédie ces années de travail à †Moses Meywēlgen, mon père de Motalava, puits de science, homme de douceur et de sagesse.

Paris, mars 2014

1 Introduction

1.1 Cadre général de mes travaux

Après une Agrégation de Grammaire (option latin-grec) et des études de linguistique africaine (option peul-touareg), j'ai entrepris d'apporter ma contribution à l'effort collectif visant à documenter la diversité des langues du monde. Tôt sensibilisé à la problématique des langues en danger et à l'urgence de les décrire, j'ai choisi de me concentrer sur la région Pacifique – zone d'extrême densité linguistique, puisque 18,5% des langues vivantes y sont parlées par 0,1% de la population mondiale (Lewis, Simons & Fennig 2013). La famille linguistique la plus représentée dans cette région est le groupe des langues *océaniques*, une branche du vaste phylum austronésien. Au sein de la famille océanique, j'ai étudié un ensemble de langues parlées au Vanuatu et aux îles Salomon, en Mélanésie insulaire (*Figure 1*).

Figure 1 – Le Vanuatu et les îles Salomon, au cœur de la Mélanésie insulaire



Ma démarche scientifique est mue par le désir d'étudier les langues dans leur vitalité, en prêtant une attention particulière à leurs diverses dynamiques. Il s'agit d'observer la manière dont les locuteurs adaptent leurs systèmes linguistiques aux contraintes de la communication, qu'il s'agisse de pressions pragmatiques, cognitives, ou sociales. Ce parti pris implique la nécessité d'observer les pratiques linguistiques dans leur idiomaticité, et donc sur le terrain, en immersion dans leur environnement social. Ma méthode implique des séjours longs et fréquents dans les communautés de

locuteurs, au cours desquels j'apprends à parler et comprendre les langues dans une démarche d'*observation participante*, en multipliant les contextes d'interaction.

Au cours de ces séjours de terrain, j'ai recueilli des données de première main sur 21 langues distinctes. Parmi elles, 7 langues sont menacées d'extinction prochaine, n'étant parlées que par une poignée de locuteurs chacune. Les autres langues sont encore utilisées par leurs communautés – dont la taille s'échelonne entre 120 et 2100 locuteurs – mais leur viabilité à long terme demeure fragile, étant largement tributaire des futures évolutions des sociétés modernes. Ainsi, la plupart de mes langues de recherche peuvent être considérées comme menacées, à moyen voire à court terme. Sachant qu'elles n'avaient fait l'objet d'aucune description jusqu'à présent, il était donc urgent de les documenter.

Du point de vue des langues d'étude, mon parcours de recherche s'articule en trois étapes principales :

- depuis 1997, l'étude approfondie de deux langues en particulier, l'*araki* et le *mwotlap*, parlées au nord du Vanuatu¹ ;
- depuis 2003, l'étude transversale des 17 langues des îles Banks et Torres, parmi lesquelles se situe le *mwotlap* [Figure 3 p.27] ;
- depuis 2005, l'élargissement de mes recherches à des langues océaniques parlées dans d'autres pays (îles Salomon, Polynésie Française).

Cet élargissement géographique s'est également accompagné d'une évolution de mes thèmes de recherche [cf. §1.2].

Du point de vue institutionnel, mes travaux se sont d'abord déroulés dans le cadre de ma thèse de doctorat à Paris-IV (1998-2001). Depuis 2002, je suis Chargé de recherche au sein du laboratoire LACITO (*Langues et Civilisations à Tradition Orale*) du CNRS. De 2009 à 2012, j'ai effectué un séjour de trois ans et demi de mobilité internationale à l'*Australian National University* de Canberra. Au cours de mes 16 premières années de carrière, j'ai partagé mon temps entre missions de terrain, archivage, enseignement, publications, collaborations, et autres formes de diffusion de ma recherche.

1.2 Problématiques de recherche

Le présent document développera les différents aspects de ces travaux. Pour l'essentiel, mes contributions à la recherche en linguistique se définissent autour des grands axes suivants :

- **documentation des langues océaniques**
recueil de données linguistiques en situation;
enregistrements audio, en particulier récits de tradition orale et musique;
archivage de données audio-visuelles et linguistiques;
constitution de corpus textuels; réalisation d'un atlas linguistique.

¹ Dans la carte de la Figure 2 p.25, l'*araki* et le *mwotlap* correspondent respectivement aux langues n° 39 et 5.

- **description des systèmes dans une perspective typologique**
travaux descriptifs à partir des données linguistiques recueillies;
analyse des catégorisations opérées par la langue, et de la variabilité à leurs frontières;
recherches en phonologie, morphologie, syntaxe, grammaire de l'espace, sémantique du Temps-Aspect-Mode, prédicats complexes...;
travail lexicographique, études de sémantique lexicale;
contribution à une typologie lexicale.
- **études aréales et comparatives**
travaux de dialectologie et typologie aréale;
application de la Méthode comparative aux langues océaniques;
réflexion théorique sur les processus historiques de divergence et de convergence entre langues apparentées;
modélisation de la généalogie des langues.
- **inscription des langues dans leur environnement socioculturel**
approche de la langue dans ses dimensions sociales :
emblématicité, diffusion, interférences par contact;
esthétique de la langue : art verbal, littérature orale, poésie, chant;
contribution à la préservation et la promotion des langues en danger.

1.2.1 Structures et dynamiques des systèmes linguistiques

Mes travaux scientifiques cherchent à décrire les **structures** des langues vivantes, mais aussi et surtout à restituer leur vitalité en observant les **dynamiques** auxquelles ces structures sont sujettes.

Dans un premier temps, l'entreprise de description linguistique peut légitimement être conçue comme ayant pour objectif d'identifier les structures stables propres à chaque système en synchronie : le descripteur définira ainsi les *catégories émiqes* de la langue qu'il étudie, ainsi que le système des *relations* existant entre ces catégories, et les *règles* qui en constituent la grammaire [§3] ou le lexique [§4]. Cependant, mes travaux descriptifs, durant la thèse et au-delà, m'ont souvent incité à voir dans chaque langue, non pas seulement un système de catégories et de règles stables, mais une architecture complexe faite de notions souvent déformables, aux frontières poreuses, sujettes à l'interprétation et à la réanalyse. Telle construction concernant la catégorie des *animés* admettra des exceptions (animés traités comme inanimés ou vice-versa) ; tel verbe transitif sera traité comme intransitif dans tel contexte ; tel phonème appartenant à une classe X présentera parfois le comportement typique d'une classe Y ; etc. Ainsi, décrire une langue, c'est pour moi tout autant analyser ses rouages et ses régularités, qu'interroger les zones de frottement entre catégories, la variabilité dans l'application de ses règles, les ambiguïtés propices à la réinterprétation.

Cette **dynamique des systèmes linguistiques** concerne d'abord la *dynamique interne* de chaque langue. Cette dynamique interne a nourri plusieurs de mes descriptions synchroniques, où j'accorde fréquemment une place aux hésitations des locuteurs dues aux conflits entre plusieurs règles, à la variation entre locuteurs, au changement linguistique en cours.

Par la suite, mes recherches comparatives sur les 17 langues des Torres–Banks m’ont permis de prendre du recul, et d’observer les dynamiques de ces langues dans leurs *dimensions spatiales et temporelles* [§5]. Ainsi, mes travaux de dialectologie et de typologie aréale mettent à jour la distribution géographique des traits linguistiques, et les effets du contact de langues. Ces distributions dans l’espace sont elles-mêmes le fruit de changements historiques, que je reconstitue à l’aide de la Méthode comparative : la dynamique des langues apparaît alors à travers ses effets diachroniques sur le long terme – en particulier sous la forme de relations généalogiques entre langues.

Une bonne partie de mes recherches cherche à interroger l’articulation entre ces deux dimensions : la dynamique interne des systèmes en synchronie, et la dynamique des familles de langues en diachronie. L’échelle d’observation est différente : d’un côté, on se situe à hauteur de locuteur, à l’instant où il construit son énoncé, en faisant jouer les structures apprises de sa langue ; de l’autre, on adopte la perspective macroscopique de l’aire linguistique ou de la famille généalogique. Par quels processus exacts la micro-variation interne à un système résulte-t-elle dans la diversification des langues ? C’est là une des questions qui animent mes recherches actuelles en linguistique.

1.2.2 Diversité des langues et écologie sociale

Mes recherches sur la dynamique des systèmes linguistiques m’ont permis de jeter une lumière nouvelle sur un concept central de la typologie linguistique : la *diversité*. La typologie des langues s’interroge sur la variété des langues, et les limites de cette variété, à l’échelle planétaire. En examinant les dynamiques internes, aréales et historiques d’un échantillon spécifique de langues, mes travaux mettent précisément en lumière les mécanismes sous-jacents à l’émergence de cette diversité linguistique.

C’est la diversité des cultures et des langues qui m’a très tôt donné l’envie de devenir linguiste. C’est elle que je souligne chaque fois que je décris les structures grammaticales, ou les catégories lexicales, des langues que j’étudie – tâchant le plus souvent de les adosser aux tendances des langues du monde afin de mieux appréhender leur singularité. Ce faisant, mes études descriptives synchroniques contribuent au programme de la typologie, qui consiste précisément à rechercher les similarités des langues humaines tout en portant une attention particulière à leur diversité.

Mais si la notion de diversité des langues est si centrale à mes recherches, c’est aussi parce qu’elle se trouve au cœur de mes observations sur le terrain. La Mélanésie présente des taux record en matière de diversité linguistique, avec près d’un millier de langues distinctes – dont une centaine pour le seul archipel du Vanuatu. Mes recherches comparatives au nord du Vanuatu (mais aussi aux îles Salomon ou en Polynésie Française) conduisent toutes à mettre en valeur ce que l’on pourrait appeler la micro-diversité des langues : il ne s’agit plus ici de souligner la singularité de telle ou telle langue par rapport aux tendances universelles, mais de mettre en exergue les mosaïques linguistiques que l’on observe sur le terrain même. Avec 17 langues distinctes pour seulement 9300 habitants (soit une moyenne de 500 locuteurs par langue), les îles Torres et Banks constituent un laboratoire privilégié pour interroger cette appétence pour le multilinguisme, ce goût de la variété des cultures et des langues, si prononcés dans cette région [cf. l’échantillon donné dans le *Tableau 13* p.99]. Cette micro-diversité,

je l'ai abordée tantôt sous la forme de descriptions synchroniques des systèmes, tantôt sous son angle sociolinguistique, tantôt dans sa dimension historique.

Tenter de comprendre la diversité des langues modernes, dans une région comme le Vanuatu où toutes les langues ont un ancêtre commun, c'est aussi soulever le problème des **processus historiques de diversification**. Quels sont donc les facteurs qui sous-tendent la divergence des langues : sont-ils cognitifs, structurels, ou purement sociologiques ? Résultent-ils (comme la représentation arborescente nous le fait croire) de l'isolement et de la perte de contact ? Sont-ils le fruit d'une différenciation délibérée et consciente, comme on l'a parfois affirmé (Thurston 1989, Thomason 2007) ? La réponse que je propose [§5.2.3] est assez paradoxale. En un mot, je conclus que la force motivant les processus de divergence est au fond la même que celle qui donne lieu à la convergence : dans tous les cas, le processus réel, à échelle humaine, c'est la **diffusion horizontale des innovations linguistiques** d'un locuteur à l'autre, en vertu d'un mimétisme servant la cohésion du groupe. Ce qui diffère d'une région du monde à l'autre, c'est l'extension plus ou moins grande des groupes sociaux de référence, chaque groupe fonctionnant comme un creuset de convergence. En Mélanésie, le groupe de référence est de l'ordre du village, et c'est surtout à cette échelle que se fixeront les innovations. L'apparence de fragmentation n'est qu'un épiphénomène, à l'échelle macroscopique, de ces stratégies locales de cohésion sociale. Au fond, la diversité si prononcée de la Mélanésie reflète la structure de ses groupes humains – structure en réseaux décentralisés, politiquement acéphales, où autorité et prestige sont distribués dans les hiérarchies de parenté plutôt que dans un lieu central du pouvoir.

Mes réflexions sur ces questions m'ont notamment conduit, ces dernières années, à remettre en cause la validité du modèle arborescent, pourtant considéré, encore aujourd'hui, comme allant de soi en linguistique historique. L'idée que la phylogénétique des langues puisse être représentée comme un arbre est une erreur, car elle repose sur des présupposés faux – en particulier, celui que la différenciation des langues résulte typiquement de la séparation de leurs locuteurs. Dans le sillage d'autres chercheurs (par ex. Bosson, Ross, Heggarty, Drinka, Garrett), j'affirme la nécessité de débarrasser la Méthode comparative du modèle fallacieux de l'arbre fondé sur la divergence, et de la doter d'outils capables d'appréhender le processus central au changement linguistique – à savoir, la diffusion des innovations à travers les réseaux sociaux. Cette prise de conscience anime aujourd'hui plusieurs chercheurs travaillant sur diverses familles de langues, élan prometteur auquel j'espère contribuer. Pour ce faire, je propose d'opérer une synthèse entre la Méthode comparative des Néo-grammairiens (Schleicher 1853, Leskien 1876), la Théorie des Ondes (Schmidt 1872), et les importants progrès de la sociolinguistique au cours du xx^e s. (cf. Le Page & Tabouret-Keller 1985, Milroy & Milroy 1985, Milroy 1987, Labov 2001).

Comprendre la diversité des langues, en somme, c'est porter attention au contexte socioculturel où elles s'inscrivent – autrement dit, à l'*écologie sociale* de ces langues. À travers des terrains fondés sur l'immersion participante, j'ai su m'intégrer moi-même à cette écologie sociale. Cette expérience intense m'a ouvert les yeux non seulement sur des attitudes linguistiques nouvelles, mais aussi sur des pratiques sociales et culturelles inédites – pratiques où prévaut, précisément, une esthétique de la diversité.

1.2.3 Panorama des productions scientifiques

Le parcours de recherche esquissé dans les pages précédentes sera illustré dans les prochaines sections du présent document. Dans le cadre de cette introduction, je propose d’abord un bref aperçu de mes publications et autres productions scientifiques, au travers desquelles j’ai mis en œuvre ces réflexions. Les références exactes de ces publications sont données dans un document séparé.¹

La première langue sur laquelle j’ai travaillé, en 1997–98, a été l’**araki** – langue du centre du Vanuatu, dont j’ai connu l’un des tout derniers locuteurs. Conscient de l’urgence de documenter cette langue, j’ai consacré à l’araki mon premier ouvrage [*Araki: A disappearing language of Vanuatu*, 2002 ▶(#29)], écrit en même temps que ma thèse sur le mwotlap. Cet ouvrage consiste en une grammaire de référence, un lexique, et une collection de textes. Je suis plus récemment revenu sur cette langue pour un article sur les verbes ditransitifs [2012 ▶(#12)].

Lors de cette année de terrain 1997–98, j’ai surtout recueilli des données sur le **mwotlap**, la principale langue des îles Banks, au nord du Vanuatu. L’attention particulière portée au mwotlap a donné lieu à ma thèse de doctorat [2001 ▶(#28)], ainsi qu’à la plupart de mes articles de 1999 à 2005, sur divers points de grammaire.² J’y ai également consacré ma seconde monographie [*La sémantique du prédicat en mwotlap*, 2003 ▶(#30)], description sémantique de son riche système Temps–Aspect–Mode [§3.4]. L’étude approfondie du mwotlap m’a permis d’appréhender un système linguistique dans toute la complexité de son architecture, mais aussi dans ses multiples dynamiques – dynamiques sociales, générationnelles, stylistiques, esthétiques. Je reviendrai sur l’apport de cette langue, que je continue à décrire encore aujourd’hui.

Le *Tableau 1* rassemble mes principales publications portant sur la description des systèmes linguistiques en synchronie – qu’il s’agisse de grammaires ou de dictionnaires, de descriptions complètes, ou d’études de cas plus spécifiques. Ces travaux portent sur l’araki et le mwotlap, mais également sur certaines langues des îles Torres et Banks (hiw, lo-toga, dorig) sur lesquelles je travaille depuis 2003. J’inclus ici les travaux en cours, sur lesquels je reviendrai dans la prospective [§8.1].

Mes recherches sur les 17 langues des îles Torres–Banks ont surtout donné lieu à des travaux comparatifs, relevant soit de la dialectologie et de la linguistique aréale, soit de la linguistique historique [§5.1] – comme le montre le *Tableau 2*. À leur tour, ces études diachroniques ont débouché sur une réflexion plus théorique sur la question de la fluidité des systèmes et de leur dynamique historique. J’y analyse les processus linguistiques et sociaux qui sous-tendent la diversification des langues, et discute de la meilleure manière de les modéliser (*Tableau 3*). Enfin, j’inclus mes publications portant sur les aspects culturels de ces communautés (musique, poésie, littérature orale), dans la mesure où ils s’inscrivent également dans cette problématique de la diversité.

¹ Les références de type (#26) renvoient aux publications reproduites dans les deux volumes associés à ce mémoire, donnant accès au texte original de mes travaux. Ces mêmes textes sont tous téléchargeables sur (<http://alex.francois.free.fr/hdr>).

² Ces premiers travaux ont été récompensés en 2004 par la *Médaille de Bronze* du CNRS.

Tableau 1 – DÉCRIRE LES STRUCTURES DES LANGUES: quelques publications, classées thématiquement

THÈME	RÉF	ANNÉE	TITRE DE LA PUBLICATION	LANGUES
Phonologie	(#1)	1999	Mouvements et clonages de voyelles en motlav: Entre phonologie et morphologie	mwotlap
	(#14)	2000 a	Vowel shifting and cloning in Motlav : historical explanation vs formal description	mwotlap
	(#9)	2010 b	Phonotactics and the prestopped velar lateral of Hiw: Resolving the ambiguity of a complex segment	hiw, dorig
Prédicats complexes	(#15)	2004 a	Chains of freedom: Constraints and creativity in the macro-verb strategies of Mwotlap	mwotlap
	(#17)	2006	Serial verb constructions in Mwotlap	mwotlap
Temps-Aspect-Mode	(#37)	2001 b	Gabarit de procès et opérations aspectuelles en motlav (Océanie)	mwotlap
	(#30)	2003 a	<i>La sémantique du prédicat en mwotlap (Vanuatu)</i>	mwotlap
Syntaxe verbale	(#2)	2000 b	Dérivation lexicale et variations d'actance : petits arrangements avec la syntaxe	mwotlap
	(#12)	2012 a	Ditransitive alignment and referential hierarchies in Araki	araki
Syntaxe de la phrase	(#16)	2004 d	Diversité des prédicats non verbaux dans quelques langues océaniques	mwotlap
Subordination	(#22)	2010 c	Pragmatic demotion and clause dependency: On two atypical subordinating strategies in Lo-Toga and Hiw (Torres, Vanuatu)	hiw, lo-toga
Sémantique	(#3)	2000 c	L'illusion des classificateurs	mwotlap
	(#5)	2004 b	La reduplication en mwotlap : les paradoxes du fractionnement	mwotlap
Systèmes spatiaux	(#4)	2003 b	Of men, hills and winds: Space directionals in Mwotlap	mwotlap
Grammaire de la langue	(#28)	2001 a	<i>Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanique du Vanuatu</i>	mwotlap
	(#29)	2002	<i>Araki. A disappearing language of Vanuatu</i>	araki
	(#7)	2005 a	A typological overview of Mwotlap, an Oceanic language of Vanuatu	mwotlap
	(#33)	[en prép.]	<i>A grammar of Hiw (Torres Islands, Vanuatu)</i>	hiw
		[en prép.]	<i>A grammar of Teanu, the language of Vanikoro (Solomon Islands)</i>	teanu
Lexique	(#36)	[en prép.]	<i>Dictionnaire mwotlap-français-anglais</i>	mwotlap
	(#35)	[en prép.]	<i>A Teanu-English dictionary</i>	teanu

Tableau 2 – ÉTUDES COMPARATIVES, ÉVOLUTION HISTORIQUE DES SYSTÈMES: quelques publications, classées thématiquement

THÈME	RÉF	ANNÉE	TITRE DE LA PUBLICATION	LANGUES
Phonologie historique	(#8)	2005 b	Unraveling the history of the vowels of seventeen northern Vanuatu languages	nord Vanuatu
	(#11)	2011 c	Where *R they all? The geography and history of *R loss in Southern Oceanic languages	Vanuatu
Morphologie historique	(#18)	2007	Noun articles in Torres and Banks languages: Conservation and innovation	nord Vanuatu
	(#21)	2009 b	Verbal aspect and personal pronouns: The history of aorist markers in north Vanuatu	nord Vanuatu
Systèmes spatiaux	(#6)	2004 c	Reconstructing the geocentric system of Proto Oceanic	nord Vanuatu
	(#25)	[s. presse b]	The ins and outs of <i>up</i> and <i>down</i> : Disentangling the nine geocentric space systems of Torres and Banks languages	nord Vanuatu
Reconstruction sémantique	(#23)	2013	Shadows of bygone lives: The histories of spiritual words in northern Vanuatu	nord Vanuatu
Lexique	(#24)	[AF 2014]	Temperature terms in Vanuatu languages	nord Vanuatu
	(#32)	[JMC/AF 2014]	<i>Atlas Linguistique de Polynésie Française – Linguistic Atlas of French Polynesia</i>	Polyn. Fr.

Tableau 3 – APPRÉHENDER LA DIVERSITÉ, MODÉLISER LA DIVERSIFICATION: quelques publications, classées thématiquement

THÈME	RÉF	ANNÉE	TITRE DE LA PUBLICATION	LANGUES
Sémantique typologique	(#19)	2008 a	Semantic maps and the typology of colexification: Intertwining polysemous networks across lang ^s	général
	(#38)	AF&MP 2013	Descriptive linguistics	général
Sociolinguistique de la diversité	(#13)	2012 b	The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages	nord Vanuatu
Processus de divergence et convergence	(#20)	2009 a	The languages of Vanikoro: Three lexicons and one grammar	Vanikoro
	(#10)	2011 b	Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence	nord Vanuatu
Modèles en linguistique historique	(#26)	[SK&AF s/p]	Freeing the Comparative Method from the tree model: A framework for Historical Glottometry	général
	(#27)	[AF s/p.c]	Trees, Waves and Linkages: Models of Language Diversification	général
Diversité culturelle	(#31)	AF&MS 2013	<i>Musiques du Vanuatu: Fêtes et Mystères – Music of Vanuatu: Celebrations and Mysteries</i>	nord Vanuatu
	(#34)	[AF, en prép.]	<i>Voices from the Volcano: Myths, Folktales and Poetry from Gaua, Vanuatu</i>	Gaua

1.3 Participation aux réseaux de la recherche

1.3.1 Aperçu de mon parcours académique

Dès l'adolescence, j'étais attiré par les langues dans leur diversité – langues anciennes ou langues lointaines, langues écrites ou langues orales. Pendant des années, mon intérêt pour les langues avait surtout consisté à en apprendre quelques-unes (espagnol, italien, catalan, arabe) et à m'initier à d'autres (russe, allemand, occitan, grec moderne, basque, gallois, chinois, japonais). Apprenant, vers l'âge de 15 ans, qu'un grand nombre des langues du monde restaient encore méconnues alors qu'elles étaient menacées d'extinction prochaine, j'avais trouvé ma vocation : linguiste – ou plutôt *ethnolinguiste*, comme je nommais alors cette synthèse que j'imaginai entre linguistique et anthropologie.

Ma découverte des langues avait été catalysée par la langue grecque, dont j'allais poursuivre l'étude pendant douze années – jusqu'à mon entrée à l'*École Normale Supérieure* de la Rue d'Ulm en 1992, puis l'*Agrégation de Grammaire* en 1995. Tout me plaisait dans l'étude du grec : l'élégance de sa grammaire, la puissance de sa littérature, et aussi la magie de remonter les millénaires pour se glisser dans des mots aujourd'hui disparus, et ainsi partager pas à pas les réflexions d'un Platon, les étonnements d'un Hérodote, les désirs d'un Ulysse, les colères d'une Antigone. Pourtant, tout en menant à leur terme mes études de Lettres Classiques, j'avais déjà décidé de consacrer mes futures recherches à l'exploration de langues à tradition orale et non encore décrites. Aiguillé par des conseils que j'avais reçus de Claude Hagège dès 1988, j'ai choisi d'acquérir une formation en linguistique en suivant la Licence en Sciences du Langage de l'Université Paris-III Sorbonne Nouvelle, et son option "Linguistique Africaine" (1992-93). En parallèle, je suivais des cours de peul et touareg à l'INALCO, et bénéficiais d'une introduction à la linguistique générale et à la typologie des langues, que Stéphane Robert et Alain Lemaréchal dispensaient à l'École Normale Supérieure.

Ma Maîtrise de Lettres Classiques portait encore sur le grec, sous un angle linguistique (*Choix aspectuel et injonction d'après l'œuvre d'Euripide*, dirigé par Anne-Marie Chanet, Paris-x Nanterre, 1994). Quant à mon mémoire de DEA (*La subordination sans marques segmentales: Formes de dépendance interpropositionnelle dans le discours*, dirigé par M-A. Morel, Université Paris-III Sorbonne Nouvelle, 1997), il a marqué mon tournant vers la typologie des langues, puisqu'il portait à la fois sur le français oral et sur un échantillon varié de langues du monde.

J'avais tôt identifié l'Océanie, et notamment le Vanuatu, comme une aire de grande diversité linguistique, et encore insuffisamment connue : plus tard, j'y consacrerai mes propres explorations. D'octobre 1997 à août 1998, j'ai ainsi effectué une année de recherches dans cet archipel¹ ; c'était le premier d'une longue série de séjours sur le

¹ Cette année de recherche avait été financée par une bourse Walter-Zellidja décernée par l'Académie Française ; et par le fonds Naudet, géré par l'École Normale Supérieure de la rue d'Ulm.

terrain, qui allait se dérouler sur une quinzaine d’années. Je travaillais d’abord dans le cadre de ma thèse de doctorat (1998-2001), effectuée à Paris-IV Sorbonne sous la direction d’Alain Lemaréchal. Depuis 2002, j’ai eu la chance de rejoindre les rangs du CNRS, et en particulier ceux de mon laboratoire, le **LACITO**. Nouvelle recrue parmi les océanistes, j’avais pour mentors Jean-Claude Rivierre et Françoise Ozanne-Rivierre : par leurs conseils, ils m’ont accompagné tout au long de mes découvertes des cultures du Pacifique ainsi que du métier de chercheur, et je leur en sais gré encore aujourd’hui. Le LACITO regroupe des gens passionnés par leur métier, dont la diversité des terrains demeure pour moi une inspiration. Durant ces années, j’ai également trouvé très formateur d’interagir avec les laboratoires proches comme le LLACAN, le CELIA/SEDYL ou le DDL de Lyon. La Fédération *Typologie et Universaux Linguistiques* a beaucoup facilité ces échanges scientifiques, à travers divers programmes auxquels j’ai participé depuis sa création [cf. §7.1.2].

Mes cercles scientifiques ont trouvé une dimension internationale depuis qu’en 1997, j’ai pris mes premiers contacts avec l’**Australian National University**. En route pour le Vanuatu pour mon premier terrain, j’avais fait un détour par Canberra pour y rencontrer Darrell Tryon, alors la principale autorité en ce qui concerne les langues du Vanuatu [cf. 2013b ▸(#39)]. Ce passage au RSPAS (Research School of Pacific and Asian Studies) m’a donné l’occasion de tisser des liens avec les chercheurs de ce prestigieux établissement : Stephen Wurm, Andrew Pawley, Malcolm Ross... Entretenus pendant quelques années à l’occasion des colloques austronésianistes et océanistes, ces liens scientifiques allaient, bien plus tard, se consolider lors d’une période de mobilité internationale (Mise à disposition) à l’A.N.U., de janvier 2009 à juillet 2012. Par chance, ce séjour coïncidait avec une période passerelle dans l’histoire de ce département : je pouvais interagir avec les membres de l’ancienne équipe historique – spécialement A. Pawley et M. Ross – tout en créant des liens solides avec la toute nouvelle équipe qui émergeait autour de Nicholas Evans. J’ai pu travailler avec des personnalités telles que Mark Donohue ou Bethwyn Evans, sans compter les nombreux doctorants du département, qu’il m’est d’ailleurs arrivé de conseiller dans leurs thèses [§7.3.3]. Particulièrement fructueuses et inspirantes, ces années ont enrichi mes connaissances sur les langues austronésiennes, australiennes, papoues, mais ont aussi attiré mon attention, par exemple, sur les liens entre linguistique et archéologie – un des points forts de l’A.N.U, souvent négligé en France. Enfin, j’ai également pu tisser des collaborations, préludes à des publications, déjà parues ou à venir [François & Ponsonnet 2013 ▸(#38); Kalyan & François, *ss presse* ▸(#26)]. Aujourd’hui, grâce au soutien sans faille de Nick Evans, je demeure officiellement affilié à l’ANU, avec le statut de “Research Associate” ; et j’espère poursuivre longtemps mes échanges scientifiques et humains avec cette université des antipodes.

1.3.2 Contributions à la vie de la recherche

S’il est vrai que mes travaux prennent typiquement la forme de publications ou d’interventions individuelles dans les colloques, ma participation à la vie scientifique va au-delà [cf. §7.1].

Certaines publications ou communications impliquent une collaboration, à l'échelle nationale et internationale, au sein de ma discipline ou au-delà : avec Jean-Michel Charpentier [§4.2], Maïa Ponsonnet ▶(**#38**), Siva Kalyan [§5.4.3], Monika Stern [§6.3.2]. Je prépare actuellement un ouvrage collectif, avec trois coéditeurs et une dizaine d'auteurs de divers pays [§8.1.4]. J'ai participé depuis 2003 à plusieurs opérations de recherche inter-laboratoires de la Fédération *Typologie et Universaux Linguistiques* [§7.1.2]. J'ai organisé le colloque austronésien 11-ICAL en 2009 (avec L. Sagart et I. Bril), et les premières rencontres internationales sur les langues du Vanuatu (avec S. Lacrampe) ; pour d'autres colloques (ALS, COOL, ICAL, SWL), j'ai été membre du Comité scientifique [§7.1.1]. J'appartiens au comité éditorial de plusieurs revues (*Oceanic Linguistics*, *Open Linguistics*) et suis relecteur scientifique dans d'autres (*Linguistic Typology*, *Linguistics*, *Studies in Language*, etc.). J'ai fondé et dirige la d'ouvrages scientifiques *Studies in the Languages of Island Melanesia*, chez l'éditeur *Pacific Linguistics* [§7.2]. Je suis consultant pour plusieurs institutions académiques (ELDP, NSF, INSHS) ; membre du Comité de sélection de l'Université de Polynésie Française ; rapporteur et jury de doctorat (U. of Oslo).

Ces contributions à la vie scientifique ont eu lieu aussi bien quand j'étais basé en France qu'en Australie. Depuis mon retour en 2012, je suis membre élu au Conseil de laboratoire du LACITO, et appartient à son équipe dirigeante, sous la nouvelle direction de Samia Naïm. Sur le plan scientifique, j'appartiens à plusieurs axes du LabEx *Empirical Foundations of Linguistics*, y compris un axe dont je suis co-responsable (avec V. Vydrine et I. Bril)¹. En parallèle, je coordonne depuis 2012 le principal séminaire de recherche du LACITO (*Problèmes d'Analyse et de Comparaison des Langues*), au cours duquel je favorise les échanges scientifiques entre chercheurs (y compris étrangers), jeunes chercheurs et doctorants du laboratoire. Enfin, j'organise une école d'été en juillet 2014, visant à former les doctorants et jeunes chercheurs sur le thème *Méthodes en linguistique : Des données empiriques aux hypothèses typologiques* [§7.3.5] ; cette école thématique, cofinancée par le CNRS, le LabEx EFL et l'INALCO, me donnera l'occasion d'inviter des typologues extérieurs à la région parisienne, y compris les Prof. Denis Creissels (DDL-Lyon) et M. Haspelmath (du Max-Planck Institute de Leipzig).

1.3.3 Au-delà de la communauté scientifique

Pour la plupart, les travaux mentionnés ici portent sur la linguistique proprement dite. Cependant, j'ai également produit diverses publications sous d'autres formats, et/ou portant sur des thèmes périphériques aux seuls questionnements linguistiques. Ces passerelles vers d'autres publics ont donné lieu, par exemple, à une publication discographique *Musiques du Vanuatu*, en collaboration avec Monika Stern (CREDO) [§6.3.2], ainsi qu'à un ouvrage en préparation, plus anthropologique, sur la littérature orale [§8.4]. J'ai coutume de diffuser mes connaissances à divers publics non spécialisés en linguistique, tantôt en collaborant avec des collègues d'autres disciplines, tantôt en

¹ Il s'agit de l'axe RT1 "Généalogie des langues : Reconstruction, classification interne et description des deux principaux phylums du monde: Niger-Congo et Austronésien" [<http://axe3.labex-efl.org/?q=fr/RT1f>].

popularisant le travail des linguistes (interviews dans les médias, conférences grand public, participation à des documentaires, présence sur internet).

Un autre destinataire de mon travail est l'ensemble des communautés linguistiques qui m'ont accueilli. J'ai ainsi confectionné 26 petits livres destinés à l'alphabétisation et la lecture dans les langues vernaculaires : abécédaires, recueils d'histoires, albums [§6.2.1] ; c'est que plusieurs de ces langues sont menacées de disparition, et tout effort pour en consolider l'usage dans la communauté est le bienvenu. À ces activités de valorisation des langues, s'ajoute un travail important d'archivage de mes enregistrements sonores (>104 h), de mes textes transcrits (sur l'archive *Pangloss*), de mes cahiers de terrain – le tout destiné aux communautés de locuteurs ainsi qu'au grand public [§6.2.2]. En somme, mes activités relèvent tout autant de la description scientifique des systèmes linguistiques que de la *documentation des langues* – deux activités qui sont parfois opposées l'une à l'autre, mais que je tente de concilier [§2.2.1].

Enfin, au travers de mes activités d'enseignement et d'encadrement de la recherche, je m'adresse également au public d'étudiants et de jeunes chercheurs. Au cours des quinze premières années de ma carrière, j'ai enseigné divers aspects de la linguistique dans plusieurs établissements [§7.3.1] : l'Université Paris-III Sorbonne Nouvelle (ILPGA), l'Université Paris-X Nanterre, l'INALCO – sans oublier l'Australian National University (ANU) de Canberra, ou encore le *LSA Summer Institute*, qui en 2011 se tenait à l'Université du Colorado (UC Boulder). Dans ces occasions, j'ai pu apprécier le plaisir d'échanger avec des étudiants d'horizons très divers, et les voir progresser dans le métier de linguiste. J'ai souvent encadré des doctorants ou jeunes chercheurs – notamment dans le cadre de mon financement *ACI Jeunes Chercheurs* (2004-2007), ou durant mon séjour à l'ANU – mais ce fut toujours dans un cadre informel.

La situation me semble mûre pour désormais formaliser ce travail de supervision. Si j'ai attendu ce jour pour solliciter ce privilège, c'est parce que je souhaitais atteindre ce moment dans une carrière où l'on a le sentiment d'avoir obtenu dans sa propre recherche de véritables résultats, suffisamment solides et substantiels pour qu'ils méritent d'être transmis à la génération suivante. L'esprit du présent document, précisément, est de présenter une synthèse de ces résultats, afin de situer mon travail dans le cadre plus large des institutions et traditions de recherche actives dans notre domaine.

1.4 Organisation de ce volume

Le présent volume présentera une synthèse de mes contributions scientifiques au cours des quinze premières années de ma carrière. En voici les principaux chapitres :

- §2 – *Documentation des langues : le terrain*
- §3 – *La description grammaticale : catégories et constructions*
- §4 – *Description lexicale et changement sémantique*
- §5 – *Linguistique historique : modéliser la diversification*
- §6 – *Les langues dans leur contexte social et culturel*

Dans la mesure où les langues océaniennes sur lesquelles j’ai choisi de travailler sont dépourvues d’écriture et de tradition grammaticale, le travail de terrain occupe une place centrale dans ma démarche empirique. J’y consacre donc le prochain chapitre, afin de présenter ma méthodologie, et mes principaux résultats en termes de documentation des langues. Les sections suivantes portent sur mes diverses publications scientifiques, organisées en domaines : la description grammaticale des systèmes en synchronie [§3] ; la description du lexique [§4] ; le travail de reconstruction historique, de comparatisme et de modélisation du changement linguistique [§5] ; l’insertion de la langue dans l’ensemble plus large des pratiques sociales, culturelles et esthétiques [§6].

Ces chapitres proprement scientifiques seront suivis d’une présentation de mes contributions à la vie de la recherche et à la formation des étudiants [§7]. Enfin, la dernière section présentera une prospective de mes projets scientifiques pour les années à venir.

2 Documentation des langues : le terrain

Avant d'être orientée vers des questionnements théoriques, la motivation première de mes recherches était de documenter des langues en danger. Dans la mesure où les langues que j'avais choisies étaient dépourvues d'écriture et de travaux antérieurs de description, la nécessité d'effectuer des enquêtes de terrain s'est imposée d'elle-même.

Cette première section présente ces enquêtes de terrain, car ce sont elles qui constituent le cadre méthodologique de mes travaux. Certes, mes années de formation initiale m'avaient fourni les outils d'analyse structurale indispensables pour formuler mes observations ; mais c'est sur le terrain, confronté aux langues dans toute leur complexité, que sont nés mes questionnements et mes problématiques de recherche. L'enquête n'était pas pour moi un moyen pour répondre à des questions préalables, mais le lieu même – dans une démarche résolument empiriste – où ces questions allaient émerger. C'est en observant comment ces langues étaient pratiquées par les acteurs sociaux dans leurs interactions quotidiennes, et en les pratiquant moi-même dans le même contexte, que j'ai pu voir se dessiner, au fil des enquêtes, les idées futures de ma recherche : la langue comme système de relations entre catégories ; la variabilité comme indice de l'ambiguïté de ces catégorisations ; la question de la convergence par contact et de la diversification des langues apparentées – pour ne mentionner que les questions les plus générales.

En somme, mon approche du terrain était délibérément dépourvue de questionnement *a priori*. Ma motivation première consistait à observer les langues et documenter les pratiques linguistiques dans leur diversité. Les problèmes de description, et la réflexion théorique sur le langage ou la société, n'allaient intervenir que dans un deuxième temps. Voilà pourquoi je commence ici par présenter les enquêtes de terrain, et à travers celles-ci, ma contribution à l'effort collectif de *documentation des langues* du monde – plus précisément, celles de la région Pacifique.¹

2.1 Le contexte géographique

L'océan Pacifique est l'une des régions du monde les plus riches linguistiquement. Comme toute richesse, celle des langues est inégalement répartie : d'un côté, la vaste zone orientale connue sous le nom de "triangle polynésien" ne compte guère qu'une quinzaine de langues, chacune étant typiquement parlée par de vastes populations (ex. 50 000 locuteurs pour le tahitien, 360 000 pour le samoan) ; de l'autre côté, la partie occidentale du Pacifique, que Dumont d'Urville nomma *Mélanésie* en 1834, compte près de 1250 langues distinctes. Si l'on met de côté le millier de langues parlées dans l'île de

¹ Le fait qu'il s'agisse de langues océaniques (austroasiatiques) n'est pas particulièrement pertinent ici. Cette appartenance généalogique a surtout du sens dans le cadre de mes travaux de comparatisme historique [§5.1].

Nouvelle-Guinée – dont la plupart sont des langues “papoues” –, il reste environ 250 langues qui sont parlées dans les seuls archipels de la Mélanésie insulaire. Parmi eux, les deux États les plus divers sont le Vanuatu et les îles Salomon – les deux pays où j’ai choisi de faire porter mes enquêtes de terrain.¹

2.1.1 Le Vanuatu, une mosaïque linguistique

Mon choix de concentrer mes recherches dans l’archipel du Vanuatu était né d’un article que j’avais lu en 1989, sur l’incroyable diversité linguistique de ce petit pays (Tryon & Charpentier 1989). Avec sa centaine de langues, dont la plupart étaient alors non décrites – et souvent menacées – le Vanuatu offrait un contexte idéal pour mon projet de contribuer à la documentation et la description des langues du monde.

Le Vanuatu compte 113 langues différentes pour 234 000 habitants, ce qui constitue la plus forte densité au monde en nombre de langues rapporté à sa population (Crowley 2000). Le décompte exact des langues diffère selon les sources : ainsi Darrell Tryon, qui avait proposé un total de 113 langues en 1976, a révisé ce chiffre à la hausse en 2010, en se fondant sur les travaux des chercheurs de ces dernières années. Il a ainsi préparé une synthèse cartographique (non publiée), que je reproduis dans la *Figure 2*, montrant un total de 125 langues vernaculaires. Cette forte diversité linguistique n’est pas une diversité génétique – puisque toutes ces langues partagent un ancêtre commun, le *proto océanien* – mais une diversité acquise au cours des 3200 ans de son peuplement, au fil d’un lent processus historique de diversification [§5.2]. Cette mosaïque de langues a largement survécu jusqu’à ce jour, même si elle commence à montrer des signes d’érosion [2012b ►(#13)].

Avant de me plonger dans les langues du pays, j’ai d’abord fait la connaissance d’une société – ou plutôt de plusieurs communautés, disséminées dans cet archipel du Pacifique Sud. C’était tout un nouveau mode d’organisation sociale que je découvrais : une nation jeune – dont l’indépendance datait de 1980 –, très nettement décentralisée, faiblement colonisée, et distribuée en une multitude de micro-sociétés formant une mosaïque complexe de cultures et de langues. Pour le moment, l’appartenance à une communauté linguistique indigène demeure un puissant marqueur identitaire dans la population du Vanuatu. Autant dans leurs pratiques sociales que linguistiques, ces groupes en constante interaction réciproque soulèvent des questions importantes sur la circulation des hommes et des idées, sur la perception de la différence et de la similitude, sur la construction d’une communauté nationale, ou encore l’articulation de l’ancien et du moderne. S’il est vrai que ces problématiques concernent surtout les collègues anthropologues avec qui je travaille, elles allaient nourrir également mes propres réflexions, qu’il s’agisse de sociolinguistique contemporaine [§6.1] ou de linguistique historique [§5.3].

¹ Le détail de mes enquêtes de terrain est donné dans mon CV.

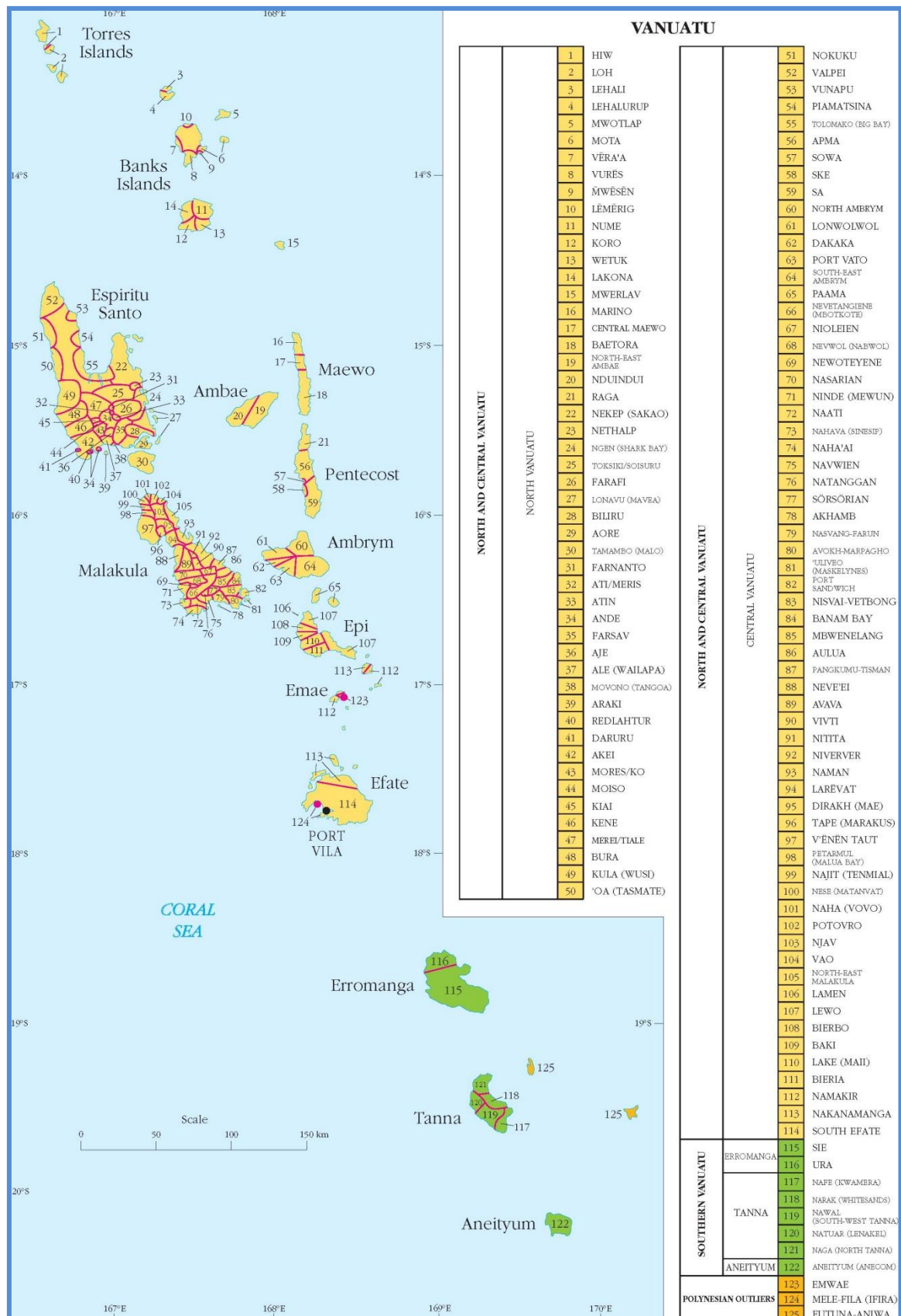


Figure 2 – Les langues du Vanuatu (source: D. Tryon, pers. com., 2010)

Anciennement connue sous le nom de “Condominium Franco-Britannique des Nouvelles Hébrides”, la République du Vanuatu comporte deux centres urbains : la capitale Port-Vila, et Luganville ou “Santo”. Pour le chercheur étranger, ces deux villes constituent un sas entre deux mondes, un centre stratégique où l’on bascule du réseau interconnecté des pays du monde au réseau interconnecté des îles de l’archipel. Ce statut très particulier rend ces villes passionnantes en elles-mêmes, car c’est là que se construit une nouvelle société politique, de nouvelles pratiques sociales — et notamment un nouveau paysage linguistique, marqué par l’hypermultilinguisme des immigrants de l’intérieur, et la montée en puissance du créole national.

Lentement mais sûrement, la population du Vanuatu cède aux sirènes de l’exode rural : entre 1999 et 2009, la proportion de la population vivant dans les villes a augmenté de 42,7% (VNSO 2009). En parallèle, la langue véhiculaire nationale qu’est le créole bislama est devenue la langue principale de 33,7% de la population en 2009 – alors que ce chiffre n’était que de 23,3% dix ans plus tôt [§6.1]. Inversement, ceux qui déclarent avoir comme langue principale du foyer l’une des langues vernaculaires a décrû, durant la même période, de 73,1 à 63,2% de la population. Dans un pays où de nombreuses langues sont fragiles car parlées par des petites communautés, cette chute de dix points en dit long sur les tendances actuelles à l’érosion linguistique. La description de ces langues en devient d’autant plus urgente.

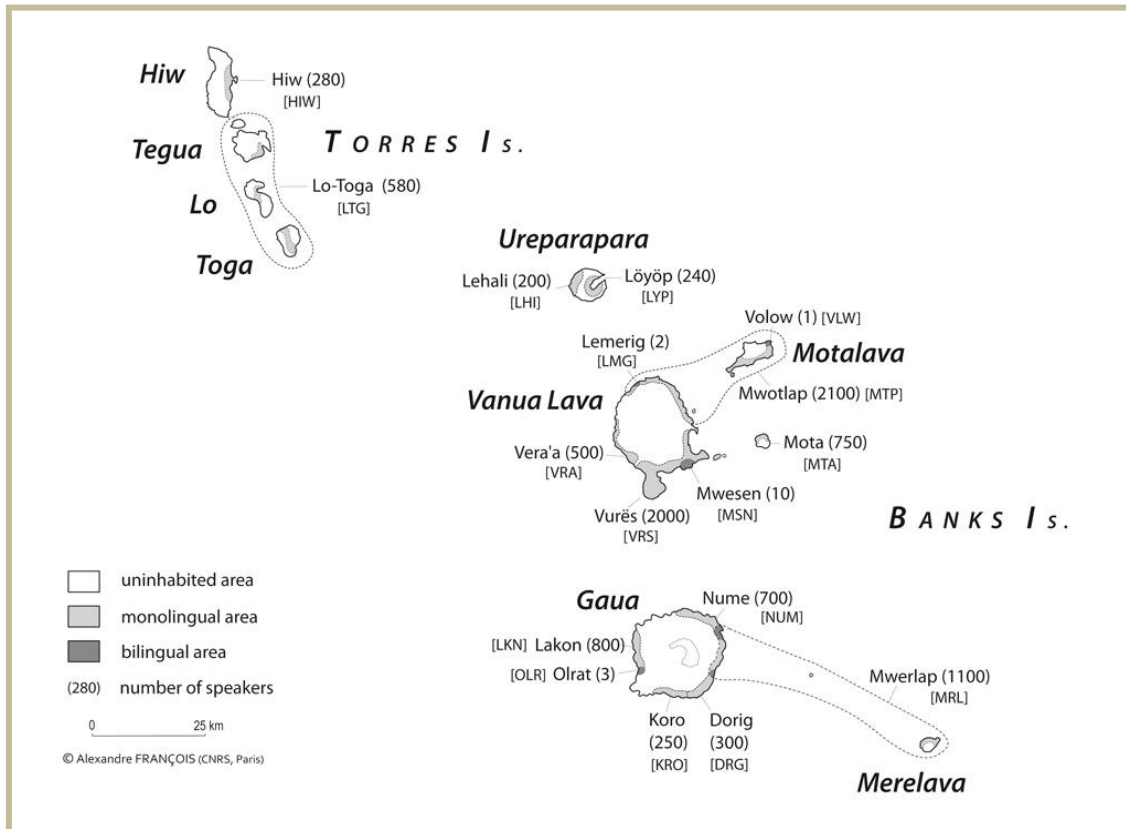
2.1.2 Mes enquêtes au Vanuatu

Mon exploration du Vanuatu a commencé dès 1997 avec ma rencontre de Lele Moli, l’un des derniers locuteurs de la langue **araki** (sud d’Espiritu Santo ; langue n°39 dans la Figure 2). À l’araki, j’ai consacré une grammaire, ou plutôt une “trilogie descriptive” à la Franz Boas, c’est-à-dire un ouvrage alliant grammaire, dictionnaire et textes [2002 ▶ (#29) ; voir aussi 2008 et 2012a ▶ (#12)].

Même s’il a donné lieu à une monographie de 388 pages, mon travail sur l’araki n’était en réalité qu’un excursus imprévu dans une année de terrain consacrée à une autre langue : le **mwotlap**. Langue principale des îles Banks au nord du Vanuatu (n°5 dans la Figure 2), parlée par environ 2100 locuteurs, le mwotlap est une langue bien vivante – au contraire, par exemple, de l’araki. Travailler simultanément sur ces deux langues m’a permis de constater la différence importante dans la nature et la qualité des données que l’on peut recueillir sur une langue, en fonction de sa situation sociolinguistique. Pour l’araki, je n’ai longtemps eu accès qu’à un seul locuteur, avec qui je devais surtout recourir à de l’éllicitation ; j’ai pu recueillir des récits – de cette personne Lele Moli, et plus tard de quelques autres derniers locuteurs de l’araki – mais je n’ai jamais pu entendre une véritable conversation dans cette langue : la communauté l’avait déjà abandonnée en faveur du tangoa voisin. La situation était très différente pour le mwotlap, langue dynamique utilisée par toutes les générations, et entendue quotidiennement dans les venelles de son île, Motalava. Certes, les migrants de l’intérieur, venus de Motalava pour vivre à la ville, présentent une tendance prononcée à abandonner leur langue pour le créole bislama, ce qui constitue une menace à terme si l’exode rural devait s’accroître. Cependant, à l’heure actuelle, la situation sociolinguistique du mwotlap est plutôt solide dans sa communauté d’origine. Cette

situation m’a permis de recueillir une abondance de données de la part de locuteurs de tous âges et de tous horizons – qu’il s’agisse de récits, de conversations, de chansons ou autres savoirs culturels.¹ S’il est vrai qu’il y a urgence à se pencher au chevet de “langues moribondes” telles que l’araki, le travail de documentation est indéniablement plus riche lorsqu’il porte sur une langue encore viable, comme le mwotlap.

Figure 3 – Les dix-sept langues des îles Torres et Banks, au nord du Vanuatu



Je n’ai pas longtemps résisté au désir de comparer le mwotlap avec d’autres langues voisines – conscient de l’éclairage essentiel que ce travail de défrichage allait m’apporter. Mon travail sur l’araki m’avait déjà donné quelques points de repère, et révélé à quel point les langues du Vanuatu pouvaient être différentes les unes des autres ; mais j’avais besoin d’aiguiser mes outils d’observation, en étudiant les langues voisines du mwotlap. Tout près de Motalava, j’ai ainsi consacré une dizaine de jours en juin 1998 à explorer l’île voisine de Vanua Lava (*Figure 3*). Là je m’initiai non seulement à la langue principale le **vurës**, mais aussi à deux langues à la vitalité menacée : le **mwesen** (une dizaine de locuteurs) et le **lernerig** (alors cinq locuteurs, aujourd’hui deux). Ce premier aperçu avait eu la vertu de soudain donner du relief à la langue mwotlap, en plaçant ses structures dans son contexte aréal. Ainsi plusieurs sections de ma thèse, écrite entre 1999 et 2001, impliquent un raisonnement comparatif entre ces langues voisines.

¹ Je reviendrai plus loin sur mes données mwotlap, et sur mon travail sur cette langue.

Cette brève expédition exploratoire de 1998 me permit également de concevoir ce qui allait plus tard venir au centre de ma recherche : la dialectologie régionale, et l'étude approfondie de la micro-diversité linguistique au nord du Vanuatu. En 2003, j'allais ainsi entreprendre une étude approfondie de la région Torres–Banks, qui allait s'échelonner sur cinq ans à raison de 2 à 3 mois par an sur le terrain. Au fil des années, j'ai ainsi pu accumuler une somme importante de données linguistiques (mais aussi, par ailleurs, ethnographiques et ethnomusicologiques) sur les 17 langues parlées dans l'archipel Torres–Banks. Le matériel relevé inclut des énoncés, conversations, interviews, récits, chants et poèmes [§2.2.3]. Ces données m'ont permis d'écrire, depuis 2004, une série d'articles de nature comparative, dans une perspective tantôt synchronique, et tantôt historique [§5].

Pour la plupart des langues, je faisais l'effort d'apprendre à m'exprimer, tâche considérablement aidée par ma connaissance préalable de la langue mwotlap [§2.2.4]. Certaines langues ont retenu particulièrement mon attention, et feront l'objet de publications spécifiques à l'avenir [§8.1]. Je pense notamment au **dorig** et au **lakon** (Gaua), au **lemerig** (Vanua Lava), au **lo-toga** et au **hiw** (îles Torres).

2.1.3 Langues océaniques hors du Vanuatu

En 2005 puis en 2012, j'ai saisi l'opportunité de me rendre dans une région mal connue et difficile d'accès, au nord des îles Torres–Banks : il s'agit de l'archipel des Santa Cruz, et plus particulièrement de l'île de Vanikoro, à l'extrême Est de l'état des Îles Salomon (cf. *Figure 1* p.9).

Ces deux missions avaient eu lieu à la faveur de projets pluridisciplinaires : en 2005, j'étais invité par Jean-Christophe Galipaud, archéologue à l'IRD, pour collaborer à l'expédition *Vanikoro–Lapérouse 2005* [§6.3.1] ; en 2012, je me joignais à une mission de géologues de l'IPGP–CNRS partis effectuer des mesures tectoniques dans cette île. Dans les deux cas, mon propre travail combinait une collaboration pluridisciplinaire (enquête sur la tradition orale concernant Lapérouse, enquête sur les perceptions locales des risques géologiques) à une enquête proprement linguistique sur les langues parlées dans l'île de Vanikoro. J'ai ainsi pu découvrir des langues océaniques étonnamment différentes de celles du Vanuatu que je connaissais. Ma surprise portait également sur la diversité interne à ce petit groupe linguistique, entre les trois langues **teanu**, **lovono** et **tanema**. Ceci allait fournir la matière à une publication [2009a ▶(**#20**)], et à une réflexion plus générale sur les processus de diversification linguistique en Mélanésie [§5.2]. Le **teanu**, principale langue de l'île, fait l'objet de plusieurs projets en cours [§8.1, 8.2].

L'île de Vanikoro comprend en outre une langue polynésienne, le **tikopia**, à laquelle j'ai consacré quelque temps sur le terrain. Mais si la famille polynésienne occupe une place particulière dans mes recherches, c'est moins par ce bref terrain sur le **tikopia**, que par le projet ambitieux d'*Atlas Linguistique de Polynésie Française* dont je suis l'un des deux coauteurs [Charpentier & François, ss presse ▶(**#32**)] [§4.2].

2.2 Recueillir la langue spontanée : méthodologie

2.2.1 Description et documentation des langues

Le terme de *documentation des langues*, popularisé depuis quelques années, est né d'un article de Himmelmann (1998), et développé dans un ouvrage collectif (Gippert, Himmelmann, Mosel 2006). Lorsque j'ai entrepris mon premier terrain en 1997, le terme n'avait pas encore été créé – et surtout, nul n'avait pas encore pensé à contraster, voire présenter comme antagonistes, le travail de *description* et de *documentation* linguistique.

En effet, Himmelmann (2006:18) souligne que ces deux facettes du travail du linguiste ne visent pas les mêmes objectifs. D'un côté, la production de grammaires et de dictionnaires s'adresse souvent à un public de linguistes, et implique une analyse abstraite des règles et principes d'une langue donnée, d'une manière qui dépendra souvent des théories linguistiques du moment. De l'autre côté, la documentation des langues vise à garder une trace durable (a “lasting record”) d'une langue telle qu'elle est réellement employée dans divers contextes de communication. La documentation met l'accent sur les productions linguistiques spontanées des locuteurs, à travers la création de corpus d'enregistrements archivables. Quant à la description des langues, elle concentre son effort sur l'analyse des systèmes ; ce sont souvent ces analyses-là qui seront publiées, plutôt que les corpus de parole qui les sous-tendent.

Personnellement, j'ai toujours considéré ces deux approches comme deux facettes inséparables de mon projet, qui allaient se nourrir l'une de l'autre : aussi n'ai-je pas souhaité choisir entre documentation et description des langues (cf. Evans 2008). Certes, mes travaux accordent une place importante à l'analyse des systèmes ; mais ils se fondent toujours sur la parole spontanée – en vertu d'un parti pris de fond qui privilégie le discours authentique, situationnellement ancré dans la réalité des interactions sociales, de préférence à l'élicitation ou aux questionnaires [§2.2.2].

En outre, j'ai été alerté au bon moment sur les meilleures pratiques en matière d'archivage des corpus linguistiques, par plusieurs collègues – en particulier Michel Jacobson alors qu'il développait l'archive des langues sonores du LACITO, la future *Collection Pangloss* ; Peter Austin, directeur de l'*Endangered Languages Documentation Project* de Londres ; ou encore Nick Thieberger, de l'Univ. de Melbourne, et fondateur, avec Linda Barwick, de l'archive PARADISEC (*Pacific and Regional Archive for Digital Sources in Endangered Cultures*). Leurs conseils, accompagnés de lectures ciblées (par ex. Jacobson 2004 ; Woodbury 2003, 2011 ; Thieberger 2012), m'ont aidé à archiver mes enregistrements de terrain, mes textes transcrits et même mes cahiers de terrain, en suivant ces meilleures pratiques autant que possible [§2.2.3, 6.2.2].

2.2.2 Questionnaires vs. énoncés authentiques

L'étape préalable à toute description consiste à réunir un corpus, qui permette par la suite d'étayer toute affirmation sur des bases empiriques. Je ne me satisfaisais pas des questionnaires préfabriqués, conçus par des linguistes pour tester tel ou tel point déjà identifié. S'il est vrai que les premiers jours d'exploration d'une langue peuvent éventuellement mettre en jeu des listes de mots isolés – parties du corps, termes de faune et

flore... – l'exercice atteint vite ses limites. On est parfois surpris par le nombre de publications, en particulier en linguistique historique, qui fondent leurs résultats sur la seule liste de Swadesh, voire sur d'autres outils intellectuellement encore plus frustes, comme les listes réduites utilisées en glottochronologie moderne (Holman *et al.* 2011). D'autres travaux de syntaxe ou de sémantique font également un usage conséquent des questionnaires. Malgré leur intérêt potentiel dans le cadre d'études de typologie – à condition d'être bien conçus, ce qui n'est pas toujours le cas – ces questionnaires ont l'inconvénient de préjuger des résultats que l'on est censé trouver. Si j'avais cherché à décrire le système aspectuel du mwotlap à l'aide du questionnaire d'Östen Dahl – pourtant un outil assez riche et élaboré – j'aurais manqué un bon nombre des morphèmes TAM de la langue, et serais passé à côté des contextes les plus pertinents. Ce ne sont là que quelques-uns des inconvénients des questionnaires. Dans le pire des cas, le terrain se réduit à une succession de séances d'élicitation fondées sur la traduction d'exemples écrits et pensés selon les catégories des langues déjà connues.

Or, je ne concevais pas mon terrain comme une simple étape nécessaire pour répondre à des questions préalables : mon intention était de recueillir la langue telle qu'elle était parlée spontanément par ses locuteurs. J'ai toujours eu une extrême exigence sur la qualité de la langue que je voulais recueillir : il me fallait recueillir des énoncés réellement prononcés, *in vivo*, par des locuteurs engagés dans des situations réelles de leurs vies. Clairement, les questionnaires préfabriqués ne remplissaient pas cette condition ; tout au plus pouvaient-ils constituer, éventuellement, une étape dans l'apprentissage de la langue.

Mon intuition à ce sujet m'était naturelle, mais avait été renforcée, au cours de ma formation de linguiste, par les enseignements d'Antoine Culioli – professeur de linguistique à l'École Normale Supérieure. S'il est vrai que j'ai eu relativement peu d'interactions personnelles avec A. Culioli lui-même, ses idées dominaient la scène linguistique parisienne durant mes années de formation. Sa *théorie des opérations énonciatives* était alors citée et commentée par la plupart des enseignants et chercheurs qui m'entouraient, ou dont je lisais les écrits... Il ressortait de ces travaux une extrême rigueur lorsqu'il s'agit de juger de la *bonne formation des énoncés* dans le discours réel. C'est ce type d'intuitions qui m'a fait comprendre, par exemple, qu'un énoncé français comme (1a) est, certes une *phrase* grammaticalement correcte, mais un *énoncé* "mal formé" – alors que d'autres apparemment similaires, comme (1b) ou (1c), seraient parfaitement acceptables :

- (1a) ?? *Tu es malade.*
- (1b) *Toi, tu es malade...*
- (1c) *T'es MA-LADE !!*

Or, la plupart des cadres d'analyse linguistique – chomskyens ou non – traiteraient ces trois énoncés comme ayant "la même structure", et donc le même statut vis-à-vis de la "grammaticalité". Ces cadres se trouvent alors incapables de comprendre les différences entre ces trois énoncés – ou pire : ils ne se donnent pas les moyens de *détecter* que (1a) pose un problème. A. Culioli m'a ainsi transmis une exigence dans l'observation, une attention permanente portée aux propriétés précises (prosodie, enchaînement discursif, structure informationnelle...) des énoncés que j'allais être conduit à décrire.

Au passage, l'influence d'A. Culioli est perceptible dans mes premiers travaux universitaires – qu'il s'agisse de ma maîtrise sur l'aspect verbal en grec ancien (1994), ou de mon DEA sur la typologie de la subordination sans marques segmentales (1997). Mon tout premier article [1998] reflète également l'attention que je portais alors, à travers les idées de Culioli, à la question de l'engagement de l'énonciateur dans son énoncé :

François, Alexandre. 1998. *Ben* en français oral, l'énonciateur (dés)engagé. In *Proceedings of the XVIth International Congress of Linguists*, edited by B. Caron. Oxford: Pergamon.

J'y proposais une analyse de la particule discursive *ben* du français parlé, particule aux multiples valeurs (hésitation, exclamation, assertion...). Selon cette analyse, *ben* servirait à encoder l'attitude de l'énonciateur vis-à-vis de son énoncé – une attitude paradoxale, où le sujet assertant s'engage personnellement sur la validité de cet énoncé, mais refuse de prendre la responsabilité sur son contenu ; *ben* y était ainsi glosé comme une particule de "désengagement énonciatif". Ce travail portant sur le français oral n'avait été possible qu'en prêtant attention à la forme précise des énoncés, et la nature exacte de leur contexte pragmatique d'énonciation. Dans mon approche, la présence de telle particule énonciative, ou la longueur d'une pause, étaient des points cruciaux dans l'observation ; or de tels détails sont rarement pris en compte dans l'approche plus classique des questionnaires de typologie, construits avant tout autour de la notion de grammaticalité.

Si mes travaux postérieurs à 1998 ont abandonné toute tentative de formalisation dérivée de la théorie de l'énonciation, j'ai pu en retrouver l'esprit à l'occasion de certains chapitres de ma monographie sur l'aspect verbal [2003a ►(#30)]. Et surtout, j'ai gardé de l'approche culiolienne une exigence d'authenticité dans mon recueil des données linguistiques. Ainsi, j'ai toujours eu du mal à raisonner sur des phrases aussi mal formées que (2)–(5) :

- (2) *The man goes.*
- (3) *A man sees a woman.*
- (4) *What do you say John said you said?*
- (5) *She eats the banana.*

Or, de telles phrases sont citées *ad nauseam*, comme si elles allaient de soi, dans nombre d'articles ou de grammaires de référence. Pourtant, elles sont toutes mal formées, pour diverses raisons (problèmes de détermination nominale, de marques aspectuelles...), et ne sont, de ce fait, *jamais prononcées* ailleurs que dans des questionnaires de linguistes. La cause exacte de ces cas de "mauvaise formation" peut porter à débat ; mais il me semble au moins indispensable à tout bon linguiste d'être sensible à ces questions, et de savoir repérer les énoncés dicibles des énoncés non dicibles.

Cette exigence d'authenticité est partagée, à divers degrés, par la plupart des linguistes de terrain, en particulier lorsqu'ils souhaitent, comme moi, recueillir la langue telle qu'elle est réellement pratiquée. Dans ce sens, mes pratiques personnelles partageaient bien des points avec le mouvement de la documentation des langues.

2.2.3 La constitution du corpus

Mon approche empirique de la description linguistique nécessitait donc d'élaborer un corpus d'énoncés. Ce travail a notamment consisté en enregistrements audio et vidéo, et en transcription de textes. Par ailleurs, j'ai trouvé particulièrement avantageux d'apprendre à parler certaines des langues sur lesquelles je travaillais [§2.2.4].

2.2.3.1 Les enregistrements

Outre les conversations d'oral spontané, la principale source de mon corpus a pris la forme d'enregistrements audiophoniques (et dans une moindre mesure, vidéo).

Entre 1997 et 2012, j'ai effectué 1096 enregistrements audio sur le terrain, soit un total de 104 heures, portant sur 23 langues différentes. Les récits de tradition orale [§2.2.3.2] constituent environ la moitié de ce total. Le reste est constitué d'autres genres de nature diverse : textes procéduraux de type instructions ; interviews (portant typiquement sur des questions d'ordre anthropologique, mais parfois aussi biographiques) ; conversations spontanées entre locuteurs natifs... Par ailleurs, à ce corpus proprement linguistique, il convient d'ajouter un ensemble important d'enregistrements musicaux, qu'il s'agisse de chants, de musique instrumentale, ou d'événements collectifs autour de la musique et de la danse [§6.3.2]. À ce corpus essentiellement audio, s'ajoutent des données vidéo : en tout une cinquantaine d'heures de rushes, pour la plupart tournées par l'anthropologue Éric Wittersheim lors de notre projet commun [§6.3.2]. Ces films incluent des scènes de village, des performances musicales et dansées, ainsi que diverses conversations en langue mwotlap.

Le *Tableau 4* donne une vue d'ensemble de mon corpus audio. Ces enregistrements ont été archivés sur *Pangloss* [<http://lacito.vjf.cnrs.fr/pangloss/>], la plateforme en ligne d'archivage sonore des langues rares que développe mon laboratoire le LACITO [§6.2.2]. Par ailleurs, en 2011 j'ai rassemblé ces enregistrements dans une médiathèque multi-média que j'ai fondée à Motalava, à destination des communautés de locuteurs [§6.2].

Tableau 4 – Enregistrements audio effectués sur le terrain (1997-2012) : Statistiques

		<i>nombre d'enregistrements</i>	<i>durée totale</i>
Par pays	<i>Vanuatu</i>	1002	97 h 25'
	<i>Îles Salomon</i>	94	6 h 46'
Par langue	<i>mwotlap</i>	551	56 h 32'
	22 autres langues	545	47 h 39'
Par genre	<i>récits</i>	389	50 h 04'
	<i>chants</i>	218	16 h 08'
	autres genres	356	29 h 15'
Total		1096	104 h 24'



2.2.3.2 *Transcription, traduction, corpus*

Au fil des années, j'ai enregistré 389 récits de tradition orale (cf. *Tableau 4*), dans 22 langues différentes. Parmi eux, 263 (= 68 %) ont fait l'objet d'une transcription sur le terrain, condition nécessaire pour être exploités en tant que corpus linguistique.

Ces textes issus de la tradition orale remplissent plusieurs fonctions. Tout en constituant un corpus de textes indispensable à l'analyse linguistique, ces enregistrements permettent également de garder une trace tangible de la très riche littérature orale qui a survécu jusqu'à nos jours dans cette région. En les recueillant, je satisfaisais ma passion – née de *l'Odyssée* ou des épopées peules – pour la littérature orale, en même temps que je préservais un patrimoine précieux aux yeux des communautés de locuteurs. Mon approche holistique impliquait de recueillir non seulement la langue elle-même – ses structures, sa grammaire, son lexique – mais également ses différents genres discursifs. Même si une langue n'est pas forcément en danger, plusieurs genres discursifs traditionnellement donnés dans cette langue (chants, poésie orale, mythologie) peuvent facilement être menacés de disparaître. Lorsqu'en 2011 j'ai donné un cours d'écriture mwotlap à des adolescents de Motalava, tous locuteurs de la langue, j'ai été effaré de constater qu'ils ignoraient jusqu'à l'existence d'Iqet, le héros mythologique fondateur de leur civilisation. Pourtant, on sait que l'ancrage des individus dans leur héritage culturel, et en particulier la connaissance de la littérature orale, sont l'un des facteurs en jeu dans le sentiment de bien-être à l'âge adulte (MNCC 2012:40). Les textes que j'ai recueillis fournissent ainsi la matière première aux nombreux livres d'alphabetisation et de lecture que j'ai réalisés au fil des années dans les diverses langues vernaculaires [§6.2.1]. Plusieurs d'entre eux ont été publiés avec une traduction : certains dans ma grammaire de l'araki, d'autres en ligne¹. D'autres textes vont être réunis sous forme d'ouvrages bilingues [§8.4].

¹ Exemple : [http://alex.francois.free.fr/AFtxt_msn-kpwet_fr.htm]. Cf. *Figure 29* p.122.

Mais en dehors de leur valeur littéraire et anthropologique intrinsèque, ces textes recueillis sur le terrain présentent également un intérêt évident pour le linguiste. Ils constituent un vaste corpus de parole spontanée, associé à des enregistrements sonores de bonne qualité, vérifiables, et organisés en discours thématiquement cohérents. Sur les 263 textes manuscrits dans mes cahiers de terrain, j'en ai d'ores et déjà recopié 168 dans le logiciel de base de données *Toolbox*, constituant ainsi un corpus informatisé aisément interrogeable. Le *Tableau 5* donne un aperçu de la taille de mes corpus, y compris le nombre de mots pour les neuf langues les mieux représentées.

Tableau 5 – Corpus textuels: Quelques statistiques

<i>Langue</i>	<i>nb textes transcrits sur le terrain</i>	<i>nb textes transcrits et informatisés</i>	<i>nb de mots dans corpus informatisé</i>
Lakon	10	8	8 876 mots
Araki	16	16	10 143 mots
Lemerig	5	4	10 470 mots
Mwesen	6	6	10 491 mots
Dorig	11	10	12 187 mots
Teanu	38	18	17 217 mots
Hiw	18	12	17 608 mots
Lo-Toga	25	14	21 313 mots
Mwotlap	103	52	99 302 mots
<i>13 autres langues</i>	31	28	...

Si l'on ajoute à ces corpus textuels mes abondantes notes issues des conversations [§2.2.5], on obtient une photographie assez riche et fidèle de la langue parlée telle qu'elle est pratiquée dans son environnement social. Pour chaque langue concernée, ce corpus écrit issu de l'oral spontané constitue la base empirique de mes observations. Dans mes descriptions grammaticales [§3] comme mes dictionnaires [§4.1], je m'attache toujours à y puiser mes exemples. Ainsi, toutes mes hypothèses linguistiques sont rendues vérifiables, en principe, grâce à un patient processus d'archivage et de diffusion de mes données – qu'il s'agisse de l'archivage des enregistrements sonores (sur *Pangloss*) ou des cahiers de terrain (sur *ODSAS*, cf. §2.2.5).

2.2.4 Pour un terrain monolingue

Pour plusieurs langues, mes enquêtes mêlaient un travail classique d'élicitation avec le recueil de corpus tel que je viens de le décrire. Lorsque mon enquête se limitait à une ou deux semaines, je me contentais parfois d'interactions dans notre langue de contact, le bislama, et recueillais les informations au cours de séances de travail. Cette approche classique du terrain a par exemple concerné les langues araki, lemerig, mwesen, mota, nume, koro, olrat, mwerlap, lovono, tanema, tikopia.

Pourtant, pour d'autres langues, j'ai voulu aller plus loin, en apprenant à converser dans la langue même. Ce fut surtout le cas pour le **mwotlap** (langue que j'ai appris à parler couramment) mais aussi pour les langues **dorig**, **lakon**, **lo-toga**, **hiw**, **vurès**, **vera'a**, et **teanu**. Sans le savoir, je mettais ainsi en œuvre l'idée du terrain monolingue

(“monolingual fieldwork”) que préconisent les anthropologues de la tradition boasienne, ainsi que certains linguistes de terrain (Everett 2001, Hale 2001:88).

Certes, l'apprentissage de la langue requiert des efforts supplémentaires au moins dans la phase initiale de l'enquête ; il implique d'effectuer des séjours longs et répétés sur le terrain, et oblige à un travail cognitif de mémorisation et d'analyse que l'on pourrait juger superflu. Pourtant, ces efforts sont vite payés de retour. Pour assurer le succès de la communication, le linguiste-apprenant est conduit à formuler ses propres hypothèses concernant les régularités de la langue, en étant attentif à tous ses aspects à la fois : phonétique, prosodie, morphologie, pragmatique. Plutôt que de simplement noter ce qu'il entend, l'enquêteur qui s'essaie à parler aura l'occasion de faire des fautes, d'être corrigé, de se reprendre jusqu'à formuler un énoncé en tous points acceptable – non seulement une phrase ‘grammaticale’, mais un énoncé *bien formé*, énonciativement et pragmatiquement [§2.2.2]. Ces moments sont cruciaux dans l'apprentissage, car c'est alors que l'apprenant ajuste les premières hypothèses de règles et généralisations qui s'étaient un temps formées dans son esprit. Cette méthode constitue un outil heuristique efficace pour comprendre la grammaire d'une langue.

Enfin, un autre avantage essentiel d'apprendre la langue ne se situe pas du côté de la production, mais de la compréhension. Lorsque l'on se contente d'étudier une langue par le truchement d'une autre – comme j'ai pu le faire pour l'araki, etc. – on est tributaire du corpus de textes que l'on a recueilli et transcrit, et les questions à poser sont largement dépendantes de ce corpus. À l'inverse, le linguiste qui apprend la langue devient soudain capable de comprendre les conversations autour de lui, et d'y participer. Les locuteurs choisiront spontanément de s'adresser à lui dans leur langue, introduisant à chaque fois nombre de tournures et de constructions nouvelles. Le “corpus” du linguiste s'enrichit de toutes ces interactions, et dépasse alors largement celui des enregistrements transcrits ; il s'accroît naturellement, au fil des conversations. Si j'ai pu consacrer un ouvrage entier à la sémantique des marques de Temps-Aspect-Mode en mwotlap [2003a ▶(#30)], c'est en me fondant sur les observations approfondies que j'avais pu effectuer dans la conversation quotidienne : mon apprentissage patient m'avait donné les moyens de comprendre les nuances sémantiques subtiles qui opposent le Parfait au Prétérit (François 2003a:104-117) ou le Futur Hodiernal au Prospectif (François 2003a:266-267). Je n'aurais jamais eu ces intuitions à partir d'un corpus purement textuel, trop peu varié en TAM, encore moins à partir d'un questionnaire.

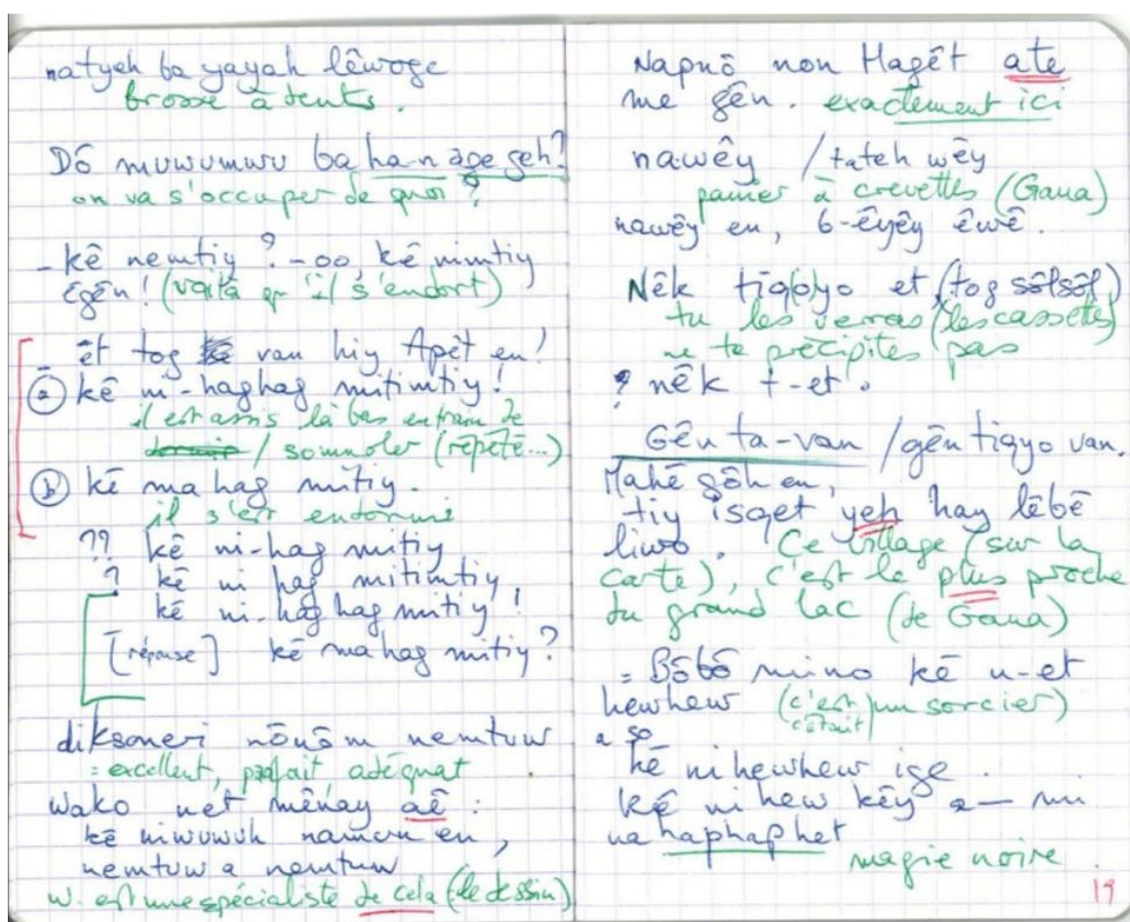
En somme, apprendre à parler, c'est aussi apprendre à comprendre la parole avec la même efficacité et les mêmes implications que les locuteurs eux-mêmes. En s'exposant ainsi au dialogue, on acquiert des intuitions personnelles sur la bonne formation des énoncés, sur les différents sens de telle expression, sur le degré de productivité, ou au contraire l'archaïsme, de telle construction. Il devient alors possible de formuler des hypothèses sur la manière dont les locuteurs eux-mêmes manipulent leur propre compétence linguistique.

2.2.5 Les cahiers de terrain

Pour ces langues que j'avais choisi d'apprendre, je partageais mon temps entre, d'une part, le travail sur les textes (enregistrement, transcription, traduction), et d'autre part,

l'apprentissage de la conversation. J'allais à la rencontre des locuteurs de tous âges pour engager la conversation, notant scrupuleusement toute nouvelle expression dans de petits carnets de terrain que je gardais en permanence sur moi. En principe, tout ce que j'ai appris durant mes années d'exploration – qu'il s'agisse du mwotlap ou d'autres langues – a laissé une trace écrite, sous la forme de **cahiers de terrain**. C'est d'ailleurs la seule condition pour que ces conversations puissent être considérées comme faisant partie intégrante de mon corpus au même titre que les textes transcrits.

Figure 4 – Extrait d'un cahier de terrain en langue mwotlap
[réf. AF-AP09-10, http://www.odsas.net/scan_sets.php?doc=104523]



La Figure 4 donne un échantillon au hasard de ces cahiers de terrain – en l’occurrence, quelques notes prises lors d’une conversation en langue mwotlap. Au fil des années, j’ai rempli sur le terrain 71 cahiers, totalisant **4156 pages** manuscrites. Afin de préserver ces documents, tous les cahiers ont été scannés, puis archivés sur ODSAS, la plateforme d’archivage de documents ethnographiques du CREDO (Centre de Recherche et de Documentation de l’Océanie, Marseille), développée par Laurent Dousset.¹

¹ Lien: [http://www.odsas.net/index.php?action=collection_category&cat=author&value=Francois,Alexandre].

2.3 Parole spontanée et approche constructionniste

Pouvoir participer aux conversations du village offre un avantage crucial pour ma recherche : tout énoncé se trouve naturellement inscrit dans son contexte pragmatique, celui auquel il est destiné. La démarche était cohérente avec le but que je m'assignais dans mon enquête, à savoir, comprendre comment les structures idiomatiques du langage sont façonnées par la pratique dynamique du discours.

Un principe sous-jacent à ma pratique de terrain a donc toujours été de noter des *énoncés en situation*, plutôt que des mots isolés. Cette exigence permet d'accéder aux mots dans des contextes bien précis, clé de leurs ambiguïtés et des polysémies que l'on abstraira plus tard. Le cheminement cognitif du linguiste sur le terrain présente alors certaines similitudes avec l'apprentissage, par le jeune apprenant, de sa propre langue : les unités de discours entendues sont toujours des énoncés, ancrés dans des situations sociales réelles, d'emblée affublés de marquage prosodique et d'intentions pragmatiques précises. Les outils que l'on apprend ainsi à maîtriser ne sont pas des morphèmes ou des mots – ce sont des formules entières, des constructions, des stratégies discursives qui se combineront de manière plus ou moins productive, plus ou moins libre.

Le terrain m'a ouvert les yeux sur un point important. L'unité fondamentale de la langue idiomatique n'est ni le mot ni le morphème, mais ce que j'appelle le **phrasème** [2003a:253, 344] : une formule ou construction éventuellement complexe, opérant un couplage forme/fonction mémorisé tel quel par le locuteur. Ces phrasèmes ne font surface qu'à travers des énoncés, événements langagiers qui sont inscrits dans un temps, un lieu et un espace social spécifiques. Ce n'est qu'après avoir formulé (d'abord dans ma thèse, 2001) ma conception du *phrasème* que j'ai découvert la **Construction grammar**, fondée par un célèbre article de Fillmore, Kay & O'Connor (1988), et théorisée plus tard par Bill Croft dans son ouvrage *Radical Construction Grammar* (Croft 2001). Or, son concept de *construction* est très proche de mon intuition du *phrasème*, unité fondamentale de l'expression.

Ce sont avant tout ces constructions, ou phrasèmes, que manipulent les locuteurs, souvent sans qu'il leur soit même nécessaire de recourir aux supposées unités minimales du type morphème. D'ailleurs, des études comme Pawley (1994) ont critiqué à juste titre la conception entièrement analytique de la langue (en anglais, "parsimonious"), car elle occulte l'importance des "formules" et des constructions routinières dans la langue, où le tout ne peut pas se déduire de ses parties.

Il serait tentant de résumer la situation en disant que je définissais mes méthodes de terrain en fonction de mes lectures théoriques sur la grammaire de construction. En réalité, ce serait déformer mon cheminement intellectuel, qui a suivi l'ordre inverse. C'est ma pratique du terrain, et l'observation empirique des processus à l'œuvre dans la conversation, qui m'a ouvert les yeux sur le rôle central joué par les phrasèmes dans la mécanique de la langue. Plus tard, mes lectures m'ont permis de reconnaître dans la *Construction grammar* une traduction adéquate des idées que j'avais eues sur le terrain.

3 La description grammaticale : catégories et constructions

La typologie des langues s'assigne comme tâche de comparer les structures des langues du monde, en s'attachant aussi bien à leurs similitudes qu'à leurs différences. Pour pouvoir mettre en œuvre un tel programme de comparaison, une première étape essentielle est de fournir de chacune de ces langues une description grammaticale. Depuis Saussure et Boas, on sait que toute langue constitue un système unique, constitué de catégories émiqes et de constructions particulières. Dans la mesure où j'ai recueilli des données abondantes sur un ensemble de langues jusque-là non décrites, ma contribution à l'effort typologique prend avant tout la forme de travaux descriptifs, visant à faire apparaître ces catégories et ces constructions.

Mes recherches se sont d'abord concentrées sur deux langues en particulier, l'araki et le mwotlap, avant de s'élargir aux langues qui l'entourent : langue des îles Torres et Banks (Vanuatu), langues de Vanikoro (îles Salomon), etc. Cet élargissement a notamment donné lieu à des travaux comparatifs et historiques, lesquels feront l'objet du *Chapitre 5*. Quant au présent *Chapitre 3*, il porte sur mes travaux de **description des systèmes grammaticaux dans une perspective synchronique**, en se concentrant généralement sur une seule langue à la fois (parfois deux : §3.4.4). Ces langues incluent surtout le mwotlap, mais aussi l'araki, le dorig, le hiw, le lo-toga. Le *Chapitre 4* traitera de la description lexicale.

Tout, dans ces langues, nécessitait d'être décrit, et j'ai fait en sorte d'aborder les principaux domaines de ces systèmes dans mes travaux. Cette variété des thèmes de recherche allait de soi, en particulier, dans les deux grammaires de référence que j'ai consacrées, l'une à l'araki, l'autre au mwotlap [§3.1]. En plus de ces monographies, les articles que j'ai publiés approfondissent certaines questions dans plusieurs domaines :

- les **systèmes phonologiques**, notamment les contraintes phonotactiques [§3.2] ;
- les *parties du discours* ou **catégories syntaxiques**, et notamment les fonctions des noms [§3.3] ;
- la description des **systèmes de Temps–Aspect–Mode** [§3.4], à la fois dans leur sémantique interne et dans leurs relations avec d'autres paramètres de la langue, comme la reduplication [§3.4.3] ou la dépendance interpropositionnelle [§3.4.4] ;
- la syntaxe des séries verbales et des **prédicats complexes** [§3.5], et la question des relations actanciennes dans les constructions **ditransitives** [§3.5.3] ;
- les **systèmes d'orientation dans l'espace** [§3.6].

Je n'offrirai ici qu'un résumé des thématiques abordées et de mes résultats ; pour le détail des démonstrations, des exemples et des références, le lecteur pourra se reporter aux ouvrages ou aux articles eux-mêmes, qui accompagnent ce volume.

Dans toutes ces études, je me suis efforcé de dégager les catégorisations émiqes de chaque langue, et d'observer les constructions qui les mettent en relation ces catégories les unes avec les autres. Chaque analyse s'efforce de modéliser la manière dont le locuteur exploite le système de sa langue pour élaborer son discours, en situation.

3.1 Décrire chaque langue comme un système

Dans un article inspiré de l’approche structurale et notamment des travaux de Gilbert Lazard, Haspelmath (2007a) affirme que les “**catégories pré-établies n’existent pas**”, réaffirmant ainsi la nécessité d’identifier les catégories de chaque langue de manière *émique* – c’est-à-dire en observant les caractéristiques internes du système plutôt qu’en se fondant sur de putatifs universaux ou des hypothèses *a priori* (cf. Evans & Levinson 2009). Décrire une langue, c’est faire émerger, par une procédure empirique, l’ensemble des catégories émiques qui sont mises en jeu par le système, ainsi que les relations et règles qui définissent leur mise en œuvre dans la construction du discours. “Analyser les langues dans leurs propres termes” : ce programme structuraliste est, selon Dryer (2006:211), au cœur de la théorie descriptive employée par la plupart des typologues aujourd’hui – approche qu’il appelle *Basic Linguistic Theory*, à la suite de Dixon (1997).¹ C’est ce programme que j’ai tenté d’appliquer dans mes divers travaux descriptifs : leur principe est toujours d’*analyser les langues dans leurs propres termes*.

En termes de publications, ces travaux incluent d’abord ma thèse de doctorat sur la langue mwotlap :

François, Alexandre. 2001 a. **Contraintes de structures et liberté dans l’organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu**. Thèse de doctorat en Sciences du Langage, Université Paris-IV Sorbonne, Paris. (1078 pages, 472 000 mots) ▶(#28)

Cette thèse fut suivie de peu par ma grammaire de la langue araki :

François, Alexandre. 2002. **Araki. A disappearing language of Vanuatu**. Pacific Linguistics, 522. Canberra: Australian National University. ▶(#29)

Commune à ces deux descriptions grammaticales était l’entreprise de décrire l’ensemble d’une langue, depuis les unités minimales du signifiant (les phonèmes) jusqu’au niveau plus élaboré de la phrase complexe et du discours.

Il n’est pas aisé de résumer une grammaire – un travail dont le principe même est d’aborder une grande diversité de thèmes, en observant les mille rouages de la grande machine. Pourtant, des lignes de force se dessinent au fil des chapitres, permettant parfois de faire émerger des caractéristiques récurrentes au sein d’un même système.

Prenons un exemple, celui de la catégorie HUMAIN. Les descriptions du mwotlap et de l’araki m’ont permis de montrer la place centrale jouée par la classe des référents humains dans l’architecture de leurs grammaires respectives. Ainsi en mwotlap, la catégorie “humain” est la seule qui soit compatible avec le marquage du nombre, lequel s’articule en quatre catégories (singulier, duel, triel, pluriel). Les référents humains sont associés à un système riche de marquage en nombre, incluant pronoms, suffixes possessifs, reduplication [§3.4.3], article nominal :

¹ Ce principe est également au cœur de l’article d’encyclopédie que j’ai coécrit avec Maïa Ponsonnet, intitulé *Descriptive Linguistics* (François & Ponsonnet 2013) ▶(#38).

- (6) **Nō-lōmgep** su en, **kē ni-van** hiy imam no-no-**n**.
 ART:SG-garçon petit TOP 3SG 3SG:AOR-aller DAT père ART:SG-POSS-3SG
 [SING] ‘Le petit garçon (il) se rendit auprès de son père.’
- (6’) **Yoge lōmlōmgep** susu en, **kōyō** (Ø-)van hiy imam no-no-**yō**.
 ART:DU garçon:DUP petit:DUP TOP 3DU n.SG:A0-aller DAT père ART:SG-POSS-3DU
 [DUEL] ‘Les deux petits garçons (ils) se rendirent auprès de leur père.’

À l’inverse, les référents non-humains – qu’ils soient animés ou non – sont incompatibles avec les contrastes de nombre, et seront toujours traités formellement au singulier – cf. (7).

- (7) **Nō-mōmō** te-le-naw en, woqse ha-**n** !
 ART:SG-poisson de-LOC-mer TOP beaucoup.de nom-3SG
 [litt. le poisson de la mer, son nom est nombreux !]
 = ‘Il existe beaucoup de noms (différents) pour les poissons.’

C’est là une des caractéristiques qui font de la catégorie sémantique “humain” une clef de la langue mwotlap. Bien d’autres rouages de la langue font appel à cette même catégorie sémantique : le traitement de l’article nominal [§3.3.2], les suffixes de possession, la combinaison avec telle ou telle préposition...

Cette logique interne au système mwotlap, en particulier l’importance donnée au critère [humain], se trouve également au centre d’un article que j’ai publié dans *Linguistic Typology* :

François, Alexandre. 2005 a. A typological overview of Mwotlap, an Oceanic language of Vanuatu. *Linguistic Typology* 9-1: 115-146. ▷(#7)

Le principe de cet article – une idée proposée par Edith Moravcsik – était d’offrir une vue synthétique des principales règles de la grammaire mwotlap, et de les situer par rapport aux tendances typologiques observées dans chaque domaine (phonologie, marquage du nombre, syntaxe des actants, séries verbales, etc.).

L’araki accorde également une place privilégiée aux référents humains : encodage du nombre (singulier-pluriel), marquage différentiel de l’objet (cf. Lazard 1984, Bossong 1985), morphologie spécifique pour encoder la possession... Dans un article plus récent, j’ai montré que le trait [$\pm$ humain] jouait également un rôle dans les règles morphosyntaxiques régissant l’encodage des propositions ditransitives [cf. §3.5.3] :

François, Alexandre. 2012 a. Ditransitive alignment and referential hierarchies in Araki. In E. van Lier (ed), *Referential Hierarchies in Three-participant Constructions*, special issue of *Linguistic Discovery* 10 (3): 97-124. ▷(#12)

Dans de tels cas de figure, la description grammaticale consiste bien à identifier empiriquement un certain nombre de catégories, et à observer comment elles interagissent les unes avec les autres. Ces catégories sont parfois explicites (*overt categories*) : c’est le cas, par exemple, du *duel*, catégorie fonctionnelle associée à un marquage spécifique, et qui consiste à assigner à un référent un nombre de deux. D’autres catégories sémantiques sont implicites (*covert categories*) : c’est le cas de la classe des

humains, qui n'est associée à aucune marque en particulier, mais qui sous-tend tout un ensemble de règles grammaticales. Ces catégories ne prennent vie qu'à travers leurs interactions : en l'occurrence, on observe que la classe des référents [-humain] n'est pas compatible avec celle des duels, alors que la classe [+humain] est distribuée dans les quatre catégories de nombres. Décrire la langue, c'est identifier les lignes de fractures entre catégories, et les principes réguliers de leurs interactions.

Ces observations permettent ensuite la comparaison typologique, celle qui confronte les langues et observe les similarités et différences dans leurs processus de catégorisation, leurs façons de “découper la réalité”. Pour reprendre le même exemple, si l'on situe les systèmes araki et mwotlap dans une hiérarchie de saillance de type Silverstein (1976) ou Bickel (2008), on constatera que la frontière *humain/non-humain* est en effet une ligne de partage (a “cut-off point”) bien attesté dans les langues, mais loin d'être universel : en effet, d'autres langues isolent plutôt des catégories telles que “animé”, et accordent au trait [humain] une importance moindre que les langues du Vanuatu que je décris ici.

Mes travaux de description grammaticale ont également pris la forme d'un ouvrage sur le système Temps–Aspect–Mode, dont je parlerai en §3.4 ; ainsi que de plusieurs articles et chapitres, portant sur une variété de thèmes. Ceux-ci font l'objet des prochaines sections.

3.2 Phonologie

La phonologie constitue un domaine privilégié pour observer la manière dont les langues construisent des systèmes structurés et cohérents en synchronie.

3.2.1 Phonologie des voyelles du mwotlap

François, Alexandre. 1999. Mouvements et clonages de voyelles en motlav: Entre phonologie et morphologie. *Bulletin de la Société de Linguistique* xciv-1:437-486. ▷(#1)

François, Alexandre. 2000 a. Vowel shifting and cloning in Motlav : historical explanation vs. formal description. In M. Klamer (ed.), *Proceedings of AFLA7 (The Seventh Meeting of Austro-nesian Formal Linguistics Association)*. Amsterdam: Vrije Universiteit. Pp.49-68. ▷(#14)

Ces deux articles montrent comment le mwotlap utilise ses voyelles comme variable d'ajustement pour satisfaire les contraintes phonotactiques du système. En particulier, le schème syllabique strict (C)V(C) du mwotlap interdit les groupes tautosyllabiques de consonnes.

Il convient de distinguer quatre processus morphologiques ^{et/ou} phonologiques portant sur les voyelles :

- **HARMONIE VOCALIQUE :**
au sein d'un lexème, une modification de la voyelle thématique entraîne parfois une assimilation partielle de la voyelle précédente, avec harmonie vocalique en trait ATR :
ex. /iplu-k/ 'mon camarade' → /ɪplu-n/ 'son camarade'

- **COPIE VOCALIQUE :**
sans disparaître, la voyelle radicale transmet son timbre à la voyelle du préfixe
– par ex. l'article /nA-/ : ex. /nA-/ + /wɔl/ → /nu-wɔl/ 'la lune'
- **TRANSFERT VOCALIQUE :**
la voyelle radicale transmet son timbre à la voyelle du préfixe, et disparaît :
ex. /nA-/ + /hinaɣ/ → /nihnaɣ/ 'une/l'igname'
- **ÉPENTHÈSE VOCALIQUE :**
au sein même du lexème, la première voyelle du radical sous-jacent est employée
épenthétiquement pour briser les groupes de consonnes initiaux :
ex. Ø + /mtiɣ/ → /mitiɣ/ 'cocotier'.

Mes articles décrivent les processus, et en proposent une analyse autosegmentale inspirée de la phonologie non-concaténative de McCarthy (1989).

Le système phonologique mwotlap séparerait consonnes et voyelles en deux plans distincts, combinés ensuite autour d'un schème contraignant (C)V(C). Une même voyelle peut se réaliser en deux positions si le schème l'exige, comme le /i/ de /mitiɣ/, la forme non-préfixée [§3.3.2] du nom 'cocotier'. À l'inverse, une voyelle sous-jacente ayant le statut de phonème "flottant" peut ne pas se réaliser du tout, comme le /a/ de l'article nominal qui ne fait pas surface dans la forme préfixée /nu-wɔl/ 'lune'. En effet, la position préradicale est l'objet d'une règle hiérarchique, selon laquelle la position est remplie *a priori* par une copie de la voyelle radicale – ex. le /u/ de /wɔl/ – chaque fois que c'est possible, auquel cas elle a priorité sur la voyelle flottante du préfixe. Les seuls cas où la voyelle flottante se réalise sont ceux où la copie est bloquée – typiquement, bloquée par un groupe de deux consonnes, qui agissent comme un "nœud barrière" pour la propagation vocalique : ex. /nA-/ + /mtiɣ/ → /na-mtiɣ/ 'cocotier' (Figure 5).

Figure 5 – Les voyelles du mwotlap:

Propagation vocalique (copie) vs. "nœud barrière" (pas de copie)



Après avoir fourni une description synchronique de ces règles, l'une de ces études [AF 2000a] s'interroge également sur leur origine historique. Je montre que le schème syllabique (C)V(C) résulte historiquement de la syncope régulière des voyelles atones à partir d'un ancien squelette austronésien caractérisé par les syllabes ouvertes et l'accent trochaïque aligné à droite : *,CVCV'CVCV, ex. *nà-βetáli 'banane' > nap.tɛl. Seules les voyelles qui étaient alors accentuées (accent primaire ou secondaire) ont survécu dans la langue moderne, moyennant des phénomènes de type umlaut (*CáCi > CɛC). Ce que l'on peut décrire synchroniquement comme des "nœuds barrière" de deux consonnes reflètent les cas où la voyelle de l'article recevait l'accent secondaire (ex. *nà-βetáli > nap.tɛl). En revanche, cette même voyelle préfixale n'a pas pu préserver son timbre

dans les cas où elle se trouvait en position prétonique : ex. *na-βúla > nowól ‘la lune’.

L’explication historique permet également d’expliquer les exceptions à ces règles : par exemple, si le nom du ‘soleil’ *na-lo* présente un vocalisme irrégulier sur l’article, c’est du fait de la disparition d’une ancienne voyelle : la comparaison avec d’autres langues de la région révèle que /na-lo/ provient non pas de **na-lóa mais de *nà-alóa, forme quadrisyllabique dans laquelle l’article *na recevait l’accent secondaire. Enfin, je compare les mérites de l’explication historique et de la modélisation synchronique ; je conclus que les deux approches sont également nécessaires, car chacune peut expliquer des faits que l’autre ne peut pas.

Ces deux études de cas ne sont pas isolées. D’abord, elles m’ont permis de poser des jalons indispensables à mon analyse de la morphologie mwotlap, et à ma description de tout le système. D’autre part, mes enquêtes sur les langues voisines m’ont fait découvrir d’autres systèmes tout aussi fascinants ; et s’il est vrai que le mwotlap était sans doute la langue la plus complexe à cet égard, mes recherches sur cette langue m’ont permis de déchiffrer les systèmes voisins d’autant plus efficacement – autant dans leur synchronie que leur diachronie. Il en est résulté un article portant sur la comparaison des systèmes de voyelles dans les langues Banks & Torres [cf. §5.1.1], lui-même le point de départ à mes études historiques sur la région [§5.2], et mes recherches de linguistique comparée.

3.2.2 Phonologie des groupes consonantiques en hiw et en dorig

François, Alexandre. 2010 b. Phonotactics and the prestopped velar lateral of Hiw: Resolving the ambiguity of a complex segment. *Phonology* 27 (3):393-434. ▷(#9)

Cette étude sur la langue hiw a pour point de départ l’observation d’une consonne originale, et typologiquement très rare : une latérale vélaire préplodée [g̠L]. Outre la description du phonème lui-même, je cherche à en établir le statut phonologique exact : s’agit-il d’une occlusive postlatéralisée /g̠^L/ ? ou d’une approximante préplodée /g̠^L/ ?

Ce problème spécifique fournit le prétexte à une étude de type méthodologique sur la question des segments complexes, et des critères pertinents pour en définir le statut. Plutôt que de mettre en jeu des observations phonétiques – en principe déconnectées des questions phonologiques – j’ai choisi l’angle des contraintes *phonotactiques* et de la hiérarchie de *sonorité* dans les groupes de consonnes. D’un côté, la langue **dorig**, parlée dans l’île de Gaua, forme ses groupes consonantiques indépendamment de la sonorité de ses composants : c’est ainsi que des mots comme /rk̠p̠^wa/ ‘femme’ ou /lk̠ɔn/ ‘Gaua’ sont bien formés en dorig, en dépit de l’“inversion de sonorité” qu’ils présentent. Par contraste, j’ai pu montrer que le **hiw** traite la hiérarchie de sonorité comme un facteur opérationnel dans ses règles de bonne formation phonotactique. Les exceptions existent (comme /wte/ ‘petit’, /ɣtiɣ/ ‘taille’) mais je montre qu’elles peuvent s’expliquer à l’intérieur du système hiw. À l’inverse, on peut montrer que la sonorité est un critère opérationnel en hiw, notamment à travers certains processus historiques de métathèse (ex. *li^wma-ɠgu ‘mes mains’ > *lm̠ɔ-k > **jm̠ɔ-k → **mj̠ɔ-k**) ou d’épenthèse (ex. *m^wiⁿdolo ‘court’ > **ɲ^wtoɟ → ɲ^wətoɟ), tous deux ayant pour fonction d’éviter les inversions de sonorité.

Or, si la sonorité est une contrainte opérationnelle en hiw (au contraire du dorig), on peut l'utiliser comme pierre de touche pour résoudre la question de la vélaire latérale / $\widehat{g}l$ /. Dès lors, des formes comme / $\widehat{mg}l$ ejɪŋə/ 'forcené' ou / $\widehat{ŋg}l$ e/ 'promontoire' sont la preuve que / $\widehat{g}l$ / est plus sonore que les phonèmes nasals : son statut phonologique n'est donc pas celui d'une occlusive */ g^l /, mais d'une *approximante latérale préplodée*, que je note / $^g l$ /. L'étude se termine par un survol typologique, montrant que ce phonème est rarissime dans le monde, et unique dans la famille austronésienne.

L'article s'achève sur une réflexion théorique sur la question de l'ambiguïté des phonèmes complexes, et la nécessité d'ancrer nos raisonnements sur les structures internes (émiques) de chaque système, sans tenir pour acquis l'existence de catégories phonologiques universelles. Par exemple, la notion de *sonorité*, parfois présentée comme universelle, n'est pas pertinente dans toutes les langues (ex. le dorig), et même quand elle l'est comme en hiw, c'est selon des modalités qui peuvent être spécifiques à une langue donnée.

3.3 Grammaire et catégorisation

La description des systèmes grammaticaux fait apparaître les manières diverses par lesquelles les langues organisent leurs lexèmes en *classes lexicales*, ou *catégories syntaxiques*. J'avais d'abord compris l'importance de ces questions, lors de mes études de linguistique africaine, en lisant l'ouvrage de Denis Creissels *Description des langues négro-africaines et théorie syntaxique* (Creissels 1991). Cette même problématique est au centre des *Parties du Discours* d'Alain Lemaréchal (1989), monographie dont la clarté et la solidité des raisonnements m'avaient d'ailleurs convaincu de choisir son auteur comme directeur de thèse.

3.3.1 Les parties du discours

L'approche structurale incite à décrire chaque langue "dans ses propres termes" (Dryer 2006, Evans & Osada 2005, Haspelmath 2007a), et le domaine des catégories syntaxiques en est un bon exemple. Au lieu de présupposer, par exemple que toute langue possède des noms et des verbes, et qu'ils remplissent partout les mêmes rôles (par ex. les verbes seraient prédicats, les noms arguments), il importe d'observer empiriquement la manière dont les lexèmes se comportent dans chaque langue.

Ces questions de catégorisation lexicale occupent une centaine de pages dans ma thèse, description grammaticale de la langue mwotlap (2001a:153-254)¹. Reposant sur une observation distributionnelle, l'inventaire des parties du discours permet de reconnaître des noms, des adjectifs, des verbes, des adverbes, etc. Chacune de ces classes lexicales est définie par un ensemble de critères (Lemaréchal 1989, Sasse 1993) : critères morphologiques d'une part, mais également critères syntaxiques, incluant l'inventaire des fonctions syntaxiques qui sont ouvertes à cette classe.

¹ Voir les pages de la thèse (pp.164-176) reproduites dans le volume joint ►(#28). L'extrait que je propose de ma grammaire du hiw (en préparation) aborde également ces questions ►(#33).

En mwotlap, il est possible de définir des tests permettant d’assigner à une racine une classe lexicale propre (nom, verbe, adjectif, etc.). Les lexèmes hybrides, appartenant à deux classes à la fois, existent (par ex. *s̄mal* ‘[N] pluie, [V] pleuvoir’, cf. anglais *rain*) mais ils sont l’exception plutôt que la norme : d’une manière générale, une racine est précatégorisée dans une classe et une seule. Cependant, dire que les parties du discours sont relativement étanches entre elles n’empêche pas qu’elles puissent partager des propriétés syntaxiques. Par exemple, noms et adjectifs, sans pour autant se confondre, peuvent tous deux fonctionner comme épithètes, ou comme modifieur de prédicat. L’exemple le plus frappant de ces chevauchements inter-catégoriels concerne leur prédicativité : verbes, adjectifs et noms se comportent de la même façon pour constituer une tête prédicative, y compris à l’aide de marques de Temps-Aspect-Mode¹ [§3.3.2].

L’analyse syntaxique conduit à poser certaines catégories lexicales originales. Ainsi, les POSTVERBES (que je nommais alors “adjoints du prédicat”, d’après Crowley 1982), qu’il faut distinguer des adverbes et autres compléments, ont pour fonction de modifier la tête à l’intérieur du syntagme prédicatif, à la manière d’une épithète dans les syntagmes nominaux. Le placement des marques TAM permet d’identifier les limites du syntagme prédicatif, et fournit un test efficace pour identifier la nature des lexèmes adverbiaux. Dans l’exemple (8), le syntagme prédicatif – en l’occurrence un syntagme à tête verbale – est encadré par la marque TAM discontinue <te-... v̄eh> ‘Potentiel’. Le contraste est net entre les adverbes spatio-temporels comme *qiyig* ‘aujourd’hui’ ou *añqōñ* ‘la nuit’, qui se trouvent toujours extérieurs au syntagme prédicatif, et les postverbes *galsi* ‘bien, correctement’, *lok* ‘encore’ etc., qui ne peuvent apparaître qu’à l’intérieur de ce même syntagme.

- (8) Dō <te- gengen **GALSI** **LOK** **SE** v̄eh> **qiyig** **añqōñ**.
 1inc.du **POT**₁- manger bien encore à.nouveau **POT**₂ auj.d’hui:futur la.nuit
 ‘Nous pourrions manger MIEUX À NOUVEAU ce soir.’

D’un point de vue distributionnel, il n’existe aucune raison de rassembler ces deux catégories syntaxiques en une seule, sous prétexte qu’elles seraient équivalentes des “adverbes” des langues européennes : de toute évidence, on a ici affaire à deux classes lexicales étanches.

3.3.2 Les fonctions des noms

Un autre cas où la distribution en classes lexicales ne va pas de soi, est celle qui concerne la vaste classe syntaxique des NOMS COMMUNS du mwotlap.

Celle-ci se divise en deux sous-catégories : (a) les *humains*, qui peuvent tels quels, sans article, constituer un syntagme nominal bien formé (ex. *tēytēybē* ‘guérisseur’) :

- (9) **MTP** <tēytēybē>_{SN} gēn.
 guérisseur DEIC:PREDIC
 ‘Le guérisseur est là.’

¹ Voir l’exemple (11’) pour les noms, et (17’) pour les adjectifs.

(b) les *non-humains*, qui ne peuvent prendre la tête d'un syntagme nominal que préfixés d'un article *nA*¹ (ex. *men* 'oiseau') :

(10)	⟨ ne -men⟩ _{SN}	<i>gēn.</i>	*⟨ <i>men</i> ⟩ _{SN}	<i>gēn.</i>
	ART-oiseau	DEIC:PREDIC	oiseau	DEIC:PREDIC
	'L'oiseau est là.'			

Chaque membre de cette seconde catégorie de noms se présente dans le discours sous deux formes : la forme préfixée (ex. *ne-men*), employée lorsque le nom est tête de syntagme nominal ; et la forme nue (ex. *men*), essentiellement attestée lorsque le nom est syntaxiquement dépendant, par exemple objet incorporé dépendant d'un verbe transitif (ex. *tēqtēq men* 'chasser les oiseaux'), ou qualifieur/possesseur dépendant d'un autre nom (ex. *n-ili men* 'plume d'oiseau').

En reprenant les termes de Lemaréchal (1989), j'ai proposé de distinguer les 'noms' proprement dits (sémantiquement non-humains, ex. *men*) et les 'substantifs', toujours humains (ex. *tēytēybē*). On pouvait alors reprendre l'analyse de Lemaréchal (1989) inspirée de Tesnière (1959), et analyser l'article *nA*- comme un opérateur de *translation substantivante* : sa fonction consiste à traduire un lexème de la catégorie des NOMS à la catégorie des SUBSTANTIFS.²

Ce contraste entre deux classes de noms est intéressant à plus d'un titre. D'abord, il illustre ici la méthode structurale, en vertu de laquelle l'assignation à une classe lexicale doit obéir à des critères distributionnels ; le contraste entre (9) et (10) oblige à définir deux classes différentes de noms en mwotlap. Ensuite, il souligne le rôle crucial joué par le trait sémantique [$\pm$ humain] dans les grammaires de ces langues océaniques [§3.1]. Enfin, il soulève la question de la fonction fondamentale des noms : car si les noms communs de type non-humain (ex. *men*), les plus courants dans la langue, n'ont pas comme fonction fondamentale de référer, quel est donc leur rôle dans le système ?

Dans cette langue, comme dans d'autres langues austronésiennes (Lemaréchal 1989, 1991), la fonction fondamentale des noms est de constituer un *prédicat qualitatif* : sans son article, un tel nom aura le rôle de modifieur (de nom ou de verbe, cf. plus haut) ou alors de prédicat.

Précisément, plusieurs de mes travaux ont apporté un éclairage sur le comportement des noms du mwotlap lorsqu'ils sont tête de prédicat. Cette question occupe un chapitre de ma monographie sur *La sémantique du prédicat en mwotlap* [AF 2003a, cf. §3.4], et a fait l'objet d'un article :

François, Alexandre. 2004 d. Diversité des prédicats non verbaux dans quelques langues océaniques. In J. François & I. Behr (eds), *Les constituants prédictifs et la diversité des langues: Actes de la Journée de la Société de Linguistique de Paris*. Mémoires de la Société de Linguistique de Paris. Louvain: Peeters. Pp.179-197. ▷(#16)

¹ La majuscule dans *nA*- signale une voyelle /a/ morphologiquement sous-jacente, mais qui s'assimile entièrement à la voyelle suivante selon des règles mentionnées en §3.2.1.

² Dans ma grammaire du hiw en préparation, rédigée en anglais, je parle respectivement de *weak nouns* et de *strong nouns* – cf. ▷(#33).

Cette étude montre que les noms, en mwotlap, peuvent non seulement constituer des prédicats sans copule (ex. 11), mais qu’ils sont également compatibles avec les mêmes marques T.A.M. (temps-aspect-mode) que les verbes ou les adjectifs – comme dans l’exemple (11’), où le nom *tita* porte la marque aspectuelle d’aoriste :

- (11) MTP Kē ⟨tita mino⟩. (11’) Kē ⟨**ni**-tita mino⟩ ēgēn.
 3sg ⟨mère ma⟩ 3sg ⟨AOR-mère ma⟩ maintenant
 ‘C’est ma mère.’ ‘Du coup, elle est devenue ma mère.’

Le mwotlap autorise tous ses noms à se combiner aux marques TAM, comme les verbes. Faut-il parler ici de conversion (ou dérivation zéro), selon laquelle le nom aurait changé de catégorie syntaxique et serait devenu un verbe ? Je réponds à cette question par la négative : dans un énoncé comme (11’), *tita* demeure bel et bien un nom, comme le montre d’ailleurs le maintien de sa syntaxe nominale. Plutôt que d’y voir un verbe sous prétexte que les marques TAM seraient toujours (universellement ?) réservées au verbe, je propose de renverser l’argument : le mwotlap prouve que le marquage aspecto-temporel n’est pas nécessairement une propriété des verbes, et peut caractériser également d’autres classes lexicales, comme les noms (mais aussi les adjectifs et les numéraux).

En somme, les langues diffèrent dans l’inventaire de leurs catégories syntaxiques, mais également dans la manière dont elles distribuent les propriétés grammaticales à ces catégories. Alors que dans la plupart des langues du monde, la faculté d’être marqué en TAM (ou *tamophoricité*, pour reprendre le terme de Tournadre 2004:8) est réservée aux verbes et éventuellement aux adjectifs, certaines langues océaniques – mais aussi certaines langues polynésiennes ou australiennes (Nordlinger & Sadler 2004) – en font également une propriété des noms.

Enfin, je montre également que la *prédicativité* et la *tamophoricité* sont deux propriétés indépendantes. Ainsi, les LOCATIFS du mwotlap sont prédicatifs, mais sont incompatibles avec les marques TAM :

- (12) MTP Tita mino ⟨Ostrelya⟩.
 mère ma Australie
 ‘Ma mère est en Australie.’
 (12’) *Tita mino ⟨mal Ostelelya⟩.
 mère ma ACP Australie
 [‘Ma mère est déjà en Australie.’]

Le Tableau 6 résume les faits en présentant les propriétés de huit catégories lexicales de la langue mwotlap : nom, adjectif, verbe, attribut, prédicat existentiel, classificateur possessif, locatif et numéral. Toutes ces classes partagent le trait de *prédicativité*, au sens où tout lexème appartenant à ces classes peut constituer la tête d’un syntagme prédicatif, sans copule de type *être*. On retrouve en mwotlap la forte propension des langues austronésiennes à “l’omniprédicativité” (cf. Lemaréchal 1989, Launey 1994). Le bas du tableau montre que les classes diffèrent cependant dans la compatibilité avec des marques TAM : un LOCATIF ne peut constituer un prédicat que *sans* marquage TAM, comme en (12) ; inversement, un VERBE ne peut le faire qu’*avec* marquage TAM ;

les NOMS, quant à eux, peuvent prédiquer aussi bien sans TAM [cf. (11)] qu’avec [cf. (11’)].

Tableau 6 – Prédicativité et compatibilité TAM en mwotlap

	Nom	Adjectif	Verbe	Attribut	Exist ^l	Poss ^{sif}	Locatif	Numéral
PRÉDICATIF	+	+	+	+	+	+	+	+
Prédicat avec TAM	+	+	+	+	–	–	–	+
Prédicat sans TAM	+	–	–	+	+	+	+	+

3.4 Sémantique du Temps-Aspect-Mode

3.4.1 La description des systèmes TAM

La description du système de marquage Temps-Aspect-Mode est souvent le parent pauvre des descriptions grammaticales des langues non-européennes. Ce constat vaut surtout pour la littérature descriptive en langue anglaise : la tradition francophone a donné lieu à des travaux riches autour de ces questions, par exemple sur les langues berbères (Galand 2003, Mettouchi 1998, 2009, Taine-Cheikh 2010), sur le wolof (Robert 1991, 1995, 1996, 2010) – ou encore, plus près du Vanuatu, les travaux des océanistes francophones sur les langues de Nouvelle-Calédonie ou de Polynésie (Moyse-Faurie 1998, 2001; Bril 1999, 2002, 2003; Vernaudeau 1999).

Inspiré par certains de ces travaux, j’ai consacré au système TAM du mwotlap le troisième volume de ma thèse – lui-même le prémice de ma monographie sur cette même langue :

François, Alexandre. 2003 a. **La Sémantique du Prédicat en Mwotlap (Vanuatu)**. Collection Linguistique de la Société de Linguistique de Paris, 84. Paris, Louvain: Peeters. ▷(#30)

J’ai mentionné plus haut cette monographie dans ma discussion sur les catégories syntaxiques [§3.3.2] : il s’agissait de déterminer quelles sont les classes lexicales compatibles avec les marques TAM (Tableau 6). Une fois ce point établi, restait à décrire les mécanismes sémantiques de ce système. Je n’en donnerai ici qu’un bref aperçu.

Bien que le phénomène soit théoriquement connu ailleurs dans le monde, un point fort du mwotlap est l’absence de référence temporelle absolue, i.e. d’indication morphologique du *temps grammatical*. Certes, les vingt-cinq marqueurs TAM de la langue ont bien pour fonction – entre autres – de coder des relations de succession / inclusion / simultanéité... entre des instants ; mais ces relations se font toujours par rapport à une situation de référence qui n’est pas nécessairement l’instant d’énonciation. Les énoncés suivants montrent ainsi que l’Accompli¹ *mal* n’encode qu’un sens aspectuel d’évène-

¹ J’adopte ici l’usage préconisé par Comrie (1976:10) et Haspelmath (2007a), d’employer la majuscule chaque fois que je désigne une catégorie propre à une langue en particulier. Ainsi “l’Accompli” .../...

ment accompli ('avoir fini P à l'instant *t* considéré'), mais reste sous-spécifié au regard de l'ancrage temporel absolu (valeurs possibles de présent, de passé, de futur ou de contrefactuel, sans aucun marquage morphologique) :

- (13) Bōbō nōnōm n-ēh leptō ? - Óòǒ, kē **mal** mat !
 aïeul ton STAT-vivre encore non 3SG ACP mort
 'Ton grand-père est toujours vivant ? – Non, il *est* mort.' ⇔ PRÉSENT
- (14) Nē-dēmdēm nonon so bōbō nonon **mal** mat.
 ART-pensée sa que aïeul son ACP mort
 'Il croyait que son grand-père *était* mort.' ⇔ PASSÉ
- (15) Nēk so van lok me l-ête itan en, bōbō **mal** mat.
 2SG PROSP aller encore VTF LOC-année autre TOP aïeul ACP mort
 'Quand tu reviendras l'année prochaine, grand-père *sera* déjà mort.' ⇔ FUTUR
- (16) Dokta so tateh tō en, togtō bōbō **mal** mat !
 médecin si non.exister CFT TOP alors:CFT aïeul ACP mort
 'Si le médecin n'avait pas été là, grand-père *serait* (sûrement) mort !' ⇔ IRRÉEL

En d'autres termes, la référence temporelle est relative et non absolue, et l'on peut dire que le mwotlap grammaticalise non pas le temps, mais l'aspect. (La seule exception à ce principe est le *Futur hodiernal*, impliquant un ancrage le même jour que l'énonciation.)

Par ailleurs, le mwotlap inclut dans son paradigme TAM la polarité de la proposition, positive ou négative. En d'autres termes, ce paradigme encode explicitement les catégories Aspect–Modalité–Polarité. Un point remarquable est l'absence de correspondance terme à terme entre TAM positifs et négatifs : ainsi, les marques affirmatives du Statif (*nE-*), du Parfait (*mE-*), Prétérit (*mE-... tō*) et de l'Aoriste (*ni-*) seront toutes niées par le Négatif Realis (*et-... te*). Au total, la langue compte dix-huit marqueurs TAM affirmatifs et sept négatifs, soit un total de vingt-cinq catégories ; elles sont reproduites dans le *Tableau 7*.

Certaines de ces catégories impliquent la référentialité du procès, et sont donc des marques *realis*, correspondant grossièrement au passé ou au présent (ex. Accompli, Statif, Parfait, Prétérit...). D'autres tiroirs impliquent au contraire un procès non référentiel, encore virtuel, et peuvent être décrits comme *irrealis* (ex. Prospectif, Futur, Potentiel, Évitatif, Prohibitif...) – c'est le domaine par excellence des relations modales, mettant en jeu des visées subjectives. Enfin, un cas intermédiaire est représenté par deux catégories, l'Aoriste et le Focus temporel, car ils peuvent porter tantôt sur des actions référentielles et donc *realis*, tantôt sur des situations encore virtuelles ; l'auditeur devra se fonder sur d'autres éléments pour interpréter correctement l'énoncé.

n'est que l'étiquette que j'ai donnée à une catégorie propre au mwotlap. Si je devais évoquer la catégorie typologique (ou plutôt le *concept comparatif*, dans les termes d'Haspelmath 2010) de l'*accompli*, j'emploierais la minuscule.

Tableau 7 – Catégories TAM du mwotlap : 18 affirmatives, 7 négatives

AFFIRMATIF		NÉGATIF	
Accompli	<i>mal ...</i>	'pas encore'	<i>et-... qete</i>
Accompli distant	<i>mal ... tō</i>		
Rémansif	<i>... laptō</i>	'ne plus'	<i>et-... si te</i>
Statif	<i>nE-...</i>	Négatif realis	<i>et-... te</i>
Parfait	<i>mE-...</i>		
Prétérit	<i>mE-... tō</i>		
Aoriste	<i>(ni-)...</i>		
Injonction forte	<i>(ni-)... tō</i>	Prohibitif	<i>tog ...</i>
Prospectif	<i>so (ni-)...</i>	Négatif futur	<i>tit-... te</i>
Futur / Futur hodiernal	<i>tE-... / tE-... qiyig</i>		
Potentiel	<i>tE-... vēh</i>	Négatif potentiel	<i>tit-... vēs-te</i>
Contrefactuel	<i>tE-... tō</i>		
Évitatif	<i>tiple ...</i>	Évitatif négatif	<i>tiple tit-... te</i>
Présentatif statique	<i>... tō</i>	Ø	
Présentatif kinétique	<i>... vatag</i>		
Focus temporel	<i>qoyo ...</i>		
Prioritif	<i>(ni-)... bah en</i>		

3.4.2 Le Gabarit standard de procès, et l'Aktionsart

Parmi les nombreuses originalités typologiques de ce système TAM, figure un mécanisme remarquable, relatif au codage de l'Aktionsart ou type de procès. Ce mécanisme, que j'ai appelé *Gabarit standard de procès*, a notamment fait l'objet d'une de mes premières publications, aux sources de ma monographie de 2003 :

François, Alexandre. 2001 b. Gabarit de procès et opérations aspectuelles en motlav (Océanie). *Actances* 11 (June 2001):145-175.

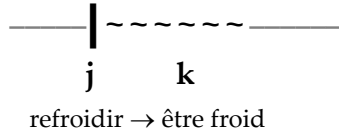
Tout se passe comme si, en mwotlap, tous les radicaux verbaux présentaient uniformément le même schéma de type de procès, consistant en l'articulation d'une première phase télique *j* (ex. *j* = 'refroidir') à une seconde phase atélique *k* (ex. *k* = 'être froid'). C'est ce schéma universel ⟨*j,k*⟩, sorte d'étalon commun à tous les radicaux, que j'appelle le Gabarit (cf. anglais *template*) ; il est représenté dans la Figure 6.

Certaines marques TAM, comme le Statif *nE-*, opèrent exclusivement sur l'état *k*. Ainsi, la forme *no-momyiy* prédique une propriété *k* 'froid' sans évoquer aucun changement d'état préalable :

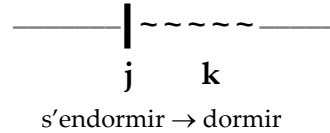
- (17) MTP Nē-bē nan **no**-momyiy.
 ART-eau ANAPH STAT-froid
 'L'eau est froide.'

Figure 6 – Le Gabarit de procès en mwotlap : tout lexème référant à un *état* peut également référer à l'événement de changement d'état, selon un type de procès unique (j,k)

ex. MOMYIY 'refroidir/être froid'



ex. M(I)TIY 's'endormir/dormir'



À l'inverse, l'Aoriste *ni-* fait partie des marques TAM qui opèrent exclusivement sur l'événement *j* de changement d'état. L'énoncé (17') relate un refroidissement qui a eu lieu à un moment *t* considéré, mais ne dit rien de l'état actuel de l'eau. Autrement dit, l'Aoriste opère sur l'événement *j* mais pas sur l'état *k* :

- (17') Nē-bē nan **ni**-momyiy.
 ART-eau ANAPH AOR-froid
 '(Soudain) l'eau refroidit.'

Quant à la marque *mE-* que j'analyse comme un Parfait, elle porte autant sur *j* que sur *k*, puisqu'elle implique un changement d'état *j* et la validation de son état résultant *k* :

- (17'') Nē-bē nan **mo**-momyiy.
 ART-eau ANAPH PRF-froid
 'L'eau a refroidi (et maintenant elle est froide).'

Alors que les deux phases sont distinguées lexicalement dans une langue comme le français, elles sont systématiquement exprimées, en mwotlap, par le même lexème (ex. MOMYIY 'refroidir → être froid'). L'auditeur identifiera la bonne phase du procès à partir des opérateurs TAM, ou bien d'autres procédés morphologiques comme la reduplication du radical verbal [§3.4.3]. Le mwotlap encode ici dans les morphèmes (les marques TAM) des informations que d'autres langues encodent dans les lexèmes (le radical verbal).

En somme, les lexèmes du mwotlap n'opposent pas des *états* à des *événements*, puisqu'ils peuvent généralement référer aux deux, sans besoin de dérivation. Ceci est vrai non seulement des verbes, mais aussi des adjectifs (ex. *momyiy*) et même des noms [§3.3.2]. Ainsi, le nom *wōh* peut encoder, sans aucune dérivation, aussi bien un prédicat statif (*k* 'être une noix de coco buvable') que l'événement qui conduit à cet état (*j* 'devenir une noix de coco buvable') :

- (18) Mey nen **nō**-wōh.
 celui là ART-coco.vert
 'Celui-là, c'est *un coco vert*.'

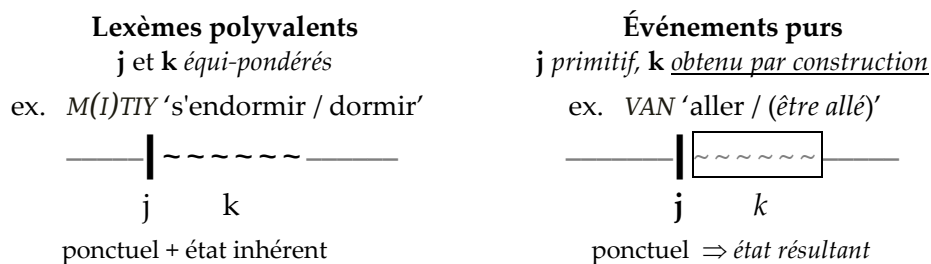
- (18') Mey nen **mal** wōh !
 celui là ACP coco.vert
 'Celui-là, c'est déjà *devenu un coco vert*.' (=il est déjà devenu buvable)

Dans la mesure où tous les noms du lexique sont ainsi, en principe, compatibles avec l'aspectualisation, il n'y a aucune raison de dire que *wōh* a été converti en verbe en (18') : on a bien affaire ici à un nom, lequel est – comme tous les noms – intrinsèquement capable de référer à un changement d'état. En d'autres termes, le Gabarit de procès en mwotlap s'applique à toutes les classes lexicales compatibles avec les marques TAM (*Tableau 6* p.49) : nom, adjectif, verbe, attribut, numéral.

Pourtant, j'ai pu montrer qu'il existe bien une dichotomie dans le lexique, fondée sur l'Aktionsart ou aspectualité lexicale. Cependant, au lieu d'opposer états à événements, elle sépare plutôt des *états-événements* et des *événements purs* :

- a) un vaste ensemble de LEXÈMES POLYVALENTS (noms, adjectifs, la moitié des verbes) pouvant fonctionner aussi bien comme ponctuels (**j**) que comme états (**k**), conformément au Gabarit de procès. Les phases **j** et **k** y sont équi-pondérées, et l'état **k** est inhérent au sémantisme lexical.
→ ex. *m(i)tiy* : 's'endormir' ~ 'dormir'
- b) un ensemble plus restreint de verbes (l'autre moitié) qui sont FONDAMENTALEMENT TÉLIQUES : leur phase ponctuelle **j** y est primitive, et se trouve obligatoirement impliquée dans les opérations aspectuelles ; ces verbes ne fonctionnent jamais comme de purs états (impliquant la seule phase **k**). Cependant, dans certains cas [Parfait / Prétérit], la phase statique **k** peut être obtenue par construction à partir de **j** : il s'agit d'un état résultant, extrinsèque au procès lui-même.
→ ex. *van* : 'aller' ⇒ (*être allé*)

Figure 7 – Les deux types d'Aktionsart en mwotlap :
lexèmes polyvalents *états-événements* (à phase statique intrinsèque) vs.
lexèmes référant à des *événements purs* (état résultant obtenu par construction)



La *Figure 7* représente cette dichotomie entre les deux types de procès (ou Aktionsarten) encodés par le lexique de la langue.

Cette démonstration a fait l'objet du chapitre conclusif de ma monographie, ainsi que d'une intervention devant la Société de Linguistique de Paris :

François, Alexandre : "Aspect et télicité en mwotlap (Vanuatu) : Un type-de-procès unique pour tous les lexèmes de la langue ?". Intervention devant la *Société de Linguistique de Paris*, Mai 2003.

Le système TAM du mwotlap présente bien d'autres points intéressants, qui ne peuvent être développés ici. Je reviendrai plus loin sur la catégorie de l'Aoriste [§5.1.2], sur laquelle j'ai également publié dans une perspective comparative.

3.4.3 Sémantique de la reduplication

En partie lié à la problématique de l'aspect et de l'Aktionsart, le domaine de la *reduplication* en mwotlap m'a également intéressé :

François, Alexandre. 2004 b. La reduplication en mwotlap : les paradoxes du fractionnement. In *Les langues austronésiennes*, edited by E. Zeitoun. *Faits de langues* n°23-24. Paris: Ophrys. Pp.177-194. ▷(#5)

Le mwotlap fait un usage intense de la reduplication. Les radicaux lexicaux, en particulier les adjectifs et les verbes, sont normalement susceptibles de se rencontrer sous deux formes : une forme simple, une forme redupliquée. Sans entrer dans le détail des règles morphologiques, on peut illustrer le phénomène à l'aide du *Tableau 8*.

Tableau 8 – Quelques exemples de reduplication en mwotlap

<i>sens</i>	<i>f. simple</i>	<i>f. dupliq.</i>	<i>sens</i>	<i>f. simple</i>	<i>f. dupliquée</i>
'assis'	<i>hag</i>	<i>haghag</i>	'éclore'	<i>dey</i>	<i>dedey</i>
'voir'	<i>et</i>	<i>etet</i>	'grand'	<i>liwo</i>	<i>lililwo</i>
'changer'	<i>leh</i>	<i>leleh</i>	'être bon'	<i>itōk</i>	<i>itōktōk</i>
'cogner'	<i>tit</i>	<i>titit</i>	'végétal'	<i>tēnge</i>	<i>tētēnge</i>
'dire'	<i>vap</i>	<i>vavap</i>	'tirer'	<i>w(e)seg</i>	<i>w(e)sewseg</i>
'faire'	<i>galeg</i>	<i>gagaleg</i>	'triste'	<i>m(a)gaysēn</i>	<i>m(a)gamgaysēn</i>
'savoir'	<i>ēglal</i>	<i>ēgēglal</i>	'répondre'	<i>v(u)lu</i>	<i>v(u)luplu</i>
'lisse'	<i>ēyat</i>	<i>ēēyat</i>	'aider'	<i>b(i)yiñ</i>	<i>b(i)yimyiñ</i>

La reduplication présente des valeurs sémantiques diverses, iconiques mais non triviales. Appliquée à un nom, la reduplication aura une valeur pluralisante, ou évoquera un référent fragmenté (ex. *ne-vet* 'pierre, rocher' → *ne-vetvet* 'pierraille, gravier'). Avec un adjectif ou un verbe, le procédé impliquera des valeurs d'intensité, de pluralité, de distributivité, d'itérativité, d'imperfectivité, de désactualisation. On comparera ainsi la forme simple *van* 'aller' (19a/b), à valeur singulative, avec sa forme redupliquée *vanvan* : cette dernière revêt un sens itératif/distributif en (19c), et une valeur potentielle/intensionnelle en (19d).

(19a) MTP Inti-k mal VAN Numea.

 fils-1sg ACP aller N.

 'Mon fils est parti à Nouméa.'

(19b) Inti-k mal VAN.

 fils-1sg ACP aller

 'Mon fils est parti / y est allé.'

(19c) Inti-k mal **VANVAN** le-pnō tegtegha.

 fils-1sg ACP aller:DUP LOC-pays différent:DUP

 'Mon fils s'est déjà rendu dans de nombreux pays.'

(19d) Inti-k mal **VANVAN**.

 fils-1sg ACP aller:DUP

 'Ça y est, mon fils marche.' (ex. bébé qui fait ses premiers pas)

C'est ainsi que les marques TAM sémantiquement habituelles ou virtuelles (y compris l'infinitif ou le nom d'agent) feront appel à la reduplication.

Tel quel, un radical simple (nom, verbe, etc.) aura tendance à représenter son référent comme une entité discrète et autonome, pleine, pourvue d'une extension finie dans l'espace ou le temps (*ne-vet* 'un rocher', *van* 'se rendre qq part [une fois]'). L'opération de reduplication présente la même notion sur le mode de la *fragmentation* : l'unité close de départ se disperse en une multiplicité d'occurrences, et perd du même coup la précision de ses limites externes. C'est ainsi qu'en fonction des contextes et des combinaisons, un seul et même procédé morphologique – non seulement en mwotlap, mais dans les nombreuses langues, austronésiennes ou non, qui connaissent ce phénomène – pourra endosser des fonctions de pluriel, d'intensif, de collectif, de diminutif, de distributif, de réciproque, d'habituel, de fréquentatif, de progressif, de conatif, d'atélique, d'intransitif, d'infinitif... Tout se passe comme si la reduplication, en fractionnant indéfiniment la notion, permettait d'atteindre à sa pureté même, alliage de virtualité et de renouvellement perpétuel.

Dans le domaine de l'Aktionsart, enfin, le mwotlap exploite la reduplication pour recatégoriser un verbe télique (*événement pur*) en un verbe polyvalent (*état-événement*) ; autrement dit, pour faire en sorte qu'un verbe ancré en *j* devienne un verbe ancré en *j+k* [cf. §0]. C'est le cas, par exemple, pour le verbe *van* : sa forme simple est un *événement pur*, comme le montre le test d'incompatibilité avec le Statif :

- (19e) *Kemem na-VAN l-ēm̃yoñ.
 1ex:pl STAT-aller:DUP LOC-église
 *Nous allons à l'église.

À l'inverse, la forme redupliquée *vanvan* appartient à la vaste classe des lexèmes polyvalents, capable de référer aussi bien à un événement *j* qu'à une propriété stative *k*. Ainsi, *vanvan* peut se combiner avec le Statif, pour référer à la propriété habituelle de *aller (régulièrement)* – d'où *savoir marcher* comme dans (19d), ou bien *fréquenter* :

- (19f) Kemem na-VANVAN l-ēm̃yoñ.
 1ex:pl STAT-aller:DUP LOC-église
 'Nous fréquentons l'église.'

Cette capacité de transformer des lexèmes d'une catégorie d'Aktionsart (*événements purs*) en une autre (*états-événements*) constitue ce que j'ai appelé "l'alchimie sémantique de la reduplication" [AF 2003a:358-361].

3.4.4 Marquage aspecto-modal et subordination

Je n'ai pas réservé la description des systèmes TAM à la seule langue mwotlap. Outre mon étude transversale des marques d'aoriste [§5.1.2], je me suis également penché sur les deux langues parlées aux îles Torres, le **lo-toga** et le **hiw**.¹

¹ Cette étude de cas formait ma contribution à l'ouvrage collectif qu'Isabelle Bril avait coordonné, *Clause hierarchy and Clause linking* (Bril 2010) – à propos des divers procédés employés par les langues du monde pour encoder les liens de dépendance entre propositions. Cet ouvrage venait
 .../...

François, Alexandre. 2010 c. Pragmatic demotion and clause dependency: On two atypical subordinating strategies in Lo-Toga and Hiw (Torres, Vanuatu). In *Clause hierarchy and Clause linking: the Syntax and pragmatics interface*, edited by I. Bril. Studies in Language Companion Series 121. Amsterdam, New York: Benjamins. Pp.499–548. ▷(#22)

Le système TAM de ces deux langues comporte, certes, certains points communs avec celui du mwotlap, mais également des différences notables. Cet article porte en particulier sur deux catégories TAM partagées par les deux langues des Torres, mais inconnues ailleurs dans les Banks : le Subjonctif, et le Parfait d'arrière-plan. Bien que sémantiquement très différentes l'un de l'autre, ces deux catégories présentent cependant un point commun : elles ont toutes les deux l'effet de rendre leur proposition syntaxiquement dépendante. Cette dépendance s'accompagne parfois de marques segmentales de subordination (marque de relative, conjonction etc.), mais pas nécessairement : dans ce dernier cas, on peut considérer que c'est la marque TAM elle-même qui crée la dépendance¹. Le but de cette (double) étude de cas était d'analyser ce qui, dans le sémantisme même de ces marqueurs, peut être à l'origine de cet effet de subordination.

J'analyse le Subjonctif de ces deux langues (hiw *on*, lo-toga *vě*) comme une marque indiquant l'absence de toute *force illocutoire*. Au contraire des autres marques TAM du système, une proposition au Subjonctif n'est pas assertée, et n'implique aucune prise en charge énonciative : elle ne fait qu'évoquer une situation virtuelle, sans la raccrocher à aucun monde particulier. En conséquence, une assertion au Subjonctif ne peut constituer un énoncé bien formé :

- (20) LTG *Nie **vě̃n** metur l-ě̃nwe mino.
 2sg SBJF dormir LOC-maison ma
 *(Qu'il dorme chez moi.

Un Subjonctif ne peut former un énoncé acceptable que de deux manières. L'une consiste à le combiner avec une intonation exclamative, le tout encodant une valeur d'*hortatif*. Dans ce cas, selon mon analyse, la force illocutoire ne provient pas du subjonctif lui-même (qui en est dépourvu), mais de sa combinaison avec un schéma prosodique spécifique :

- (20') Nie **vě̃n** metur l-ě̃nwe mino!
 2sg SBJF dormir LOC-maison ma
 'Qu'il dorme chez moi !'

L'autre option consiste à raccrocher cette proposition à une autre proposition, elle-même pourvue d'une force illocutoire (assertion, injonction etc.). C'est ainsi que le Subjonctif peut, à lui seul, encoder la valeur d'hypothèse dans un énoncé conditionnel :

lui-même couronner un programme de la Fédération *Typologie et Universaux du Langage*.

¹ Ce raisonnement m'était en partie inspiré des travaux de Stéphane Robert sur l'aoriste en wolof, et ses effets de dépendance (Robert 1991, 1995, 1996).

- (20") *Nie* **vēn** *metur* *l-ēñwe* *mino*, *noke* *të* *ke* *melekelake*.
 2sg SBJF dormir LOC-maison ma 1sg PROSP 1sg ravi
 ‘Qu’il dorme chez moi, (et) j’en serai ravi.’

Le Parfait d’arrière-plan (*Background Perfect*, en abrégé BkPf) prend la forme *ve... ti* en hiw, et *ve... si* en lo-toga. Dans les deux langues, ce marqueur BkPf contraste avec le Parfait normal (hiw *nē...ti*, lo-toga *na...si*) en fonction du degré d’activation discursive du prédicat. Lorsque le prédicat est rhématique, il sera marqué par le Parfait standard :

- (21) LTG *Sesē* **na** *hag* **si** !
 ta.sœur PRF₁ s’asseoir PRF₂
 ‘Ta sœur a accouché !’

À l’inverse, le Parfait d’arrière-plan sera toujours corrélé à une information ancienne, thématisée ou préconstruite :

- (21') *Sesē* **ve** *hag* **si** : *ne* *teñwën* *hitë* *ne* *legëvine*?
 ta.sœur BkPf₁ s’asseoir BkPf₂ ART homme ou ART femme
 ‘Ta sœur a accouché (comme nous le savons) : c’est un garçon, ou une fille ?’

Comme pour le Subjonctif, le Parfait d’Arrière-plan contient dans ses gènes un élément qui rend la proposition subordonnée à une autre. Cette fois-ci, ce qui crée la dépendance n’est pas l’absence de force illocutoire, mais l’absence de rhématicité. La proposition se trouve dé-rhématisée par la marque de BkPf, ce qui explique l’impossibilité de constituer à elle seule un énoncé bien formé. À l’inverse, le BkPf sera la forme attendue dans toutes les propositions thématisées ou subordonnées, comme les relatives restrictives :

- (21") *Resiga* **ve** *hag* **si** *mat* *nwule* *wulë* *l-ēñwe* *më*.
 celle BkPf₁ s’asseoir BkPf₂ ACP rentrer à.nouveau LOC-maison sa
 ‘Celle (qui) a accouché est déjà rentrée chez elle.’

Le relativiseur (*pe*) existe, mais il est ici optionnel : en effet, le seul emploi du Parfait d’arrière-plan suffit à créer la relation de subordination.

J’ai proposé de formuler le point commun entre ces deux cas de figure en parlant de “privation énonciative” (*pragmatic demotion*). Qu’il s’agisse du Subjonctif ou du Parfait d’arrière-plan, dans les deux cas on a affaire à une marque TAM dont l’effet est de dépouiller la proposition d’une propriété sémantique dont elle aurait besoin pour constituer, à elle seule, une phrase énonciativement bien formée :

- le Subjonctif ôte à la proposition sa *force illocutoire*, et l’empêche donc de construire un acte de langage
- le Parfait d’Arrière-plan ôte à la proposition sa *rhématicité*, et l’oblige donc à se raccrocher à un élément au premier plan rhématique.

Ces deux langues des Torres font l’objet de travaux descriptifs en cours [§8.1].

3.5 Prédicats complexes et relations actancielles

3.5.1 Macroverbes et verbes sériels en mwotlap

Le mwotlap fait un usage intense de la sérialisation verbale, au point que j’y ai consacré deux articles. L’un d’eux était un chapitre dans un ouvrage collectif (Bril & Ozanne-Rivierre 2004) sur les prédicats complexes dans les langues océaniques :

François, Alexandre. 2004 a. Chains of freedom: Constraints and creativity in the macro-verb strategies of Mwotlap. In *Complex predicates in Oceanic languages: Studies in the dynamics of binding and boundness*, edited by I. Bril & F. Ozanne-Rivierre. Empirical Approaches to Language Typology. Berlin: Mouton de Gruyter. Pp.107-143. ▷(#15)

Les articles théoriques et typologiques concernant la sérialisation verbale (Foley & Olson 1985, Durie 1997, Crowley 2002, Bril 2004, Aikhenvald 2006) ont montré combien le phénomène de sérialisation verbale peut varier d’une langue à l’autre. Tous s’accordent à distinguer deux principaux types, parfois dénommés *core-layer serialization* et *nuclear-layer serialisation* (Foley & Olson 1985) :

1. *core-layer serialisation* : structures sérielles dans lesquelles la récursivité se situe au niveau du syntagme verbal, voire de la proposition dans son ensemble
→ ex. ⟨Le garçon PRIT la pierre⟩ ⟨LANÇA en l’air⟩.
2. *nuclear-layer serialisation* : structures sérielles dans lesquelles la récursivité se situe au niveau du noyau lexical verbal
→ ex. Le garçon ⟨PRIT LANÇA⟩ la pierre en l’air.

Il est en effet indispensable de distinguer entre ces deux types syntaxiques ; l’idée même de les regrouper dans une seule catégorie “série verbale” pose d’ailleurs problème, et ne semble pas toujours justifiée. Crowley (1987) a montré qu’une même langue – en l’occurrence le paama du Vanuatu – peut avoir les deux structures. C’est aussi le cas en araki ; aussi ai-je choisi, dans ma grammaire [2002], de les distinguer terminologiquement, en réservant le terme de *verb serialisation* à la seconde structure dite ‘nuclear’ ; quant à la première, je l’ai décrite sous le nom de *clause chaining* [▷(#29), pp.189-200]. Le hiw, langue des Torres que je suis en train de décrire, présente également les deux constructions.

Le mwotlap ne possède qu’une seule construction sérielle : la structure dite *nuclear*, dans laquelle un seul et même syntagme verbal peut comporter deux, trois voire quatre lexèmes verbaux – tout en n’ayant jamais plus qu’un seul sujet et un seul objet. En voici un exemple, tiré d’une célèbre chanson d’amour mwotlap :

- (22) MTP (nêk) ⟨lak tēy yoyōñ ēwē⟩ no.
(2sg) danser tenir être.calme être.bien 1sg
(‘Tu) danse(s) avec moi bien doucement.’

On peut d’abord montrer qu’il s’agit bien d’un cas prototypique de sérialisation : par exemple, la chaîne verbale ne pourra être associée qu’à un seul sujet, un seul objet, mais aussi une seule marque TAM ou négation, ces dernières se trouvant en facteur commun pour l’ensemble de la série verbale. Pourtant, une construction sérielle

comme (22) soulève de nombreuses questions, auxquelles mes deux articles ont tenté de répondre.

Les quatre verbes de la série sont-ils sur le même plan, ou bien peut-on identifier une tête et des modifieurs ? J’ai montré qu’il était légitime de considérer le premier verbe comme la tête du syntagme, celle qui détermine en premier lieu les relations syntaxiques avec les arguments, et en particulier avec le sujet. Quant aux autres verbes, j’ai montré qu’ils s’insèrent en réalité dans une position syntaxique de modifieur, que j’ai appelée “adjoint du prédicat” (en empruntant le terme “adjoint” à Crowley 1982), ou “postverbe”. Cette position est ouverte non seulement aux VERBES, mais également aux ADJECTIFS et aux NOMS – sans oublier les POSTVERBES dont c’est la seule fonction, comme *galsi* ‘bien’, *lok* ‘encore’ ou *se* ‘à nouveau’ dans l’ex. (8) p.46.

Corollaire de ce point, les séries verbales du mwotlap ne sont pas “symétriques” (pour reprendre le terme d’Aikhenvald 2006) : alors qu’il existe une entière liberté de choix pour le premier verbe V1, en revanche la position postverbale implique des restrictions pour le choix de V2. Seul un sous-ensemble des verbes est compatible avec cette position. En outre, pour le même verbe, on observe des différences entre les positions V1 et V2 : différences morphologiques parfois (ex. *ēglal* ‘savoir’ en V1 devient *vēglal* en V2), et surtout différences sémantiques. Prenons l’exemple du verbe *tēy* ‘tenir’. En position de tête du syntagme verbal (V1), *tēy* aura son sens lexical littéral, celui de ‘tenir à la main’, à l’exclusion d’autres sens :

- (23) No ⟨mē- **tēy**⟩ ni-ḡilas nōnōm.
 1sg PRF- tenir ART-lunettes tes
 ‘J’ai tes lunettes dans la main / **sur les yeux*.’

Pourtant, que ce même verbe se retrouve sérialisé en position de V2, et il perd cette restriction sémantique. Ainsi, (23’) pourrait signifier ‘danser avec tes lunettes *sur les yeux* ~ *dans mon sac*, etc.’ :

- (23’) No ⟨ma- *lak* **tēy**⟩ ni-ḡilas nōnōm.
 1sg PRF- danser tenir ART-lunettes tes
 ‘J’ai dansé avec tes lunettes (dans la main / sur les yeux...).’

On peut proposer que le lexème *tēy*, en position postverbale, a généralisé son sémantisme à une signification comitative ‘avec’ ; c’est d’ailleurs elle qui est illustrée dans (22) ci-dessus. On pourrait également formuler le changement en considérant que le postverbe *tēy* s’est “grammaticalisé” en un morphème d’applicatif, de type *applicatif instrumental-comitativ*, comme il en existe en kinyarwanda par exemple (Lemaréchal 1998:195).

Un exemple comme (23’) concentre précisément les problématiques qui m’ont intéressé à propos de ces constructions. S’agit-il toujours d’une série verbale, ou bien a-t-on déjà basculé dans la grammaticalisation en applicatif ? Doit-on voir dans ces deux *tēy* une seule et même unité lexicale, ou bien deux homophones ? Mon argument est précisément qu’il n’y a pas lieu de choisir, car l’ambiguïté n’est pas seulement une illusion de linguiste : elle existe également pour le locuteur lui-même. Lichtenberk (1991), en discutant précisément l’aspect *graduel* de la grammaticalisation, a ainsi proposé le terme d’*hérérosémie*, “where the different but related meanings of a given

morpheme are associated with distinct grammatical contexts” (Enfield 2006:297). Mes articles décrivent le continuum de grammaticalisation typique des constructions sérielles du mwotlap. À un extrême, on trouve des constructions à haut degré de compositionnalité, où les deux verbes apparaissent sémantiquement autonomes (ex. *yow tig* <bondir être.debout> = ‘sauter à pieds joints’). À l’autre extrême, on trouve des constructions complexes dont le premier mot est un verbe, suivi par un postverbe n’ayant qu’un lien ténu avec un lexème verbal.

Parfois, il est possible d’identifier son origine verbale, mais l’emploi en tant que verbe est archaïque, ou attesté seulement dans les langues voisines, et souvent sémantiquement éloigné. Par exemple, le mwotlap a un postverbe *kal* ‘en montant’ :

- (24) MTP Nok so yem **kal** hag le-mtig en.
 1sg PROSP grimper (monter) en.haut LOC-cocotier ANAPH
 ‘Je vais grimper / monter en haut du cocotier.’

Or on peut montrer que cette forme *kal* est étymologiquement liée à un ancien verbe **kal* ‘avancer’. Ce dernier n’existe plus en tant que verbe, mais est reflété dans la forme redupliquée *kakal* ‘ramper, marcher à quatre pattes’ ou dans l’ancien composé *kalbat* <‘avancer dedans’> ‘entrer’. Se pose alors la question de la meilleure analyse en synchronie : dans un cas comme (23)-(23’), on pouvait parler d’une hétérosémie entre deux valeurs de *tēy* (V1 ‘tenir à la main’, V2 ‘avoir, être avec’), et l’analyse en termes de série verbale était acceptable. Mais dans le cas de (24), a-t-on une série verbale ? Dans la mesure où *kal* n’est plus un verbe en synchronie, mais se trouve uniquement employé en position postverbale, on ne peut plus voir ici une construction sérielle, mais plutôt un syntagme verbal complexe (ou “macro-verbe”) constitué d’un verbe *yem* et de son postverbe *kal*.

Mon étude explorait ce continuum, en l’illustrant de divers exemples : certains étaient des cas synchroniquement ambigus comme *tēy*, compatibles avec une analyse sérielle ; d’autres montraient combien une ancienne sérialisation avait pu se grammaticaliser aujourd’hui en des structures différentes. À chaque fois, on retrouvait la même structure primitive, dans laquelle V1 est la tête du syntagme et garde toujours ses caractéristiques de verbe – alors que V2 montre des tendances à la réanalyse : tel verbe était devenu applicatif, tel autre était grammaticalisé en préposition, en marque TAM... En dehors des tendances centrifuges qu’on vient de voir (V2 tend à s’autonomiser), on observe aussi des tendances intégratives, où le complexe V1-V2 tend à se co-lexicaliser avec un sens non-compositionnel : ainsi, *dēm veteg* <penser-laisser> s’est lexicalisé avec le sens ‘pardonner’ ; *leg leh* <épouser-changer> est le terme pour ‘se remarier’, etc.

Cette analyse a donné lieu à une réflexion plus théorique sur la productivité des structures grammaticales, et leur propension à enfanter de nouvelles structures. Les séries verbales du mwotlap sont un domaine privilégié pour observer la vie des constructions (au sens de la *Construction grammar*, cf. §2.3) et les multiples formes de réinterprétation ou réanalyse auxquelles les locuteurs se livrent dans leur pratique de la langue.

3.5.2 Constructions sérielles et relations actancielles

Les constructions sérielles du mwotlap sont la clef de bien des rouages du système ; je leur ai ainsi consacré une seconde publication. Il s'agit d'un chapitre dans l'ouvrage collectif dirigé par Dixon & Aikhenvald, et qui a fait suite à ma participation au workshop de typologie organisé en 2003 au RCLT de Melbourne.

François, Alexandre. 2006. Serial verb constructions in Mwotlap. In *Serial Verb Constructions: A cross-linguistic typology*, edited by R.M.W. Dixon & A. Aikhenvald. Explorations in Linguistic Typology. Oxford: Oxford University Press. Pp.223-238. ▷(#17)

Outre les processus de réanalyse, grammaticalisation et colexicalisation présentés ci-dessus, la sérialisation verbale pose la question des relations actancielles. En effet, une structure sérielle de type *nuclear-layer*, telle que celle du mwotlap, concatène plusieurs têtes lexicales verbales au sein d'un seul verbe composé ; ce *macro-verbe* se comporte alors comme un verbe, doté d'un seul sujet et d'un seul objet. Mais comment est déterminé le schéma actanciel (*argument structure*) de ce composé ? Si la sérialisation combine un verbe monoactanciel à un verbe biactanciel, quel sera le statut du composé ? Comment fonctionnent les arguments dans une construction comme (22) ci-dessus [*'2sg danser tenir être.calme être.bien 1sg'*] ?

Cette étude m'a permis de relever les régularités dans la distribution des actants. Ainsi, l'exemple (23') déjà cité combine une tête monoactancielle *lak* 'danser' à un modifieur biactanciel *tēy* 'tenir' :

- (23') No ⟨ma- *lak* **tēy**⟩ ni-ḡilas nōnōm.
 1sg PRF- danser tenir ART-lunettes tes
 'J'ai dansé avec tes lunettes.'

Le macroverbe résultant est biactanciel : la combinaison du verbe intransitif *danser* avec un transitif *tenir* résulte dans un verbe '*danser avec*' lui-même transitif. Tout se passe comme si le second verbe avait ajouté une place d'actant au premier – d'où l'analyse possible en terme d'applicatif.

J'ai proposé de formaliser le processus, à l'aide d'une algèbre simple. En (23'), c'est le même actant qui *danse* et qui *tient* les lunettes : je l'encode comme *x* ; le seul autre actant (les lunettes) est nommé *y*. En adoptant un ordre conventionnel ⟨S·V⟩ pour les verbes monoactanciels et ⟨A·V·O⟩ pour les verbes biactanciels, je traduis (23') ainsi :

$$x \cdot V_1 \times x \cdot V_2 \cdot y = x \cdot [VP] \cdot y$$

Le mwotlap présente d'autres combinaisons, moins triviales que celle-ci. Ainsi, une structure typologiquement très rare est celle où deux verbes sont sérialisés alors qu'ils ne partagent aucun actant :

- (25) Na-lo ⟨me- **hey** **simsim**⟩ n-aes.
 ART-soleil PRF- briller fondre ART-glaçon
 'Le soleil ⟨a fait fondre⟩ le glaçon.' [*litt. a brillé–fondre le glaçon*]

Dans de telles structures, la combinaison de deux verbes monoactanciel à sujets distincts (x et y) engendre un macroverbe biactanciel ($x\cdot[VP]\cdot y$), dont l'objet y correspond au sujet du second verbe. Fonctionnellement parlant, il s'agira toujours d'une construction *causative non-agentive*, dont la formule est la suivante :

$$x\cdot V_1 \times y\cdot V_2 = x\cdot[VP]\cdot y$$

Parmi les cas fréquents, figure aussi celui où le sujet sous-jacent du verbe V_2 n'est pas un argument de V_1 ni un référent externe, mais *le contenu propositionnel* représenté par V_1 et ses arguments. C'est ainsi que la plupart des adjectifs du mwotlap (une sous-classe des verbes) peuvent fonctionner en position postverbale, comme un équivalent d'adverbe. En (26), le sujet de *maymay* 'fort' n'est ni le requin ni sa victime, mais l'événement exprimé par V_1 (*la morsure*) :

- (26) Na-bago <mi- **ñit** **maymay**> no.
 ART-requin PRF- mordre (être).fort 1sg
 'Le requin <m'a mordu violemment>.'

Ce cas de figure, que j'ai proposé d'appeler *event-argument serialization*, peut se rendre en encodant comme E l'événement correspondant au premier verbe. Pour un énoncé comme (26), on obtient alors :

$$x\cdot V_1\cdot y \times E\cdot V_2 = x\cdot[VP]\cdot y$$

En conduisant le même type d'analyse actancielle sur mon corpus mwotlap (99 300 mots, cf. *Tableau 5* p.34), j'ai identifié 11 différentes règles de fusion des actants. Même si je ne peux toutes les illustrer ici, elles sont citées dans le *Tableau 9*.

Tableau 9 – Séries verbales en mwotlap : Les onze règles de fusion actancielle

	$E\cdot V_2$	$x\cdot V_2$	$y\cdot V_2$	$x\cdot V_2\cdot y$	$z\cdot V_2$	$x\cdot V_2\cdot z$	$z\cdot V_2\cdot y$	$y\cdot V_2\cdot z$
$x\cdot V_1$	$x\cdot[VP]$	$x\cdot[VP]$	$x\cdot[VP]\cdot y$	$x\cdot[VP]\cdot y$				
$x\cdot V_1\cdot y$	$x\cdot[VP]\cdot y$	$x\cdot[VP]\cdot y$	$x\cdot[VP]\cdot y$	$x\cdot[VP]\cdot y$	$x\cdot[VP]\cdot z$	$x\cdot[VP]\cdot z$	$x\cdot[VP]\cdot z$	*

3.5.3 Syntaxe de la ditransitivité en araki

Je me suis de nouveau intéressé aux relations actancielle dans une autre langue, l'araki, et sur un tout autre sujet : les constructions ditransitives.

François, Alexandre. 2012 a. Ditransitive alignment and referential hierarchies in Araki. In E. van Lier (ed), *Referential Hierarchies in Three-participant Constructions*, special issue of *Linguistic Discovery* 10 (3): 97-124. ▷(#12)

Cette recherche s'inscrivait dans le projet EuroBABEL "Referential Hierarchies in Morphosyntax" (autour de Balthasar Bickel et Anna Siewierska), et plus particulièrement dans l'étude typologique des constructions ditransitives, à travers la problématique question des hiérarchies de saillance.

Dans ses travaux, Haspelmath (2005, 2007b) formule clairement le parallèle entre les questions d'*alignement primaire* affectant les verbes monotransitifs (alignement accu-

satif vs ergatif) et celles d'*alignement secondaire* affectant les deux actants non-agentifs des schémas triactanciels de type 'donner'. Dans un tel schéma, on peut nommer A l'agent, T le thème (*la chose donnée*) et D le destinataire du don. La question est alors la suivante : quel argument, de T ou de D, sera aligné avec le patient P de l'énoncé mono-transitif ?

- alignement INDIRECTIF : $\langle P \Leftrightarrow T \rangle$ le patient des monotransitifs s'alignera avec le thème T du verbe ditransitif (cf. français *J'ai envoyé le chèque à ton frère*)
- alignement SECONDATIF : $\langle P \Leftrightarrow D \rangle$ le patient des monotransitifs s'alignera avec le destinataire D du verbe ditransitif (cf. anglais *They provided us with water*)
- alignement NEUTRE : $\langle P \Leftrightarrow D, T \rangle$ le patient des monotransitifs s'alignera avec les deux non-agents du verbe ditransitif (cf. anglais *They gave us some water*.)

Certaines langues adopteront préférentiellement un alignement plutôt que l'autre. En araki, les trois verbes *sle* 'donner', *varai* 'dire' et *vse* 'montrer' présentent un cas intéressant d'alternance régulière entre les deux premiers alignements, indirectif vs secondatif.

(27) **ARK** Nam= sle $\langle \text{MARA VARIRI} \rangle_D$ $\langle \text{lo hanhan} \rangle_T$.
 1sg:R= donner enfant petit:PL LOC repas
 'J'ai donné leur repas aux enfants.' [alignement SECONDATIF]

(28) Na= pa= sle $\langle \text{-ko} \rangle_T$ $\langle \text{sa-n RĀMĀRE} \rangle_D$.
 1sg:Irr= FUT= donner 2sg:OBJ DAT-CSTR monstre
 'Je vais te donner à un monstre !' [alignement INDIRECTIF]

Or j'ai montré que cette alternance dépend des propriétés relatives de D et de T sur une échelle de saillance à la Silverstein (1976) ou Bickel (2008). Parmi D et T, l'actant le plus haut sur cette échelle prend la fonction d'objet syntaxique, autrement dit s'aligne sur P ; l'actant de rang inférieur sera périphérisé dans une fonction oblique. Ainsi en (27), le thème (*hanhan*) est inanimé, et le destinataire (*mara variri*) est animé : c'est donc ce dernier qui sera encodé comme objet, tandis que T se trouve périphérisé à l'aide d'une préposition locative. Inversement, en (28), les deux non-agents sont animés : mais D est une 3^e personne tandis que T est une 2^e personne, et c'est donc le thème qui l'emporte.

La hiérarchie de saillance à l'œuvre dans ses constructions combinent deux paramètres : le trait [$\pm$ humain] d'une part, et d'autre part le trait [$\pm$ local], c'est-à-dire membre de la paire *locuteur-interlocuteur* ('local' dans la terminologie d'Andrews 1985) ou 3^e personne ('non-locale', cf. la *non-personne* de Benveniste 1946). Ces deux paramètres se combinent pour donner la hiérarchie (29) :

(29) [$\pm$ humain] [$\pm$ local] > [$\pm$ humain] [-local] > [-humain]

J'ai constaté une grande régularité des résultats chaque fois que D et T appartiennent à deux catégories distinctes parmi ces trois : un humain sera toujours favorisé par rapport à un non-humain ; un pronom de 1^e ou 2^e personne sera toujours favorisé par rapport à une 3^e.

Lorsque D et T appartiennent à la même catégorie, on constate en revanche une variation libre. Ainsi, je n’ai pas constaté de hiérarchie entre les deux premières personnes :

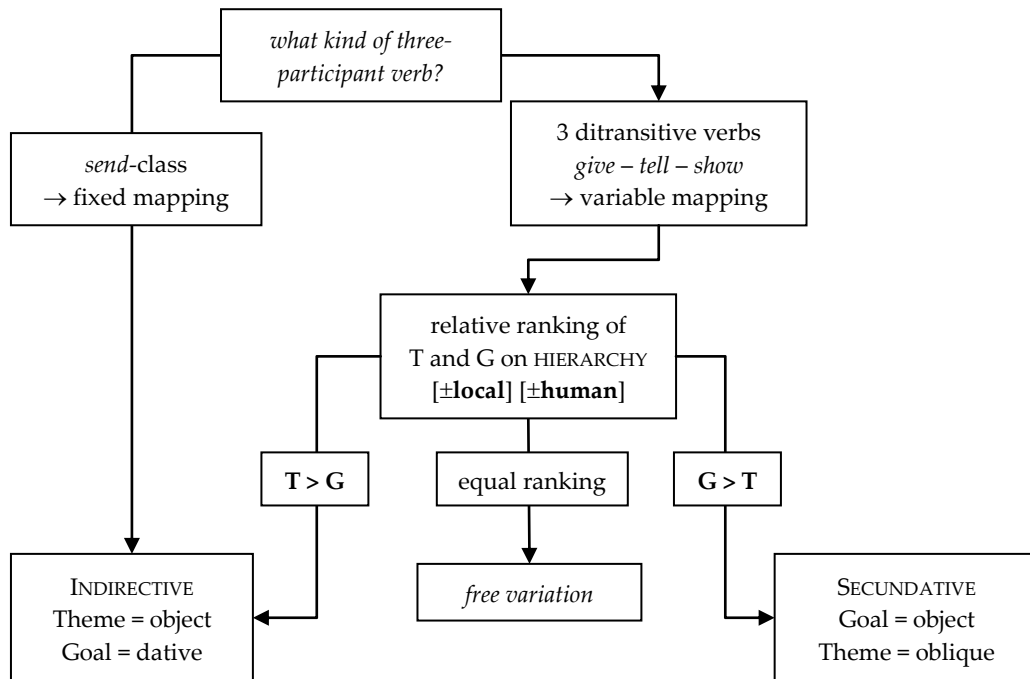
- (30) *Āpāla*-ku mo= je varai **-ko**_T <isa-ku>_D.
 ami-1sg 3:R= NEG dire -2sg:OBJ DAT-1sg
 = *Āpāla*-ku mo= je varai **-ā**_D <**ini**-ko>_T.
 ami-1sg 3:R= NEG dire -1sg:OBJ OBL-2sg
 ‘Mon ami ne m’a jamais parlé **de toi**.’

Et la variation est également libre entre deux 3^e personnes de même animéité :

- (31) *Āra* sa om= je vsei *raā-m*_D <in ***raāla-m***>_T?
 CAUS quoi 2sg:R= NEG montrer père-2sg OBL ami-2sg
 = *Āra* sa om= je vsei ***raāla-m***_T <*sa-n raā-m*>_D?
 CAUS quoi 2sg:R= NEG montrer ami-2sg DAT-CSTR père-2sg
 ‘Pourquoi ne montres-tu pas **ton ami** à ton père ?’

La syntaxe des structures ditransitives en araki est donc complexe, faite tantôt de régularités fixes, tantôt de variation libre. Je l’ai résumée en une seule figure, sous la forme d’un “diagramme de décision” (*decision chart*).

Figure 8 – Syntaxe des verbes triactanciels en araki : un diagramme de décision



directionnels, ou bien l’ambiguïté est-elle toujours de mise ? En outre, sachant qu’un énoncé mwotlap ne peut pas combiner deux directionnels, comment se fait le choix du bon directionnel lorsqu’il existe plusieurs candidats ?

En mwotlap, le directionnel s’insère à la droite du groupe verbal, dans la position des syntagmes locatifs. La fonction commune des directionnels est d’orienter l’événement en question dans une certaine direction :

- (24) MTP Nok so yem kal **hag** le-mtig en.
 1sg PROSP grimper (monter) en.haut LOC-cocotier ANAPH
 ‘Je vais grimper / monter en haut du cocotier.’

Je montre que le système joue sur trois types d’orientation :

- *orientation vers un participant* :
 contraste **me** ‘ventif, vers ici, vers EGO’ ≠ **van** ‘itif, vers là, vers non-EGO’
- *orientation définie topologiquement* :
 direction sur l’axe vertical (haut/bas) ou un axe topologique (dedans/dehors)
- *orientation définie géocentriquement* :
 direction sur le plan horizontal, en se repérant aux 4 quadrants conventionnalisés
 (vers l’est ≠ vers l’ouest ≠ vers l’intérieur de l’île ≠ vers la mer)

Or cette étude a montré, tests à l’appui, que ces trois types d’orientation correspondait à trois stratégies, lesquelles étaient hiérarchisées entre elles. Ainsi, si un mouvement se déroule sur un axe vertical en l’absence d’un participant saillant, c’est la stratégie topologique que l’on emploiera. Je l’ai remarqué un jour que ma tasse était posée sur une table, et que mon hôte me proposa d’y verser du thé :

- (32) Nok so sey se **hōw** la-bankēn nōnōm.
 1sg PROSP verser encore en.bas LOC-tasse ta
 ‘Je t’en ressers (*en bas*) dans ta tasse.’

Quelques minutes plus tard, la même scène se produisit, alors que je tenais ma tasse à la main. Mon interlocuteur eut alors recours à la stratégie personnelle, en utilisant le directionnel *van*:

- (32') Nok so sey se **van** la-bankēn nōnōm.
 1sg PROSP verser encore vers.non.ego LOC-tasse ta
 ‘Je t’en ressers (*vers toi*) dans ta tasse.’

Il m’expliqua alors que l’emploi de *hōw* ‘vers le bas’ n’aurait pas convenu ici : ma présence dans la scène était si saillante qu’elle imposait l’emploi du couple *me/van*. Je formule cette observation en disant que, dans une situation qui pourrait être décrite en usant de la stratégie personnelle autant que de la stratégie topologique, c’est la stratégie personnelle (référence à un participant) qui l’emporte.

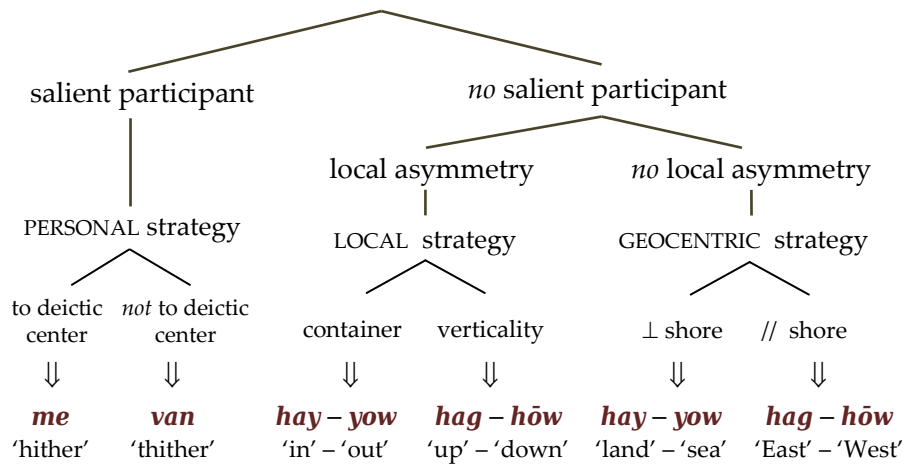
J’ai fait la même démonstration avec d’autres combinaisons. J’en ai inféré une nouvelle hiérarchie entre les trois stratégies directionnelles :

- (33) stratégie PERSONNELLE > TOPOLOGIQUE > GÉOCENTRIQUE

En d'autres termes, chaque fois que la référence spatiale par la médiation d'un participant (*stratégie personnelle*) est possible, c'est elle qui sera employée à l'exclusion des autres, comme en (32'). De même, dans une situation où le vecteur peut être encodé soit géocentriquement (ex. *hag* 'vers l'Est'), soit topologiquement (ex. *yow* 'vers l'extérieur'), alors c'est la stratégie topologique qui aura priorité. Quant à l'option géocentrique, elle n'a lieu que par défaut, lorsque les autres stratégies ne sont pas disponibles ou cognitivement activées : c'est le cas pour les vecteurs placés sur un plan horizontal, orienté vers un point dans l'espace plutôt que vers une personne.

J'ai synthétisé ces résultats dans la *Figure 9*, également un diagramme de décision. Son principe est de retracer le processus cognitif du locuteur dans sa décision d'utiliser le bon directionnel spatial.

Figure 9 – L'architecture de l'espace en mwotlap : Choix de la stratégie directionnelle



J'ai ainsi décrit non seulement le système de directionnels spatiaux du mwotlap, mais également ceux des 16 autres langues des îles Banks et Torres. Ce fut d'ailleurs là l'occasion de mettre à jour la diversité des langues de la région. J'y reviendrai dans mon chapitre comparatif [§5.1.4].

3.7 Prospective

Mon travail de description des systèmes linguistiques se poursuivra au cours des prochaines années. Il prendra la forme d'articles spécifiques, mais surtout d'au moins deux grammaires de référence, déjà commencées : l'une pour la langue teanu, l'autre pour la langue hiw. Je reviendrai sur ces projets dans le dernier chapitre [§8.1].

4 Description lexicale et changement sémantique

S'il est vrai que la grammaire se taille souvent la part du lion dans les pratiques académiques de linguistique historique, je n'oublie pas le domaine vaste et passionnant du lexique. L'étude du vocabulaire peut prendre plusieurs formes, qu'il s'agisse de rédiger des dictionnaires, de composer un atlas linguistique, ou de me pencher sur un thème précis, dans une perspective soit synchronique soit diachronique. Par ailleurs, j'ai également travaillé sur des questions plus théoriques de typologie lexicale, et souhaite les poursuivre à l'avenir.

4.1 Le travail lexicographique

L'une des formes prises par le travail de description linguistique est celle du **dictionnaire**. C'est d'ailleurs souvent cet objet particulier qui retient l'attention des communautés de locuteurs : sur le terrain, les gens ne me demandent jamais des nouvelles de ma thèse, de ma grammaire ou de mes articles sur leurs langues ; en revanche, on s'enquiert régulièrement des progrès de mon "dictionnaire". Pourtant, dans le milieu académique de la description linguistique et de la typologie, il faut bien admettre que ce genre d'ouvrage ne bénéficie pas du prestige des grammaires, et n'est pas autant encouragé. Ce relatif déficit de reconnaissance, combiné à l'immensité du travail nécessaire à la complétion d'un dictionnaire, expliquent sans doute pourquoi je n'ai pas donné la priorité à ce type de travail au cours de mes premières années. C'est une situation à laquelle j'espère remédier prochainement, d'autant que le travail lexicographique est l'une des tâches les plus agréables et passionnantes de mon travail de linguiste.

4.1.1 Dictionnaires publiés et en préparation

À ce jour, le seul travail lexicographique que j'aie publié est un lexique de la langue **araki**. Ce fut d'abord une partie de ma monographie sur cette langue [2002 ▶(#29)] avant de devenir un lexique à part entière [2008c], publié en ligne en format LexiquePro.

François, Alexandre. 2008 c. *An Araki–English–French dictionary*, 2e édition. Publication en ligne [<http://alex.francois.free.fr/AF-Araki.htm>].

Si j'ai pu publier ce lexique assez vite, c'est aussi parce que mon corpus sur la langue araki – parlée par une poignée de locuteurs, avec qui je n'ai pu m'entretenir que rarement – était plutôt réduit. Ce lexique araki de 104 pages contient 765 entrées principales ; il couvre l'essentiel du vocabulaire quotidien, mais ne prétend pas à l'exhaustivité.

C'est à une échelle différente que je travaille, depuis 1998, sur un projet de dictionnaire de la langue **mwotlap** :

Figure 10 – Extrait du dictionnaire mwotlap en préparation ▷(#36) : entrée *eh***eh (n-eh)** [neh] N. <Mus>

- (1) chant, chanson. ▷ **Na-l̄ne eh yoyoy me gēn.** *J'entends une chanson qui s'approche.* ▷ **Mahē kē so ni-tōytōy, kē ni-se n-eh, n-eh non Iqet.** *Tout en balayant, elle se mit à chanter une chanson, un chant en langue d'Iqet.*

Lexique du chant [n-eh*]

n-eh	chant, chanson
na-he~₂	chant en l'honneur de qqn
tow	composer
se₂	chanter
qētlen̄	chanter [‡]
hēgeg	entonner
nō-tōti eh*	incipit ; genre
na-vasig	couplet
ēleg goy	chanter la coda
til kal	monter vers les aigus
til tēqēl	descendre vers les graves

Les chants coutumiers

[n-eh* liwo]

Reconnaissables à leurs lentes mélodies vibrées et solennelles, les chants coutumiers (**n-eh liwo**, **n-eh kastom**) ne sont pas récités en langue mwotlap, mais dans le dialecte littéraire dit langue d'Iqet (cf. **Iqet***), impénétrable au commun des mortels. Ils sont composés par des poètes professionnels, grâce à leur savoir-faire poétique (cf. **towtow***), et aussi un certain talent surnaturel (**na-man***).

Ces chants peuvent être chantés *a cappella* (**qētqētleñ**) ou accompagnés de percussions (ex. **na-wha***, **no-koy***) ; dans ce dernier cas, ce sont plutôt des chants à danser (**n-eh ba-laklak**). Certains chants importants durent jusqu'à vingt minutes ; c'est notamment le cas des odes de prestige, composées en l'honneur d'un notable (cf. **he~₂***). Ils sont réservés aux grandes occasions, d'importance religieuse ou politique.

- ♦ **n-eh liwo** ~ **n-eh kastom**. N. <Poét> litt. "grande chanson" : chant traditionnel, chant coutumier (opp. **n-eh stiriñban**, etc.).

♦ La tradition poétique distingue de nombreux styles pour ces chants, définis par une mélodie typique et par un incipit spécifique (cf. **nō-tōti eh**). On peut citer : **na-maleñ**, **na-wlēwlē liwo**, **na-wlēwlē togolgol**, **ne-leq**, **nē-wēt**, **na-syēsye**, **rōvinēvēsēgme**, **no-towhiy**, **vawelop**, **wōisēwōvēvē**. À ces styles de chants, il faut ajouter les divers styles de danses qui leur correspondent (cf. **laklak**).

♦ **n-eh qētqētleñ** N. litt. "chant à chanter" : chant *a cappella*, chanté par une personne seule (opp. **n-eh ba-laklak** 'chant à danser', interprété par plusieurs chanteurs et accompagné de musique). ♦ Désigne en particulier certains genres de chants, ex. **vawelop**.

♦ **n-eh stiriñban** N. chanson moderne, jouée par les jeunes gens à la guitare ou au banjo (cf. angl. *String band*).

Les chansons modernes

[n-eh* stiriñban]

Au contraire des chants coutumiers, les chansons modernes (**n-eh stiriñban**, ou simpl. **n-eh**) sont accessibles à tous : composées en langue mwotlap, parfois émaillées de passages en pidgin, ces mélodies populaires expriment surtout les désirs et les amours des jeunes gens, entre quinze et trente ans. Toujours joyeuses même si elles racontent des chagrins d'amour, elles sont chantées lors des fêtes de mariage (**ne-leg**) ; mais elles sont aussi connues de tout le monde, et s'entendent à tous moments de la journée dans le village. Souvent, les paroles relatent les frasques de jeunesse d'une personne précise, encore présente dans le village, et plus âgée.

- (2) (*par ext*) mélodie, musique. ♦ Sém. **N-eh** peut désigner toute musique chantable, c-à-d. mélodique. Mais la plupart des musiques traditionnelles de Mwotlap sont des rythmes à danser ; elles sont désignées par le terme **na-wha**, ou par le nom de la danse. Il n'y a pas d'autre terme générique pour désigner la 'musique'.

[<*na āsi ; cf. Mota as 'a song']

François, Alexandre. (en prép.). *Dictionnaire culturel mwotlap–français–anglais, Vanuatu*.

[version 2011 en ligne: <http://bit.ly/AF-Mwotlap-dictionary>]

▷(#36)

La version actuelle du manuscrit, disponible en ligne, compte 312 pages, soit 1705 entrées rédigées, en format trilingue (mwotlap–français–anglais). J’ai conscience que ceci ne correspond qu’à une portion – peut-être la moitié – du vocabulaire mwotlap que j’ai recueilli sur le terrain. La *Figure 10* présente un extrait de ce dictionnaire mwotlap, et permet d’avoir une idée de l’ouvrage que je vise [un plus long extrait est donné en (#36)]. Acheter cet ambitieux projet est l’un de mes projets prioritaires pour les prochaines années.

À moyen terme, j’aimerais produire les dictionnaires d’autres langues sur lesquelles j’ai des données suffisantes – par exemple le hiw, le lo-toga, le lemerig, le lakon. De même, j’ai commencé un dictionnaire de la langue **teanu** :

François, Alexandre. (en prép.). *A dictionary of Teanu (Vanikoro, Solomon Islands)*.

▷(#35)

Sur environ 1300 entrées prévues (items lexicaux extraits de mon corpus), j’en ai rédigé 614 à ce jour. La *Figure 11* ci-dessous présente un extrait de ce dictionnaire [voir aussi l’extrait plus long fourni en (#35)].

4.1.2 Principes généraux de lexicographie

Quelle que soit la langue, je ne me contente jamais d’une nomenclature bilingue du type *liste de mots*. Si le travail lexicographique m’intéresse, c’est parce qu’il permet d’identifier et définir des catégories sémantiques particulières, souvent propres à une langue. La *Figure 11* présente une entrée au hasard, extraite de mon dictionnaire teanu–anglais en préparation.

Figure 11 – Extrait du dictionnaire teanu–anglais en préparation : entrée *botongo*

botongo VT [mbo'toŋo] (s.o.) interpose (s.th., oneself) between A and B, so as to prevent A (typically an agentive participant in motion) from reaching B (typically a static target). ♦ Synt. The grammatical object is sometimes A (yielding such meanings as ‘stop, prevent’), sometimes B (yielding such meanings as ‘protect, take care’).

(1)[object=A] block, stand in the way of ⟨s.th.⟩. ▷ **U-botongo ero etapu, u-sali!** *Stop blocking the water (with your hands), let it go!*

(2)[object=A] stop, prevent ⟨s.o.⟩ from doing s.th. (**me** or **nara** +Irr.); forbid. ▷ **Li-botongo ene me ne-le.** *They're preventing me from going there.* ▷ **U-botongo ini nara kape i-koie ne moe.** *Make sure he doesn't come inside the house.*

(3)[non-sg object] separate ⟨people⟩, pull

apart. ▷ **Ka ni-aiu ni-vio ni-botongo da.** *(two men fighting) So I rose and separated them.*

(4)[object=B] impede access to ⟨s.th.⟩. ▷ **Vilisao i-botongo se mataiko ponu.** *The tornado was blocking the passage.*

(5)[serialised after posture V] be in a position, whether deliberately or not, to impede access to ⟨s.o., s.th.⟩; hence (sit+) on ⟨s.th.⟩; (sit+) besides or with ⟨s.o., s.th.⟩ as a way to protect. ▷ **Bara ba-te ba-botongo okoro ponu.** *I'm afraid you may be sitting on my knife.* ▷ **Van' ni-wene ni-botongo nara kape le-punuo ñi.** *[I lie I block it] I sleep on (my money) so nobody can steal it.* ▷ **I-te i-botongo menu.** *She's sitting with the baby (to look after him).*

(6)[object=B] protect ⟨s.o., s.th.⟩ from a potential danger. ▷ **Dapa li-woi nuduro ne touro, me i-botongo temaka.** *They put*

up taboo signs on the seashore, to protect the area (from poachers).

- (7) (hence) take care of, look after ⟨s.o., s.th.⟩. ▷ **Dapa Paiu li-botongo dapa France.** *The villagers of Paiou were looking after the French.* ▷ **Kupa pi-kamai monone ne pe-ko me u-botongo.** *We brought this chest for you to look after.* ▷ **Ni-la piene ono me ne-botongo ne-mini kaipa.** *I'm recording your language so I can take care of it for you.* ▷ **U-botongo men' one!** *Take care of my kid!*

- (8) (espec) [reflexive construction] take care of ⟨o.s.⟩, be cautious. ▷ **U-botongo eo u-ejau !** *Take good care of yourself!*

- (9) [serialised after action V] do s.th. in favour of ⟨s.o.⟩. ▷ **Ni-vet' piene ni-botongo eo.** *I stood up for you.* [lit. *I spoke I protected you*]

botongo ADV (rare) [after verb] (do s.th.) so as to cover, close, protect+. ▷ **Li-dai ñe vilo li-asai botongo.** *We wrap it around a stick and tie it with a stitch* [lit. *stitch cover*].

Ma conception structuraliste – chaque lexème présente un contour sémantique spécifique, unique à sa langue – donne lieu à des entrées aussi détaillées que possible. Une attention particulière est portée à la description précise de la polysémie : chaque acception possible d'un terme doit être explicitée et illustrée – même si cela n'empêche pas, éventuellement, de proposer une signification nucléaire ou invariante (voir les deux premières lignes de l'entrée *botongo*, avant le sens n°1). Je reviendrai plus loin sur mon intérêt sur la polysémie, et ce que j'ai appelé *colexification* [§4.4].

J'accorde une grande importance à la qualité et l'authenticité des exemples choisis. Ceux-ci sont tirés de mes divers corpus de parole spontanée, qu'il s'agisse de conversations ou de textes enregistrés et transcrits [§2.2.3].

Même si mes dictionnaires sont classiquement articulés par lexème, j'ai sur la langue en général, et sur le lexique en particulier, une perspective *constructionniste*. Ceci se manifeste, dans mes dictionnaires, sous la forme de petites notes syntaxiques, associant tel sémantisme avec telle configuration. Dans l'exemple de *botongo* cité ici, je note par exemple que ce verbe peut se traduire 'avec' ou 'auprès de' (sens #5), mais spécifie dans quel contexte précis – en l'occurrence, dans une série verbale dont le premier verbe est un verbe de posture ('assis', 'couché', etc.). Ainsi, plutôt qu'une traduction simpliste qui se situerait au niveau du lexème { *botongo* = 'auprès de' }, je place ma définition au niveau de la construction dans son ensemble. De même, la définition (#2) indique les éléments de la construction 'empêcher qqn de faire qqch' (conjonctions *me* ou *nara*, suivies de l'Irrealis/Subjonctif). Par ailleurs, les crochets pointus correspondent systématiquement à l'objet syntaxique – convention particulièrement utile pour un lexème tel que *botongo*, qui peut porter sur des participants sémantiquement différents (tantôt l'agresseur A, tantôt l'agressé B).

À ces notes syntaxiques, s'ajoutent parfois des remarques sémantiques ou méta-linguistiques. Ainsi, 'hence' marque un lien logique dans la progression sémantique ; il sert souvent à signaler un cas dérivé du sens mentionné précédemment, mais suffisamment notable (parce que fréquent dans le discours, associé à des connotations spécifiques, etc.) pour être traité à part. Des symboles comme 'rare', 'freq.', '†' (obsolète) ou '‡' (rare, stylistiquement marqué) signalent la perception de tel ou tel sens dans le discours. Dans les définitions ou les exemples, des traductions littérales permettent de retracer certains cheminements sémantiques lorsqu'ils ne vont pas de soi.

Dans mon travail lexicographique, j'avoue m'inspirer de mes dictionnaires bilingues préférés : le *Robert & Collins Super Senior* (anglais-français), les Dictionnaires grec-français de Bailly, ou le grec-anglais de Liddell & Scott... Certes, ce choix fait parfois courir le risque d'une surcharge d'informations et de parenthèses de toute sorte ; mais il a l'avantage de condenser en peu de place toute la palette de propriétés (formelles, sémantiques, pragmatiques) que j'ai pu trouver associées à tel ou tel lexème.

Enfin, mes dictionnaires présentent d'autres caractéristiques encore, qui apparaissent dans l'échantillon de la *Figure 10*. Par exemple, j'enrichis occasionnellement certaines entrées d'images, ou encore d'encadrés, de deux types. Les encadrés lexicologiques fournissent une liste de termes appartenant à un champ sémantique donné. Les encadrés culturels apportent une information de type encyclopédique, éclairage utile pour permettre au lecteur d'appréhender non seulement la dénotation d'un terme, mais ses connotations pragmatiques ou culturelles : il s'agit de capturer ce que Robert (1997, 2003) appelle *l'épaisseur du langage*, "un lieu complexe où le linguistique s'associe à la fois à du linguistique et à de l'extralinguistique". Ces deux types d'encadrés apparaissent dans l'entrée *eh* 'chant' de mon dictionnaire mwotlap : l'un est lexicologique (*Lexique du chant*), et deux encyclopédiques, associés à deux sous-entrées (*Chants coutumiers* ; *Chansons modernes*). L'idée de ce travail lexicographique est d'aller plus loin que ne l'aurait fait une banale traduction du type "*eh* = chant, chanson".

Je compose mes dictionnaires dans le logiciel Toolbox (SIL). Au prix de patients efforts, j'ai trouvé un moyen d'en assouplir la structure afin de le personnaliser en fonction de mes désirs. Mes premiers pas dans la programmation, en 2000, ont précisément consisté à ajouter au logiciel (et à ses scripts CCT) de nouveaux champs qui n'étaient initialement pas prévus : "sens littéral", "contexte syntaxique", "sujet typique", "lien lexicologique avec sens précédent", "encadré culturel", "étymologie", etc.

4.2 L'atlas linguistique de Polynésie Française

Un projet lexicographique de taille est l'*Atlas Linguistique de Polynésie Française*, dont je suis un des deux coauteurs :

Charpentier, Jean-Michel & Alexandre François. (ss presse) *Atlas Linguistique de Polynésie Française – Linguistic Atlas of French Polynesia*. Berlin: Mouton de Gruyter. ISBN 978-3-11-026035-9. 2480 pages. ▷(#32)

Cette publication sous presse, prévue pour Juin 2014, a un statut paradoxal dans mon parcours. En effet, la Polynésie Française n'est pas le lieu usuel de mes recherches, et à la vérité, c'est mon coauteur Jean-Michel Charpentier (LACITO, aujourd'hui retraité) qui s'est chargé de tout le recueil des données sur le terrain – travail qui lui a pris plusieurs années, de 2004 jusqu'à sa retraite en 2010. Ma contribution a consisté à reprendre les données brutes – listes de mots, cahiers de terrain – que m'a confiées mon collègue, les constituer en une vaste base de données, et les transformer en autant de cartes linguistiques que d'entrées dans notre base de données lexicale.

Au total, cet atlas linguistique compte 2251 entrées – et donc 2251 cartes, que j'ai dessinées, complétées, vérifiées. Pour ce faire, j'ai dû apprendre la programmation, et

consacrer de longues périodes à travailler sur ce vaste projet – une bonne quinzaine de mois au total, répartis sur plusieurs années. Si l'on ajoute à cela la rédaction conjointe des 150 pages de chapitres d'introduction, la patiente compilation de 120 pages d'index trilingues (français, anglais, tahitien), ainsi que la mise en page finale, il faut bien en conclure que cet atlas linguistique de 2480 pages aura pris une place importante, bon gré mal gré, dans mes dernières années de recherches.

Le principe de cet ouvrage, inspiré de la tradition des atlas linguistiques, est de documenter la diversité lexicale des langues et des dialectes dans le vaste archipel de Polynésie Française. Ce dernier archipel se réduit souvent, dans l'imaginaire collectif et parfois même dans les représentations locales, au seul groupe des îles de la Société, et de sa langue le **tahitien**. Celui-ci est en effet largement dominant, ayant toujours été la langue la plus puissante démographiquement, depuis les royautes anciennes jusqu'aux sociétés modernes où l'exode rural a concentré les Polynésiens de l'archipel dans la seule île de Tahiti. Certes, le tahitien est aujourd'hui détrôné par le français, la langue coloniale ; mais dans ce bras de fer entre français et tahitien, on a tôt fait d'oublier les six autres langues de l'archipel, dont le sort est d'autant plus fragile qu'elles sont parlées parfois dans des îles ou atolls qui se sont massivement dépeuplés au cours des dernières générations.

Figure 12 – Les langues de Polynésie Française (en rouge, les points d'enquête de l'atlas)

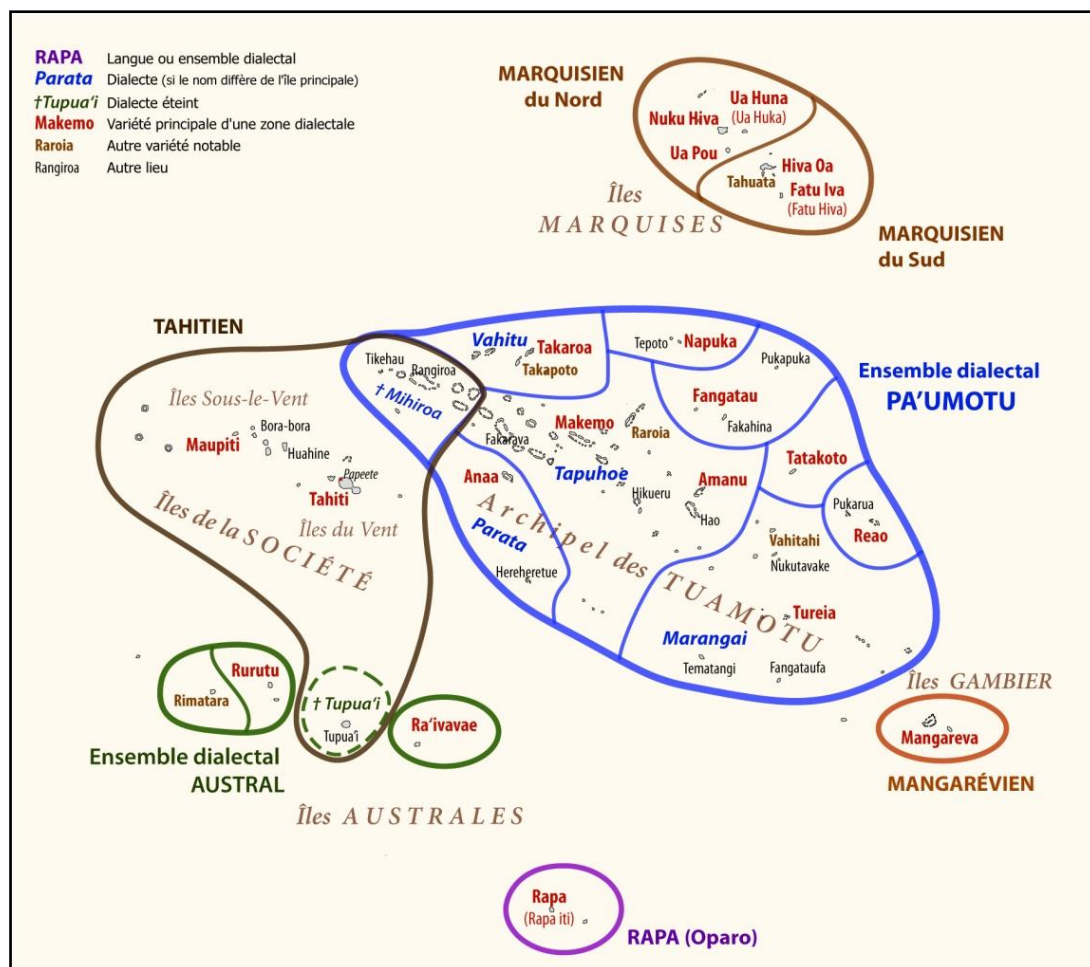
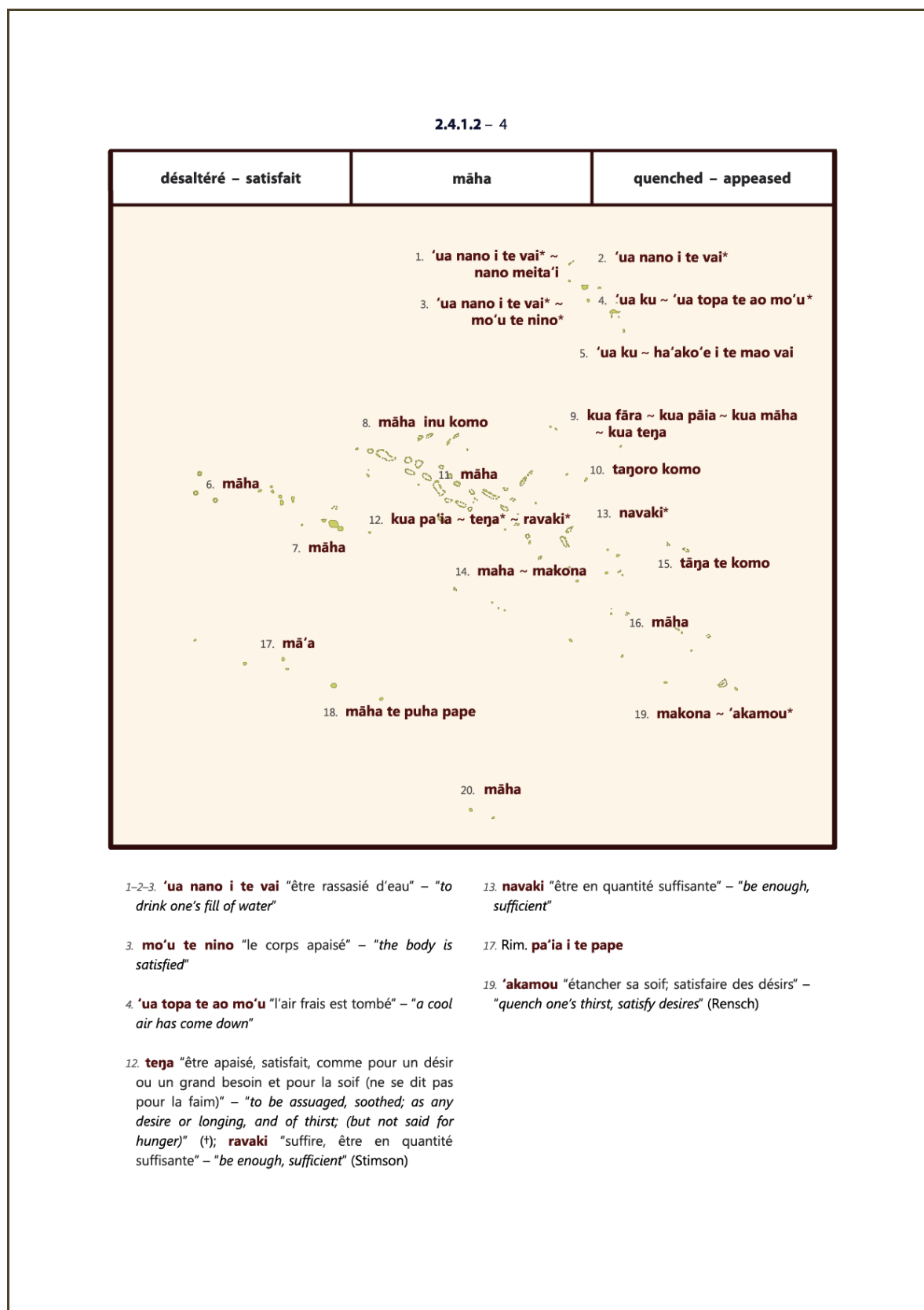


Figure 13 – Une carte extraite de l'Atlas Linguistique de la Polynésie Française



Notre atlas s'attache à documenter les sept langues de l'archipel – **tahitien, austral, rapa, pa'umotu, mangarévien, marquisien du nord, marquisien du sud** (voir la carte de la Figure 12). Du fait de la forte dialectalisation interne de ces langues – en particulier le pa'umotu, dispersé dans les nombreux atolls des Tuamotu – Jean-Michel Charpentier a choisi non pas sept points d'enquête, mais vingt. Pour chacun de ces points, il s'agissait de recueillir l'équivalent des 2251 entrées du questionnaire préparé pour l'atlas. La Figure 13, une carte tirée de l'ouvrage, illustre le principe de l'atlas. Comme on le voit, le format est très différent d'un dictionnaire, et s'inscrit dans la tradition des atlas linguistiques. Force est de constater la variété des formes que l'on peut entendre à travers le vaste archipel de Polynésie Française. Les notes, présentées en-dessous de chaque carte, incluent des nuances sémantiques, notes bibliographiques, précisions diverses (ex. 't': forme obsolète ; *Rim.*: variante dans l'île de Rimatara...).

En dépit de son caractère périphérique par rapport à la cible principale de mes recherches qu'est la Mélanésie, cet atlas linguistique de Polynésie Française reste largement cohérent avec l'esprit de mon travail. D'abord, ces années passées à côtoyer les langues de Polynésie Française m'ont donné l'occasion de me familiariser avec la branche polynésienne de l'arbre des langues océaniques, et juger de leur proximité – ou plutôt, de leur éloignement – avec leurs lointaines cousines du Vanuatu. Par ailleurs, le principe de cet atlas est de préserver la mémoire de la diversité linguistique tant qu'elle existe encore – pour le bien de la communauté scientifique, autant que pour celles des locuteurs. Enfin, ce quadrillage dialectologique aux mailles serrées, que j'ai contribué à mettre en lumière dans cet atlas, présente bien des analogies avec le type d'observations fines que je menais, en parallèle, avec les langues du Vanuatu.

4.3 Études de sémantique lexicale

En parallèle avec mes publications portant sur la grammaire, j'ai parfois eu l'occasion de réfléchir à quelques questions posées par le lexique. J'évoque ici brièvement ces travaux.

4.3.1 Sémantique lexicale et variations syntaxiques

François, Alexandre. 2000 b. Dérivation lexicale et variations d'actance : petits arrangements avec la syntaxe. *Bulletin de la Société de Linguistique* 95 (1): 15-42. ▷(#2)

Cet article, un des premiers que j'aie écrits, est né alors que je m'attélais à mon dictionnaire de mwotlap, au retour de mon premier terrain. L'une des entrées les plus intrigantes était un morphème *goy* [ɣɔj], sorte d'enclitique qui se combine productivement à droite des verbes pour en dériver un macro-verbe aux propriétés nouvelles.

Le problème de départ était la multiplicité des sens possibles de ce morphème *goy*. Avec certains verbes, il ajoutait le sens bénéfactif 'protéger, (faire V) en guise de protection' ; avec d'autres, il créait un sens détrimental 'interdire, empêcher, s'opposer'. Cette polysémie, typologiquement rare, est étonnamment parallèle avec celle du verbe teanu *botongo*, présenté dans la Figure 11 p.71 ; et ce sont en partie les mêmes interprétations sémantiques que je proposais pour le suffixe *goy*. Ce dernier, cepen-

dant, ajoutait des sens que le *teanu* n'a pas : par exemple *se* 'chanter' → *se goy* 'chanter en écho', ou *lak* 'danser' → *lak goy* 'danser sur (une musique)'.

En parallèle avec ces glissements sémantiques, une caractéristique de ce dérivatif *goy* est de modifier la structure actancielle de son verbe d'attache. Un verbe intransitif se trouvera transitivé ; un verbe transitif modifiera son orientation secondaire (au sens de Lemaréchal 1989:201) en la réorientant vers un nouvel objet. Ainsi, alors que *sal* 'suspendre <qqch>' est orienté vers l'objet suspendu, en revanche *sal goy* 'marquer sa propriété privée sur <sa maison>' en y suspendant un signe ne peut prendre comme objet syntaxique que la maison, et non plus l'objet suspendu.

J'ai proposé d'analyser ce morphème *goy*, dans certains cas, comme un applicatif. Pour certains verbes, plusieurs interprétations sont possibles, suggérant que la réorientation actancielle peut cibler des participants différents.¹ L'article observe ainsi les interférences entre glissements sémantiques et restructurations syntaxiques.

4.3.2 Termes de température

François, Alexandre. ss presse a. Temperature terms in Vanuatu languages. In Maria Koptjevskaja Tamm (ed.). *Linguistics of Temperature*. Amsterdam, New York: Benjamins.

La linguiste Maria Koptjevskaja Tamm (Univ. of Stockholm), spécialiste de typologie lexicale, m'a invité à participer à un volume collectif (sous presse) sur le lexique de la température à travers les différentes régions du monde. Dans ce court article, je montre que la structuration de ce domaine lexical dans les langues du nord du Vanuatu est généralement simple, articulée autour de deux termes *chaud/froid*, mais que certaines langues proposent trois termes. Ainsi, le hiw distingue entre deux adjectifs : *məwmiw* '(objet) froid tel que ressenti par un sujet dans son corps' (ex. eau froide, vent frais...) vs *maetit* '(objet) froid au toucher' (ex. viande froide). Le mwerlap a un verbe *malas* 'avoir froid' sans lien étymologique avec l'adjectif *mmərir* 'froid' ; etc.

L'article examine la syntaxe des termes de température. Je montre qu'une partie du sémantisme se trouve logée non dans le lexème, mais dans son appartenance à une catégorie syntaxique : ainsi, le sème de CHANGEMENT D'ÉTAT, que le français encode par des contrastes lexicaux (*chaud* vs *chauffer*), doit être assigné, dans les langues océaniques considérées, parmi les propriétés de la partie du discours ADJECTIF. En effet, tous les adjectifs sont compatibles avec une interprétation stative ou dynamique, en fonction de la marque TAM avec laquelle ils sont associés (cf. François 2003a:48-53).

Je détaille également les différents schémas actanciels (*case frames*) associés aux prédicats de température. Dans ces langues, les énoncés expérientiels comme 'j'ai froid' suivent généralement une construction dans laquelle le sujet est un inanimé, un nom *le froid* ; quant à l'expérient, il se retrouve en position d'objet, si bien qu'on a littéralement une phrase comme *le froid me fait*.

Enfin, je pose la question des extensions métaphoriques des termes de température dans ces langues. Certes, il y en a quelques-unes (ex. *chaud* = 'épicé') ; mais lorsque l'on

¹ Cf. les participants A et B dans la définition du *teanu botongo* (Figure 11 p.71).

connaît les nombreuses extensions figuratives de ces adjectifs dans des langues comme l'anglais (polysémies de *cold*, *cool*, *warm*, *hot*...), alors le plus spectaculaire dans les langues du Vanuatu est plutôt la quasi absence de telles extensions. L'article contraste cette pauvreté sémantique avec, par exemple, la vaste polysémie attestée, dans ces mêmes langues du Vanuatu, pour des oppositions du type *lourd* vs *léger* (*lourd* = 'important', 'majeur', 'solennel', 'raffiné', etc.).

4.3.3 Étymologie et changement sémantique

Avec mes études de Lettres Classiques et de linguistique indo-européenne, j'avais forcément pris goût aux délices de l'étymologie. Mettant à profit des connaissances pointues en phonologie et morphologie historique, cette discipline permet de reconstituer l'histoire des mots. On y découvre souvent des liens sémantiques insoupçonnés entre des mots apparentés : que l'on songe, par exemple, aux liens qui unissent les mots français *divin*, *jovial*, *midi* et *journal*... Cette science, qui a fait ses preuves parmi les langues à tradition écrite, mérite d'être reconnue davantage pour les lumières qu'elle peut également apporter à l'histoire des langues à tradition orale.

À travers mon travail lexicographique sur les langues mélanésiennes, j'ai souvent l'occasion d'observer des glissements de sens, des associations sémantiques plus ou moins cachées entre les mots – tantôt au sein d'une même langue, tantôt en comparant un idiome à l'autre. Aussi ne me fis-je pas prier lorsque Robert Mailhammer, alors mon collègue à l'Australian National University, me proposa de participer à un *symposium étymologique* :

François, Alexandre. "Exploring word histories in Vanuatu languages". First *Etymological Symposium* [resp. R. Mailhammer]. Australian National University, Avril 2010.

À l'issue de ces rencontres, l'organisateur mit sur pied un volume collectif sur l'étymologie. J'ai saisi cette occasion pour y présenter une anthologie de quelques reconstructions que j'avais pu faire dans les îles du nord du Vanuatu.

François, Alexandre. 2013 a. Shadows of bygone lives: The histories of spiritual words in northern Vanuatu. In R. Mailhammer (ed.), *Lexical and structural etymology: Beyond word histories*. Studies in Language Change. Berlin: DeGruyter Mouton. Pp.185-244. ➤(#23)

J'ai choisi d'illustrer les pouvoirs de l'étymologie à travers un domaine particulièrement passionnant pour qui étudie la Mélanésie : celui de la spiritualité et de la religion animiste ancestrale. Ainsi, tout en mettant en œuvre les outils de la linguistique, je pouvais en même temps profiter de l'occasion pour rassembler un grand nombre de données d'ordre ethnographique que j'avais pu recueillir lors de mes années de terrain – enquêtes où les questionnements proprement linguistiques se complètent souvent de recherches littéraires, anthropologiques ou musicologiques [§6].

Cette étude commence par expliciter la méthodologie de la reconstruction lexicale, autour notamment de la notion de changement phonologique régulier. J'avais en tête les rapprochements souvent hasardeux opérés par certains anthropologues dès qu'il s'agit de proposer une étymologie, souvent en dépit de la vraisemblance ou de toute idée de régularité phonologique. L'article propose alors une centaine de reconstructions lexicales d'intérêt ethnographique. Leur intérêt n'est pas tant la protoforme elle-

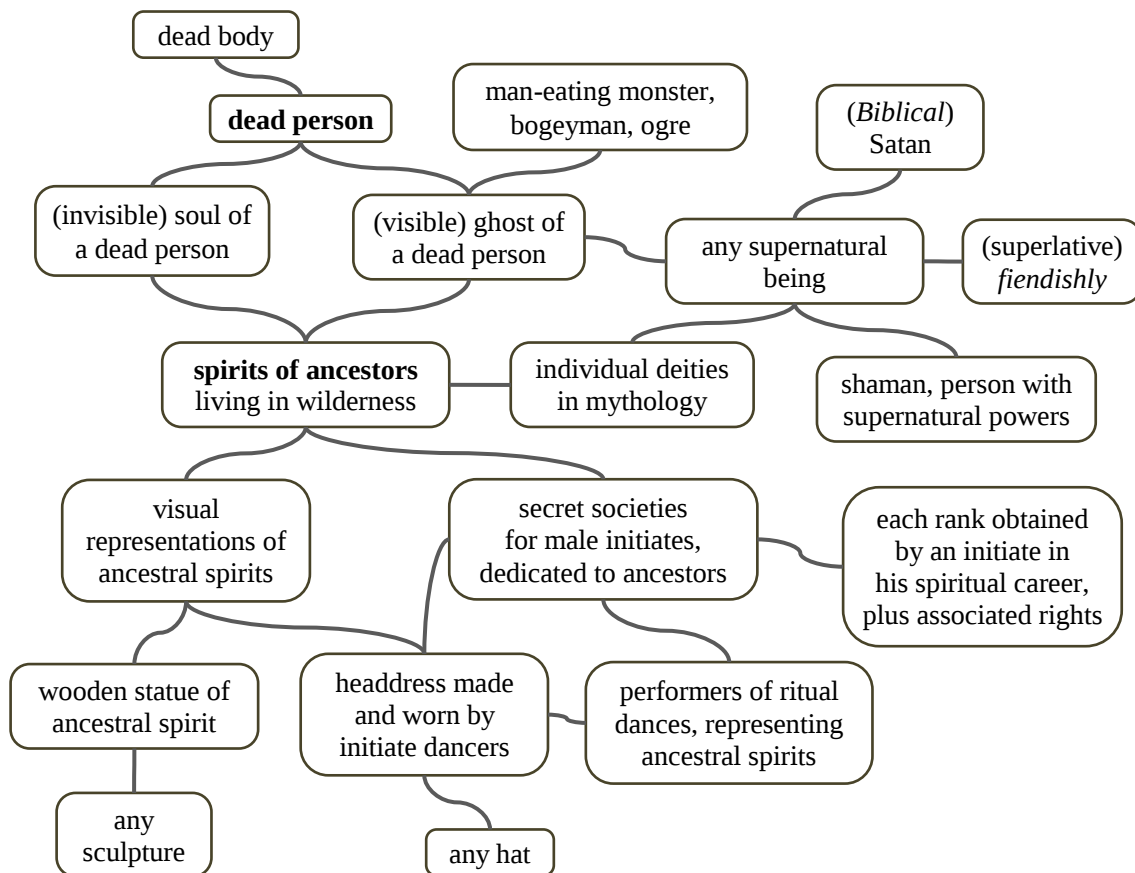
même, que le rapprochement qu'elle permet entre des termes qu'on ne penserait pas nécessairement à rapprocher.

Pour ne prendre que quelques exemples, l'article montre comment la reconstruction historique permet de rapprocher le mot lo-toga *huk^{wə}* et le dorig *βsvk*, qui désignent tous deux un ancien système politique fondé sur une hiérarchie de rangs honorifiques. Les deux termes reflètent une protoforme *su^mb^we désignant 'le chef suprême', et que l'on retrouve dans l'araki *supe* 'chef suprême, Dieu'.

De même, le nom du cimetière en mwotlap *ne-teq* [nɛ'tɛk^w] est le reflet régulier d'un étymon proto-océanien *ta^mbu 'sacré, intouchable, interdit' (lui-même à l'origine de l'anglais *taboo*, emprunté par James Cook au tahitien *tapu*). La même racine *ta^mbu présente des reflets réguliers dans le mota *tap* 'jour de deuil et de silence' ou le mwerlap *tɔm* 'signe marquant un endroit comme tabou, interdit soit aux humains soit aux esprits maléfiques (particulièrement employé pour protéger la maison d'un nouveau-né)'.

Autre exemple, la racine *[a]tamate 'défunt, fantôme' présente, dans les langues du nord du Vanuatu, une vaste polysémie (Figure 14), que certaines langues préservent intégralement, d'autres partiellement.

Figure 14 – Carte sémantique montrant la polysémie maximale associée aux reflets de la racine *[a]tamate 'défunt, fantôme, esprit+' dans les langues Torres–Banks



Les méthodes rigoureuses de l'étymologie permettent bien d'autres rapprochements sémantiques, particulièrement éclairants quant à la reconstruction des idéologies et pratiques anciennes dans la région.

4.4 Approche typologique de la colexification

La typologie des langues consiste à comparer les systèmes linguistiques entre eux du point de vue de leurs structures. L'approche typologique a toujours été, pour moi, l'horizon de mes publications. S'il est vrai que mes travaux portent généralement sur la description d'une langue en particulier, ou la comparaison de quelques langues proches, je m'adresse généralement à un public de typologues, au sens large. Les termes que j'emploie, les problématiques qui m'animent, sont toujours formulés de telle manière que d'autres linguistes, non spécialistes du domaine océanien, puissent se sentir concernés, et pourquoi pas, intégrer mes observations à leurs propres travaux.

Quant à rédiger moi-même des travaux de typologie, c'est là une possibilité que j'aimerais réaliser plus souvent. M'attaquer à la comparaison, pour telle ou telle étude typologique, de langues aussi diverses que le basque, le chinois, le russe, l'arabe, le grec ancien, le japonais..., voilà qui m'offrirait le plaisir de ressusciter les souvenirs de mes anciens apprentissages ; et j'aimerais pouvoir donner à mes réflexions sur le langage une ampleur plus universelle qu'elles n'ont jusqu'à présent. Pourtant, j'ai souvent été détourné de cet objectif par le sentiment qu'il était plus judicieux, "dans un premier temps", d'accorder la priorité à la *description* de mes abondantes données : nul autre que moi ne pouvait faire un sort aux notes manuscrites que j'avais recueillies sur le terrain, et je ne voudrais pas trahir la cause en m'éloignant trop de ma Mélanésie d'adoption.

Cette position de principe, cependant, souffre une exception : un article que j'ai consacré à une réflexion générale sur les principes d'une typologie lexicale.

François, Alexandre. 2008 a. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In *From Polysemy to Semantic change: Towards a Typology of Lexical Semantic Associations*, edited by M. Vanhove. Studies in Language Companion Series, 106. New York, Amsterdam: Benjamins. Pp.163-215. ➤(#19)

L'ensemble du volume était né des recherches d'un programme de la Fédération *Typologie et Universaux du Langage*, dirigé par Martine Vanhove (LLACAN), et consacré à la *Typologie des rapprochements sémantiques*. Inspiré par les discussions en son sein, j'ai tenté d'aborder une question centrale à la typologie lexicale que nous voulions encourager (cf. Koptjevskaja-Tamm, Vanhove & Koch 2007) : comment rendre les langues comparables entre elles ? En effet, l'un des principes du structuralisme saussurien est la thèse que chaque langue structure le sens (grammatical ou lexical) à sa manière, selon ses propres catégories [§3.1]. Or, cette vision complique l'entreprise typologique : comment comparer les lexiques des langues, s'ils sont entre eux incommensurables ? Le problème est central à la typologie, qui doit trouver un terrain commun entre les langues, malgré la singularité de leurs catégorisations (Haspelmath 2007a, 2010). Mon article propose deux outils : le concept de *colexification*, et l'usage des *cartes sémantiques*.

4.4.1 Pour une typologie de la colexification

Cet article propose d'aborder le sémantisme du lexème polysémique non comme un bloc indivisible, mais comme une collection de *sens* atomiques (qu'on appelle parfois des *acceptions* d'un mot). Ainsi, le mot anglais *straight* peut-il être vu comme embrassant un ensemble de sens possibles : 'rectiligne', 'franc', 'immédiat', 'hétérosexuel', etc. Une fois isolés, ces atomes de sens pouvait être comparés d'une langue à l'autre, afin de voir s'ils étaient associés de la même manière dans d'autres langues. Par exemple, on dira que le français *droit* associe également 'rectiligne' et 'direct' (*droit au but*), mais n'inclut pas le sens 'hétérosexuel'. En somme, les deux mots *straight* et *droit* occupent chacun une portion différente de l'espace sémantique ; en identifiant quels atomes ils ont en commun et lesquels les différencient, on peut cartographier avec précision leur contour sémantique.

Je proposais alors un nouvel outil conceptuel, le terme de *COLEXIFICATION* – à savoir le fait, pour une langue, d'associer entre eux deux atomes de sens, en permettant de les exprimer (les "lexifier") avec le même signifiant. On dira donc que l'anglais *straight* colexifie les sens 'rectiligne' et 'hétérosexuel', ou que le français *droit* colexifie les sens 'rectiligne' et 'situé à droite' (ex. *le côté droit*). Ces deux outils – atomes de sens d'une part, et colexification entre ces atomes d'autre part – sont le point de départ d'une méthode possible en typologie lexicale. On peut ainsi chercher à savoir quelle est la prévalence, dans les langues du monde, de la colexification 'rectiligne' et 'honnête' : est-ce une association rare ou fréquente ? est-elle distribuée aréalement ? On peut également partir d'un sens particulier, et se demander quels sont les autres sens qui sont typiquement colexifiés avec lui ; on peut chercher à identifier quels *schémas de colexification* (ou si l'on veut, quelles "polysémies") sont les plus fréquemment attestés.

Par exemple, en Europe, la notion de 'lumière' se trouve souvent colexifiée avec l'idée de 'connaissance' ou 'compréhension' (cf. *éclaircissement*, *lucide*, *lumineux*, les *Lumières*; angl. *bright*, *brilliant*, *enlighten*, etc.). Or on peut montrer que les mêmes métaphores se retrouvent dans d'autres familles linguistiques : ainsi le proto-océanien **marama* '(lumière de la) lune' a donné le marquisien *maʔama* 'intelligence' (Charpentier & François f/c) ; au Vanuatu, l'étymon **marani* 'lumière du jour' est présent dans **lolo-marani* /intérieur-jour/ signifiant 'intelligent, lucide, savant...' [2013a ▸ (#23), p.208], etc. Tous ces exemples de colexification méritent d'être recueillis, décortiqués, comparés d'une langue à l'autre, pour les lumières qu'ils pourraient nous livrer quant à l'existence, par exemple, de métaphores universelles. Le concept de colexification, simple mais efficace, ouvre la voie à un nouveau domaine de recherche potentiellement passionnant.

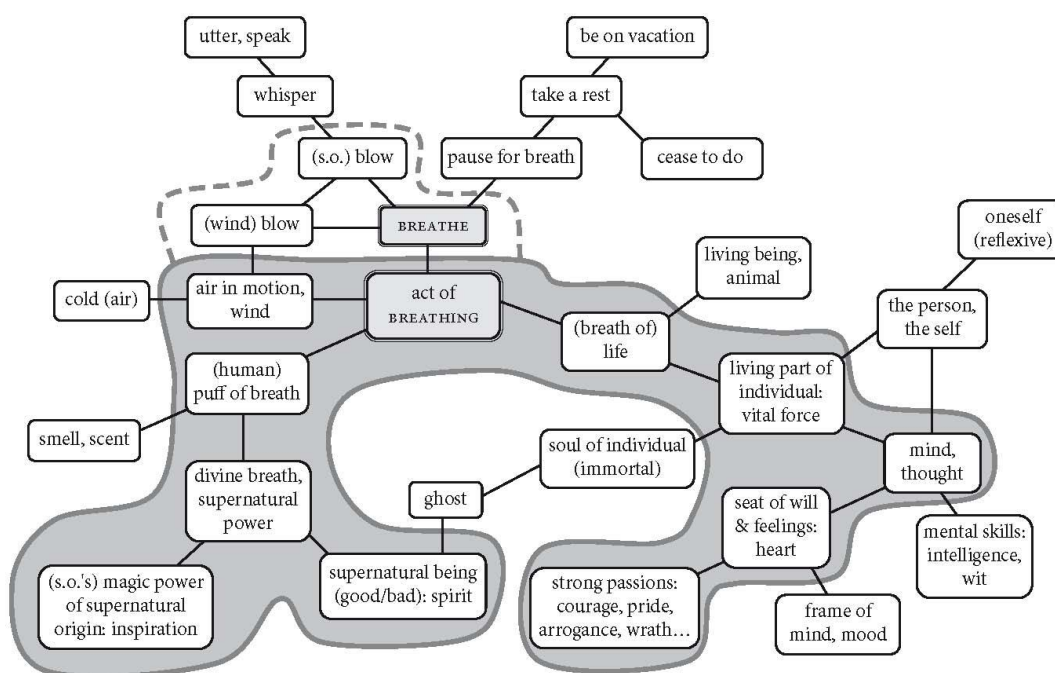
4.4.2 Les cartes sémantiques

Dans une seconde étape, mon article propose de visualiser les schémas de polysémie sous la forme de *cartes sémantiques*. Ce faisant, je reprends un outil employé fréquemment en typologie grammaticale (cf. Haspelmath 2003), et l'adapte au projet nouveau de typologie lexicale. En partant d'une notion sémantique quelconque prise comme *pivot* de la carte, je propose une méthode, basée sur la comparaison empirique des langues, pour identifier les différents sens en jeu, ainsi que les liaisons logiques entre eux.

La dernière partie de l'article consiste à illustrer la faisabilité de cette méthode, à partir d'un exemple pris au hasard. J'ai choisi de prendre comme pivot la notion de *respiration*, conçue d'abord comme l'acte physique d'inspirer et expirer de l'air par les poumons. J'ai ensuite examiné, un par un, les différents cas de colexification attestés dans un échantillon d'une quinzaine de langues de diverses familles linguistiques. Ainsi, on constate une colexification récurrente de *respirer* avec *se reposer*, voire *prendre des vacances*. On trouve parfois une chaîne sémantique *respirer* – *souffler* – *murmurer* – *parler*. Mais c'est surtout le domaine nominal qui fournit les polysémies les plus riches : par exemple, le *souffle* signifie parfois aussi le *vent*, ou le *parfum*. Mais c'est surtout le lien du souffle et de la vie qui déclenche toute une cascade de colexifications : bien des langues (grec, latin, arabe, russe, inuit...) associent le *souffle* et l'*âme*, et l'*esprit* dans ses diverses manifestations (siège des sentiments, des émotions, de l'intelligence, de l'humour). Enfin, la notion d'une *âme* éventuellement dissociée du corps ouvre la voie au domaine spirituel (du latin *spīritus* 'respiration') : fantômes (les 'esprits'), êtres surnaturels, déités diverses, se trouvent parfois associés à la racine *respirer*.

La Figure 15 présente la carte sémantique, potentiellement universelle, que cette comparaison des langues permet d'élaborer. En surimpression, cette même carte fait figurer le contour de l'*ensemble isolexique* associé au nom *pneuma* du grec classique.

Figure 15 – Ensemble des colexifications associées au nom grec *pneuma* (en gris), adossé à la carte sémantique universelle liée à la notion de RESPIRATION



Au cours des prochaines années, je compte revenir sur le domaine passionnant de la polysémie lexicale, et sur les applications possibles de la notion de colexification [§8.2]. Une piste qui m'intéresse particulièrement consisterait à mettre en valeur la contribution de ce modèle à l'étude des phénomènes aréaux [§8.2.2].

5 Linguistique historique : modéliser la diversification

Mes recherches ont souvent pris la forme d'études synchroniques portant sur la description d'un point particulier d'une langue – qu'il s'agisse de grammaire [§3] ou de lexic [§4]. Mais une autre approche, au cours de mes recherches, a pris la forme de travaux de comparaison. Cette démarche, qui implique le plus souvent des langues généalogiquement¹ proches entre elles, se situe entre la dialectologie et la description aréale. Par la force des choses, ce type de présentation a tendance à mettre en valeur la *diversité* des langues entre elles, et à poser la question des processus de *diversification*.

5.1 Études comparatives au nord du Vanuatu

Si l'on excepte mon excursus polynésien [§4.2] ou le détour par Vanikoro [§5.2.2], mon principal terrain d'enquête pour explorer la question de la diversité linguistique correspond aux îles Banks et Torres, au nord du Vanuatu. La majorité de mes publications depuis 2005 – comme il apparaît dans le *Tableau 2* p.16 – portent en effet sur la comparaison des 17 langues encore parlées dans cette région.

Le terme de “comparaison” peut recevoir ici plusieurs interprétations. Dans certains cas, mon travail a d'abord consisté à rassembler, autour d'un thème précis, mes données de première main portant sur ces 17 langues. On est là dans la continuité des études descriptives mentionnées au chapitre 3, à ceci près que ces études “horizontales”, pour ainsi dire, décrivent un phénomène non pas à travers une ou deux langues, mais à travers toute une région. En même temps, ces études aréales adoptent également une **perspective comparatiste** proprement dite, au sens de la linguistique historique. En effet, dans la mesure où toutes les langues du Vanuatu descendent d'un même ancêtre océanien, il est souvent possible de proposer des hypothèses diachroniques précises pour expliquer le développement historique de telle ou telle structure. Ainsi, observer la diversité des langues modernes sur un point précis, a pu souvent donner lieu à des reconstructions historiques portant sur les formes ou sur les systèmes. Je prendrai ici quelques exemples.

5.1.1 Les systèmes de voyelles

François, Alexandre. 2005 b. Unraveling the history of the vowels of seventeen northern Vanuatu languages. *Oceanic Linguistics* 44 (2): 443-504. ▷(#8)

Cette publication a tenu une place importante dans le développement de mes recherches sur le comparatisme historique dans la région septentrionale du Vanuatu. J'avais déjà décrit le système des voyelles du mwotlap [§3.2.1] – intéressant surtout par ses

¹ J'adopte ici la recommandation formulée par Haspelmath (2004:222), d'employer le terme *généalogique* en lieu et place de *génétique* lorsqu'il s'agit de relations historiques entre langues.

règles complexes d’harmonie, copie ou épenthèse vocalique en lien avec les principes phonotactiques de la langue. La dimension historique était déjà présente dans cette première étude du mwotlap, et m’avait donné des outils pour élargir mes hypothèses. Quant à l’article de 2005, il s’attelait à un programme plus vaste : décrire pour la première fois les systèmes vocaliques des 17 langues des îles Torres–Banks, et proposer une théorie explicative unifiée pour expliquer leur développement historique.

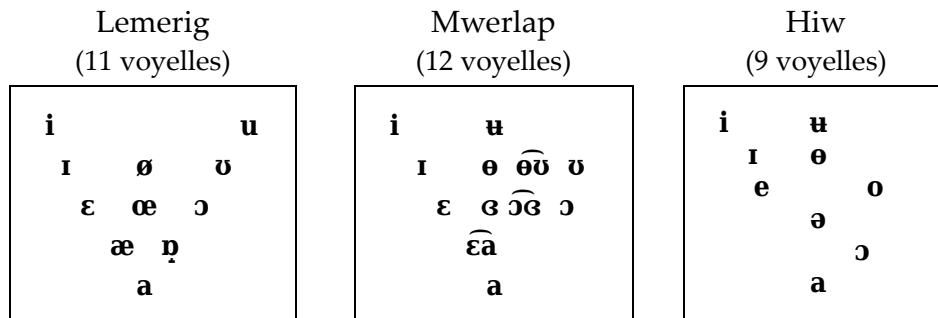
La première chose qui frappe l’observateur est la diversité des inventaires vocaliques dans la région. Le *Tableau 10* [2005b:445; 2011b:194] reproduit les systèmes vocaliques pour chacune des langues de la région. J’y distingue les monophthongues brèves des diphtongues ou des phonèmes longs.

Tableau 10 – Les 17 langues Torres–Banks et leurs inventaires de voyelles

<i>langue</i>	<i>nom</i>	<i>inventaire des phonèmes vocaliques</i>	<i>nb V</i>
HIW	<i>Hiw</i>	i i e a ə ɔ o u	9
LTG	<i>Lo-Toga</i>	i e ɛ a ə ɔ o u + ie ie ia oa oo	8+5
LHI	<i>Lehali</i>	i i ɛ æ a ə ɔ u	10
LYP	<i>Löyöp</i>	i i ɛ æ a œ ø y ɔ u + ie	10+1
VLW	<i>Volow</i>	i i ɛ a ɔ u	7
MTP	<i>Mwotlap</i>	i i ɛ a ɔ u	7
LMG	<i>Lemerig</i>	i i ɛ æ a ɔ œ ø ɔ u	11
VRA	<i>Vera’a</i>	i i ɛ a ɔ u	7
VRS	<i>Vurës</i>	i i ɛ a œ ø y ɔ u + ia	9+1
MSN	<i>Mwesen</i>	i i ɛ a ɔ u	7
MTA	<i>Mota</i>	i e a o u	5
NUM	<i>Nume</i>	i i ɛ a ɔ u	7
DRG	<i>Dorig</i>	i i ɛ a ɔ u + a:	7+1
KRO	<i>Koro</i>	i i ɛ a ɔ u + ea	7+1
OLR	<i>Olrat</i>	i i ɛ a ɔ u + i: i: ɛ: a: ɔ: u: u:	7×2
LKN	<i>Lakon</i>	i i ɛ æ a ɔ u + i: i: ɛ: æ: a: ɔ: u: u:	8×2
MRL	<i>Mwerlap</i>	i i ɛ a ɔ ɔ u ɔ u + ea oa oo	9+3

Les systèmes de voyelles s’échelonnent du plus simple (5 voyelles en mota) au plus complexe : 11 timbres différents en **lemerig** ; si l’on y ajoute les diphtongues, on trouve 12 phonèmes en **mwerlap** et 13 en **lo-toga** ; et pour peu qu’on veuille compter les voyelles longues comme des phonèmes à part, on obtient un maximum de 16 phonèmes vocaliques en langue **lakon**. Ces systèmes présentent donc une grande diversité, non seulement dans leur nombre de voyelles, mais dans leur organisation interne : on comparera ainsi l’impressionnante symétrie du lemerig avec les configurations boiteuses du mwerlap ou du hiw (*Figure 16*).

Figure 16 – Les systèmes vocaliques modernes sont plus ou moins symétriques.



En dépit de la diversité frappante des systèmes contemporains, je tâchai de leur donner une explication historique : sachant que l’ancêtre commun à toutes ces langues, le proto-océanien, avait un système simple de 5 voyelles /i e a o u/, comment pouvait-on expliquer ces développements à la fois radicaux dans chaque langue, et étonnamment divers d’une langue à l’autre ? J’ai pu montrer que partout, le principe historique qui sous-tendait le développement des systèmes vocaliques était le même : il s’agit d’un processus de type *métaphonie* – similaire à l’*umlaut* des langues germaniques – par lequel une voyelle posttonique, avant de disparaître, a systématiquement influencé le timbre de la voyelle tonique qui précédait. On a donc suivi des processus du type *vāti > *væt ‘quatre’, *vātu > *vɔt ‘pierre’, *vūro > *vɔr ‘volcan’, etc. : la voyelle tonique s’est partiellement assimilée à la voyelle atone qui suivait, tantôt se fermant ou s’ouvrant, tantôt s’arrondissant ou s’antériorisant, en fonction des caractéristiques de cette dernière voyelle. J’ai proposé de décrire le phénomène comme une “hybridation vocalique”, car tout se passe comme si la paire de voyelles *(C)á(C)u s’était hybridée en une nouvelle voyelle *ɔ.

À partir d’un système originel à cinq voyelles, les combinaisons de voyelles caractérisant les pieds *CV₁CV₂ s’élevaient à 5×5 =25. Certaines de ces combinaisons ont fusionné, pour ainsi dire, dans la même voyelle hybride : ainsi dans la plupart des langues de la région, *(C)í(C)a et *(C)é(C)i ont donné le même résultat *(C)ɪ(C). Cependant, même si le processus d’hybridation n’a pas donné naissance à 25 nouvelles voyelles, il reste que les langues modernes affectées par le processus métaphonique – toutes les langues Torres–Banks, sauf le mota – présentent toutes un inventaire de voyelles plus riches que leur ancêtre. Pour ce qui est de la diversité des systèmes telle qu’on peut la constater, elle est due au fait que la métaphonie a donné des résultats différents d’une langue à l’autre : ainsi, un ancien */a/ tonique suivi d’un */u/ (comme dans *vātu) est reflété régulièrement par /ɔ/ en **mwesen**, /ø/ en **löyöp**, /œ/ en **lernerig**, /ɛ/ en **mwotlap**, /a:/ en **dorig**, une diphtongue /ɛa/ en **koro**, etc. Le phénomène est spectaculairement régulier à l’intérieur de chaque langue, au point de confirmer le principe néogrammairien de la régularité des changements phonétiques, fondement de la Méthode comparative. Mais cette régularité interne s’accompagne d’une grande diversité des résultats dans les langues modernes – diversité qualitative des timbres, et diversité structurale des inventaires de phonèmes.

Cette étude historique des systèmes de voyelles a constitué un jalon important dans ma recherche. D’abord, elle m’a permis de coucher sur le papier les nombreuses règles

de correspondance régulières impliquant les voyelles – au minimum (5×5 règles) ×(17 langues), sans parler des autres complexités abordées dans l'article. Ensuite, elle m'a permis de reconstruire avec un haut degré de certitude plusieurs centaines de protoformes pour l'ancêtre commun de la région, le proto Nord Vanuatu ; ces reconstructions ont été publiées pour partie dans le même article [AF 2005b, pp.489-502], et pour partie dans des publications ultérieures [par ex. AF 2013a ▶(#23)].

Enfin, ces recherches ont également constitué mes premiers pas dans ma réflexion théorique sur les processus précis de diversification linguistique entre des langues apparentées – réflexion qui allait se poursuivre au cours des années suivantes [§5.2].

5.1.2 Morphologie historique des marques d'aspect

François, Alexandre. 2009 b. Verbal aspect and personal pronouns: The history of aorist markers in north Vanuatu. In *Austronesian historical linguistics and culture history: A festschrift for Bob Blust*, edited by A. Pawley & A. Adelaar. Pacific Linguistics, 601. Canberra. Pp.179-195. ▶(#21)

Cette étude de morphologie historique est une passerelle entre mon travail de phonologie historique [§5.1.1] et mes recherches sémantiques sur le Temps-Aspect-Mode [§3.4]. En effet, j'y compare les 17 langues du nord Vanuatu, en observant la manière dont elles encodent une catégorie aspectuelle particulière, que j'appelle "aoriste".

Du point de vue fonctionnel et sémantique, cette catégorie aspectuelle a des points communs avec certaines marques qui ont pu être décrites sous le nom d'*aoriste* pour des langues berbères (Galand 2003) ou pour le wolof (Robert 1991, 1995). Cette marque TAM inclut des valeurs de narratif, d'injonctif, de subjonctif, et un fonctionnement typique d'une *forme dépendante* du point de vue aspectomodal. Ayant décrit l'aoriste du mwotlap dans ma monographie sur les catégories TAM dans cette langue [AF 2003: 165-199], j'ai cherché à savoir si la même catégorie TAM existait dans les langues voisines. Il apparaît que les six langues des Banks du Sud (Gaua et Merelava) n'ont pas cette catégorie sémantique, mais que cette dernière est partagée par les onze autres langues plus au nord (y compris, donc, le mwotlap).

Ma seconde question concernait la morphologie liée à cette catégorie. En effet, la forme de l'aoriste implique une consonne /k/ dans certaines langues, une voyelle /a/ dans d'autres, parfois une syllabe /ni/ voire une marque *zéro*, sans que le lien entre ces formes ne soit très clair. Le travail de comparaison – y compris au sens du comparatisme historique – m'a permis de reconstruire un ancien paradigme d'aoriste à flexion personnelle, avec les quatre formes **^hgu* (>/k/) 'Aoriste:1sg', **u* (>Ø) 'Aoriste:2sg', **ni* (>/ni/) 'Aoriste:3sg', **(k)a* (>/a/) 'Aoriste:non.sg'. Je conclus que ces marques aspectomodales fléchies sont très probablement le reflet d'un ancien paradigme austronésien de pronoms personnels (**ku* '1sg', **mu* '2sg', **ña* '3sg'). Ce paradigme de formes pronominales "légères" se serait spécialisé dans les propositions dépendantes de type subjonctif ou narratif – d'où sa grammaticalisation comme marque d'aoriste – tandis que le prédicat principal se serait systématiquement associé aux formes lourdes des pronoms personnels, devenues standard.

5.1.3 Les articles des noms

Les linguistes spécialistes des langues de Mélanésie furent choqués par le décès inopiné du prolifique linguiste Terry Crowley, en 2005. Un volume fut destiné à lui rendre hommage, auquel j'ai tenu à participer.¹ Or l'un des (nombreux) intérêts que je partageais avec ce dernier concernait l'évolution historique des articles nominaux dans les langues du Vanuatu. C'est sur ce thème que je choisis de faire porter mon chapitre dans cet ouvrage :

François, Alexandre. 2007. Noun articles in Torres and Banks languages: Conservation and innovation. In *Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley*, edited by J. Lynch, J. Siegel & D. Eades. Creole Language Library, 30. New York: John Benjamins. Pp.313-326. ▷(#18)

Ce chapitre présente la morphologie et le fonctionnement des articles nominaux dans les 17 langues Torres–Banks. J’ai déjà évoqué ici l’article nominal *na*- de la langue mwotlap, qu’il s’agisse des règles de morphophonologie qui le concernent [§3.2.1] ou de sa raison d’être syntaxique [§3.3.2]. Dans l’ensemble, les langues Torres–Banks possèdent des articles similaires, généralement de même origine (<proto-océanien **na*), mais avec des règles de morphologie propre, que je ne détaillerai pas ici. Le chapitre a d’abord consisté à recenser les formes de l’article, non seulement au singulier, mais également son comportement au pluriel, et la question du marquage en nombre. La comparaison des langues de la région a révélé un point important : l’existence dans certaines langues non pas d’un article mais de deux, voire trois.

D’abord, plusieurs langues ont préservé un article personnel **i*, bien connu des austronésianistes (cf. Lemaréchal 1991, 2010). Cet article – que le mwotlap a perdu – précède les noms propres de personne et certains noms humains, lorsqu’ils se trouvent en fonction d’actant : ex. (mwesen) **e** *Tevêt* ‘David’, **e** *maranag* ‘le chef’ [voir aussi le dorig **i** *vvi-k* ‘ma mère’ en (34a)]. Pour ces termes, l’absence de l’article n’est possible qu’en fonction vocative : *maranag!* ‘chef!’ Autrement dit, pour les langues qui l’ont gardé, cet article personnel consiste à translater des *noms personnels*, qui à eux seuls n’ont de fonction que vocative, en des syntagmes nominaux susceptibles de référer. Le principe est donc parallèle à la fonction de l’article des noms communs (généralement *<*na*), lequel transforme en syntagme référentiel des lexèmes qui ont la fonction primaire est *modifieur* [cf. §3.3.2] :

- nom personnel : ⟨fonction primaire = *vocatif*⟩
+ article personnel ***i**
⇒ syntagme nominal référentiel (actant)
- nom commun : ⟨fonction primaire = *modifieur*⟩
+ article commun ***na**
⇒ syntagme nominal référentiel (actant)

¹ Cet auteur laissa également quatre grammaires de langues de Malekula (Vanuatu), que John Lynch publia à titre posthume. À la demande de la revue scientifique KITLV, j'en ai écrit un compte rendu extensif (*review essay*) en 2007.

Enfin, une caractéristique de cinq langues des îles Banks (vurès, mota, nume, dorig, koro) est l'existence de deux articles pour les noms communs : **na* et **wo* [cf. isoglosse #5 dans la Figure 25 p.111]. Le plus étonnant est la manière dont ces deux articles sont syntaxiquement répartis :

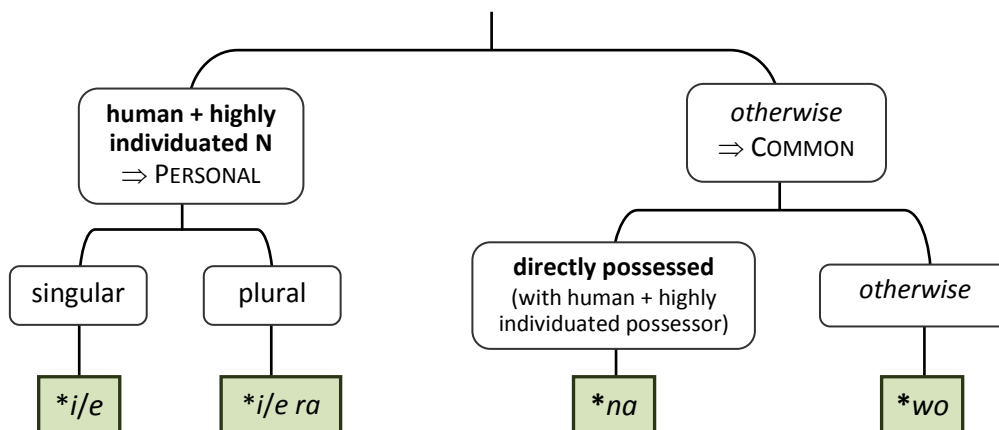
- l'article **na*, qui dans les langues voisines est requis pour tous les noms communs, est ici réservé aux **noms communs possédés de manière inaliénable par un possesseur humain référentiel** (et donc hautement individué).
- l'article **wo* couvre tous les autres cas :
 - (1) noms communs non possédés,
 - (2) noms communs possédés de manière aliénable,
 - (3) noms communs possédés de manière inaliénable par un possesseur [-humain]
 - (4) noms communs possédés de manière inaliénable par un possesseur [+humain] [-référentiel].

Cette distribution idiosyncratique est partagée par les cinq langues en question ; l'exemple (34a-c) l'illustre pour la langue dorig :

- (34a) DRG **na** ssa i vvi-k
 ART:POSS.INAL nom ART:PERS mère-1SG
 'le nom de ma mère' (possesseur [+hum] [+référ])
- (34b) **ɔ** ssa rkɸ^wa
 ART:DEFAULT nom femme
 'un nom de femme' (possesseur [+hum] [-référ])
- (34c) **ɔ** ssa wasɲin neɲ
 ART:DEFAULT nom lien DÉM
 'le nom de cet endroit' (possesseur [-hum] [+référ])

Au bout du compte, le système des articles dans les langues Torres-Banks compte un maximum de quatre éléments : **na*, **wo*, **i/e* et son pluriel **i/e ra*. Dans les langues – comme le dorig – qui les reflètent tous les quatre, l'architecture de ce système peut être représentée sous la forme d'un diagramme de décision (Figure 17).

Figure 17 – Architecture maximale des systèmes d'articles dans les langues Torres-Banks



5.1.4 Les systèmes de directionnels géocentriques

J'ai présenté en §3.6 mon étude synchronique de l'espace en mwotlap. Celle-ci fut suivie de deux études comparatives sur le même sujet, celui des systèmes de directionnels spatiaux, en particulier géocentriques, dans les langues océaniques. Le premier de ces articles – qui suivit immédiatement mon étude centrée sur le mwotlap – consistait en une comparaison de 16 langues, distribuées dans l'arbre généalogique océanien de manière à observer les tendances générales à l'échelle de la famille :

François, Alexandre. 2004 c. Reconstructing the geocentric system of Proto Oceanic. *Oceanic Linguistics* 43-1 (jun 2004): 1-32. ▷(#6)

Plus récemment, j'ai mené une étude comparative sur ce thème, mais à un grain d'observation beaucoup plus précis. En effet, il s'agissait de comparer, dans tous leurs détails, les systèmes géocentriques des 17 langues Torres et Banks :

François, Alexandre. (ss presse b). The ins and outs of *up* and *down*: Disentangling the nine geocentric space systems of Torres and Banks languages. In *The Languages of Vanuatu: Unity and Diversity*, ed. by A. François; S. Lacrampe; S. Schnell & M. Franjeh. Studies in the Languages of Island Melanesia. Canberra: Asia Pacific Linguistics Open Access. 74 pp. ▷(#25)

Tout en portant sur des questions similaires, les deux études présentaient une différence d'échelle : d'un côté, 16 langues prises comme un échantillon représentatif des 450 langues océaniques ; de l'autre côté, les 17 langues du nord du Vanuatu.

5.1.4.1 La reconstruction du système proto-océanien

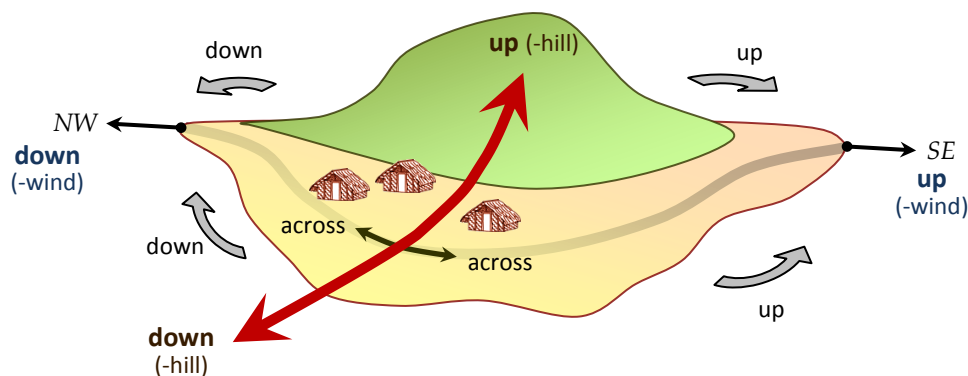
Ma première étude s'appuyait sur des publications préalables, en particulier sur les études de cas qui constituaient l'ouvrage collectif *Referring to space: Studies in Austro-nesian and Papuan languages* (Senft 1997). Parmi celles-ci figurait notamment une synthèse sur les langues de Nouvelle-Calédonie rédigée par Françoise Ozanne-Rivierre (Ozanne-Rivierre 1997b) ; mon mentor au LACITO, c'est d'ailleurs elle qui m'avait ouvert les yeux, à travers un précédent article (Ozanne-Rivierre 1997a), sur le grand intérêt des systèmes spatiaux dans les langues océaniques. En comparant systématiquement 16 langues de cette famille, j'ai proposé de reconstituer le système géocentrique du proto-océanien, l'ancêtre commun à toutes ces langues. Ce système est représenté dans la *Figure 18*.¹

Décrire ce système en détail ici serait trop long, et je me contenterai de l'essentiel. Toutes les langues océaniques connues construisent leur système sur un unique axe cardinal, orienté SE ≠ NO. Partout, l'axe est lexifiée à l'aide d'un contraste *up/down*, où *up* = 'vers le sud-est', et *down* = 'vers le nord-ouest'. J'ai pu montrer que cet axe était fondé sur le sens du vent dominant, les alizés, qui du fait de la force de Coriolis, soufflent en direction du nord-ouest dans cette partie du monde (zone entre l'équateur

¹ J'emploierai les gloses en anglais (*up/down*, etc.) parce qu'elles fournissent un équivalent très pratique pour traduire les directionnels océaniques, qui n'ont pas vraiment d'équivalent français.

et le Tropique du Capricorne). Le contraste vertical haut/bas était ainsi employé par les navigateurs océaniques pour lexifier un contraste entre *up* = '(en remontant) face au vent' et *down* = '(en descendant) sous-le-vent' ; ce que la langue anglaise fait également à travers les termes *upwind* et *downwind*, respectivement.

Figure 18 – Le système des directionnels géocentriques en proto-océanien



Comme le montre la Figure 18, l'axe cardinal lexifié *up*[wind] et *down*[wind] était employé pour les longues distances, ce que Bowden (1997) appelle échelle navigationnelle : distances entre les îles, ou entre villages éloignés dans une même île. C'est ainsi que dans de nombreuses langues océaniques, on "monte" vers le sud, et l'on "descend" vers le nord. Dans ma reconstruction du proto-océanien, ainsi que dans les langues conservatrices de la famille (Nouvelle-Calédonie, marquisien, etc.), l'axe cardinal n'est d'ailleurs employé qu'à cette échelle – un peu comme les langues européennes n'utilisent leur système cardinal (N-S, E-W) que pour les grandes distances. Quant à l'échelle locale, par exemple au sein du village, elle ne recourt normalement pas aux points cardinaux. Elle utilise là un système très différent, construit autour d'un contraste topographique 'vers la mer' ≠ 'vers l'intérieur des terres'. Ce dernier est traduit de diverses manières dans les langues océaniques, mais le plus fréquent est d'employer les directionnels verticaux *up/down* : simplement, alors que ces deux termes désignent des points cardinaux à grande échelle, à l'échelle locale ils désignent des points topographiques qui n'ont rien à voir avec l'axe des vents. Je désambiguïse parfois ce cas étonnant d'homophonie, en glosant une paire *up*[wind] ≠ *down*[wind], et l'autre paire *up*[hill] ≠ *down*[hill] ; pourtant, dans les langues océaniques conservatrices, les directionnels sont bien homophones (ex. en mota, *sage* = 'up' = 'vers le SE' = 'vers l'intérieur des terres').

Quant à l'autre axe du système local, il n'est pas non plus cardinal. Dans les langues conservatrices – et dans ma reconstruction du proto-océanien – le système local comporte en effet un second axe orthogonal par rapport à l'axe terre/mer. Cet axe n'est que dérivé du premier, et il est formellement symétrique, au sens où les directions des deux côtés seront lexicalement identiques entre elles : c'est un *axe transversal non-différencié* (Palmer 2002). De chaque côté on emploiera le même terme, que dans la Figure 18 je glose 'across' ; il ne signifie rien d'autre, si l'on veut, que 'en suivant une direction parallèle au littoral'.

5.1.4.2 Étude comparative des systèmes géocentriques aux Torres–Banks

Quoique publiée début 2014, cette reconstruction historique avait été rédigée en 2002/2003. Je ne connaissais alors, de première main, que le système du mwotlap, qui était assez différent de ma reconstruction océanienne ; plus proches de ce système reconstruit étaient les lointains systèmes de Nouvelle-Calédonie ou des îles Marquises, que je ne connaissais que par mes lectures. Retournant dans les îles Banks en juillet 2003, j’eus la surprise de découvrir que les cinq langues parlées dans l’île de Gaua présentaient exactement le même système que celui que j’avais reconstruit pour le proto-océanien (*Figure 18*). Ainsi, les structures directionnelles anciennes n’avaient pas seulement survécu à l’autre bout du Pacifique, mais à quelques lieues à peine de Motalava, au système pourtant bien différent. Aussi décidai-je d’inclure les données spatiales parmi les nombreuses caractéristiques que je souhaitais documenter dans les langues du nord du Vanuatu. Ce sont ces données, patiemment récoltées entre 2003 et 2011, qui m’ont permis d’écrire ma synthèse comparative actuellement sous presse.

L’observation minutieuse des systèmes de directionnels employés dans les 17 langues Torres–Banks n’a révélé l’existence ni de 17 systèmes différents, ni d’un seul système homogène, mais de **neuf systèmes** au total. Je ne parle pas des formes lexicales des directionnels, qui changent d’une langue à l’autre, mais de leurs oppositions structurales. Mon étude consiste donc en une micro-typologie des systèmes géocentriques employés dans cette région, en observant particulièrement les cas de polysémie ou, si l’on préfère, de *colexification* [§4.4.1].

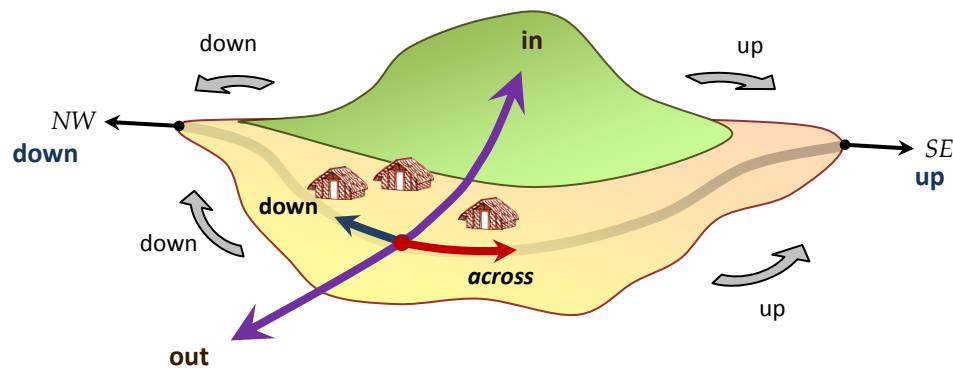
Ainsi, prenons le cas des quatre langues de Vanua Lava. Leur système a gardé des traits du système ancestral (*Figure 18*), notamment à l’échelle navigationnelle ; en revanche, son système local – visible au centre de la *Figure 19* – est très différent.

L’énoncé (35), tiré d’un conte en langue lemerig (système de Vanua Lava), est un cas typique d’emploi des directionnels référant aux quatre directions géocentriques à l’échelle locale. Alors que l’axe mer/terre implique un couple symétrique IN/OUT, en revanche sur l’axe parallèle à la côte, on voit que DOWN ne contraste pas avec UP, mais avec un morphème **wël** qu’il convient de gloser ‘ACROSS/SE’ :

- (35) LMG Ti m-sëk. Sëk lu **wël** nē, sëk lu
 3sg PFT-chercher chercher dehors ACROSS/SE là chercher dehors
sōw nē, sëk lu **row** nē, sëk lu **sar** nē —
 (DOWN/NW) là chercher dehors OUT là chercher dehors IN là —
 ti ‘esgō’ qāl’ä.
 3sg trouver NEG
 ‘Il se mit à chercher. Il chercha côté *sud-est*, chercha côté *nord-est*, chercha côté
mer, chercha côté *montagne*... Mais il ne trouva rien.’ [LMG.Qet.072]

Ce n’est là qu’un exemple parmi les nombreuses dissymétries que l’on observe parmi les directionnels spatiaux de certaines langues des Torres–Banks.

Figure 19 – Système des directionnels géocentriques dans les langues de Vanua Lava



Mon article consiste d’abord à rassembler en une seule synthèse des données portant sur 17 langues différentes, à raison de 18 vecteurs dans chaque langue. Ainsi, le *Tableau 11* compare quelques systèmes pour les différents sens qu’ils associent à leur directionnel *up* (‘vers le haut’). Malgré un évident air de famille, on constate partout des différences en ce qui concerne les appariements (colexifications) entre vecteurs.¹

Tableau 11 – Diversité des systèmes de directionnels dans les langues Torres–Banks :
Les différentes colexifications du directionnel *UP* dans cinq systèmes

LANGUE	SENS TOPOLOGIQUE	SENS GÉOCENTRIQUES	
		échelle locale	échelle navigat ^{lle}
5 lg de Gaua	‘vers le haut’ + ‘vers l’intérieur’	‘vers le sommet de l’île’ + ‘vers l’intérieur de l’île’	‘vers le sud-est’
4 lg de Vanua Lava	‘vers le haut’	—	‘vers le sud-est’
langue mwotlap	‘vers le haut’	‘// à la côte, vers le sud-est’	‘vers le sud-est’
langue löyöp	‘vers le haut’	‘vers le sommet de l’île’ + ‘// à la côte, vers le sud-est’	‘vers le sud-est’
langue hiw:	<i>up1</i>	‘vers le haut’	‘vers le sommet de l’île’
	<i>up2</i>	—	‘vers l’intérieur de l’île’
			‘vers le sud-est’

Afin de rendre intelligible cette abondance de données, j’ai choisi de les organiser en vertu d’hypothèses historiques. Ainsi, sachant que le système de Gaua est conservateur d’un système ancien (*Figure 18*), quelles sont les évolutions que l’on doit reconstituer pour expliquer les systèmes plus innovants, comme ceux de Vanua Lava, du mwotlap, du löyöp, du hiw ? Comment est-on passé historiquement de la *Figure 18* à la *Figure 19* ?

¹ La langue hiw présente un système particulièrement complexe – notamment du fait de l’existence de deux directionnels susceptibles de correspondre à *up* (*vēn* ‘*up1*’ ≠ *ag* ‘*up2*’).

Je montre que la diversité que l'on constate entre les neuf systèmes modernes peut en réalité s'expliquer par l'application, totale ou partielle, d'un petit nombre d'innovations historiques, en particulier les trois suivantes¹ :

- la généralisation (totale ou partielle) de l'axe cardinal SE/NW à toutes les échelles, et l'abandon (total ou partiel) de l'axe indifférencié parallèle à la côte ('across')
- la relexification (totale ou partielle) de l'axe terre/mer à l'aide d'un contraste IN/OUT en remplacement de l'axe UP/DOWN
- la relexification (dans 3 langues, dont le hiw) du terme vertical 'up' à l'aide d'une forme innovante.

Une fois identifiées ces innovations, il devient alors possible de retracer, d'une manière somme toute assez simple, dans quelles langues modernes elles sont reflétées – c'est ce que représente le *Tableau 12*.

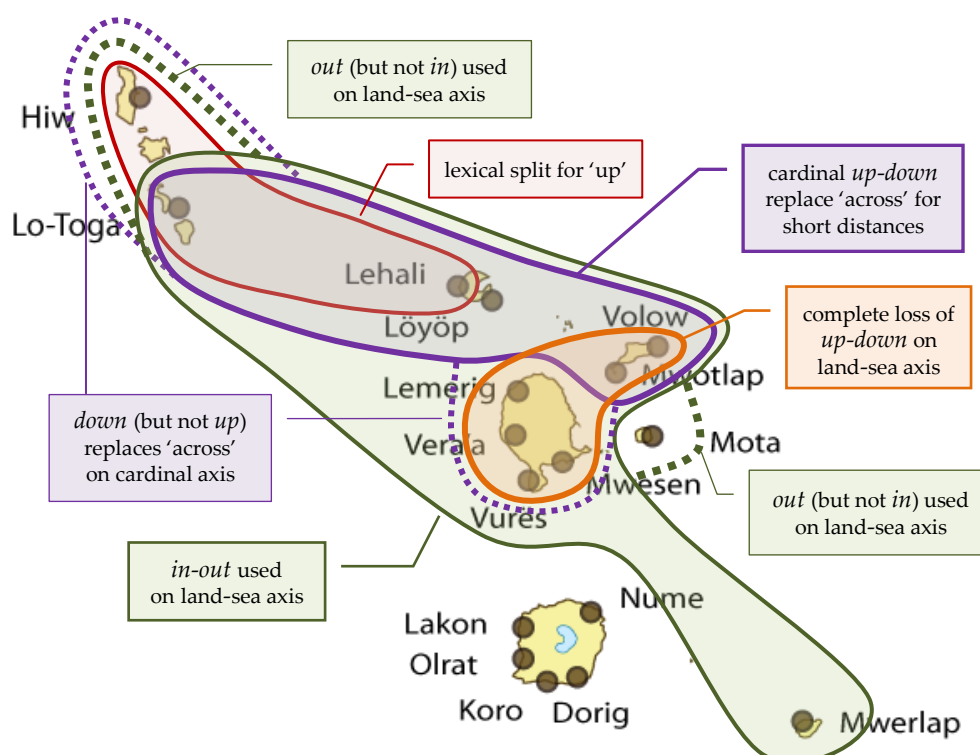
Tableau 12 — Principales innovations reflétées par les systèmes géocentriques modernes des langues Torres–Banks.

	PARALLEL TO SHORELINE		ORTHOGONAL TO SHORELINE			
	DOWN replaces across (NW)	UP replaces across (SE)	IN used for 'inland'	OUT used for 'seawards'	total loss of UP-/DOWN-hill	lexical split of UP
5 Gaua lgs	—	—	—	—	—	—
Mota	—	—	—	+	—	—
Mwerlap	—	—	+	+	—	—
4 Vanua Lava lgs	+	—	+	+	+	—
Mwotlap	+	+	+	+	+	—
Volow	+	+	+	+	+	—
Löyöp	+	+	+	+	—	—
Lehali, Lo-Toga	+	+	+	+	—	+
Hiw	+	—	—	+	—	+

Outre son intérêt pour une typologie de l'espace, cette étude historique constitue aussi parfaitement les processus historiques de diversification dans la région. Bien que les 17 langues des Torres–Banks partagent toutes un ancêtre commun, trois millénaires de divergence ont donné lieu à une mosaïque linguistique moderne. Or, cette mosaïque ne résulte pas d'un isolement total, puisque la plupart des innovations sont partagées entre langues adjacentes. Ce qu'on observe, c'est le résultat d'innovations historiques qui se sont propagées par contact d'une langue à l'autre, mais de manière souvent incomplète, dessinant des isoglosses entrecroisées (*Figure 20*) – voir aussi §5.4.

¹ Les deux premières innovations avaient déjà été identifiées dans mon article sur la famille océanienne [§5.1.4.1].

Figure 20 – La diversité moderne des systèmes géocentrique dans les îles Torres–Banks reflète l’accumulation d’innovations historiques qui se sont diffusées entre langues voisines selon des zones qui s’entrecroisent



Cette conclusion nous conduit au thème de la prochaine section : les processus historiques de diversification des langues, et la meilleure manière de les modéliser.

5.2 Comprendre la diversification des langues

5.2.1 Contexte

Mes différentes études comparatives sur les langues du Vanuatu septentrional mettent en jeu une dialectique récurrente entre l’homogène et l’hétérogène, entre l’unité ancienne et la diversité moderne. Plusieurs de mes travaux consistent à déconstruire la multiplicité des langues et des termes d’aujourd’hui pour remonter à une forme d’unité originelle : c’est le cas des recherches étymologiques, qui tissent des liens entre des mots séparés par le temps ; mais aussi des études de phonologie ou de morphologie historique, ou encore mes recherches sur les systèmes de directionnels spatiaux : à chaque fois, la variété des formes modernes autorise une explication unifiée, et la reconstruction d’un proto-système ancestral.

Plus j’observais ainsi le kaléidoscope des langues des îles Banks–Torres, plus j’étais en mesure de reconstituer les multiples cheminements qu’avaient pris les processus de différenciation au cours du temps. Les langues avaient modifié leurs consonnes et enrichi leurs inventaires de voyelles, elles avaient tantôt perdu tantôt gagné des catégories de pronoms ou de classificateurs possessifs, elles avaient modifié parfois

considérablement leur système spatial ou la signification de leurs étymons, pourtant jadis communs à toute la région. Le résultat moderne était un patchwork de 17 langues dépourvues d'intercompréhension.

Tout en poursuivant mon entreprise de description et de reconstruction de tous ces changements historiques, j'ai voulu prendre du recul, et m'interroger sur les principes qui sous-tendent ces processus. Pourquoi les langues changent-elles ? Plus précisément, pourquoi une communauté homogène à une époque finit-elle par se fragmenter en une mosaïque de micro-communautés linguistiquement différentes ? Les causes de ces changements sont-elles toutes sociales, ou cognitives, ou proprement linguistiques ? Cette fragmentation linguistique est-elle corrélée, comme on pourrait le croire, avec une isolation croissante ? ou bien a-t-elle lieu, paradoxalement, en parallèle avec des relations continues de contact de langues ? Enfin, cette apparence de fragmentation linguistique est-elle réelle, ou bien ne concerne-t-elle que certaines strates de la langue plutôt que d'autres ? Et quel modèle employer pour synthétiser le mieux ma reconstruction de l'histoire de ces langues : peut-on reconstruire un *arbre généalogique* des langues des îles Torres-Banks – et si non, pourquoi ?

Ces problèmes m'ont particulièrement préoccupé durant les années 2008-2013. Ils ont dominé ma période de recherches à l'Australian National University (2009-2012), et m'ont conduit à formuler des idées innovantes, parfois radicales, sur ces questions.

5.2.2 Unité et diversité dans une même île : Vanikoro

Bien que l'essentiel de ma recherche ait pour objet les îles du nord du Vanuatu, mes toutes premières réflexions sur la question de la diversification eurent en fait pour théâtre l'île de Vanikoro, dans l'archipel voisin des Santa Cruz, aux îles Salomon. Je m'étais rendu dans cette île à l'occasion de deux expéditions pluridisciplinaires auxquelles j'avais été convié [§2.1.3, 6.3.1].

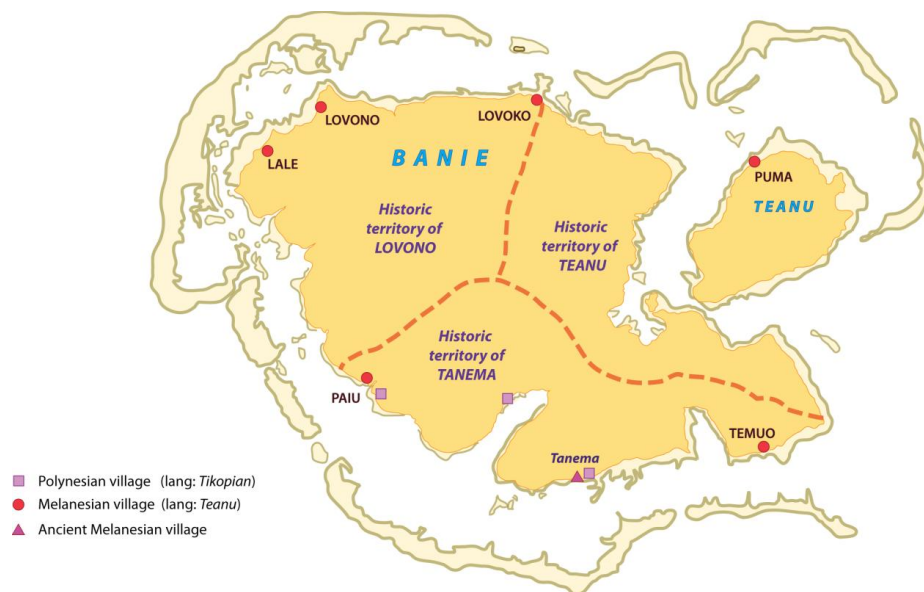
François, Alexandre. 2009 a. The languages of Vanikoro: Three lexicons and one grammar. In *Discovering history through language: Papers in honour of Malcolm Ross*, edited by B. Evans. Pacific Linguistics, 605. Canberra: Australian National University. Pp.103-126. ▷(#20)

Quatre langues sont parlées dans l'île de Vanikoro. L'une d'entre elles, le **tikopia**, est parlée par une communauté polynésienne immigrée d'environ 500 personnes ; celles-ci sont originaires de l'île de Tikopia, située à 200 km à l'est de Vanikoro, à l'extrême Est des îles Salomon. La langue tikopienne est une exclave issue de la branche polynésienne de la famille océanienne, et n'est arrivée dans la région que tardivement, au cours du dernier millénaire. La majorité de la population de Vanikoro est mélanésienne, et descend probablement des premiers navigateurs océaniques qui occupèrent la région, il y a environ 3200 ans.

Or, cette population mélanésienne parle trois langues différentes : le **teanu**, le **lovono** et le **tanema**. Le teanu s'est aujourd'hui généralisé à toute cette population, et compte donc environ 800 locuteurs. Même si la population locale a gardé le souvenir des anciennes subdivisions entre les trois tribus rivales (*Figure 21*), le *XX^e* siècle a bouleversé la donne, au point que le teanu a désormais remplacé les deux autres langues ancestrales de l'île. Au cours de mes enquêtes de 2005 et 2012, j'ai constaté que

le lovono et le tanema ne subsistent plus qu'à l'état de souvenir auprès de quelques derniers locuteurs. En 2012, le lovono n'était plus connu que par 4 personnes âgées, et le tanema par un seul homme, Lainol Nalo. Conscient de l'urgence de documenter ces deux langues, j'ai réussi à recueillir suffisamment de données – des phrases, quelques textes – pour permettre d'esquisser une comparaison avec le teanu, la langue dominante que je connaissais mieux.

Figure 21 – Territoires des trois anciennes tribus de Vanikoro



Ces trois langues m'ont frappé par leur caractère innovant, et ce, à double titre. Premièrement, les langues de Vanikoro sont incroyablement différentes des autres langues océaniques que je connaissais – en particulier, elles ne partagent rien avec les langues des Torres-Banks pourtant voisines, et n'ont ni le lexique ni la morphologie typiques des langues océaniques conservatrices. Les langues de Vanikoro apparaissent si "aberrantes" – pour emprunter le terme de Grace (1990) ou Pawley (2006) – que j'ai même un temps douté qu'elles fussent austronésiennes : étant parlées à la limite de la zone qu'on considérait alors 'papoue', la question de leur affiliation génétique était réelle.¹ S'il est vrai que, selon ma conclusion, il s'agit sans doute bien de langues austronésiennes, elles ont connu suffisamment de remplacement lexical et morphologique pour devenir méconnaissables.

Ma seconde surprise était le degré extrême de divergence y compris entre des langues très proches. Ainsi, bien que Ross & Næss (2007) aient proposé de regrouper les langues de Vanikoro dans le même sous-groupe généalogique que les langues Reef – SantaCruz voisines, j'ai toujours eu beaucoup de mal à leur trouver un quelconque point commun. Cette divergence ne concerne pas seulement les relations entre Vanikoro et ses voisines immédiates, mais également celles qui sépare les trois langues

¹ Cf. mon exposé à Oxford en 2006: "Are Vanikoro languages really Austronesian?"

de Vanikoro entre elles. Lorsque l'on sait que ces langues partagent un ancêtre commun, on peut être frappé du degré de divergence entre leurs signifiants – comme dans les énoncés (36)-(37) :

- (36) TEA a- ko u- ka u- katau ke^mba ?
 LVN nu- pu ku- ma ku- ki ^ugema ?
 TNM ^ugo- po ^ugo- loma ^ugo- ie ^uga^mbe ?
 2SG:R- dire 2SG:IR- venir 2SG:IR- suivre 1EX.DU
 'Veux-tu venir avec nous ?'
- (37) TEA pi- te ne sekele iupa, pi- wowo uo.
 LVN nupe- lu ne amenoja iemitore, nupe- ŋoa upie.
 TNM tei- o ini vasaŋola ake^ugamuto, ti- oa uva.
 1EX.PL:R- rester LOC jardin POSS:1EX.PL 1EX.PL:R- planter igname
 'Nous étions dans notre jardin, nous plantions des ignames.'

Certes, les trois langues ont un air de famille. Mais leur degré de différenciation lexicale est notable : ainsi, leurs taux de vocabulaire commun calculés sur une liste de type Swadesh (Tryon & Hackman 1983:481) se situent entre 51,1 et 57,3%. Autrement dit, dans une île peuplée de quelques centaines d'habitants, les taux sont comparables au taux le plus bas (51,1%) relevé entre les deux langues romanes les plus éloignées, le français et le valaque (Dyen *et al.* 1992).

C'est alors que prit forme ce qui allait devenir un thème central dans ma réflexion. Il s'agit du paradoxe en vertu duquel les langues que j'étudie présentent à la fois une forte hétérogénéité dans les formes de leurs mots, et une extrême homogénéité dans leurs structures – manifestée ici par l'alignement parfait des morphèmes et de leurs gloses. Tout l'article, d'ailleurs, est organisé selon ce diptyque, symbolisé par le titre "*three lexicons and one grammar*". D'une part, j'y illustrais le haut degré de différenciation des formes (dû au changement phonétique ou au renouvellement lexical), mais d'autre part, j'y mettais à jour la parfaite unité structurale de ces trois langues à tous les étages de la catégorisation (phonologie, syntaxe, parties du discours, polysémie lexicale, polyfonctionnalité grammaticale, phraséologie, etc.). Au-delà de la description des faits, ce paradoxe constituait un problème théorique qu'il me fallait résoudre.

La littérature mentionne parfois des situations similaires, où des langues lexicalement différentes présentent un alignement quasi parfait de leur morphosyntaxe. Ainsi, Gumperz & Wilson (1971) montrent que les trois langues – deux indo-aryennes, une dravidienne – parlées dans le village de Kupwar, en Inde, se sont côtoyées si longtemps, et si intensément, qu'elles en sont venues à aligner parfaitement leurs grammaires, tout en gardant distincts leurs lexiques. De même, dans le cas de Vanikoro, il est probable que ce soit la situation permanente de contact entre les trois langues, qui explique leur parfait isomorphisme, au morphème près. Cependant, la différence principale avec Kupwar, est que dans le cas de Vanikoro, les trois langues proviennent d'un seul ancêtre commun. Ainsi, alors qu'à Kupwar le phénomène le plus remarquable était la *convergence* de trois langues originellement différentes, le cas présent est l'opposé. Ce qui est vraiment spectaculaire ici, ce n'est pas tant le parfait isomorphisme

des structures (qui découle logiquement de la proximité de ces langues, à la fois héritage génétique et convergence par contact) mais le degré de *divergence* observé entre les trois langues au niveau des formes des mots.

On ne saurait se contenter d'expliquer cette divergence des formes par l'hypothèse simpliste de l'*isolement* des locuteurs, au sens d'un isolement physique du type migration. Certes, la tradition orale mentionne des conflits anciens entre les trois tribus, mais elle fait également état d'alliances et d'intermariages, conditions toujours propices au contact linguistique. Clairement, la divergence des formes (lexèmes ou morphèmes) avait donc eu lieu *en dépit* d'une situation de contact soutenu, dans une configuration qui aurait plutôt, dans d'autres régions du monde, donné lieu à un maintien de l'homogénéité originelle. Il fallait identifier la force centrifuge à l'œuvre dans les langues de cette région, force capable de créer de l'hétérogène et de la diversité à partir d'un état autrefois homogène.

Dans cet article sur Vanikoro, je discutais le concept d'*esoterogeny*, proposé par Thurston (1989) et repris par Ross (1996, 2001). Le terme désigne la tendance bien connue, chez certains groupes sociaux, à promouvoir des codes langagiers "ésotériques" connus d'eux seuls, et opaques aux groupes extérieurs – sur le mode des argots ou sociolectes restreints de nos contrées. Thurston et Ross expliquent certains processus de diversification des langues en Mélanésie par essentiellement le même type de phénomène sociolinguistique. À mon tour je reprends cette analyse, tout en me démarquant de la description de ces auteurs en termes de "processus *conscient* de différenciation". Il semble plus adéquat de déceler dans ce phénomène un effet de "main invisible" (pour reprendre le terme de Keller 1989, 1994), non délibéré, en vertu duquel les attitudes sociales dans cette région du monde finissent par donner lieu à la configuration observée : homogénéité des structures, et renouvellement intense des formes.

Cette réflexion, esquissée avec les langues Vanikoro, s'est affinée par la suite, lorsque je me suis posé les mêmes questions au sujet des langues que je connaissais mieux – celles du nord du Vanuatu.

5.2.3 Divergence et convergence parmi les langues Torres-Banks

Les observations que j'avais pu conduire lors de mon détour par Vanikoro ont été confirmées dans les îles du nord du Vanuatu. Il ne s'agissait pas de trouver des similarités grammaticales entre ces deux régions, car les structures observées sont en réalité fort différentes ; plutôt, il s'agit de constater que chacune de ces deux régions – Vanikoro d'un côté avec ses 3 langues, les îles Torres-Banks de l'autre avec leurs 17 langues – présentait une homogénéité structurale interne particulièrement marquée, alors même qu'elles constituaient des mosaïques linguistiques fort hétérogènes du point de vue des formes des mots.

Les éléments du diagnostic initial sont similaires à la situation décrite plus haut pour Vanikoro. Les 17 langues du nord du Vanuatu partagent toutes un ancêtre commun – le proto-océanien – et sont parlées par des communautés qui ont manifestement entretenu des relations soutenues de *contact* (commerce, mariages inter-insulaires, etc.) au cours des 3200 ans de leur développement historique. Ces deux circonstances – origine commune, contact permanent – sont les deux facteurs qui

expliquent le fort isomorphisme structural qu'on peut observer entre les langues modernes (ordre des mots, catégories syntaxiques, organisation sémantique du lexique ou de la grammaire, etc.). Pourtant, au lieu de résulter aujourd'hui dans une homogénéité linguistique, ce que l'on constate est au contraire une hyper-fragmentation : 17 langues sans intercompréhension, dont les vocabulaires se sont fortement dissimilés au cours du temps. Le *Tableau 13*, en donnant les équivalents du même énoncé (transcrit en API) dans les 17 langues, illustre la configuration typique que l'on observe sur le terrain :

Le *Tableau 13* permet d'avoir un bon aperçu de la situation linguistique dans la région. Il ne fait nul doute que les langues ont un air de famille, et l'on trouvera aisément des similitudes de proche en proche dans les marques personnelles, les possessifs, l'article, la négation... Mais on sera également frappé par le haut degré de différenciation dans les formes, des lexèmes (ex. 'savoir', 'parole'...) autant que des morphèmes grammaticaux. Et pourtant, ces 17 langues partagent essentiellement les mêmes structures, le même ordre des mots et les mêmes catégorisations et règles : partout la négation est bipartite, partout le premier élément de cette négation se place entre le pronom sujet et le verbe, partout son second élément se place entre l'adverbe et l'objet, partout le possessif suit le nom...

Tableau 13 – Homogénéité des structures, hétérogénéité des formes au nord du Vanuatu : un échantillon

<i>HIW</i>	sisə	tati	əjmə ⁹ Len	wu ⁹ Lɔχ	k ^w e	i	nə	məŋa	=ta
<i>LTG</i>	nihə	tat	lɔlmərən	ʊrβe	k ^w ε	e	nə	βəχəβaχə	mətə
<i>LHI</i>	kəj	tətne	χlal	χalse	k ^w ɒ		n-	βap	munχen
<i>LYP</i>	kīɛj	tɛ	χilal	tʃəjmat	tʃɛk ^w ɛ		n-	βaβap	ŋm ^w ɔniɛn
<i>VLW</i>	ʊgij	ɛt	χilal	χalsi	tɛ ⁹ gb ^w ɛ		n-	χatχat	njɔnχim
<i>MTP</i>	kij	ɛt	ɪχlal	χalsi	k ^w ɛtɛ		nɔ-	hohole	nɔnɔnχim
<i>LMG</i>	tær	ɪ	χɒlɒl	ʔormaʔ	ʔæ.kiʔis		n	tɛktɛk	mɔχɔt
<i>VRA</i>	ˈdir	ɪʔ	lamai	ɛntɛχ	ʔim		m	tɪktɪk	mu ⁿ di
<i>VRS</i>	nɪr	χiti	χilal	wareχ	tɛn		ɔ	k ^w ak ^w	namɔχynɪn
<i>MSN</i>	nɪr	ɛtɛ	lɪli	maŋtɛ	βɪs		ɔ	χatle	mɔχɔnin
<i>MTA</i>	nira	χate	χlala	mantax	tuk ^w ɛ		o	βaβae	naŋm ^w unina
<i>NUM</i>	nɪr	βitis	χil	liŋliŋi	mi		u	luwluw	namχin
<i>DRG</i>	nɪr	sɔwse	βɪɾɪχɪl	taβul	tɛ		na	lɪɟa	-χɪm
<i>KRO</i>	nɪr	tɪ	rɔŋ	taβul	wɔs.mɛlɛ		ɔ	βalβalaw	namɪχɪn
<i>OLR</i>	nɪj	tɪ	rɔŋ	βɪli:	wɔs.mɛlɛ			ususra:	mɔtʃ
<i>LKN</i>	χɪ:	atɪ	rɔŋ	kɛrɛ	aβɔh.malɛ			ɛlɟa	-nχɪtʃ
<i>MRL</i>	kɛr	ti	βalχɛar	mɪnmɪm	tɪk ^w ɪtɛa		nɔ-	liŋɪ	-χɛan
	3pl	PAS.ENC ₁	savoir	bien	PAS.ENCORE ₂	[OBL]	ART	parole	POSS:1incl.pl

‘Ils ne connaissent pas encore très bien notre langue.’

5.2.3.1 Pourquoi la divergence ?

Après une première esquisse dans un volume d'hommages à mon directeur de recherche Alain Lemaréchal [AF 2010a], j'ai surtout publié mes résultats dans un article de revue internationale :

François, Alexandre. 2011b. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1 (2): 175-246. ➤(#10)

J'y examine les trois principales voies par lesquelles les 17 langues des Torres-Banks se sont différenciées autour du temps :

- [pp.192-198] CHANGEMENT FORMEL RÉGULIER, au niveau du système phonologique (cf. l'éclatement des systèmes de voyelles, §5.1.1)
- [pp.198-203] CHANGEMENT FORMEL IRRÉGULIER, propre à tel ou tel lexème. Ainsi dans certaines langues, le nom de la 'langouste' ne reflète pas un étymon **ura* mais une forme **ira*, au vocalisme inattendu (cf. lemerig [n-ær], reflet régulier de **ira* mais pas de **ura*).
- [pp.203-210] REMPLACEMENT LEXICAL : quand un lexème est remplacé, pour un sens donné, par un lexème non apparenté. Ainsi le nom du 'crabe de cocotier', reconstituable comme **ndaeru*, a été remplacé en löyöp par /n-ⁿdø-ŋiēj/, litt. 'feuilles de pandanus'. J'affirme que ce processus de remplacement lexical peut s'expliquer par des processus de changement sémantique attestés ailleurs dans le monde.

J'ajoute là un argument important, à savoir que ces trois types de changement linguistiques sont universellement attestés. Par conséquent, ce qu'il faut expliquer n'est pas tant la nature qualitative des changements eux-mêmes (changement formel, remplacement lexical par changement sémantique), mais leur densité géographique. Car ce qui est propre à cette région, c'est le maillage extrêmement serré de la diversité linguistique, avec 17 langues pour une population de 9400 âmes – l'une des plus fortes densités linguistiques au monde [2012b ➤(#13)]. Les taux de similarité lexicale (44,5%) sont encore plus bas que les minimums des familles romanes ou germaniques, pourtant parlées par de bien plus vastes populations.

Mon hypothèse est alors la suivante. Nul n'est besoin de supposer un phénomène particulier du type *ésotérogénie* que l'on définirait, comme Thurston (1989), comme une action "consciente" de complexification d'une langue par ses locuteurs pour la rendre opaque aux non-initiés. Les processus de changement linguistique au Vanuatu sont de même nature qu'ailleurs : le changement émerge naturellement, par glissements spontanés (glissements sémantiques, micro-changement phonétique, etc.), lors des interactions langagières entre individus. Si de telles innovations individuelles existent également dans les sociétés européennes modernes, la particularité de ces dernières est la tendance au nivellement linguistique à l'échelle de vastes populations. Les innovations linguistiques en Europe ont typiquement un taux élevé de déperdition, au sens où telle tournure innovatrice, propagée auprès d'un cercle d'une centaine de personnes (un village, un groupe social...) aura tendance à disparaître rapidement ; ou alors, si elle se fixe dans l'usage de ce groupe – comme c'est l'usage de tel jargon sociolectal, vocabulaire propre à tel groupe urbain... – elle ne sera généralement pas en mesure d'affecter

l'usage standard de la langue majoritaire. Cette dernière (le français standard, l'allemand standard etc.) préservera typiquement son apparente homogénéité, et continuera de transcender les micro-lectes en son sein. La langue standard, du fait d'une loyauté généralisée des individus vis-à-vis d'un vaste réseau de locuteurs connectés par les médias de masse, résistera davantage, par effet d'inertie, à la propagation des innovations. Bien sûr les changements existent également dans ces macro-langues, mais ils se propagent lentement et difficilement dans la population, et beaucoup ne viennent jamais grossir le fleuve de l'usage *mainstream*.

Les paramètres qui définissent l'*écologie linguistique* en Mélanésie sont bien différents de ce que connaissent les grandes langues modernes de communication. Certes, la différence est en partie technique, au sens où les langues mélanésiennes ne sont traditionnellement associées ni à des médias de masse, ni à l'écriture – technologies qui ont contribué à l'uniformisation linguistique de vastes ensembles géographiques au cours des derniers siècles. Inversement, en Mélanésie, on ne dispose traditionnellement d'aucun mode de transport terrestre tel que le char ou le cheval ; on se déplace exclusivement à pied, dans des îles volcaniques si escarpées que le rayon d'action d'un individu au jour le jour ne dépasse guère un rayon de 10 à 20 km.

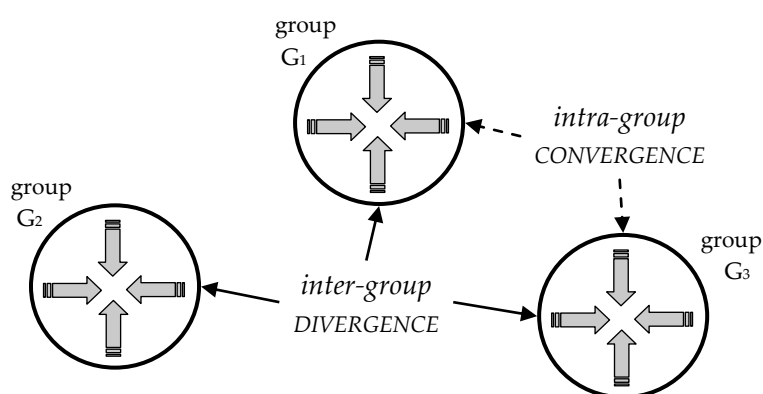
À ces facteurs techniques ou géographiques, s'ajoutent des arguments de nature sociopolitique et idéologique. La Mélanésie se caractérise par l'absence de tout pouvoir politique centralisé – au contraire, par exemple, des royaumes de Polynésie (Sahlins 1963, Pawley 2007) – et par une grande tolérance, voire une appétence, pour l'hyperdifférenciation entre groupes sociaux. Dans cette société essentiellement égalitaire et horizontale, dépourvue de hiérarchies ou de classes sociales, les catégories sociologiquement pertinentes ne sont pas, comme en Europe, définies sur des critères économiques ou socioprofessionnels, mais identifiées par un ancrage exclusivement territorial : le groupe auquel l'individu appartient sera toujours associé à une aire géographique spécifique. Intervient alors la question de l'*échelle* où se définit le groupe social le plus pertinent, celui auquel le sujet locuteur tend à rattacher principalement son action : est-ce le hameau, l'île, l'archipel, la nation ? Dans la Mélanésie traditionnelle, l'échelle pertinente est manifestement celle du village, éventuellement d'un groupe de trois ou quatre villages voisins avec lequel l'individu interagit au quotidien. C'est à cette communauté-là (que l'on pourrait appeler "*micro-communauté*" si l'on ne craignait l'ethnocentrisme) que le sujet mélanésien aura tendance à s'identifier, et partant, à aligner ses pratiques culturelles et linguistiques.

Dès lors, qu'une innovation linguistique émerge et se diffuse au sein de mon village, il est probable que je l'adopte vite, en vertu du mimétisme bien connu qui caractérise les expressions linguistiques de loyauté à un groupe (Le Page & Tabouret-Keller 1985). La propagation du changement est d'autant plus rapide qu'elle a besoin, pour être stabilisée, de ne cibler qu'un cercle de quelques dizaines de locuteurs. À l'inverse, si à l'occasion d'une visite dans une communauté voisine, je suis témoin d'une innovation lexicale, la pression que je ressentirai pour aligner mes propres pratiques linguistiques sera bien moindre : car il ne s'agit plus là de mon groupe social, celui auquel je suis censé ressembler. La diffusion d'innovations linguistiques d'une communauté à l'autre existe bel et bien, et nous verrons qu'elle est d'ailleurs fré-

quente ; mais elle n'est pas nécessaire pour que cette innovation puisse rentrer dans l'usage *mainstream* d'une langue particulière.

En somme, si cette région de Mélanésie présente aujourd'hui cette exceptionnelle densité linguistique et ce taux élevé de différenciation lexicale, ce n'est dû ni à une plus grande fréquence des innovations, ni à un désir "conscient" de générer de l'opacité et de la "complexité". La mosaïque que l'on observe s'explique par l'*écologie sociale et linguistique* de la région, et en particulier la structuration de la société en petites communautés ancrées territorialement. Chacune de ces communautés opère comme un foyer de convergence entre locuteurs, d'une manière qui résulte typiquement – mais pas nécessairement – dans une divergence par rapport aux communautés voisines. Car tel est le paradoxe central : **le processus social véritablement pertinent est la CONVERGENCE entre sujets locuteurs ; l'effet de DIVERGENCE entre communautés linguistiques n'est qu'un épiphénomène** (cf. Figure 22).

Figure 22 – La divergence globale résulte d'événements de convergence locale :
Les locuteurs adoptent des innovations locales au sein de leur réseau social, accroissant indirectement leurs différences avec les groupes voisins



5.2.3.2 Pourquoi la convergence ?

Voilà donc l'explication que je propose à une question lancinante des travaux portant sur la diversité linguistique en Mélanésie : la différenciation des langues a pour origine non pas un "désir de diverger" (cf. *esoterogeny* de Thurston) mais un *désir de converger*, une tendance spontanée à imiter les pratiques culturelles et linguistiques de ceux dont on se sent proche. Ce qui distingue les zones à forte diversité linguistique telle que la Mélanésie, des zones fortement homogènes, ce n'est ni le type de changement linguistique en jeu, ni le processus de diffusion d'individu à individu ; c'est plutôt la *dimension typique du cercle social* qui fonctionne comme foyer de convergence pour les individus. Cette donnée, au passage, n'est pas linguistique à proprement parler, mais de nature sociopolitique, économique, et idéologique.

Il reste cependant un aspect du paradoxe de départ qui n'est pas résolu. En effet, le *Tableau 13* p.99 illustre le haut degré de différenciation lexicale entre les langues, mais il montrait également à quel point elles étaient parfaitement parallèles du point de vue des structures grammaticales. Une partie de l'article [AF 2011b:211-228] est d'ailleurs consacrée à apporter d'autres illustrations de ce point. Mais comment expliquer un tel

isomorphisme structural ? On ne peut pas tout expliquer par l'héritage d'un ancêtre commun, et il est indéniable qu'une bonne part de cet isomorphisme moderne est dû aux relations de *contact* entre communautés linguistiques.

En elle-même, cette convergence par contact n'est pas problématique : l'article cité de Gumperz & Wilson (1971), entre autres, nous rappelle que c'est un phénomène courant. Si paradoxe il y a, c'est dans l'observation que les phénomènes de *convergence* par *mimétisme* – qui sont la clef de tout – opèrent à deux échelles contradictoires :

- les innovations qui affectent la FORME SONORE du message (changements phonétiques, renouvellements lexicaux) ont un rayon géographique de propagation réduit – typiquement celui du village ou de la communauté proche
- les innovations qui n'affectent pas la forme sonore du message, mais seulement sa STRUCTURE (structures syntaxiques, organisation du lexique, polysémies, phraséologie) ont un rayon géographique de propagation généralement bien plus étendu – typiquement l'archipel des îles Torres-Banks dans son ensemble.

Comment expliquer cette différence de traitement ? J'ai fait alors l'hypothèse que ce contraste était lié à la manière dont ces deux volets de la langue sont accessibles à la conscience du locuteur. Les innovations lexicales ou phonétiques affectent la forme sonore du message, et sont cognitivement assez saillantes pour subir un processus de filtrage au moment de la diffusion : la forme nouvelle est perçue comme nouvelle, et ne sera aisément adoptée par le locuteur que si elle remplit des conditions d'ancrage dans son groupe social proche, foyer de convergence. À l'inverse, les innovations affectant les structures ne sont pas accessibles à la conscience du locuteur : elles sont donc libres de se propager de sujet à sujet à travers le vaste réseau social de l'archipel, sans autre filtrage que l'efficacité dans la communication.

Au bout du compte, ce qui apparaissait initialement comme un contraste entre divergence et convergence, se trouve donc résolu. Dans tous les cas, le changement linguistique implique des phénomènes de convergence entre locuteurs. Ce qui change, c'est l'échelle géographique et sociale à laquelle opère cette diffusion – cercle étendu pour certaines innovations, cercle rapproché pour d'autres. Plus une innovation sera cognitivement perçue comme nouvelle, plus son rayon de propagation dépendra de la taille de l'unité sociale fonctionnant comme foyer de convergence.

Quant à l'effet de diversification des langues, il ne s'agit que d'une conséquence indirecte, à l'échelle macroscopique, de ces phénomènes de micro-diffusion à l'œuvre dans les interactions entre locuteurs.

5.2.3.3 Diffusion de ces résultats

J'ai d'abord présenté mes réflexions à plusieurs cercles de linguistes :

- | |
|---|
| <p>Sept 2007. "Parallel meanings, divergent forms in the North Vanuatu Sprachbund. Diffusion or genetic inheritance?". Communication au 7^{ème} Colloque de l'Association Internationale de Linguistique Typologique (ALT7), tenu à Paris, du 25 au 28 septembre 2007.</p> <p>Oct 2008. "Structural parallelism and diversity of forms across the languages of north Vanuatu: One or several communities?". Communication à l'atelier: <i>Écologie du changement linguistique</i>. LACITO-CNRS, Villejuif, 23 oct 2008. [resp. F. Jacquesson]</p> |
|---|

- | |
|--|
| <p>Oct 2008. “Innovation of forms, preservation of concepts: The paradox of constructional inertia in north Vanuatu languages”. Communication à l'atelier: <i>Grammatical Constructions in Time and Space</i>. University of Oslo, Oslo, 30-31 oct 2008. [resp. Åshild Næss]</p> <p>Jul 2009. “Local words, shared ideas. Lexical divergence and structural homogeneity among north Vanuatu languages”, Conférence annuelle de l'ALS (<i>Australian Linguistics Society</i>), Melbourne, Australie. [CONFÉRENCIER INVITÉ—KEYNOTE ADDRESS]</p> <p>Aug 2009 — “Lexical diversification and structural parallelism in north Vanuatu: Two pressures in conflict” – Linguistic Seminars, University of Newcastle [resp. Bill Palmer].</p> <p>Nov 2010 — “Language divergence and convergence in north Vanuatu: Competing social forces and their linguistic correlates” – Universität Leipzig [resp. Balthasar Bickel].</p> |
|--|

Ce cycle de présentations, au cours desquels j’ai pu bénéficier des commentaires de mes collègues, s’est achevé par la publication de mon article de 2011 (cité en §5.2.3.1). En même temps qu’il tentait de présenter et discuter ce paradoxe, cet article constituait une synthèse de 14 années de recherches sur ce groupe de langues. En effet, le problème que je soulevais mettait en branle la totalité de ce que j’avais pu observer dans cette région : ma réflexion y mobilisait des connaissances sur le changement phonétique ou morphologique, sur les catégories syntaxiques, sur la description sémantique du lexique... À cela s’ajoutaient des observations cruciales de nature sociale concernant la transmission des langues, les traditions de multilinguisme, les modalités du contact linguistiques, le goût de la diversité culturelle, les perceptions locales de l’homogène et de l’hétérogène – toutes observations constitutives de ce que j’y ai appelé “language ecology”. J’ai mis beaucoup du mien dans cette publication.

5.3 Reconstituer les processus historiques de diffusion

M’étant rendu compte du rôle clef joué, dans l’évolution des langues proches, par les processus de diffusion, je décidai de m’intéresser à ce problème de plus près.

5.3.1 La disparition de la consonne *R

Au détour d’un article dans lequel je décrivais le système phonologique de la langue hiw et un phonème latéral rare [AF 2010b ▶(#9), cf. §3.2.2], je me suis un jour intéressé à la consonne du proto-océanien d’où elle provenait – une rhotique notée *R. On ignore la prononciation exacte de cette consonne, si ce n’est qu’elle devait autrefois se distinguer de l’autre rhotique, notée *r – peut-être un contraste de type *r [r] ‘r roulé’ vs *R [r] ‘r battu’ ? Mais l’intérêt de ce phonème *R n’est pas tant de reconstruire ses propriétés phonétiques, que d’observer sa distribution géographique. En effet, Geraghty (1990) avait émis l’hypothèse que si l’on recherchait les reflets de ce proto-phonème *R dans les langues océaniques modernes, on observait un étrange schéma de distribution : plus on allait vers le sud-est de la Mélanésie insulaire (des îles Salomon vers le Vanuatu, la Nouvelle-Calédonie, et Fidji), moins les *R étaient préservés. J’ai voulu à mon tour tester cette hypothèse, en profitant des riches données accumulées au cours de ces dernières décennies, par moi-même et par d’autres chercheurs, sur les langues du Vanuatu. Il en est résulté un nouvel article.

François, Alexandre. 2011 c. Where *R they all? The geography and history of *R loss in Southern Oceanic languages. *Oceanic Linguistics* 50 (1):140-197. ▷(#11)

Mes résultats confirmaient l'hypothèse de Geraghty, et de manière spectaculaire. Prenant en compte plus d'une centaine d'étymons incluant la consonne *R, j'ai pu constater qu'en effet, plus une langue était parlée vers le sud du Vanuatu, plus elle avait de chances d'avoir perdu *R dans ces lexèmes, alors que la consonne survivait le plus souvent au nord. L'exemple (38) illustre ce principe avec le nom du poulpe, *kuRita. Il en donne les reflets dans les langues modernes, ordonnées géographiquement du nord (Hiw) au sud (langue lelepa, dans l'île d'Efate) :

(38) POc *kuRita 'octopus, squid' >

HIW ^glitə; LTG ɣəritə; LHI n-wejɛt; LYP n-wujɛt; VLW ni-wijit; MTP na-wjit;
 LMG wiriʔ; VRA wiriʔi; VRS wirit; MSN wirit; MTA wirta; NUM wirit; DRG writ;
 OLR wurit; LKN wirit; MRL ni-wirɛāt ♦ Sungwadia wita; Raga ɣuita; Hukua huita;
 Sakao neð; Araki huirā; Nguna wiita; Lelepa wiit.

Loin d'être chaotiques, les données sont cohérentes. Le losange ici situé entre le mwerlap (île de Merelava), et le sungwadia (île de Maewo) signale la frontière de rétention de *R pour ce mot. Au nord de ce point, toutes les langues retiennent le *R (sous la forme de [r] ou [j] ou [^gL]) ; au sud, toutes les langues reflètent sa disparition.

Or le plus étonnant, est que le schéma nord-sud est toujours confirmé, alors même que l'emplacement exact de la ligne de fracture diffère selon les lexèmes. Fondée sur l'observation d'une centaine d'étymons et de leurs reflets modernes, la *Figure 23* synthétise les résultats en une seule carte, en révélant les 15 lignes de fractures observées à travers le Vanuatu. Chaque ligne correspond à un ou plusieurs étymons – par exemple la ligne #7 correspond à la ligne de partage pour l'exemple (38) ci-dessus, *kuRita 'poulpe'.

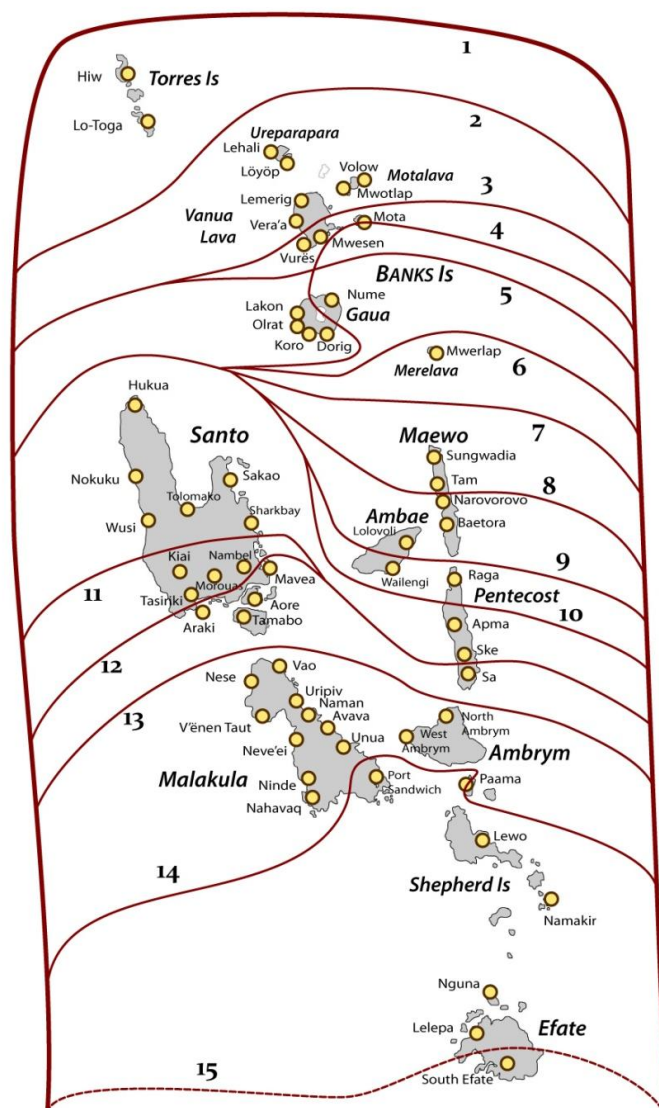
5.3.2 Diffusion lexicale, diffusion dialectale

Si ces résultats sont instructifs, ce n'est pas tant pour ce qu'ils nous disent du destin de cette consonne *R en particulier – car j'admets qu'il s'agit là, de prime abord, d'une préoccupation réservée à quelques historiens initiés aux langues océaniques. Le problème est surtout dans l'interprétation de ces résultats en termes de modèle historique. En effet, si tout le monde s'accorde à considérer les langues du Vanuatu comme descendues d'un ancêtre commun – le proto-océanien –, les océanistes diachroniciens (en particulier les travaux de John Lynch) ont toujours privilégié le *modèle arborescent* pour penser l'histoire de ces langues. Telle proto-langue se serait divisée en deux descendants, eux-mêmes séparés en plusieurs langues héritières, et répandues sur le territoire à la faveur de migrations d'une île à l'autre...

Je reviendrai plus bas sur les problèmes posés par les représentations cladistiques (arborescentes) en phylogénétique des langues [§5.4.1]. Dans l'article en question, j'explique que la distribution des reflets de *R rend très peu vraisemblable, et donc caduque, toute tentative de résumer l'histoire linguistique du Vanuatu sous la forme d'un arbre, censé représenter des phases de migrations. Le type de phénomène illustré

ici résulte bien plus probablement d'un processus de *diffusion* d'un changement linguistique à travers des réseaux de locuteurs déjà établis dans l'archipel. Plutôt que des déplacements de populations – interprétation fréquente mais naïve des arbres génétiques – ce qu'on voit ici c'est le résultat d'un changement phonétique irrégulier (la perte aléatoire de *R dans certains lexèmes) tel qu'il s'est diffusé d'île en île, au cours d'interactions de contact entre populations.

Figure 23 – Distribution géographique des reflets de *R au Vanuatu du nord et centre: Toutes les isoglosses séparent les langues selon une pente nord-sud.



La Figure 23 suggère que ce changement phonétique est parti de quelque part au sud de cette carte, et s'est propagé vers le nord. Les langues les plus touchées par l'effacement de *R sont celles du sud, les plus proches de l'épicentre, et les plus épargnées sont celles des îles Torres tout au nord, les plus éloignées. Le phénomène correspond à un cas de *diffusion lexicale* (Krishnamurti 1998), au sens où un changement phonétique affecte le lexique lexème après lexème, et non pas d'un seul coup au niveau phonologique. Il est très comparable au fameux *Rhenish fan*, ce dégradé dialectal en

zone germanique en vertu duquel les mutations consonantiques augmentent progressivement à mesure qu'on avance vers le sud-est.

Au-delà de la dialectologie spécifique de cette consonne du proto-océanien, j'avais trouvé là un moyen de reconstituer des schémas géographiques de diffusion du changement linguistique à une date ancienne. Parce que la diffusion lexicale opère lexème après lexème, elle offre un outil particulièrement raffiné pour observer le chemin emprunté autrefois par la diffusion des innovations. La *Figure 23* peut être vue comme une cartographie des réseaux sociaux à l'époque (sans doute très ancienne) de la perte de *R : tel un marqueur radioactif permettant de visualiser l'histoire ancienne de la Terre, la phonologie historique nous donne ici le moyen de reconstituer les réseaux sociaux du Vanuatu préhistorique : liens privilégiés entre le sud de Santo et le sud de Pentecôte, relations des îles Banks avec Maewo mais pas avec Santo... Appuyée par la puissance de la Méthode comparative, la dialectologie historique offrait une fenêtre inespérée sur l'histoire précise des réseaux sociaux dans l'archipel.

5.4 Modéliser les processus de diversification

Je ne souhaitais pas m'arrêter en si bon chemin. Les résultats de mon étude historique du *R m'ont encouragé à me lancer dans la dialectologie des îles que je connaissais le mieux, les îles Torres–Banks. Cela faisait des années que, à la faveur de mes recherches de terrain ou de mes études de cas sur telle ou telle question, j'avais pris l'habitude de cartographier la distribution des isoglosses historiques au nord du Vanuatu. Cependant, j'ai longtemps été freiné, dans la publication de mes découvertes, par mon excessive loyauté envers le *modèle arborescent*.

5.4.1 Problèmes avec le modèle arborescent

Comme la plupart des historiens des langues, je n'imaginais pas le remettre en question. Depuis mes études de reconstruction indo-européenne, je tenais pour acquis que la généalogie des langues ne pouvait être représentée que sous la forme d'un dendrogramme. D'ailleurs, la plupart des diachroniciens du monde, y compris parmi les austronésianistes (ex. Blust, Lynch, Marck, Sagart...) continuent d'en faire amplement usage,¹ sans guère laisser de place au doute. En outre, remettre en question les bases de la linguistique comparée aurait couru le risque de me faire passer pour un adepte des théories du linguiste australien Dixon (1997), à juste titre controversées, selon lequel certaines familles linguistiques seraient incompatibles avec la Méthode comparative des Néo-grammairiens. Je compris plus tard que la mauvaise réputation de cette théorie était due au style particulier de son auteur, et des simplifications parfois grossières qu'il s'autorisait ; mais cela ne devait pas m'empêcher d'avoir mon propre regard critique.

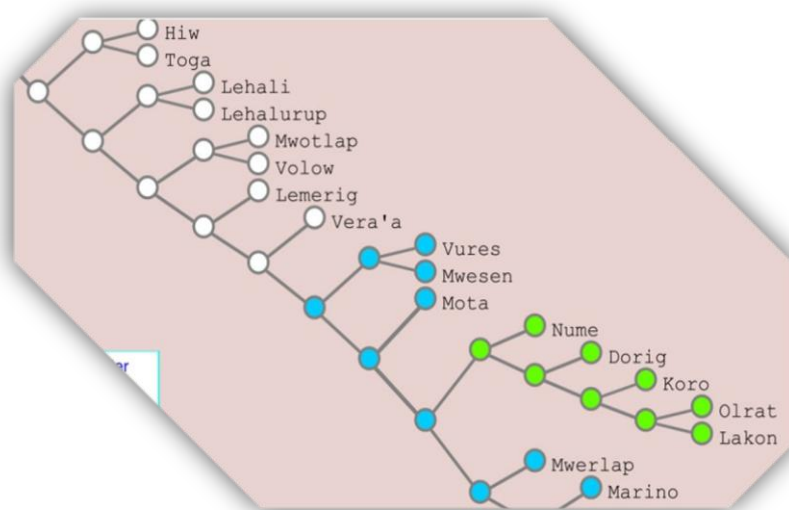
¹ Je n'allais découvrir que plus tard les réserves émises par M. Ross ou P. Geraghty.

Je m'étais donc un temps efforcé de faire entrer mes propres données de terrain dans le cadre d'un arbre généalogique – comme en témoigne, par exemple, cet exposé que je présentai en 2004 au congrès des Océanistes :

François, Alexandre (2004). Subgrouping hypotheses in North Vanuatu. *6th International Conference on Oceanic Linguistics (COOL6)*. University of the South Pacific, Port Vila, Vanuatu.

À cette occasion, je présentai une possible classification généalogique des langues du nord du Vanuatu (*Figure 24*). Je la décrirais comme la moins mauvaise hypothèse que j'aie pu formuler, après plusieurs mois de recherches et de tentatives infructueuses, dans le cadre d'une approche cladistique de la phylogénétique.

Figure 24 – Une proposition hypothétique de classification généalogique pour les langues Torres–Banks (colloque COOL6, 2004)



Mais cet exposé fut également pour moi l'occasion de discuter les problèmes que je rencontrais. Sur le moment, je m'inquiétais d'avoir commis moi-même des erreurs d'analyse, ou de manquer de données, ce qui pouvait expliquer que j'aie du mal à me satisfaire de mes résultats. J'allais plus tard comprendre que mes doutes n'étaient en réalité que la conséquence des présupposés, à mon avis illégitimes, du modèle arborescent. Ainsi, le principe des innovations partagées (principe de Leskien 1876) est censé définir des groupes de langues discrets, proprement emboîtés les uns dans les autres ; mais plus j'examinais mes données de terrain, plus je constatais que **les isoglosses historiques se croisaient constamment**, et empêchaient donc de définir des sous-groupes généalogiques discrets. L'arbre de la *Figure 24* n'était pas malhonnête : je pouvais étayer chaque nœud par un nombre non négligeable d'innovations partagées, ce qui aurait sans doute suffi pour le faire admettre dans une publication.

Si je n'ai pas publié cet arbre, c'est parce que j'étais tiraillé d'un doute lancinant : n'avais-je pas – consciemment ou non – sélectionné les isoglosses qui correspondaient le mieux à mon hypothèse du moment ? Que faire des nombreuses innovations qui se trouvaient à cheval sur plusieurs nœuds, incompatibles avec cet arbre ? (par exemple, une innovation que partageraient le mwesen, le mota et le nume, mais ni le vurès ni le dorig...) Au cas par cas, il était toujours possible de trouver une explication honnête

pour éliminer telle ou telle isoglosse ‘aberrante’, et retomber sur mes pattes : il pouvait s’agir d’un effet du “contact”, d’un “drift”, d’une innovation parallèle, d’un cas de réversion du changement... Mais loin de me rassurer, cette constatation me causait plutôt de la contrariété : il était donc si facile d’élaborer un scénario, en toute bonne foi, qui permît d’éliminer les faits gênants et ne retenir que ceux qui convenaient à un modèle préconçu ? Ce n’était pas là ma conception de la science : les constructions formelles devaient émerger des observations empiriques, sans que l’interprétation des faits ne soit subordonnée à des axiomes définis *a priori*.

Aussi en vins-je, au fil des ans, à remettre en question les présupposés même du modèle arborescent. Prenons, par exemple, le cas des isoglosses que j’appelle ici aberrantes (pour un adepte du dendrogramme), c’est-à-dire à cheval sur plusieurs nœuds sans pouvoir être assignées à aucun nœud ancestral en particulier. Je finis par comprendre que si ces isoglosses sont perçues comme un problème, c’est en vertu d’un préjugé étroitement associé au modèle arborescent : c’est l’idée que, dans un arbre, les branches divergeant d’un nœud correspondent nécessairement à un événement de **division des populations** de locuteurs (par migration, ou d’autres formes de séparation) – car c’est de cette division, et de l’isolement qui en découle, qui expliquerait typiquement les phénomènes de divergence linguistique. Dans ce cas, et seulement dans ce cas, les isoglosses aberrantes peuvent en effet être conçues comme une anomalie : car une fois séparées, les populations ne sont pas censées pouvoir continuer à partager des innovations. Si elles partagent de nouveaux traits linguistiques après leur séparation, alors ce phénomène sera étiqueté “contact” et sera exclu du signal généalogique.

Mais le problème devient très différent dès lors qu’on s’aperçoit que cet axiome, selon lequel la généalogie des langues reposerait fondamentalement sur des événements de séparation et de divergence, n’est qu’un préjugé sans fondement. C’est ainsi que j’en suis venu à une remise en cause radicale du modèle arborescent en généalogie des langues.

5.4.2 Réhabiliter la Théorie des Ondes

Ainsi, si j’avais eu du mal à rendre mes données compatibles avec le modèle arborescent utilisé par tous les diachroniciens, ce n’était pas un problème de mes données, ou de leur interprétation : c’était un vice de conception inhérent au modèle lui-même.

À mesure que je comprenais où se logeait l’erreur, je fus en mesure de cibler mes lectures, et compris que je n’étais pas le premier à formuler de tels doutes. À peine le Néo-grammairien Schleicher avait-il théorisé le *Stammbaum* en 1853, que Johannes Schmidt et Hugo Schuchardt lui avaient opposé leur *Wellentheorie*, ou “Théorie des Ondes” (Schmidt 1872). Bloomfield, en 1933, en avait proposé une synthèse convaincante (1933:311). Bien d’autres critiques de l’arbre ont fleuri çà et là, jusqu’aux années récentes [cf. citations dans AF f/c c ►(#27)]. Cependant, ces objections ont toujours eu du mal à faire le poids face à l’omniprésence du modèle arborescent, au triomphe insolent. La résurgence des études de phylogénétique informatisées ces dernières années ne fait que renforcer le poids de ce modèle, souvent sans discussion de ses présupposés.

En quelques mots, j'affirme que le modèle de l'arbre est incapable de traiter une configuration pourtant on-ne-peut-plus banale : le cas où **les isoglosses s'entrecroisent**. Toutes les études de dialectologie, tous les atlas historiques, montrent que l'intersection d'isoglosses est la situation normale, attendue, par défaut. Pour ne prendre qu'un exemple, le changement phonétique *ka > /ʃa/, censé être une innovation caractéristique du français par opposition à l'ensemble occitan (cf. fr. /ʃato/ vs languedocien /kastel/ 'château'), ne peut en réalité pas être assigné au nœud 'français' dans l'arbre, car le changement est attesté en nord-occitan (limousin /ʃahtew/), et inversement il est ignoré des dialectes oïl du nord tels que normand, picard, wallon (cf. picard /kato/ dans *Cateau-Cambrésis*). La zone délimitée par la palatalisation chevauche donc les autres isoglosses censées définir l'ensemble d'oïl. On pourrait citer des centaines de cas de ce genre, dans les langues indo-européennes ou ailleurs, où les isoglosses historiques se croisent au lieu de former des ensembles discrets [cf. Anttila 1985:305; Kalyan & François f/c ►(#26)].

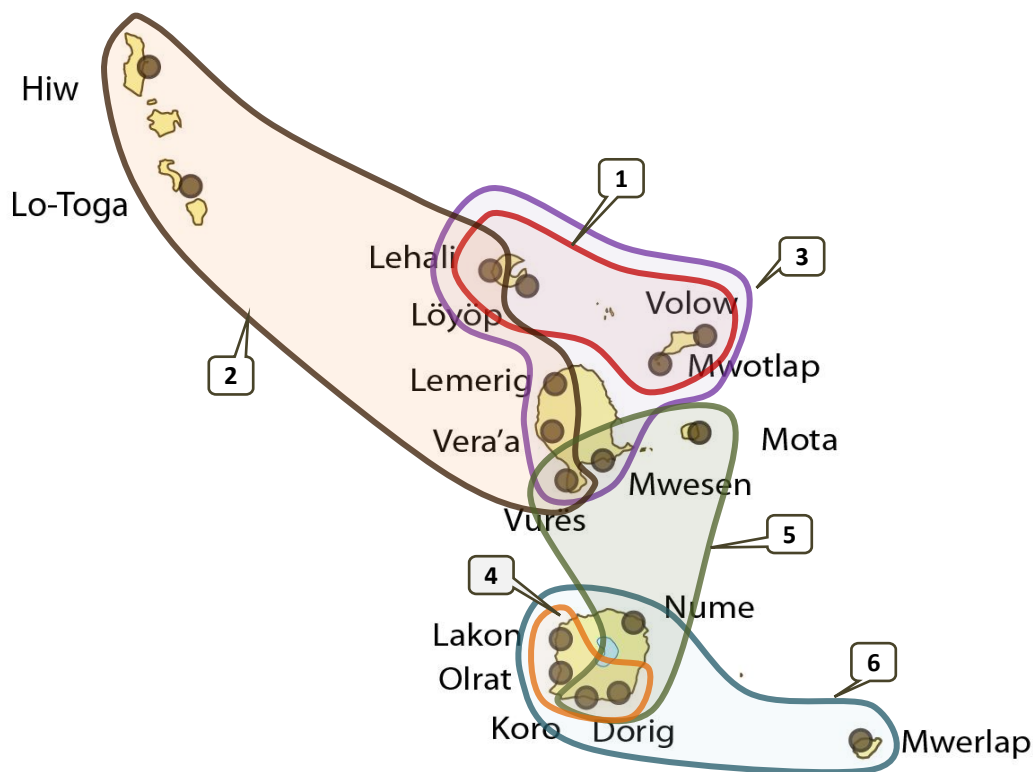
Si la question des intersections d'isoglosses est bien connue des dialectologues – notamment autour des questions de chaînes dialectales ou de continuums – en revanche elle semble peu émouvoir les phylogénéticiens des langues, qui continuent de fonctionner sur les axiomes fictifs du modèle arborescent. L'une des exceptions à ce constat est Malcolm Ross qui, dans sa thèse publiée en 1988 sur les langues océaniques occidentales (Western Oceanic), a introduit la notion de *linkage* (1988:8), que je propose de rendre en français par "chaînage" (ou "chaînage généalogique"). Un chaînage est une famille dont les membres sont mutuellement reliés généalogiquement par une série d'innovations partagées de proche en proche, mais distribuées en chevauchement, de telle manière que cette famille ne puisse être représentée sous la forme d'un dendrogramme.

La situation typique d'un chaînage généalogique peut être illustrée avec quelques exemples venus des langues Torres-Banks. Le *Tableau 14* cite au hasard six innovations partagées par au moins deux langues (pour les abréviations de la colonne de droite, voir la *Figure 3* p.27). Si l'on représente ces innovations sur une carte, on obtient des isoglosses qui s'entrecroisent – voir la *Figure 25* [voir aussi la *Figure 20* p.94].

Tableau 14 – Quelques innovations historiques dans les langues Torres-Banks

#	TYPE DE CHANGEMENT	INNOVATION	LANGUES
1	chang ^t phonétique régulier	*r > [j]	LHI, LYP, VLW, MTP
2	chang ^t phonétique lexical	*late > *lete 'couper en deux'	HIW, LTG, LHI, LMG, VRA, VRS
3	chang ^t morphologique	suffixe possessif *-yi pour 3sg	LHI, LYP, VLW, MTP, LMG, VRA, VRS, MSN
4	chang ^t morphologique	métathèse des pronoms triels *yi ⁿ da-tolu → *tolu-yi ⁿ da '1inc:tri'	DRG, KRO, OLR, LKN
5	chang ^t morphosyntaxique	article *wo pour noms aliénables	VRS, MSN, MTA, NUM, DRG, KRO
6	innovation lexicale	*we ⁿ de remplace *lulu 'blanc'	NUM, DRG, KRO, OLR, LKN, MRL

Figure 25 – Quelques isoglosses historiques dans les langues Torres–Banks
(correspondant aux innovations du Tableau 14)



Une carte comme la Figure 25 représente pour moi à la fois un aboutissement, et un commencement. C'est un aboutissement, parce qu'elle synthétise plusieurs années passées sur le terrain à recueillir des données, à les interpréter en synchronie – à quoi s'ajoute le travail proprement comparatif de reconstruction historique, permettant de reconstruire l'origine commune des formes modernes. Pour ne prendre qu'un exemple, la *métathèse des pronoms triels* dans les langues de Gaua (#4) n'est pas observable telle quelle : comme le montrent les formes de pronoms d'inclusif triel ci-dessous, la métathèse morphologique en question (inversion du pronom *ɣiⁿda 'incl' et du chiffre *tolu 'trois', par rapport à la forme conservatrice *ɣiⁿda-tolu) ne peut être reconstituée qu'après un travail de reconstruction de protoformes, lui-même fondé sur la connaissance des changements phonétiques réguliers dans ces langues. C'est ainsi, par exemple, qu'il devient possible de détecter la présence de l'ancien pronom de triel *ɣiⁿda (< POc *kita) dans la forme /tʃilɾtʃ/ du lakon moderne :

- | | | | |
|---------|-----------------------|--|-----------------------|
| – Nume | <i>ɳdotul</i> | < * ⁿ datólu < *ɣi ⁿ da-tolu | [FORME CONSERVATRICE] |
| – Dorig | <i>tolɣm</i> | < *tolu-ɣi ⁿ da | [MÉTATHÈSE] |
| – Koro | <i>tɪlm</i> | < *tɪl ⁿ d < *tolí ⁿ da < *tolu-ɣi ⁿ da | [MÉTATHÈSE] |
| – Oirat | <i>tʃilɾt</i> | < *tɪlɾtʃ < *tɪl ⁿ d < *tolí ⁿ da < *tolu-ɣi ⁿ da | [MÉTATHÈSE] |
| – Lakon | <i>tʃilɾtʃ</i> | < *tɪlɾtʃ < *tɪl ⁿ d < *tolí ⁿ da < *tolu-ɣi ⁿ da | [MÉTATHÈSE] |

S'il est vrai que ce travail d'identification des innovations partagées repose sur plusieurs années de recherches, il marque également un nouveau point de départ. Que faire des résultats tels que ceux du *Tableau 14* ou de la *Figure 25* ? Si je m'arrêtais à ce stade, ces cartes resteraient anecdotiques ; or j'avais le sentiment que l'on pouvait aller plus loin. C'est ainsi que je me lançai dans l'aventure de la *glottométrie historique*.

5.4.3 La glottométrie historique

Le désir d'aboutir à une synthèse de mes reconstructions historiques a inspiré nombre de mes travaux pendant la période de 3 ½ ans que j'ai passée à l'Université Nationale Australienne (2009-2012). C'est à cette époque que j'ai rédigé et publié mes réflexions sur le paradoxe divergence/convergence, à Vanikoro [§5.2.2] ou au nord du Vanuatu [§5.2.3], ainsi que mon exploration dialectologique de la consonne *R [§5.3].

J'eus la chance, durant cette période, de faire la connaissance de Siva Kalyan, étudiant à l'A.N.U., et ancien élève de l'université de Princeton (New Jersey). Ce qui commença sous la forme de conversations entre collègues au sujet des questions que soulevaient les langues du nord du Vanuatu, s'est vite transformé en un projet de collaboration. Siva était prompt à comprendre mes résultats et mes problèmes – en particulier la question des intersections d'isoglosses. Partageant assez tôt mon insatisfaction avec le modèle arborescent, il mit à profit ses intuitions en modélisation mathématique, et me présenta des propositions de représentation. Il s'ensuivit une collaboration fructueuse de plusieurs semaines en 2011, au cours desquelles nous échangeons des idées pour raffiner nos intuitions réciproques, et définir un modèle qui tînt la route. Comme nous étions inspirés par la *dialectométrie* de l'école de Salzburg (Goebel 2006), je décidai de baptiser notre modèle *glottométrie historique*.

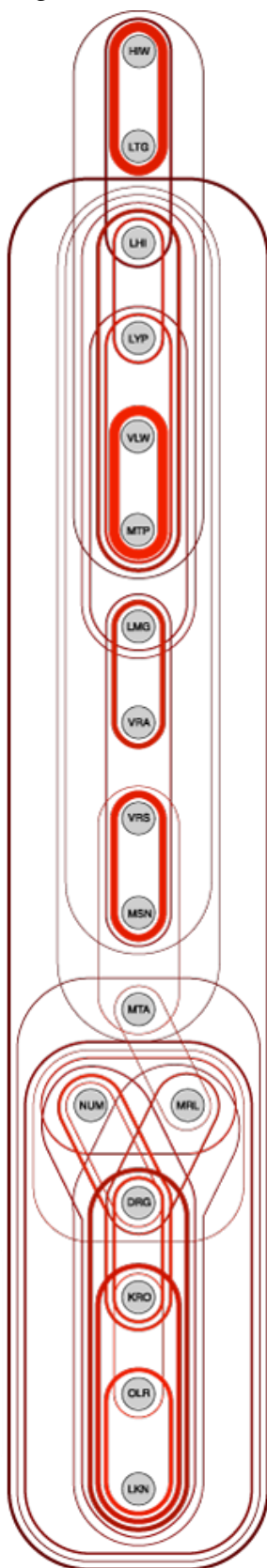
En quelques mots, la glottométrie historique consiste à observer les innovations partagées dans leur distribution géographique (à l'instar de la *Figure 20* p.94, ou de la *Figure 25* p.111). Cependant, au lieu d'observer des isoglosses individuelles, il s'agit de recueillir un grand nombre d'innovations – plusieurs dizaines, voire plusieurs centaines – sur une aire donnée, et d'observer les configurations les plus fréquentes. Si l'on convient d'appeler "sous-groupe généalogique" (pour reprendre le *genetic subgroup* des phylogénéticiens) tout groupement de langues ayant partagé une ou plusieurs innovations, alors il devient possible d'observer quels sous-groupes sont attestés le plus souvent, et lesquels le sont rarement.

Dans cette perspective, j'ai exploité la base de données diachroniques (dans l'esprit du *Tableau 14* p.110) que je m'étais construite au fil des ans, et j'en ai extrait un total de 473 innovations, de divers types : changement phonétique régulier (au niveau du phonème) ou irrégulier (au niveau du lexème), changement morphologique, innovation syntaxique, remplacement lexical. Cette matrice de 473 innovations a servi de base à des calculs visant à observer quelles langues partagent quelles innovations, et partant, quels groupements généalogiques se trouvent le plus souvent attestés.

Nos publications expliquent la nature des calculs. Pour tout sous-groupe, on peut calculer un taux de *cohésion* k_G (en angl. *cohesiveness*) :

$$k_G = \frac{\text{number of supporting innovations}}{\text{total number of relevant innovations}} = \frac{p}{(p + q)}$$

Figure 26 – Diagramme glottométrique des langues Torres-Banks



Le nombre total des innovations “pertinentes” pour le calcul de la cohésion d’un groupe G inclut, d’une part, le nombre p de *supporting innovations*, qui en confirme la validité (innovations qui mettent en jeu l’ensemble des langues de G) ; et d’autre part, le nombre q d’innovations qui le contredisent (innovations mettant en jeu certains membres de G mais pas tous, avec d’autres langues extérieures à G). En d’autres termes, le taux de *cohésion* exprime un degré de confirmation du groupe, comparé au degré de contradiction.

En combinant ce taux de cohésion k_G avec le nombre absolu d’innovations partagées (ε), on peut ensuite estimer la *solidité* (en angl. *subgroupiness*) du sous-groupe ($\zeta = \varepsilon \times k$). Je renvoie aux publications pour davantage de détails.

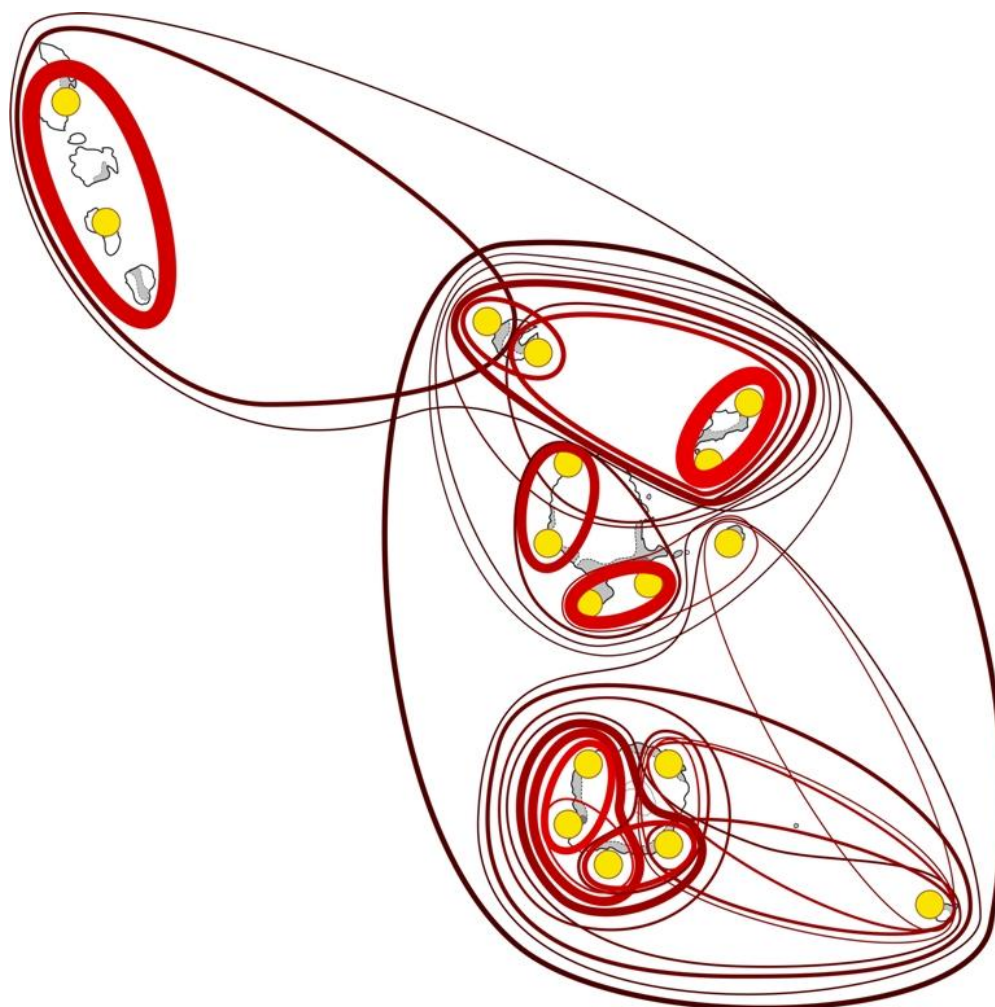
Ainsi, le sous-groupe généalogique formé des six langues des Banks du sud (correspondant à l’isoglosse #6 dans la Figure 25 p.111) est établi sur la base de $\varepsilon = 7$ innovations exclusivement partagées. Par ailleurs son taux de cohésion est $k = 0,43$: ceci signifie que, parmi toutes les innovations pertinentes à ce sous-groupe, 43% ont impliqué ces six langues ensemble, venant ainsi confirmer la validité du sous-groupe en question. Le taux de solidité de ce groupe est donc $\zeta = \varepsilon \times k = 7 \times 0,43 = 3,00$. Ce dernier chiffre n’est interprétable que relativement aux autres groupes de la région. Sachant que le taux de solidité des sous-groupes de la région a une valeur médiane de $\zeta = 0,3$ (avec un maximum de 12,82 pour le groupe mwotlap-volow), le taux de 3,00 est élevé (il est plus élevé que 90% des sous-groupes).

La présentation la plus synthétique de ces résultats glottométriques peut prendre la forme soit d’un diagramme, soit d’une carte. La Figure 26 présente le diagramme glottométrique de la région Torres-Banks. Seuls les sous-groupes ayant une solidité ζ supérieure à 1 sont représentés. L’épaisseur du trait est fonction de la *solidité* du sous-groupe ; le degré de rouge est fonction de sa *cohésion*. Un tel diagramme représente la structure généalogique de cette famille de langues, capable de refléter – bien mieux qu’un arbre – la réalité des processus historiques de diffusion des innovations. Les mêmes résultats peuvent être présentés sous forme de carte (Figure 27).

Cartes et diagrammes glottométriques constituent le type d’aboutissement que je recherchais. Au lieu de constituer des faits anecdotiques séparés, les nombreuses observations que j’avais pu faire sur le terrain, les longs moments passés à reconstruire l’histoire de tel paradigme de pronoms ou tel

processus de grammaticalisation, voilà enfin que j’avais trouvé un moyen – grâce à mes interactions avec Siva Kalyan – de les réunir en une seule synthèse pour leur donner sens. En effet, les résultats de la glottométrie faisaient soudain la lumière sur la **structure des réseaux sociaux d’échange** à époque ancienne (au cours des trois derniers millénaires). S’il est vrai que j’avais pu acquérir des intuitions, sur le terrain, quant au degré de similarité et de différence entre langues voisines, la glottométrie me donnait le moyen de substituer à ces intuitions des faits empiriques, quantifiables, falsifiables et visualisables.

Figure 27 – Carte glottométrique des langues Torres-Banks



Ces diagrammes, par exemple, soulignent la forte proximité généalogique du hiw et du lo-toga d’une part, du mwotlap et du volow d’autre part, ou encore du vurès et du mwesen – points dont les locuteurs ont souvent l’intuition. Ils montrent très clairement l’existence d’une chaîne dialectale tout au long des côtes de Gaua. Ils confirment que les habitants de Mota, en dépit de leur proximité géographique avec Motalava, ont historiquement davantage développé leurs échanges avec les communautés à l’ouest ou au sud. Enfin, ces résultats révèlent l’existence de deux grands sous-groupes, les

Torres et les Banks, mais accordent à la langue lehali un statut généalogiquement hybride (situation qu'un dendrogramme serait incapable de représenter).

Toutes ces observations sont précieuses, parce qu'elles sont porteuses de sens également pour les non-linguistes : archéologues, anthropologues, ethnomusicologues – voire pour les communautés elles-mêmes.

5.4.4 Diffusion de ces résultats

Nous avons présenté nos réflexions à plusieurs occasions :

- A. François, Oct 2011
 “A diffusionist approach to language history in Vanuatu: Clearing off the trees, exploring the bush”. *International Workshop on the Languages of Vanuatu*, Australian National University, Australia, 20-23 Oct 2011.
- A. François & S. Kalyan, Dec 2011
 “Language history in Vanuatu: The epic failure of the tree model”. *Centre for Research on Language Change*, Australian National University, Canberra. [resp. Harold Koch].
- A. François, Apr 2012
 “Linguistique historique au Vanuatu: La défaite du modèle arborescent” – *Méthodes en linguistique historique: Rencontre Austronésianistes–Africanistes*, LLACAN, Villejuif [resp. Laurent Sagart].
- S. Kalyan & A. François, Feb 2013
 “Freeing the Comparative Method from the tree model: A framework for Historical Glottometry”. *International multidisciplinary workshop Let's talk about trees*. Ōsaka, Japan. [resp. R. Kikusawa]
- A. François, Apr 2013
 “Linguistique historique au Vanuatu: La défaite du modèle arborescent” – *Linguistique comparative historique au XXI^e siècle : enjeux théoriques et méthodologiques*, INALCO, Paris [resp. Konstantin Pozdniakov].
- A. François & S. Kalyan, Aug 2013
 “Subgrouping without trees: A glottometric study of the northern Vanuatu linkage” – *21st International Conference on Historical Linguistics (ICHL21)*, University of Oslo, Norway.
- S. Kalyan & A. François, Oct 2013
 “Historical Glottometry: A wave-model approach to subgrouping” – *Australian Linguistic Society (ALS) annual conference*, Melbourne, Australia.
- S. Kalyan & A. François, Feb 2014
 “Historical Glottometry: A wave-model approach to subgrouping” – *Berkeley Linguistic Society (BLS) annual conference*, Berkeley, CA.

Deux publications, actuellement sous presse (acceptées), présentent ces recherches. L'une est un article écrit à deux, et qui discute notamment des aspects techniques du modèle glottométrique que nous proposons :

Kalyan, Siva, & Alexandre François. ss presse. Freeing the Comparative Method from the tree model: A framework for Historical Glottometry. In *Let's talk about trees: Tackling Problems in Representing Phylogenetic Relationships among Languages*, edited by R. Kikusawa & L. A. Reid. Senri Ethnological Studies. Osaka: National Museum of Ethnology. ▷(#26)

La seconde publication – mon chapitre dans un prochain *Handbook of Historical Linguistics* – est davantage consacrée à une discussion théorique sur la nature exacte des relations généalogiques entre les langues :

François, Alexandre. ss presse c. Trees, Waves and Linkages: Models of Language Diversification. In <i>The Routledge Handbook of Historical Linguistics</i> , edited by C. Bown & B. Evans. New York: Routledge.	▷(#27)
--	--------

J'y montre que la différence entre *Tree model* et *Wave model* va bien au-delà d'un simple choix de représentation : en réalité c'est toute la conception du changement linguistique, et de la diversification des langues au cours du temps, qui est en jeu.

Qu'il s'agisse de cette publications ou de nos interventions dans les colloques, nous expliquons toujours d'abord quels sont les problèmes inhérents aux approches cladistiques (arborescentes) de la généalogie des langues, avant de plaider pour une perspective diffusioniste. Un autre point crucial, sur lequel nous prenons soin d'insister (cf. le titre de notre article coécrit), est que notre rejet du modèle de l'arbre n'implique en aucun cas un rejet de la Méthode comparative. Au contraire de Dixon (1997) ou de Schuchardt (1885) lui-même, qui rejetaient en bloc les acquis des Néogrammairiens, je maintiens que la Méthode comparative demeure un outil indispensable à la reconstruction historique – en particulier les notions de régularité du changement phonétique, la nécessité de travailler sur les correspondances régulières entre langues plutôt que sur les ressemblances, ou encore le principe de Leskien (1876) qui établit les groupements généalogiques sur la base des innovations partagées, à l'exclusion des rétentions. Non seulement ces outils restent indispensables, mais c'est même sur eux que je m'appuie pour démontrer l'invalidité du modèle de l'arbre, trop souvent considéré comme indissociable de la Méthode comparative elle-même.

Je compte développer plus avant le modèle de la Glottométrie historique au cours des prochaines années, et créer une dynamique collective sous forme de collaborations, en visant une application à d'autres familles linguistiques [§8.3].

6 Les langues dans leur contexte social et culturel

La section précédente l’a montré : comprendre l’évolution des langues, et les processus historiques de diversification, implique de les replacer dans le contexte de leur *écologie sociale*. Or cette dernière ne se manifeste pas seulement dans le langage, mais aussi dans tout un ensemble de pratiques d’ordre social, culturel, esthétique, auxquelles je voue également un grand intérêt. Plusieurs de mes activités et publications rendent compte de ma motivation à documenter les cultures, les arts, les savoirs collectifs des communautés qui m’ont accueilli.

6.1 Un regard sociolinguistique

En Mélanésie, les sujets locuteurs, agents du changement et de la diffusion, appartiennent à des réseaux sociaux aux dimensions multiples.

Le réseau le plus déterminant dans les pratiques quotidiennes, et le plus saillant à la conscience, est celui que forme la communauté linguistique – typiquement, un territoire contigu incluant un à quatre villages. Mais d’autres réseaux s’y superposent : celui des relations de parenté, qui communément s’étendent sur plusieurs îles ; celui des relations interpersonnelles, réseau dont les limites, selon les individus, ira tantôt en-deçà, tantôt au-delà des frontières de la communauté. À ces réseaux traditionnels, il convient aujourd’hui d’ajouter ceux qui sont nés de la colonisation, et de la construction de l’état moderne du Vanuatu : l’individu appartient à telle ou telle dénomination chrétienne, à tel parti politique, à tel réseau d’amis connus au lycée de la ville...

Or, cette stratification des appartenances sociales multiples engendre des pratiques linguistiques complexes, que j’ai souhaité analyser dans une revue internationale de sociolinguistique :

François, Alexandre. 2012 b. The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214:85–110. ➤(#13)

J’y montre que les pratiques linguistiques au nord du Vanuatu subissent des pressions contradictoires. Premièrement, les attitudes héritées de l’époque précoloniale donnent lieu à une apparente fragmentation des langues – fragmentation discutée dans les pages précédentes, et illustrée dans le *Tableau 13* p.99 : ancré socialement dans sa communauté locale, l’individu aura tendance à aligner ses usages linguistiques sur ceux de ses pairs immédiats, favorisant ainsi des phénomènes de micro-diffusion, et par conséquent, de diversification des langues. L’article suggère que si cette mosaïque linguistique a pu se maintenir jusqu’à nos jours, c’est grâce à un “multilinguisme égalitaire”, idéologie en vertu de laquelle, dans ces sociétés politiquement égalitaires, aucune langue ne jouit de plus de prestige qu’une autre.

Deuxièmement, du fait d'un encouragement généralisé à l'exogamie, ces mêmes locuteurs se trouvent en contact permanent avec des locuteurs d'autres langues, y compris au sein de leur famille. Le multilinguisme est monnaie courante, et l'on passe aisément d'une langue à l'autre. Certes, les *formes des mots*, et autres propriétés de la langue accessibles à la conscience, auront tendance à être associées à telle ou telle communauté en particulier ; leur caractère socialement emblématique les empêchera souvent de passer le filtre du contact [§5.2.2]. Mais il n'en est pas de même des *structures linguistiques* (grammaticales ou lexicales), moins accessibles à la conscience : celles-ci cèdent facilement à la tentation du calque, à la "quête de codes traduisibles mot pour mot" (Gumperz & Wilson 1971: 165). Le résultat que l'on constate est un *isomorphisme structural* généralisé à l'échelle de toute la région.

Enfin, une troisième strate intervient, celle de la construction de l'état moderne du Vanuatu. Ce dernier dispose d'une langue nationale : le **bichelamar** (ou bislama), pidgin/créole¹ à lexifieur anglais, probablement né sur les plantations de sucre du Queensland à la fin du XIX^e s. (Tryon & Charpentier 2004). Pidgin dans les zones rurales, créole dans les villes où il est appris comme langue première, le bichelamar est la langue véhiculaire de tout l'archipel. Or, pour les générations de l'exode rural venues suivre des études ou chercher du travail en ville, c'est la seule langue aujourd'hui. Le résultat est une forte progression du bichelamar aux dépens des langues vernaculaires. Le *Tableau 15* – extrait de mon article, et calculé d'après les données brutes du recensement (VNSO 2009) – révèle, en l'espace d'une décennie, un recul de 10 points dans l'usage des langues vernaculaires (73,1 % de la population en 1999 → 63,2 % en 2009). Par un effet de vases communicants, l'usage du créole a progressé de 10 points (23,3 → 33,7%) durant la même période.

Tableau 15 – Langue principale du foyer familial, par région :
pourcentages comparant les recensements de 1999 et 2009

Province (N au S)	1999			2009		
	Vernac.	Bichel.	autre	Vernac.	Bichel.	autre
TORBA	90,6	8,3	1,1	85,6	13,8	0,6
SANMA	60,1	36,2	3,7	51,1	46,5	2,4
→ dont Luganville	23,8	67,2	9,0	14,0	81,9	4,1
PENAMA	94,1	5,3	0,6	91,8	7,6	0,6
MALAMPA	83,0	16,0	1,0	74,4	24,8	0,8
SHEFA	50,4	39,2	10,4	39,7	53,4	6,9
→ dont Port Vila	31,2	52,4	16,4	22,4	67,8	9,8
TAFEA	95,6	3,6	0,8	91,2	8,0	0,8
<i>National, rural</i>	85,3	13,3	1,4	77,1	21,7	1,2
<i>National, urbain</i>	29,3	56,4	14,3	20,5	70,9	8,6
NATIONAL	73,1	23,3	3,6	63,2	33,7	3,1

¹ Le bislama correspond à ce que Bakker (2008) appelle *pidgincreole*.

Le résultat de ces diverses forces explique la situation particulière que connaît aujourd'hui le Vanuatu : d'un côté, une grande densité en nombre de langues, qui ont su se transmettre au fil des générations jusqu'à ce jour ; mais de l'autre, une menace grandissante de disparition prochaine pour les langues les plus vulnérables.

6.2 Agir pour la préservation des langues

Depuis le premier article que j'avais lu sur les langues du Vanuatu (Tryon & Charpentier 1989), je savais à quoi m'en tenir : nombre des langues de l'archipel étaient menacées d'extinction prochaine.

6.2.1 Donner aux langues leurs premières publications

De fait, avant même d'atteindre l'île de Motalava en 1997, alors que je me trouvais encore à l'étape de Luganville (Santo), j'avais été abordé par un homme âgé du nom de Lele Moli : celui-ci se présentait comme le dernier locuteur de la langue araki. Même si j'allais identifier plus tard cinq ou six autres locuteurs, il fallait se rendre à l'évidence : les jours de l'araki étaient comptés. Aussi ne me fis-je pas prier deux fois pour accepter la proposition de Lele Moli : recueillir ce qu'il était possible de sa langue, avant qu'elle ne tombe tout à fait dans l'oubli. Après quatre semaines d'entretiens quotidiens, suivis quelques mois plus tard d'une visite dans son île d'Araki au sud de Santo, j'avais réuni suffisamment de données pour en écrire une description, incluant grammaire, lexique, textes [AF 2002]. Quelques années plus tard – après le décès de Lele Moli – les familles d'Araki m'encouragèrent à rééditer des textes dans la langue de leurs ancêtres : c'est ainsi que je proposai une nouvelle édition de mon lexique, et un recueil de textes destiné à la lecture :

François, Alexandre. 2008 c. *An Araki-English-French dictionary*, 2e édition. Publication en ligne [<http://alex.francois.free.fr/AF-Araki.htm>].
 — 2008 d. *Sorosoro māñan Raki*. Recueils de contes issus de la tradition orale d'Araki, unilingues araki, à usage local. Illustrations de Sawako François. Édition à tirage limité, chez l'auteur. 32 pages. 2^{ème} édition.

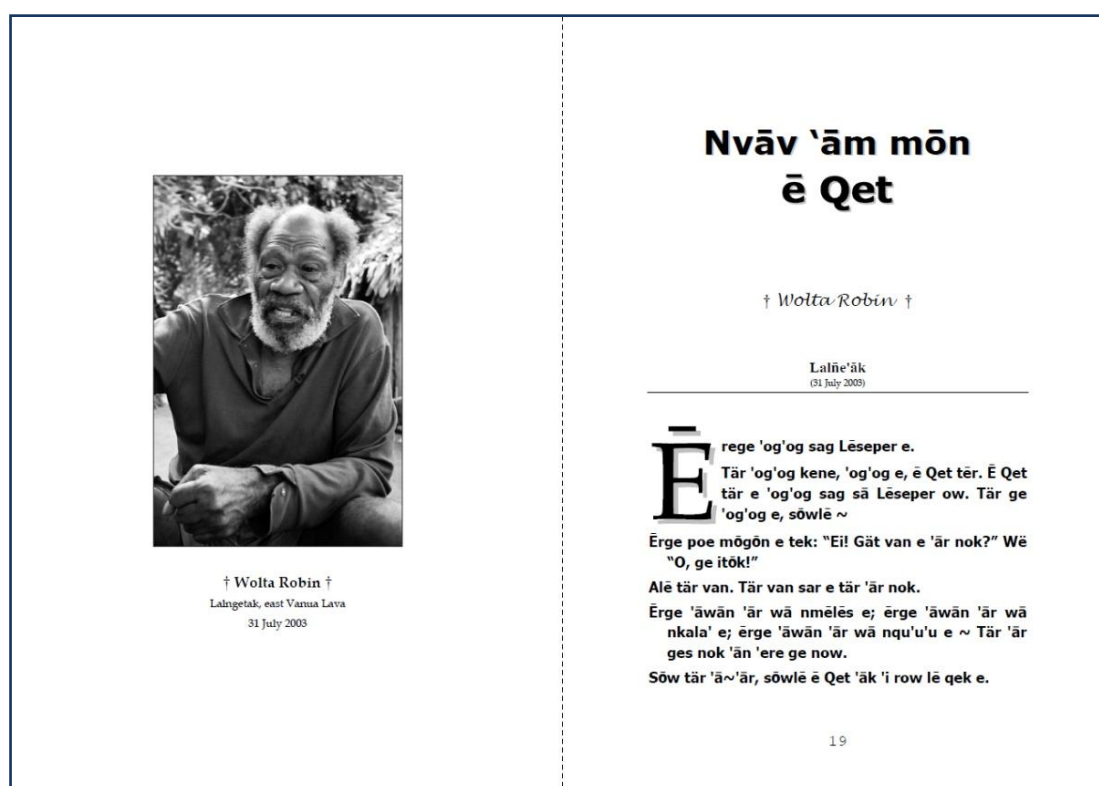
Même si la région Torres-Banks (TORBA) a été relativement épargnée jusqu'à présent par l'expansion récente du créole [Tableau 15], plusieurs de ses langues sont fragiles, voire moribondes. Ainsi, les îles Banks comptent trois langues parlées par une poignée de locuteurs chacune : le mwesen, l'olrat et le lemerig [Figure 3 p.27] ; le volow est déjà éteint. Pour ces trois langues, j'ai créé des recueils de textes (Figure 28) :

François, Alexandre. 2003. *O olñevu ta turmō ta Mēsēn* ['Récits anciens de Mwesen']. Recueils de contes issus de la tradition orale de Mwesen, unilingues mwesen, à usage local. Illustrations de Sawako François. Édition à tirage limité, chez l'auteur. 32 pages.
 — 2006. *Nvāv 'ām 'a Lēmērig* ['Histoires de Lemerig']. Recueil de contes issus de la tradition orale de Lemerig (nord Vanua Lava, Vanuatu), unilingues lemerig, à usage local. Illustrations de Sawako François. Édition à tirage limité, chez l'auteur. 44 pages.

— 2011. *Ususraa pule maraw men Ōlrat* [‘Histoires anciennes d’Olrat’]. Recueil de contes issus de la tradition orale d’Olrat (Gaua, Îles Banks), unlingue en olrat, à usage local. Illustrations de Sawako François. Publié par l’Alliance Française de Port-Vila. 32 pp.

Du fait de la situation de déperdition avancée de ces trois langues, la création de tels ouvrages a surtout une portée symbolique ; mais ce travail de mémoire est particulièrement apprécié des communautés (cf. Thieberger 2002). Plus réaliste, peut-être, est la création de livrets similaires pour les langues qui sont encore activement parlées. Ce qui manque le plus à ces dernières, c’est la pratique de la lecture et de l’écriture : scolarisés dans les écoles françaises ou anglaises, les jeunes n’apprennent pas à écrire leur propre langue, et associent donc l’écriture aux langues coloniales, ou au bislama.

Figure 28 – Extrait du recueil de textes en lemerig, figurant le portrait de †Wolta Robin, le dernier conteur de cette langue



J’ai souhaité participer à la préservation des langues menacées, de plusieurs manières. D’abord, j’ai travaillé à la création d’orthographe adaptées aux phonologies de ces langues, en concertation avec les communautés concernées – impliquant parfois des réunions plénières de tout un village. Ensuite, afin de diffuser les conventions que nous avons choisies, j’ai réalisé des livrets destinés à l’alphabétisation des locuteurs. À l’aide de mon épouse, mais aussi de certains locuteurs passionnés par la préservation de leur patrimoine (je pense à Edgar Woleg Howard), j’ai ainsi créé, au fil des ans, plus d’une vingtaine de livrets dans mes diverses langues d’étude [cf. échantillons (#43)–(#46)].

Ces livrets incluent des abécédaires illustrés, ainsi que des recueils de textes que j'ai recueillis et transcrits sur le terrain.¹ En 2000, j'ai également coécrit avec Edgar Howard un livre en langue mwotlap, destiné aux écoles :

François, Alexandre & Edgar Woleg Howard. 2000. *Bulsal, dam galsi me lèklek* ['Suis-moi bien, ne te perds pas']. Livre de lecture en mwotlap, à destination des écoles. Illustrations de Sawako François. 2^{ème} édition. Alliance Française de Port-Vila. 28 pages.

Certains ont été reproduits par moi-même et distribués sur place ; d'autres ont bénéficié du soutien de l'Alliance Française et de la Communauté Européenne pour être imprimés à plus grande échelle, et distribués dans les écoles des villages. Il m'est arrivé de croiser, dans les îles, des écoliers qui avaient emprunté ces petits ouvrages à leur bibliothèque scolaire – vision qui m'a rempli d'aise. Cependant, je dois bien avouer que je ne trouve pas, sur place, tous les relais que j'espérais. Bien souvent, les freins proviennent des enseignants eux-mêmes, qui donnent la priorité, dans leur enseignement, aux langues coloniales – y compris lorsque les autorités ministérielles, à la capitale, professe son désir de développer l'enseignement des langues vernaculaires.

6.2.2 Archivage et diffusion de mes enregistrements

La préservation des ressources existantes dans mes langues d'études ne passe pas seulement par l'écrit : pour ces langues à tradition orale, il est de bon aloi de vouloir en conserver une trace sonore. C'est ainsi que j'ai entrepris, en 2007, d'archiver mes enregistrements de terrain.

Mes enquêtes m'ont permis de réunir, au fil des ans, plus de 104 heures de matériau sonore, dans 23 langues différentes [Tableau 4 p.32]. La majorité (>85%) se trouve d'ores et déjà archivée sur le site de la *Collection Pangloss* – la plateforme d'archivage des langues rares développée par le LACITO (Michailovsky *et al.*, ss presse).² Il s'agit tantôt de l'enregistrement audio brut, tantôt d'un texte transcrit accompagné de sa traduction [Figure 29]. Fin 2013, j'ai obtenu une bourse de 'valorisation' du LabEx EFL, visant précisément à enrichir le nombre de mes textes présents sur Pangloss, et synchronisés – c-à-d. présentant un alignement phrase à phrase entre l'enregistrement et sa transcription (format intéressant mais long à préparer) ; ces textes devraient être mis en ligne courant mars 2014. Certains récits ont également été publiés sur mon site personnel de popularisation scientifique³.

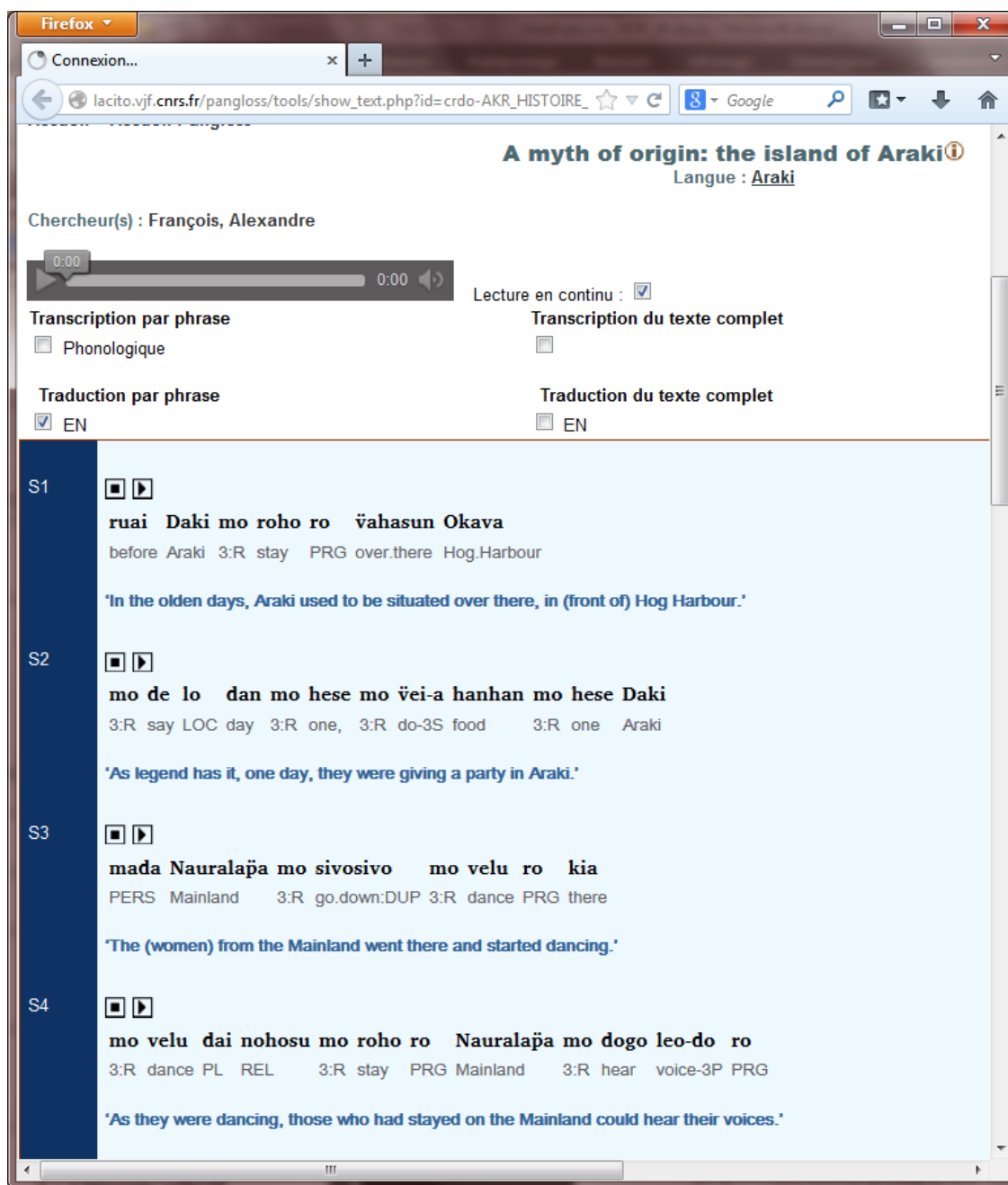
Excepté une poignée de pièces que les locuteurs m'ont demandé de ne pas publier, mes enregistrements sont donc ouvertes au grand public, en libre accès. Ce public inclut de futurs linguistes ou chercheurs, qui souhaiteraient avoir accès aux textes ou aux enregistrements. Bien sûr il inclut également, en principe, les membres des communautés concernées. Cependant, au jour d'aujourd'hui, seuls les locuteurs émigrés, soit à la ville soit à l'étranger, ont accès à l'internet ; les îles du nord sont encore dépourvues d'ordinateurs, et même d'électricité.

¹ Tous ces livres sont accessibles en ligne : [http://alex.francois.free.fr/AF-literacy-f.htm].

² Liens : [lien Pangloss] ou [lien CoCoOn] ou [lien OLAC].

³ Exemple : [http://alex.francois.free.fr/AFtxt_mtp-gecko_fr.htm].

Figure 29 – Un mythe d’origine en langue araki, consultable en libre accès sur le site *Pangloss* consacré aux archives sonores des langues rares



Aussi pris-je la décision, en 2011, de créer dans l’île de Motalava un centre culturel régional multimédia, première au Vanuatu. Avec l’aide de l’Alliance Française de Port-Vila (Georges Cumbo), j’ai offert à cette communauté des îles Banks un ordinateur portable avec son chargeur solaire, ordinateur dans lequel j’avais rassemblé tous mes enregistrements numériques du Vanuatu, accompagnés de leurs métadonnées¹.

¹ Voir les explications : [http://alex.francois.free.fr/AF-audio-library-f.htm].

6.3 Documenter les pratiques esthétiques

6.3.1 La littérature orale

Certes, mes enregistrements de récits issus de la tradition orale constituent un riche corpus linguistique, dans lequel je vais librement quérir les authentiques énoncés qui me permettront d'illustrer telle règle de grammaire, telle hypothèse lexicale. Un texte est aussi un événement langagier, construisant sa référence et sa temporalité, un discours doté de sa structure interne et de sa prosodie – autant de raisons qui incitent le linguiste à s'y intéresser.

Ceci dit, j'ai toujours eu le sentiment que ces récits avaient une signification intrinsèque – une valeur anthropologique, historique, et littéraire, qui méritait mon attention. Il est d'ailleurs notable que les locuteurs eux-mêmes – ou faudrait-il, les conteurs et leurs auditeurs – tiennent en haute estime ces contes, légendes, mythes qu'ils acceptent de me raconter. C'est ainsi que la plupart des livrets que j'ai créés pour la lecture en langues vernaculaires [§6.2.1] sont des recueils de contes et récits traditionnels. Certains d'entre eux sont d'ailleurs de beaux ouvrages illustrés, qui enchantent leur public :

François, Alexandre. 2003. *Tog tog i van en* ['Il était une fois'] Recueil de contes issus de la tradition orale de Mwotlap, unilingue mwotlap, à usage local. Illustrations de Sawako François. Édition à tirage limité, chez l'auteur. 78 pages. [→ <http://bit.ly/Togtogivanen>] ▶(#45)

François, Alexandre. 2012. *Liatevo iepiene ñe piene akapa* ['Histoires anciennes dans nos langues']. Recueil de contes et légendes issus de la tradition orale de Vanikoro (Îles Salomon). Unilingue en teanu, lovono et tanema, à usage local. Illustrations de Sawako François. Édition à tirage limité, 42 pp. [→ <http://bit.ly/Vanikoro>] ▶(#47)

Également sur le thème de la “tradition orale”, figure un chapitre concernant les légendes historiques que j'ai recueillies à Vanikoro. En effet, il se trouve que ma visite dans cette île avait eu lieu à l'instigation de Jean-Christophe Galipaud (IRD-Nouméa), dans le cadre d'une vaste mission archéologique autour de la figure de Lapérouse. En effet, c'est dans cette île isolée du Pacifique que le navigateur français Lapérouse, envoyé de Louis XVI pour émuler le Capitaine Cook, avait fait naufrage en 1788. Or, on savait depuis 1826 que plusieurs survivants de l'événement avaient vécu quelques temps dans l'île de Vanikoro. L'objectif de cette mission pluridisciplinaire, d'ailleurs fortement médiatisée en France, était de rechercher les souvenirs matériels ou immatériels de cette rencontre entre marins français et insulaires.

Dans ce contexte, ma mission de linguiste consistait à interroger la tradition orale. À ma grande surprise, il existe bien une “légende des Français”, transmise de génération en génération, qui relate ces événements lointains. Cabotant de village en village, apprenant la langue teanu au passage, j'ai pu recueillir en tout 14 versions de ce récit. Partout, l'histoire de deux navires pris dans la tempête s'était teintée de merveilleux, et revêtu une dimension épique ; le récit des premiers Blancs arrivés sur l'île en était d'autant plus instructif et passionnant. À la demande des organisateurs de l'expédition

2005, j’ai relaté mes découvertes, et les réflexions qu’elles m’inspiraient, dans un chapitre d’un ouvrage collectif, destiné aux passionnés d’Histoire.

François, Alexandre. 2008 b. Mystère des langues, magie des légendes. In *Le mystère Lapérouse ou le rêve inachevé d’un roi*, edited by A. Salomon. Paris: de Conti – Musée national de la Marine. Pp.230-233.

6.3.2 Musique et poésie

J’avais depuis toujours été attiré par la diversité du monde en matière esthétique et, en particulier, musicale. Ainsi, dès avant mon premier terrain, j’avais suivi un séminaire de l’ethnomusicologue Jean-Michel Beaudet à Paris-X Nanterre, visant à préparer les étudiants à effectuer des enregistrements de la meilleure qualité possible sur le terrain.

Dès mon arrivée dans l’île de Motalava, je fus frappé par sa richesse musicale – qu’il s’agisse de danses coutumières, de chants *a cappella*, de chansons de mariages, ou d’hymnes religieux. En parallèle avec mes recherches linguistiques et mes enregistrements de littérature orale, je ne manquais pas une occasion d’explorer les codes et genres musicaux [cf. *Figure 10* p.70], d’enregistrer tout événement festif dont j’étais témoin, voire d’apprendre moi-même des poèmes chantés. La période de Noël 1997 fut particulièrement riche à cet égard – mais aussi les mois que je passais dans la même île en 1998, 2003, 2004... Mon intérêt pour la musique allait me poursuivre lors de mes explorations d’autres îles du Vanuatu.

Désireux de tirer parti de ce violon d’Ingres, et de faire un sort à mes plus beaux enregistrements, je décidai de monter un projet pluridisciplinaire sur la musique au Vanuatu. C’est ainsi que j’obtins une bourse de 80 000 € du Ministère de la Recherche, dite “ACI (Action Concertée Incitative) Jeunes Chercheurs”, pour la période 2004-2007. J’avais invité à bord deux post-doctorants déjà familiers du Vanuatu : Monika Stern, une jeune ethnomusicologue qui avait consacré sa thèse à l’île de Pentecôte ; et Éric Wittersheim, anthropologue, réalisateur de documentaires autour de la construction politique de l’état moderne du Vanuatu.

Le projet s’intitulait *Rythmes à danser, poèmes à chanter en Mélanésie: Esthétique, transmission et impact social des arts musicaux au Vanuatu*, et visait à découvrir la variété des formes musicales et poétiques dans les îles de l’archipel, largement sous-documentées jusqu’alors. Le rapport final détaillé de ce projet est disponible en ligne :

François, Alexandre, Monika Stern & Éric Wittersheim. 2008. Rythmes à danser, poèmes à chanter en Mélanésie: Esthétique, transmission et impact social des arts musicaux au Vanuatu. Rapport final du projet “ACI Jeunes Chercheurs” (2004-2007) du Ministère de la Recherche. Paris: LACITO, Centre National de la Recherche Scientifique. 55 pp. [\[lien\]](#)

Les fruits de ce projet ont été multiples. Le financement nous a permis de conduire plusieurs missions de terrain, grâce auxquelles nous avons recueilli diverses formes de musique – sans oublier les données linguistiques et littéraires, en particulier poétiques, qui venaient enrichir mon corpus. Notre moisson fut riche, notamment par la découverte d’instruments dont on ignorait l’existence ou que l’on croyait disparus dans la région (membranophone, flûte de bambou, rhombe, arc à bouche) ; par l’observation

des relations complexes entre genres poétiques, répertoires musicaux et styles de danses ; ou encore, par l'étude que j'ai faite du dialecte des chants dans les îles Banks, langue poétique à part entière qu'il fallait déchiffrer pour en mettre à jour toutes les beautés.

Nous avons publié des articles (ex. Stern 2007) et participé à des colloques, donné des exposés à des étudiants en linguistique ou en anthropologie :

- François, Alexandre. 2010. "*Holem taet kastom: A disc project from north Vanuatu*". Communication lors de l'atelier *Study Group for the Music and Dance of Oceania*, ICTM (International Council of Traditional Music). ANU School of Music, Australian National University, Canberra.
- François, Alexandre. 2010. "Song tradition and poetic language in north Vanuatu: Exploring the mystery". Workshop *Language and verbal art*, Dept of Linguistics, RSPAS, Australian National University [resp. Tony Woodbury].
- François, Alexandre. 2011. "Mysterious words: The language of sung poetry in northern Vanuatu". Cours invité dans le séminaire *Languages of the Pacific* (postgraduate). Linguistics, CHL, Australian National University [resp. N. Evans].
- François, Alexandre. 2012. "Les mystères de l'antique: L'art de la poésie chantée au nord du Vanuatu", cours invité dans le séminaire de Master *Tradition Orale en Mélanésie*. Université de Nouvelle-Calédonie [resp. J. Vernaudon].

Mais la véritable récompense de ces années d'exploration a pris la forme de deux objets. D'abord, Éric Wittersheim a réalisé un documentaire autour de la composition d'un chant à Motalava, et autour de nos recherches de terrain :

- Wittersheim, Éric. 2009. *Le Salaire du Poète – The Poet's Salary*. Film documentaire (59'). Paris: CNRS-LACITO. Prix Bartók 2009 du meilleur documentaire ethnomusicologique (*Festival Jean Rouch du film ethnographique*, Paris).

Par la suite, après des années de préparation, Monika Stern et moi-même avons publié en novembre 2013 un album CD, présentant une sélection des 41 pièces musicales les plus significatives parmi nos enregistrements :

- François, Alexandre & Monika Stern. 2013. ***Musiques du Vanuatu: Fêtes et Mystères–Music of Vanuatu: Celebrations and Mysteries***. Disque d'enregistrements de terrain, 1997–2011. Label INÉDIT. Paris: Maison des Cultures du Monde.

En mars 2014, cet album a reçu l'honneur d'un "Coup de coeur" de l'Académie Charles Cros, dans la catégorie *Musiques du monde*.

La publication écrite la plus importante issue de ce projet pluridisciplinaire est d'ailleurs liée à cet album. En effet, il se trouve que le label *Inédit* de la Maison des Cultures du Monde propose une formule intéressante, selon laquelle, d'une part, les pièces musicales peuvent être achetées sur internet, et d'autre part, le livret accompagnant le CD est offert à tous en libre accès. Si le livret imprimé qui accompagne le CD est réduit à une limite de 24 pages, ce n'est pas le cas du livre électronique (ebook) que l'on peut télécharger.

Ce livret, long de 128 pages au total, bilingue français-anglais, comporte non seulement les explications des morceaux du disque, mais inclut en réalité une synthèse détaillée de nos travaux sur la musique et la poésie. Aussi peut-on le considérer, d'une

certaine manière, comme un livre à part entière :

François, Alexandre & Monika Stern. 2013. *Musiques du Vanuatu: Fêtes et Mystères – Music of Vanuatu: Celebrations and Mysteries*. Livre électronique, accompagnant le CD du même nom. Label *Inédit*, W 260147. Paris: Maison des Cultures du Monde. 128 pp. ▷(#31)

La Figure 30 présente un extrait de cette publication : une page évoquant l’art de la poésie chantée. Le livret est reproduit en ▷(#31).

Figure 30 – Extrait du livre *Musiques du Vanuatu: Fêtes et Mystères*

Nombre de chants mettent en scène des chefs d’autrefois, dotés de pouvoirs magiques et de feuilles sacrées, les biceps ornés de prestigieux bracelets en canines de cochons. Ce monde de la “coutume” antique est d’ailleurs cohérent avec les aspects archaïsants de la langue poétique elle-même, qu’il s’agisse de ses sonorités ou de son choix de termes littéraires. Ainsi, les mots “homme” et “femme”, trop prosaïques, sont remplacés dans les chants par des termes rares et anciens, référant originellement aux hommes de haut rang et à leurs épouses – comme si les hommes étaient tous des “sires” (*wegut*) et les femmes des “dames” (*mōter*). De même, les maisons deviennent des “palais” (*gemel*), et la richesse se célèbre en cochons, en cordons de coquillages ou en nattes.

Le chant suivant, dédié par le poète au dignitaire qui le lui a commandé, n’omet aucun détail de l’apparat du Grand Homme dans la société ancienne : bracelets d’ivoire en canines de cochon recourbées, symboles de grand prestige et de richesse, si nombreux qu’ils cliquent sur ses biceps ; piédestal en pierres ; monnaie de coquillages :

*Debout, tu m’écoutes
le cliquetis du pouvoir retentit
le long de ton bras
oui tu fais cliqueter tes bracelets d’ivoire
oui tu fais résonner les pierres de ton piédestal
ta voix est venue jusqu’à moi*

*la foule est unie pour te voir
ces mots te conduiront à la monnaie sacrée*
[Ode de prestige, Motalava]

Ici, la poésie chantée joue le rôle de monument – que les poètes de Motalava désignent comme *namawlōn* “mémorial”. À travers ce chant, le commanditaire du poème se trouve commémoré pour la postérité, tel un prince florentin sublimé par un héroïque portrait.

La fascination pour la nature

Maints poèmes chantent les éléments naturels, avec une attention redoublée lorsque ceux-ci sont déchaînés : déluges, raz-de-marée, cyclones, volcans en éruption, offrent par excellence des thèmes pour le chant.

Certaines de ces évocations littéraires font parfois référence à des événements historiques réels, dont le poète fut lui-même témoin : ainsi, le poème *Le cyclone*, entendu à Gaua en 2003, aurait été composé après le typhon Wendy qui frappa la région en 1972 ; on peut en lire le poème dans l’analyse de la pièce ③. Autre exemple, le chant *Tremble la terre*, recueilli à Hiw en 2007, relate un séisme mémorable qui eut lieu en 1997 :

*C’était la nuit noire dans l’île
Étendu j’étais sur mon lit
Au loin je sens soudain la terre qui vrombit
Grondements sourds partout sur l’île
grondements sourds autour de l’île (...)*
[Tremble la Terre, newēt, Hiw]

Au cours des prochaines années, je souhaite accorder plus de place encore à mes recherches sur la littérature orale. Après avoir étudié la poésie, je souhaite me concentrer désormais sur la riche tradition narrative et mythologique, dont j’ai enregistré bien des joyaux au cours de mes années de terrain. Je travaille notamment sur un recueil plurilingue de récits traditionnels de l’île de Gaua, en cours de finalisation [§8.4].

6.4 Valorisation, popularisation

Nombre des publications mentionnées dans cette section s’adressent à un public ouvert plutôt qu’aux seuls experts. Je ne suis pas convaincu, d’ailleurs, que ce type de dichotomie (expert/profane) soit si pertinente, car on a plutôt affaire, ici comme ailleurs, à un continuum. Tel musicologue lira les pages sur l’organologie des instruments en expert, mais redeviendra profane en lisant la section sur la poésie. Tel anthropologue lira certains récits coutumiers avec un œil de connaisseur, et en appréciera d’autres en tant que simple lecteur. Il en va de même, d’ailleurs, de mes publications scientifiques : mon article de *Phonology* sur la phonotactique du hiw [§3.2.2] s’adresse autant aux phonologues qu’aux océanistes ; mon recueil d’étymologies [§4.3.3] parlera aux spécialistes du changement sémantique, mais aussi aux anthropologues de la Mélanésie, voire à tous les lecteurs curieux des cultures du Pacifique.

Ce refus de segmenter mon auditoire en catégories étanches se manifeste de plusieurs manières. Dans mon style de rédaction, je m’efforce toujours de rendre mon sujet accessible, voire attrayant, à un spectre large de lecteurs potentiels, y compris étudiants et non-initiés. On m’a tantôt cité ce point comme une qualité, tantôt comme un défaut : certains relecteurs s’impatientent lorsque, par souci de pédagogie, je rappelle les principes de la hiérarchie de sonorité, ou celui de la régularité des correspondances phonétiques, réputés déjà connus des lecteurs spécialistes. C’est un équilibre à trouver. Je garde présent à l’esprit le rôle que le chercheur du CNRS joue au sein de la cité – hussard noir d’une République où le savoir doit prévaloir sur les préjugés, et la raison sur l’obscurantisme.

Ces motivations expliquent la place que j’accorde aux activités de “vulgarisation” (ou comme on le dit en anglais, de *popularisation*) des recherches dans mon domaine¹. Je contribue à des expositions temporaires ou permanentes dans des musées. Je participe à des films documentaires, et donne des interviews dans la presse écrite (*le Monde*, *Libération*, *Historia*, *der Spiegel*), radio ou télévisée. J’interviens dans des conférences publiques – généralement à propos des langues en danger ou des cultures du Pacifique. Je diffuse mes connaissances sur différents sites internet : sur Canal-U, Academia.edu², Scribd³, Wikipedia⁴. Mon site personnel de popularisation scientifique⁵ donne libre accès à mes publications, exempliers, archives sonores, photographies de terrain, livrets d’alphabétisation, dictionnaires, textes de tradition orale – documents pour lesquels je reçois régulièrement des réactions positives de la part d’internautes, de collègues, ou de locuteurs des langues en question.

¹ Cf. la section “Valorisation” dans mon CV.

² Statistiques sur Academia.edu (depuis Février 2012): 935 *profile views*, 4852 *document views*.

³ Statistiques sur Scribd.com (depuis Janvier 2012): 4185 *document views*.

⁴ Sur Wikipedia, j’ai effectué 2862 modifications, dont 983 sur WP francophone et 1672 sur WP anglophone (depuis 2006).

⁵ Site: [http://alex.francois.free.fr/]. Statistiques (depuis 2004) : 150 260 *pageloads*, 41 309 visites.

7 Participation à la vie scientifique

7.1 Contributions à la recherche

7.1.1 Organisation d'événements scientifiques

J'ai donné au total 29 communications dans des colloques internationaux : langues océaniques (COOL), langues austronésiennes (ICAL), typologie des langues (ALT), linguistique historique (ICHL), etc. En 2009, j'ai été l'orateur principal (*keynote speaker*) de la Conférence annuelle de l'ALS (Australian Linguistics Society) :

Jul 2009 — "Local words, shared ideas. Lexical divergence and structural homogeneity among north Vanuatu languages", Conférence annuelle de l'ALS (**Australian Linguistics Society**), Melbourne, Australie. [Conférencier invité]

En 2008-2009, en compagnie de Laurent Sagart et Isabelle Bril, j'ai organisé la 11^e édition du *Colloque international des langues austronésiennes*, 11-ICAL. Cet événement scientifique d'importance s'est tenu à Aussois, en juin 2009, et a réuni 156 participants pour 140 communications. Les orateurs invités étaient Robert Blust, Nikolaus Himmelmann, et Aone van Engelenhoven. Entre autres fonctions, j'ai confectionné le livre de résumés (150 pages), créé et maintenu le site officiel du colloque¹.

Un peu plus tard, j'ai organisé (avec l'aide de Sébastien Lacrampe) le premier *International Workshop on the Languages of Vanuatu*, qui s'est tenu à l'Australian National University en octobre 2011. La rencontre a rassemblé 32 chercheurs, réunis pour la première fois pour présenter leurs recherches sur les langues du Vanuatu.

Je mentionnerai plus loin l'École thématique *Méthodes en Linguistique*, que j'organise en juillet 2014 au nom de mon laboratoire LACITO [§8.5].

Par ailleurs, j'ai appartenu aux comités scientifiques (sélection de résumés) de plusieurs colloques internationaux : *International Conference on Oceanic Linguistics* (COOL8, Nv-Calédonie, 2007; COOL9, Newcastle, 2013) ; *Syntax of the World's Languages* (SWL4, Lyon) en 2010 ; *Annual meeting of the Australian Linguistic Society* (ALS 2011) ; *International Conference on Austronesian Linguistics* (ICAL12, Bali) ; etc.

7.1.2 Séminaires

Dès 2003, j'ai activement pris part à plusieurs opérations de recherche inter-laboratoires sous l'égide de la Fédération *Typologie et Universaux Linguistiques* :

- Typologie de la dépendance interpropositionnelle (resp. I. Bril) ;
- Typologie phonologique et changements diachroniques (resp. A. Rialland) ;
- Typologie des pronoms personnels (resp. Y. Moñino) ;

¹ Cf. lien vers une [version archivée du site](#), et du [livret de résumés](#).

- Typologie des rapprochements sémantiques (resp. M. Vanhove).

Certaines de ces opérations ont donné lieu à des ouvrages collectifs (Vanhove 2008, Bril 2010), dans lesquels j’ai contribué.

Durant ma présence à l’A.N.U., j’ai participé activement à plusieurs séminaires (Social Cognition group ; Grammar Writing Group ; Phonetics and Phonology), et j’ai été responsable de l’organisation de certaines saisons du Grammar Writing Group.

Les titres de mes interventions dans les séminaires – quelle qu’en soit la nature ou le lieu – apparaissent dans mon Curriculum Vitae joint à ce dossier.

En décembre 2012, j’ai créé au LACiTO un séminaire de description des langues et de typologie, que j’anime une fois par mois. Intitulé *Problèmes d’Analyse et de Comparaison des Langues*, ce groupe réunit doctorants, postdoctorants et chercheurs confirmés autour d’une présentation. À l’occasion, j’invite des chercheurs étrangers, comme en décembre 2013 où nous avons entendu Paul Heggarty, spécialiste du quechua au Max Planck Institute de Leipzig.

7.1.3 Administration de la recherche

De novembre 2012 à décembre 2013, j’ai assisté régulièrement au Conseil de laboratoire de mon laboratoire le LACiTO, en tant que membre nommé par l’équipe de l’ancien directeur François Jacquesson. En décembre 2013, à l’occasion des élections internes, j’ai été élu au même poste.

Par ailleurs, suite à la fin du mandat de François Jacquesson, la nouvelle directrice Samia Naïm (élue pour 2014-2019), m’a invité à prendre part au comité de direction du laboratoire – comité également composé de Jean-Michel Roynard, Isabelle Bril et Marc-Antoine Mahieu. Nous nous réunissons régulièrement pour prendre les décisions nécessaires à la bonne marche de l’équipe.

7.1.4 Comités de sélection

J’ai évalué des dossiers de candidatures (bourses, postes) à la sollicitation de plusieurs organismes français ou étrangers :

- | | |
|---|--|
| 2003 — Examen de candidature pour “Les systèmes de dénomination : Origine, Diversité, Evolution, Devenir”, programme interdisciplinaire du département SHS du CNRS.
2009 — Examen de candidature pour le <i>National Science Foundation (NSF)</i> , USA.
2009 — Examen de candidatures pour le <i>Endangered Languages Documentation Programme (ELDP)</i> , basé au SOAS, Londres.
2011 — Examen de candidature pour le <i>Wissenschaftskolleg zu Berlin</i> . | |
|---|--|

En 2012, j’ai été également rapporteur au Comité de sélection de l’Université de Polynésie Française, pour un poste en linguistique des langues polynésiennes.

7.2 Activités éditoriales

7.2.1 Comptes rendus

Une contribution à la vie de la recherche prend la forme de comptes rendus sur des ouvrages. J'en ai publié sept [▷(#40)–(#42)], y compris un compte rendu étendu (*review essay*) au sujet de quatre grammaires de Terry Crowley [2007 ▷(#42)].

7.2.2 Comités de lecture, consultance

Je suis régulièrement sollicité par les revues scientifiques pour remplir le rôle de relecteur (comité de lecture anonyme) :

- *Linguistic Typology*
- *Linguistics*
- *Amerindia*
- *Journal of the Polynesian Society*
- *Studies in Language*
- *Language Documentation & Conservation*
- *Journal de la Société des Océanistes*

De même pour certains éditeurs d'ouvrages :

- *Mouton de Gruyter*
- *Pacific Linguistics*
- *Benjamins*
- University of Hawaii (manuel de linguistique)

7.2.3 Édition scientifique

Je suis membre du Comité scientifique éditorial (Editorial advisory board) des revues *Oceanic Linguistics* (U. of Hawaii) et *Open Linguistics* (Mouton de Gruyter).

Par ailleurs, j'ai fondé en 2012, et dirige en tant que responsable éditorial, la collection d'ouvrages scientifiques *Studies in the Languages of Island Melanesia*¹. Il s'agit d'une série d'ouvrages linguistiques (dictionnaires, grammaires, recueils de textes, collections d'articles) avec comité de lecture. Elle appartient à l'éditeur *Pacific Linguistics* – ou plutôt sa version en accès libre, *Asia Pacific Linguistics Open Access*. Notre catalogue compte déjà trois ouvrages :

Stebbins, Tonya (with the assistance of Julius Tayul). 2012. ***Mali (Baining) dictionary***. *Studies in the Languages of Island Melanesia* (SLIM-01), A-PL 001. Canberra: Asia Pacific Linguistics.

Dryer, Matthew. 2013. ***A grammatical description of Kara-Lemakot***. *Studies in the Languages of Island Melanesia* (SLIM-02), A-PL 009. Canberra: Asia Pacific Linguistics.

Lovegren, Jesse ; Alice Mitchell ; Natsuko Nakagawa. (ss presse). ***The Wala language of Malaita, Solomon Islands***. *Studies in the Languages of Island Melanesia*. Canberra: Asia Pacific Linguistics.

¹ Site: [<http://pacling.anu.edu.au/materials/Studies-in-the-Languages-of-Island-Melanesia.htm>].

Il convient d’y ajouter deux ouvrages en préparation, y compris celui que je coédite [§8.1.4] :

François, Alexandre; Sébastien Lacrampe; Stefan Schnell & Michael Franjeh (eds). (en prép). ***The Languages of Vanuatu: Unity and Diversity***. Studies in the Languages of Island Melanesia, 3. Canberra: Asia Pacific Linguistics.

Scoditti, Giancarlo. (en prép.) ***Texts from the Kula area (Papua New Guinea)***. Studies in the Languages of Island Melanesia. Canberra: Asia Pacific Linguistics.

7.3 La formation des étudiants

7.3.1 Enseignements

En parallèle avec ma recherche, j’ai souvent eu l’occasion de transmettre mes connaissances sous la forme de cours, dans diverses universités de la région parisienne :

1999-2002 — Trois années de monitorat en Sciences du Langage à l’**ILPGA** (Institut de Linguistique et de Phonétique Générales et Appliquées), **Université Paris-III Sorbonne Nouvelle** [64 h/an] :

Introduction à la Linguistique Générale (CM 1 h + TD 1 h, deug1 "Découverte") – 26 h, 115 étudiants répartis en deux groupes.

Introduction à une langue du Pacifique : le motlav (cours d’option deug2 + Licence FLE) – 38 h, 45 étudiants.

Français langue étrangère : Expression orale (niveau A, moyens-forts + forts) – 60 h, 25 étudiants.

Français langue étrangère : Phonétique articulatoire du français (niveau B) – 24 h, 35 étudiants.

2002-2003 — Séminaire de Maîtrise : *Tradition orale et mémoire à l’épreuve du terrain*. Maîtrise d’Ethnologie et de Sociologie Comparative, **Université Paris-x Nanterre**. 18 h/an.

2007-2008 — Cours de Master I : *Initiation à l’enquête de terrain*. Master de Langues Océaniques, **INALCO** (Institut des Langues et Civilisations Orientales). 39h/an.

2007-2008 — Cours de Doctorat : Animation du stage doctoral organisé par le **CELIA** (Centre d’Études des Langues des Indiens d’Amérique) [resp. F. Queixalós]. Supervision des séances d’élicitation autour d’un thème : “La prédication”.

Durant mes années de mobilité internationale, j’ai dispensé des cours à l’**Australian National University**, en 2011 :

2011 — *Study of a language family: Oceanic*. Undergraduate & graduate course taught at the **Australian National University**, Semester 2 (July–Nov. 2011). 13 weeks, 39 hours.

La même année 2011, j’ai été professeur invité au célèbre *Summer Institute* de la Linguistic Society of America (LSA) [<http://verbs.colorado.edu/LSA2011>], qui se tenait à Boulder, CO cette année-là. Le thème en était “Language in the World”, et j’ai été sollicité pour mon expertise sur les langues océaniques :

2011 — *Topics in Oceanic languages*. Intensive undergraduate & graduate course taught at the **2011 LSA Summer Institute**, Univ. of Colorado @Boulder. July 2011. 4 weeks, 16 hours.

Ces enseignements m’ont donné l’occasion de fréquenter des étudiants d’horizons différents : étudiants de 1^{er}, 2^e ou 3^e cycle ; ‘undergraduates’ et ‘graduates’ ; linguistes ou anthropologues ; étudiants originaires de France métropolitaine, de Nouvelle-Calédonie ou d’autres pays, étudiants australiens, ou américains...

7.3.2 Interventions invitées

Outre ces cours réguliers, pour lesquels j’étais responsable du groupe, j’ai souvent été sollicité pour des interventions dans les séminaires des collègues (voir la section *guest lectures* dans mon CV). Parfois, ces cours avaient lieu loin de mon domicile, et c’est l’université d’accueil qui m’invitait à donner un exposé dans le séminaire des doctorants :

- | | |
|----------|---|
| Aug 2009 | — “Lexical diversification and structural parallelism in north Vanuatu: Two pressures in conflict” – Linguistic Seminars, University of Newcastle , Australie [resp. Bill Palmer]. |
| Nov 2010 | — “Language divergence and convergence in north Vanuatu: Competing social forces and their linguistic correlates” – Universität Leipzig [resp. Balthasar Bickel]. |
| Mai 2012 | — “Les mystères de l’antique: L’art de la poésie chantée au nord du Vanuatu” – <i>Tradition Orale en Mélanésie</i> . Université de Nouvelle-Calédonie . [resp. J. Vernaudeau] |
| Nov 2013 | — “The ins and outs of <i>up</i> and <i>down</i> : Disentangling the nine geocentric space systems of Torres and Banks languages” – Linguistic seminar, Stockholms Universitet , Suède [resp. Ljuba Veselinova]. |

D’autres interventions de ce type sont prévues pour les mois à venir.

7.3.3 Direction d’étudiants

N’étant pas habilité, je n’ai bien sûr jamais pu diriger officiellement les recherches des étudiants. Cependant, il m’est arrivé plus d’une fois d’encadrer les travaux de certains doctorants dont le sujet relevait de mon expertise.

C’est ainsi que j’ai pu conseiller Agnès Henri, puis Benjamin Touati, deux étudiants inscrits auprès de mon propre directeur de thèse Alain Lemaréchal (Paris-IV Sorbonne), et qui ont fait porter leur thèse sur des langues du Vanuatu – respectivement, le sungwadia et le sakao. Agnès Henri est aujourd’hui maître de conférences à l’INALCO, et je continue de l’encadrer – d’autant plus qu’elle étudie désormais le mwerlap, une langue des îles Banks sur laquelle j’ai moi-même travaillé.

J’ai également eu l’occasion de conseiller ou encadrer plusieurs doctorants au département de Linguistique de l’A.N.U. C’est notamment le cas de :

- Fanny Cottet *The Phonetics of Mbaham, West Papua*
- Tom Honeyman *Discourse in Fas, PNG*
- Yusuf Sawaki *A description of Wooi, West Papua*
- Christian Döhler *A grammar of Kómzo, PNG*
- Maïa Ponsonnet *Semantics of Emotions in Dalabon, northern Australia*

J’ai particulièrement encadré **Sébastien Lacrampe**, dont la thèse porte sur la description du lelepa, une langue du Vanuatu central. J’étais membre de son *supervisory panel*, plus précisément l’un de ses deux *supervisors* : dans le système

australien, il ne s'agit pas du directeur de thèse proprement dit (le *chair*, Mark Donohue), mais d'un rôle de conseiller tout au long du travail de thèse. Outre nos rendez-vous portant sur la thèse elle-même, j'ai eu beaucoup d'autres occasions d'interagir avec S. Lacrampe, et de lui faire profiter de mon expérience. Ainsi, nous avons organisé ensemble les premières rencontres internationales sur les langues du Vanuatu [§7.1.1] ; et nous collaborons aujourd'hui à la préparation d'un ouvrage collectif sur cette région, ouvrage dont nous sommes coéditeurs tous les deux [§8.1.4].

S. Lacrampe n'est pas le seul doctorant pour qui la relation d'encadrement nous a donné l'occasion d'une publication en commun. Ainsi, mes discussions fertiles avec Maïa Ponsonnet, autre doctorante de l'ANU (avec Nick Evans pour *chair*), ont donné naissance à une publication en commun – un article de méthodologie définissant l'approche structurale en linguistique descriptive :

François, Alexandre & Maïa Ponsonnet. 2013. **Descriptive linguistics**. In J. McGee & R. Warms (ed.), *Theory in Social and Cultural Anthropology: An Encyclopedia*, Vol.1, 184-187. SAGE.

(#38)

Enfin, s'il est un étudiant de l'ANU pour qui mon encadrement a porté ses fruits, c'est **Siva Kalyan**. Notre travail en commun ne se situait pas dans le cadre d'un doctorat ; plutôt, c'est lui qui est venu vers moi pour me proposer une collaboration scientifique. Il s'agissait d'un échange : je lui expliquais mes vues sur le changement historique et lui montrait mes données ; et lui se proposait de m'aider à définir un modèle mathématique susceptible d'en rendre compte le mieux possible. De ces conversations de 2011, est née la Glottométrie Historique [§5.4.3], et une relation de collaboration plus que de supervision. Aujourd'hui, S. Kalyan est inscrit en thèse à l'ANU, sous le titre *A quantitative study of linkages in Western Oceanic* : il s'agit d'appliquer notre modèle glottométrique à un ensemble linguistique nouveau, décrit par Malcolm Ross. Même si Bethwyn Evans dirige officiellement cette thèse, j'y remplis le rôle de conseiller externe.

7.3.4 Jury de thèse

En 2012-13, à la demande de l'Université d'Oslo, j'ai (avec Ulrike Mosel) été rapporteur et jury de la thèse de Anders Vaa, intitulée : *A grammar of Engdewu, an Oceanic language of the Solomon Islands*. Mon rôle consistait, d'une part, à rédiger un rapport détaillé sur la thèse ; et d'autre part, de participer à la soutenance (*disputation*) qui s'est tenue en novembre 2013. Ayant le rôle de contradicteur (*opponent*), je devais débattre durant environ 1h30 avec le candidat, conformément à la tradition académique norvégienne.

7.3.5 L'école thématique

Je suis le principal organisateur d'une "école thématique" de linguistique, intitulée *Méthodes en linguistique : Des données empiriques aux hypothèses typologiques*. Mise en place avec l'aide de Samia Naïm et Jean-Michel Roynard, cette école d'été du LACITO se déroulera sur une semaine en juillet 2014 à la station océanographique du CNRS de Roscoff (Finistère). Elle est co-financée par la Formation permanente du CNRS, le LabEx *Empirical Foundations of Linguistics*, et l'INALCO.

Destinée autant aux chercheurs qu'aux étudiants, cette école thématique consistera à expliciter le cheminement intellectuel du linguiste au moment d'interpréter les données recueillies sur le terrain. Nous aborderons une variété de domaines, y compris la phonologie, la morphologie, la syntaxe nominale et verbale, la linguistique historique, la sémantique grammaticale et lexicale. Les linguistes qui interviendront incluent Martin Haspelmath (Max Planck Institute, Leipzig), Denis Creissels (DDL, Lyon), Alexis Michaud (MICA, Hanoi), Guillaume Jacques (CRLAO), Martine Vanhove et Mark van de Velde (LLACAN), ainsi que plusieurs chercheurs du LACITO (S. Naïm, I. Bril, C. Pilot-Raichoor, A. François). La *Figure 31* donne le programme prévisionnel de cette école.

Figure 31 – Programme de l'école thématique *Méthodes en linguistique*

École thématique organisée par le LACITO (CNRS, France)
Comité d'organisation: A. François <francois@vjf.cnrs.fr> — S. Naïm <snaim@vjf.cnrs.fr> — I. Bril <ibril@vjf.cnrs.fr>

Méthodes en linguistique : Des données empiriques aux généralisations typologiques						
7 au 12 Juillet 2014 — Roscoff						
PROGRAMME PROSPECTIF — Janvier 2014						
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi
	<i>TYPOLOGIE ET UNIVERSAUX</i>	<i>PHONOLOGIE</i>	<i>CORPUS ET ARCHIVAGE</i>	<i>SÉMANTIQUE</i>	<i>SYNTAXE-SÉMANTIQUE</i>	
Matin 1 (9h - 10h30)	M. Haspelmath "The relation between typological linguistics and language description"	G. Jacques "Groupes consonantiques: une étude panchronique"	A. Michaud "Collecte de données pour un corpus partageable"	M. Vanhove "Méthodes en typologie lexicale et sémantique"	D. Creissels "Possession prédictive et prédication existentielle"	S. Naïm A. François I. Bril
Matin 2 (10h50 - 12h20)	D. Creissels "Régularités typologiques et diachronie"	A. Michaud "L'analyse des systèmes de tons"	G. Jacques "Génération automatique des paradigmes"	S. Naïm & C. Pilot "Les relations spatiales"	I. Bril "Coordination, subordination and information structure"	Table ronde "Préparer un terrain": Projets des participants
DÉJEUNER						
	<i>MORPHOSYNTAXE</i>	<i>LINGUISTIQUE HISTORIQUE</i>		<i>SYNTAXE-SÉMANTIQUE</i>	<i>SYNTAXE ET DISCOURS</i>	
Après-midi 1 (14h - 15h30)	M. Van de Velde "La possession adnominal: une approche canonique"	A. François "Généalogie des langues et processus de diffusion"	<i>EXCURSION ~ QUARTIER LIBRE</i>	M. Haspelmath "Ditransitive constructions in typological perspective"	M. Vanhove "Syntaxe et prosodie du discours rapporté"	
Après-midi 2 (15h50 - 17h20)	I. Bril "NP conjunction: additive, inclusory and comitative"	C. Pilot-Raichoor "Typologie et linguistique historique"		A. François "La catégorie du nombre et l'aspect verbal"	M. Van de Velde "Comportement grammatical des noms propres"	
SOIRÉE						

7.3.6 Autres formes d'encadrement

Outre les cas d'encadrement doctoral à proprement parler – impliquant relecture de chapitres, conseils aux étudiants, etc. – je trouve d'autres occasions d'échanger mon expérience ou mes connaissances avec les jeunes générations. Par exemple, le séminaire *Problèmes d'Analyse et de Comparaison des Langues* que j'organise [§7.1.2], même s'il est ouvert à tous, est typiquement l'occasion d'échanges avec les doctorants; en particulier, j'ai veillé à ce que chacun des doctorants du LACITO aient l'occasion d'y donner au moins un exposé.

L'école thématique que j'organise en juillet 2014 [§8.5] sera aussi particulièrement tournée vers les étudiants et doctorants désireux de développer un projet de terrain. De même, j'ai souvent l'occasion de contribuer à des séminaires doctoraux, aussi bien au LACITO qu'au LLACAN ou au SEDYL-CELIA, nos laboratoires partenaires au sein du CNRS [§7.3.2]. J'ai l'intention de participer, en juin 2014, au séminaire doctoral du SEDYL.

Enfin, il m'est arrivé d'encadrer, non pas des étudiants, mais de jeunes chercheurs. Je pense particulièrement au projet pluridisciplinaire que je dirigeais de 2004 à 2007, une "ACI Jeunes Chercheurs" portant sur les arts musicaux du Vanuatu [§6.3.2]. Alors que j'étais chercheur statutaire au LACITO-CNRS, j'avais voulu faire bénéficier de cette bourse deux jeunes docteurs alors non statutaires, à savoir l'ethnomusicologue Monika Stern et l'anthropologue Éric Wittersheim. Notre collaboration a donné lieu à de belles réalisations. Mais surtout, elle a permis à ces deux collègues de trouver un emploi de chercheur dans les années qui ont suivi la fin de notre projet : M. Stern est désormais chargée de recherche CNRS au CREDO de Marseille (Centre de Recherche et de Documentation sur l'Océanie) ; et É. Wittersheim est maître de conférences à l'EHESS. J'ai la faiblesse de croire que c'est notre projet commun, entrepris sous mon égide, qui a ouvert à mes deux collègues les portes de ces institutions.

8 Projets

Ces premières années de recherche ne m'ont permis, j'en ai bien conscience, de rendre justice qu'à une partie des données que j'ai recueillies sur le terrain. Il est bien des pages de notes, dans mes carnets manuscrits, qui mériteraient de faire l'objet d'un nouvel article, d'une section à part dans une grammaire, ou d'une belle entrée de dictionnaire. Il faut dire que la région de Mélanésie que j'ai choisie, à travers son goût pour la diversité des langues, offre des trésors inépuisables à quiconque souhaiterait observer le changement phonétique ou sémantique, le renouvellement des systèmes, l'adaptation des structures linguistiques aux diverses pressions en jeu dans le discours.

Mon premier désir, au cours des prochaines années, est de poursuivre – que dis-je, redoubler – mes efforts pour mettre en forme et diffuser mes observations sur les langues de mon terrain. S'il est vrai qu'un tel programme pourrait à lui seul remplir encore des années, je comprends également l'importance de prendre du recul, et de proposer, à partir de mes observations empiriques, des hypothèses plus générales sur le langage. C'est là une démarche que j'ai déjà commencée, notamment à travers mes réflexions sur le changement dans les langues, ou sur les processus historiques de diversification, et la meilleure manière de les modéliser. J'entends poursuivre cet élargissement thématique au cours des années à venir.

À court terme, c'est-à-dire pour l'année 2014, j'ai une idée précise de mes projets :

- terminer l'*Atlas linguistique de Polynésie Française* [§4.2]
- écrire la version finale de l'article sur les directionnels spatiaux [§5.1.4.2]
- mener à bien l'école thématique LACITO de juillet 2014 [§7.3.5]
- intervenir dans l'école doctorale du SEDYL [§7.3.6]
- participer à APLL7 (*Austronesian and Papuan Languages and Linguistics*) en mai 2014, et à SWL4 (*Syntax of the World's Languages*) en septembre [§8.1.2]
- publier l'ouvrage collectif sur le Vanuatu [§8.1.4]
- mettre en place un blog scientifique autour de la Glottométrie historique [§8.3]
- achever le recueil de textes de Gaua [§8.4].

Mais la plupart de mes projets scientifiques s'étendent au-delà de cette année 2014. Ils prennent la forme de publications en cours de préparation, mais aussi de désirs de collaboration ou de futures orientations pour mes recherches.

8.1 Poursuivre la description des langues

8.1.1 Grammaires de référence

Pour les prochaines années, j'ai choisi de me concentrer sur la rédaction de deux monographies grammaticales, que j'ai d'ores et déjà commencées : l'une portant sur le **teanu**, l'autre sur le **hiw**. D'autres descriptions grammaticales sont prévues, sous d'autres formes (cf. §8.1.3).

Le **teanu**, principale langue de l'île de Vanikoro (îles Salomon), présente une grande originalité par comparaison avec les autres langues océaniques [§5.2.2]. Le décrire apporterait beaucoup aussi bien aux études historiques sur cette famille, qu'aux recherches en typologie des langues. Je n'y ai pour l'instant consacré qu'une publication, incluant une esquisse grammaticale pour le **teanu** et même les deux autres langues de l'île [2009 ▷(#20)]. À ce jour, mon manuscrit de la description du **teanu** (grammaire + lexique) compte 172 pages. Pour la partie lexicale, voir l'extrait donné dans la *Figure 11* p.71, ainsi que dans le volume joint [▷(#35)].

François, Alexandre. (en prép.). *A grammar of Teanu, the language of Vanikoro (Solomon Is.)*.

Le **hiw** est également une langue océanique très innovante ; mais même si **teanu** et **hiw** sont 'voisines' géographiquement, les deux langues n'ont aucune similarité l'une avec l'autre. Suite à plusieurs séjours de terrain dans l'île de Hiw, entre 2004 et 2011, j'ai recueilli un corpus conséquent dans cette langue. Tout en lui consacrant plusieurs articles [*Tableau 1* p.15], ainsi que diverses communications dans des colloques, j'ai commencé à en rédiger une grammaire :

François, Alexandre. (en prép.). *A grammar of Hiw, Torres Islands (Vanuatu)*.

▷(#33)

À ce jour, le manuscrit de ma grammaire de **hiw** compte 195 pages ; j'en fournis un extrait (pp.84-121) parmi les documents complémentaires à ce dossier [▷(#33)]. Je souhaite me consacrer à la rédaction de cette monographie dès que mes autres publications me le permettront.

La seule description que j'aie écrite sur le **mwotlap** l'a été en langue française. Hélas, ceci signifie qu'elle est peu accessible à un grand nombre de linguistes, océanistes ou typologues. À terme, j'envisage d'écrire une grammaire de référence de la langue **mwotlap** en langue anglaise. Dans la mesure où cette future grammaire incorporera les résultats de mes recherches sur le Temps-Aspect-Mode, ceci permettrait de renforcer la représentation du domaine TAM, qui ne bénéficie pas toujours de la place qu'il mérite dans les grammaires anglophones des langues du Pacifique.

Pour l'instant, j'envisage mes grammaires dans le format classique du livre. Cependant, féru de nouvelles technologies, je suis à l'affût des progrès techniques en matière de *grammaticographie électronique* (cf. Nordhoff ed. 2012) : je souhaiterais pouvoir, par exemple, lier les sections de ma grammaire aux sections concernées dans mon corpus de textes – en m'inspirant notamment de la grammaire du South Efate par Thieberger (2006).

8.1.2 Études spécifiques

Outre les grammaires, j'ai pour projet d'écrire des études de cas concernant des questions plus spécifiques, sous la forme d'articles dans des revues, ou chapitres dans des ouvrages collectifs. La liste de ces publications est ouverte ; je ne citerai ici que quelques pistes parmi les thèmes que j'ai déjà envisagés :

- *Synchrétisme pronominal et tournures impersonnelles en **teanu** (Vanikoro, îles Salomon)*
Les catégories de pronoms personnels de la langue *teanu* sont classiquement océaniques : trois nombres *singulier–duel–pluriel*, et contrastes de clusivité à la 1^e personne. Cependant le système présente deux originalités. D’une part, la langue possède un pronom *idi* dédié à la référence impersonnelle et générique. D’autre part, le paradigme de préfixes sujets présente, au non-singulier, des syncrétismes typologiquement rares : (a) *1excl=2* ; (b) *1incl=3=impersonnel*. Je propose de décrire la sémantique de ces deux catégories idiosyncrasiques.¹
- *Nombre nominal et pluractionnalité en **hiw** et **lo-toga** (îles Torres, Vanuatu)*
Le *hiw* et le *lo-toga* (îles Torres) associent leurs syntagmes nominaux à référents humains à des contrastes de nombre : *singulier–duel–pluriel*. Plus original dans la famille austronésienne, ces deux langues emploient un supplétisme lexical pour les verbes, pour encoder le nombre de l’actant principal (ex. *hiw* *tō* ‘aller:SG’ vs *vën* ‘aller:PL’). Avec ses 29 paires lexicales basées sur le nombre verbal, le *hiw* présente d’ailleurs un record typologique. Les deux langues des Torres diffèrent dans la liste des paires de verbes concernées, même si elles en partagent une dizaine. Alors que le *lo-toga* est plutôt classique en alignant les référents duels avec la forme *plurielle* du verbe, le *hiw* est plus original car il aligne le duel avec les singuliers. La notion de pluractionnalité en *hiw* est donc basée sur un contraste typologiquement rare entre deux catégories {singulier–duel} vs {pluriel}.
- *Le système possessif du **hiw** (îles Torres, Vanuatu) : l’écroulement des structures*
Le *hiw* garde encore les traces d’un ancien système de possession, reflétant la complexité typique des langues océaniques : contrastes d’inaliénabilité, et parmi les possessions aliénables, contrastes entre les relations génitatives, tels qu’encodés par des classificateurs possessifs (génitif de nourriture, de boisson, de vêtement etc.). Pourtant, ce système conservateur n’existe plus aujourd’hui que dans le registre soutenu de la langue ; le registre familial a massivement simplifié les contrastes, au point d’encoder désormais toutes les relations génitatives à l’aide d’un seul paradigme de formes.
- *Morphologie du marquage de l’objet en **hiw** (îles Torres, Vanuatu)*
En *hiw*, langue accusative à ordre SVO, les propositions transitives présentent une forme *a priori* assez classique de Marquage différentiel de l’objet (DOM) en fonction du trait [±humain]. Alors que les objets [-humain] sont construits directement, les objets sémantiquement [+humain] sont introduits par un morphème /i/. Or ce dernier morphème a un statut morphologique ambigu. Dans la plupart des énoncés, ce /i/ semble se comporter comme une préposition, d’ailleurs homophone avec la préposition /i/ encodant l’oblique ou l’instrumental. Pourtant, certaines propriétés morphologiques, en particulier des règles de sandhi avec certains verbes, incitent à l’analyser comme un enclitique, voire un suffixe verbal : ex. /et^gLakə/ ‘choisir’ + /i/ + /-kə/ ‘2sg’ → /et^gLakə =ti-kə/ ‘te choisir’. En réalité, les tests suggèrent que ce marqueur d’objet doit être analysé tantôt comme une marque de cas morphologiquement orientée à droite (préposition), tantôt comme un enclitique orienté à gauche (applicatif?). La réponse la plus appropriée diffère selon le registre stylistique – probablement le signe qu’une réanalyse de ce morphème est en cours.²

¹ Je présenterai ce travail à *Syntax of the World’s Languages* (SWL4, Pavia, Sept 2014).

² Je présenterai ce travail à *Austronesian and Papuan Languages and Linguistics* (Londres, Mai 2014).

8.1.3 Un ouvrage de dialectologie aréale

Écrire une grammaire de référence offre l'avantage de se concentrer sur une langue en particulier, et d'explorer son architecture interne dans ses moindres détails. Cet exercice intellectuel est à la fois passionnant et indispensable. Or, il se trouve que j'ai recueilli des données sur une vingtaine de langues, qui n'avaient jamais été décrites jusqu'à présent. Est-ce à dire que je devrais envisager de rédiger des grammaires détaillées de chacune de ces langues ? En réalité, ce ne serait ni possible, ni nécessairement souhaitable, d'autant que leurs structures sont souvent proches les unes des autres [cf. §5.2.3.2].

Une solution à ce problème pourrait consister dans la rédaction d'une monographie synthétique – de type *dialectologie aréale* – qui rassemblerait mes observations sur l'ensemble de la région du nord du Vanuatu. La formule exacte reste à définir, et chaque option a des avantages et des inconvénients. Il pourrait s'agir d'une compilation d'esquisses grammaticales taillées selon un modèle unique, à l'instar de l'ouvrage de Ross, Lynch & Crowley (2002) sur les langues océaniques. Ou bien d'une grammaire pan-dialectale similaire à celle d'Evans (2003) pour le Bininj Gun-wok – avec le problème qu'on a quand même affaire à des langues distinctes et non des dialectes. La piste la plus probable serait un volume organisé thématiquement, autour de chapitres tels que *Inventaires phonologiques*, *Règles phonotactiques*, *Catégories lexicales*, *Déixis et orientation spatiale*, *Sérialisation verbale*, etc. ; chaque chapitre inclurait une comparaison des langues Torres–Banks sur tel ou tel point considéré.

Un tel projet de longue haleine – voire de toute une vie – me permettrait de coucher noir sur blanc l'essentiel de mes observations sur ces langues de Mélanésie. Si la description typologique de leurs structures s'accompagne de notes historiques, un tel ouvrage de synthèse pourrait également venir étayer mes travaux de comparatisme (Glottométrie historique) sur la même région [§5.4.3, 8.3].

8.1.4 Ouvrage collectif sur les langues du Vanuatu

Suite à la rencontre que j'ai organisée à l'A.N.U., en novembre 2011, sur les langues du Vanuatu, le projet est né de publier un recueil d'articles, visant à décrire certains points grammaticaux des langues du Vanuatu. Cet ouvrage collectif, dont je suis le responsable (avec trois co-éditeurs), mettra en valeur la diversité linguistique de l'archipel :

François, Alexandre; Sébastien Lacrampe; Stefan Schnell & Michael Franjeh (eds). (en prép).
The Languages of Vanuatu: Unity and Diversity. Studies in the Languages of Island Melanesia,
 4. Canberra: Asia Pacific Linguistics Open Access.

À titre indicatif, la table des matières de cet ouvrage inclut les chapitres suivants :

- (the editors) Introduction: *Unity and Diversity of Vanuatu languages*
- Nick Thieberger: The historical relationship between Erromango and Efate
- Dorothy Jauncey: Oral narratives on Malo
- Murray Garde: Numerals in South Pentecost languages
- Cindy Schneider & Andrew Gray: Language, dialect, variety? the case of Suru Kavian
- Benjamin Touati: The article prefix in Sakao
- Alex François: Space directionals in Torres & Banks languages [→§5.1.4.2]

- Agnès Henri: Clausal complementation in Sungwadia (Maewo)
- Liz Pearce: Completive and terminative in Unua
- Michael Franjeh: The 3sg possessor suffix in North Ambrym
- Stefan Schnell: Emergent subject-TAM markers in Vera'a
- Pete Budd: Valence increasing and instrumental shift in Bierebo

La plupart de ces chapitres se trouvent actuellement dans leur dernière phase, celle de révisions des auteurs après relecture anonyme. Nous espérons publier l'ouvrage avant la fin de l'année 2014, dans la collection *Studies in the Languages of Island Melanesia* [§7.2.3] dont je suis également responsable.

8.2 Recherches en sémantique lexicale

Les prochaines années devraient me permettre de redoubler l'effort lexicographique. Dès 2014 devrait paraître enfin notre *Atlas Linguistique de Polynésie Française* [§4.2], après plusieurs années de travail ; il ne me reste guère que la mise en page à terminer. Mes divers projets de dictionnaires présentés en §4.1.1 progresseront, du moins autant que le leur permettront mes travaux de description grammaticale.

8.2.1 Typologie de la colexification

À plus long terme, j'aimerais me donner l'occasion d'approfondir mes réflexions sur l'étymologie, le changement sémantique et la polysémie. Mon article de 2008 sur la *colexification* [§4.4] s'est trouvé plutôt isolé dans mon parcours, dans la mesure où je n'avais écrit ni d'article précurseur, ni vraiment de suite. Or, mon approche méthodologique a reçu des commentaires positifs dans les comptes rendus (ex. Newman 2010, Traugott 2010, McConvell & Ponsonnet 2013) ; et plusieurs collègues m'ont également contacté ces dernières années (Jacques François, Eugenio Luján, John Newman, David Wilkins) pour me faire part de leur souhait d'adopter cette approche dans leurs propres travaux, ou de la recommander à leurs étudiants. Ces échos positifs que j'ai reçus sur ce travail m'encouragent à aller plus loin : il peut s'agir de reprendre la réflexion méthodologique là où je l'avais laissée, ou encore de l'illustrer avec de nouvelles polysémies – dans l'esprit de la base de données des rapprochements lexicaux présentée par Zalizniak *et al.* (2012).

Une continuation logique consisterait à réunir quelques chercheurs ou étudiants intéressés par ce thème, et superviser des travaux sur ces questions de typologie lexicale. Il serait facilement concevable, par exemple, qu'un doctorant choisisse de consacrer sa thèse à la typologie des colexifications attestées autour d'une notion bien choisie : par exemple, quelles significations les langues du monde associent-elles typiquement au mot *cœur* ? au *chemin* ? à l'idée de *tourner* ? de *tomber* ? d'*ouvrir* ? J'aimerais trouver un moyen de lier plus étroitement ma fascination personnelle pour la lexicographie, l'étymologie, et les travaux de typologie lexicale. Pour ce faire, j'aimerais collaborer au projet de plateforme de dictionnaire électronique que coordonne Martin Haspelmath (Max-Planck Institute), sous le nom *Open-access publication of linked dictionaries for cross-linguistic comparison* ; et voir dans quelle mesure le format électronique permettrait des progrès dans la typologie de la sémantique lexicale.

8.2.2 Polysémies et phénomènes aréaux

Également au cœur de mes intérêts, figure la question de l'emprunt des schèmes de colexification, par contact de langues. Ce phénomène, que l'on peut appeler *calque sémantique* ou *relexification*, est bien connu en créolistique, ou en acquisition de langue seconde. Il s'agit, pour le locuteur d'une langue L1 apprenant une langue L2, du processus cognitif par lequel il aura tendance à calquer les extensions sémantiques de L1 sur la langue L2 (et parfois l'inverse, L2 influençant sa propre pratique de L1).

C'est ainsi, par exemple, que la polysémie de la racine océanienne *roŋo 'entendre, ressentir' s'est trouvée calquée, en pidgin/créole bislama parlé au Vanuatu, dans la forme *harem* (< angl. *hear him*) 'entendre, ressentir' : on a ici un cas, d'ailleurs classique (Camden 1979; Lefebvre 2008, 2009), de mot créole dont le signifiant est emprunté à la langue lexifiante (ici l'anglais) tandis que le signifié est issu du substrat vernaculaire du pidgin. Ce type de convergence dans les structures sémantiques est on ne peut plus fréquent dans les langues du monde, mais trop peu étudié à mon goût. J'aimerais l'explorer davantage – à partir de mes propres données, ou d'autres langues. Une occasion où je pourrai présenter certains faits sera le nouveau programme de la Fédération TUL, intitulé *Pidgins et créoles en contact* (resp. I. Léglise, B. Migge).

Cette réflexion pourrait donner lieu à une publication, et/ou à des travaux collectifs, avec des collègues intéressés. Pour étudier ce thème, je pense me rapprocher notamment de Maria Kotjevskaja-Tamm (Univ. of Stockholm), qui avait organisé un workshop intitulé précisément *Lexicon in contact: contact-induced structural isomorphism in the lexicon* – workshop au sein de la Societas Linguistica Europea, à laquelle je n'avais pas pu me joindre. Cette réflexion constituerait, d'ailleurs, une suite logique au travail que j'ai mené, ces dernières années, sur les phénomènes d'isomorphisme structural entre langues en contact [§5.2.3], y compris au niveau des structures lexicales.¹

8.3 Développer la Glottométrie historique

Dans les années à venir, je continuerai à rédiger de nouvelles études horizontales, visant à comparer les langues du nord du Vanuatu sur tel ou tel point précis : faits de phonologie, structures possessives, syntaxe des actants, etc. Sans même nécessiter de nouveaux terrains, je suis loin d'avoir épuisé les données que recèlent mes cahiers de terrain.

Cependant, dans un avenir plus proche, c'est la *Glottométrie historique* que je souhaite continuer à développer [§5.4.3]. Je prévois d'écrire un nouvel article de revue scientifique avec mon coauteur Siva Kalyan. Ce dernier vient d'ailleurs de s'inscrire pour une seconde thèse à l'Australian National University, afin d'appliquer notre méthode glottométrique aux langues océaniques de l'ouest (Western Oceanic). Même si ses superviseurs sont localement Bethwyn Evans et Simon Greenhill, je continue d'être son consultant externe. En parallèle, Siva et moi-même travaillons à une base de

¹ Le terme de *structural isomorphism* dans le titre du workshop était précisément emprunté à mon article sur la convergence au Vanuatu [2011b:211] ►(#10).

données glottométrique en ligne, ainsi qu'un site de référence (blog scientifique) consacré à la Glottométrie historique.

Ce dernier site, encore en construction, aura pour but de fédérer les collègues linguistes intéressés par l'idée d'appliquer ce modèle non-cladistique sur leurs propres données ; plusieurs d'entre eux se sont déjà manifestés, notamment lors des colloques où nous avons présenté nos recherches. À en juger par les réactions que je reçois à mes travaux, il semble que le domaine de la linguistique historique ressente aujourd'hui un besoin de se renouveler en adoptant des approches non-cladistiques, en lien avec l'essor de l'analyse des réseaux sociaux (*social network analysis*). J'entends contribuer à coordonner cet élan collectif en organisant une rencontre internationale sur cette question – en commençant par un atelier au sein d'une conférence existante. Ceci pourrait se faire à travers l'axe "*Généalogie des langues*" que je coordonne au sein du LabEx EFL [note 1 p.19]. Si l'intérêt dans la Glottométrie historique se confirme au cours des prochains mois, je pense coordonner un nouveau projet – au sein du LabEx ou en dehors – dédié spécifiquement à cette problématique. Un livre édité, ou un numéro spécial de revue, en sera alors le débouché logique.

En parallèle, j'espère former des étudiants ou jeunes chercheurs dans les prochaines années, et leur proposer d'intégrer l'approche glottométrique à leurs recherches en diachronie des langues, sur n'importe quelle famille de langues du monde. Je souhaite conseiller et superviser leurs travaux.

Enfin, je construirai des collaborations multidisciplinaires sur la question des structures des réseaux sociaux. Ainsi, j'ai accepté de participer au consortium interdisciplinaire MiCMAC (*Migrations Contemporaines au Vanuatu : Modélisation et Analyse des Circulations*), en cours de création, qui sera coordonné par la géographe et ethnobotaniste Sophie Caillon (CNRS, Centre d'Écologie Fonctionnelle & Évolutive). Son but est de faire collaborer plusieurs disciplines (géographie, anthropologie, mathématiques et théorie des réseaux sociaux) autour de la reconstruction des circulations et réseaux dans cette partie de la Mélanésie, en particulier aux îles Banks. Je serai ravi d'acquérir une formation plus solide dans la représentation des réseaux sociaux. De mon côté, je pourrai montrer combien la géolinguistique en général, et la Glottométrie historique en particulier, sont capables d'apporter une contribution intéressante à l'histoire des peuples sans écriture.

8.4 Littérature orale

Dans l'immédiat, mon souhait est de mener à son terme un recueil de textes que j'ai commencé à préparer, autour de la tradition orale dans l'île de Gaua. Il s'agit d'une sélection de 21 textes culturellement importants, issus de la tradition orale de Gaua :

François, Alexandre. (ss presse). ***Voices from the Volcano: Myths, Folktales and Poetry from Gaua, Vanuatu***. World Oral Literature Series. Cambridge: Open Book Publishers. ISBN 978-1-906924-82-9. [→ <http://www.openbookpublishers.com/product.php/110>] ►(#34)

Cet ouvrage présentera ces histoires sous forme bilingue (vernaculaire–anglais), dans les six langues parlées dans les îles Gaua, accompagnées de commentaires de nature

littéraire ou anthropologique. Le *Tableau 16* en présente la table des matières ; les titres en gras signalent les textes donnés dans le volume joint ►(#34).

Tableau 16 – Ouvrage en préparation : 21 textes issus de la tradition orale de Gaua

	<i>Titre</i>	<i>Narrateur</i>	<i>langue</i>
1	<i>The Boy who turned into a Flying Fish</i>	Derik Bogo	Koro
2	<i>On the People of Gaua</i>	Rokatson Haris	Mwerlap
3	<i>The Origin of the Great Lake</i>	Nikson Wevalēs	Lakon
4	<i>The Curse of the Eel</i>	† Maten Womal	Olrat
5	<i>The Woman who became Volcano</i>	John Collection	Dorig
6	<i>The Tragic Revenge of the two Sisters</i>	Charles Maklën	Dorig
7	<i>The Legend of Kpwat and his eleven Brothers</i>	Moses Stiven Wëting	Lakon
8	<i>The Bird on the Great Lake</i>	Wini Rovalēs	Lakon
9	<i>The Ogress and her Four Sons</i>	Maten Womal	Olrat
10	<i>Wenagon the Trickster</i>	Bresli Wisiw	Nume
11	<i>The Barracuda Boy</i>	Selwyn Wesur	Dorig
12	<i>The Mysterious Pixies</i>	John Star	Dorig
13	<i>The Orphan from the Volcano</i>	Klera Romalēv	Lakon
14	<i>In Love with an Eel Man</i>	Banabas Womal	Koro
15	<i>The Origin of Dances</i>	Nikson Wevalēs	Lakon
16	<i>The Battle between Gods and Men</i>	Banabas Womal	Koro
17	<i>The Prestige of Hawks</i>	Nelson Vagēl	Lakon
18	<i>The Guardian of the Underworld</i>	Charles Maklën	Dorig
19	<i>A Ship to Queensland</i>	† Maten Womal	Olrat
20	<i>Song of the Great Hurricane</i>		Olrat
21	<i>The Elegy of Lake and Sea</i>		Lakon

J’ai déjà transcrit et traduit la plupart de ces textes ; il me reste à rédiger l’introduction de certains chapitres, et de l’ouvrage dans son ensemble. Le principe de ce livre a d’ores et déjà été accepté par Mark Turin (Univ. of Cambridge), directeur du *World Oral Literature Project* [<http://www.oralliterature.org>], qui l’intégrerait à sa collection ‘World Oral Literature Series’. Les éditeurs universitaires *Open Book Publishers* ont également accepté de publier le manuscrit, et l’annoncent pour cette année 2014.

Si cette expérience est concluante, je pense réaliser d’autres volumes similaires. Ainsi, j’ai d’ores et déjà sélectionné 18 textes enregistrés dans les îles de Vanua Lava, Motalava et Ureparapara (en tout neuf langues représentées), en vue d’un futur volume *Folktales, myths and legends from northern Banks islands*. D’autres volumes pourraient suivre, puisant parmi les 389 récits de tradition orale que j’ai recueillis sur le terrain [*Tableau 4* p.32].

Par ailleurs, un travail que j’ai toujours souhaité de faire, consisterait en une analyse structurale de ces récits. Il n’est sans doute pas raisonnable d’envisager des analyses détaillées pour tous les textes que j’ai. Pourtant, en transcrivant ces centaines de récits, j’ai remarqué nombre de phénomènes intéressants concernant leurs structures narratives, la récurrence de tel ou tel motif, ou encore de nombreuses similarités – avec variations – entre des textes recueillis dans des îles différentes. Ces observations

me font penser aux cours de littérature orale de Dominique Casajus (INALCO), et les différentes études folkloristiques que j’y avais découvertes, de Vladimir Propp à Levi-Strauss, en passant par le catalogue Aarne–Thompson. À terme, j’aimerais faire un sort aux idées que j’ai accumulées autour de mes textes – même si j’ignore encore s’il s’agira d’une communication dans un colloque de linguistes ou d’anthropologues, d’un article ou d’un chapitre de livre, ou simplement d’analyses accompagnant les contes chaque fois que je les publierai.

J’aimerais aussi publier, dès que possible, un article dans une revue d’anthropologie ou de musicologie, à propos de la langue poétique des îles Banks. J’ai particulièrement étudié ce sujet au cours de la préparation de mon livre sur la musique du Vanuatu [cf. *Figure 30* p.126], et il serait de bon aloi que ces réflexions prennent la forme d’un article.

8.5 Vie de la recherche et formation

Au cours des prochaines années, j’ai bien sûr l’intention de poursuivre mes diverses contributions à la vie scientifique, autant à l’échelle de mon laboratoire que sur la scène nationale et internationale. Mes activités d’encadrement de la recherche et des étudiants seront précisément facilitées si je suis habilité à diriger leurs travaux.

8.5.1 Séminaires

Je participe à de nouveaux séminaires de la Fédération TUL à partir de 2014, notamment celui consacré aux pidgins et créoles en contact (resp. I. Léglise) : j’y apporte mes données de première main relatives au bichelamar, le créole national du Vanuatu [§6.1]. Les autres séminaires sont en cours d’ouverture au moment où j’écris.

8.5.2 Enseignements

Il est prévu qu’à la rentrée 2014, j’assure avec Isabelle Bril un enseignement à l’INALCO, portant sur les langues austronésiennes et leur contribution à la typologie des langues. Ce séminaire pour étudiants de master et de doctorat s’intitulera *Structures grammaticales et évolution historique des langues austronésiennes*, et se déroulera sur 50 heures.

À plus long terme, j’aimerais donner un type de cours que j’ai découvert en Australie (A.N.U.) ainsi qu’aux États-Unis (Summer Institute du LSA) : il s’agit de ce qu’on appelle *Field Methods class*, une formation au travail de terrain. L’idée n’est pas de donner de simples conseils aux étudiants, mais d’illustrer par la pratique le travail d’éllicitation, en découvrant une langue nouvelle en compagnie d’un locuteur natif d’une langue inconnue du public, présent dans la salle de classe. J’avais dirigé un séminaire similaire en 2008 pour le Stage doctoral du SEDYL, avec A. Babaliyeva, une locutrice du lezgi. Un tel cours me permettrait de renouer avec le plaisir du terrain tout en découvrant de nouvelles langues, mais aussi de transmettre mon approche de la description linguistique auprès des nouvelles générations de linguistes.

8.5.3 Supervision d'étudiants

Je serai à l'affût de tout étudiant qui souhaiterait me choisir comme directeur de recherches. Parmi les sujets possibles en linguistique, ma préférence scientifique irait à quatre thèmes de recherche en particulier :

- description synchronique du système grammatical d'une langue encore non décrite
- études de sémantique lexicale ; en particulier, exploration typologique d'un domaine particulier de polysémie lexicale, en utilisant des approches du type carte sémantique ou étude de colexification [§4.4, 8.2.1]
- utilisation de la Méthode comparative en linguistique historique, combinée à une modélisation non-cladistique de la généalogie des langues – dans l'esprit de la Glottométrie historique [§5.4, 8.3]
- étude de phénomènes aréaux, impliquant des langues généalogiquement apparentées ou non.

La supervision des thèses d'étudiants prendra logiquement sa place dans mon effort pour promouvoir la linguistique de terrain et de la diversité des langues.

8.6 À plus long terme

Mes voyages au Vanuatu sont actuellement interrompus du fait du moratoire sur la recherche qu'a imposé en 2013 le Centre Culturel du Vanuatu, à l'initiative de son directeur Marcellin Abong. Pour le moment, cette interruption n'affecte pas trop mon travail, car je me concentre sur l'exploitation des corpus que j'ai déjà acquis au cours de mes années de terrain. Mais si ce moratoire devait se prolonger – comme ce fut le cas, au Vanuatu, entre 1984 et 1994 (Taylor & Thieberger 2011) – j'envisagerais alors de renouveler mon ancrage géographique.

En effet, après avoir consacré une vingtaine d'années aux recherches en Océanie, j'aimerais un jour décrire des langues d'une toute autre famille linguistique. Je suis attiré par plusieurs familles linguistiques, aussi bien dans la zone Pacifique (familles papoues, australiennes) qu'ailleurs (Afrique de l'Ouest, Amériques) : j'aimerais étudier une région de forte densité linguistique comme le Vanuatu, mais offrant davantage de diversité génétique. Je choiserais une langue peu décrite, menacée, aux caractéristiques typologiques nouvelles pour moi : il peut s'agir d'une langue à tons, d'une grammaire présentant des structures ergatives, des marques casuelles, des idéophones ou des marques d'évidentiels... Les langues du monde savent se parer de bien des charmes pour éveiller le désir du linguiste.

Il est aujourd'hui trop tôt pour faire un choix, et à ce stade je dois d'abord achever mes projets en cours, qui me retiendront encore plusieurs années – au moins jusqu'à l'horizon 2020. Cependant, je garde à l'esprit ce projet d'élargir un jour mon terrain à une autre région du monde, et à des langues et des cultures nouvelles. Où que j'aille alors dans le monde, je garderai toujours la nostalgie de ces belles années où j'explorais les îles de la Mélanésie.

9 Abréviations

9.1 Langues

Les abréviations de noms de langues apparaissent dans la *Figure 3* p.27, et sont ici répétées.

DRG	dorig	LTG	lo-toga	NUM	nume
HIW	hiw	LYP	lōyōp	OLR	olrat
KRO	koro	MRL	mwerlap	VLW	volow
LHI	lehali	MTA	mota	VRA	vera'a
LKN	lakon	MTP	mwotlap	VRS	vurēs
LMG	lemerig	MSN	mwesen	POC	proto-océanien

9.2 Gloses

Les énoncés sont glosés en fonction des conventions de Leipzig. La liste ci-dessous inclut les gloses plus spécifiques.

ACP	Accompli	IRR	Irrealis
ANAPH	anaphorique	LOC	locatif
AOR	Aoriste	N.SG	non-singulier
ART	article	POSS	(classificateur) possessif
ART:PERS	article personnel	POSS.INAL	possession inaliénable
BkPf	parfait d'arrière-plan (Background perfect)	POT	potentiel
CFT	Contrefactuel	PRF	Parfait
CSTR	<i>construct suffix</i>	PROSP	Prospectif
DEIC	déictique	R	Realis
DU	duel	SBJF	Subjonctif
DUP	réduplication	TOP	topicalisateur
		VTF	ventif (angl. 'hither')

10 Références bibliographiques

- Aikhenvald, Alexandra (2006). Serial verb constructions in a typological perspective. In Robert M.W. Dixon & Alexandra Aikhenvald (eds.), *Serial Verb Constructions: A cross-linguistic typology* 1-87. Oxford: Oxford University Press.
- Ameka, Felix K., Alan Dench, & Nicholas Evans, eds. 2006. *Catching Language: The Standing Challenge of Grammar Writing*. Trends in linguistics: Studies and monographs, 167. Berlin: Walter de Gruyter.
- Andrews, Avery. 1985. The major functions of the noun phrase. In *Language Typology and Syntactic Description*, vol. 1: *Clause structure*, edited by T. Shopen. Cambridge: Cambridge University Press. Pp.62–154.
- Bakker, Dik. 2008. Pidgins versus creoles and pidgincreoles. In *The handbook of Pidgin and Creole studies*, edited by S. Kouwenberg & J.V. Singler. Oxford: Wiley–Blackwell, 130-157.

- Benveniste, Émile. 1966 [1946]. Structure des relations de personne dans le verbe. In *Problèmes de Linguistique Générale*, vol. 1. Tel. Paris: Gallimard. Pp.225-236.
- Bickel, Balthasar. 2008. On the scope of the referential hierarchy in the typology of grammatical relations. In Greville Corbett & Michael Noonan (eds), *Case and grammatical relations: Papers in honor of Bernard Comrie*, 191 - 210. Amsterdam: Benjamins.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Boas, Franz. 1966 [1911]. *Introduction to Handbook of American Indian Languages*. Lincoln: U. of Nebraska Press.
- Bosson, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den Neuiranischen Sprachen*. Tübingen: Narr.
- Bowden, John. 1997. The meaning of directionals in Taba. In Senft (ed.), 251-268.
- Bril, Isabelle. 1999. Mode, temps et aspect en nêlêmwa. *Actances* 10:47-66.
- . 2002. *Le nêlêmwa (Nouvelle-Calédonie): Analyse syntaxique et sémantique*. Edited by A. Bensa & C. Moyse-Faurie. Langues et Cultures du Pacifique, 16. Paris: Peeters.
- . 2003. Quantification, aspect et modalité: phénomènes de portée et d'échelle, quelques exemples en nêlêmwa. In S. Robert (ed.), *Perspectives synchroniques sur la grammaticalisation. Polysémie, transcatégorialité et échelles syntaxiques*. Louvain: Peeters. Pp.53-68.
- . 2004. Complex nuclei in Oceanic languages: contribution to an areal typology. In Bril & Ozanne-Rivierre (eds), 1-48.
- Bril, Isabelle (ed.) 2010. *Clause hierarchy and Clause linking: the Syntax and pragmatics interface*. Amsterdam, New York: Benjamins.
- Bril, Isabelle, & Françoise Ozanne-Rivierre (eds). 2004. *Complex predicates in Oceanic languages: Studies in the dynamics of binding and boundness*. Berlin: Mouton de Gruyter.
- 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.
- Creissels, Denis. 1991. *Description des langues négro-africaines et théorie syntaxique*. Grenoble: Ellug.
- . 1995. *Éléments de syntaxe générale*. Linguistique Nouvelle. Paris: Presses Universitaires de France.
- Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. Longman Linguistics Library. London: Longman.
- . 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Crowley, Terry. 1982. *The Paamese language of Vanuatu*. (Pacific Linguistics, B-87, Canberra: Australian National University.
- . 1987. Serial verbs in Paamese. *Studies in Language* 11:35-84.
- . 1995. Melanesian languages: Do they have a future? *Oceanic Linguistics* 34 (2):327-344.
- . 2000. The language situation in Vanuatu. *Current Issues in Language Planning* 1 (1):47-132.
- . 2002. *Serial Verbs in Oceanic: A Descriptive Typology*. Linguistics. Oxford: Oxford University Press.
- Dixon, R.M.W. 1997. *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.
- . 2010. *Basic Linguistic Theory*. 3 vols. Oxford: Oxford University Press.
- Drinka, Bridget. 2013. Phylogenetic and areal models of Indo-European relatedness: The role of contact in reconstruction. *Journal of Language Contact* 6, 379-410.
- Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and Basic Linguistic Theory. In Ameka, Dench & Evans (eds), 206-234.
- Durie, Mark. 1997. Grammatical structures in verb serialization. In *Complex predicates*, edited by A. Alsina, J. Bresnan & P. Sells. Stanford: CSLI publications. Pp.289-354.
- Dyen, Isidore, Joseph B. Kruskal, & Paul Black. 1992. *An Indo-European classification: A lexico-statistical experiment*. Transactions of the American Philosophical Society. Philadelphia.

- Enfield, Nick. 2006. Heterosemy and the grammar-lexicon trade-off. In Ameka, Dench & Evans (eds), 297-320.
- Evans, Nicholas. 2003. *A Pan-dialectal Grammar of Bininj Gun-Wok (Arnhem Land): Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.
- 2008. Review of *Essentials of language documentation* (Gippert et al. 2006). *Language Documentation & Conservation* 2 (2):340-350.
- 2010. *Dying words: Endangered languages and what they have to tell us*. The Language Library. Malden, MA: Wiley-Blackwell.
- Evans, Nicholas, & Stephen Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32:429-448.
- Evans, Nicholas, & Toshiki Osada. 2005. Mundari: The myth of a language without word classes. *Linguistic Typology* 9 (3):351-390.
- Evans, Nicholas, & David Wilkins. 2000. In the mind's ear: the semantic extensions of perception verbs in Australian languages. *Language* 76 (3):546-592.
- Everett, Dan. 2001. Monolingual field research. In Newman & Ratliff (eds), 166-188.
- Fillmore, Charles, Paul Kay, & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64, 3:501-538.
- Foley, William A., & Mike Olson. 1985. Clausehood and verb serialization. In *Grammar inside and outside the clause. Some approaches to theory from the field*, edited by J. Nichols & A. C. Woodbury. Cambridge: Cambridge University Press. Pp.17-60.
- François, Alexandre → voir liste des publications, sur document séparé.
- Galand, Lionel. 2003. L'aoriste berbère, l'aspect et les valeurs modales. In *Mélanges David Cohen*, edited by J. Lentin & A. Lonnet. Paris: Maisonneuve & Larose. Pp.235-246.
- Garrett, Andrew. 2006. Convergence in the formation of Indo-European subgroups: Phylogeny and chronology. In Peter Forster & Colin Renfrew (eds.), *Phylogenetic methods and the prehistory of languages* 139-151. Cambridge: McDonald Institute for Archaeological Research.
- Geraghty, Paul A. 1983. *The History of the Fijian languages*. Oceanic Linguistics Special Publication, 19. Honolulu: University of Hawaii Press.
- 1990. Proto-Eastern Oceanic *R and its reflexes. In Jeremy H. C. S. Davidson & George Bertram Milner (eds.), *Pacific Island languages: Essays in honour of G. B. Milner* 51-93. London, Honolulu: University of London, School of Oriental and African Studies.
- Gippert, Jost; Nikolaus Himmelmann & Ulrike Mosel. 2006. *Essentials of language documentation*. Studies and Monographs, 178. Berlin: Mouton de Gruyter.
- Goebel, Hans. 2006. Recent Advances in Salzburg Dialectometry. *Literary and Linguistic Computing* 21, 411-435.
- Grace, George W. 1990. The "aberrant" (vs. "exemplary") Melanesian languages. In *Linguistic change and reconstruction methodology*, vol. 45, edited by P. Baldi. Trends in Linguistics: Studies and monographs. Berlin: Mouton de Gruyter. Pp.155-173.
- Gumperz, John J., & Robert Wilson. 1971. Convergence and Creolization: A Case From the Indo-Aryan/ Dravidian Border in India. In *Pidginization and Creolization of Languages*, edited by D. Hymes. Cambridge: Cambridge University Press. Pp.151-168.
- Hale, Kenneth. 2001. Ulwa (Southern Sumu): the beginnings of a language research project. In Newman & Ratliff (eds), 76-101.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The new psychology of language* 211-243. New York: Erlbaum.
- 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3 (1):1-21.
- 2004. How hopeless is genealogical linguistics, and how advanced is areal linguistics? *Studies in Language*, 28 (1):209-223.

- 2007a. Pre-established categories don't exist—consequences for language description and typology. *Linguistic Typology* 11:119-132.
- 2007b. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14 (1):79-102.
- 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86 (3):663-687.
- Haspelmath, Martin, & Uri Tadmor. 2009. *Loanwords in the World's Languages: A comparative handbook*. Berlin: Mouton de Gruyter.
- Himmelman, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36:161-195.
- 2006. Language documentation: What is it and what is it good for? In Gippert, Himmelman & Mosel (eds), 1-30.
- Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, & Dmitry Egorov. 2011. Automated dating of the world's language families based on lexical similarity. *Current Anthropology* 52 (6):841-875.
- Jacobson, Michel. 2004. Corpus oraux en linguistique de terrain. *Traitement automatique des langues (TAL)* 45(2):63-88.
- Keller, Rudi. 1989. Invisible-hand theory and language evolution. *Lingua* 77 (2):113-127.
- 1994. *On language change: the invisible hand in language*. London: Routledge.
- Koptjevskaja-Tamm, Maria, Martine Vanhove, & Peter Koch. 2007. Typological approaches to lexical semantics. *Linguistic Typology* 11-1: 159-185.
- Krishnamurti, Bh. 1998. Regularity of sound change through lexical diffusion: A study of s > h > Ø in Gondi dialects. *Language Variation and Change* 10, 193-220.
- Labov, William. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Le Page, Robert B., & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Lefebvre, Claire. 2008. Relabelling: A Major Process in Language Contact. *Journal of Language Contact*, 91-111.
- 2009. The contribution of relexification, grammaticalisation, and reanalysis to creole genesis and development. *Studies in Language* 33, 277-311.
- Launey, Michel. 1994. *Une grammaire omniprédicative: Essai sur la morphosyntaxe du nahuatl classique*. Sciences du Langage. Paris: CNRS.
- Lazard, Gilbert. 1984. Actance variations and categories of the object. In *Objects. Towards a theory of grammatical relations*, edited by F. Plank. London, New York: Academic Press. Pp.269-292.
- Lemaréchal, Alain. 1989. *Les parties du discours, Syntaxe et sémantique*. Linguistique Nouvelle, Paris: Presses Universitaires de France.
- 1991. *Problèmes de sémantique et de syntaxe en Palau*. Sciences du Langage. Paris: CNRS.
- 1996a. Classificateurs possessifs et relationalité des noms. *SCOLIA* 8, 71-93.
- 1996b. Connexion, dépendance et translation : "Boîte noire" et théorie tesnièreenne de la relation syntaxique. In Gertrud Gréciano & Helmut Schumacher (eds.), *Lucien Tesnière, syntaxe structurale et opérations mentales (Actes du colloque franco-allemand, Strasbourg 1993)* 91-100. Tübingen: Max Niemeyer.
- 1998. *Études de morphologie en f(x,...)*. Bibliothèque de l'Information Grammaticale. Louvain-Paris: Peeters.
- 2010. *Comparative Grammar and Typology. Essays on the Historical Grammar of the Austronesian Languages*. Orbis Supplementa, 35. Louvain: Peeters.
- Leskien, August. 1876. *Die Declination im Slawisch-Litauischen und Germanischen*. Leipzig: Hirzel.
- Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International. [http://www.ethnologue.com]

- Lichtenberk, Frantisek. 1991. On the gradualness of grammaticalization. In *Approaches to grammaticalization*, vol. 1, edited by E. C. Traugott & B. Heine. Philadelphia: Benjamins. Pp.37-80.
- . 2000. Inclusory pronominals. *Oceanic Linguistics* 39 (1):1-32.
- Lynch, John. 2000. Linguistic subgrouping in Vanuatu and New Caledonia. In Bill Palmer & Paul A. Geraghty (eds.), *Proceedings of the Second International Conference on Oceanic Linguistics (SICOL)*, vol. 2: *Historical and descriptive studies* 155-184. Canberra: University of the South Pacific.
- Lynch, John; Ross, Malcolm & Crowley, Terry. 2002. *The Oceanic languages*. Richmond: Curzon.
- Malchukov, Andrej, Martin Haspelmath, & Bernard Comrie. 2010. Ditransitive constructions: A typological overview. In *Studies in Ditransitive constructions. A comparative handbook*, edited by A. Malchukov, M. Haspelmath & B. Comrie. Berlin: Mouton de Gruyter. Pp.1-64.
- McCarthy, John. 1989. Linear order in Phonological representation. *Linguistic Inquiry* 20.1:71-99.
- McConvell, Patrick & Maïa Ponsonnet. 2013. Results and Prospects in the Study of Semantic Change: A Review of From Polysemy to Semantic Change (2008). *Journal of Language Contact* 6(1):180-196.
- Mettouchi, Amina. 1998. Aspect et négation: Remarques sur l'inaccompli et la négation en anglais et en berbère (kabyle). *Cahiers Chronos* 2:191-205.
- . 2009. Mood and Modality in Berber. In *Modals in the Languages of Europe: A reference work*, edited by B. Hansen & F. de Haan. Berlin–New York: Mouton de Gruyter. Pp.431-456.
- Michailovsky, Boyd, Martine Mazaudon, Alexis Michaud, Séverine Guillaume, Alexandre François, & Evangelia Adamou. (ss presse). Documenting and Researching Endangered Languages: The Pangloss Collection. *Language Documentation & Conservation*. U. of Hawaii.
- Milroy, James & Milroy, Lesley. 1985. Linguistic change, social network and speaker innovation. *Journal of linguistics* 21, 339-384.
- Milroy, Lesley. 1987. *Language and social networks*. (Language in Society, Oxford: Blackwell.
- MNCC. 2012. Alternative Indicators of Well-being for Melanesia. Vanuatu Pilot Study Report. Port Vila, Vanuatu: Malvatumauri National Council of Chiefs.
- Moyse-Faurie, Claire. 1998. Relations actanciellles et aspect en drehu et en xârâcùù (Nouvelle-Calédonie). *Actances* 9:135-145.
- . 2001. L'aspect "transitionnel" du faka'uvea (wallisien). *Actances* 9:135-145.
- Newman, John. 2010. From polysemy to semantic change: Towards a typology of lexical semantic associations (review). *Language* 86 (2):473-475.
- Newman, Paul, & Martha Ratliff, eds. 2001. *Linguistic fieldwork*. Cambridge: Cambridge University Press.
- Nordhoff, Sebastian, ed. 2012. *Electronic Grammaticography*. Language Documentation & Conservation Special Publication, 4. Honolulu: University of Hawai'i Press.
- Nordlinger, Rachel, & Louisa Sadler. 2004. Nominal tense in crosslinguistic perspective. *Language* 80:776-806.
- Ozanne-Rivierre, Françoise. 1992. The Proto-Oceanic consonantal system and the languages of New Caledonia. *Oceanic Linguistics* 31 (2):191-207.
- . 1995. Structural Changes in the Languages of Northern New Caledonia. *Oceanic Linguistics* 34 (1):45-72.
- . 1997a. Systèmes d'orientation: quelques exemples austronésiens. In *Diversité des langues et représentations cognitives*, edited by C. Fuchs & S. Robert. Paris/Gap: Ophrys. Pp.81-92.
- . 1997b. Spatial references in New Caledonian languages. In Senft (ed.), 84-100.
- . 2004. Spatial deixis in Iaai (Loyalty Islands). In *Spatial deixis in Oceanic languages*, edited by G. Senft. Pacific Linguistics. Canberra: Australian National University. Pp.127-137.
- Palmer, Bill. 2002. Absolute spatial reference and the grammaticalisation of perceptually salient phenomena. In G. Bennardo (ed.), *Representing space in Oceania: Culture in language in mind*. Pacific Linguistics 523. Canberra: Australian National University. Pp.107-157.

- Pawley, Andrew. 1993. A language which defies description by ordinary means. In *The role of theory in language description*, vol. 69, edited by W. Foley. Trends in Linguistics. Berlin: Mouton de Gruyter. Pp.87-129.
- 2006. Explaining the aberrant Austronesian languages of Southeast Melanesia: 150 years of debate. *Journal of the Polynesian Society* 115 (3):213-256.
- 2007. 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. Communication au colloque *Migrations*, Sept. 2007, Porquerolles (Var, France). 38 pp. [<http://www.vjf.cnrs.fr/lacito/colloque/diaporamas/Pawley.pdf>]
- Pulgram, Ernst. 1961. The nature and use of proto-languages. *Lingua* 10, 18-37.
- Robert, Stéphane. 1991. *Approche énonciative du système verbal: Le cas du Wolof*. Edited by S. Auroux. Sciences du Langage. Paris: CNRS.
- 1995. Aoristique et mode subordonnatif : liens entre aspect et prédication. In *Langues et langage. Problèmes et raisonnement en linguistique (Mélanges offerts à Antoine Culioli)*, edited by J. Bouscaren, J.-J. Franckel & S. Robert. Linguistique Nouvelle. Paris: Presses Universitaires de France. Pp.373-389.
- 1996. Aspect zéro et dépendance situationnelle: L'exemple du Wolof. In *Dépendance et intégration syntaxique (Subordination, coordination, connexion)*, edited by C. Müller. Linguistische Arbeiten. Tübingen: Niemeyer. Pp.153-161.
- 1997. Variations des représentations linguistiques : des unités à l'énoncé. In *Diversité des langues et représentations cognitives*, edited by C. Fuchs & S. Robert. Paris/Gap: Ophrys. Pp.25-39.
- 2003. L'épaisseur du langage et la linéarité de l'énoncé : vers un modèle énonciatif de production. In *Parcours énonciatifs et parcours interprétatifs: Théories et applications*, edited by A. Ouattara. HDL. Paris/Gap: Ophrys. Pp.255-274.
- 2010. Clause chaining and conjugations in Wolof. In Bril (ed.), 469-498.
- Ross, Malcolm. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In *The Comparative Method Reviewed: Regularity and irregularity in language change*, edited by M. Durie & M. Ross. Oxford: Oxford University Press. Pp.180-217.
- 1997. Social networks and kinds of speech-community event. In Roger Blench & Matthew Spriggs (eds.), *Archaeology and language 1: Theoretical and methodological orientations* 209-261. London: Routledge.
- 2001. Contact-induced change in Oceanic languages in North-West Melanesia. In *Areal diffusion and genetic inheritance: problems in comparative linguistics*, edited by A. Aikhenvald & R. M. W. Dixon. Oxford: Oxford University Press. Pp.134-166.
- Ross, Malcolm, & Åshild Næss. 2007. An Oceanic Origin for Äiwoo, the Language of the Reef Islands? *Oceanic Linguistics* 46 (2):456-498.
- Sahlins, Marshall D. 1963. Poor Man, Rich Man. Big Man, Chief: Political types in Melanesia and Polynesia. *Comparative Studies in Society and History* 5 (3):285-303.
- Sasse, Hans-Jürgen. 1993. Syntactic categories and subcategories. In *Syntax: An International Handbook of Contemporary Research*, edited by J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann. Pp.646-686.
- Schleicher, August. 1853. Die ersten Spaltungen des indogermanischen Urvolkes. *Allgemeine Monatsschrift für Wissenschaft und Literatur* 1853, 786-787.
- Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Hermann Böhlau.
- Schuchardt, Hugo. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Oppenheim.
- Senft, Gunter, ed. 1997. *Referring to space. Studies in Austronesian and Papuan languages*. Oxford Studies in Anthropological Linguistics. Oxford: Oxford University Press.
- Silverstein, Michael. 1976. Hierarchy of Features and Ergativity. In R.M.W. Dixon (ed.), *Grammatical Categories in Australian Languages*, 112-171. New Jersey: Humanities Press.

- Stern, Monika. 2007. Les identités musicales multiples au Vanuatu. In *Identités musicales*, vol. 20. Cahiers d'ethnomusicologie. Genève: Adem. Pp.165-190.
- Taine-Cheikh, Catherine. 2010. The role of the Berber deictic and TAM markers in dependent clauses in Zenaga. In Bril (ed.), 355–398.
- Taylor, John, & Nicholas Thieberger, eds. 2011. *Working Together in Vanuatu: Research Histories, Collaborations, Projects and Reflections*. Canberra: ANU Press.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Thieberger, Nicholas. 2002. Extinction in whose terms? Which parts of a language constitute a target for language maintenance programmes? In D. Bradley & M. Bradley (eds), *Language endangerment and language maintenance*. New York: Routledge. Pp.310-328.
- . 2006. *A grammar of South Efate: an Oceanic language of Vanuatu*. Oceanic Linguistics Special Publication No. 33. Honolulu: University of Hawai'i Press.
- Thieberger, Nicholas, ed. 2012. *The Oxford Handbook of Linguistic Fieldwork*. Oxford: Oxford University Press.
- Thomason, Sarah G. 2007. Language contact and deliberate change. *Journal of Language contact* Thema 1, 41-62.
- Thurston, William R. 1989. How exoteric languages build a lexicon: Esoterogeny in Western New Britain. In *VICAL 1: Oceanic Languages, Papers from the Fifth International Conference on Austronesian Linguistics, part 2*, edited by R. Harlow & R. Hooper. Auckland: Linguistic Society of New Zealand. Pp.555–579.
- Tournadre, Nicolas. 2004. Typologie des aspects verbaux et intégration à une théorie du TAM. *Bulletin de la Société de Linguistique de Paris* 99 (1):7-68.
- Traugott, Elizabeth. 2010. Review of Martine Vanhove (ed.), *From polysemy to semantic change: Towards a typology of lexical semantic associations*. *Linguistic Typology* 14 (2010):292-299.
- Tryon, Darrell. 1976. *New Hebrides Languages: An internal classification*. Pacific Linguistics, C-50. Canberra: Australian National University.
- . 2002. Buma. In *The Oceanic languages*, edited by J. Lynch, M. Ross & T. Crowley. Curzon Language Family Series 1. London: Curzon. Pp.573-586.
- . 2007. Language contact in the languages of the Te Motu Province, Solomon Islands. Paper read at 7th International Conference on Oceanic Linguistics (COOL7), July 2007, at Nouméa.
- Tryon, Darrell, & Jean-Michel Charpentier. 1989. Les problèmes linguistiques à Vanuatu. In *Renaissance du Pacifique*, n° spécial de la revue *Ethnies*, 4 (8-9-10), 13-17.
- . 2004. *Pacific Pidgins and Creoles: Origins, Growth and Development*. Berlin-New York: Mouton de Gruyter.
- Vernaudon, Jacques. 1999. Valeurs aspectuelles de quatre marqueurs du tahitien. *Actances* 10: 67-90.
- VNSO. 2010. *2009 National Census of Population and Housing: Summary Release*. Vanuatu National Statistics Office, 31 August 2009.
- Waters, Sophia. 2010. The Semantics of French discourse particles *quoi* and *ben*. In *Selected Papers from the 2009 Conference of the Australian Linguistic Society*, edited by Y. Treiss & R. De Busser. Melbourne: Australian Linguistic Society.
- Woodbury, Anthony C. 2003. Defining documentary linguistics. *Language documentation and description* 1 (1): 35-51.
- . 2011. Language documentation. In *The Cambridge Handbook of Endangered Languages*, edited by P. K. Austin & J. Sallabank. Pp.159-186.
- Zalizniak, Anna A., Maria Bulakh, Dmitriy Ganenkov, Ilya Gruntov, Timur Maisak, & Maxim Russo. 2012. The catalogue of semantic shifts as a database for lexical semantic typology. In *New directions in lexical typology*, edited by M. Koptjevskaja-Tamm & M. Vanhove. Special issue of *Linguistics*, 50, 3. Pp.633–669.

