

DES VALEURS EN HÉRITAGE

Les isomorphismes sémantiques et la reconstruction des langues

Alexandre FRANÇOIS
(CNRS, LACITO)

francois@vjf.cnrs.fr

1 Typologie, comparatisme et contact de langues

1.1 *Les isomorphismes sémantiques*

La typologie des langues consiste à observer la récurrence, d'une langue à l'autre, de faits de structure linguistique. Il peut s'agir de structures morphosyntaxiques, mais aussi, par exemple, de *configurations sémantiques*. Désignons, par ce terme, l'ensemble des valeurs sémantiques propres à une forme particulière (morphème, lexème, phrasème), autrement dit son signifié ou sémème. Une fois qu'une telle configuration a été mise à jour par l'observation d'une forme dans une langue, elle peut devenir objet d'étude en soi, indépendamment de son support formel, et comparée d'une langue à l'autre. Par exemple, on observera qu'une langue L_1 possède un terme X doté d'une certaine polysémie ; puis l'on rencontrera la même polysémie dans une autre langue du monde L_2 , tout en constatant son absence dans une langue L_3 . On dira alors que L_1 et L_2 , mais pas L_3 , partagent une même configuration sémantique, qu'elles sont localement (c'est-à-dire, sur ce point précis) “isomorphes”.

Dans cette démarche, on ne compare pas le matériau phonologique, qui peut être différent, mais les faits de structure sémantique : de ce point de vue, cette première

phase d'observation relève davantage de la typologie que du comparatisme. Cependant, dans un second temps, on peut chercher à interpréter les isomorphismes observés entre les deux langues L_1 et L_2 ; c'est là que la linguistique historique peut entrer en jeu.

Globalement, si une même configuration sémantique se retrouve à l'identique dans deux langues L_1 et L_2 , trois types d'explications sont possibles :

- *explication typologique* : cette similitude peut être le fruit du hasard – ou pour être plus précis, d'une convergence typologique, par innovation parallèle
- *explication aréale* : il peut s'agir d'un fait de diffusion aréale, par contact de langues
- *explication génétique* : une configuration sémantique identique entre deux langues apparentées peut avoir été héritée de leur ancêtre commun.

Pour chaque cas observé, la question sera de savoir comment départager ces trois explications. Un premier critère résidera dans la nature des relations entre les deux langues que l'on compare. De toute évidence, l'explication aréale n'a de sens que si L_1 et L_2 appartiennent en effet à une même aire linguistique, ou sont entrées en contact au cours de leur histoire. De même, l'explication génétique ne fonctionne que si l'on sait par ailleurs que L_1 et L_2 sont apparentées, et descendent d'un même ancêtre. Le *Tableau 1* récapitule les différentes explications possibles, en fonction de la relation entre L_1 et L_2 .

Tableau 1 – Explications possibles d'une similitude structurale entre deux langues L_1 et L_2 , en fonction de leurs liens génétique et aréal

| | <i>aire linguistique différente</i> | <i>même aire linguistique</i> |
|--|---|---|
| <i>langues génétiquement non apparentées</i> | <ul style="list-style-type: none"> ▪ convergence typologique | <ul style="list-style-type: none"> ▪ convergence typologique ▪ diffusion aréale |
| <i>langues génétiquement apparentées</i> | <ul style="list-style-type: none"> ▪ convergence typologique ▪ héritage génétique | <ul style="list-style-type: none"> ▪ convergence typologique ▪ diffusion aréale ▪ héritage génétique |

Comme on le voit, la situation la plus ambiguë est celle des langues à la fois apparentées génétiquement, et coexistant dans la même aire linguistique. Or, c'est précisément sur un tel cas de figure que nous allons nous pencher (section 2).

1.2 Structures sémantiques et reconstruction historique

Supposons donc que l'on observe, entre deux langues L_1 et L_2 apparentées et en contact, une configuration sémantique identique – que celle-ci soit portée, ou non, par des formes apparentées. La question que souhaite poser cet article est de savoir *dans quelle mesure cette configuration sémantique peut être attribuée à un ancêtre commun.*

Cette question soulève deux types de difficultés. D'une part, il faut d'abord invalider, quand c'est possible, les deux autres explications concurrentes : celle de la simple convergence, et celle de la diffusion par contact. Ainsi, l'originalité typologique de la structure entrera en ligne de compte pour évaluer la probabilité d'une convergence de hasard (cf. Wilkins 1996) : plus une configuration sémantique sera idiosyncrasique, plus sa récurrence dans deux langues proches aura des chances d'être due à une histoire commune – que celle-ci remonte à un ancêtre commun, ou bien à une situation de contact. Quant à départager ces deux dernières hypothèses, j'en reparlerai plus loin (§2.4).

L'autre difficulté que soulève cette problématique est de savoir s'il est concevable de faire du comparatisme en travaillant sur des valeurs, plutôt que des formes. La question est paradoxale, tant il est vrai que la linguistique comparée est profondément enracinée dans la phonétique historique. Ce que recherche le comparatiste, c'est avant tout la récurrence des mêmes formes, ou plutôt celle de couples signifiant-signifié. Et bien entendu, cette démarche est indispensable pour établir les apparentements entre les langues et l'histoire précise de chaque terme.

Pourtant, cet article propose de discuter un cas de figure particulier. Il s'agit d'un groupe de langues dont on sait déjà – précisément en appliquant la méthode comparative classique – qu'elles sont étroitement apparentées. Or, à côté des diverses proto-formes que l'on peut reconstruire, il arrive parfois que l'on achoppe sur une difficulté : c'est lorsque les langues modernes s'accordent parfaitement sur une même structure sémantique, alors qu'elles divergent sur le plan formel. Dans ce cas, et selon certaines conditions que je discuterai, il peut être légitime de reconstruire une proto-forme *X, à signifiant provisoirement indéterminé, et caractérisée par sa configuration sémantique.

En somme, si une structure sémantique donnée se retrouve en plusieurs langues d'un même groupe génétique, et qu'il est possible d'écarter les hypothèses du hasard ou de la diffusion aréale, alors cette configuration doit être reconstruite au niveau de l'ancêtre de ce groupe. En émettant ce type d'hypothèse, on ne fait que transférer, sur le plan de la structuration du sens, la rigueur logique qui a déjà fait ses preuves dans le domaine de la comparaison des formes et de la morphosyntaxe.

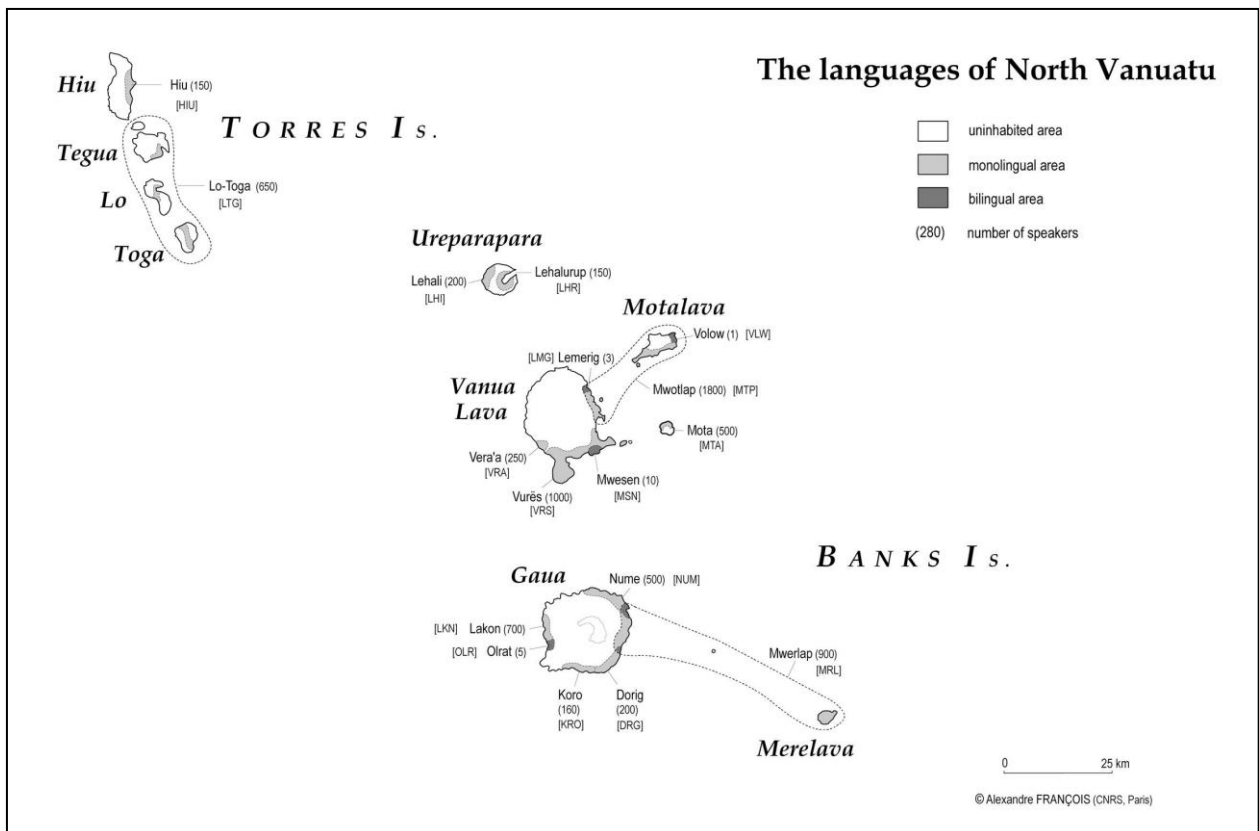
2 Diversité des formes, concordance des valeurs au nord Vanuatu

Ces questionnements concernant la reconstruction des structures sémantiques trouvent leur origine dans le travail d'observation et de description linguistique que j'ai entrepris en Mélanésie, plus précisément au nord du Vanuatu.

2.1 Les dix-sept langues du nord Vanuatu

Cette région du monde se caractérise par un émiettement linguistique extrême, avec pas moins de cent dix langues pour tout l'archipel du Vanuatu. Dans les seuls groupes des îles Banks et Torres au nord du Vanuatu, j'ai pu recueillir des données sur dix-sept langues distinctes, pour une population de 7500 habitants tout au plus (*Figure 1*).

Figure 1 – Carte des dix-sept langues parlées au nord du Vanuatu



Malgré l'absence d'intercompréhension, ces dix-sept langues appartiennent clairement à un même ensemble génétique : il s'agit de langues austronésiennes, appartenant toutes au sous-groupe océanien, et plus précisément à un ensemble cohérent d'une centaine de membres environ, nommé "Nord-Centre Vanuatu" (NCV). Ces dix-sept langues présentent tant de similitudes entre elles, que de futures recherches pourraient proposer de les regrouper en un clade spécifique à l'intérieur du groupe NCV, le groupe "Nord Vanuatu" (François 2005). Ainsi, en fonction du degré de précision voulu et de l'ampleur des comparaisons, "l'ancêtre commun" à ces langues des Banks et Torres pourra être, selon les cas, le "proto Nord Vanuatu" (PNV, hypothétique), le "proto Nord Centre Vanuatu" (PNCV), ou encore le proto-océanien (POc). Par ailleurs, ces communautés

ont toujours entretenu des relations de contact à travers leurs échanges économiques et matrimoniaux, si bien que les exemples de diffusion linguistique sont nécessairement nombreux (cf. §2.4).

2.2 Des langues à la fois diverses et isomorphes

Or, si l'on met de côté un bon nombre de morphèmes et de lexèmes dont l'apparement est évident, ces langues possèdent également toute une part de vocabulaire qui présente une double propriété : d'un côté, une différenciation marquée du matériau formel ; de l'autre, une extrême similarité des structures. Le *Tableau 2* présente ainsi le même énoncé traduit dans quelques-unes de ces langues¹ :

Tableau 2 – Correspondances terme à terme pour un même énoncé, dans quelques langues du nord Vanuatu

| | “Non, je n'ai pas bien dormi.” | | | | | |
|----------------|--------------------------------|------|------------------|---------------------|---------------------|------------------|
| | NÉG.EXIST, | 1sg | NÉG ₁ | dormir | bien | NÉG ₂ |
| <i>hiw</i> | təʔo, | nəkə | tati | miti ⁹ L | wu ⁹ Ləʔ | |
| <i>lo-toga</i> | tətəʔe, | nekə | tətə | mətər | ərβe | |
| <i>lehali</i> | tetye, | nə | tət | mutuj | ɣalse | tə |
| <i>mwotlap</i> | tateh, | nək | et | mitij | ɣalsi | te |
| <i>lemerig</i> | niβ, | nœ | ɪ | miʔir | ʔərmaʔ | ʔæ |
| <i>vera'a</i> | ɣitay, | nə | ɪʔ | miʔir | mintey | rus |
| <i>vurës</i> | ɔdiaŋ, | nə | ɣutu- | muβut | warey | |
| <i>mwesen</i> | eneŋ, | na | ete | muβut | maŋte | |
| <i>mota</i> | tayai, | nau | ɣate | matir | mantay | |
| <i>nume</i> | bek, | na | βeta | ŋɔŋɔr | liŋliŋi | mi |
| <i>dorig</i> | ɔbek, | na | t- | ŋɔr | taβul | timi |
| <i>olrat</i> | taya, | na | tɪ | ŋɔj | βilɪ: | wus |
| <i>lakon</i> | ta, | na | tɪ | ŋɔ: | kere | aβuh |
| <i>mwerlap</i> | tɪyɪ, | nə | ti- | mətər | mɪmɪn | tə̄a |
| <i>bislama</i> | NOGAT, | MI | NO | SLIP | GUT | |

Si l'on prête attention aux formes, ce tableau frappe plutôt par son hétérogénéité. En dépit de diverses similitudes (pronom personnel, formes du verbe “dormir”...) qui donnent à ces langues un air de famille certain, on observe une variation formelle parfois radicale, d'une langue à l'autre, pour des termes fonctionnellement équivalents. Ainsi, les

¹ Les transcriptions sont en API. L'ordre adopté est géographique, du nord-ouest au sud-est (cf. *Figure 1*). Pour le cas du *bislama* (dernière ligne du tableau), voir la discussion *infra*.

deux formes négatives, qu'il s'agisse de la négation prédicative (glosée /NÉG.EXIST/) ou de la négation associée au verbe (/NÉG₁/ ou /NÉG₁...NÉG₂/), présentent une variation dialectologique extrême. De même, l'adverbe interne au groupe verbal (glosé "bien") revêt des formes spectaculairement différentes d'une langue à l'autre.

Cependant, cette hétérogénéité au niveau des formes ne trouve pas son équivalent au niveau des structures, qui sont remarquablement parallèles dans toute la région.

C'est d'abord le cas pour les structures syntaxiques. Le *Tableau 2* révèle un parallélisme rigoureux dans l'ordre des mots ; les langues qui possèdent une négation discontinue la placent de part et d'autre du syntagme prédicatif ; et ce dernier se présente invariablement avec l'ordre {tête prédicative + adverbe interne}, à l'exclusion de tout argument ou circonstant.

Le parallélisme structural se retrouve également sur le plan fonctionnel et sémantique. Ainsi, quelle que soit sa forme phonologique, l'adverbe du *Tableau 2* présentera partout les mêmes contours sémantiques : "[accomplir une action] correctement, tout à fait, en la portant jusqu'à son degré qualitatif d'accomplissement". Les contextes d'emploi de cet adverbe seront rigoureusement identiques dans toutes les langues de la région. De même, dans chacune de ces langues, le mot-phrase négatif (fr. "Non") se confond avec le prédicat existentiel négatif (fr. "il n'y a pas") ; et ce terme, ici glosé /NÉG.EXIST/, est doté d'une configuration polysémique qui se retrouve à l'identique dans toutes ces langues : { *non ; il n'y a pas, ne pas exister ; être absent ; sans succès ; ce n'est pas grave, ça va* }. Quant au morphème de négation associé au prédicat verbal, il combine sa fonction de polarité négative à une valeur aspecto-modale de Realis : celle-ci embrasse, à elle seule, les valeurs de statif, présentatif, parfait, prétérit et aoriste, que distinguent les énoncés affirmatifs (François 2003: 35, 316). Pour ce morphème comme pour les autres, les caractéristiques sémantiques et morphosyntaxiques seront essentiellement les mêmes dans toute la région.

Ainsi, dans la grande majorité des cas, la description fonctionnelle et sémantique que l'on pourrait donner d'un mot pour une de ces langues conviendrait également pour toutes les langues voisines... à la forme du mot près. En forçant le trait, on pourrait présenter ces systèmes comme structurellement *isomorphes*, au sens où la catégorisation du sens, dans le lexique comme dans la grammaire, est en tous points parallèle d'une langue à l'autre². Tout se passe comme si l'aire linguistique du nord Vanuatu employait un seul et même "archi-système" linguistique, constitué des mêmes unités de sens, chaque langue ne différant de ses voisines que par le matériau phonologique qu'elle

² On trouve bien sûr des exceptions à ce principe d'isomorphisme généralisé. En particulier, le domaine où il se vérifie le moins est sans doute l'architecture des systèmes aspecto-modaux, en fait assez variables d'une région à l'autre.

associe à ces unités.

Le degré d'isomorphisme structural entre deux langues n'est pas nécessairement fonction du degré d'apparement de leurs formes : ainsi, les structures du lemerig et du lakon demeurent très parallèles, en dépit d'une nette divergence lexicale. Plus remarquable encore, on retrouve les mêmes catégories fonctionnelles (les mêmes "unités de sens") lorsque le matériau lexical provient d'une autre famille linguistique : c'est ainsi que le bislama, pidgin à base anglaise qui sert aujourd'hui de langue véhiculaire dans la région, épouse systématiquement les structures syntaxiques ou sémantiques des langues du substrat mélanésien (Keesing 1991). Comme le suggère la dernière ligne du *Tableau 2*, le bislama fonctionne *de facto* comme un nouvel avatar de cet "archi-système" régional, avec cette seule différence, que les moules structuraux sont remplis par du matériau phonologique d'origine anglaise plutôt qu'austronésienne³. C'est ce que certains créolistes de tendance substratiste nomment *relexification* : les valeurs sémantiques héritées du substrat demeurent inchangées, seul le matériau phonologique est emprunté à la langue lexifiante⁴.

Au bout du compte, si la divergence des formes explique l'absence d'intercompréhension spontanée entre les langues, en revanche, du point de vue structural, leur concordance quasi parfaite garantit leur traductibilité sans déperdition – que ce soit entre deux vernaculaires, ou entre un vernaculaire et la langue véhiculaire. Cet isomorphisme généralisé a pour corollaire la cohérence d'un univers de référence, associé à une aire culturelle fondamentalement homogène. Pour employer une formulation de type sociologique, on peut dire que ces communautés mélanésiennes, tout en utilisant la forme du signifiant comme marqueur identitaire, préservent et favorisent leur unité régionale en organisant leur univers notionnel selon les mêmes catégories.

2.3 *Les structures sémantiques récurrentes : rétention d'un ancêtre commun ?*

Cet isomorphisme généralisé des langues du nord Vanuatu constitue un fait de synchronie. Mais comment en est-on arrivé là ? Quel processus évolutif est susceptible d'aboutir à une telle concordance entre les langues d'une même région, qu'elles semblent toutes taillées selon le même patron, alors que leur lexique est si différent ?

³ Pour ne prendre qu'un seul exemple tiré du *Tableau 2*, le bislama /nogat/ (< angl. *no got*) présente exactement les mêmes valeurs que son équivalent dans les langues vernaculaires : { *non ; ne pas exister ; être absent ; sans succès ; ce n'est pas grave...* }.

⁴ Lefebvre (1998: 9) définit la relexification comme un "processus de substitution lexicale au cours duquel la seule information adoptée de la langue-cible, dans l'entrée lexicale, est l'information phonologique".

Une chose est certaine, vue l'idiomaticité des schémas sémantiques, on ne saurait voir là des cas de simple convergence typologique – l'autre nom du hasard en linguistique. Restent alors deux explications possibles (cf. *Tableau 1*) : d'une part, l'héritage génétique partagé ; d'autre part, la diffusion aréale par contact. Je vais les examiner successivement.

2.3.1 *Des valeurs en héritage*

Parler d'un héritage commun des valeurs sémantiques, c'est faire l'hypothèse que l'isomorphisme structural que l'on observe en synchronie peut s'expliquer par l'existence de ces mêmes configurations au niveau de l'ancêtre commun à ces langues. On proposerait ainsi, par exemple, d'attribuer à la proto-langue ancestrale aux langues modernes – disons, le “proto Nord Vanuatu” (PNV) – les propriétés structurales présentes dans le *Tableau 2* (ordre des constituants, propriétés morphosyntaxiques et sémantiques de chaque terme, etc.).

Admettons provisoirement – j'en discuterai plus loin – que cette proposition puisse être démontrée. Elle présenterait alors un grand intérêt : car elle permettrait, au moins au titre d'hypothèse de travail, de reconstruire des états anciens de langue en reconstituant non seulement des proto-formes, mais également des proto-valeurs sémantiques. Bien entendu, par définition, chaque proto-signifié n'avait d'existence que s'il était lexifié par un proto-signifiant unique, lequel idéalement devrait pouvoir être reconstruit par la même occasion. Oui mais voilà, il peut arriver que les données suggèrent fortement l'existence d'un terme *X unique pourvu de certaines propriétés sémantiques, mais ne permettent pas de reconstruire la proto-forme elle-même. Par exemple, on peut raisonnablement supposer, au vu du *Tableau 2*, que le PNV possédait un adverbe interne de prédicat signifiant “correctement, tout à fait...” ; mais quant à en déterminer la forme phonologique, voilà qui semble impossible au vu des formes modernes.

D'une certaine façon, la démarche ici proposée renverserait l'ordre usuel des priorités du comparatiste, lequel fait généralement porter tous ses efforts sur la reconstruction des formes, avant de leur attribuer – quand il prend le temps de le faire vraiment ! – un ensemble ordonné de valeurs. En réalité, la même rigueur méthodologique (principe des innovations partagées, des correspondances régulières...) devrait également permettre de reconstruire certains proto-signifiés, en tant que propriété structurale de la proto-langue que l'on cherche à caractériser. Quant à déterminer la forme phonologique de ces configurations sémantiques (leur proto-signifiant), c'est bien sûr une démarche souhaitable, mais qui doit parfois être remise à plus tard, en la subordonnant, par exemple, à l'acquisition ultérieure de nouvelles données.

2.3.2 *Changement des signifiants, conservation des signifiés*

Il est vrai que la linguistique historique a déjà coutume de reconstruire des propriétés typologiques, notamment syntaxiques, des proto-langues. Mais généralement, la notion de préservation des faits de structures va nécessairement de pair avec celle de leur support formel, qu'il s'agisse de marques segmentales, séquentielles ou autres. Dans le cas que j'analyse ici, l'hypothèse envisagée est plus hardie : c'est l'idée que la transmission d'une langue au cours du temps puisse consister à modifier les formes tout en préservant intactes les valeurs sémantiques.

Ce n'est pas là la conception usuelle du changement sémantique. Dans beaucoup de langues du monde, on observe généralement que tout changement formel – une forme X étant concurrencée, sur son propre territoire sémantique, par une forme Y – a pour corollaire une reconfiguration des contours sémantiques de chaque mot, et une redistribution de leurs sèmes. Par exemple, dans certains registres du français, le sème “condition naturelle qui différencie les hommes des femmes” tend actuellement à être lexifié par le mot *genre* plutôt que le mot *sexe*. Par un effet structural évident, les contours notionnels de ces deux mots s'en trouvent redessinés, le champ sémantique de *sexe* cédant du territoire à celui de *genre*. C'est ainsi qu'un changement lexical implique normalement une redéfinition des configurations sémantiques.

Pourtant, il semble que les langues du nord Vanuatu suivent parfois un mécanisme évolutif différent. D'un côté, les facteurs universels de changement lexical – glissements métaphoriques, recherche de l'expressivité, tendance sociale à la diversification et à “l'ésotérogénie” (Ross 1996)... – motivent bel et bien, dans ces langues, un renouvellement interne du vocabulaire ; mais, d'un autre côté, le découpage des valeurs sémantiques tel qu'il existe au niveau de l'archi-système régional semble avoir, pour ainsi dire, assez de pouvoir contraignant pour imposer à tout nouveau terme un ensemble de valeurs prédéfini. C'est ainsi que les formes des mots peuvent changer, tout en laissant intactes, au bout du compte, les unités de sens. On peut comparer ce phénomène, *mutatis mutandis*, à la “relexification” observée dans les pidgins et créoles : les configurations sémantiques sont héritées telles quelles, seul change le matériau phonologique qui vient se couler dans le moule préexistant.

Ce phénomène de *relexification interne* peut être illustré par l'évolution d'un des directionnels spatiaux en dorig. Dans les langues du nord Vanuatu, le directionnel signifiant “vers le bas” présente une polysémie en fonction du contexte de référence spatiale (François 2004) : [axe vertical] “vers le bas, par terre” ; [sur l'île] “vers la côte” ; [en pirogue] “en s'éloignant de l'île” ; [échelle géographique] “en allant vers le nord-ouest” – sans parler de certaines valeurs métaphoriques de “vers le bas” (action qui ralentit, nombre qui diminue...). La plupart des langues des îles Torres et Banks

lexifient cet ensemble polysémique avec des formes que l'on peut reconstruire en un étymon PNV **suwe* “en bas, vers le bas” (François 2005: 499). Or, la langue dorig – comme d'ailleurs le nume, le koro et l'olrat – a remplacé **suwe* par une forme innovative *rɔr*. Son étymologie peut être reconstruite : PNV **roro* “profond, bas”. Dans la mesure où le trait sémantique à la source du changement est la notion de profondeur ou faible altitude, la langue dorig a sans doute commencé par élargir le sémantisme de **roro* de “être bas” à “être en bas”. On peut imaginer que cet élargissement de **roro* n'a d'abord remplacé que l'acception verticale de **suwe*, tandis que les autres valeurs (ex. “vers le nord-ouest”), peu transparentes, ont dû continuer pendant un temps à être codées par **suwe*.⁵ À ce stade-là de l'évolution, on aurait assisté en effet à une véritable reconfiguration sémique, le sémème de **suwe* cédant du terrain à celui de **roro*.

Pourtant, cette reconfiguration n'est pas ce que l'on observe aujourd'hui. Au contraire, on s'aperçoit que le *rɔr* du dorig contemporain a, d'une part, entièrement perdu les anciennes valeurs de **roro* (“profond” se dit aujourd'hui *leβsɔrsɔr* ; “bas” se dit *ta:βtin*), et d'autre part, entièrement endossé celles de l'ancien **suwe* {vers le bas ; vers la côte ; vers le large ; vers le nord-ouest ; action qui ralentit...}. Au bout du compte, on retrouve exactement les mêmes catégorisations sémantiques qu'au départ, comme si la forme **roro* avait tout simplement quitté son moule sémantique originel pour venir occuper la place de **suwe*, sans même en modifier la signification. Aujourd'hui, les contours sémantiques de *rɔr* (< **roro*) en dorig épousent étroitement, par exemple, ceux de *hɔw* (< **suwe*) en lakon voisin : les deux termes sont devenus parfaitement équivalents et traduisibles d'une langue à l'autre, et n'ont gardé aucune trace sémantique de leurs différentes étymologies.

Au bout du compte, tout se passe comme si les reconfigurations de signifié qui “auraient dû” se produire au cours d'un tel changement lexical avaient été contrecarrées, en cours d'évolution, par la rigidité des catégorisations notionnelles préexistantes, ces unités de sens propres à l'archi-système qui transcende toutes les langues de la région. Alors que le signifiant apparaît perméable à l'innovation et au remplacement lexical, le signifié semble résister davantage à toute évolution, en vertu d'un effet d'inertie qu'entretiennent continuellement les interactions interlangues.

Voilà comment, dans cette région du monde, l'évolution des langues combine typiquement changement de signifiant (tendance à la divergence lexicale) et transmission intacte des signifiés (préservation d'une traductibilité optimale interlangues). C'est en tout cas de cette manière que l'on peut justifier l'hypothèse, *a priori* paradoxale, selon laquelle un groupe de langues peut hériter d'un ancêtre commun une

⁵ C'est exactement la situation que l'on trouve dans la langue iaai, parlée à Ouvéa (François 2004: 28).

configuration sémantique donnée, tout en remplaçant l'étymon original par des formes innovantes.

2.3.3 Quelques candidats à la reconstruction sémantique

Cette démarche de reconstruction sémantique permettrait de formuler des hypothèses historiques pour une bonne partie du vocabulaire de ces langues. Le *Tableau 3* présente ainsi un échantillon de quatre termes qui ont en commun deux propriétés :

- un signifié à la fois complexe (idiosyncrasique) et identique d'une langue à l'autre
- des signifiants trop hétérogènes pour permettre de reconstruire une proto-forme.

Chaque série de formes est accompagnée d'une brève glose, ainsi que de son équivalent exact dans le pidgin bislama (§2.2).

Tableau 3 – Quelques candidats à la reconstruction sémantique

| | *X ₁ Focus temporel | *X ₂ Restrictif | *X ₃ Verbe interrogatif | *X ₄ Déclaratif- conjonction |
|----------------|--------------------------------------|-------------------------------|--|---|
| <i>bislama</i> | JAS | NOMO | MEKWANEM | SE |
| <i>hiw</i> | takə | ŋ ^w utujə | taβə | təm |
| <i>lo-toga</i> | akə | wərəŋə | ʈaβə | tɛ |
| <i>löyöp</i> | pɛj | wɛβan | antʃe | tə |
| <i>lehali</i> | jak | ɛwwe | danha | dæ |
| <i>volow</i> | ḡb ^w ɔjə | ɣɛwɪ | aŋtɛ | gə |
| <i>mwotlap</i> | ḡp ^w ɔjə | ɪwɪ | akteɣ | sə |
| <i>lemerig</i> | mak | ḡp ^w ɔɣər | səβɪɪ | wœ |
| <i>vera'a</i> | mak | ɣiβa | dasiβiɛ | sə |
| <i>vurës</i> | kara | ɣɛm | daɣɛɪ | βita |
| <i>mwesen</i> | kara | ɣəp | naɣsɛ | (wə)ta |
| <i>mota</i> | ḡp ^w ara | ɣap | saβai | was |
| <i>nume</i> | ḡp ^w ar | am | dɛsɛ | si |
| <i>dorig</i> | ḡp ^w ra | wər | (d)aksaβ | (k)ak |
| <i>koro</i> | ḡp ^w ara | wər | dasβa | (k)ak |
| <i>olrat</i> | tak | wəj | ɣasβa | ka |
| <i>lakon</i> | lak | wə: | βaha | sa |
| <i>mwerlap</i> | k ^w ɛr | ɣəm | dɪ | si |

Voici, résumées, les propriétés morphosyntaxiques et sémantiques de chacune de ces séries (dénommées ici par leur équivalent bislama).

- JAS : proclitique aspecto-temporel dit “*Focus temporel*” (François 2003: 199-216). Localise un événement dans le temps, en sorte que cette localisation se trouve focalisée contrastivement (“c'est alors [et pas avant] que P”). Employé comme passé immédiat (ex. “je viens juste de le voir”), présent inaugural (“c'est la première fois que je le vois”) ou futur dilatoire (“c'est seulement alors que je le verrai”).
- NOMO : adverbe à valeur de restrictif quantitatif ou d'atténuateur qualitatif : “ne... que ; seulement”. Accompagne les prédicats notionnellement orientés vers la faible quantité (ex. “juste petit”, “juste léger”, “juste quatre”...). Atténue le prédicat “bien” avec effet de politesse (“c'est bien seulement” = “tout va bien”).
- MEKWANEM : verbe interrogatif, non analysable, compatible avec trois interprétations : dynamique agentif (“X fait quoi ?”), dynamique patientif (“qu'arrive-t-il à X ?”), statif (“X est comment ?”).
- SE : particule déclarative introduisant le discours rapporté (avec ou sans verbe introducteur⁶), fonctionnant également comme complémentateur universel (“que”), et comme conjonction de but.

Pour le linguiste qui cherche à remonter à l'ancêtre commun de ces systèmes, l'absence d'indice permettant de proposer une proto-forme plausible pour chacune de ces séries ne devrait pas empêcher d'émettre des hypothèses sur le plan sémantique. Étant donnée la concordance quasi parfaite de ces termes d'une langue à l'autre, il semble légitime de supposer que le proto-système était caractérisé par des structures sémantiques analogues à celles qui sont attestées de nos jours. Dans ce cas, à côté de reconstructions authentiques (ex. **suwe* “en bas”...), on obtiendrait au moins des semi-reconstructions, du type * X_1 “*Focus temporel*”, * X_3 “*verbe interrogatif*”, etc.

Malgré l'apparence d'inachèvement, on ferait ainsi avancer à grands pas le travail de reconstruction historique. Il ne resterait plus ensuite qu'à acquérir des données manquantes – soit à l'intérieur de cette région, soit à l'extérieur – pour compléter ces hypothèses de structure par des hypothèses de forme, la première étape du travail favorisant nécessairement la seconde.

2.4 Et la diffusion ?

2.4.1 L'interprétation diffusionniste

La section précédente vient de montrer la possibilité, tout du moins théorique, de traiter un isomorphisme structural entre langues d'une même famille – même en l'absence de

⁶ Les sept dernières langues du tableau emploient un morphème différent pour introduire le discours direct sans verbe introducteur.

formes apparentées – comme un cas de préservation d'une structure sémantique présente au niveau de leur ancêtre commun. On a vu les effets importants qu'une telle hypothèse peut comporter en matière de reconstruction historique.

Pourtant, il demeure une objection possible à un tel raisonnement, et c'est l'argument de la diffusion aréale. En effet, les connaissances historiques et anthropologiques dont on dispose sur le nord Vanuatu montrent clairement que les îles Banks et Torres constituent une aire de contacts culturels, et partant linguistiques, particulièrement intenses. Dès lors, comment savoir si telle similitude structurale entre ces langues est due à la rétention de traits propres à leur ancêtre commun, ou bien à des phénomènes de contact ?

Prenons le cas du Focus temporel (*Tableau 3*). On peut imaginer que l'ancêtre commun de ces langues (le PNV) ne possédait pas une telle catégorie, mais par exemple, deux morphèmes en distribution complémentaire : l'un opérant dans le domaine *realis*, l'autre dans le domaine *irrealis*. Puis serait apparue, dans une de ces langues, la tendance à fusionner ces deux opérations en une seule, donnant naissance à ce que je décris comme "Focus temporel". Pour peu que cette tournure émergente ait paru fonctionnellement plus efficace, elle aura pu se diffuser, de proche en proche, tant et si bien que toutes les langues de la région auront fini par adopter cette nouvelle catégorie aspectuelle.

Que ce dernier scénario soit plausible ou non, sa simple possibilité théorique apporte un contre-argument à l'idée que l'on devrait expliquer tout cas d'isomorphisme par la rétention de structures appartenant à la proto-langue. En fait, il peut aussi bien s'agir de convergence par contact.

2.4.2 Une solution : élargir le champ d'observation

En somme, la concordance structurale forte que l'on observe aujourd'hui dans ce groupe de langues doit être attribuée, pour partie, à un héritage génétique commun – impliquant donc des structures que l'on est en droit de reconstruire au niveau de la proto-langue – et pour partie, à des phénomènes ultérieurs de convergence par diffusion aréale. Comment faire le départ entre ces deux interprétations, pour chaque structure considérée ?

Tant que l'on demeure à l'intérieur d'une même aire de contacts, et qu'on ne dispose pas de documents historiques anciens, la solution à cette énigme sera probablement impossible à trouver, car tout parallélisme de structure pourra s'expliquer aussi bien, en principe, par diffusion que par héritage génétique. Pourtant, une solution est possible, suggérée par le *Tableau 1* (§1.1) : elle consiste à élargir le champ d'observation à d'autres langues du même groupe génétique, mais appartenant à une aire linguistique distincte. Dans le cas du nord Vanuatu, les déterminismes géographiques et historiques

permettent de considérer, pour l'essentiel, la zone des Banks et des Torres comme formant un “Sprachbund” homogène, et toute langue parlée en dehors de cette zone, comme appartenant à une aire différente.

Ainsi, supposons qu'une structure aujourd'hui récurrente dans le nord Vanuatu ne trouve son équivalent nulle part ailleurs dans la famille océanienne : on n'aura alors aucun moyen de savoir s'il s'agit là d'une innovation datant de l'ancêtre commun PNV (*isomorphisme par héritage*) ou bien si elle est apparue localement après la fragmentation dialectale, et répandue par diffusion au sein de cette aire linguistique (*isomorphisme par contact*). À l'inverse, imaginons maintenant que cette structure soit également attestée dans d'autres langues apparentées, et restées historiquement à l'écart de notre aire : ce sera alors un argument solide pour éliminer l'hypothèse aréale, et reconstruire la structure en question au niveau de la proto-langue.

Voici quelques exemples empruntés à l'araki (François 2002), une langue NCV parlée au sud de l'île de Santo, environ 200 km au sud des îles Banks, assez loin pour être considérée externe au Sprachbund du nord Vanuatu :

- en araki, la marque aspectuelle *pa* consiste à focaliser un événement dans le temps, comme passé immédiat, présent inaugural, ou futur dilatoire. C'est la preuve que la configuration dite “*Focus temporel*” doit être reconstruite au niveau de leur proto-langue commune, le PNCV (et donc aussi de sa langue-fille le PNV).
- l'araki possède un verbe-interrogatif inanalysable *ro* “faire quoi, être comment, qu'arrive-t-il”. Ce fait de structure peut donc être reconstruit avec certitude au moins au niveau du PNCV.
- l'araki possède un verbe *re* “dire” qui introduit le discours rapporté, fonctionne comme complémenteur universel et conjonction de but. Certes, la relative fréquence d'une grammaticalisation {“dire” > complémenteur} dans le monde pourrait suggérer qu'il pourrait s'agir là d'une simple convergence typologique entre les langues du nord Vanuatu et l'araki. Mais si d'autres langues du groupe NCV présentent exactement la même propriété structurale, alors il y a de fortes probabilités pour qu'une telle structure polyfonctionnelle doive être reconstruite dès le niveau de la proto-langue (le PNCV).
- enfin, le prédicat existentiel négatif se dit *motfêre* (François 2002: 65), lequel fonctionne comme mot-phrase (“non”) et signifie également “[faire] sans succès”, “ce n'est pas grave”, etc. Il en va d'ailleurs de même pour *tæe* en teanu (langue océanienne parlée à Vanikoro, îles Salomon) : ceci suggère de reconstruire la catégorie /NÉG.EXIST/, avec sa polysémie, non seulement au niveau du PNCV, mais même du proto-océanien...

3 Conclusion

En somme, pour peu que l'on réussisse à croiser et départager les trois types d'interprétation – typologique, génétique et aréale – il doit être possible de formuler des hypothèses historiques portant sur la structuration du sens dans des états de langue anciens. Cette démarche, ne l'oublions pas, s'applique d'abord aux proto-formes que l'on peut reconstruire. Mais de façon légèrement plus hardie, j'espère avoir montré qu'elle peut également enrichir le travail du comparatiste lorsque les cartes du signifiant ont été brouillées par des phénomènes de relexification interne. C'est du moins là une discussion qui méritait d'être commencée.

4 Références

- DURIE, Mark & Malcolm ROSS (éds). 1996. *The comparative method reviewed: Regularity and irregularity in language change*. Oxford: Oxford University Press.
- FRANÇOIS, Alexandre. 2002. *Araki. A disappearing language of Vanuatu*. Pacific Linguistics, 522. Canberra: Australian National University.
- . 2003. *La sémantique du prédicat en mwotlap (Vanuatu)*. Collection Linguistique de la Société de Linguistique de Paris, 84. Paris, Louvain: Peeters.
- . 2004. “Reconstructing the geocentric system of Proto Oceanic”. *Oceanic Linguistics* 43-1, 1-32.
- . 2005. “Unraveling the history of the vowels of seventeen northern Vanuatu languages”. *Oceanic Linguistics* 44-2, 443-504.
- KEESING, Roger M. 1991. “Substrates, calquing and grammaticalization in Melanesian Pidgin”. In E. TRAUGOTT & B. HEINE (éds), *Approaches to grammaticalization I*, 315-342. Philadelphia: Benjamins.
- LEFEBVRE, Claire. 1998. *Creole genesis and the acquisition of grammar: The case of Haitian creole*. Cambridge: Cambridge University Press.
- ROSS, Malcolm. 1996. “Contact-induced change and the comparative method: Cases from Papua New Guinea”. In M. DURIE & M. ROSS (éds), 180-217.
- WILKINS, David. 1996. “Natural tendencies of semantic change and the search for cognates”. In M. DURIE & M. ROSS (éds.), 264-304.