

# Les premières expansions austronésiennes

Michel Ferlus

► **To cite this version:**

Michel Ferlus. Les premières expansions austronésiennes. The Eleventh International Conference on Austronesian Linguistics / Onzième Conférence Internationale de Linguistique Austronésienne, Jun 2009, Aussois, France. halshs-00952334

**HAL Id: halshs-00952334**

**<https://halshs.archives-ouvertes.fr/halshs-00952334>**

Submitted on 26 Feb 2014

**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.

## Les premières expansions austronésiennes

Michel Ferlus  
*Chercheur indépendant*

### 0. Introduction

Trois disciplines interviennent dans les recherches sur l'origine et l'expansion des langues austronésiennes: la linguistique (par définition), l'archéologie et la génétique. Selon la théorie dominante actuelle des linguistes, Taiwan est le centre de diversification et d'expansion des langues austronésiennes. Les autres spécialistes ajustent leurs interprétations sur les présupposés des linguistes. Toutefois, un examen attentif des arguments linguistiques permet de déceler certaines faiblesses : non prise en compte de l'origine continentale des langues austronésiennes, conceptions phylogéniques erronées, sous-estimation du rôle de la durée dans l'espace restreint de Taiwan et, par dessus tout, trop de conclusions hâtives et péremptoires. En interprétant différemment les mêmes données que les linguistes nous proposerons une nouvelle hypothèse sur les premières expansions austronésiennes.

AN	austronésien
MP	malayo-polynésien
PAN	proto austronésien
PMP	proto malayo-polynésien
TK	tai-kadai

### 1. Etat des lieux

Les spécialistes situent l'arrivée des premiers locuteurs formosans venant du continent vers 4000-3500 av. notre ère. Le 'Out of Taiwan' se serait produit un peu plus tard entre -3000 et -2000. Ce qui signifie que pendant au moins cinq mille ans le formosan des origines se serait diversifié en quelques dix taxons sur un espace restreint tandis qu'un seul taxon, le MP aurait occupé tout le reste de l'espace austronésien en quatre millénaires. Malgré les incertitudes sur les dates, il y a une contradiction flagrante.

Les spécialistes ne tiennent pas compte des langues AN disparues du continent asiatique. C'est comme si les langues indo-européennes disparaissaient de l'Ancien Continent et ne subsistaient que par l'anglais, l'espagnol, le portugais et le français en Amérique. Les linguistes supposeraient-ils une diversification sur place !

Les conceptions phylogéniques de l'AN sont plus ou moins basées sur le modèle du *Stammbaum*. Or, nous avons la conviction que ce modèle ne convient pas à l'aire linguistique de l'Asie Orientale où les contacts entre langues et entre familles de langues ont une grande importance. D'autre part, ce modèle devient très vite inopérant lorsque des langues apparentées sont maintenues en contact dans un espace restreint.

La phylogénie AN repose sur le vocabulaire de base, ce qui est normal, et sur les confusions phonétiques dont il est souvent difficile d'affirmer si elles sont héritées ou empruntées.

---

\* Je remercie cordialement Frédéric Pain (Université Catholique de Louvain, Belgique), spécialiste de linguistique de l'Asie du sud-est, qui a relu ce texte avec la plus grande attention.

En réexaminant la propagation du vocabulaire des nombres dans les langues formosanes et les relations entre l'AN et le tai-kadai, nous proposerons un autre modèle des premières expansions austronésiennes.

## 2. Les nombres AN en MP et en formosan

Le traitement du vocabulaire des nombres est exemplaire des différences entre l'approche phylogénique (les nombres sont hérités) et l'approche diffusionniste (les nombres sont empruntés). En un mot, les nombres en formosan sont-ils hérités ou empruntés ? Notre point de vue est que les nombres sont en partie empruntés. Ils sont de ce fait de bons marqueurs des influences entre langues. On a donc choisi d'aborder le problème des premières expansions AN par l'étude des nombres en formosan parce qu'il nous est apparu que ce paradigme pourrait permettre de mettre en évidence des mouvements de populations insoupçonnés. Nous partirons du travail salutaire effectué par Sagart sur la genèse des nombres AN.

Laurent Sagart [2004, 2008] a démontré comment en MP et dans certaines langues formosanes les nombres '7', '8' et '9' se sont formés par composition à partir de '5+2', '5+3' et '5+4' [2004: 415-419]. Ce mode de formation est encore visible en pazeh, langue conservatrice du nord-ouest de l'île. Ensuite, Sagart [2004: 417] montre comment à partir des formes explicites du pazeh se déduisent les formes courtes du PMP par les procédés de troncation syllabique et d'amuïssement consonantique. Reprenons les exemples de l'auteur en les complétant et en les réorganisant selon nos besoins :

		Pazeh		PMP
		<i>composition</i>	<i>réduction</i>	
'5'		xaseb		(*lima)
'6'	(5+1)	xaseb-uza	xasebuza	(*enem)
'7'	(5+2)	xaseb-i-dusa	xasebidusa	*pitu
'8'	(5+3)	xaseb-i-turu	xasebuturu	*walu
'9'	(5+4)	xaseb-i-supat	xasebisupat	*siwa
'10'		(isit)		(*puluq)

L'origine de PMP '6' et '10' est obscure ; PMP '5' provient de 'main'. Pour une meilleure compréhension de ces phénomènes, les lecteurs sont invités à se reporter aux articles cités de Sagart.

Les nombres PMP de '5' à '10' ont donc deux origines possibles : soit ils se sont formés par la réduction de formes composées préservées en pazeh ('7', '8' et '9'), soit ils résultent d'une introduction expliquée ('5') ou non explicable ('6' et '7'). Cette séquence des nombres de '5' à '10' reconstruite en PMP est irrégulièrement présente dans les langues formosanes (voir **tableau 1**). L'ordre est '7', '5', '6', '8', '9', '10' du plus représenté ('7' est dans seize langues de la liste) au moins représenté ('10' est dans trois langues). La séquence s'étale selon un gradient orienté en gros du nord-ouest au sud-est de l'île de Taiwan (voir **tableau 2**).

Dans la théorie de Sagart [2008, plate 5.1], les innovations se sont propagées à partir du nord-ouest, en tournant autour de l'île dans le sens lévogyre jusqu'au sud-est, selon l'ordre cumulatif '7' '5' '6' '8' '9' '10' : sur un substrat conservateur apparaît l'innovation de '7' (\*pitu). Sur ce nouvel état de langue qui se propage apparaît l'innovation de '5' (\*lima). Sur l'état de langue suivant apparaît l'innovation de '6' (\*enem) et ainsi de suite... Les expansions de ces innovations en chaîne finissent par atteindre le sud-est de Taiwan, zone où elles se cumulent. Puis, toujours selon Sagart, des locuteurs porteurs de ces innovations

quittent l'île (vers 2000 avant notre ère) selon deux directions pour donner naissance au MP et TK. La branche MP connaîtra le destin que l'on sait, tandis que la branche TK s'arrimera au continent. Cette sorte de course de relais des innovations de nombres n'a pas d'équivalent dans le monde. Le linguiste ne peut être que sceptique devant ce modèle où les innovations se propagent à sens unique. D'autre part, il se pose la question de savoir pourquoi le MP qui serait né sur l'île de Taiwan n'y est plus représenté aujourd'hui. Pour résoudre ces problèmes, nous allons proposer un autre modèle des premières expansions austronésiennes.

En inversant le sens de la propagation nous envisagerons une arrivée par le sud de populations de langue MP venant du nord des Philippines (Luçon de toute évidence). Le corollaire étant que le MP s'est formé hors de Taiwan.

Le **tableau 2** montre assez clairement une progression en tenaille à partir du sud-est de Taiwan où est représenté le maximum d'innovations, jusqu'au nord-ouest où se sont maintenues des langues conservatrices. Cette disposition est encore assez visible aujourd'hui malgré les millénaires, ce qui montre une certaine stabilité des populations jusqu'au début de la sinisation.

Du point de vue maritime, le sud-est est le point d'arrivée naturel des embarcations venant de Luçon (voir **tableau 3**). C'est cette voie que les Espagnols des Philippines ont suivi lors de leur brève implantation (1626-1642) dans l'île : ils ont remonté la côte est et se sont installés au nord-est qui offrait un arrière-pays.

En résumé, notre hypothèse est que les formes MP des nombres de '5' à '10' sont arrivées du nord des Philippines et se sont répandues irrégulièrement parmi les langues formosanes à partir du sud-est de l'île de Taiwan. Il est probable qu'il y a eu plusieurs vagues d'arrivées. Cette hypothèse s'oppose donc aux hypothèses en cours qui placent l'émergence du MP dans l'île. Question : d'où venait la forme de MP implantée au nord des Philippines qui a introduit les nombres '5'-'10' parmi les langues formosanes au point de laisser croire aux linguistes qu'ils sont le résultat de changements sur place ? C'est là qu'il faut se tourner vers les langues de la famille TK.

### 3. Le tai-kadai et l'austronésien

Les langues de la famille tai-kadai / kradai peuvent être réparties en cinq groupes [Ostapirat 2005] :

Kra / kadai : gelao, qabiao, lachi, buyang

Hlai / hli

Kam-sui : kam / dong, mulam, maonan, sui, ai-cham, mak, then, lajia

Be : ong be / lingao

Tai ou tai-yay :

Nord (*Northern Tai*) : buyi, zhuang, saek

Centre (*Central Tai*) : tày, nung

Sud-Ouest (*Southwestern Tai*) : ahom, khamti, shan, tay blanc, tay noir, tay dèng, tay yo, lao, siamois

La place de la famille TK parmi les familles linguistiques de l'Asie Orientale a toujours été un sujet de controverses. Le TK a d'abord été classé dans le sino-tibétain avant de devenir une famille autonome. L'existence d'un vocabulaire commun au TK et à l'AN a conduit Benedict [1942, 1975] à regrouper ces deux familles dans un phylum austro-thai, et donc à les considérer comme génétiquement apparentées. Cette hypothèse, quoique féconde en son temps, n'a jamais été franchement admise par les spécialistes. Elle a été réfutée par Reid [1988] et par Thurgood [1994] arguant que le vocabulaire en commun était le résultat d'emprunts.

Récemment, Sagart [2004, 2005], naviguant à contre courant a carrément annexé le TK à l'AN en le plaçant à un embranchement inférieur au niveau des taxons formosans. Malgré son audace, cette nouvelle hypothèse mérite un examen attentif. Remarquons d'abord que cette hypothèse est en droite ligne du 'Out of Taiwan'. En dépit de l'importance du vocabulaire partagé avec l'AN, la plus grande partie du vocabulaire fondamental TK (parties du corps, animaux, ...) ne se retrouve ni en AN ni dans les autres familles linguistiques voisines. La parenté génétique étant exclue, il faut envisager un phénomène d'emprunt du MP vers le TK. Où, quand et dans quelles circonstances auraient eu lieu ces emprunts. Examinons d'abord la place et les problèmes des langues TK.

L'examen des langues TK montre que cette famille a eu une histoire phonétique complexe et troublée. En dehors du vocabulaire bien réel qu'elle partage avec l'AN, on peut y déceler une strate môn-khmer (vocabulaire du riz et de la riziculture) et une strate sinitique. Haudricourt [1968] a mis en évidence un ancien phénomène d'appauvrissement sémantique en TK. Alors que les langues d'Asie possèdent trois ou quatre mots pour nommer les états du riz (riz sur pied, riz non décortiqué, riz décortiqué, riz cuit), le TK n'en possède qu'un seul. Cet appauvrissement doit avoir son origine dans une situation de contact précaire, le TK se serait mal transmis en passant d'une population à une autre.

Le problème de la place du TK dépend d'abord de la conception phylogénique sous-jacente. La plupart des austronésianistes privilégient inconsciemment le modèle indo-européen du *Stammbaum* négligeant le fait que dans l'aire extrême-orientale les phénomènes de contact et d'emprunt sont tout aussi déterminants que la part héritée de la proto langue. Les emprunts sont parfois si bien intégrés qu'ils peuvent être pris pour du vocabulaire hérité. Déceler les strates d'emprunts nécessite une analyse en profondeur et des reconstructions exactes à tous les niveaux. En vérité, dans l'aire extrême-orientale, le vocabulaire emprunté est aussi important que le vocabulaire hérité dans la formation des familles et des groupes linguistiques.

Examinons quelques mots du vocabulaire partagé entre l'AN et le TK. Les données proviennent de Sagart [2004, 2008] et de Ostapirat [2005]. Le buyang est une langue conservatrice du groupe kra / kadai, qui a préservé des sesqui-syllabes. Le proto tai utilisé ici correspond au *Southwestern Tai* de Fang Kuei Li [1977].

PAN	PMP	Buyang	PT	
maCa	mata	m <sup>o</sup> ta <sup>54</sup>	ta: <sup>A</sup>	'œil, <i>eye</i> '
maCay	matay	m <sup>o</sup> te: <sup>54</sup>	taj: <sup>A</sup>	'mourir, <i>die</i> '
kuCu	kutu	qa <sup>o</sup> tu <sup>54</sup>	–	'pou, <i>louse</i> '
Cumay	–	ta <sup>o</sup> mε <sup>312</sup>	<sup>h</sup> miaj <sup>A</sup>	'ours, <i>bear</i> '
(qayam)	manuk	m <sup>o</sup> nuk <sup>11</sup>	nok	'oiseau, <i>bird</i> '
daNum	danum	(ʔo:ŋ <sup>42</sup> )	nam <sup>C</sup>	'eau, <i>water</i> '
bulaN	bulan	–	βlian <sup>A</sup>	'lune, <i>moon</i> '
(RaCep)	lima	ma <sup>312</sup>	–	'cinq, <i>five</i> '
enem	enem	nam <sup>54</sup>	–	'six, <i>six</i> '
pitu	pitu	tu <sup>312</sup>	–	'sept, <i>seven</i> '
walu	walu	m <sup>o</sup> du <sup>11</sup>	–	'huit, <i>eight</i> '
Siwa	siwa	va <sup>11</sup>	–	'neuf, <i>nine</i> '

On peut constater, chaque fois que cela est possible, que les formes TK (buyang et PT) correspondent directement au PMP. Des arguments phonétiques et sémantiques le prouvent. Par exemple l'apicale t dans 'œil' et 'mourir' en TK est en accord avec PMP \*t, tandis que le PAN atteste \*C. Rappelons que les proto phonèmes PAN \*C et \*t se sont confondus dans

\*t en PMP. Les formes ‘oiseau’ et ‘cinq’ proviennent des mêmes racines en PMP, buyang et PT tandis qu’elles ont une autre origine en PAN. Ces quelques exemples suffisent à montrer que la parenté du vocabulaire partagé entre le TK et l’AN est à placer au niveau du PMP, c’est à dire à un embranchement inférieur de l’AN. Les linguistes sont aujourd’hui d’accord sur ce point. La parenté d’origine du TK et de l’AN, défendue dans l’austro-thai de Benedict, s’exclue au profit de l’explication par un contact linguistique. Pareillement, le retour sur le continent, imaginé par Sagart [2005], de l’avatar MP de l’AN est peu vraisemblable malgré une démonstration rigoureuse. Nous défendons ici l’hypothèse d’un ancien contact linguistique au cours duquel le TK aurait emprunté du vocabulaire à la branche MP de l’AN. Où et quand se serait produit ce contact ? Il faut donc imaginer un scénario cohérent capable de rendre compte de ces événements. Quel scénario de contact peut-on envisager pour expliquer les emprunts MP en TK ?

Les langues TK ont toujours été continentales, par contre les langues AN bien qu’actuellement insulaires sont originaires du continent. Le berceau d’origine et de consolidation de l’AN est à situer vers le bas Yangtsé (le fleuve bleu). De là, les langues AN se seraient propagées le long des côtes vers le sud, et probablement aussi vers le nord. C’est à partir des côtes du Fujian qu’elles ont commencé à coloniser Taiwan en plusieurs vagues dans la période 4000-3500 avant notre ère [Bellwood 2005a]. Nous pensons que le MP s’est individualisé dans le Guangdong, autour du delta de la Xijiang (la rivière des perles). C’est de là que seraient partis les navigateurs MP pour aller coloniser le nord des Philippines (voir **tableau 4**) vers 3000 avant notre ère. Ensuite, le MP du Guangdong aurait été recouvert par une poussée du TK originel venant de l’intérieur. Nous pensons que cette phase a dû se produire au moment de la conquête du sud de la Chine par les Qin (2<sup>ème</sup> siècle av. JC). Ce scénario explique pourquoi le vocabulaire commun au TK et à l’AN appartient au lexique fondamental. Par la suite, la sinisation aurait éliminé le TK des régions côtières du Guangdong sauf de Hainan. Le TK des origines seraient en fait du MP tai-kadai-isé à un degré tel que des linguistes ont pu être trompés sur son appartenance phylogénique.

Les MP installés à Luçon sont remontés vers Taïwan entre -2000 et -3000, influençant les langues formosanes mais sans faire souche dans l’île. Vers le sud, ils ont continué la colonisation des Philippines. Le laps de temps de mille ans séparant l’arrivée des premiers AN à Taiwan et l’arrivée des MP à Luçon correspond à la première longue pause observée par Blust [1999: 73-77].

#### 4. Austronésien et archéologie

Mettre en relation des faits archéologiques et des faits linguistiques a toujours été un exercice difficile voire périlleux. Nous allons analyser des faits archéologiques concernant Taiwan à la lumière de l’hypothèse de l’arrivée des MP par le sud.

La culture Dapenkeng est la plus ancienne de Taiwan, en dehors bien sûr du changpinien des chasseurs-cueilleurs. Elle est originaire du Fujian et se caractérise par le faciès de la poterie cordée et incisée (*cord-marked and incise pottery*) qui évoluera vers 2500-2000 en deux styles régionaux, le faciès de la poterie cordée (*cord-marked pottery*) et celui de la poterie lissé-rouge (*red-slipped pottery*) [Bellwood 2005b: 135-4]. La poterie lissé-rouge est supposée avoir supplanté la poterie cordée dans le sud-est de Taiwan, en particulier à Chaolaiqiao [Bellwood 2008: 25]. Constatant que la poterie lissé-rouge est également représentée aux Philippines et à Bornéo, les archéologues, toujours inspirés par le ‘Out of Taiwan’, font tout naturellement venir ce faciès de Taiwan. Se basant sur les présupposés des linguistes, ils attribuent l’expansion de la poterie lissé-rouge à l’expansion des MP de Taiwan vers Luçon. Pour nous, il est clair que la poterie lissé-rouge a été

introduite à Taiwan par le sud-est, point d'accostage naturel de la navigation venant des Philippines. Blust [1999: 76] pensait justement que Lanyu (Botel Tobago / Îles orchidées), située au sud-est de Taiwan, devait avoir été peuplée à partir des Philippines. Il y a une concordance significative entre les faits linguistiques et les faits archéologiques.

La possibilité d'une avancée timide des Formosans vers les Batanes [Bellwood 2005c: 7] et vers la plus proche des Ryukyu [Blust 1999: 76, d'après Pearson 1969] n'est pas à exclure.

Citons un argument botanique : Pejros et Schnirelman [1998: 384] notent que la variété de riz *Oriza indica* a atteint Taiwan par le sud et *Oriza japonica* par le nord. Là encore, on est amené à penser que *Oriza indica* a été introduite par les MP.

## 5. Austronésien et génétique

Les austronésianistes attendent beaucoup de la génétique – discipline en plein développement – pour confirmer leurs hypothèses. À l'instar de la linguistique et de l'archéologie, nous allons montrer que les preuves par la génétique de l'expansion AN dépendent de l'interprétation des faits linguistiques. Cependant, étant peu au fait de la génétique nous nous limiterons à quelques remarques.

Les généticiens ont observé une diversité génétique plus grande dans le nord-ouest que dans le sud-est de Taiwan [Sanchez-Mazas & als 2005]. Cette situation peut recevoir des explications opposées : (i) seule une partie des populations du nord-ouest est allée peupler le reste de l'île (hypothèse courante), (ii) les populations du sud-est sont venues par une autre voie (hypothèse défendue ici). En fin de compte, les généticiens ne peuvent apporter de réponse sûre sur le lieu d'origine des Formosans. Précisons que cette impossibilité n'a rien à voir avec la qualité scientifique des travaux des généticiens. Par ailleurs, il a été également remarqué une individualité génétique bien caractérisée de chaque groupe formosan, et une différence marquée entre les populations des montagnes du centre et celles des plaines de l'ouest [Lin & als 2005: 242] ce qui dénoterait une origine multiple des Formosans. Ce fait pourrait plaider en faveur d'une double origine, l'une par le nord-ouest (les premiers Formosans) et l'autre par le sud-est (les MP).

Il est tout à fait compréhensible que la génétique ne puisse apporter de réponse claire au peuplement de Taiwan. Les langues ne sont pas liées génétiquement aux locuteurs et les mouvements de populations sur plusieurs millénaires obscurcissent le tableau.

## 6. Hypothèse sur les premières expansions austronésiennes

En rassemblant les faits linguistiques et archéologiques, on peut proposer le schéma suivant :

- Le foyer des langues austronésiennes au néolithique doit se situer sur le continent dans la région côtière du Yangtsé. À une date indéterminée, les Austronésiens ont commencé à migrer le long des côtes vers le sud jusqu'au Guangdong, et probablement aussi vers le nord.

- Vers 4000-3500 BC, début de la colonisation de Taiwan par plusieurs vagues de petits groupes de locuteurs AN venant de différents endroits des côtes du Fujian.

- Au cours de leur migration côtière vers le sud les langues AN continentales changent et se simplifient phonétiquement. L'ancêtre du MP s'individualise dans la région du Guangdong.

- Vers 3000, les MP colonisent le nord des Philippines. Ils y introduisent la poterie lissé-rouge.

- En longeant le chapelet des îles Babuyan et Batanes, les locuteurs MP atteignent Taiwan par le sud-est. Le MP influence les langues formosanes mais ne s'impose pas dans l'île.

- Dans le même temps, les locuteurs MP colonisent les Philippines d'où partira la grande expansion vers le Pacifique. Mais là, c'est une autre histoire.

- Un peu avant notre ère, le MP continental est recouvert par la langue ancêtre du taikadai en laissant un substrat de vocabulaire fondamental. Ce vocabulaire partagé sera la cause de l'illusion d'une parenté génétique entre les deux familles.

## 7. En conclusion

Si l'hypothèse de l'arrivée des MP par le sud est valide il faudra revoir certaines idées acquises : Le 'Out of Taiwan' devra être changé en 'Out of Guangdong'. Néanmoins, il faut souligner que l'expression 'Out of Taiwan' relève d'une attitude scientifique sérieuse, contrairement à des expressions comme 'Austronesian Express-train' [Diamond 1988] et 'Taiwan's gift to the world' [Diamond 2000] qui relèvent plutôt du folklore scientifique.

En un mot, Taiwan n'est pas sur la grande voie de l'expansion austronésienne vers le Pacifique. Taiwan serait plutôt une zone refuge pour les langues, comme le Caucase, les hauts plateaux du Centre Vietnam ou l'Afrique de l'est. C'est une chance, car cette situation a permis la préservation des langues AN les plus archaïques.

Pourquoi les Formosans n'ont-ils pas poursuivi leur expansion vers les Philippines et ont été devancés par les MP ? Indirectement, une réponse a été apportée par Blust [1999: 73-77] en voulant expliquer la première longue pause. Selon cet auteur, c'est le retard des techniques de navigation qui aurait retenu les Formosans dans l'île. Les premières embarcations étaient des radeaux de bambou à voiles et des canoës monoxyles qui n'étaient pas adaptés à la navigation hauturière. L'introduction du canoë à balancier aurait permis aux Formosans de passer de Taïwan aux Philippines. Nous pensons que les nouvelles techniques de navigation ont dû être introduites par les MP, d'ailleurs le mot désignant le canoë à balancier n'est attesté que dans les langues MP. Le 'courant noir' (*black current*) qui coule vers le nord a autant facilité la navigation des MP vers Taïwan qu'il a gêné celle des Formosans dans le sens inverse.

Les contraintes techniques ne sont qu'un aspect des différences culturelles et d'autres explications sont envisageables. Les premiers Formosans étaient essentiellement des côtiers pratiquant une économie de subsistance basée sur la cueillette, la pêche et une agriculture sommaire. Ils devaient être organisés en petits groupes. Isolés sur leur île, ils n'ont pu recevoir les innovations du continent. De telles sociétés sont en général condamnées à la stagnation. Les ancêtres des malayo-polynésiens ont bénéficié des progrès techniques et agricoles. Ils devaient former des sociétés plus larges et étaient plus entreprenants. La recherche de produits rares et le commerce ont dû compter dans leur détermination à chercher des terres nouvelles. Ce ne sont que des supputations mais c'est la seule façon d'expliquer les différences d'attitude entre les Formosans et les MP.

En réalité, la véritable expansion austronésienne qui devait envahir le Pacifique jusqu'à l'île de Pâques n'a commencé qu'avec le 'Out of Guangdong', c'est à dire quelques mille ans plus tard que le peuplement de Taïwan.

Il serait souhaitable que les spécialistes des trois disciplines impliquées dans les recherches austronésiennes, la linguistique, l'archéologie et la génétique, essayent d'analyser leurs données à la lumière des hypothèses présentées ici.



## Références succinctes

- Bellwood, Peter. (1983) 'New perspectives on Indo-Malaysian prehistory', *Bulletin of the Indo-Pacific Prehistory Association* 4: 71-83.
- Bellwood, Peter. (1988) 'A Hypothesis for Austronesian Origins', *Asian Perspectives* 26(1): 107-117.
- Bellwood, Peter. (1997) *Prehistory of the Indo-Malaysian Archipelago*, revised 1985' edition. Honolulu: University of Hawai'i Press.
- Bellwood, Peter. (2004) 'The Origins and Dispersals of Agricultural Communities in Southeast Asia' in Ian Glover & Peter Bellwood (eds) *Southeast Asia from Prehistory to History*, pp. 21-40.
- Bellwood, Peter. (2005a) 'Austronesian Prehistory in Southeast Asia: Homeland, Expansion and Transformation' in *The Austronesians: Historical and Comparative Perspectives*, pp.103-118. Canberra: ANU E Press.
- Bellwood, Peter. (2005b) *First Farmers*. Oxford, Blackwell Publishing. [Chap 7: The Spread of Agriculture into Southeast Asia, 128-145]
- Bellwood, Peter & Eusebio Dizon. (2005c) 'The Batanes Archeological Project and the "Out of Taiwan" Hypothesis for Austronesian Dispersal', *Journal of Austronesian Studies* 1(1): 1-31.
- Bellwood, Peter & Eusebio Dizon. (2008) 'Austronesian Cultural Origins: Out of Taiwan, via the Batanes Islands, and onwards to Western Polynesia' in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Past Human Migrations in East Asia*, pp. 23-39.
- Benedict, Paul K. (1942) 'Tai, Kadai and Indonesian: a new alignment in south east Asia', *American Anthropologist* 44: 76-87.
- Benedict, Paul K. (1975) *Austro-Thai: Languages and Culture*. HRAF Press.
- Blust, Robert. (1988) 'The Austronesian Homeland: A Linguistic Perspective', *Asian Perspectives* 26(1): 45-67.
- Blust, Robert. (1999) 'Subgrouping, Circularity and Extinction: Some Issues in Austronesian Comparative Linguistics', in Elizabeth Zeitun & Paul Jen-kuei Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, pp. 31-94.
- Chang, Kuang-chih. *Fengpitou, Tapenkeng and the Prehistory of Taiwan*. Yale University Publication in Anthropology n° 73.
- Diamond, Jared Manson. (1988) 'Express train to Polynesia', *Nature* 336: 307-308.
- Diamond, Jared Manson. (2000) 'Taiwan's gift to the world', *Nature* 403: 709-710.
- Dyen, Isidore. (1963) 'Position of Malayo-Polynesian languages of Formosa', *Asian Perspectives* 7(1-2): 261-271.
- Fang Kuei Li. (1977). *A Handbook of Comparative Tai*. Oceanic Special Publication N° 15. The University Press of Hawaii.
- Ferrell, Raleigh. (1969) *Taiwan Aboriginal Groups: Problems in Cultural and Linguistic Classification*. Academia Sinica, monograph 17. Taiwan.
- Gray, Russell D. & Fiona M. Jordan. (2000) 'Language trees support the express-train sequence of Austronesian expansion', *Nature* 405: 1052-1055.
- Haudricourt, André Georges. (1968) 'Les arguments géographiques, écologiques et sémantiques pour l'origine des Thai', *Readings on Asian Topics, Papers read at the inauguration of the Scandinavian Institute of Asian Studies, 16-18 September 1968*, Scandinavian Institute of Asian Studies Monograph Series 1: 27-34. Copenhagen.
- Hung, H.-C. (2005) 'Neolithic interaction between Taiwan and northern Luzon', *Journal of Austronesian Studies* 1(1): 109-133. Taiwan.
- Li, Fang kuei, cf. Fang Kuei Li.
- Lin, Marie & als . (2005) 'Genetic diversity of Taiwan's indigenous Peoples' in Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas *The Peopling of East Asia*, pp. 230-247. London & New York: Routledge.
- Lu, Tracey L.-D. (2005) 'The Origin and Dispersal of Agriculture and Human Diaspora in East Asia' in Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas *The Peopling of East Asia*, pp. 51-62. London & New York: Routledge.
- Oppenheimer, Stephen J. & Martin Richards. (2001a) 'Fast trains, slow boats, and the ancestry of the Polynesian islanders', *Science Progress* 84: 157-181.

- Oppenheimer, Stephen J. & Martin Richards. (2001b) 'Polynesian origins: slow boat to Melanesia', *Nature* 410: 166-167.
- Ostapirat. Weera. (2005) 'Kra-Dai and Austronesian' in Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas *The Peopling of East Asia*, pp. 107-131. London & New York: Routledge.
- Pawley, Andrew. (2002) 'The Austronesian Dispersal: Languages, Technology and People' in Peter Bellwood & Colin Renfrew (eds) *Examining the farming/language dispersal hypothesis*, pp. 251-273. Cambridge, McDonald Institute for Archaeological Research.
- Pearson, Richard J. (1969) *Archeology of the Ryukyu Islands: A Regional Chronology from 3000 B.C. to the Historic Period*. Honolulu, University of Hawai'i Press.
- Peiros, Ilia & Victor Schnirelman. (1998) 'Rice in Southeast Asia: A regional interdisciplinarity approach' in Roger Blench & Matthiew Spriggs (eds) *Archeology and Langage II*, pp. 379-389. London: Routledge.
- Reid, Lawrence A. (1988) 'Benedict's Austro-tai Hypothesis – An Evaluation', *Asian Perspectives* 26(1): 19-34.
- Sagart, Laurent. (2004) 'The Higher Phylogeny of Austronesian and the Position of Tai-Kadai', *Oceanic Linguistics* 43(2): 411-444.
- Sagart, Laurent. (2005) 'Tai-Kadai as a subgroup of Austronesian' in Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas *The Peopling of East Asia : Putting Together Archeology, Linguistics and Genetics*, pp. 177-181.
- Sagart, Laurent. (2008) 'The expansion of *Setaria* farmers in East Asia: a linguistic and archeological model', in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Past Human Migrations in East Asia*, pp. 133-157.
- Sanchez-Mazas, Alicia, Ludmilla Osipova, Jean-Michel Dugoujon, Laurent Sagart & Estella Poloni. (2008) 'The GM genetic polymorphism in Taiwan aborigines' in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds) *Past Human Migrations in East Asia*, pp. 313-333.
- Starosta, Stanley. (1995) 'A Grammatical Subgrouping of Formosan Languages' in Paul Jen-kuei Li & als. (eds) *Austronesian Studies related to Taiwan*, pp. 683-726.
- Thurgood, Graham. (1994) 'Tai-Kadai and Austronesian: The Nature of the Relationship', *Oceanic Linguistics* 33(2): 345-368.
- Tsang Cheng-hwa. (2005) 'Recent Discovery at the Tapengkeng Culture Sites in Taiwan' in Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas *The Peopling of East Asia*, pp. 63-77. London & New York: Routledge.
- Wolf, John. (1995) 'The Position of the Austronesian Languages of Taiwan within the Austronesian Group' in Paul Jen-kuei Li & als. (eds) *Austronesian Studies relating to Taiwan*, pp. 521-583. Academia Sinica: Taiwan.
- Zeitoun, Elizabeth. (2004) 'Les langues austronésiennes: situation géolinguistique', *Faits de langues, Revue de Linguistique* 23-24: 11-21.
- Zhang Chi & Hsiao-chun Hung. (2008) 'The Neolithic of Southern China – Origin, Development, and Dispersal', *Asian Perspectives* 47(2): 299-329.

### Ouvrages collectifs :

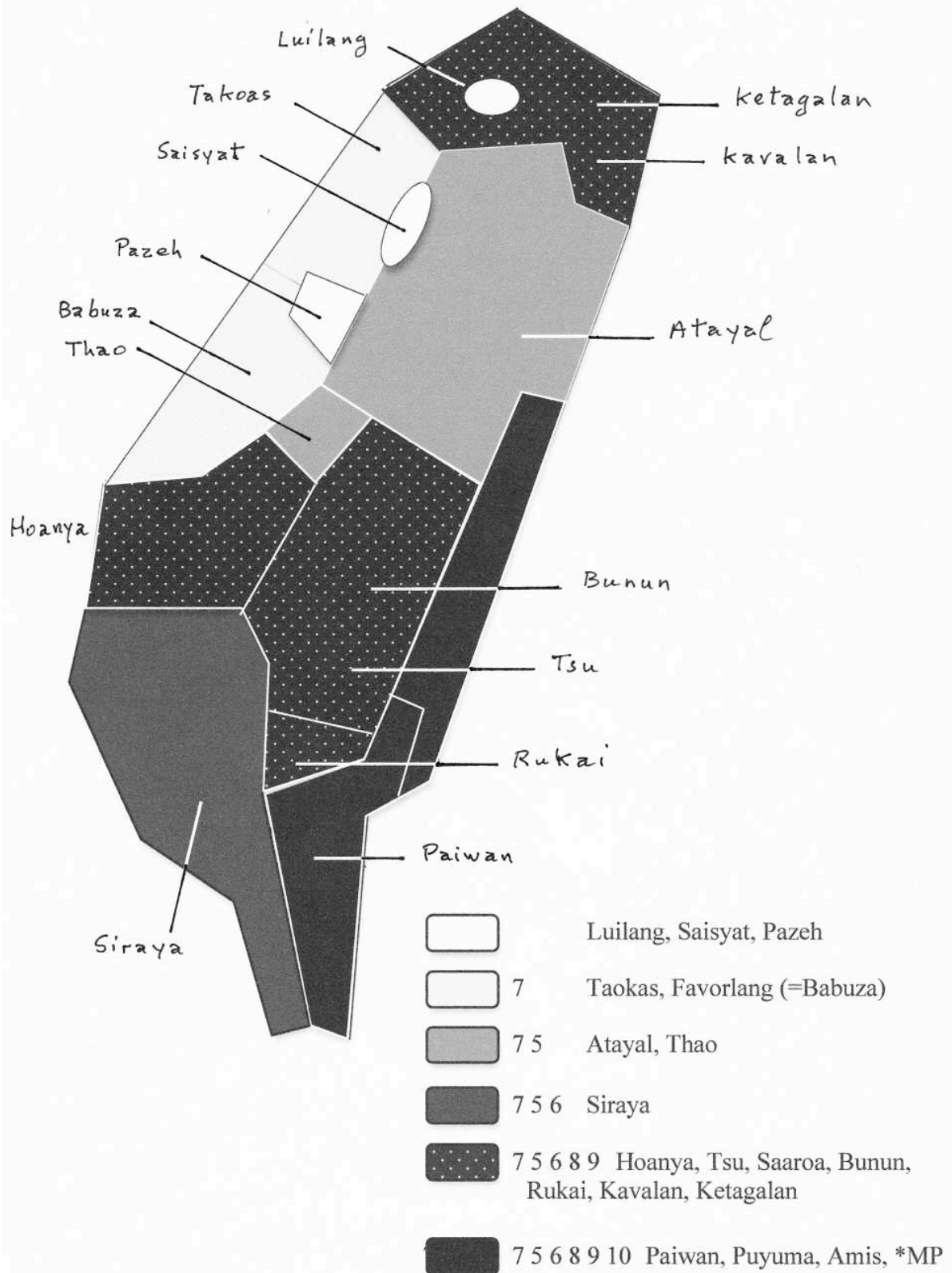
- Bellwood, Peter, James J. Fox & Darrell Tryon (ed.). (1995/2006) *The Austronesians: Historical and Comparative Perspectives*. Canberra: ANU E Press.
- Li, Paul Jen-kuei & als. (1995) *Austronesian Studies relating to Taiwan*. Symposium Series of the Institute of History and Philology, Academia Sinica 3. Taipei.
- Li, Paul Jen-kuei. (2004) *Selected Papers on Formosan Languages*, 2 vols. Languages and Linguistics Monograph series, C3. Taipei: Institute of Linguistics, Academia Sinica.
- Sagart, Laurent, Roger Blench & Alicia Sanchez-Mazas. (2005) *The Peopling of East Asia*. London & New York: Routledge.
- Sanchez-Mazas, Alicia, Roger Blench, Malcolm D. Ross, Ilia Peiros & Marie Lin. (2008) *Past Human Migrations in East Asia. Matching archeology, linguistics and genetics*. London & New York: Routledge.
- Zeitoun, Elizabeth & Paul Jen-kuei Li. (1999) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*. Taipei: Academia Sinica.

**Tableau 1** : Les nombres de ‘5’ à ‘10’ dans les langues formosanes / *Numerals ‘5’-‘10’ in Formosan languages.*

Langues	5	6	7	8	9	10
Luilang	(na)lup	(na)tsulup	innai	patulunai	satulunai	isit
Saisiat	rrasu	saiyusa	saiyuseaha	makaspat	ra:ha	ranpon
Pazeh	xasep	xasebuza	xasebidusa	xasebuturu	xasebisupat	isit
Favorlang	achab	nataap	naito	maaspat	tannacho	zchiett
Taokas	hasap	tahap	yweto	mahalat	tanasu	(ta)isid
Atayal	imagal	cziu?	pitu?	spat	qeru?	lpuu
Sediq	lima	mataro	pito	maspat	maŋali	maxal
Thao	rima	ka-turu	pitu	kahspat	tanacu	maqcin
Siraya	rima	nəm	pīttu	kuixpa	matuda	saat
Hoanya	Lima	(mi)nun	pīto	(mi)alu	(a)sia	(miata)isi
Papora	nema	(ne)nom	pitu	mahal	(me)siya	(me)tsi
Tsou	eímo	nómə	pítu	vóeu	sío	máskə
Saaroa	(k)ulima	(k)ənəmə	(k)upito	(k)ualo	(k)usia	(ku)ma:ʔə
Kanakanabu	rima	nəm	pitu	(h)a:ru	si:ya	ma:nə
Bunun	hima <sup>?</sup>	nuum	pitu <sup>?</sup>	vau <sup>?</sup>	siva <sup>?</sup>	mas <sup>?</sup> an
Rukai	Lima	eneme	pītu	valu	baŋatə	maŋeale
Paiwan	lima	enem	pitju	alu	Siva	puluq
Puyuma	Lima	nem	pitu	waLu	iwa	puLu
Amis	lima	‘enem	pitu	falu	siwa	polo
Kavalan	rima	‘nem	pitu	waru	siwa	betin
Ketagalan	tšjima	anəm	pitu	wasu	siwa	labatan
Proto M-P	*lima	*enem	*pitu	*walu	*siwa	*puluq
Proto An	*lima	*enem	*pitu	*walu	*Siwa	*puluq

*Sagart 2008, table 5.2, p. 147.*

**Tableau 2** : Répartition des innovations des nombres '5'-'10' dans les langues formosanes  
 / Distribution of changes of numerals '5'-'10' in Formosan languages



**Tableau 3** : Voie maritime naturelle de Luçon à Taiwan / *Natural sea route from Luzon to Taiwan*

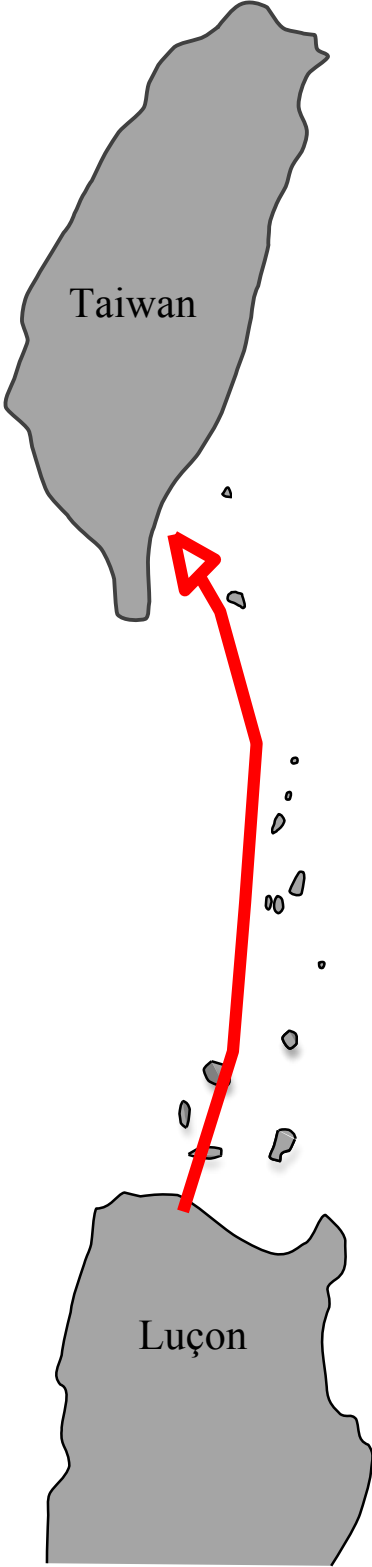
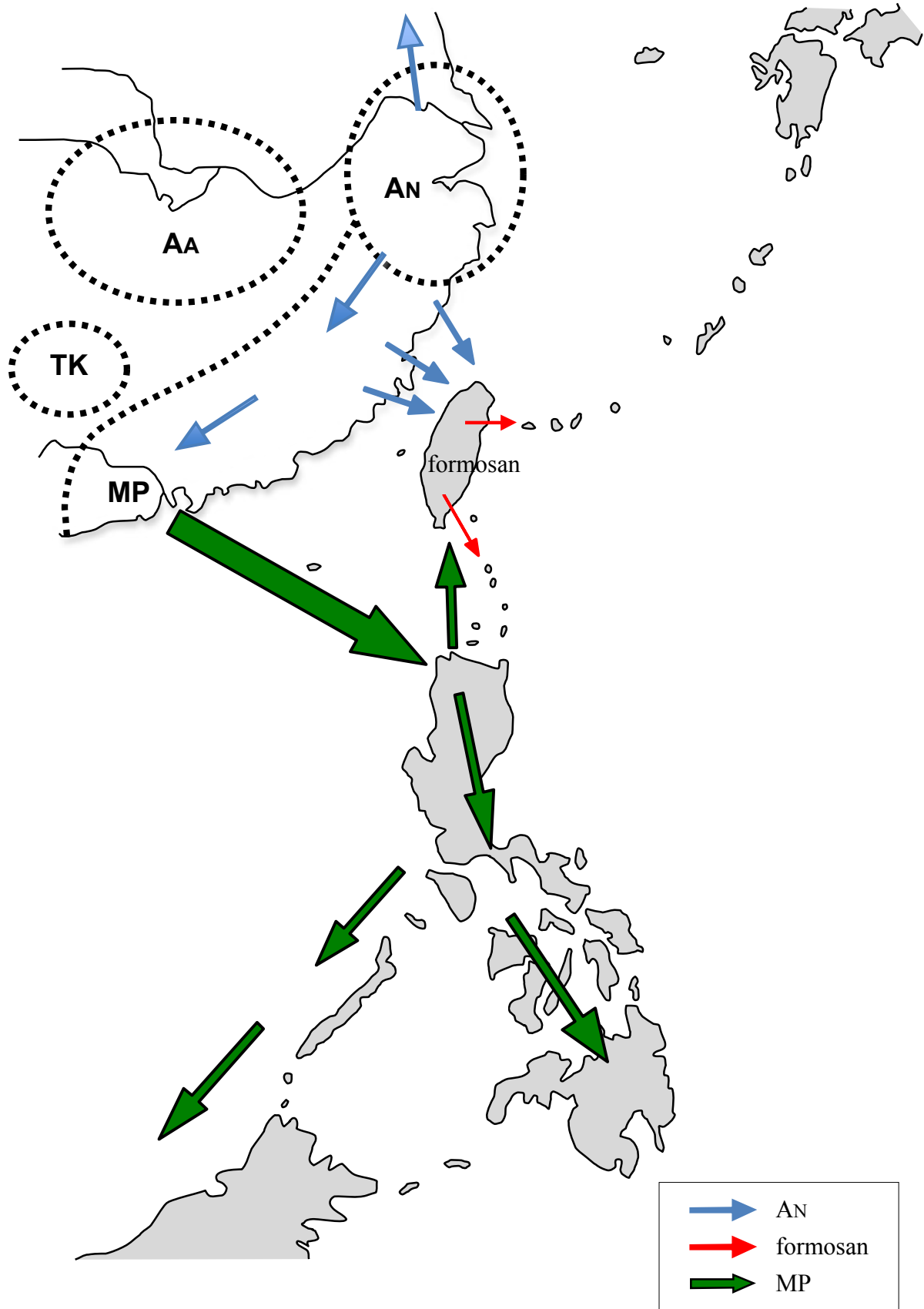


Tableau 4 : De l'AN au formosan et au malayo-polynésien / *From AN to Formosan and MP*



## The first Austronesian Expansions

### 0. Background

According to the dominant theory of linguists, Taiwan would be the center of diversification and expansion of the Austronesian languages. The other specialists, archaeologists and geneticists, fit their interpretations on the presuppositions of linguists. By differently interpreting the data put forward by the linguists, a new hypothesis on the first Austronesian expansions can be proposed.

AN	Austronesian
MP	Malayo-Polynesian
PAN	Proto Austronesian
PMP	Proto Malayo-Polynesian
TK	Thai-Kadai

### 1. State of the art

Specialists date the arrival of the first Formosan speakers from the Continent around 4000-3500 BC. The ‘Out of Taiwan’ would have occurred a bit later around 3000-2000 BC. During at least five thousand years, the original Formosan would have split into ten taxons on a restricted area, while a single taxon, the Malayo-Polynesian, would have occupied all the remainder of the Austronesian space for at least four milleniums. It is contradictory.

The phylogenic conception of AN is more or less based on the *Stammbaum* model. Our conviction is that this model is unappropriate for the Far East Asian linguistic area where contacts between languages and families of languages have a great importance.

A re-examination of the propagation of numerals in Formosan languages and the relationships between AN and Tai-Kadai, I will propose a new model of the first Austronesian expansions.

### 2. Austronesian numerals in MP and in Formosan

The treatment of numerals is significant of the differences between the phylogenic approach (numerals are inherited) and diffusionnist one (numerals are borrowed). My point of view is that the Formosan numerals were partly borrowed and are then good markers of the influences between languages. I will use the work carried out by Sagart on the genesis of AN numerals.

Laurent Sagart [2004, 2008] showed how in MP and in some Formosan languages the numerals ‘7’, ‘8’ and ‘9’ were composed as ‘5+2’, ‘5+3’ et ‘5+4’ [2004: 415-419]. Analytic forms still exist today in Pazeh, a conservative language in Northeast Taiwan. Then, Sagart [2004: 417] showed how, on the basis of explicit forms of Pazeh, the short PMP forms result.

	Pazeh		PMP
	<i>composition</i>	<i>réduction</i>	
‘5’	xaseb		(*lima)
‘6’ (5+1)	xaseb-uza	xasebuza	(*enem)
‘7’ (5+2)	xaseb-i-dusa	xasebidusa	*pitu
‘8’ (5+3)	xaseb-i-turu	xasebuturu	*walu
‘9’ (5+4)	xaseb-i-supat	xasebisupat	*siwa
‘10’	(isit)		(*puluq)

This sequence of numerals ‘5’–‘10’ reconstructed in PMP is irregularly present in Formosan (see **table 1**). It is arranged from the Northwest to the Southeast of the island according to the rank ‘7’, ‘5’, ‘6’ ‘8’ ‘9’ ‘10’ (see **table 2**).

In Sagart’s theory [2008, plate 5.1], the innovations spread from the Northwest turning around the island in a laevogyrous direction to the Southeast : on a conservative substratum the innovative form ‘7’ (**\*pitu**) was formed, then ‘5’ (**\*lima**), then ‘6’ (**\*enem**), etc. , to the Southeast area where all innovations are totalized. From there, speakers carrying these innovations left the island (around 2000 BC or before) following two directions to give rise to MP and TK. This sort of cumulative relay race of innovations was never observed anywhere in the world. The linguist as I am can only be skeptical.

By reversing the direction of the propagation, I postulate arrivals of MP speakers by the south coming from the northern part of the Philippines. **Table 2** shows a progression of innovations as a pincer movement starting from Southeast Taiwan (maximum innovations), to the northwest (where one can locate the conservative languages). Moreover, from a nautical point of view, southeast is the natural landing point from Luçon (see **table 3**).

My hypothesis is that the MP numerals from ‘5’ to ‘10’ were introduced in Taiwan to the Southeast from the North of the Philippines and spread irregularly among the Formosan languages. But, from where did the MP language established north of the Philippines come from ? It is here necessary to examine the TK family.

### 3. Tai-Kadai and Austronesian

The Tai-Kadai/Kradai family can be divided into five groups [Ostapirat 2005] :

Kra / Kadai : Gelao, Qabiao, Lachi, Buyang.

Hlai / Hli.

Kam-Sui : Kam / Dong, Mulam, Maonan, Sui, Ai-cham, Mak, Then, Lajia

Be : Ong Be / Lingao.

Tai or Tai-Yay :

Northern Tai : Buyi, Zhuang, Saek.

Central Tai : Tày, Nung.

Southwestern Tai : Ahom, Shan, White Tai, Black Tai, Tai Daeng, Lao, Siamese.

TK and AN share a short, but significant, list of core lexicon. This led Benedict [1942, 1975] to conceive his Austro-Thai phylum. Recently, Sagart [2004, 2005] annexed TK to the AN family placing the first in a lower branching, on the same level as the Formosans taxons. Against these theories, I assert that TK is an autonomous family and that the shared lexicon with AN is the result of borrowings. Where and when would this phenomenon have occurred ?

Let’s now examine some words of the shared lexicon between the AN and TK. The data are taken from Sagart [2004, 2008] and Ostapirat [2005]. Buyang is a sesqui-syllabic language of the Kra/Kadai group. Proto Tai (PT) correspond to Li’s Southwestern Tai [1977].

PAN	PMP	Buyang	PT	
maCa	mata	m <sup>o</sup> ta <sup>54</sup>	ta: <sup>A</sup>	‘œil, eye’
maCay	matay	m <sup>o</sup> te: <sup>54</sup>	taj: <sup>A</sup>	‘mourir, die’
kuCu	kutu	qa <sup>o</sup> tu <sup>54</sup>	–	‘pou, louse’
Cumay	–	ta <sup>o</sup> mɛ <sup>312</sup>	<sup>h</sup> miaj <sup>A</sup>	‘ours, bear’
(qayam)	manuk	m <sup>o</sup> nuk <sup>11</sup>	nok	‘oiseau, bird’
daNum	danum	(ʔo:ŋ <sup>42</sup> )	nam <sup>C</sup>	‘eau, water’
bulaN	bulan	–	ʔlian <sup>A</sup>	‘lune, moon’
(RaCep)	lima	ma <sup>312</sup>	–	‘cinq, five’



enem	enem	nam <sup>54</sup>	–	‘six, <i>six</i> ’
pitu	pitu	tu <sup>312</sup>	–	‘sept, <i>seven</i> ’
walu	walu	m <sup>o</sup> đu <sup>11</sup>	–	‘huit, <i>eight</i> ’
Siwa	siwa	va <sup>11</sup>	–	‘neuf, <i>nine</i> ’

One remark, when possible, that TK forms (Buyang and PT) directly correspond to PMP. Eg. TK apical **t** in ‘eye’ and ‘to die’ is in agreement with PMP **\*t**, while PAN attests **\*C**. Let’s recall PAN proto phonemes **\*C** and **\*t** merged into PMP **\*t**. Words for ‘bird’ and ‘five’ come respectively from the same roots in PMP, Buyang and PT while they have another origin in PAN. These examples are sufficient to show that the relationship of the shared vocabulary between TK and AN is to be placed on the level of PMP. I defend here the hypothesis of an ancient linguistic contact during which TK would have borrowed vocabulary from the MP branch of AN. Where and when would this contact have occurred ?

I think that MP was formed in Guangdong, around the delta of Xijiang (the Pearl River). Then, the MP would have been overflowed by a TK thrust from the inland. This stage may have occurred at the time of the Qin conquest of South China (2nd century BC).

The MP speakers settled in Luçon went up towards Taiwan between -2000 and -3000, having an influence on the Formosan languages but without rooting in the island. The one thousand years or so separating the first AN arrival in Taiwan and the arrival of MP in Luçon corresponds to the first long pause observed by Blust [1999: 73-77].

#### 4. Austronesian and archeology

Let’s analyze archaeological facts relating to Taiwan in the light of the hypothesis of the arrival of MP from the South.

The oldest culture of Taiwan is Dapenkeng originating in Fujian. It is characterized by the cord-marked and incise pottery which will evolve around 2500-2000 BC into two regional styles: the cord-marked pottery and the red-slipped pottery [Bellwood 2005b: 135-4]. The red-slipped pottery is supposed to have supplanted the cord-marked pottery in Southeast Taiwan, site of Chaolaiqiao [Bellwood 2008: 25]. As the red-slipped pottery is also represented in the Philippines and Borneo, archaeologists suppose that this style was brought from Taiwan into Philippines by the MP speakers. For us, it is clear that the red-slipped pottery was introduced in Taiwan from southeast, natural place for landing from the Philippines.

There is a significant agreement between the linguistic facts and the archaeological facts.

Botanical argument : Pejros and Schnirelman [1998: 384] note that the variety of rice *Oriza indica* reached Taiwan from the south and *Oriza japonica* from the north. I then think *Oriza indica* was introduced by MP speakers.

#### 5. Austronesian and genetics

Genetics is a rapidly developing discipline. Conclusions of geneticists on the Austronesian peopling of Taiwan depend on the linguists’ presuppositions. As in linguistics and archaeology, the facts can be interpreted in a way as well as in its opposite.

#### 6. Hypothesis on the first austronesian expansions

Putting together the linguistic and archaeological facts, we can propose the following scenario:

- The neolithic Austronesian homeland must be located in the coastal area of the Yangtse. At an unspecified date, Austronesian peoples started to migrate along the coasts towards the south to Guangdong and probably also towards north.

- Around 4000-3500 BC, beginning of the colonization of Taiwan by several waves of small groups of AN speakers coming from various places of the coasts of Fujian.

- During their coastal migration towards the south the continental AN languages changed and were phonetically simplified. The ancestor of MP language emerged in the area of Guangdong.

- Around 3000 BC, the MP speakers colonized the north of the Philippines. They introduced the red-slipped pottery.

- While going along the chain of Babuyan and Batanes islands, MP speakers reached Taiwan from the southeast. The MP influenced the Formosans languages but did not root in Taiwan.

- A little before our era, the continental MP was covered by the ancestor of Tai-Kadai languages and left a substratum of core lexicon. This shared vocabulary will be the cause of the illusion of a genetic relationship between the two families.

## 7. In conclusion

If the hypothesis of the arrival of MP from the south is valid some received ideas will have to be re-examined: The 'Out of Taiwan' will have to be changed into the 'Out of Guangdong'. Nevertheless, it should be stressed that the expression 'Out of Taiwan' concerns a serious scientific attitude contrary to expressions like 'Austronesian Express train' [Diamond 1988] or 'Taiwan's gift to the world' [Diamond 2000] which rather belong to a scientific folklore.

Taiwan is not on the great way of the Austronesian expansion through the Pacific. Taiwan would rather be a refuge area for languages. It is a chance, because this situation allowed the preservation of the most conservative AN languages. Why didn't the Formosans continue their expansion towards the Philippines and were preceded by the Malayo-Polynesians ?

Blust [1999: 73-77] explained that the first long pause was due to the necessary improvement of nautical techniques. Formosans first sailed with rafts or dugout canoes should have awaited the invention of the outrigger canoe to face the open seas.

I think that we have to consider two types of societies : The first Formosans were primarily coastal practising a subsistence economy based on gathering, fishing and elementary agriculture. They were isolated on their Island and could not receive the technical innovations from the Continent. In contrast, the MP formed broader societies and were more undertaking. The trade and search for rare products had to hope in their determination to found new lands.

The true Austronesian expansion which was to invade the Pacific until the Easter Islands only started with the 'Out of Guangdong', i.e. a few thousand years later that the settlement of Taiwan.

We'd like specialists in the three fields involved in Austronesian researches, linguistics, archaeology and genetic, to try to analyze their data in the light of the hypothesis presented above.