



**HAL**  
open science

## Le vocabulaire austroasiatique du riz : origine et expansion

Michel Ferlus

► **To cite this version:**

Michel Ferlus. Le vocabulaire austroasiatique du riz : origine et expansion. The 20th Anniversary Meeting of the Southeast Asian Linguistics Society, Jun 2010, Zurich, Suisse. halshs-01116287

**HAL Id: halshs-01116287**

**<https://shs.hal.science/halshs-01116287v1>**

Submitted on 13 Feb 2015

**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.

# Le vocabulaire austroasiatique du riz : origine et expansion

Michel Ferlus

*Chercheur indépendant (retraité du CNRS)*

## 1. Introduction

La particularité de la terminologie du riz en Asie Orientale est d'offrir une large séquence de dénominations selon la phase de préparation : “riz en général / riz sur pied”, “plant de repiquage”, “riz non décortiqué / paddy”, “riz décortiqué” et “riz cuit”. De ce paradigme, il sera retenu “riz non décortiqué / paddy” et “riz décortiqué” pour représenter les deux principaux termes de la séquence “riz brut (général, non transformé)” et “riz élaboré (décortiqué, prêt pour la consommation)”. Le comparatisme dans les langues de la famille austroasiatique (AA), en particulier mon-khmer (MK), montre qu'il existe plusieurs séquences terminologiques.

Les premières traces d'utilisation du riz mises au jour en Chine, dans la région du bassin du Yangtsé, datent d'environ dix mille ans (Pour des informations récentes sur ces problèmes, on peut consulter Fuller 2008 et Higham 2002). Avant sa domestication, le riz sauvage était cueilli et consommé par des populations de collecteurs. Cette phase de cueillette a pu s'étaler sur des millénaires. Chaque région linguistique devait avoir son propre vocabulaire pour nommer la plante. Quand la culture du riz est apparue, puis s'est propagée, les anciens collecteurs devenant progressivement agriculteurs pouvaient soit emprunter le mot “riz” à la population donneuse, soit conserver le mot de leur propre langue. Ce dernier cas peut expliquer la diversité des noms du riz. Dans la réalité il y a eu probablement combinaison des deux processus. Les perfectionnements des phases de préparation du riz, broyage ou décortiquage, cuisson à l'eau en brouet ou cuisson saturée, ont dû jouer un rôle dans l'enrichissement et la diversification terminologique. De plus, la constante régression de l'espace des langues mon-khmer sous les poussées du chinois, des langues tibéto-birmanes et en dernier des langues thai-kadai, a eu pour conséquence la disparition de nombreuses langues et une certaine concentration des langues survivantes, et donc des noms du riz, sur un espace réduit.

Quand la culture du riz s'est répandue dans des régions où la plante sauvage n'existait pas, le nom du riz de la langue donneuse s'est répandu avec la culture de la plante. Nous verrons comment dans ce cas de figure, une racine unique s'est propagée jusqu'aux extrémités de l'Eurasie.

A la suite de ces considérations, on comprend pourquoi il n'est pas possible de reconstruire une séquence unique au niveau du proto mon-khmer.

---

La première version de ce texte a été présentée à *The 20th Anniversary Meeting of the Southeast Asian Linguistics Society*, Zurich, Suisse, 10-11 juin 2010. Le texte présent est la traduction de la version améliorée publiée : “The Austroasiatic Vocabulary for Rice: its Origin and Expansion”, *Journal of the Southeast Asian Linguistic Society* 3.2: 61-76. (2010)

## 2. Les noms du riz dans les langues mon-khmer

On a répertorié la séquence “riz non décortiqué”, “riz décortiqué” et “riz cuit” dans environ quatre-vingt langues et dialectes de quinze groupes de la famille austro-asiatique. Les langues munda sont citées mais ne sont pas prises en compte. Sur ces trois termes, on n’a retenu que “riz non décortiqué” et “riz décortiqué”, fondamentalement “riz brut” et “riz élaboré”, car ces termes nous ont semblé les plus pertinents pour une étude scientifique sur l’histoire de l’émergence du riz, de sa mise en culture et de sa transformation pour la consommation. Le terme “riz cuit” relève d’une autre sémantique et n’interfère pratiquement pas avec les deux précédents.

On a pu mettre en évidence sept racines principales pour “riz brut” et deux racines pour “riz élaboré”, l’association des deux formant plus d’une douzaine de paires.

### 2.1 Paires principales “riz paddy – riz décortiqué”

On a sélectionné des exemples de paires dans les langues les plus représentatives pour notre démonstration, laissant de côté les autres pour un usage occasionnel.

Note: Les latinisations, les translittérations et certaines transcriptions anciennes sont citées en italique.

	“riz paddy, riz brut”	“riz décortiqué, riz élaboré”	
(1) khasi	<i>kba</i> [kba:]	<i>khaw</i> [k <sup>h</sup> aw]	(Bars 1973)
danaw (palaungique)	ba:	ko:	(Luce 1965)
kenieng (khamouique)	pa:	rko:	(Ferlus)
semelai (aslien)	baba:	<i>bras</i> (< AN)	(Kruspe 2004)
semnam (aslien)	ba:ʔ	bijə:n	(Burenhult 2009)
PROTO FORME	*k.ʔa:	*r.koʔ	(Diffloth 2005)
(2) mnong gar	ba:	p <sup>h</sup> ɛ:j	(Guilleminet 1959)
stieng	ba:	p <sup>h</sup> ɛ:j	(Haupers 1991)
PROTO BAHNARIQUE SUD	*ʔa:	*p <sup>h</sup> ɛ:	(Sidwell 2000)
(3) da-ang (palaungique)	<sup>h</sup> ŋɔ:	rko:	(F.)
riang (palaungique)	ŋoʔ <sup>-</sup>	koʔ <sup>-</sup>	(Luce 1965)
samtao (angkouique)	<sup>h</sup> ŋoʔ	ŋkuʔ	(F.)
*waique	hŋɔʔ	rŋkoʔ	(Diffloth 1980)
khamou	<sup>h</sup> ŋɔʔ	rŋkoʔ	(F.)
ksing mul (khamouique)	ŋɔ: kaw	həko:	(F.)
hat / ödu (khamouique)	ŋa:w	ŋkəw	(F.)
PROTO FORME	*s.ŋɔʔ	*r.koʔ	(F.)

La reconstruction de la présyllabe de \*s.ŋɔʔ est étayée par deux langues de la branche angkouique, le angkú *s’ngaw* [s.ŋaw] (Scott & Hardiman 1900) et le pesin *si ŋa* [s.ŋa] (Yán & Zhōu 1995).

(4) laven	cɛh	p <sup>h</sup> ɛ:	(F.)
lave	cɛh	p <sup>h</sup> ɛ:	(F.)
PROTO BAHNARIQUE OUEST	*cɛh	*p <sup>h</sup> ɛ:	(Sidwell 2003)

Le bahnarique ouest est très uniforme.

(5) vietnamien	<i>luá</i> [lua <sup>3</sup> ]	<i>gəo</i> [ɣa:w <sup>4</sup> ]	
muong	lɔ: <sup>3</sup>	kaw <sup>3</sup>	(F., Trần Trí Dõi)
ruc	alɔ: <sup>3</sup>	kaw <sup>3</sup>	(F., Trần Trí Dõi)
arem	ala: <sup>?</sup>	ŋkɔ: <sup>?</sup>	(F., Trần Trí Dõi)
PROTO VIET-MUONG	*c.lɔ?	*r.ko?	(F.)

En viet-muong, “riz brut” est un emprunt au chinois : chinois archaïque \***C.lu?** > chinois ancien \***dao** > *dào* 稻 “plant de riz” (Sagart 1999). Une connection à date très ancienne entre \***C.lu?** et “taro” en austroasiatique (khamou **sro?**) est possible.

(6) somre	ha:l	rək <sup>h</sup> o:	(Martin, <i>non pub.</i> )
chong	ha:j	k <sup>h</sup> o:	(Siripen 2001)
PROTO PEARIQUE	*ha:l	*r.k <sup>h</sup> o:	(Headley 1985)

## 2.2 Paires où “paddy” provient du nom d’une autre plante comestible

	“riz paddy, riz brut”	“riz décortiqué, riz élaboré”	
(7) souei	srɔ:	həŋkaw	(F.)
(8) pacoh	trɔ:	asə?	(Watson 1979)
(9) kantou	arɔ:	amə:j	(F.)
PROTO KATOUIQUE	*srɔ:	-	(Sidwell 2005)

Dans les langues katouïques la racine \***srɔ:** “riz non décortiqué” est associée à trois racines pour “riz décortiqué”, dont deux, **asə?** et **amə:j**, qui ne sont pas attestées dans les autres groupes linguistiques mon-khmer.

(10) vieux môn	<i>sro</i> ’ [sroʔ]	<i>sno</i> ’ [s.ŋuʔ]	(Shorto 1971)
môn moyen	<i>sro</i> ’ [srɔʔ]	<i>snu</i> [s.ŋuʔ]	(Shorto 1971)
môn moderne	<i>sro</i> ’/swa’ [sɔʔ]	<i>snu</i> [haoʔ]	(Shorto 1971)
nyah-kur	c <sup>h</sup> ro:ʔ	hŋkɔ: (<khmer)	(Theraphan 1984)
PROTO MON	*sroʔ	*s.ŋəʔ	(Diffloth 1984, Ferlus 1983)
(11) vieux khmer	<i>srū</i> / <i>srū</i> [sru:]	<i>raŋko</i> [rŋko:]	
khmer moderne	srɔw srūv	<i>’aŋkar</i> [ʔəŋkɑ:]	
PROTO KHMER	*sru:	*rŋko:	(Ferlus 1992)

En katouïque, khmer et môn, le terme “riz brut” provient du nom du “taro” (Ferlus 1996) ; plus précisément, le terme désignant le taro a été utilisé pour nommer le riz. Ces deux plantes, quoique dissemblables, ont une histoire commune et partagent la même niche culturelle, ce qui a rendu possible le glissement du nom du “taro” à celui du “riz”. Des botanistes pensent que le riz était une plante sauvage des fossés à taro.

(12) sora (munda)	sərɔ:	roŋko:	(Zide & Zide 1976)
-------------------	-------	--------	--------------------

Remarquons que la paire en sora est la même qu’en khmer, mais cela n’est dû qu’à un simple parallélisme et non à une origine commune. La racine \***r.ko?** “riz décortiqué” est attestée dans plusieurs langues munda (gtaʔ, remo, gutob, juang).

(13) mlabri	ju:k	ju:k t <sup>h</sup> irba:ʔ	(Rischel 1995)
-------------	------	----------------------------	----------------

Le terme générique **ju:k** rappelle le proto thai (*Southwestern Tai*) \***fuak**, lao **juak**<sup>D1</sup> (ຢອກ) “cœur (comestible) de la fausse tige du bananier”, à rapprocher du khmer **ce:k**

et du semelai (aslrien) **tyək** “banane”. Donc, en mlabri, le riz a été désigné d’après le nom du cœur de bananier. Il convient de rappeler que dans l’histoire de l’alimentation en Asie du Sud-Est, les bananiers ont été consommés pour leur rhizome, longtemps avant qu’ils le soient pour leur fruit. Le fruit du bananier sauvage est plein de graines et n’est pas consommable par les humains. Après des millénaires de reproduction par bouturage, le bananier sauvage a perdu sa capacité de reproduction sexuée, et le fruit est devenu nourriture.

(14) thin (khamouique)	sa: k <sup>h</sup> uʔ	sa: p <sup>h</sup> i:	(F.)
(15) phay (khamouique)	sa: ŋk <sup>h</sup> u:ʔ	sa: ŋk <sup>h</sup> ɔ:ʔ	(F.)
PROTO LUA	*sa:		(F.)

Les dialectes thin (Thaïlande) et phay (Laos) sont deux variantes exonymiques d’une même langue, le lua [luaʔ]. Le terme générique **sa:** (aussi dans **sa: ci:n** “riz cuit”) peut être rapproché du khamou **hrna:** (< \***srna:**) “rizière” qui en serait le dérivé infixal. Contrairement au terme “paddy” en katouique, môn, khmer et mlabri dont l’origine peut être suivie, le sens du supposé \***sa:** reste obscur.

### 2.3 Paires “riz paddy – riz décortiqué” recomposées.

Dans quelques langues mineures, les termes de base sont des composés dont le premier élément provient de “riz cuit, nourriture” dans la langue voisine.

(16) pong (viet-muong)	kə:m <sup>2</sup> təkə: <sup>3</sup>	kə:m <sup>2</sup> ko <sup>3</sup>	(F.)
------------------------	--------------------------------------	-----------------------------------	------

Le terme générique **kə:m<sup>2</sup>** (ton de série basse) est un emprunt au thô (Làng Lõ) **ɣɣ:m<sup>2</sup>** formé sur le vietnamien *còm* **kə:m<sup>1</sup>** “riz cuit” par hypercorrection fautive de l’initiale (viet **k-** interprété par **ɣ-**). Notons que viet *còm* est un emprunt au chinois *gān* 澆 “eau de rinçage du riz, cuire” (chinois archaïque \***kam** “cuire, préparer la nourriture”). Le déterminatif **təkə:<sup>3</sup>** signifie “grand, principal” (viet *cái* “mère, femme”), tandis que **ko<sup>3</sup>** provient de \***r.koʔ**.

(17) maleng brô (viet-muong)	ca:w <sup>3</sup> təkək <sup>7</sup>	ca:w <sup>3</sup> manə: <sup>56</sup>	(F.)
------------------------------	--------------------------------------	---------------------------------------	------

Le terme générique a deux origines possibles, le vietnamien *cháo* “brouet de riz”, plus probablement le lao **ca:w<sup>C1</sup>** (ຈ້ວ) “(riz) cuit à l’eau”.

### 2.4 Formes marginales

On n’a pas tenu compte de quelques langues mineures, hu, u, mən é et lua phalok (angkuic), mang et bolyu : influencées par les langues voisines ou mal situées dans les classifications, elles n’offrent pas suffisamment de possibilité pour le comparatisme. Citons-les pour mémoire (en graphies d’auteurs) :

	“riz paddy, riz brut”	“riz décortiqué, riz élaboré”	
hu	θák	ŋk <sup>h</sup> óʔ	(Svantesson 1991)
u	sáŋ	ŋk <sup>h</sup> ù	(Svantesson 1991)
mən é	hək <sup>35</sup>	taʔ <sup>35</sup>	(Yán & Zhōu 1995)
lua phalok	hək	hək ku:	(F.)
mang	lam <sup>35</sup>	bak <sup>35</sup>	(Yán & Zhōu 1995)
bolyu	te <sup>53</sup>	ʔja:ŋ <sup>55</sup>	(Edmondson 1995)

## 2.5 Le cas des langues Thai

Le vocabulaire du riz dans les langues thai est entièrement d'origine mon-khmer. Le mot général **k<sup>h</sup>aw<sup>C1</sup>** (exception: **k<sup>h</sup>a:w<sup>C1</sup>** en thai-siamois) provient de la racine **\*r.koʔ** “riz décortiqué” très répandue dans les langues mon-khmer.

	“riz non déc.”	“riz décortiqué”
thai-siamois	k <sup>h</sup> a:w <sup>C1</sup> pliak <sup>D1</sup> (ข้าว เปลือก)	k <sup>h</sup> a:w <sup>C1</sup> sa:n <sup>A1</sup> (ข้าว สาร)
lao	k <sup>h</sup> aw <sup>C1</sup> piak <sup>D1</sup> (ເຂົ້າ ເປືອກ)	k <sup>h</sup> aw <sup>C1</sup> sa:n <sup>A1</sup> (ເຂົ້າ ສານ)

Les états du riz sont nommés par des composés, **k<sup>h</sup>a:w<sup>C1</sup> pliak<sup>D1</sup>** “riz brut”, formé de **k<sup>h</sup>a:w<sup>C1</sup>** suivi du déterminant **pliak<sup>D1</sup>** “décortiquer, écorcer”. Pareillement pour **k<sup>h</sup>a:w<sup>C1</sup> sa:n<sup>A1</sup>** “riz décortiqué” où **sa:n<sup>A1</sup>** provient du proto mon-khmer **\*sa:l** “peler”, kenieng **sa:l**, khamou **ha:l** (**h-** < **\*s-**). Rappelons le saek **yaw<sup>6</sup> sa:l<sup>2</sup>** (Gedney 1993) où **\*l** final est préservé. Notons que le **r** (ຣ) final du thai-siamois écrit n'est qu'une fausse étymologie. Dans **k<sup>h</sup>a:w<sup>C1</sup> suk<sup>D1</sup>** “riz cuit”, **suk<sup>D1</sup>** signifie “cuit”.

## 2.6 Essai de synthèse

On a sélectionné sept racines principales, et leurs variantes, pour “riz brut” : **\*k.ɓa:** (aussi **\*ɓa:**), **\*s.ŋoʔ**, **\*cɛh**, **\*sroʔ** (aussi **\*sro:** et **\*sru:**), **\*ha:l** et **\*sa:** (voir la carte à la fin de l'article). La racine **\*c.lɔʔ** du viet-muong est un emprunt au chinois archaïque. Ces termes qui couvrent le riz dans son sens le plus général (plante, grain non transformé), moins spécifique que “riz décortiqué”, sont la survivance des dénominations les plus anciennes du temps où le riz n'était qu'une plante sauvage de cueillette. Par un processus universel, les états du riz travaillé ont dû être désignés par l'adjonction de déterminants au terme de base, comme on l'observe aujourd'hui en thin/phay, pong et maleng brô. L'usure phonétique naturelle des expressions d'utilisation fréquente a réduit les termes composés à leur dernier élément.

Curieusement, il n'y a que deux racines principales à l'origine des termes spécifiant l'état élaboré du riz, la forme **\*r.koʔ** et ses variantes très largement répandue dans tout le domaine mon-khmer, et **\*p<sup>h</sup>ɛ:** limitée au groupe bahnarique. Les autres termes sont trop localisés pour être pris en compte. On peut penser que ces termes signifiaient à l'origine “peler, piler, décortiquer” même si, pour l'instant, ce n'est pas confirmé dans le vocabulaire des langues mon-khmer. Signalons au passage que dans les langues munda les formes issues de **\*r.koʔ** désignent le riz non élaboré.

Le décorticage avec un pilon serait une technique découverte longtemps après la mise en culture du riz et qui se serait propagée à partir de son centre d'innovation. Cette découverte pourrait être liée à l'invention du grand pilon qui permet un bon décorticage du riz sans en écraser le grain (Ferlus 2009). Le plus ancien mode de consommation du riz devait être le brouet dans lequel le grain concassé, ou grossièrement décortiqué, était cuit dans un excès d'eau. A cet ancien mode a succédé la cuisson saturée à l'eau où le riz décortiqué et l'eau sont dans des proportions déterminées pour que les grains cuits adhèrent sans coller. Aujourd'hui, la cuisson en brouet utilise du riz décortiqué.

Si la cuisson en brouet est la plus simple et la plus ancienne, la cuisson saturée à l'eau, en revanche, nécessite un réel savoir-faire qui ne peut que résulter d'une innovation localisée dans le temps et l'espace et qui s'est propagée avec le nom du riz

décortiqué dans l'aire des consommateurs. A notre avis, l'invention du pilon, le décortilage et la cuisson saturée sont des processus liés d'origine, ce qui pourrait expliquer le faible nombre de racines mon-khmer pour “riz décortiqué”.

En définitive, l'existence de série de termes pour désigner les différents états du riz, si caractéristique des langues de l'Asie Orientale, serait la conséquence d'une familiarité millénaire entre les humains et le riz. Le rétrécissement de l'espace des locuteurs mon-khmer laisse penser qu'il a pu exister d'autres termes. Une analyse de la carte à la fin de cet article peut donner une idée de la distribution des termes exclusivement dans les derniers siècles, après l'intrusion des langues tibéto-birmane et thai-kadai. Avant cela, toute tentative de restitution ne peut être qu'hasardeuse.

Ces réflexions impliquent deux conséquences : premièrement, on ne peut reconstruire de racine unique pour chaque état du riz ; deuxièmement, la mise en culture du riz s'est produite dans l'aire des langues austroasiatiques.

### 3. La dispersion du riz et de son nom (étymon \*C.rac)

Quand le riz domestiqué a migré vers des régions où la plante sauvage n'existait pas, le nom du riz dans la langue donneuse s'est propagé avec la culture de la plante. Cette situation concerne l'étymon austroasiatique \*C.rac (C: présyllabe consonantique) dont les formes filles ont essaimé dans toute l'Eurasie (la justification de cette reconstruction sera expliquée au §.5.). Ce cas de figure remarquable contraste avec la situation dans les langues mon-khmer actuelles qui préservent une large diversité d'appellation du riz (voir § 2.1 ci-dessus).

Cette racine est attestée en chinois par *li* 糲 < chinois archaïque \*m.rats “grain brut (millet ou riz)” (Sagart 2003) et en tibétain par *'bras* “riz (terme général)”. On va examiner à grands traits la propagation de l'étymon d'origine, \*C.rac, dans les deux directions, d'abord vers l'Est, le monde austronésien, puis vers l'Ouest jusqu'à l'Europe par les plaques indiennes et iraniennes. La plupart des termes issus de \*C.rac présentent le même gabarit phonétique sesqui-syllabique malgré des évolutions phonétiques divergentes. Dans les langues formosanes de l'aire austronésienne, la vibrante médiale -r- peut se changer en l, g, h ou #. La comparaison des anciennes formes, chinois archaïque \*m.rats, tibétain *'bras*, proto austronésien \*b.Ras (voir §.3.1) et sanskrit-indus \*vrijhi (voir §.3.3) montre que la présyllabe \*C- était probablement une unité labiale.

Structure générale :

présyllabe labiale	vibrante r (ou l/g/h/#)	voyelle	finale palatale
1	2	3	4

#### 3.1 Propagation vers l'Est

L'étymon proto austroasiatique \*C.rac “riz (en général)” a évolué vers le proto austronésien \*b.Ras “riz décortiqué” (cf. Wurm 1975) qui s'est fragmenté en de multiples formes dont le malay-indonésien *běras* est la plus représentative. On donne ici une courte liste de vocables cités dans leur graphie de publication (Revel 1988) pour “riz paddy” et “riz décortiqué” suivis de leur analyse selon les quatre éléments du gabarit syllabique.

		1	2	3	4
		“riz non décortiqué”		“riz décortiqué”	
proto austronésien		<b>*b.Ras</b>	b.	R	a s
Formose: atayal	pagai	buax	bu		a x
	rukai	bərat'	bə	r	á t'
	puyuma	vərás	və	r	á s
	kavalan	bəyás	bə	y	á s
	paiwan	vat	v		a t
Luzon: tagalog	pálay	bigas	bi	g	a s
	ivatan	paray	pa	r	a y
	kallahan	begah	be	g	a h
Palawan: batak	paráy	bəgas	bə	g	a s
Visayas: hanunóo	pāray	binugas (infixé)	bi nu	g	a s
Bornéo: kadazan	parai	wagas	wa	g	a s
	kayan	bahah	ba	h	a h
malay-indonésien	<i>padi</i>	<i>béras</i>	<i>be</i>	<i>r</i>	<i>a s</i>

On peut observer dans une même langue des cas de cohabitation entre des formes filles issues de **\*b.Ras** par des voies différentes, par exemple en batak : *paráy* “riz paddy” et *bəgas* “riz décortiqué”. Dans l’aire austronésienne, “riz non décortiqué” est fréquemment *padi/paday* et on peut se demander si cette forme n’est pas aussi issue de **\*b.Ras**. Ces problèmes de phonétique historique ne peuvent être résolus dans les limites de cette étude. L’aire austronésienne est une région d’intense circulation maritime et il y est difficile de suivre les cheminements du vocabulaire. L’aire austroasiatique / mon-khmer contraste par sa stabilité terrienne.

### 3.2 Propagation vers l’Ouest : la plaque indienne

Lors de sa propagation vers l’Ouest, l’étymon **\*C.rac** a traversé la plaque indienne avec ses trois familles linguistiques (munda, dravidien et indo-aryen), où le mot a subi de nombreuses fragmentations, puis la plaque iranienne d’où il s’est répandu vers l’Occident.

Le mot est attestée en sanskrit védique (c. 1500 AEC) par *vr̥hi* < sanskrit-indus **\*vr̥jhi** “riz (général)” que Michael Witzel (1999: 13) attribue à une strate ṛgvedique para-munda acquise lors du séjour des indo-aryens au Pendjab (Indus), strate qui présente de similitudes morphologiques avec le munda. Certains spécialistes ont trop vite simplifié para-munda en munda ! La connaissance du terme “riz” en sanskrit nous dispense d’avoir à le rechercher dans les langues indo-aryennes de l’Inde. Remarquons dès à présent, et avant les développements ultérieurs, la similitude de gabarit entre proto austronésien **\*b.Ras**, indus **\*vr̥jhi** et latin *oryza*.

		1	2	3	4
proto austronésien	<b>*b.Ras</b>	b.	R	a	s
indus	<b>*vr̥jhi</b>	<i>v</i>	<i>r</i>	<i>ī</i>	<i>jhi</i>
latin	<i>oryza</i>	<i>o</i>	<i>r</i>	<i>y</i>	<i>za</i>



Les différences vocaliques ne sont pas un obstacle au rapprochement des mots. Des correspondances entre la voyelle **a** et des voyelles fermées, **i** ou **ĩ**, sont attestées dans des aires plus réduites :

waique **ʔma:r**, lamet **ma:r**, bahnarique **mi:r** “champ”  
 môn **ɗaik** (< **ɗa:k**) *ɗāk*, khmer **tik** *dik*, vieux khmer *dik* “eau”  
 môn **<sup>h</sup>laʔ** *sla*, khmer **slək** *sljk*, vieux khmer *slik* “feuille”

Les langues dravidiennes attestent deux catégories de noms du riz. Des formes complètes : tamil *ariçi* et *virīgi* (Bloch 1925), telugu *arise* (Burrow 1961) “riz pilé” qui s’accordent bien à *\*vri̯jhi*. Des formes tronquées, tamil, telugu *vari*, tamil, tulu *ari*, tulu *bār* “paddy”. Le malgache *vary* provient du dravidien (Ottino 1975). Pour expliquer la différence entre forme complète et forme tronquée, il faut se tourner vers les langues munda. On sait aujourd’hui que la différence essentielle entre les langues munda et les langues mon-khmer réside dans l’accentuation du mot (Donegan & Stampe 1983) : accent descendant en munda et accent montant en mon-khmer. Dans le sesqui-syllabe mon-khmer l’accent est montant ; la pré-syllabe est réduite et inaccentuée tandis que les composants de la syllabe principale sont pleinement réalisés. Dans le correspondant munda du sesqui-syllabe, la pré-syllabe s’allonge et devient une syllabe avec voyelle significative, tandis que la syllabe principale se simplifie par la finale. Sachant que les langues munda ont dû acquérir ces caractéristiques mélodiques au contact du dravidien, il est facile d’expliquer la genèse des formes tronquées et d’en placer les composants dans le gabarit général des mots pour “riz” depuis le sanskrit jusqu’aux langues européennes modernes.

Donnons quelques exemples tirés de Bloch (Bl.) et de Burrow (Bu.) dont les numéros sont ceux du *Dravidian Etymological Dictionary*. Selon les termes, le sens peut être “riz non décortiqué”, “riz décortiqué”, ou peut désigner une variété de riz.

		1	2	3	4
proto austroasiatic	<b>*C.rac</b>	C.	r	a	c
chinois archaïque	<b>*m.rats</b>	m.	r	a	ts
proto austronésien	<b>*b.Ras</b>	b.	R	a	s
indus	<i>*vri̯jhi</i>	<i>v</i>	<i>r</i>	<i>ĩ</i>	<i>jhi</i>
sanskrit	<i>vri̯hi</i>	<i>v</i>	<i>r</i>	<i>ĩ</i>	<i>hi</i>
hindi	<i>biryani</i> (suffixé)	<i>bi</i>	<i>r</i>	<i>ya</i>	( <i>ni</i> )
tamil (Bl.)	<i>virīgi</i>	<i>vi</i>	<i>r</i>	<i>ĩ</i>	<i>gi</i>
tamil (Bl.)	<i>ariçi</i>	<i>a</i>	<i>r</i>	<i>i</i>	<i>çi</i>
tamil (Bu. 178)	<i>arici</i>	<i>a</i>	<i>r</i>	<i>i</i>	<i>ci</i>
tamil, telugu (Bu. 4306)	<i>vari</i>	<i>va</i>	<i>r</i>	<i>i</i>	
telugu (Bu. 2991)	<i>nivari</i> (préf.)	( <i>ni</i> ) <i>va</i>	<i>r</i>	<i>i</i>	
tamil, tulu (Bu. 178)	<i>ari</i>	<i>a</i>	<i>r</i>	<i>i</i>	
tulu (Bl.)	<i>bār</i>	<i>bā</i>	<i>r</i>		
malgache	<i>vary</i>	<i>va</i>	<i>r</i>	<i>y</i>	

On peut constater une diversification phonétique poussée en dravidien, mais, contrairement aux langues austronésiennes, la médiale **-r-** est bien préservée tandis que le mot s’est simplifié par la finale dans quelques langues.

### 3.3 Propagation vers l'Ouest : la plaque persane

Dans son article, ancien mais bien documenté, Jules Bloch (1925) rattache l'ancien grec *oruza* ὄρυζα, moderne *ruzi* ρύζι, à l'ancien persan qu'il reconstruit \**wrinjhi*. Le persan moderne *birinj* est passé dans toutes les langues du Caucase.

Le mot s'est répandu dans les langues européennes à partir du grec. Le gabarit syllabique est toujours respecté ; parfois le premier élément peut s'effacer.

		1	2	3	4
proto austroasiatique	* <b>C.rac</b>	C.	r	a	c
chinois archaïque	* <b>m.rats</b>	m.	r	a	ts
proto austronésien	* <b>b.Ras</b>	b.	R	a	s
indus	* <i>vrījhi</i>	v	r	ī	jhi
sanskrit	<i>vrīhi</i>	v	r	ī	hi
ancien persan	* <i>wrinjhi</i>	w	r	i	njhi
persan	<i>birinj</i>	bi	r	i	nj
ancien grec	<i>oruza</i>	o	r	u	za
grec	<i>ruzi</i>		r	u	zi
latin	<i>oryza</i>	o	r	y	za
espagnol, portugais	<i>arroz</i>	a	rr	o	z
italien	<i>riso</i>		r	i	so
ancien français	<i>ris</i>		r	i	s
français	<i>riz</i> [ri]		r	i	z
allemand	<i>reis</i>		r	ei	s
néerlandais	<i>rijts</i>		r	ij	ts
anglais	<i>rice</i> [rais]		r	i	ce

Dans les langues du Caucase, Jules Bloch (1925: 46) signale en outre des termes concernant des céréales comestibles qui présentent le même gabarit : andi *perinj*, agoul *burunz*, géorgien *brič* “grau (de céréale)” ; scythe ancien *wrujaka*, “seigle”.

### 3.4 Quelles voies de propagation ?

Comment et par quelle voie un nom du riz, de l'étymon \***C.rac**, originaire du centre de la Chine, a-t-il été emprunté par l'indo-aryen du Pendjab au deuxième millénaire avant notre ère ? On pourrait penser que le mot a été transmis par le munda, branche de l'austroasiatique, mais cette racine n'y est pas attestée (cf. Zide 1976: 1299-1310) sinon marginalement comme le santali *baḍraṣ* “variété de riz de montagne”. Quand des comparaisons sont possibles, le vocabulaire munda du riz montre une parenté avec celui de la branche mon-khmer. La transmission n'a donc pas pu se faire par le munda. En revanche, l'étymon \***C.rac** est présent en dravidien et la grande variété des formes filles montre son ancienneté dans cette famille linguistique. Nous pensons que c'est le dravidien qui a été l'intermédiaire dans la propagation du nom du riz entre l'austroasiatique du centre de la Chine et l'indo-aryen du Pendjab. Le rôle du dravidien a été occulté par son recul devant les poussées du munda et de l'indo-aryen.

Une autre solution est possible, quoique moins probable, le nom du riz aurait pu voyager par le nord de l'Himalaya en traversant le territoire du Tibet.

L'Inde est très riche en variétés sauvages de riz et pendant longtemps les botanistes ont cru qu'elle était le berceau le plus ancien de la domestication du riz

avant que cette place ne soit reconnue à la Chine (de Candolle 1883: 309-311). Cependant, le riz cultivé en Inde est la variété *Oriza sativa indica*, tandis que celui cultivé en Chine est l'*Oriza sativa japonica*. Si le nom du riz en Inde est originaire de Chine, cela signifie que sa dénomination est passée d'une variété sur une autre. Ce processus est bien connu en sémantique historique. Il n'est pas rare qu'un terme change de sens pour un désigné voisin. On a vu plus haut comment en katouique, en môn et en khmer le nom du taro est devenu celui du riz. Dans le Caucase, le nom persan du riz peut devenir celui du seigle. En occitan le nom du maïs *mil* provient du nom du millet. Plus surprenant, en mlabri (langue khamouique), le terme général **ju:k** pour riz rappelle le proto thai **\*fuak**, lao **juak**<sup>D1</sup> "cœur (comestible) du bananier".

Pour conclure, nous pensons que les populations pré-néolithiques de l'Inde devaient consommer l'*indica* sauvage de cueillette ; la culture du riz et sa dénomination seraient venues de populations austroasiatiques d'avant l'expansion des langues munda.

La propagation du nom du riz depuis l'aire iranienne jusqu'à l'Europe est assez difficile à suivre. Les tableaux présentés ici ne font que comparer les formes et ne prétendent pas en tracer la route avec précision. Il y a déjà plus de trois quarts de siècle, Jules Bloch (1925: 38) écrivait :

Les noms européens et presque tous les noms sémitiques du riz se rattachent au grec. D'où venait le mot grec lui-même? A cette question, la réponse qui se présente d'abord est que le mot a dû venir avec le grain, d'Extrême-Orient, et sans doute plus particulièrement de l'Inde, le premier pays d'Extrême-Orient que les anciens ont connu, où d'ailleurs le riz pousse à l'état sauvage et a été cultivé dès une haute antiquité.

Aujourd'hui, le problème est toujours au même point.

#### 4. Autres connections possibles

Il y a des correspondances possibles entre l'étymon **\*C.rac** et des noms de céréales dans certaines langues indo-européennes.

Proto indo-européen **\*p<sup>h</sup>ūr** "blé", grec *pūrós* "grain de blé", sanskrit *pūra* "sorte de pâte". Anciens dialectes indo-européen **\*wru<sup>h</sup>yo-** "seigle", vieil anglais *ryge*, anglais *rye*, vieux haut allemand *rokko*, allemand *roggen* (Gamkrelidze 1995: 564-8). Toutes ces attestations rentrent assez bien dans le gabarit de l'étymon.

		1	2	3	4
proto austroasiatique	<b>*C.rac</b>	C.	<b>r</b>	<b>a</b>	<b>c</b>
proto indo-européen	<b>*p<sup>h</sup>ūr-</b>	p <sup>h</sup> ū	r	-	
grec	<i>pūrós</i>	pū	r	ó	s
sanskrit	<i>pūra</i>	pū	r	a	
anciens dialectes i.-e.	<b>*wru<sup>h</sup>yo-</b>	w	r	u	g <sup>h</sup> yo-
scythique	<i>wru<sup>h</sup>aka</i>	w	r	u	ǰaka
vieil anglais	<i>ryge</i>		r	y	ge
vieux haut allemand	<i>rokko</i>		r	o	kko

Ces comparaisons peuvent paraître audacieuses, mais il faut remarquer que les différences phonétiques et sémantiques impliquées sont de même nature que celles observées dans l'aire dravidienne. Nous ne faisons qu'exposer le problème.

## 5. Origine du nom du riz de l'étymon \*C.rac

Nous allons à présent nous expliquer sur le choix de \*C.rac pour représenter l'étymon primordial du nom du riz qui s'est répandu hors de son foyer cultigène.

Remarquons d'abord que la finale **-c** (dans **-p -t -c -k**), fréquente en mon-khmer, correspond structurellement à **-ts** (dans **-p -t -ts -k**) du chinois archaïque. Les sinologues interprètent **-ts** comme **-t** + *suffixe -s*. A notre avis, la combinaison **-ts** est bien trop fréquente pour pouvoir être expliquée uniquement par la suffixation. Il est probable que le chinois connaissait la finale **-c** qui se serait confondue très tôt avec **-ts**. Toutefois, on ne peut savoir si cette finale supposée est héritée ou empruntée.

Le plus ancien mode de cueillette du riz consiste à prélever les grains en serrant l'épi dans la main et en tirant vers le haut ; le procédé survit chez certaines minorités montagnardes de l'Asie du Sud-Est. Dans plusieurs langues mon-khmer cette action précise est désignée sur la base de la racine \*ru:c : proto pong \*pru:c<sup>7</sup>, kenieng ru:c, chong rù:c, muong k<sup>h</sup>u:c<sup>7</sup> < \*kru:c, arem aruəc, vietnamien rút "retirer, extraire" ; le lao ສູນ hu:t<sup>8</sup> < \*ru:t (pas de palatale finale dans les langues thai-kadai) est un emprunt au mon-khmer. Il n'y a pas de mot dans les langues européennes car le procédé n'est pas praticable avec le blé dont l'épis adhère solidement à la tige.

D'une manière générale, on peut observer une tendance en mon-khmer à nommer les actions d'enlèvement ou de raclage le long de choses longues par des mots en initiale **r-** et en finale palatale **-c**. Par exemple, on a en khamou : ra:c "déchirer", <sup>h</sup>ra:c "enlever l'écorce d'un arbre", <sup>h</sup>ro:c "effiler, amincir des liens", ro:c "vider des boyaux (en pinçant entre les doigts)", etc.

Des exemples plus précis vont justifier notre reconstruction \*C.rac "paddy" sur laquelle est bâtie une partie de cette étude. Le mlabri (Nan, Thaïlande) atteste krec, dans krec ju:k "moissonner" (Rischel 1995). Or cette langue a connu le changement a:/a > ε:/ε pour une partie de son vocabulaire, par exemple :

<i>Khamou</i>	<i>Mlabri</i>	
ma:m	mε:m	"sang"
ba:r	bε:r	"deux"
ga:ŋ	gε:ŋ	"maison"

Ces correspondances nous autorisent à reconstruire \*rac à un niveau plus ancien du mlabri. Le nyah kur, langue proche parente du môn, atteste ręc "moissonner". En comparant cette forme avec le môn ròt rat "moissonner, couper (des céréales)" (Shorto 1962) on a pu reconstruire régulièrement \*rac (Diffloth 1984, Ferlus 1983).

La reconstruction \*rac, en mlabri et en môn, correspond à la syllabe principale de l'étymon primordial \*C.rac, et ce malgré l'écart de temps qui les sépare. Le fait que \*rac désigne la moisson à la faucille ne fait que traduire l'évolution des techniques.

Résumons quelques correspondances :

		1	2	3	4
proto austroasiatique	*C.rac	C.	r	a	c
mlabri	*rac (> ręc)		r	a	c
proto môn	*rac (> ròt)		r	a	c
chinois archaïque	*m.rats	m.	r	a	ts
proto austronésien	*b.Ras	b.	R	a	s

sanskrit	<i>vr̥hi</i>	<i>v</i>	<i>r</i>	<i>ī</i>	<i>hi</i>
persan (ancien)	<i>*wriṅḥi</i>	<i>w</i>	<i>r</i>	<i>i</i>	<i>ṅḥi</i>
latin	<i>oryza</i>	<i>o</i>	<i>r</i>	<i>y</i>	<i>za</i>

En conclusion :

Le rapport de **\*rac** “cueillir les grains à la main le long de la tige” à **\*C.rac** “riz (général)” est celui d’une base verbale à son dérivé nominal. Ce rapprochement nous permet d’affirmer que la signification d’origine de l’étymon **\*C.rac** n’était rien d’autre que : “céréale cueillie en arrachant les grains le long de la tige avec les mains”.

## RÉFÉRENCES

- Bloch, Jules. 1925. Le nom du riz. *Etudes Asiatiques*, publiées à l’occasion du 25<sup>e</sup> anniversaire de l’EFEO (PEFEO 19): 37-47. Paris: G. Van Oest.
- Burrow, T. & M.B. Emeneau. 1961. *A Dravidian Etymological Dictionary*. Oxford: Clarendon Press.
- de Candolle, Alphonse (1883). Origine des plantes cultivées. Paris: Germain Baillière et Cie. [Reprint: Editions Jeanne Laffitte, Marseille]
- Diffloth, Gérard. 2005. The contribution of Linguistic Palaeontology to the Homeland of Austro-Asiatic. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds), *The Peopling of East Asia*, pp. 77-80.
- Donegan, Patricia Jane & David Stampe. 1983. Rhythm and the Holistic Organization of Language Structure. *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Linguistic Department of the Ohio State University.
- Ferlus, Michel. 1988. Les langues austroasiatiques. In Nicole Revel (sous la direction de), *Le riz en Asie du Sud-Est (Atlas du vocabulaire de la plante)*, 3 livrets, 310 p. + 64 p. + 78 cartes. Paris: École des Hautes Études en Sciences Sociales. Articles: 81-94, Vocabulaires: 327-31, Cartes n° 26 à n° 34.
- Ferlus, Michel. 1996. Du taro au riz en Asie du Sud-Est, petite histoire d’un glissement sémantique. *Mon-Khmer Studies* 25: 39-49.
- Ferlus, Michel. 2009. A Layer of Dongsonian Vocabulary in Vietnamese. *Journal of Southeast Asian Linguistic Society* 1: 95-109.
- Ferrell, Raleigh. 1969. Taiwan Aboriginal Groups: *Problems in Cultural and Linguistics Classification*. Academia Sinica Monograph 17. Taiwan.
- Fuller, Dorian Q., Ling Qin & Emma Harvey. 2008. Evidence for a late onset of agriculture in the Lower Yangtse région and challenges for an archeobotany of rice. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin (eds), *Past Human Migrations in East Asia*, pp. 40-83. London & New York: Routledge.
- Gamkrelidze, Thomas V. & Vjačeslav V. Ivanov. 1995. *Indo-European and the Indo-Europeans*, part I *The Text*, part II *Bibliography, Indexes*. Berlin / New York: Mouton de Gruyter.
- 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.

- Harlan, J.R. 1987. *Les plantes cultivées et l'homme*. Presses Universitaires de France.
- Higham, Charles. 2002. Languages and Farming Dispersals: Austroasiatic Languages and Rice Cultivation. In Peter Bellwood & Colin Renfrew (eds), *Examining the Farming / Language Dispersal Hypothesis*, pp. 223-232. University of Cambridge: McDonald Institute for Archeological Research.
- Ottino, Paul. 1975. L'origine dravidienne du vocabulaire du riz et de certains termes de riziculture à Madagascar. *Annuaire des pays de l'Océan Indien* 2: 103-121.
- Pejros, Ilia & Victor Schnirelman. 1998. Rice in Southeast Asia: A regional interdisciplinary approach. In Roger Blench & Matthiew Spriggs (eds), *Archeology and Language II*, pp. 379-389. London: Routledge.
- Sagart, Laurent. 1999. *The Roots of Old Chinese*. Amsterdam/Philadelphia: John Benjamins.
- Sagart, Laurent. 2003. The Vocabulary of Cereal Cultivation and the Phylogeny of East-Asian Languages. *Bulletin of the Indo-Pacific Prehistory Association* 23 (Taipei papers, vol. 1): 127-136.
- 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. London & New York: Routledge.
- Witzel, Michael. 1999. Substrate Languages in Old Indo-Aryan (R̥gvedic, Middle and Late Vedic). *Electronic Journal of Vedic Studies* 5(1): 1-67.
- Wurm, S.A. & B. Wilson. 1975. *English Finderlist of Reconstructions in Austronesian Languages*. Pacific Linguistics C-33. Canberra: The Australian National University.
- Zide, Arlene R.K. & Norman H. Zide. 1976. Proto-Munda cultural vocabulary: evidence for early agriculture. In Philip N. Jenner & als (eds) *Austroasiatic Studies* part II, pp. 1295-1334. Honolulu: University of Hawaii Press.

#### **OUVRAGES COLLECTIFS**

- Bellwood, Peter. 2005. *First Farmers*. Oxford: Blackwell Publishing. [Chap 7: The Spread of Agriculture into Southeast Asia, 128-145]
- Revel, Nicole (ed.). 1988. *Le riz en Asie du Sud-Est (Atlas du vocabulaire de la plante)*, 3 livrets: articles, vocabulaires et planches. Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- 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.

## Les sept racines principales pour “riz non décortiqué” dans les langues mon-khmer

