



HAL
open science

Old Chinese and Proto-Austronesian

Laurent Sagart

► **To cite this version:**

Laurent Sagart. Old Chinese and Proto-Austronesian. CAMAC (Conference on Asia-Mainland/Austronesian Connections), May 1993, Hawai'i, Honolulu, United States. hal-00614416

HAL Id: hal-00614416

<https://hal.science/hal-00614416>

Submitted on 11 Sep 2011

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.

CONFERENCE ON ASIA-MAINLAND/AUSTRONESIAN CONNECTIONS
University of Hawai'i at Manoa, May 10-13, 1993

OLD CHINESE AND PROTO-AUSTRONESIAN

Laurent Sagart

Centre de Recherches Linguistiques sur l'Asie Orientale, Paris,
and
Social Science Research Institute, Honolulu.

This paper presents an updated argument for a genetic relationship between Old Chinese (OC) and Proto-Austronesian (PAM). In response to criticism of my earlier work, a more restrictive approach to lexical comparison has been adopted. Particular attention has been paid to the Formosan material. Evidence of congruence in the domains of morphology (§. II), the basic lexicon (§. III), and sound correspondences (§. IV, V), as well as some elements of extra-linguistic evidence (§. VI) will be presented.

I. BACKGROUND

A. PROTO-AUSTRONESIAN.

Proto-Austronesian, the ancestor language of all living AM languages, is believed by a majority of Austronesianists to have been spoken on Formosa or neighboring areas of the China coast (see Blust 1988b for evidence from linguistic paleontology). Even though Formosa may well be the location from which the AM language family began its outward spread, the earliest AM speakers in Formosa would have migrated there by boat from the Chinese coast: Chang (1989) equates the speakers of PAM with the neolithic cultures of Ta-Pen-K'eng in northern Taiwan and Fu-Kuo-Tun on Quemoy island (a few miles off the coast of Fujian), both C14-dated to the late 5th millenium BC; Bellwood (1991) cites Hemudu in Hangzhou Bay, about 1000 kms. north of the northern tip of Taiwan, as representative of the type of neolithic society of the China coast which could have sent settlers to Taiwan. The earliest C14 dates for Hemudu are slightly anterior to 5000 BC. Since outside Taiwan the earliest dates which may be associated with AM-speaking cultures are posterior to 3000 BC, the first migration outside Taiwan, and the subsequent break-up of PAM into its daughter languages, may be dated to approximately 3000 BC (Bellwood 1991). Because the comparative method reconstructs proto-languages as they appeared immediately before the time they began to diverge, the date of 3000 BC is thus taken as representative of reconstructed PAM.

A proper understanding of subgrouping is essential in deciding which words are assignable to the proto-language. I will assume the view of higher-order Austronesian subgrouping argued for in Blust (1988a: 16-7):

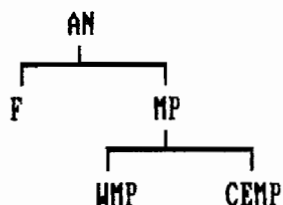


Fig. 1: higher-order subgrouping of AN languages, after Blust (1988:17, simplified). F = Formosan: one or more primary subgroups in Taiwan; MP = Malayo-Polynesian: all extra-Formosan AN languages; WMP: Western Malayo-Polynesian: the MP languages of Borel Tobago, the Philippines, the Marianas, Palau, Sulawesi and satellites, Borneo, Sumatra, Java, Bali, Lombok and Western Sumbawa, Malaya, Indochina and Madagascar; CEMP = Central-Eastern Malayo-Polynesian: all other MP languages.

In Blust's work, reconstructions are made at different levels: proto-WMP (level 3), proto-MP (level 2), and PAN (level 1). We will be concerned mostly with those items that are reconstructible at the highest level (PAN). Although Blust is not committed as to whether the Formosan languages form one or three primary branches of AN (Tsouic, Atayalic, Paiwanic), in practice he treats Formosan as a single witness in determining the level of assignment of reconstructed items, due to the presumed antiquity of contact and importance of mutual borrowings between Formosan languages. As a result, only items attested simultaneously in and outside of Formosa qualify for the PAN level. I follow that procedure.

Reconstructed AN items are either words or 'roots'. 'Roots' are defined (Blust 1988) as meaning-associated submorphemic strings, of CV(C) shape -mostly CVC-, occurring as the final syllable of at least four reconstructed or isolated words in AN languages. Blust (1988a) lists 231 such roots. While comparison routinely makes use of reconstructed words, use of roots is less well established. Yet there seems no reason to exclude AN roots from external comparison. On the contrary, roots (at least those which are non-onomatopoeic and can be shown to belong to the PAN level) represent some of the most characteristic sound-meaning associations in the language. Blust (1988:54) suggested that "attention to the root could open up new possibilities in exploring the external relationships of the Austronesian languages". I follow his suggestion.

Although in Blust's published work only reconstructed words are assigned to different reconstruction levels, the same can be done meaningfully with roots. I regard a root as PAN if it includes both Formosan and non-Formosan items. I take it that a PAN root occurred as the final syllable of at least one PAN word of the same meaning. Since only the last syllable of AN words is reflected in Chinese, comparing a PAN root with a Chinese word is not very different from comparing a PAN word with a Chinese word.

Only reconstructions assignable to the earliest proto-language, PAN, will be used as comparanda in this paper. The cognate sets and reconstructions are mainly drawn from the literature, in particular the lists in Blust (1980, 1983-84, 1985, 1989) and Tsuchida (1976). In a few instances, reconstructed items assignable to lower-level languages (such as Proto-Malayo-Polynesian, Proto-Western Malayo-Polynesian, Proto-Philippine, etc.) because they seemed to lack Formosan cognates, have been raised to the PAN level by myself after one

or more Formosan cognates were identified (see in section III the cognate sets under *siraQ1 'salt', *SulaR 'snake', *tuDuR 'sleep', *liqeR 'neck'; *kupit 'shut, close', *puluh 'gather'); a 'candidate root': -baŋ 'broad', 'wide', has likewise been elevated to the status of PAN root. A new PAN root (as far as I know) has been identified: -qem 'cloud, cloudy'.

B. OLD CHINESE

The oldest known texts written in a form of Chinese are the oracular inscriptions of the late Shang dynasty (ca. 1300-1100 BC). It was formerly believed that the Shang civilization, and the Chinese civilization in general, ultimately stemmed from a more westerly neolithic culture, the Yangshao tradition, whose bearers had gradually expanded eastward in early neolithic times (Chang 1968). However, as a series of neolithic sites as ancient as Yangshao were discovered on the Pacific seaboard, notably at Dawenkou, Qingliangang and other sites in south Shandong and north Jiangsu, it became clear (in agreement with the Shang tradition of their eastern origins) that a continuity exists between these sites, in particular Dawenkou, and the Shang civilization (Chang 1980: 347, 354). Chang (1980: 345) listed continuities in material culture between Shang and the Lungshanoid cultures on the coast of Shandong, and suggests the possibility that the rulers of the Shang dynasty could have been a conquering group from the East. In line with Chang's argument, pottery signs at Dawenkou have been identified as early Chinese characters (Qiu Xigui 1978).

Although identification of the Shang oracular language as an early form of Chinese is certain, the phonetic information carried by the oracular script is too scanty for reconstruction. OC as a reconstructed language represents the language of the early and mid Zhou dynasties, ca. 1100-600 BC.

Two OC reconstructions are given for each Chinese word: Li Fang-kuei's (Li 1971, 1976) and, following Li's between round brackets, a slightly modified version of Baxter's system (Baxter 1992, with emendations as outlined in my forthcoming review in *Diachronica*, Fall 1993). Baxter's system includes much of the progress made in Old Chinese since Karlgren: the distinction between l- series and t- series, the division 2 -r- paradigm, and the elimination of final voiced stops are among the most conspicuous. The same features have been selected in several recent reconstruction systems independently elaborated by scholars in China (Zhengzhang), the ex-USSR (Starostin) and the USA (Schüssler). Baxter's system also incorporates his own statistically-based reanalysis of the rhyming system. It is in my opinion the best approximation of OC pronunciation currently available. In what follows, the sound correspondences between OC and PAN will be formulated in terms of Baxter's system, with my modifications. Li's transcriptions are provided for comparison.

C. GENETIC RELATEDNESS.

It is generally considered that morphology is stable and non borrowable (unless when affixed words are borrowed), because it is internally strongly structured and resistant to change, both internal and external. For that reason the diagnostic value of morphological congruence in determining genetic relationships is particularly high. Sound correspondences by themselves do not necessarily spell genetic relationship, since short, intense episodes of language contact (say, during a colonial episode) may result in one language borrowing a set

of words from the other, with sound correspondences; the words borrowed in such contexts are, however, usually of a non-basic, cultural kind. Conversely, the sharing of elements of the basic vocabulary without sound correspondences may indicate borrowing in situations of prolonged and intimate contact. Sound correspondences are a strong proof of genetic relationship if they can be shown to exist between elements of the basic vocabulary: the presence of sound correspondences then argues against a prolonged contact situation, and the basic vocabulary, against a short contact situation: only genetic inheritance remains as a possible explanation for the simultaneous presence of both.

It might be argued that the time depths involved are too great for sound correspondences to be still observable between the daughter languages of a proto-language so old as the one I have in mind. Nothing could be further from the truth. If the proto-language ("Proto-Sino-Austronesian", or PSAN) was spoken around 5000 BC., when rice cultivation and timber houses, both known to the speakers of PSAN (see §. III), begin to appear in the archaeological record, some 6000 years of cumulated linguistic history are needed for the development of PSAN into PAN and OC: a time depth slightly less than that of the Indo-Iranian branch of IE.

It might also be argued that to rely exclusively on semantic proximity and the regularity of sound correspondences in making cognate decisions is an overly mechanistic approach to sound change as implemented in the lexicon: etymological dictionaries of languages with a well-documented history are teeming with sporadic processes (analogies, hypercorrections, contaminations, blends, onomatopoeic retentions, and so forth) which compete with regular sound change in determining the shape of certain words. While this is true in principle, I cannot see how, especially in the case of languages without a written history, such sporadic developments can be detected, if not against the backdrop of a well-understood history of phonetic changes. There is no serious methodological alternative to requiring regular sound correspondences in comparative work.

I shall then present these two kinds of linguistic evidence: morphological congruence (section II) and sound correspondences in the basic vocabulary (section IV).

II. MORPHOLOGICAL CONGRUENCE BETWEEN OC AND PAN

In general, OC and PAN agree in possessing a mostly derivational morphology making use of prefixes, infixes and suffixes. Unlike Dakota, whose infixes are inserted after the first vowel of a stem (Moravcsik 1977: 113-4), or PIE, whose 'nasal infix' was inserted before the last segment of a root (Meillet 1937: 215-6), both PAN and OC insert their infixes before the first vowel of a word stem.

Evidence is presented below that at least one prefix and two infixes are shared by OC and PAN.

A. PAN ~~-ar-~~ : OC ~~-r-~~.

PAN and OC share a derivational infix (PAN ~~-ar-~~, OC ~~-r-~~) whose core function was to derive verbs of distributed action and nouns of distributed objects.

1. PAN -ar- 'distributed action' : OC -r- 'distributed action'

'Distributed action' is defined as temporally or spatially distributed: the action either occurs repeatedly (true iterative), or its intensity undergoes rapid variations during the course of the event, or the action occurs in several locations, either at once or in succession. Onomatopoeia for sounds in which a certain acoustic pattern is repeated, or sounds coming from several sources at once, or sounds coming from a moving source, fit naturally here. Derived verbs of distributed action are characteristically more kinetic than their word-base.

A verbal infix /ar/ 'distributed action' is reconstructible at the PAN level from Formosan and Philippine reflexes. At least one verb has been reconstructed in infix form in PAN: (B1) k-ar-uSkuS 'to scrape' (see section III for a Chinese cognate, also r-infix).

FORMOSA: PAIWAN /ar/ 'distributed action':

Ferrell (1982) /ar2/ 'do indiscriminately, on all sides':

Egli (1990:68) /ar/ 'Infix des Bald-dort-bald-da Tuns; Infix des stürmischen Tuns':

k/ar/a-kim 'to search everywhere' (kim 'search'; Egli lists k-ar-im 'look everywhere')

dj/ar/a-djulāt 'move feet constantly' (djulat 'a pace')

dj/ar/a-djiqes 'splash water on all sides' (djiqes 'splash')

t/ar/a-tekēl 'drink everything in sight' (tekēl 'to drink')

b/ar/usbus 'to chatter on and on' (busbus 'very light drizzling rain')

dj/ar/akadjak 'stomp one's feet' (djakadjak 'leg with which one kicks')

Ferrell (1982) /ar1/ ('having sound or quality of; involving use of')

b/ar/engbeng 'make buzzing or humming sound' (engbeng onom.)

k/ar/uḍkuḍ 'to have sound of hoeing' (kuḍkuḍ 'hoe')

k/ar/uskus 'have sound of scraping' (kuskus 'scraper')

s/ar/eLseL 'make buzzing sound' (eLseL onom.)

FORMOSA: AMIS /ar/ distributed action:

Examples from Fey (1986)

p/ar/okpok 'to gallop'

p/ar/ekpek 'to march'

t/ar/iktik 'to walk lightly'

h/ar/eghey 'sound of waves or approaching vehicles or planes'

g/ar/awgaw 'noise of many voices'

PHILIPPINES: TAGALOG /al/ 'distributed action' (data from Panganiban 1973):

d/al/akdak 'sowing of rice seeds or seedlings for transplanting' (dakdak 'driving in of sharp end of stakes into soil')

g/al/ukgok 1. 'trembling from anger or cold' 2. 'noise of stomach movement when beginning to be hungry'

g/al/uygoy 1. 'trembling due to coldness' 2. 'vibration of fat flesh when an obese person moves; bobbing up and down of female breasts when a woman runs or jumps'.

b/al/usbos 'noise of falling water' (busbos 'pierced, perforated') d/al/ugdog 'beating of drums, sound of drum beats'

OLD CHINESE: -r- 'distributed action'

Distributed actions designated by more than two r-infixed synonyms are: scratching/scraping (4), sewing (3), accumulating/stocking up (5), hesitating/walking hesitatingly (5).

Some minimal pairs are:

講 gak (gak) 'speak frankly and brusquely' : 謬 grak (g-r-ak) 'insist; dispute, contest'
洗 sidx (sij?) 'wash' : 洒 sridx (s-r-ij?) 'sprinkle, cleanse'
合 gap (gop) 'join' (two things together) : 洽 grap (g-r-op) 'unite' (more than 2 things)
嗒 gap (gap) 'talkative' : 聒 grap (g-r-ap) 'sound of many people talking at the same time'
磬 khgh (khegts) 'musical stones' : 鐃 khriq (kh-r-eg) 'sound of bells etc.'

Morphology had died out in the standard language by the time of Middle Chinese (ca. 500 AD), but r-infixation survived in varieties of northern Mandarin spoken in Shaanxi, Shaanxi and Inner Mongolia:

MODERN CENTRAL NORTHERN MANDARIN, INNER MONGOLIA (Ih Ju League, Li 1991): -l- (regular reflex of OC -r-) 'distributed action':

pai3 'to shake' : pə? lai3 'to oscillate, swing'

pan4 'to mix' : pə? lan4 'to stir'

pan1 'to run' : pə? lan1 'to run on all sides'

kup3 'to roll' : kuə? lup3 'to do somersaults'

xua4 'to draw, paint' : xuə? la4 'to scribble'

NOTE: in the notation [ə?] the glottal stop is not phonetic, but serves to indicate the shortness of the preceding neutral vowel.

2. PAN -ar- 'distributed object' : OC -r- 'distributed object'

Distributed objects form sets or series. Names of distributed objects are count nouns.

Typical nouns of distributed objects are: names of animal or plant species (collectives); names of paired or multiple bodyparts; names of paired objects (tongs, pincers, chopsticks, shoes, etc.) or objects having a repetitive pattern, for instance toothed objects (rakes, combs, harrows and the like). The function 'distributed object' of /ar/ existed in PAN, as shown by the convergence of Paiwanic and Philippine languages:

FORMOSA: PAIWANIC /ar/ 'distributed object'

Paiwan (Ferrell 1982):

d/ar/epdep 'medium-small yellow glass beads'

q/ar/ungqung 'Forrestia chinensis'

t/ar/ivtiv 'Beilschmiedia erythrophloia; Cinnamomum randaiense'

tj/ar/ungul 'small lobster or shrimp'

v/ar/aLa-valan 'Dace (fish sp.); Leuciscus macropus'

Amis (Fey 1986):

t/ar/odo' 'fingers, toes'

k/ar/ot 'harrow'

k/ar/iwkiway 'hermit crab'

c/ar/ofacof 'fog, light cloud'

Pazeh (Ferrell 1968):

b/ar/uzak 'pig; wild pig' (cf. AtaSq bziok, AtCi bazuk, Sed bOyak 'wild pig').

PHILIPPINES: TAGALOG /al/ 'distributed object'

b/al/ugbog 'dorsal vertebrae'

k/al/iskis 'scales of fish'

b/al/atbat 'horizontal bamboos along roof frame, across rafters'

k/al/aykay 'rake'

g/al/udgod 'series of tops or summits of mountain ranges'

d/al/andan 'Citrus aurantiaca' (orange tree)

OLD CHINESE -r- 'distributed object'

In OC -r- is rare in names of substances or unique objects (sun, moon, etc.). Its prime function in nouns was to mark distributed objects. Objects designated by more than two r-infixed synonyms are: chopsticks (3), street/alley (3), young shoot (3), pig (4).

A statistic (to appear) bearing on OC names of human and animal bodyparts in Karlgren's 'Grammata Serica Recensa', where the overall ratio of names of plural bodyparts (eyes, teeth, fingers, horns, etc.) to names of unique bodyparts (head, mouth, heart, tail, etc.) is roughly 2 : 1, gave the following result: out of 36 bodypart names carrying -r-, 32 were names of plural bodyparts, a statistically very significant indication of the correlation between -r- and names of distributed objects.

Some minimal pairs illustrating the role of -r- in forming distributed nouns may be cited:

齊 dzid 'equal, uniform' : 儕 dz-r-id 'class, group of equals'

合 gəp 'join, unite' : 禘 g-r-əp 'name of a sacrifice to ancestors taken collectively'

眼 gənɣ 'to protuberate' : 眼 g-r-ənɣ 'the eyeballs'

As with distributed verbs, r-infixation in nouns survived in some peripheral Mandarin dialects, and scattered examples are also found in southern dialects:

MODERN CENTRAL NORTHERN MANDARIN: INNER MONGOLIA (Ih Ju League, Li 1991): -l- (regular reflex of OC -r-) 'distributed object':

təu1 'hood, hanging pocket' → tə? lau1 'heavy clusters of fruits hanging from tree branches'

thuŋ2 'buttocks' → thuə? luŋ2 'buttocks'

tʂhiag3 (earlier kh-) 'lungs' : khə? laŋ3 'lungs'

thau4 'set, series' → thə? lau4 'tirade, long string of words'

xəŋ4 'alley, small street' → xə? laŋ4 'idem'

khu3 'nave of a wheel' → khuə? lu3 'wheel(s) of a car'

CANTONESE: -l- 'distributed object' (fossilized)

k-l-ək tei

armpit-under

armpit (OC 腋 [766d] kak 'armpit')

k-l-ɔk theu

angle-NOUW SUFFIX

'corner, angle', cf. Cant. 角 [1225ac] kək 'horn, corner'

B. PAN -in- 'patient nominalizing infix' : OC -j- 'non-agent nominalizing infix'

PAN -in- 'patient nominalizing infix': (Starosta, Pawley and Reid 1982) : "the N affected by V-ing":

hau 'to fine' : h-n-au 'person who is fined' (Atayal);

karat 'to write' : k-in-arat 'book, paper' (Saisiyat);

pa-orip 'cause to be alive' : p-in-a-orip 'domestic animals' (Amis);

batak 'to stretch, pull' : b-in-atak 'sugar candy stick' (Tagalog);

tapa 'smoking; to dry' : t-i-apa 'smoked food' (Maranao);

faisen 'to ask' : i f-in-aisen 'the question' (Chamorro).

Note the reduction of the infix to -i- in Maranao.

OC -j- (= yod) 'non-agent nominalizing infix'

割 kat (kat) 'to cut, castrate' : 羯 kjat (k-j-at) 'gelded sheep'

脫 duat (lot) 'to peel off' : 蛻 ruat (l-j-ot) 'insect exuviae'

銘 mig (meg) 'inscribe, engrave, inscription' : 名 m-j-ig (m-j-eg) 'written word, name'

投 dug (do) 'to throw' : 投 d-j-ug (djo) 'k.o. lance or spear'.

斷 duanx, tuanx (don?, ton?) 'to cut, cut off' : 脔 djuanx, tjuanx (d-j-on?, t-j-on?) 'cut meat, slice'

駢 bien (bin) 'to associate, join together' : 嬖 bjien (b-j-in) 'mate, wife' / 皇 guag (wag) 'august' :

王 gwjag (w-j-ag) 'king' / 句 kug (ko) 'curved' : 痾 kjug (k-j-o) 'crooked spine'

稜 læg (rɨg) 'angular, lozenge-shaped' : 菱 ljæg (r-j-ig) 'water chestnut (Trapa bicornis)'

丹 tan (tan) 'red' : 旗 tjan (t-j-an) 'banner of red silk'

峨 gar (ɣaj) 'solemn' : 儀 gjar (ɣ-j-aj) 'decorum, proper demeanor, ceremony, rule'.

旁 baŋ (baŋ) 'side' : 房 bjag (b-j-ag) 'side-room'

屏 biŋ (beŋ) 'screen' : 屏 bjig (b-j-eg) '[screened place:] latrine'

火 hmarx (hmij?) 'fire' : 焚 hmjærx (hm-j-ij?) 'land cleared by fire'

C. PROTO-AUSTRONESIAN ma- 'stative/intransitive verb prefix'

OLD CHINESE M- 'stative/intransitive verb prefix'

AUSTRONESIAN

Atayal (Egerod 1980):

m-: 'prefix forming an active indicative of intransitive verbs'
 m-buloq 'to be blind'
 m-busuk 'to be drunk'
 m-takui 'to fall'
 m-ʔuyay 'to be hungry'
 m-guju 'to be afraid'
 m-sayux 'to be ashamed'

Paiwan (Ferrell 1982):

ma- ('be affected by; be in condition of (without own volition)'
 ma-kan 'to be(come) eaten' (kan 'eat')
 ma-tjani 'to fall down'
 ma-sadau 'to receive a gash wound'
 ma-quɟal 'to be rained upon (quɟal 'rain')
 ma-qadau 'to be bothered by the heat of the sun (qadau 'sun')
 ma-lum 'to be(come) ripened'

ma- [non intentional] 'be object/recipient of action (usually involves an outside agent); be in a state of':

keLang 'understand, know about' : ma-keLang 'be(come) known about'
 langeda 'hear' : ma-langeda 'be(come) heard; be audible'
 adjuq 'leave behind' : ma-adjuq 'be(come) left behind'
 gutsguts 'scratch' : ma-gutsguts 'be(come) scratched'

Tagalog: ma- 'stative prefix'

ma-kita 'be seen' ma-init 'become hot'
 ma-tapus 'be finished' ma-lutoq 'get cooked'

OLD CHINESE ACTIVE/TRANSITIVE : INTRANSITIVE/STATIVE ALTERNATIONS

TRANSITIVE : INTRANSITIVE

見 kianh (kens) 'to see' : 現 gianh (N-kens) 'to appear'
 覆 phjəkw (phjuk) 'turn over' : 復 bjəkw (N-phjuk) 'to return'
 別 pjat (pjat) 'to separate' : 別 bjat (N-pjat) 'to take leave'
 會 kuadh (kots) 'to assemble (trans.)' : 會 guadh (N-kots) 'to assemble' (intr.)
 合 kap (kop) 'to put together' : 合 gəp (N-kop) 'to agree'
 折 tjat (tjat) 'to break, to bend (trans.)' : 折 djat (N-tjat) 'to bend' (intr.)

ACTIVE : STATIVE

敗 pradh (prats) 'to defeat' : 敗 bradh (N-prats) 'to be defeated'
 壞 kuradh (krujs) 'to destroy, ruin' : 壞 guradh (N-krujs) 'to be ruined'
 中 trjəwuh (trjujs) 'hit the centre' : 仲 drjəwuh (N-trjujs) 'to be in the middle; middle brother, 2nd of 3'
 斷 tuanh (tons) 'to cut, sever' : 斷 duanh (N-tons) 'to be broken off, cut off from; to

cease'

- 屬 tjuk (tjok) 'to assemble' : 屬 djuk (N-tjok) 'to be connected with'
著 trjak (trjak) 'to place, put something in a certain place' : 著 drjak (N-trjak) 'to occupy a fixed position'
解 krigx (kre?) 'separate, untie' : 解 grigx (N-kre?) 'loosened'
汜 phjamh (phjla,olms) 'to overflow, inundate' : 汜 bjamh (N-phjla,olms) 'to be thrown about, float about'
降 krəŋwh (krugs) 'to go down, send down' : 降 grəŋwh (N-krugs) 'to lie down, submit to'
繫 kigh (keks) 'to attach' : 繫 gigh (N-keks) 'be attached to'
檢 kjiamx (krjam?) 'to restrain' : 儉 gjiamx (N-krjam?) 'poor, in dire straits'
夾 kriap (krep) 'to press between' : 狹 griap (N-krep) 'narrow, pressed on both sides'
偌 krid (krij) 'to accompany' : 諧 grid (N-krij) 'to harmonize with'

Li's reconstruction of voiced obstruent initials in the intransitive or stative set is a backward projection of the MC dictionary forms which show these initials. Pulleyblank (1973), followed by Baxter (1992), supposes voiceless initials in OC undergoing secondary voicing in MC under the influence of a prefix h-. There is clear evidence, however, that the prefix responsible for intransitive and stative voicing was a nasal, which I write as N- : in a study of Chinese loanwords in the Mien dialect of Yao, Downer (1973: 14-16) cited the following Mien forms:

tshe? (PY *tshək) 'to pull down, pull apart'
dze? (PY *nts(h)ək) 'to be cracked, as earth'

khɔi1 (PY *khɔiA) 'to open' (transitive: as in 'open the door')
gɔi1 (PY *ŋk(h)ɔiA) 'to open' (intransitive: as in 'the heart opens: be happy' and 'to blossom')

gaat? (PY *ŋkhat) 'thirsty'

As Downer saw, these are Chinese loanwords: cf. 拆 thrak (th-r-ak) 'to split, be rent; fissure', 開 khəd (khij) 'to open', 渴 khat (khat) 'thirsty'. They are regarded by Downer as forming part of the earliest stratum of Chinese loans into Mien. The last item has a White Miao cognate, cited by Downer as nqhe_ 'thirsty'. At least in this case, it would seem that the borrowing was from Chinese into PMY.

Downer supposes a PY prefix (prenasalization) capable of turning active verbs into intransitive ones. He seems unaware that a similar process exists in Chinese, and cannot explain why his prefix only affects Chinese loanwords.

The MC dictionaries do not list voiced-initial readings reflecting N- prefixed allomorphs for the three Chinese words under discussion (though a second reading g-j-at (? N+g-j-at) exists for 渴 in the meaning 'dried up, of a lake'), yet it must be assumed that such forms did exist and were borrowed by PMY, and later PY: there is no other explanation for the fact that the "Mien nasal prefix" does not seem to ever affect native Mien words, only Chinese

loanwords.

III. EVIDENCE FROM THE LEXICON

Evidence is presented in this section that OC and PAM share some lexical items, and that the set of shared words is not of a kind usually taken as indicative of borrowing.

In order to reduce the risk of chance resemblances, the following methodological controls have been applied to the lexical comparisons:

- first, strict requirements have been applied to the circumstances of attestation of the lexical material under comparison, as a partial hedge against late innovations. Specifically, on the Chinese side, only words attested in Han or pre-Han epigraphy, literature, or lexicographical works (in other words, words which were demonstrably part of the language in the period 1400 BC - 200 AD) are considered for comparison. This is a highly restrictive move, which excludes all words attested only in such widely used lexicographical sources as *Guang Ya*, *Guang Yun*, *Yu Pian* and *Ji Yun*, the latter a prominent source in my earlier work (though naturally, all OC reconstructions are based on the MC readings in these works). Care has also been taken to exclude rare words. On the AN side, only reconstructions (words and roots, in the sense of Blust 1988a) assignable to the earliest (PAM) level on the subgrouping assumptions of Blust (see paragraph I.A) are included. This is also a highly restrictive move, one which in effect narrows the list of eligible items to those attested both inside and outside of Formosa, altogether perhaps no more than a few hundred items in currently available lists.
- Second, comparisons involving onomatopes and words of expressive origin have not been considered for comparison. This excludes 62 of the 231 roots in Blust (1988a), plus some words which do not contain any root, for instance (B1) tuRtuR 'dove, coo of the dove', cf. OC 雉 t-j-uj 'kind of dove', Lat. turtur 'dove'.
- Third, close semantic matchings have been required. Where semantic shifts are needed, evidence that the shift occurred in other languages (often Indo-European languages, based on C. D. Buck's 1949 'Dictionary of selected synonyms in the principal Indo-European languages') is presented.

In continuation of my earlier practice, unless otherwise specified, the sound material on both sides is entirely accounted for by the sound correspondences presented in section IV.

As a result of applying these controls, only a small number of comparisons (56 altogether) can be presented here. They relate for the most part to universal semantic notions: bodyparts and bodily functions (1-7), the physical world (8-11), animals and plants (12-17), deictics (18-19), qualities (20-26) and actions (27-52). A few comparisons involving cultural notions of great antiquity (53-56) are also cited.

Note: in the citations below, capitalized vowels are stressed. OC prefixes and suffixes are separated from the word stems by means of '+', and infixes by hyphens. AN reconstructions

are preceded by the indication of their source and level of reconstruction. Abbreviations for sources are: 'B' = Blust (1970, 1980, 1983-84, 1985, 1989); 'Br' = Blust (1988a); 'D' = Dempwolff (1938); 'S' = Sagart, this paper; 'T' = Tsuchida (1976). Roots are signaled by the symbol 'r' following the source abbreviation, e.g. 'Br', 'Sr'. Reconstruction levels are: '1' = PAN; '2' = PMP; '3' = PWMP; '1+' indicates a level 2 or 3 reconstruction raised to level 1 in the present paper on the strength of one or more Formosan cognates. Abbreviations for specific articles and monographs are: 'AE1' = Blust (1980); 'AE2' = Blust (1983-83); 'AE3' = Blust (1985); 'AE4' = Blust (1989); 'ART' = (Blust 1988a); 'JWGL' = Chou Fa-Kao (1974); 'LWFL' = Tsuchida (ms.); 'RPTP' = Tsuchida (1976). Chinese characters are followed by *Grammata Serica Recensa* number (GSR, Karlgren 1964) between square brackets.

1. HEAD

(D1+) quluH1, consisting of (D2) qulu 'head', Dahl (PAN = level 2) quluH1 'head', plus Puyuma qhuru-yu 'head' (Wolff, p.c.); evidence for final -H1 is provided by final -h in Aklanon and Cebuano reflexes (Dahl 1981:46-7).

首 [1102a] hrjəgwɔx (?) (hl-j-u?) 'head' (Shi Jing, Shu Jing, Zhou bronze, cf. JWGL 1194). Note: 首 is phonetic in GSR 1048: e.g. 1048a: 道 dəgwɔx (lu?) 'way, road', though this is not recognized by Karlgren. OC 首 was first compared to Dempwolff's *qulu by Peyros and Starostin (1984).

2. BRAIN

(T1, RPTP:172) punuq 'brain, head; forehead': Sar punuʔu, Tso pnuu 'brain', RukBd punu, RukMg t-punu-a, RukTn punu', RukMn ta-punu-a 'forehead', Pai punuq, Ami punuq, Bunun punuh 'brain', Paz punuʔ 'head', Tag pu:noʔ 'leader'. Additional evidence for this word in Dyen (1992:188) who cited more Philippine cognates in the meanings 'chief, head of a group', and 'beginning, source, origin'. Blust (1992:15) criticized this reconstruction, arguing that "none of the Philippine forms cited here can be said to be associated with body parts" and claiming that "the whole extra-Formosan set almost certainly reflects *puqun 'beginning, source, origin; base of a tree'". Blust's principal reason for supposing that such forms as Tag., Hlg pu:noʔ 'leader, head of a group' are unrelated to the Formosan set but are a metathesized form of (D2) *puqun 'beginning, source, origin; base of a tree' is that they are also used in the meaning 'trunk of a tree'. Blust's suggestion of a metathesis thus has some plausibility. However, two Philippine words reflecting *puqun definitely show semantic association with 'head' as a bodypart: Ceb pu:noʔ, Bkl po:noʔ 'sores on the head'. The connection between the Formosan and extra-Formosan sets is thus certain. If Blust were right about these forms reflecting *puqun 'trunk of a tree, etc', it is quite unclear how their semantics could have evolved. In contrast, the semantic change from 'head' to 'leader' is so common (It. capo, Fr. chef, both 'chief, leader', from Lat. caput 'head'; Mand. tou2tou2 'leader', from tou2 'head'; Eng. 'head of state', 'head official', etc.), and the semantic extensions of 'head' as 'beginning, source, origin' so widespread (Mand. tou2 'head', also 'beginning, first', etc.; cf. also Eng. 'head of a river') that assuming a semantic development from PAN punuq 'brain, top of the head' poses no particular difficulty. I will assume that PAN punuq 'brain' underwent semantic shifts to 'head' and 'beginning, source, origin', and ultimately to 'base of a tree'. (D2) puqun 'base of a tree; beginning, source, origin' is best treated as the metathesized form of (T1) punuq.

腦 [1244f] nagux (nu?) 'brain' (Zuo Zhuan: Xi 28; Zhan Guo Ce; Li Ji; Shuo Wen, written with key 21 匕 instead of 130 肉); 'head, skull' (Huai Nan Zi).

3. NECK/GULLET

(D1+) 𤑔 [ilqeR 'neck': Fav ri (Happart 1650 [1896]; Marsh 1980 [1977]), Kan ə?əre, RukBd la:?, RukMg əra:, RukTn ə?ə, RukMn la?ə (data from Tsuchida's LWFL), Amis li?el (Fey 1986), Tg li?ig, Ml leher, all 'neck'.

咽 [370h] ?in (?in) 'gullet' (Zhan Guo Ce).

4. BREASTS/MILK

(T1, RPTP:229) nunuH1 'breasts, milk': Sar nuunu?-a, Tso nun?u 'breasts of a woman; to milk, milk'; Paz mu-nunuh 'to suck milk', Mlg nUnu 'breasts of woman'.

乳 [135a] njugx (n-j-o?) 'milk, suckle' (Zuo Zhuan); 'nipple' (Zhou Li); the Zuo Zhuan (Xuan:4) cites a non-infixed equivalent in the Chu dialect, read as though from OC nugx (no?) (Jing Dian Shi Wen 15).

5. TO VOMIT/SPIT OUT

(D1+) u(n)taq 'to vomit', consisting of the PMP forms under (D2) [lu(n)taq 'to vomit', plus the Formosan set assembled by Tsuchida (RPTP:169), including: Atayal m-utaq, Tsou tra-vto, Rukai ut-taa, Paiwan m-utaq, Puyuma utaqh+an 'vomit' and other forms.

吐 [62d] thagx (tha?) 'to eject from the mouth' (Shi Jing); thagh (tha?+s) vomit' (Guang Yun). On the reconstruction of the initial of this word as OC hl- by Baxter (1992), see my comments under 8. EARTH below.

6. TO SLEEP

(D1+) tuDuR 'to sleep', consisting of Bunun (Nihira 1983) tudul 'sleep', ma-tudul 'sleepy, drowsy'; Favorlang atorro (Happart 1650 [1896]; Marsh 1980 [1977]); plus (D2) tuDuR 'to sleep' (Tg. tulog, Jav. turu, Ho. turi, Sa'a me-uru 'to sleep'); and PPh (Zorc) tu:DuR 'sleep'.

睡 [31d] djuarh (d-j-oj+s) 'to sleep' (Guo Ce); 'to sleep while sitting' (Shuo Wen).

7. TO THINK

(B1, AE2) nemnem 'think': PAI ki-nemnem 'think, cogitate, pay close attention to'; KAN nemnEm 'mental faculties'; ILK pa-nemnem-en 'consider, fix the mind on, think on', ILONGOT ninnin, MOTA nonom 'to think'.

念 [670a] niamh (ni+ts) 'think of' (Shi Jing, Shu Jing, Zhou bronze, cf. JWGL 1378). With unexplained tone correspondence: 恇 [667q] njəmx (n-j-iə?) 'think' (late Zhou inscriptions; Hou Han Shu).

8. EARTH

(Br1) -taq 'mud; earth, ground', incl. PUY litaH 'dirt', AMIS pota? 'soil, soft mushy dirt', sota? 'earth, dirt; mud; land'; also (B2, AE4) buRtaq 'mud, earth'. Add Favorlang ta 'territory, land or dwelling-place, the common road or street, the ground'.

土 [62a] thagx (tha?) 'soil, earth, land' (Shi Jing). Used in graphs as general signfic for 'earth'. Based on an isolated TB comparandum, Bodman (1980:102) reconstructed this word

as OC hla: . Baxter (1992:793) followed, reconstructing hla?, and extending the same treatment to another word in the same series: 吐 [62d] thagx,h (tha?(+s)) 'eject from the mouth; vomit', Baxter hla?(s). The treatment of the initial of these two words as OC hl- by Bodman and Baxter is absolutely contraindicated by the internal Chinese evidence: other words in the same phonetic series have MC initials which can only reflect OC dental stops, showing the series cannot be a lateral one: 肚 [M29270] 'belly', MC tuoB, unambiguously reflecting OC t-; and 社 [62j] MC zjaB, unambiguously reflecting OC d-.

9. TO FLOW/WATER(S)

(Br1+) -luR 'to flow', consisting of (Br3) -luR, 'to flow' (incl. (D3) aluR 'pond, stream', BKL sulog 'flow of water'), plus AMIS ?alol 'to be carried away by water; to drift, of a boat; to move from one place to another, usually with reference to the water' (Fey 1986: part of the gloss is found under the entry for "sonol"). Blust (1988b:52) observed that no word for 'river' is reconstructible at the PAN level. He suggested that a word for 'water', or 'flowing water' may have been used instead.

水 [576al hurjidx (hl-j-uj?) 'river' (Shi Jing, Shu Jing); 'body of water' (Shu Jing); 'water' (as substance) (Yi Jing: 地上有水). The graph is a drawing of running water. This word is placed in the OC Zhi 脂 category by Tung (1967) and Chou (1970) with ambiguous support from Ode 92; based on Odes 104 and 183, Tang (ms) and Baxter (1992) place it in the Wei 微 category. I follow Tang and Baxter. Both meanings of 'water' (as substance) and 'river'; 'body of water' are derivable from an original meaning 'to flow': cf. IE wedor 'water', from a root -wed 'wet; to flow'; most of the words for 'river' in IE langs. are derived from roots meaning 'flow, run' (Buck 1949:34,41).

10. SALT

(D1+) siraQ1 'salt', consisting of (D3) sira[q] 'condiment, salt' (Tg. sila? 'culinary preparation', TB, Ho sira 'salt'); plus Amis tsirah (Ogawa and Asai 1935). Note also Bunun ?aslla? (Ferrell 1969:98), Thao qtl:La? (Li Fang-kuei, as cited by Ferrell 1969:98). TsoDt sIru 'salt' (LWFL) may also belong here, but if so final -u is unexplained (-o expected). In contrast to PAN qasiN, siraQ1 seems to refer to salt as substance rather than taste.

鹵 [71al lagx (ra?): 'salt' (Zhou bronze, cf. JWGL 1488); 'salty' (Yi Jing); 'saltlands' (Zuo Zhuan, Rang 25); 'salt' (natural, as opposed to processed: 『鹽, 天生曰鹵, 人生曰鹽』: Shuo Wen). Used in graphs as general signfic for 'salt'.

11. CLOUD/CLOUDY

(Sr1) -qem 'cloud, cloudy': Amis to?əm 'clouds' (Fey 1986); RukBd əmE:mə 'cloud' (LWFL); Bunun lohəm (Ogawa and Asai 1935), luhəm (Nihira 1983) 'cloud'; this root also includes a PWMP word, reconstructed here on the basis of a cognate set first assembled by K. Wulff (1942, No 53): (3) q-al-emqem 'overcast' (Sundanese haləmhem 'overcast', Ceb. ?alum?um 'for the sky to be overcast'). Note: Saisiyat imim (Ogawa and Asai 1935), əməm (Li 1978) 'cloud' is possibly related but absence of initial ?- is irregular. Note also PPh DaGqum 'cloudy', (Zorc and Charles 1971) with discrepant vowel.

陰 [651y] ?jəm (?-(r)j-um) 'cloudy; dark, shade; northern slope of a height (not receiving sunlight)' (Shi Jing); 'the Yin or dark cosmogonic principle' (Yi Jing); [651a] id. 'cloudy' (Da Dai Li Ji); [651x] id. 'cloudy' (Shuo Wen: the signfic is 云 'cloud').

12. SNAKE

(D1+) SulaR 'snake', consisting of (D3) ulaR 'snake' (Ja ula, Ml ular, Ho ulatra 'snake'); plus RukBd solAo, RukMg surA:, RukTn soA?a, RukMn ?ola?a 'snake' (forms from Tsuchida's LWFL). Note: Tsuchida (RPTP:143) reconstructed PHM bul[æ]y 'snake' on the basis of Kan vunAy, Sar vuli?i, Tso fkoi and Itb volay 'snake'. Since outside of Formosa this etymon is only represented in Itbayaten, a northern Philippine language, it cannot be assigned to level 1 because of the risk of borrowing from or into Formosan.

蛇 [41] djar (l-j-aj) 'snake' (Shi Jing); 虺 [4a] thar (hlaj) 'snake' (Shuo Wen).

13. WORM

(D1+) [lulej 'worm', consisting of (D2) ulej 'worm, larva' and Puyuma qhured 'worm, insect' (Wolff, p.c. 1992).

蟪 [458j] rinx (l-j-in?) 'earth-worm' (Xun Zi); db1. 蚓 [371c] id., same meaning (Meng Zi).

14. EGG

(T1, RPTP:172) qiCeluR 'egg': Kan ?icuuru, Pai qacilu, Tag ?itlog, Sa sa-olu 'egg'; see also (D2) teluR 'egg'.

卵 [179a] luanx (l+lon?) 'egg' (Zuo Zhuan); 橢 [11c] thuarx (hloj?) 'oval' (Chu Ci).

15. BEAK/PECK

(B1, AE1) tuktuk 'beak of a bird; to peck': PAI tjuktjuk 'beak (of bird)', tj-m-uktjuk 'to peck; to strike (as a snake)'; KAP 'bill, beak (of a bird)'; WBM tuktuk 'beak of a bird'; (...); to peck', etc. Note also Proto-Atayalic g-um-atuk 'to peck' (Li 1981), and (D3) pa(n)tuk 'to peck' (Ja patuk 'beak').

啄 [1218b] tuk (tok) or truk (t-r-ok) 'to peck up' (Shi Jing); 嚼 [1224n] truk (t-r-ok) 'to peck' (Guo Ce); trjugh (t-rj-ok+s) 'beak' (only Han time text example).

16. EDIBLE FERN

(B1, AE1) lukuC 'parasitic plant sp.': Puy rukuT 'Asplenium nidus', Pai lukuts 'parasitic long-leaved plant with edible fruit — grows on trees: Asplenium nidus', CHM lu-luhot 'a vine: Maytenus thompsonii', MARUM ge-lukut, BATU BELAH lukot 'moss', IBAN lukut 'moss, lichen, some algae', MAL lukut 'long moss, in contrast to mere lichen' (in Brunei, Sarawak: 'large fern, Platycerium spp., found on trees'), SND lukut 'duckweed, moss', BM lukut 'parasite or creeper that grows on durian trees, etc., and which is used medicinally', UMA luku? 'epiphyte', LI'O luku 'tree with edible fruit'. Add Amis lokot 'edible fern growing on trees' (Fey 1986).

蕨 [381d] kujat (k-j-ot) 'an edible fern, Pteridium aquilinum' (Shi Jing).

17. HUSK, CHAFF

Dyen (PHM ? = level 1, 1992:189) epa ('husk, chaff'?): Atayal (Sq) pa-qi?, Pangasinan qepa, Toba Batak hopa-h-opa 'rice husk', Buli (Halmahera) afa 'chaff'. Also (D3) qempa 'husks, chaff' (consisting of Ml empa, NgD hampa, both 'husks, chaff').

藎 [182*, M47739] phjag (ph-j-a) 'husks, chaff of wheat' (Shuo Wen).

18. THIS

(T1, RPTP:135) *iniH2 'this', incl. Sar ka-ni?i 'this', na-ani 'here', Tso eni 'this', RukBd ku-ini, RukMg k-inii, RukTn ak-ini 'that (near hearer)', Puy ini?, Ami ini, Sai h-ini, Paz ini?, Ata qa-ni, Sed ?ini? 'this', Ulawā ini 'that'. For more extra-Formosan cognates see (D2) ini 'this (demonstrative pronoun)'.

爾 [359a] njidx (n-j-ij?) deictic adjective 'this' (Shi Jing: Da Ya, Sang Rou); deictic pronoun 'this' (Gong Yang Zhuan); 'thus' (Meng Zi); 'you' (Zuo Zhuan, Lun Yu, Meng Zi). On these and other references, see Yang Shuda's Ci Quan. Note also Cantonese ni (high level tone) 'this'.

19. OPPOSITE/THAT (FAR)

(T1, RPTP:160) S6i(m)pa[rR] 'opposite shore': Kan m-u-a-siparə, Sar m-u-u-siparə 'walk across the river', i-isiparə 'the other side of the river'; add Bunun sipal 'opposite' (Mihira 1983); plus the items under (D3) i(m)pa[r] 'other side' and Zorc and Charles (PPH, 1971) iMpa(rR).

彼 [25g] pjarx (p-j-aj?) 'far demonstrative' (as opp. to 此, close demonstrative): 'he, him, that, these, they, etc.' (Shi Jing, Yi Jing).

20. HIGH

(Br1) -kaw 'high, tall', including (B2, AE1) laŋkaw 'high, tall', raised to level 1 in ART (p.105), on the strength of unspecified Formosan material.

高 [1129a] kagw (kaw) 'high' (Shi Jing); 僑 [1138i] k-j-agw (k-j-aw) 'high, lift the head' (Zhuang Zi).

21. BROAD

(Sr1) -bag 'broad', consisting of candidate root -bag 'broad' (ART:73), based on 3 WMP items (probably including (D3) bagbag 'broad, wide' and (B3) lak(e)bag 'broad, wide'), supplemented with Atayal (Egerod) labag 'broad', Sediq llabag 'wide' (Li 1981).

旁 [740f'] bag (bag) 'broad, wide' (Shu Jing); 'wide' (Xun Zi); 'extensive, vast' (Shuo Wen); also 'side' (Shu Jing, Zuo Zhuan), and in derivatives: 房 b-j-ag 'side-room', 傍 bag+s 'at side of'. The primary meaning is 'broad, wide', with semantic extension to 'side' through 'width', as in Lat. *latus* 'side', orig. 'spread out, wide', and English 'side', from OE *sid* 'wide, broad, long'.

22. SALTY

(T1, RPTP:128) q2asiN 'salt': Kan ma-?aini, Sar m-?aiLi 'salt' Tso m-aski, Puy hasil 'salty', Mlg h-asina 'salt, savour', To m-ahi 'sour, astringent'. Other MP items under (D2) asin 'salt', in particular Fi m-asi-m-asi-a, Fu, Sm masi 'fermented breadfruit'. The semantics in Tso, Puy, Mlg, Fi, Fu, Sm point to the taste, rather than to the substance, of salt.

辛 s jin (s-j-in) 'hot-tasting, pungent, bitter' (Shi Jing, Shu Jing); describes the taste of onions, garlic, etc. Semantic parallels: Lithuanian *suurus* 'salt', from an IE word meaning 'sour' (Buck 1949:15.38); Atayal tqmtux, qpmtux 'bitter, salty' (Egerod).

23. OVERCAST

(Br1) -Dem 'dark, overcast': incl. (B1, AE3) quDem 'dim, obscure, dark' (PAI quzem 'dim, dull-colored', PUY Huz-Huzem 'black', BAL urem 'unclear, misty, half-dark, twilight'); as well as other PMP and PWMP reconstructions and isolated words. To these add the cognate set assembled by Tsuchida (T1, RPTP:228) under DZəmDZəm 'dark', incl. Sar səəsəm-a 'dark', Tso cməcmə 'rainclouds', RukTn mə-D-a-məDEmə, Bun mə-dumdum 'dark', Tonga lo:lo: 'absolutely dark'. Or (Br1) -lem 'dark', incl. (B1, ART) qiS(e)lem 'dark color', (D3) kelem 'dark', raised to level 1 in ART on the strength of unspecified Form. evidence., and various other reconstructions and isolated words, notably from Paiwan and Puyuma.

灑 [656*, M42452] drjəm ([d,l]-rj-[i,u]m) 'long spell of overcast weather' (Shuo Wen); 曇 [M14172] dəm ([d,l][i,u]m) 'clouds covering the sky' (Shuo Wen Xin Fu; Yu Pian).

24. BENT I

(Br1) -kukl 'bent, crooked'. Includes Amis ʔotkok 'to huddle, as when cold'; (B2, AE3) dugkuk 'hunched over' and numerous other PWMP and PMP reconstructions as well as isolated words.

曲 [1213a] khjuk (kh-(r)j-ok) 'bend, bent' (Shi Jing, Shu Jing), 'crooked, unjust' (Zuo Zhuan); with 'distributed object' -r-: 角 [1225a] k-r-uk (k-r-ok) 'horn, the horns' (Shi Jing); with voiced initial (stative M- prefix): 屈 [1214a] g-j-uk (M+k-(r)j-ok) 'compressed, bent, curved, of the body' (Shi Jing); 'curl, twist, of hair' (Shi Jing).

25. BENT II

(Br1) -kul 'curl, bend', incl. Amis fitokol 'dried and bent over, though still alive (trees)', (B2, AE2) sigkul 'bent, of the arm', (D3) Degkul 'curved bodypart, bend a limb', plus various WMP reconstructions and isolated words. Add Favorlang (Happart 1650 [1896]) bokkol 'deformity of going bent', Pazeh (Ferrell 1968) ti-bukul 'crooked, bent'.

卷 [226a] gwjian (M+k-rj-on?) 'bend, curved' (Shi Jing); '(rolled hand:) fist' (Li Ji; also written as 拳 [226g] id.); 捲 [226d] kuhjian (kh-rj-on) 'crooked wood' (Guo Ce); 'bowl made of curving wood' (Meng Zi); 鬢 [226e] khujian (kh-rj-on) or gwjian (M+kh-rj-on) 'curly hair of foreigners' (Guang Yun); [226*, M22259] gwjian (M+k-rj-on) 'sickness of crooked hands' (Guang Yun).

26. BENT III

(Br1) -kut 'hunched, bent over': Amis 'okot 'huddle, as when cold'; (B2, AE2) sikut 'bent, of the arm'; etc.

屈 [496k] khujət (kh-j-ut) 'bend' (Zuo); 'subdue' (Shi).

27. TO ASCEND

(T1, RPTP:155) a(n)Dak 'ascend': Sar mə-asakə, Tso ocʔo 'to ascend the hill', Kyp alak 'to raise the price', Mlg andra 'look upwards', Fi z-andra 'rise (of the sun or moon)'.

陟 [916ac] trjək (t-rj-ik) 'ascend' (Shi Jing, Shu Jing, Zhou bronze, cf. JWGL 1816). With irregular voicing correspondence.

28. TO STRIKE

(B1, AE4) basbas 'hit, strike', incl. AMIS facfac 'to hit, strike, as a child or a water

buffalo', and BKL bashas 'to whip (beat, hit)'.
伐 [387a] bjat (b-j-at) 'to strike; hew, cut; attack' (Shi Jing).

29. TO PULL OUT, UPROOT

(Br1) -buC 'weed, pluck, pull out', including three PAM items: (D2) butbut 'pull out', (D2) Rabut 'pull out', both raised to PAM level and listed as *buCbuC 'pluck, pull out', *Ra(m)buC 'to weed, pull out' by Blust (1988a:86-7) on the strength of unspecified Formosan evidence; and (B1) u(m)buC 'pluck, pull out'. The supporting evidence for that word is also unspecified. The root includes several PMP and PWMP reconstructions as well as isolated items.

拔 [276h] briadh (b-r-ot+s, cf. Baxter 1992:396) 'pull up, uproot' (Shi Jing, Shu Jing, Yi Jing); badh (bot+s) 'thinned out (as a forest, some trees having been pulled up)' (Shi Jing).

30. TO GRASP/TO CATCH

(Br1) -gem 'grasp in the fist': incl. (B2, AE1) gem 'hold in the fist', and (D2) gemgem 'make a fist', raised to level 1 in ART on unspecified evidence. Add Rukai wagam'gem 'to squeeze' (Ferrell 1969).

擒 [651n] gjəm (g-j-[i,u]m) 'catch' (Guo Yu); same word, written as [652*, M12183] id. 'catch in the hand' (Yi Qie Jing Yin Yi); 禽 [651m] id. 'catch, capture; bird' (Zhou bronze, cf. JMGL 1838; Zuo Zhuan); 'animal' (Yi Jing). Semantic development from 'catch' to 'game' to 'animal, bird' cf. Wang Li (1982:685).

31. TO ENCLOSE/TO CONTAIN

(Br1+) -kem 'enclose, cover, grasp', consisting of (Br3) -kem 'enclose, cover; grasp' (add (D3) kemkem 'hold closed, contain'), supplemented with Pazeh mu-hakem 'carry in arms' (Ferrell 1968).

衾 [651h] khjəm (kh-(r)j-i+m) 'coverlet' (Shi Jing); 衾 [615a] kəm 'cover' (Mu Tian Zi Zhuan); 龕 [M48839] khəm 'to contain' (揚 Yang and 越 Yue dialects acc. to Fangyan); 'niche for the effigy of a god' (only Tang time text example); also, with unexplained voicing correspondence: 函 [643a] gəm 'to envelop, contain' (Shi Jing); 'cuirass' (Meng Zi); 含 [651l'] gəm (g+i+m) 'to contain' (Shu Jing, Yi Jing); 衿 [651h] gjəmh (g-(r)j-i+m+s) 'single shroud' (Yi Li).

32. TO EMBRACE, GRASP, CLASP

(Br1) -kep 'seize, grasp, embrace', incl. (B1) dekep 'embrace' (the supporting material for this word has not been published), (D3) kepkep 'grasp, hold', FIJ moko 'embrace, clasp in the arms', etc.

挟 [638l] giap (gep, with unexplained initial voicing) and tsiap (tsep, probably fr. s+kep) 'grasp, hold' (Shi Jing); 'clasp under arm' (Meng Zi); read as tsiap (tsep, from s+kep) 'encompass, embrace' (Shi Jing); 夾 [638a] kriap (k-r-ep) 'be on both sides of' (Shi Jing); 'press between' (Zuo Zhuan).

33. TO ENCIRCLE, WRAP AROUND

(Br1) -kes 'encircle, wrap firmly around', incl. (B3, AA) bakes 'belt', restated as (B1,

ART) ba(R)kes 'k. o. belt' on unspecified evidence; and other PMP and PWMP reconstructions and isolated items, such as KAN balikes 'k. o. woven belt'; HLG waykus 'belt, sash, waist-band', etc.

繫 [279d] kiat (ket), giat (? N-ket) 'to wrap a cord around an object and measure it' (Chu Ci); 縶 [279k] khat (khet), giat (? N-khet) 'girdle' (Zhuang Zi).

34. TO SCRAPE I

(B1, AE4) kiSkis 'scrape off', incl. KAVALAN kiskis 'to shave', AMIS kiskis 'scrape off a surface, strip', YANDENA (CMP) kiki 'scour, scrub off'. Or: (T1, RPTP:130) ki0ki0 'shave, scrape', incl. Kan k-um-akisikisi, Sar k-um-isikisi, Tso m-ru-?si?si 'to shave fur', Bun kiskis-un 'to shave beard', At kə-kiskis 'a file' Mlg kiki-ana 'to be grawed (sic: gnawed ?), scraped'. Add Amis (Fey 1986) kickic 'to cut flesh off the bones'.

搨 [313*, M12833] khriat (kh-r-[i,elt]) 'scratch, scrape' (Shuo Wen, Guang Yun, Yu Pian). Note: Yu Pian and Guang Yun give a second reading kriat (k-r-[i,elt]).

35. TO SCRAPE II

(B1, AE4) kuSkus (sic; apparently a typo for kuSkus) 'scrape': AMI koskos 'scratch one's body', PAI kuskus 'a scraper'; CEB kukhu 'scrape, scatch off something that sticks to a surface with a thrusting motion'. An infix form (B1, AE4) k-ar-uSkus 'scrape', paralleling the Chinese form, is independently reconstructible at the PAN level.

刮 [302n] kwat (k-r-ot) 'scrape' (Zhou Li).

36. TO CLOSE, SHUT

(B1+) kupit 'close, shut', consisting of (B3, AE2) kupit 'close, shut', plus Amis kopit 'close the eyes' (Fey 1986).

閉 [412ab] pit (pit) 'shut, close' (Shi Jing, Shu Jing, Zuo Zhuan); 閉 [405n] pjidh (p-j-it+s) 'shut, close' (Shi Jing).

37. TO SLIP OFF, LOOSE

(Br1+) -lus 'to slip off' consisting of (Br3) -lus 'slip off' (includes such meanings as 'loose', 'fit loosely', 'slip off, free from grip', slipping freely off and on', etc.), and Bunun (Nihira 1983) luslus 'slip, slide', Ami (Fey 1986) ploc 'to pull off, disengage, as to take off a ring', pololoc 'slide down out of position, as for a baby to slide down too low on its mother's back'.

脫 [324m] duat (lot), thuat (hlot) 'peel off' (Lie Zi); 'take off (as clothes)' (Guo Yu); (slip off :) 'escape, disappear' (Lao Zi); 脫 [324e] thuadh (hlot+s), sthjuadh (hl-j-ot+s), ruat (l-j-ot) 'exuviae of insects or reptiles' (Zhuang Zi); 說 [324q] sthjuadh (hl-j-ot+s) 'let loose' (Shi Jing); 脫 [324d] thuadh (hlot+s), duadh (lot+s) 'withdraw, flee' (Shi Jing); and a series of items with semantics such as 'careless, easy, leisurely, pleased, glad', all deriving from 'loose, comfortable', see 324a,m,o.

38. TO GNAW

(B1, AE4) gitgit 'gnaw', incl. Amis gitgit 'gnaw on bones or wood', TB gitgit 'moth'. Note also root (Br1) -geC 'angry, gnash the teeth'.

齧 [279e] giat (get) 'gnaw, crunch in the teeth' (Guan Zi, Li Ji).

39. TO SHOOT

(T1, RPTP:168) panaq2 'shoot': Kan mu-pana?-ə, Sar u-a-pana, Tso pono (pna-a), RukBd[LL] pAna, RukMg u-pnaA, RukTn ua-panA, RukMn u-pana 'to shoot with a bow', Pai panaq, Ami panaq 'arrow', Puy p-an-anah, Bun panah-un, Sai p-um-ana? 'to shoot with a bow', Paz mu-pana? 'to throw at, to hit', To[Cl] fana, shoot with bow'; Malayo-Polynesian cognates were assembled by Dempwolff under the reconstruction (D2) panaq 'shoot, shooting weapon': Tag pana?, Jav panah 'arrow', Ml, MgD panah 'shooting weapon', Fi vana, Sa'a hana 'to shoot', Fu fana 'bow', Sm fana 'to shoot'.

弩 [94z] nagx (na?) 'crossbow' (Zhou Li, Mo Zi, Zhuang Zi); 弩 [94d'] id. 'flint arrow-head' (Shu Jing). An hapax legomenon character: 弩, whose meaning is certainly related to the shooting of arrows and whose pronunciation clearly involves the stem na(?), appears on a Western Zhou bronze. It has been interpreted by Xu Zhongshu (Xu 1934:427) as an early form of 弩 'crossbow' above. The similarity in sound with words for 'crossbow' in some South-East Asian languages (Northern Tai naaB1; Vietnamese na (Sac tone); etc.) has led some scholars (Benedict 1975:109; Norman and Mei 1976:293) to suppose that the name of the crossbow, and the crossbow itself, were borrowed by the Chinese from their southern neighbors. This is unnecessary and unsupported. Use of crossbows by southern peoples is not recorded in Chinese historical works until the Han dynasty, while Chinese crossbows are known to have existed at least since the Warring States period. A passage in Shu Jing: 『若虞機張，往省括于，則釋』(書。太甲上), where the word 機 is understood by most commentators as designating a part of the crossbow mechanism, implies that the crossbow existed in the early archaic period, supporting Xu Zongshu's identification of an early form of the character 弩. From the point of view of the history of techniques, Needham (Science and Civilization in China, Vol. 4, part 2, section 27, p. 224) regards the crossbow as "a characteristically Chinese invention". Indigenous development of Chinese crossbows from simple bows (definitely part of the tool-kit of the Neolithic Chinese, judging from widespread finds of arrowheads in sites all across China, cf. Chang 1986), or from bow snares (such as those used by the Amis, see Fey 1986:217) is the most plausible explanation. Indigenous origin of the Chinese crossbow is supported by the continued use, through the entire archaic period, of a word-stem na? (all with phonetic 奴, a subset of GSR 94) for words relating to the shooting of arrows. The semantic scatter in this set of words (flint arrowhead; crossbow) suggests 'shoot' as original meaning.

40. TO CHEW

(B1, AE4) paqpaq 'chew': Amis pa?pa? 'chew'; MAR papa? 'bite, chew'; note also (D3) se(m)paq 'to chew, chew up'.

哺 [102i'] bagh (N+pa?+s : Mandarin [pu3] reflects the bare stem: OC pa?) 'to chew, have food in the mouth' (Zhuang Zi; Shuo Wen). An infix variant [M3361] 'to chew' reflecting OC pjaq (p-j-a?) is listed in MC dictionaries: Guang Yun, Ji Yun (Ji Yun also ph-j-a?).

41. TO BRAID

(Br1) -pid 'braid, wind together', incl. Puy Hupiz 'to plait rattan or bamboo', Amis ?opir 'braid of hair or rope', (B3, AE3) apid 'braid' plus several isolated WMP words.

纜 [221*] pianx (p[le, iln?]) 'braids of hemp or wheat stalks' (Shuo Wen); 纜 [246e] id.

'plait (the hair)' (Shi Ji, reading after Ci Yuan); also, with intransitively voiced initial: 辮 [Z19*] bianx (ble, iln?) 'plait the hair into braids' (Liang Shu).

42. TO GATHER, BE TOGETHER, ALL

(D1+) pulug 'gather, together', consisting of (D3) pulug 'gather' (Tg pulog, TB pulug 'gathered', Ja pulug 'to harvest', NgD pulug 'assembled'), plus Amis polog 'all, all together' (Fey 1986: the examples include verbal uses as 'to gather, get together'.)

同 [1176a] dug (log) 'to be the same, join, unite' (Shi Jing, Shu Jing, Zhou bronze, cf. JWGL 1035); also [1176] id. 同 'all' (Shu Jing); 銅 [1176d] id. (alloy:) 'bronze' (Zuo Zhuan).

43. TO RISE

(B1, AE4) sakat 'rise, climb up': AMIS cakat 'arise, go up'; BKL sakAt 'go upstairs'; climb, scale, mount, ascend'; MAR sakat 'step up, rise up'; NGADHA (CMP) saka 'ascend, mount (as a horse)'; 'rise (as water in a stream)'; BURU (CMP) saka-h 'ascend to'; saka-t 'up'. Note also (D2) apkat 'to lift, lift up'.

揭 [313n] kjat (k-j-at) 'to rise, raise, lift' (Shi Jing). Also read as gjiat (g-rj-at), kjiat (k-rj-at), kh-j-iat (kh-rj-at).

44. TO CHOP

(B1, AE4) saksak 'hack, chop into pieces' (incl. Amis cakcak 'cut into pieces with a machete, as sweet potatoes', CEB saksak 'chop something into pieces (...)', MLG sasaka 'the half', roa sasaka 'divided into halves, cut into two equal parts').

斫 [798o] tsrjak (ts-rj-ak) 'to cut off' (Gong Yang Zhuan), 'cut in twain' (Mathews). Note: the signific is 斤 'axe', not 刀 'knife', implying cutting done with an axe.

45. TO STICK INTO/TO PIERCE

(Br1) -sek 'insert, stick into a soft surface': (B1, ART) pasek 'wooden pin, wedge; drive in a stake' (based on (D3) pasek 'Pflock', supplemented with unspecified Formosan cognates, which may include Amis pacek 'nail'); and (B1, ART) Cesek 'insert, force into a soft surface', a reconstruction based on unspecified evidence.

刺 [868d] tshjik (tsh-j-ek), tshjigh (tsh-j-ekts) 'pierce, prod, stab' (Shi Jing; Zhou bronze, cf. JWGL 927; Meng Zi); 棘 [868a] tshjigh (tsh-j-ekts) 'thorn' (Shang inscription; Shuo Wen, no text example); 策 [868e] id. 'to prick' (He Guan).

46. TO CRAM/TO STOP UP

(Br1) -sek 'cram, crowd'. This root includes one PAN item, (D2) seksek 'to stuff' (Tg siksik 'sausage', Ho sesika 'stuffed', Fi soso 'to stuff'), for which a Formosan cognate may be cited: Amis cekcek 'to press, crowd together' (Fey 1986). Other reconstructions having root -sek are (B2, AE3) be(n)sek 'overcrowded', (B3, ART) dasek 'press together, compress' etc. Isolated words under this root come mostly from WMP languages (but FIJ oso 'narrow, crowded; strait').

塞 [908a] sak (sik) 'stop up, block' (Shi Jing); 'fill' (Meng Zi); Modern Chinese 'to stop up, to block. To cork. A cork or stopper' (Mathews).

47. TO INSERT

(B1, AE4) selsel 'insert, cram in': Amis celcel 'slip through a crowded place'; JAV sesel 'that which is pushed/inserted into something; push in'; BAL 'be crowded into any place' (clothes etc.).

晉 [378a] tsjinh (ts-j-in+s) 'insert' (Zhou Li); 摺 [378f] id. 'insert' (Li Ji); and 進 [379a] id. 'introduce' (Shu Jing).

48. TO SUCK

(Br1) -sep 'sip, suck', incl. (B1, AE2) qesep, q-um-esep 'suck' (PAI qetep 'swallow liquid when masticating something solid, as sugarcane', q-m-etep 'to "eat" the liquid portion of something, as sugarcane', TIR esef 'suck, sip', etc.); and (D2) sepsep 'to suck' (Tg sipsip, TB sopsop, Ja sesep, Sa'a toto 'to suck'), the latter in fact a PAM item, cf. the cognate set assembled by Tsuchida (T1, RPTP:129): 0əp0əp 'to suck': RukBd 0əpE0əpə, RukMg u-0əpE0əpə, RukTn ua-0əpE0əpə, RukMn səpəsəpə, Pai t-m-əptəp, Ami mi-cepcep, Bun supsup-un, Sa toto 'to suck'. A doublet form with rounded vowel is also reconstructible from Amis and Philippine reflexes: (B1, AE4) supsup 'sip, suck' (Amis copcop 'to suck, as a mosquito sucking blood', BKL supsup 'sip, suck', MAR sosop 'suck').

0晉 [6680] tsəp (ts[i,ulp] 'to sting and suck, as mosquitoes' (Lie Zi); 囉 [679*] id. 'to suck, lick' (Han time text example); 'to suck' (modern Mandarin: Mathews).

NOTE: Note: this comparison may involve an onomatopoeic element, although Blust (ART) does not treat root -sep as onomatopoeic.

49. TO WASH

(B1, AE4) SawSaw 'wash, rinse': Amis sa(w)saw 'generic word for "wash"', Tag hawhaw 'rinsing' (said of clothes).

漂 [1134g] tsagux (tsaw?) 'to wash' (Li Ji); 縹 [1134i] id. 'to bleach' (sc. silk) (Li Ji).

50. TO TIE

(B1, AE4) SikeC 'tie, attach to', incl. Amis siket 'be attached to, joined to', MAR iket 'to tie'. Note: this is identical with (D2) iket 'to tie, make a knot'.

結 [393p] kit (kit) 'to tie, knot' (Shi Jing, Yi Jing, Zuo Zhuan).

51. TO TREMBLE/TO SHAKE

(Br1) -ter 'shiver, tremble'; also (D3) ke(n)Te[r] 'tremble, vibrate'. NOTE: the meaning 'shiver' (from cold) is peripheral in this root. It occurs in only one item: (B1, AE1) terter 'shiver, tremble', and there only in a Maranao reflex, SAI man-tete being just 'tremble'. Another Formosan reflex may be added, also without 'shiver' semantics: AMI terter 'to be shaken or jolted' (Fey 1986).

振 [455p] tjian,h (t-j-in(+s)) 'to shake' (Shi Jing); 震 [455s] id. 'to shake' (Shi Jing, Shu Jing). Both verbs are used transitively and intransitively.

52. TO HAMMER, POUND

(B1, AE3) tuqtuq 'to hammer, pound, crush', incl. Pai tjuqtjuq 'blacksmith's hammer', tj-m-ugtjuq 'to hammer', ILK toto 'to crush, pound, hammer, beat (bark of trees), MLG toto

'the act of pounding with a pestle', LETI (CMP) 'knock, pound, stamp'; or (T1, RPTP:133) tutuH1 'to beat', incl. Kan m-aka-tutu 'to hit with the fist', Tso m-uutu (tutv-a) 'to strike or hit (in general)', Paz kalu-batu m-tutih 'to strike with a stone', etc. See (D3) tutu 'Stampfen' (TB, Ja, Ho tutu 'Stampfen') for WMP cognates of tutuH1 and/or tuqtuq.

槌 [1090r] təgux (tu?) 'to hammer, beat, pound' (Shi Jing: Xiao Ya 197.2: reading after Jing Dian Shi Wen; also Li Ji, Shuo Wen).

NOTE: this comparison may involve an onomatopoeic element.

53. RICE GRUEL/COOKED RICE

(Br1) -buR 'rice gruel; to mix'; this root includes (D2) buDbuD 'pulverize, gruel' (MAL, TB bubuk 'rice gruel'), modified to (D1) buRbuR 'porridge' in ART (see also AE2:49) on unspecified evidence; Also PPh (Zorc) buRbuR 'cook wet rice'. Root -buR also includes such items as HAN lAbug 'rice gruel', KEL menebhur 'rice gruel', TBL libol 'rice soup'. Add Bunun (Nihira 1983) hibul 'to mix'.

飯 [262il] bjanx,h (b-j-[a,oln?]) 'cooked rice or millet' (Li Ji). Note: the Yu Pian says this word was written with phonetic 弁 (indicating rounded vowel, b-j-on? rather than b-j-an?) in the Zhou Shu, a section of the Shu Jing.

54. RICE AS FOOD

(D3) imay 'rice' (Ml imai 'cooked rice', TB eme 'unhusked rice'). Wolff (pers. comm., 1992) cited Amis hemay, Kuv qmay, Puy sumay 'cooked rice', Ibg ammay, Ceb humAy, Simalungun Batak omE, TB eme 'Oryza sativa', reconstructing PAN (= level 1) semay (in standard orthography: Semay) 'rice as food' despite phonological irregularities.

米 [598a] midx (mij?) 'peeled grain, rice' (Shang inscriptions, Shu Jing, Zuo Zhuan, Shuo Wen, etc.).

55. HOUSE

(T1, RPTP:232) Rumaq1 'house': Sar ruma?-a 'lair or den of beasts', Tso rmoo 'house', Pai umaq 'id.', Puy rumaH, Ami lumaq, Bun lumah, Paz xuma?, Sa'a lume 'house'. For other extra-Formosan cognates (TB, Ml, NgD, Ho) see (D2) Rumaq 'house'.

廡 [103il] mjagx (m-j-a?) 'large house' (Guan Zi, Shi Ming); 'side buildings or covered hallways around a yard' (Chu Ci, Shuo Wen).

56. TORCH/FIRE

(T1, RPTP:154) DZamaR 'torch': Paiwan zama-n 'torch', Ami lamal 'fire', Tonga ama 'to fish by torch-light', Samoa lama 'torch', and other MP items under (D2) damaR 'resin, torch, light', q.v. Other Formosan forms show irregular correspondences, suggesting borrowing: in LWFL are found Bunlsh dmal 'lightning' (irregular first vowel, damal expected); RukBd DAmara, RukMg DmArə, RukTn DamAə RukMn damara 'moon' (-r- irreg.); also Sar Lamar 'to burn' (cited by Tsuchida 1976:77; s- expected). Other possibly related forms are: Kavalan zamaR 'fire'; Ketagalan (Basay dial.) namaD, (Trobiauan dial.) zamal 'fire' (cited by Li 1992:282); Favorlang ramal 'a burnt field' (Happart 1650 [1896]); Siraya lamag 'burning' (cited by Ferrell 1969:88; but I could not find it in Van der Vlis' vocabulary).

火 [353a] hmərx (hmij?) 'fire' (Shang insc., Shi Jing, Shu Jing); 燄 [356b] hmjərx (hm-j-ij?) 'blazing fire' (Shi Jing); 炬 [164s] id. 'fire' (Zhou Li).

Although the scarcity of cultural items in the above list is not necessarily meaningful (it is perhaps the case that non-cultural items tend to predominate in the lexicon of reconstructed languages such as PAN), it is noteworthy and significant that a majority of items in the list are not of a kind that would normally lend itself to borrowing in a limited contact situation. Note also that the list contains not only basic vocabulary items but also over 50 p.c. of verbs, a word class believed to resist borrowing well. Also noteworthy, though somewhat enigmatic at this point, is the frequent association of verbs and PAN roots.

Missing from the list are personal pronouns and numerals, two very stable word classes in IE languages. This is not abnormal in the East Asian context: avoidance of personal pronouns is a widespread cultural characteristic in East Asia, resulting in a high rate of replacement for this word class. Numerals also are less stable in East Asia than they are in the IE languages: numerals from 2 to 10 were borrowed by Thai from Chinese and from Yi by Gelao (Edmondson and Thurgood 1992); Chinese numerals from 1 to 10 were borrowed by Miao-Yao.

IV. SOUND CORRESPONDENCES.

The above comparisons illustrate the correspondence of OC monosyllabic word stems with the final syllable of PAN words, both C1V(C2) in structure: affixes, whether OC or PAN, have been disregarded. The ancestor language ('Proto-Sino-Austronesian', or PSAN) is assumed to have been polysyllabic like PAN. The shift to Chinese monosyllabism occurred through the loss of non-final syllables, a typological shift recently at work in Huihui, a Chamic language on Hainan (Haudricourt 1984) and currently in progress in various AA languages (Ferlus 1979, Diffloth 1980). Correspondences between PAN and my OC reconstructions, slightly modified from Baxter (1992), are stated below for each position. Comparison numbers between rounded brackets involve ambiguously reconstructed Chinese vowels or consonants.

PAN	OC	Comparison No.
C1:		
-p-	p(h)-	17, 19, 36, 40, 41
-t-	t(h)-	5, 8, 15, 51, 52
-k-	k(h)-	16, 20, 24, 25, 26, 31, 32, 33, 34, 35, 43, 50
-q-	ʔ-	3, 11
-b-	b-	21, 28, 29, 53
-D-	d-	6, (23)
-D-	t-	27 !
-g-	g-	30
-m-	(h)m-	54, 55, 56
-n-	n-	2, 4, 7, 18, 39
-ŋ-	ŋ-	38
-l-	(h)l-	1, 9, 12, 13, 14, (23), 37, 42
-r-	r-	10
-s-	s-	22, 46
-s-	ts-	44, 45, 47, 48
-S-	ts-	49

U:

-a- : -a- 5, 8, 10, 12, 17, 19, 20, 21, 28, 39, 40, 43, 44, 49, 54, 55
-a- : -i- 27, 56
-i- : -i- 18, 22, (34), 36, (41)
-i- : -e- (34), 38, (41)
-u- : -u- 1, 2, 9, 26, 52,
-u- : -o- 4, 6, 14, 15, 16, 24, 25, 29, 35, 37, 42, 53
-e- : -i- (7), (23), (30), 31, 46, (48), 51
-e- : -e- 32, 33, (34), 45
-e- : -i- 3, (7), 13, 47, 50 (preceding dentals, except No. 7)
-e- : -u- 11, (23), (30), (48) (preceding labials)

C2:

-p : -p 32, 48
-t : -t 26, 36, 38, 43
-C : -t 16, 29, 50
-s : -t 28, 33, 37
-S : -t 34, 35
-k : -k 15, 24, 27, 44, 45, 46
-q : -? 2, 5, 8, 40, 52
-q1 : -? 55
-q2 : -? 39
-Q1 : -? 10
-H1 : -? 1, 4, 52
-H2 : -? 10
-m : -m 8, 11, 23, 30, 31
-l : -n 25, 47
-N : -n 22
-r : -n 51
-R : -n(?) 3, 14, 53
-d : -n? 41
-j : -n? 13
-g : -g 21, 42
-R : -j(?) 6, 9, 12, 19, 56
-y : -j? 54
-u : -u(?) 20, 49
-0 : -0 17

With adjustments due to the change in OC transcription system, these correspondences are by and large those proposed in my earlier work. The interested reader is referred to my recent publication (Sagart 1993) for a discussion of the sound changes involved. A short outline will be given here.

Correspondences for the C1 position show a basic agreement in consonantal voicing.

Voicelessness among OC sonorants is assumed to be secondary, perhaps the trace of a lost voiceless prefix. The OC aspiration contrast among voiceless stops is thought to reflect a PSAN contrast of singleton vs. cluster among voiceless consonants in word-internal position. That PSAN contrast, lost in Formosan, is continued in PMP by nasal accretion. Only in a small number of cases can MP nasal accretion be explained as an innovation on the basis of reconstructible PAN forms.

Each of the four PAN vowels has two correspondents in Chinese, one open, one close. It is noteworthy that at least three of these four pairs of Chinese vowels (i-e, u-o, i-a) are recognized ablaut sets in OC: both members of the same pair frequently alternate in word families. With the fourth pair (i-e), the lower vowel is often replaced by (i) preceding a front ending, and by (u) when preceding a labial ending.

With final consonants, nasals and voiceless stops correspond between OC and PAN. The PAN voiced stops are presumed inherited from PSAN, and evolved to clusters of homorganic nasal + glottal stop in Chinese. The sounds symbolized as PAN *-q, *-Q and *-H merged as *-ʔ in OC, aligning with *-ʔ from PSAN voiced stops to form the Tone B category. Support for my earlier proposal that PSAN final -s changed to -h in OC, later to become the tone C category, is missing from the present list of comparisons: this makes it doubtful. The present list of comparisons provides support for a correspondence of PAN -s to OC -t.

That a coherent, non-stratified system of sound correspondences accounts for the phonetics of the above comparisons is an indication that the common vocabulary is not shared by OC and PAN as a result of a long period of intimate contact between them.

V. SOUND CORRESPONDENCES IN THE BASIC VOCABULARY.

A minimal argument for sound correspondences in the basic vocabulary, based on 6 basic vocabulary items listed in Table 1, is presented here. Ten sound correspondences interlock to account for the segments in the final syllables of the PAN reconstructions and in the OC word-bases. Apart from the correspondences of PAN -t- to OC th-, and PAN -D- to OC d-, which must be stipulated if only the comparisons in Table 1 are considered (support for these has, however, been provided in section IV), each sound correspondence is exemplified twice. The correspondences are listed in Table 2.

PAN		OC	
1. quluH1	'head'	首	hl-j-u? 'head'
2. nunuH1	'breasts, milk'	乳	n-j-o? 'milk, suckle; nipple'
3. punuq	'brain'	腦	nu? 'brain'
4. -taq	'mud; earth, ground'	土	tha? 'earth'
5. SulaR	'snake'	蛇	l-j-aj 'snake'
6. tuDuR	'to sleep'	睡	d-j-ojts 'to sleep'

Table 1. Some lexical comparisons between PAN and OC.

C1	V	C2
-l- : (h)l- 1,5	-u- : -u- 1,3	-H1 : -ʔ 1,2
-n- : n- 2,3	-u- : -o- 2,6	-q : -ʔ 3,4
-t- : th- 4	-a- : -a- 4,5	-R : -j 5,6
-D- : d- 6		

Table 2. Some sound correspondences between OC and AM.

VI. EXTRA-LINGUISTIC EVIDENCE

As outlined in section I, independent work by specialists of Chinese and Austronesian cultural history points to the east China coast as a key area in the emergence of both the Chinese and AM cultures.

Looking back in time from the early dynastic period, a continuous cultural sequence connects the Shang civilization, the predynastic lords in the Shangqiu area, the Shandong Longshan culture of the late 3rd millennium BC, and the neolithic cultures of Dawenkou and Qingliangang in south Shandong and north Jiangsu, whose earliest sites have been dated around 5000 BC.

Further to the south another continuous sequence leads from the historical AM-speaking societies of Taiwan back to the late 5th millennium BC ceramic sites at Ta-P'en-K'eng in northern Taiwan and across the Taiwan straits, at Fu-Kuo-Tun on Quemoy island (Chang Kwang-Chih 1989), and beyond, to neolithic sites in south-eastern coastal China where, as in Hemudu of north Zhejiang, rice was grown, pile dwellings built, and domestic pigs, dogs and water buffaloes kept around 5000 BC.

Convergences of many kinds exist between these two coastal sequences, the southern Hemudu-Fuguodun-Formosa sequence ancestral to the AM culture, and the northern Dawenkou-Shandong Longshan-Shang sequence which is one of the main strands in the development of the Chinese culture.

Based on evidence of archaeological continuity between them, the Manjing Museum group (1978) and Pearson (1983:137) have argued that the neolithic cultures of coastal south Shandong, Jiangsu and north Zhejiang are regional variants of the same culture, called by them the Qingliangang culture. Pearson (1983:124) regards Hemudu as "the forerunner of the Ch'ing-lien-kang [= Qingliangang] culture in terms of general adaptation and pottery styles". With respect to the division between southern and northern variants of the Qingliangang culture, Pearson also notes that "cultivation equipment appears to have developed earlier in the south and then spread to the north" (1983:138).

Another trait shared by the two traditions is tooth extraction. Skulls unearthed at the neolithic site of Dawenkou in south Shandong had had their upper lateral incisors extracted between 12-13 and 18-21 years of age in both males and females (Yan Yan 1972). Extraction of upper lateral incisors as a puberty rite is widespread among the modern AM populations of Formosa, for instance Bunun (Mihira 1983, under "bintus"), Atayal (Turner and Lien 1984),

Saisiyat, Tsou (Ferrell 1966). It is also attested archaeologically (Chang 1957).

Last but not least, physical anthropology provides evidence of close physical proximity between populations associated with these two traditions. Of particular interest is Pietrusewsky's multivariate investigation of 53 series of human skulls from Asia and the Pacific (Pietrusewsky 1992). The study showed Formosan aboriginals, represented by 36 Atayal skulls, to be closest, in a cluster analysis of distances, to Chinese of Taiwan and Hainan Island, AND SHANG DYNASTY CHINESE FROM ANYANG IN NORTHERN CHINA. Based on these results alone, Pietrusewsky (1992:326) suggested "an older and possibly separate ancestry of these groups [i.e. Formosan represented by Atayal, modern Chinese from the southeast coast, and Shang dynasty Chinese, L.S.], one which was distinct from the more northerly Asiatic mainland mongoloid groups such as Japanese, Korean and Chinese" ('Chinese' at the end of the citation refers to a series of skulls of soldiers of the Mandchu army, who were quite likely Mandchu themselves).

VII. CONCLUDING REMARKS.

Based on the morphological argument detailed in section II, and on the argument from sound correspondences in the basic vocabulary presented in section IV, the conclusion must follow that OC and PAN are genetically related. Note that the arguments from morphology and from sound correspondences in the basic vocabulary are independent from one another, and that each one, taken separately, would suffice to demonstrate the relationship. As shown in section VI, such a relationship also makes good sense in terms of archaeology and physical anthropology.

The same conclusion had been reached earlier in this century by Conrady (1916, 1923) and Wulff (1942) on the basis of fragmentary evidence. Progress, both qualitative and quantitative, made in our knowledge of OC and PAN since the time of Wulff and Conrady has allowed the insight of these scholars to rest on firmer methodological ground.

The claim made here of a genetic unity between OC and PAN is not exclusive of other links between either language and other East Asian language groups: Austric, Sino-Tibetan, Austro-Tai are all plausible constructs, though the evidence for each is unequal and diverse in nature. It is implied, however, that any valid relationship between either OC or PAN and another language is valid for both OC and PAN. For instance, if Chinese and the TB languages are truly related, then the TB and AN groups must also be related. A systematic comparison of TB and AN has never been attempted, yet *prima facie* evidence exists for a TB-AN unity.

In morphology, PAN and TB share a causative prefix: PAN pa-, TB p(ə)- (Empeo pē-, Mikir pe-, pi-, pa-, Dimasa p'a-, p'u-, Bodo fa-, fi-), pa- also a causative prefix in AA; and a nominalizing suffix: PAN -en (e.g. *ka 'to eat', *ka-en 'food'), TB -n (e.g. WT za 'to eat', za-n 'food'). The WT prefix m- had, according to Wolfenden (1929: 26 sq.), the power to derive adjectives and intransitive verbs, much like PAN ma- and OC M-. There is also some evidence that the TB medials -r- and -y- were infixes (von Koerber 1935, Miller 1958, Gong 1977) and that some of their functions (in particular as described for WT by von Koerber) were shared with the corresponding Chinese and AN infixes.

In lexicon, potential PAN-TB cognates (some of which are shared with Chinese and/or other Asian language groups) are listed in Table 3:

tongue	PAN	Dilaq	PTB m-lyak, s-lyak
brain	PAN	punuq	PTB nuk
head	PAN	quluH1	** lu
breasts	PAN	nunuH1	PTB nuw = nəw
mother	PAN	ina	PTB m-na
day	PAN	daqaNi	PTB ney
slip off	PAN	-lus	PTB g-lwat 'free, release'

Table 3: some potential PAN - TB cognates. PTB forms from Benedict (1972); ** lu 'head' in Lushai, Lakher, Laizo, Komren, Khumi (Matisoff 1991).

Detailed studies will be necessary to investigate this and other connexions.

REFERENCES

- Baxter, W. H. III (1992) A Handbook of Old Chinese phonology. Trends in Linguistics Studies and Monographs 64. Berlin: Mouton de Gruyter. 922 pages.
- Bellwood, P. (1991) "The Austronesian Dispersal and the origin of languages". Scientific American (July 1991). pp. 88-93.
- Benedict, P.K. (1972) Sino-Tibetan: a conspectus. Cambridge: University Printing House. 230 p.
- Benedict, P.K. (1975) Austro-Thai: language and culture. HRAF Press. 490 p.
- Blust, R. (1970) Proto-Austronesian addenda. Oceanic Linguistics Vol. IX, 104-62.
- Blust, R. (1980) Austronesian Etymologies. Oceanic Linguistics Vol. XIX numbers 1 and 2, 1-181.
- Blust, R. (1983-84) Austronesian Etymologies II. Oceanic Linguistics Vol. XXII-XXIII numbers 1 and 2, 29-149.
- Blust, R. (1985) Austronesian Etymologies III. Oceanic Linguistics Vol. XXV numbers 1 and 2, 1-123.
- Blust, R. (1988a) Austronesian root theory. Amsterdam: John Benjamins.
- Blust, R. (1988b) The Austronesian Homeland: a linguistic perspective. Asian Perspectives XXVI, 1: 45-67.
- Blust, R. (1989) Austronesian Etymologies IV. Oceanic Linguistics Vol. XXVIII number 2, 111-80.
- Blust, R. (1992) The position of the Formosan languages: Method and theory in Austronesian comparative linguistics. In Papers for International Symposium on Austronesian studies relating to Taiwan, Institute of History and Philology, Academia Sinica, pp.21-55.
- Bodman, M. (1980) "Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship". In Van Coetsem and Waugh (eds.) Contributions to historical

- linguistics, 34-199. Leyden: E.J. Brill.
- Buck, C. D. (1949) A dictionary of selected synonyms in the principal European languages. Chicago: Chicago University Press.
- Chang Kuang-chih (1957) Yuanshan chutu de yi ke ren chi. Kaogu Renlei Xuekan 9, 10:146-7
- Chang Kuang-chih (1968) The archeology of ancient China. Newhaven and London: Yale University Press. 2ème édition, 1968.
- Chang Kuang-chih (1986) The archeology of ancient China. Newhaven and London: Yale University Press. 4ème édition, 1986. 450 p.
- Chang Kuang-chih (1988) "Shang civilization". Newhaven and London: Yale University Press.
- Chang Kuang-chih (1989) "Xin Shiqi shidai de Taiwan Haixia" [the Taiwan straits in neolithic times]. Kaogu 261: 241-50.
- Chou, Fa-kao (1978) "Lun Shang-ku In ho Ts'ieh-Yun In". Journal of the Institute of Chinese Studies of the University of Hongkong, Vol. III, No. 2, 321-457.
- Chou, Fa-kao (1974) Chin-Men Ku-Lin [Etymological dictionary of ancient Chinese bronze inscriptions]. 16 vols. Hongkong: The Chinese University Press.
- Conrady, A. (1916) Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen. Aufsätze zur Kultur- und Sprachgeschichte E. Kuhn gewidmet, 475-504. München.
- Conrady, A. (1923) Neue austrisch-indochinesische Parallelen. Hirth anniversary volume (Asia Major introductory volume), 23-66.
- Dahl, O. C. (1981) Early phonetic and phonemic changes in Austronesian. Instituttet for sammenhengende kulturforskning Serie B: Skrifter LXIII. Oslo: Universitetsforlaget. 175 p.
- Dempsolff, O. (1938) Vergleichende Lautlehre des Austronesischen Wortschatzes. Band 3: Woerterverzeichnis. Beihefte zur Zeitschrift für Eingeborenen Sprachen, 19. Berlin: Dietrich Reimer.
- Downer, G. (1973) Strata of Chinese loanwords in the Mien dialect of Yao. Asia Major XIII, 1: 1-33.
- Dyen, I. (1992) The internal and external classification of the Formosan languages. In Papers for International Symposium on Austronesian studies relating to Taiwan, Institute of History and Philology, Academia Sinica, pp. 162-96.
- Edmondson, J. and G. Thurgood (1992) Gelao reconstruction and its place in Kadai. Paper presented at the 25th ICSTLL, 14-18 Oct. 1992, Berkeley.
- Egerod, Søren (1988) Atayal-English Dictionary. Scandinavian Institute of Asian Studies monograph 35. London and Malmo: Curzon Press.
- Egli, H. (1990) Paiwangrammatik. Wiesbaden: Harrassowitz. 349 p.
- Ferrell, R. (1966) The Formosan tribes: a preliminary linguistic, archaeological and cultural synthesis. Bulletin of the Institute of Ethnology, Academia Sinica, 22: 97-127.
- Ferrell, R. (1968) The Pazeh-Kahabu language. Bulletin of the Dept. of Archaeology and Anthropology, National Taiwan University, No. 31-32: 73-96.
- Ferrell, R. (1969) Taiwan Aboriginal groups: problems in cultural and linguistic classification. Monograph No. 17, Institute of Ethnology, Academia Sinica. Nankang:

Academia Sinica.

- Ferrell, R. (1982) *Paiwan dictionary*. Pacific Linguistics C73, Canberra, A.N.U.
- Fey, V. (1986) *Amis dictionary*. The Bible Society in the Republic of China.
- Gong, Huang-chenq (1977) "Ku Tsang-wen ti y chi ch' i hsiang-kuan wen-t' i" [the classical Tibetan y and its related problems]. *BIHP* 48, 2: 205-28.
- Happart, Gilbertus (1650 [1896]) *Wordboek det Favorlang'sche Taal*. Published in English translation as "Happart's Favorlang vocabulary" in W. Campbell (ed.) *The articles of Christian Instruction in Favorlang-Formosan Dutch and English*. London; Kegan Paul, Trench and Trübner, 1896, pp. 122-99.
- Haudricourt, A.-G. (1984) *La tonologie des langues de Hainan*. *Minzu Yuwen* 1984: 4 17-25.
- Karlgren, B. (1964) *Grammata Serica Recensa*. Göteborg: Elanders Boktryckeri Aktiebolag.
- Li Fang-kui (1971) *Shang-ku In Yen-chiu*. *Tsinghua Journal of Chinese Studies, New Series* IX, numbers 1 and 2: 1-61.
- Li Fang-kui (1976) *Chi-ke Shang-ku Sheng-mu Wen-ti*. In Szu Ch' i-liang et al. (eds), *Tsung-tung Chiang Kung Shih-shih Chou-nien Chi-nien Lun-wen-chi*, 1143-50. Taipei: Academia Sinica.
- Li, Paul J.-K. (1978) "A Comparative vocabulary of Saisiyat dialects". *Bulletin of the Institute of History and Philology, Academia Sinica* 49, 2: 133-99.
- Li, Paul J.-K. (1981) "Reconstruction of Proto-Atayalic Phonology". *Bulletin of the Institute of History and Philology, Academia Sinica* 52, 2: 235-301.
- Li, P. J.-K. (1992) "Formosan vs. non-Formosan features in some Austronesian languages in Taiwan". In *Papers for International Symposium on Austronesian studies relating to Taiwan*, Institute of History and Philology, Academia Sinica, pp. 268-84.
- Li Zhiguo (1991) "Yi Meng fangyan de 'fenyinci'". *Fangyan* 1991, 3: 206-10.
- Marsh, Mikel A. (1980 [1977]) *Favorlang-Pazeh-Saisiat: a putative Formosan subgroup*. PhD dissertation, Washington State University. Ann Arbor and London: University Microfilms International.
- Matisoff, J. (1991) *New goals in ST linguistics'*. Paper presented at the 24th ICSTLL (Bangkok, 7-11 Oct. 1991)
- Meillet, A. (1937) *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette.
- Miller, R. A. (1958) *The Tibeto-Burman Infix System*. *Journal of the American Oriental Society* 78, 3: 193-204.
- Moravcsik, Edith A. (1977) *On rules of infixing*. Bloomington: Indiana Linguistics Club.
- Manjing Museum (1978) "Changjiang xiayou xinshiqi shidai wenhua ruogan wenti de tanxi" [an investigation of some questions relating to the neolithic cultures of the lower Yangzil]. *Men Wu* 4: 46-57.
- Mihira, Y. (1983) *A Bunun Vocabulary: 2nd edition*. Privately published.
- Norman, J., and Tsu-lin Mei (1976) *The Austroasiatics in ancient south China: some lexical evidence*. *Monumenta Serica* Vol. XXXII, 274-301.
- Ogawa, Naoyoshi, and Erin Asai (1935) *The myths and traditions of the Formosan native tribes*. Taipei.
- Panganiban, Jose V. (1973) *Diksyunario Tesauro Pilipino-Ingles*. Lungsod Quezon:

Manlapaz.

- Pearson, R. (1983) "The Ch'ing-lien-kang Culture and the Chinese Neolithic". in Keightley (ed.) The origins of Chinese civilization. Berkeley: University of California Press. 617 p.
- Peyros, I., and S. Starostin (1984) Sino-Tibetan and Austro-Tai. Computational Analyses of Asian and African Languages 22, 123-7.
- Pietrusewsky, M. (1992) "Taiwan aborigines, Asians and Pacific Islanders: a multivariate investigation of skulls". In P. Naylor (ed.) Austronesian studies: papers from the 2nd Eastern Conference on Austronesian Languages. Ann Arbor. Pp. 319-60.
- Pulleyblank, E. (1973) Some new hypotheses concerning word families in Chinese. Journal of Chinese Linguistics 1, 1:111-25.
- Sagart, L. (1993) "Chinese and Austronesian: evidence for a genetic relationship". Journal of Chinese Linguistics 21, 1: 1-62.
- Starosta, S., A. Pawley and L. Reid (1982) The evolution of Focus in Austronesian. In Halim, Carrington and Hurm (eds.) Papers from the Third International Conference on Austronesian Linguistics, vol.2, 145-70. Pacific Linguistics, C-75 (1982).
- Tang Zuofan (ms) Shanggu Yin Shouce [Manual of Old Chinese pronunciation]. No indication of editor.
- Tsuchida, S. (ms.) List of words of Formosan Languages. Manuscript, dated 1971.
- Tsuchida, S. (1976) Reconstruction of Proto-Tsuoic phonology. Study of languages and cultures of Asia and Africa monograph series 5. Tokyo.
- Tung, T'ung-ho (1967) Shang-ku In-Yun Piao Kao Taipei: Institute of History and Philology monograph series B, NO 21. 249 p.
- Qiu Xigui (1978) Hanzi xingcheng wenti de chubu tansuo. Zhongguo Yuwen 1978, 3: 162-71.
- Turner, Christy G. II and Chao-Mei Lien (1984) Diachronic differences in Taiwan dental morphology. Bulletin of the Indo-Pacific Prehistory Association, 5: 74-82. Canberra: Australian National University.
- Von Koerber, H. N. (1935) Morphology of the Tibetan language. Los Angeles and San Francisco: Suttonhouse. 230 p.
- Wang Li (1982) Tongyuan Zidian [Dictionary of word families]. Beijing: Shangwu. 695 p.
- Wulff, K. (1942) Ueber das Verhaeltnis des Malayo-Polynesischen zum Indochinesischen. Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelser XXVII, 2. København: Ejnar Munksgaard. 157 p.
- Yan Yan (1972) "Dawenkou xin shiqi shidai rengu de yanjiu baogao". Kaogu Xuebao 1972, 1: 92-122.
- Zorc, R. D., and M. Charles (1971) Proto-Philippine Finder List. Cornell: typescript.