SINO-TIBETO-AUSTRONESIAN: AN UPDATED AND IMPROVED ARGUMENT
Laurent Sagart

To cite this version:

HAL Id: halshs-00085059
https://halshs.archives-ouvertes.fr/halshs-00085059
Submitted on 11 Jul 2006

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.
SINO-TIBETO-AUSTRONESIAN:
AN UPDATED AND IMPROVED ARGUMENT

L. Sagart
CNRS, Paris, France

contents

Sino-Austronesian, version I .................................................................................................... 1
Sino-Austronesian, version II ................................................................................................... 1
Higher AN subgrouping (adapted from Ho 1998) ................................................................... 2
Support for ECL: six innovations shared exclusively by PMP and Formosan East Coast
languages .................................................................................................................................. 2
Some PAN and PECL reconstructions ..................................................................................... 2
Sino-Tibeto-Austronesian lexical comparisons ....................................................................... 5
sound correspondences ............................................................................................................. 7
morphology .............................................................................................................................. 8
references ............................................................................................................................... 13

Sino-Austronesian, version I

(Sagart 193, 1994)
1. Blust's higher AN subgrouping
2. Austronesian reconstructions by Dempwolff and Blust. MP material deemed essential.
3. only Chinese and Austronesian related
4. sound correspondences between final syllable of AN words and Chinese
   monosyllables.
5. some morphology
6. criticisms: too little basic vocabulary; Tibeto-Burman outside the relationship.

Sino-Austronesian, version II

1. East Coast Linkage (ECL) is one primary branch of PAN; Malayo-Poynesian forms
   one clade within ECL.
2. all Austronesian reconstructions either include Formosan material or are based on
   Formosan exclusively. MP material not essential.
3. Tibeto-Burman is reintegrated: relationship is between Sino-Tibetan and
   Austronesian.
4. some modifications in the sound correspondences; in addition, initial consonant of
   "lost" penultimate syllable of AN words governs Chinese syllable type.
5. more morphology
6. more basic vocabulary
Higher AN subgrouping (adapted from Ho 1998)

Support for ECL: six innovations shared exclusively by PMP and Formosan East Coast languages

- PAN *C => *t (Siraya, Bunun, Amis, Kavalan, Basay-Trobiawan, PMP: Ferrell 1969)
- PAN *N => *n (Kanakanabu, Bunun, Amis, Kavalan, Basay-Trobiawan, PMP: Ho 1998)
- 'bird' PAN *ayam replaced by *manuk (Basay-Trobiawan, PMP)
- 'banana' PAN *belbel replaced by *pu(n)ti (Basay-Trobiawan, PMP: Li 1992:277)
- 'to count' PAN *supeR replaced by *bilang (Kavalan, Basay-Trobiawan, PMP: Li 1992:277)
- pang-V instrumental construction (Amis, PMP: Starosta 2001)

Some PAN and PECL reconstructions

**body hair**
PAN gumiN ‘body hair’. Puyuma (Cauquelin 1991) gumul 'body hair, feather'; Saisiat (Ogawa and Asai 1935:6) komol ‘pubic hair’. The analogy of PAN *kumis ‘beard’ may or may not have played a role in the reflexion of PAN *g- as k- in the Saisiat form (the Saisiat reflex of PAN g- is uncertain).

**bone**
PAN kukut ‘bone’. Dempwolff reconstructed UA *kukut ‘joint, articulation’ based on corresponding forms in Javanese and Ngaju Dayak. Add Favorlang *o’ot ‘bone’ (PAN *k goes to zero in Favorlang).

**foot**
PAN kakay 'foot'. Atayal kakai, Sediq qaqay, Thao kakai, all 'foot' (Ferrell 1969).
palm of hand
PAN dapa 'palm of the hand', consisting of Blust's PWMP (1980 # 112) Da(m)pa 'palm of the hand, sole of the foot' plus Atayal (Egerod 1980) rapa? 'palm of the hand'.

horn, antler
PAN (q)uRuŋ 'horn/antler': Pazeh ?uxúŋ; Sediq uruŋ (Ferrell 1969), oloŋ (Pecoraro 1977); Yami ?urúŋ, all 'horn' (Pazeh, Sediq and Yami forms cited from Ferrell 1969: 136); add Kanakanabu ta?uruŋu, Saaroa tauruŋu 'pygmy deer' (Tsuchida ms), both with prefixed ta- and final echo vowel; Long Nawang uheng, Uma Juman hueng 'horn, antler'.

Note: Blust (1973) reconstructed PAN *(q)uReng 'horn, antler', based on the Pazeh, Sediq and Yami forms above, plus reflexes in two WMP languages: Long Nawang uheng, Uma Juman hueng 'horn, antler'. I reconstruct /u/ in the final syllable because Formosan and Extra-Formosan languages belonging to three primary branches (Pazeh-Sediq; Kanakanabu-Saaroa; and Yami) agree on /u/. I assume the Long Nawang and Uma Juman forms exhibit a secondary (WMP) u > o dissimilatory change; Blust's decision to reconstruct /e/ in the last syllable reflects the greater weight given Extra-Formosan languages in his practice.

Note: Dempwolff had reconstructed "Ur-Austronesisch" *suŋu 'horn'. Tsuchida (1976:165) reconstructed another form in the same meaning: "Proto-South-Formosan" *seqųŋu, with reflexes in Saaroa, Saisiat and Paiwan. He regarded Dempwolff's *suŋu as a metathesized form of his own *seqųŋu. Both *seqųŋu/suŋu and *(q)uRuŋ appear to have been PAN words; their referents must have been different, perhaps as 'horn' (hollow, permanent, non-branching) vs. 'antler' (solid, shed annually, branching).

chicken
PAN kuka 'chicken, fowl': Paiwan koka, Rukai (Taramakau) koka, Siraya kuka.

Note: several Formosan languages have prefixed elements ta-, tari-, taru-, toro-: Siraya takuka, tahuka, Rukai (Maga) torokuka, Kanakanavu tariku:ka, Saaroa torokooka, etc.

Note: Li (1992: 271; 1993:305) reconstructed "Proto-Formosan" *teRakuk, based on Bunun, Puyuma, Kavalan and Amis forms, and interpreting Tsou tro?ua, Kan tarikuka and Saaroa turukuuka as containing a suffixed formative -a of no particular function. The forms reflecting teRakuk, without final -a, are found exclusively in languages of the East coast, a well-known diffusion zone. It is more natural to suppose that a word for 'chicken' something like *teRakuk was reanalyzed as containing root -kuk 'sob, cackle', and was for that reason changed to Li's *teRakuk in a language of the East coast; and that the word then spread to other East coast languages. The alternative is to suppose that a suffixed -a of unknown nature was added to teRakuk independently in proto-Tsouic, in Taramakau Rukai, in Siraya and in Paiwan.

cave, hole
PAN b[e]luŋ 'cave, hole'. Paiwan beruŋ 'hole', Tsou (Tung) feóŋo 'cave, hole', Pazeh waruŋ 'cave, hole'.
cloud
PAN (root) -qem 'cloud': Amis toʔəm 'clouds' (Fey 1986); Rukai (Budai) əmə:ma 'cloud' (Tsuchida ms); Bunun lohom (Ogawa and Asai 1935), luhum (Nihira 1983) 'cloud'; this root also includes a WMP word *q-al-emqem reconstructed here on the basis of a cognate set first assembled by K. Wulff (1942, No 53): Sundanese halömhöm 'overcast', Ceb. ?alumʔum 'for the sky to be overcast'.

to carry
PECL baba 'carry on back'. Dempwolff reconstructed baba 'to bring along' (incl. several items meaning 'to carry on back'). Add Amis fafa 'to carry on back' (child), Basai baba 'carry on back', Trobiawan uvava 'id.'.

to cut off, short

to drown, disappear
PAN Nemes 'to drown, disappear'. Dempwolff reconstructed *lemes 'to drown, immerse'. Add Puyuma (Cauquelin 1991) lemes 'disappear, destroy, to end'.

to fall
PAN (root) -luR 'to fall'. Dempwolff reconstructed quluR 'drop slowly' (Tg 'fall'). Add Atayal (Egerod) tarui < taluR 'to fall down', with prefixed ta- 'agentless action'.

to follow
PECL d[u]NuR 'to follow'. Blust (1980 #109) reconstructed (3) duluR 'accompany, go together with'. Add Siraya dmyllough 'to follow' (Adelaar 1999:321), equivalent to d-m-ilux (with short i) in Adelaar's restored notation.

to put together
PECL puluN 'to gather', consisting of Dempwolff's puluN 'to gather', plus Ami polon 'all, all together' (Fey 1986; the examples include verbal uses as 'to gather, get together'.)

to speak, say
PAN kawaS1, kaw(a)kawaS1 'speak, say': Pazeh ka:kawas 'speak, say' (Ferrell 1969: 313), Basai makaukawas 'say', kawakawase 'story', Trobiawan kawas 'story'.

to sleep
PAN (root) -zem 'close the eyes, sleep'. Based on NW Formosan (ma)-qizem 'sleep' (Saisiat maʔərəm 'sleep', Pazeh ʔidəm, ma-ıdəm 'sleep': Ferrell 1968:90) plus WMP ke(zZ)em 'close the eyes' (Blust 1973) and WMP peZem 'close the eyes' (Blust 1973), dbl. of the preceding.

to ruin

broad
PAN (root) -baŋ 'broad', consisting of 'candidate root' -baŋ 'broad' (Blust 1988:73), based on 3 WMP items (probably including Dempwolf's baŋbaŋ 'broad, wide' and Blust's PWMP lak(e)baŋ 'broad, wide'), plus Atayal (Egerod) labaŋ 'broad', Sediq (Li 1981) Ilbaŋ 'wide'.

**hot**
PECL qaiŋet 'hot'. Dempwolf reconstructed qaNet 'hot' (Dempwolf 1929); 'warm breath' (Dempwolf 1938). Add Siraya (Matthew Gospel) ainɡit 'fever' (Adelaar, p.c., 1999).

**old**
PAN (root) -daŋ. Dempwolf reconstructed guDaŋ 'adult, grown-up'. Add Puyuma iDaŋ 'old' (inferred from maranai DaiDaŋ 'oldest, Huang 2000a: 49); Rukai maraDaŋ 'old' (Ferrell 1969).

**thick**

**loincloth**
PAN sabuk 'loincloth'. Dempwolf reconstructed sabuk 'loincloth'. Add Atayal (Egerod 1980) habuk 'loincloth'; Taroko (Pecoraro 1977) xabuk 'belt'. Cf also Paiwan ibuk 'man's upper garment', m-ibuk 'to wear' (Ferrell).

**to bury, tomb**
PAN (root) -buN 'to bury, cover with earth'. Blust (1988) reconstructed a root -bun 'heap, pile, cover with earth; collect, assemble', based on MP evidence, including Dempwolf's ta(m)buN 'heap of earth'. Add Atayal (Egerod) ?bul 'to bury; tomb'; Mayrinax g-um-bul 'to bury'. This Atayal evidence indicates that the root was -buN at PAN level, changing regularly to -bun in MP.

**Sino-Tibeto-Austronesian lexical comparisons**

In italics: words in Swadesh's 100-words list; in bold caps: words in Yakhontov's 33-words list.

<table>
<thead>
<tr>
<th>basic vocabulary</th>
<th>PAN or PECL</th>
<th>Old Chinese</th>
<th>Tibeto-Burman</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. body hair</td>
<td>gumuN</td>
<td>眉 (mu [r] (eyebrow))</td>
<td>mul (Moshang kemul)</td>
</tr>
<tr>
<td>2. <strong>BONE</strong></td>
<td>kukut</td>
<td>骨 (kut)</td>
<td>(s-)nuk</td>
</tr>
<tr>
<td>3. brain</td>
<td>punuq</td>
<td>脑 (nu)</td>
<td>Gyarong tkru</td>
</tr>
<tr>
<td>4. elbow</td>
<td>siku(H₂)</td>
<td>肘 (t-r-ku)</td>
<td>Lushai lu</td>
</tr>
<tr>
<td>5. female breast</td>
<td>numuH₁</td>
<td>乳 (no)</td>
<td>nuw</td>
</tr>
<tr>
<td>6. foot</td>
<td>kakay</td>
<td>扶 (pa)</td>
<td>pa</td>
</tr>
<tr>
<td>7. head</td>
<td>quluH₁</td>
<td>頭 (hlu)</td>
<td>Lushai lu</td>
</tr>
<tr>
<td>8. palm of hand</td>
<td>dapa</td>
<td>扶 (pa)</td>
<td>Tib. m-na</td>
</tr>
<tr>
<td>9. pus</td>
<td>nanq</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. mother</td>
<td>ina(-q)</td>
<td>女 (nra)</td>
<td></td>
</tr>
</tbody>
</table>

5
| 11. **EGG** | qiCelU | ꜕Ca-lo[r]? | twiy < t-l-? |
| 12. **HORN, ANTLER** | (q)uRung | ꜕k-rok | rung=rwang |
| 13. leech | Limatek | ꜕tik | P-Loloish lay/1 'python' |
| 14. snake | SulaR | ꜕m-la[r] | |
| 15. worm | [ Julej | ꜕lit? F? | |
| 16. cloud, cloudy | -qem | ꜕qim | Bur. um' |
| 17. earth | -taq | ꜕tha? | Tib. 'dag pa a mud' |
| 18. moon | qiNaS | ꜕s-la | s-la |
| 19. **SALT** | siRaH | ꜕ra? S! | la! |
| 20. sunlight | siNaN | ꜕lang | Bur. lang 'to be light' |
| 21. water | daNum | ꜕t-hlim? (liquid, juice) | g-liy |
| 22. wind | bali | ꜕s-hwat-s S! | ba |
| 23. cave, hole | b[e]lun | ꜕long S! | Jingpo kinti, lun33 |
| 24. **YEAR** | kawaS | ꜕s-hwat-s S! | b |
| 25. carry | MBAH | ꜕m-pa?-s | ? PS *waa F! |
| 26. chew | paepaq | ꜕pit | Lepcha tultul |
| 27. close, shut | kupit | ꜕to[r,n]1; ꜕s-lo[r] | Jingpo kot |
| 28. cut off, short | [p,b]utul | ꜕m-k(h)ut | mit 'extinguish' (fire) |
| 29. dig | -kut | ꜕m-hlim? | twiy < t-l-, lwiy 'to flow' |
| 30. drown, disappear | Nemes | ꜕t-l[u]r!? (water) | twiy < t-l-, lwiy 'to flow' |
| 31. fall | -luR | ꜕to[r]1 | |
| 32. flow > water, river | qaluR 'to flow' | ꜕hu[r]1? (water, river) | |
| 33. follow | duNuR | ꜕s-lo[r] | |
| 34. grasp, embrace | -kep | ꜕m-kep, ꜕s-kep, ꜕k-r-ep | |
| 35. hold sth in fist | gemgem (in fist) | ꜕g-im (in mouth) | gam 'put into mouth' |
| 36. hunt | qaNup | ꜕Ca-lap | |
| 37. lick | dilaq | ꜕m-le? | m-lyak |
| 38. meet | Cebung | ꜕buŋ S! | ? PS pung 'assemble' |
| 39. open | -kaq | ꜕khe? | Jingpo kha < -k < parted, open |
| 40. put together | pulung | ꜕long | Tib. r-ko, Gyarong ka rkos |
| 41. ruin, damage | r[i]bas | ꜕bet-s | Kuki-Naga d-kew |
| 42. scrape I | kuSkuS | ꜕k-rot | kut |
| 43. scrape II | ku[C][ku][C] | ꜕k-rot | nup |
| 44. sink | -neb | | |
| 45. sleep | -zem | ꜕tshim? | Tib. gzim, Dhimal dzim |
| 46. **speak, say** | kawaS | ꜕m-kw-r-at-s; ꜕wat S! | Tib. s-go |
| 47. think | nemnem | ꜕nim-s | Tib. s-nya-pa 'to think' |
| 48. vomit, spit | utaq | ꜕tha? | (m-)tuk V! |
| 49. wash | basaq | ꜕s(r)u? | Lushai shuk, Luoba cuk |
| 50. wrap around (belt) | -kes | ꜕ket | |
| 51. bent, crooked | -kuk | ꜕N-(h)r(k)ok | kük |
| 52. broad | -bang | ꜕bang | |
| 53. curled, bent | -kul | ꜕N-(h)r[o][r,n] | kuar (PS) |
| 54. dark | -lem | ꜕+lhin?, ꜕hlim? | BL wiy |
| 55. far | ma-dawiN | ꜕wa[r,n]! V! | Bur. kaw: (heavy tone) 'rise up, swell, bulge' |
| 56. high, tall | -kaw | ꜕kaw | |
| 57. **hot** | qa(i)nget | ꜕nget | |
| 58. old, adult, grownup | -dan | ꜕drang? | |
| 59. sharp | Cazem | ꜕tšim | [660a] tšim |
60. thick -tul  
61. THIS di  

[ITALIC9, Canberra, January 8-11, 2002]

**cultural vocabulary**

1. husked rice beras  
2. grain of cereal -may  
3. chicken bï I? Tib. 'di 'this'  
4. cage, enclosure kurung  
5. broom CapuH  
6. stopper, plug sensen  
7. to bury, tomb -buN 'to bury'  
8. loincloth, robe sabuk  
9. to plait, braid -pid  
10. to shoot panaq  

**sound correspondences**

*initial consonants (PAN final syllable initial : Chinese root initial : TB)*

p- : p(h)- : p-  
leech, earth, vomit, thick, short  

k- : k(h)- : k-  
hold in fist or mouth  

m- : (h)m- : m-  
body hair, drown  

n- : n- : n-  
brain, breast, pus, mother, think, shoot, sink  

n- : (h)n- : n-  
hot,  

l- : (h)l- : l-  
head, snake, head, flow, lick, put together, fall, wind, cave, worm  

R- : r- : r-  
horn, salt, husked rice  

w- : (h)w- : w- (Tib. g-)  
year, far, say  

s- : s- : ?  
wash, stopper  

z- : ts- : ?  
sharp, sleep, wink  

*final consonants (PAN final consonant: Chinese root final consonant : TB)*

-0 : -0 : -0  
palm of hand, chicken, carry, this  

-k : -k : -k  
leech, crooked, loincloth  

-t : -t : -t  
bone, hot, dig, close  

-p : -p : -p  
hunt, grasp  

-ng : -ng : -ng  
put together, broad, cage, meet, sunlight, cave, old  

-m : -m : -m/p  
horn, stopper  

-m : -m : -m/p  
warm, think, hold in fist or mouth, dark, cloud  

-H  : -0 : -0  
head, female breast, elbow, salt, broom  

-q : -q : -k  
brain, pus, earth, lick, vomit, chew, shoot, wash, open  

-l- - [r]- -r-  
curled, thick,  

-R : -r : -y  
dog, snake, egg, flame, flow, fall, follow  

-S : -t : -0  
say, year, scrape I, moon  

-s : -t : -t (a.) -t (else) husked rice, drown, wrap around, ruin  

-N : -n : -y- -l  
body hair, far, tomb
### STAN PAN : Chinese examples

<table>
<thead>
<tr>
<th>STAN</th>
<th>PAN : Chinese examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>u (before labials)</td>
<td>-u- : -ï- water</td>
</tr>
<tr>
<td>u (elsewhere)</td>
<td>-u- : -u- head, brain, elbow, bone, body hair, dog, flow, thick, dig, meet, tomb</td>
</tr>
<tr>
<td>o (before labials)</td>
<td>-o- : -a- hunt</td>
</tr>
<tr>
<td>o (elsewhere)</td>
<td>-o- : -o- breast, egg, horn, fall, put together, curl, crooked, cut off, cage, cave</td>
</tr>
<tr>
<td>a (before y)</td>
<td>-a- : -ï- grain</td>
</tr>
<tr>
<td>a (elsewhere)</td>
<td>-a- : -a- palm, mother, snake, year, salt, earth, vomit, shoot, speak, broad</td>
</tr>
<tr>
<td>æ</td>
<td>-æ- : -e- chicken, lick, ruin, open</td>
</tr>
<tr>
<td>e (after grave cons.)</td>
<td>-e- : -e- grasp, wrap around, drown, hot</td>
</tr>
<tr>
<td>e (elsewhere)</td>
<td>-e- : -i- think, leech, worm, sleep</td>
</tr>
<tr>
<td>i (open syll.)</td>
<td>-i- : -i- this</td>
</tr>
<tr>
<td>i (closed syll.)</td>
<td>-i- : -i- plait, close</td>
</tr>
<tr>
<td>a</td>
<td>-æ- : -æ- dark, sink, hold in fist, stopper, sharp</td>
</tr>
</tbody>
</table>

### syllable types

This applies to comparisons with polysyllabic AN members (excluding reduplications)

<table>
<thead>
<tr>
<th>PAN penultimate syllable initial</th>
<th>Chinese syllable type</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop (except q), or zero</td>
<td>(non-division 3)</td>
<td>bone, brain, horn, close, put together, spit</td>
</tr>
<tr>
<td>other initials (including q)</td>
<td>(division 3)</td>
<td>elbow, head, palm, leech, snake, water, drowned</td>
</tr>
</tbody>
</table>

### tones (PAN word ending : Chinese word ending/tone : TB word ending)

<table>
<thead>
<tr>
<th>PSTAN</th>
<th>AN</th>
<th>Chinese</th>
<th>TB</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0</td>
<td>-0</td>
<td>tone A</td>
<td>-0</td>
<td>palm of hand, chicken, carry, this</td>
</tr>
<tr>
<td>-m, -ng, -l, -N, -w, -R etc.</td>
<td>id.</td>
<td>tone A</td>
<td>id.</td>
<td>eyebrow, meet, put together, think, broad, curled, high, thick, needle, cage, follow, snake</td>
</tr>
<tr>
<td>-q</td>
<td>-q</td>
<td>? = tone B</td>
<td>-k</td>
<td>brain, pus, earth, lick, vomit, wash, open</td>
</tr>
<tr>
<td>-h, -?</td>
<td>-H₁, -H₂</td>
<td>? = tone B</td>
<td>-0</td>
<td>elbow, female breast, head, broom, salt</td>
</tr>
<tr>
<td>-mh, -nh, -lh, -Nh, -wh, -jh, -rh, -Rh, -m&quot;, -g&quot;, -l&quot;, -N&quot;, -w&quot;, -j&quot;, -r&quot;, -R?</td>
<td>-h or ? is lost</td>
<td>? = tone B</td>
<td>-h or ? is lost (&gt;? TB tonogenesis)</td>
<td>water, cut off, far, egg, flow</td>
</tr>
<tr>
<td>voiced stops: -b, -d, -j, -g</td>
<td>id.</td>
<td>nasals+ B</td>
<td>nasals, -r</td>
<td>worm, plait</td>
</tr>
<tr>
<td>-p, -t, -C, -k</td>
<td>id.</td>
<td>tone D (-p, -t, -t, -k)</td>
<td>-p, -t, -t, -k</td>
<td>bone, leech, close, dig, bend, etc.</td>
</tr>
<tr>
<td>-ɛ (?)</td>
<td>-s</td>
<td>tone D (+t)</td>
<td>-s (+/a_ ) ~ -t</td>
<td>husked rice, drown, wrap around, ruin</td>
</tr>
<tr>
<td>-s</td>
<td>-S</td>
<td>tone D (+t)</td>
<td>-0</td>
<td>year, say, scrape I, moon</td>
</tr>
</tbody>
</table>

Chinese tone C (去聲) is secondary, from root+s suffix.

### morphology

Better preserved in TB than in Chinese.

<table>
<thead>
<tr>
<th>sV-Verb valency increaser</th>
<th>AN</th>
<th>Chinese</th>
<th>TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>pV-Verb causative</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
**sV-Verb > valency increaser**
A prefix PAN Si- : OC s- : TB s- allows a verb stem to take into its argument structure a NP standing for a 'third' participant with situational roles such as causer, beneficiary, instrument, etc. The Austronesian Si-V construction is known as 'Instrument focus' but its semantics are considerably more complex. Huang (1991) characterizes the Si- construction in Atayal as 'circumstantial voice'. She notes (1991:45) that one characteristic of circumstantial voice is 'increased transitivity'. I cite here examples with a transitive/causative character, because the semantic difference between prefixed and non-prefixed forms can be apprehended directly through simple lexical glosses.

AN Si-Verb > 'transitive/causative'
- Atayal ɲumu 'to be afraid' : s-ɲumu 'to frighten'
- Paiwan k/m/avuL 'beg' : si-kavuL 'cause someone to beg'
- Bunun ma-bahliv 'to buy' : is-bahliv 'to sell'

Old Chinese s-Verb 'transitivizing'
- *m-lun-s > zywinH 'to be pliant, obedient' : *s-lun > zwin 'to tame'
- *m-lik 'to eat' : s-lik-s 'to feed'

TB s-Verb 'transitivizing'
- Tibetan 'bar 'to burn, catch fire, be ignited' : s-bar-pa 'to light, to kindle, to inflame'
- Gyarong rong 'to see' : s-rong 'to show'
- Boro gi 'be afraid of, fear' : si-gi 'frighten'
- Proto-Loloish¹ (C)-no2 'to awake' : s-ño2 'to awaken' (tr.)

**pV-Verb 'causative'**
A verbal prefix **pV- making causatives out of noncausatives is widely attested in TB and in AN. It has not been observed in Chinese.

AN: pa-Verb 'causative'
- Amis kaǝn 'eat' : pa-kaǝn 'feed'
- Puyuma kan 'to eat' : pa-kan 'to feed'
- Atayal biru 'to write' : p-biru 'to let write'
- Thao kan 'eat' : pa-kan 'cause to eat, feed'

TB: *pV-Verb 'causative'
- Jingpo: ran 'be separated' : pǝ-ran 'to separate'
- Boro: ɡonŋ 'to bend' (intr.) : bo-khoŋ 'to make bent'

¹ Bradley 1979
**m- intransitive**

AN m-/um- + transitive verb > intransitive verb

The Austronesian 'Actor Focus' (AF) marker is a nasal affix m- (prefix) or -um- (infix) depending on root shape; in Starosta’s ergative interpretation of AN grammar (Starosta 1991, 1994), all verbs in Actor Focus are intransitive, and the function of m-/um- is to derive intransitive verbs from transitive verbs. Constructions in which AF verbs (marked with m-/-um-) take both a patient and an agent are antipassive (pseudo-transitive): such sentences translate naturally into transitive sentences in accusative languages. For that reason, French or English translations fail to capture the intransitive character of verbs marked with m-/-um-. This intransitive character of AF verbs appears from the fact that their situational patient, if there is one, is encoded in an oblique case, genitive or locative.

TB m-Verb > intransitive

In Gyarong and Xide Yi, this prefix is a nasal which takes the place of articulation of the root initial. Examples (Lin Xiangrong 1993: 193 for Gyarong; Dai Qingxia 1998:86 for Xide Yi):

Gyarong

- ka-tʃop ‘to set fire to’ : kə-ndʒop ‘to catch fire’
- kɛ-p’ek ‘to split open’ : kə-mbjk ‘to be rent’
- kɛ-tʃ’op ‘to break’ : kə-mdʒop ‘broken’
- kɛ-ktʃk ‘to wipe off’ : kə-ŋglk ‘to fall’

Xide Yi

- ti₃₅ ‘to suspend’ (tr.) : ndi₃₅ ‘to be suspended’ (intr.)
- tu₃₅ ‘to burn’ (tr.) : ndu₃₅ ‘to burn’ (intr.)
- po₃₁ ‘to roll’ (tr.) : mbọ₃₁ ‘to roll’ (intr.)
- tʃi₃₅ ‘to pull off’ (tr.) : ndʒi₃₅ ‘to fall off’
- to₃₃ ‘to give to drink’ : ndọ₃₁ ‘to drink’

In these examples, the N- prefix has voiced a following voiceless stop. In many TB languages (such as Tibetan, Kiranti, Bahing, Vayu and Bodo-Garo), the nasal element has been lost and only voicing of the root initial remains. Preceding initials other than voiceless stops, however, the intransitive prefix maintains itself as m- (WT), ma- (Jingpo), mV- (Boro), and prefixed forms are easy to detect outside of minimal pairs. Examples (Wolfenden 1929:30 for WT and Jingpo, Bhattacharya 1977: 184, 328-330 for Boro):

WT

- m-gu-ba ‘to rejoice, be glad’
- m-nar-ba ‘to suffer, be tormented’
- m-nam-ba ‘to smell (intr.), stink’
- m-nal-ba ‘to sleep’
- m-nab-ba ‘to dress oneself’

---

2 Gyarong ka-, kɛ- and kə- are verb prefixes for controllable (ka-, kɛ-) and non-controllable (kə-) actions.
Kachin  
ma-nam 'to smell' (intr.)
ma-den 'to grow, to expand'
ma-ni 'to laugh'

Boro 
mō₂-nam₁ 'to spread smell'
mō₂-zom₁ 'close the fist'
mī₂.ni, 'to laugh'

The original form of the intransitive prefix was therefore m- or ma-, with a variant N- before voiceless stops.

Chinese N-transitive verb > intransitive verb
Like Tibetan and other TB languages, Middle Chinese (mid-first millennium CE) has some contrasting pairs of transitive verbs with voiceless stop initials and intransitive verbs with voiced stop initials:

MC  
khjowk 'to bend' (tr.): gjowk < *Nkh- 'bent, curved; bend the body'
trjuvwngH 'to hit, place in the middle': drjuvwngH < Ntr- 'to be in the middle'

That the voiced embers of these pairs eralier had prenasalized initials is shown by Miao-Yao. Proto-Miao-Yao (Wang and Mao 1995) borrowed some of these forms from Late Old Chinese (before MC), with prenasalization and no voicing: *ŋkhök 'curved', *ŋŋŋA 'middle'. Evidently Chinese had the same intransitive prefix N- as Gyarong and Xide Yi.

**pV-Noun > Verb

AN: pa-Noun > Verb
Atayal: suzit 'a whistle': p-suzit 'to whistle, blow an instrument'
Taroko: bais 'husband, wife, mate': p-bais 'to form a pair'
Amis sual 'words, speech': pa-sual 'to tell, explain'
Paiwan tutu 'milk': pa-tutu 'to feed milk'
Puyuma tigami 'letter': pa-tigami 'write a letter'
Thao apuy 'fire': p-apuy 'start a fire'

TB: pV-Noun > Verb
Burm. rum 'gathering place': p-rum 'assemble, crowd together'
Tib. źo 'milk': b-ző-ba 'to milk'
Mikir chor 'pair': pa-chor 'to marry' (trans.)

Chinese p-Noun > Verb
Perhaps, from Wuxiang 武鄉 a Jin dialect of Shanxi:
  pa-sur 'to card, comb'

Where sur is etymology the same word as 梳(shu) MC srjo 'a comb'.

**Verb-en > 'thing Verb-ed/thing to be Verb-ed'

PAN Verb-en > 'thing Verb-ed/thing to be Verb-ed'

Atayal niq 'eat': niq-un 'eaten thing'
Paiwan kan 'to at': kan-en 'food'
Amis ajik 'to sweep': aajik-en 'place to sweep'
TB Verb-n > 'thing Verb-ed'

Tib. za-ba 'to eat': zan 'food, fodder, pap, porridge'.
skyi-ba 'to borrow': skyin-pa 'a thing borrowed, a loan'
rduz-ba 'to misrepresent, to delude, to falsify': rduzn-pa 'falsehood, fiction, lie'

**Sa-Verb > instrument of action

PAn Sa-Verb > instrument of action

Rukai to?to? 'to peck': sa-to?to? 'beak'
Amis senat 'to harrow': sa-senat 'a harrow'
Pazeh m-bazu 'wash clothes': sa-bazu 'soap, soapberry'

TB sa-Verb > instrument of action

Gyarong ka-top 'to hit' (ka- verb prefix): sa-top 'instrument for hitting'
Tib. 'bud' to blow': sbud-pa 'bells'
Kachin nan 'to light, illumine, as with a torch': sha-nan 'a torch'

Old Chinese s-Verb > instrument of action

Old Chinese *t-iin 'to steam': s-t-iin-s 'earthen-ware pot for steaming rice'
*lat(-s) 'to pull': *s-hlat 'leading-string'
*s-jet 'bite, gnaw': *s-jet 'wooden wedge between the teeth of a corpse'

**-ar- distributed action; distributed object.

This infix was inserted between the root initial and first vowel of a stem. Attached to verbs of action it indicated that the action was distributed in time (occurring over several discrete occasions), or in space (involving several agents/patients/locations); attached to stative verbs it was simply intensive; attached to nouns it indicated a referent distributed in space, i.e. having double or multiple structure.

AN -ar- distributed action/object; intensive

Paiwan k-ar-akim 'to search everywhere' (kim 'search')
k-ar-akap-an 'sole of foot'
Puyuma D-ar-ukap 'palm of hand'
Bunun d-al-apa 'sole of foot' (PAN *dapa 'palm of hand')
Amis p-ar-okpok 'to gallop'
t-ar-odo' 'fingers, toes'
k-ar-ot 'harrow'
Pazeh b-ar-anhan 'urn'
d-ar-unudun 'gong, drum'
Rukai Budai ma-D-ar-argDang 'hot' (weather)
m-ar-eDang 'old' (compare Puyuma maidang 'old')
Tagalog d-al-akdak 'sowing of rice seeds or seedlings for transplanting' (dakdak 'driving in of sharp end of stakes into soil')
k-al-aykay 'rake'

3 Gyarong has another prefix sa- expressing causative meaning in verbs (see above).
Malay  ketap 'to bite teeth' : k-er-etap 'to bite teeth repeatedly'

TB -r- distributed action/object

Burm. pok ‘a drop (of liquid)’ : prok ‘speckled, spotted’
    phu 1 ‘to protuberate’ : pru3 ‘to protuberate, as the eyes’
    pwak ‘to boil up and break, as boiling liquid’ : prwak ‘id’
    khwe2 ‘curve, coil’ : khrwe2 ‘surround, attend’

    the body; to appear on the body, of pimples’

Chinese -r- distributed action/object; intensive.

Fuzhou  niaʔi₁₃ ‘to blink’ : niaʔi₁₁-liA/ɪ₃ ‘to blink repeatedly’
    tsʰi₃₅₅ ‘to turn, rotate, as a machine’ : tsʰi₃₃-li₃₅₅ ‘to rotate in alternating
        directions, as an electric fan’
    pau₂₁₃ ‘to protrude’ : pa₁₁-liu₂₁₃ ‘to protrude much’
    tsiŋ₅₅ ‘to stare’ : tsi₃₃-liŋ₅₅ ‘to stare fixedly’
    tʰaʔ₅₅ ‘pile or stack of (paper)’ : tʰa₁₁-laʔ₃₃ ‘a pile of paper, irregularly
        stacked’

references

    language. In: Elizabeth Zeitoun and Paul Jen-kui Li (eds.) Selected papers from the Eighth
    University Dept. of Publications.
Blust, R. (1973) Additions to 'Proto-Austronesian addenda' and 'proto-Oceanic Addenda with
    cognates in non-Oceanic Austronesian languages', N°2. Working Papers in Linguistics
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. (1989) Austronesian Etymologies IV. Oceanic Linguistics Vol. XXVIII number 2,
    111-80.
    and Malmö: Curzon.
d'Extrême-Orient.
Dai Qingxia 1998
Dempwolff 1929
Dempwolff, O. (1938) Vergleichende Lautelehre des Austronesischen Wortschatzes. Band 3:
    Reimer.


