The Reconstruction of Proto-South-Sulawesi

R.F. Mills

Archipel, Année 1975, Volume 10, Numéro 1
p. 205 - 224

III. 1) Roger F. Mills (University of Michigan) compares here the various languages (or groups of languages) of the South Celebes : Makassar, Bugis, the languages of Massénrémpulu', Mandar, the languages of the Pitu Ulunna Salu, and the "Sa'dan", for whom he clearly proves the close affiliation, only slightly clouded by the sporadic influence of an eventual substrate which may have some connection with the languages of central Celebes. The author then fixes the discernable phonetic correspondances by comparison of related words in these various languages, and by way of a conclusion proposes a highly-probable reconstruction of a part of the vocabulary of the original common language which he designates by the name of "Proto-South-Sulawesi".

THE RECONSTRUCTION OF PROTO SOUTH SULAWESI

by R. F. MILLS

1. Using data collected in the field and from published sources, I have reconstructed the proto-language, Proto South Sulawesi (PSS), from which the present-day languages of the area can be shown to have descended. The languages — or better, dialect groups — include Buginese, Makassarese, Mandar, Sa'dan Toraja, Pitu Ulunna Salo, Seko and Massenrempulu'. (1) These seven constitute a minimum, for we do not know the exact number of dialects within each group, and it is quite possible that "dialects" may yet be found which ought to be considered separate languages. In an area like South Sulawesi, where all the languages have been in prolonged and intimate contact with each other, the decision as to what is a "dialect" and what is a "language" is often difficult.

2. PSS is reconstructed with the following sound system:

Voiceless stops and affricate:	p t c k
Voiced stops and affricate:	b d j g
Nasals:	m n ñ ng

(1) Hereafter abbreviated: Bug., Mak., Mdr., Sad., PUS, Mass.; SekoL = Seko, Lemo dialect, SekoP = Padan dial., DuriC = Duri Cakke dial., DuriK = Kalosi dial. In the Appendix, exs. are cited also from the language of Mamuju (Mmj.) which seems to be a mixture of PUS and Mdr. Other abbreviations: AN, Austronesian; PAN, Proto Austronesian; IN, Indonesian; SSul, South Sulawesi; Ml., Malay; BI, Bahasa Indonesia; OJ, Old Javanese; Jav., modern Javanese; BWB, Matthes 1874; MWB, Matthes 1859; SWB, van der Veen 1940; A&K, Adriani and Kruijt 1914.

Fricatives :	s, z
Continuants :	w r, l y R
Vowels :	i e ĕ a u o

(The symbol *R represents the voiced velar fricative usually symbolized with Greek gamma; regarding *z and *R see §§4d and 5c).

We also tentatively reconstruct glottal stop (symbolized *q in PSS and in the modern languages) and *h, both of very limited distribution. In intervocalic position only, the following nasal clusters are reconstructed : *mp mb nt nd ñc ñj ngk ngg ns nz.

3. Several of the proto-phonemes are retained with little or no change. They are listed here, without further discussion ; examples of each will be found in the Appendix.

a. Initial and intervocalic PSS *p t m n ng r l s > /p t m n ng r l s/ in all the languages. (2)

b. Initial and intervocalic PSS *g > Seko /k/, all others /g/.

c. PSS *y (attested only intervocalically) > Mak. Seko /y/, Bug. Duri PUS /j/; Mdr. /y/ or /j/, and Sad. /y/, /j/ or ∅, depending on dialect. The Sad. dialect recorded in van der Veen's (1940) dictionary shows *y > ∅ regularly, but van der Veen 1929 discusses other dialects with /y/ or /j/, as well as PUS dialects with /y/.

d. PSS *a > /a/ in all languages (but see §5a regarding *a before certain final consonants).

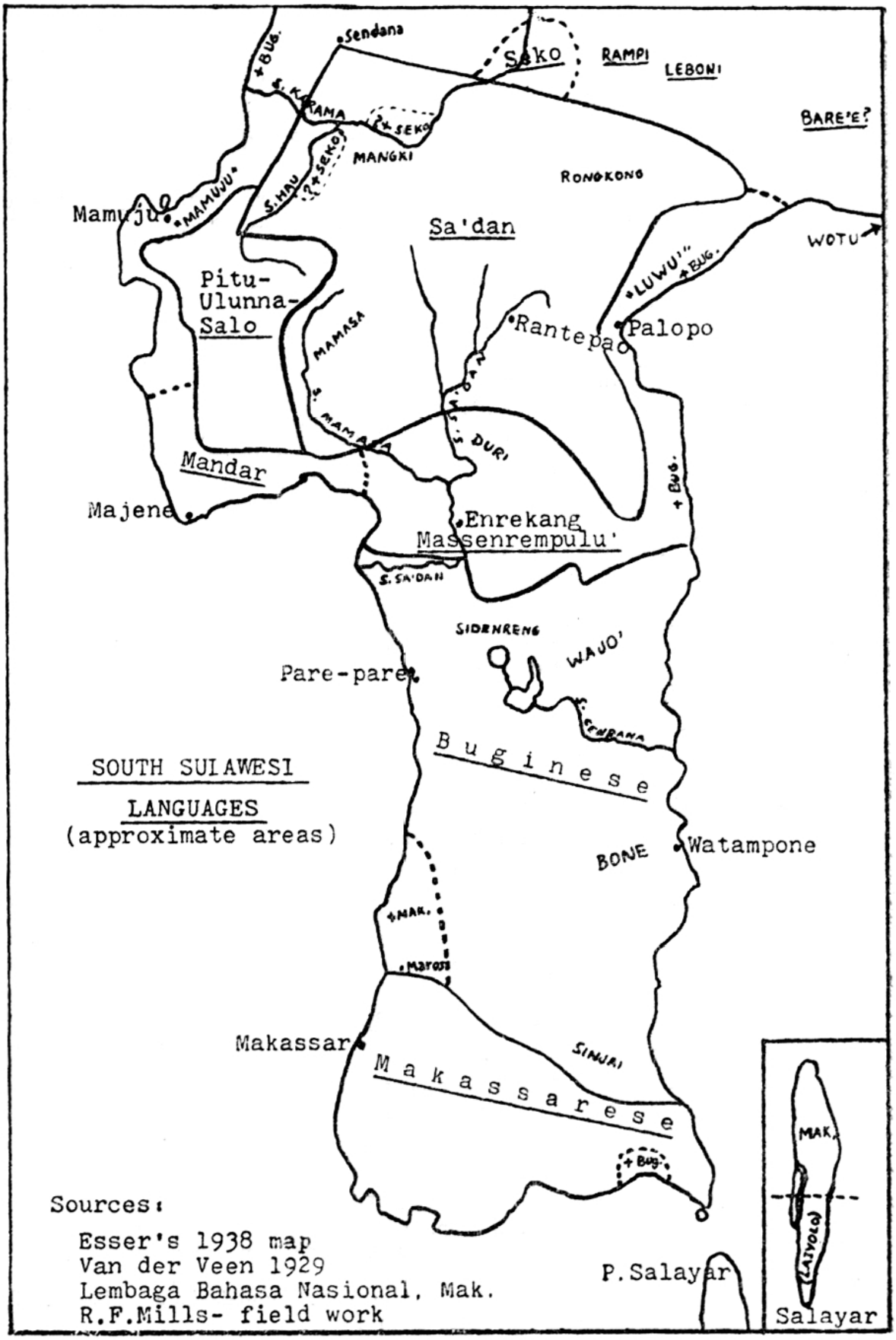
e. PSS *i > /i/ in all; occasionally /e/, especially in closed syllables.

f. PSS *u > /u/ in all; occasionally /o/, especially in closed syllables. (3) In the Mdr/PUS area there is an unusual secondary development whereby /u o/ > /i e/ before a final consonant; see §5b.

---

(2) In Bug., PSS initial *p < ∅ occasionally, e.g. ile, mile 'choose' < *pile. This may be an *um*-form, thus *p-um-ile > *mile (note the same rule affecting initial labials with *-um-* in OJ), then reinterpreted as prefix *m-* plus base *ile*. Some exs. like uro 'quail' (BI puruh) could be loans; yet others, like ase 'padi' must be native, and the loss of *p is inexplicable. Regarding *r, a frequent reflex in PUS is /h/; this change affects r < PSS *d as well as r < *r.

(3) In many of the modern languages, the pronunciation of /i/ and /u/ varies between a high-close [i], [u] and mid-close [e], [o]. The three dictionaries all list frequent doublets, e.g. Bug. ingēq, engēq 'remember'.


Sendana
Seko
RAMPI
LEBONI
BARE'E?
MANGKI
RONGKONG
Mamuju
MAMUJU
Sa'dan
WOTU
LUWU'
Pitu-
Ulunna-
Salo
MAMASA
Rantepao
Palopo
Mandar
DURI
Majene
Enrekang
Massenrempulu'
SIDENRENG
WAJO'
Pare-pare
Buginese
BONE
Watampone
Maros
Makassar
SINJAI
Makassarese
MAK.
(Ayoreo)
P. Salayar
Salayar
SOUTH SULAWESI
LANGUAGES
(approximate areas)
Sources:
Esser's 1938 map
Van der Veen 1929
Lembaga Bahasa Nasional, Mak.
R.F.Mills- field work

g. PSS *e and *o > /e/, /o/ in all. Reconstructible with certainty only as finals, where *e < PAN *-ay or *-ih, PSS *o < PAN *-aw or *-uh. (4) In non-final syllables, PSS *e and *o can usually be traced to PAN *i and *u, or else represent borrowings — forms with /e/ probably from the Moluccan area where that vowel is the regular reflex of PAN *ĕ; forms with /o/ from Central Sulawesi, where /o/ also reflects PAN *ĕ.

h. PSS *mp nt ngk > Bug. Mdr. SekoL /pp/ /tt/ /kk/, Mak. Mass. SekoP /mp/ /nt//ngk/. Sad. shows both developments; at an earlier stage, this probably represented a dialect split, but nowadays there is considerable free variation — evidence, perhaps, of a sound change in progress. Since in Bug. morphophonemics, nasal assimilate to a following voiceless stop (e.g. asĕng 'name', asĕkku 'my name'), instances of Bug. /mp nt ngk/ *corresponding to the same sounds in other languages* must be taken as signs of outside influence.

i. PSS *mb nd ngg > Bug. /mp /nr/ /ngk/,(5) SekoP/mm/ /nn/ /ngk/, Mak. /mb/ /nr/ /ngg/, Mdr. Sad. Mass. PUS SekoL /mb/ /nd/ /ngg/. (Examples for *ngg are lacking in SekoL, but we should expect /ngk/ there). Again, morphophonemic evidence in Bug. allows us to mark Bug. /mb nd ngg/ as signs of outside influence; likewise /nd/ in Mak.

j. PSS *ns > /ss/ in all, though van der Veen 1929 cites examples of /-ns-/ in some PUS dialects.

4. The remaining consonants and vowels, and the final consonants of PSS, require special comment.

a. *PSS *é (reflecting PAN *ĕ).* The reflexes of PSS *e are not in themselves remarkable: /ĕ/ in Bug., /a/ in all the others, with instances in every language of irregular /i e u o/ due to assimilation to neighboring vowels. It is the change in the consonants following *ĕ which is of interest, and which represents one of the distinctive features of SSul languages — that is, the development of geminate, or long, consonants. Since gemination of consonants following PAN *ĕ is found in a number of languages — Madurese, Sangirese, scattered languages of the Philippines, perhaps in Old Javanese (see Ras

(4) I use Dempwolff's symbol *h rather than Dyen's *q to avoid any possible confusion with my *q for glottal stop. Likewise I use the pepet (ĕ) instead of Dyen's *e to avoid confusion with my *e*, which represents a low or mid front vowel.

(5) A&K (p. 156) claim the same development in Mdr. Campalagiang dial., but all the exs. they cite could be borrowings from Bug.

1968-70) — it is not clear whether this is an example of parallel development, or a feature retained from some earlier stage, Proto-Indonesian or even Proto Austronesian. A possible explanation for the PSS situation is that the morpheme structure rule (inherited from PAN) forbidding *ĕ in a *final* open syllable was generalized to forbid *ĕ in *any* open syllable. It is possible too, I think, to connect gemination with the development of a fixed penultimate stress in PSS. Thus while the vowels *i e a u o, when stressed, developed long allophones, some peculiarity in the nature of *ĕ prevented this, and the length feature came to be associated with the following consonant instead. For example, PAN *talu 'defeated' > PSS *talo phonetically (presumed) [tá:lo], whereas PAN *tĕlu 'three' > pre-PSS *tĕlu, phonetically not [tĕ:lu] but [tĕl:u]. With reanalysis of the syllable structure from CV.CV to CVC.CV, this resulted in PSS *tĕllu. (6) And although gemination in PSS was completely predictable and so non-phonemic, with the development of geminates from other sources — from the voiceless nasal clusters, or from consonant clusters at morpheme boundaries — and with the change of *ĕ to other vowels, the present-day languages apparently contrast single and geminate consonants — e.g. Sad. bale 'fish' vs. balle 'tell a lie'. Synchronically, however, it is probably more economical to analyze them as clusters.

In Seko, the geminates appear to have been simplified back to single consonants (the data are unclear on this point); but they are retained in all the other languages. Further, even where the reflex of *ĕ is irregular, the following geminate is usually retained. PSS geminate voiceless stops, nasals and continuants are reflected as: single consonants in Seko, true geminates in all the others.(7) PSS geminate voiced stops are reflected as: single stops in Seko (but no examples for *gg), preglottalized /qb/, /qd/ (further > /qr/ in Mak.) and /qg/ (very rare) in all the other languages. Since *w and *y did not occur after *ĕ (reflecting another PAN morpheme structure rule), it follows that they are never found geminated. The apparent

---

(6) Apparently a CVC syllable had the same value, for purposes of the stress, as a C$\bar{V}$, and a prenasalized consonant after *ĕ satisfied the same condition. The phenomenon is not unheard of elsewhere — consider Latin.

(7) I.e., phonetically, the articulators are in contact longer for a geminate than for a simple stop, nasal or continuant. /rr/ is strongly trilled, /r/ is a single tap. Note that Matthes' system of transliteration in MWB and BWB distinguished -qC- from -CC-; but modern Bug. knows no such contrast, and t is marginal in Mak. SWB also makes the distinction, as do some informants; but in view of the many doublets, I suspect dialect mixture. Just one example: SWB pappa' or pa'pa' 'flat'.

/yy/ in Bug. iyya 'he, she' can be explained as a prefix /i/, a fossilized personal article (as in "I La Galigo"), plus /ia/, parallel with /iaq/ 'I' (< i + aq) and /iko/ 'you' (<i + ko).

b. *Reflexes of PSS *k.* Regularly > /k/ in Mak. Sad. Mass. and PUS. In Bug. we find both /k/ and ∅; in Mdr. and Seko /k/ /q/ or ∅. Since the items with ∅ most often belong to the so-called basic vocabulary, loss is probably the "native" development, but inter-language and inter-dialect borrowing have reintroduced a phonemic /k/ into these languages.

Thus, what would otherwise have been a completely regular change *k > ∅ has been obscured and, to the linguist's confusion, a multiple reflex is found. In Bug., the original /k/ sometimes reappears in derived forms, e.g. ita, mita 'to see' (PSS *kita) but pakita 'vision, sight'.

c. *The PSS palatals *c j ñ ñc ñj.* These have been retained as palatals in Bug. Mak. and Mdr. (with *ñc regularly > Bug. Mdr. /cc/, *ñj > Bug. /ñc/); elsewhere they have merged with *s d n ns nd respectively. But even in Bug. Mak. and Mdr., there are frequent doublets and irregular cognates showing these same mergers. Since Bug. Mak. and Mdr. are the three languages most likely to have had extensive contact with Ml. or Jav. (and indeed with each other), one could argue that their palatals are due to outside influence, and that PSS in fact had lost them.

d. *Voiced stops *b d, continuants *w z.* Multiple reflexes of *b d occur, but aside from presumed borrowing, morphophonemics also provide a possible explanation. The reflexes, in summary: PSS initial *b > /b/ most often in all the languages. ([8]) In Bug., initial b/w alternate morphophonemically (see below); further, Sad. Mass. PUS and Seko also have the reflex of a secondary /w/ (> Sad. ∅, PUS Seko h-) which must have been distinct from PSS *w-. PSS intervocalic *b > /w/ regularly, and this /w/ further > Sad. ∅, PUS, Seko /h/. /b/ occasionally is found. The development in Sad. must have been *b > w > h > ∅

Initial and intervocalic PSS *d > either /d/ or /r/. Only Mak. shows /r/ with fair regularity; in the other languages, the reflex is quite mixed. Several PAN etyma reconstructed with *d or *ḍ show /r/ in all SSul languages; these are reconstructed to PSS *(dr)

([8]) To be perfectly honest, the correct Seko reflex is unclear. However, b < *b outnumbers h < w < *b; further, *h* is often (though not exclusively) found in words of unknown cognacy, or for which a relatable form with *w* occurs in one of the Toraja languages, Bare'e, Rampi, Leboni etc.

— i.e. *d or *r — since it it not possible to determine when the shift took place.

PSS initial *w > Bug. Mdr. Sad. Mass. /u/ (phonetically [uw ...]) or /w/; the sequence uw- > uh- in PUS and Seko (note the different development of w < *b). The only two examples of *w- in Mak. are contradictory: bani 'bee', (Bug. uwani), and akaq 'root' (Mdr. uakeq, PUS uhakaq). The total number of examples is low in any case due to the rarity of PSS (and PAN) *w-. PSS intervocalic *w > Bug. Mak. Mdr. Mass /w], PUS /w, h, b, v/ and Sad. ∅ or /w/ depending on dialect; Seko /h/. Interestingly, there are dialects of Bug. (Sinjai) and Mak. (Konjo, Salayar) where w > h also.

In Bug., there is a synchronic rule to the effect that initial /w/ and /r/ change to /b/ and /d/ following the glottal stop of the prefix /maq-/. From an historic point of view, however, the rule must be stated just the reverse: /b/ and /d/ change to /w/ and /r/ in absolute initial position, but are retained after /q/. Thus we have maqbere 'give' < were, or maqdangĕng 'be friends, partners in business' < rangĕng 'friend, associate'. This is assumed to be the regular development; there are exceptions, of course, and many doublets — e.g. dappiq 'nearby' > either sidappiq or sirappiq 'next to each other, side by side'. The base dappiq has probably arisen by analogy from some such form as "maqdappiq" 'be near'. Yet it is the synchronic rule which seems to be winning out, since nowadays some forms which reflect historic *r (< PAN *r or *R) also undergo this r/d alternation — e.g. maqdaiq 'carry on a raft' < raiq (note PAN *Rakit), or even maqdoti 'make bread' based on M. roti, itself a loan word from Indic. The Cendana dialect of Mandar described by Adriani and Kruijt shows just the opposite conditioning: /b/ and /d/ occur only initialy, but change to/w/ and /r/ when a vowel comes to precede them, e.g. boyang 'house' > diwoyang 'at home'. Doublets in Sad. suggest that it once followed the Bug. pattern, but no longer does. Probably the other languages too once had similar rules governing b/w and d/r alternations; when the rules ceased to operate, analogy and borrowing obscured the regular alternations, and created new phonemic contrasts.

PSS *z is reconstructed in intervocalic position only. After *ĕ, the reflex seems to be: Bug. -ss-, Mdr. Sad. Mass PUS -qd-, but there are only two clear examples. Both show irregularities in Mak. and Seko, so that a clear reflex in those languages in not available. After other vowels, *z is clearly reflected as Bug. -s-, all others -r-. PSS *nz, also attested only in two uncertain examples, apparently

merged with *nd; if the merger took place in pre-PSS times, then *nz need not be reconstructed for PSS. PSS *z reflects Dempwolff's PAN *g, which in turn represented part of the old RLD Law, whose reflexes were Batak g (-k), Ml. d (-t), Jav. r, Tag. l (-d), Fijian c, (9) and Polynesian h. Dempwolff proposed that it was a palato-velar stop — a point which I will not argue here. On the basis of the SSul reflexes, PSS *z would seem to have been a dental which combined the features of stop and continuant; thus, perhaps, an affricate, phonetically [dz] and systematically the voiced counterpart of *s.

5. *Final consonants.* I posit the following consonants in final position in PSS: *p t k m n ng r l s R. There can be no argument over *k m n ng r l s, since each is retained in one or another of the modern languages; likewise *R which is clearly reflected in Duri and Bug. (as will be shown). Yet all the SSul languages have sharply reduced the number of permitted finals: Bug. with only /q/ and /ng/ has the most extreme development. Sad. PUS and Mass, dialects are slightly less extreme, with /q k n ng/ (some PUS dialects are reported with /m/ also; Seko lacks /n/; while Mass.-Duri adds /h/). Mdr. has /q n ng r l s/, although there can be doubt about the status of /l/, and /n/ is apparently in process of merging with /ng/ (10). Mak. has /q ng r l s/, the continuants having been preserved by the addition of an echo-vowel plus /q/ sequence (11).

The morphophonemic alternations of /-q/, however, indicate the earlier presence of other consonants (unfortunately, data on this point are lacking for PUS and Seko):

Bug. -q > /r s k/ before -ĕng and -i.

Sad. -q > /r s/ before -an.

Duri -q > /t/ (probably /r/ too) before -an.

Mak. -q > /k/ before -ang and -i.

---

(9) Fijian orthographic *c* represents a voiced interdental fricative, like *th* in English *the*.

(10) In the dial. ("Balanipa") for which I collected data. A&K report /q m n ng/ for Mdr.-Campalagiang, but gave no exs. for the crucial /m/. Mdr.-Majene (Pelenkahu 1967) has /q ng r l s/.

(11) Synchronically, one might say that Mak. has only final /q ng/. But the continuants are clearly present in underlying forms, for the suffixes /-ang/ and /-i/ are added not to the base plus echo-vowel form, but to the base only. E.g. lémbaraq 'carry on a shoulder-pole' (note stress on the underlying penult) > lembárang 'a shoulder-pole'.

Mdr. -q > Ø before -an/-ang, but both /q/ and Ø are found in examples with -i ([12]).

a. The "inserted" /t/ of Duri is strong evidence for the reconstruction of PSS final *t, but two other bits of evidence also support it. First is the quite regular correspondence of Bug. /-ĕq/ to a sequence of /a/ plus dental (t d n r l s) in languages like Jav. or Tag. which distinguish PAN *a + dental from PAN *ĕ + dental. Second is the correspondence (mostly in Mdr., but occasionally in the other languages) of /-eq/ and /-iq/ to, respectively, PAN *at/ad and *ĕt/ĕd. Altogether, the evidence suggests that PSS had a contrast in final position between stops *t and *k at a minimum.

b. PSS final *p is tentatively reconstructed on the basis of the change *ĕ > /u/ in a handful of examples involving known cognates in PAN with *-ĕp. The presence of a labial seems to me the only possible conditioning factor for such a change, and is supported by the same change (*ĕ > u) before *-m. Otherwise, PSS *-p is neither retained, nor reflected in the morphophonemics of any SSul language. (Final *p is reflected everywhere as /q/, final *m as /n/ in Sad. PUS (my data), Mass., /ng/ in Bug. Mak. Seko, /n/ or /ng/ in Mdr. As mentioned above, Adriani and Kruijt and van der Veen both report dialects of PUS with /-m/, in examples which do not appear to be borrowings.

c. The continuant *R (presumably a voiced velar fricative) is reflected as follows : *R > Duri (Cakke dial.) /h/, Duri (Kalosi dial.) Ø, /r/ in Mdr. and Mak., /q/ elsewhere. The sequence *-aR is distinguished from *-ar in Bug. in that *-aR > /aq/ while *-ar > /e/ and /o/ preceding the reflex of *-R in most languages, but remain as /i/ and /u/ preceding that of *-r. Possibly this last change could be ascribed to PSS, but the lack of clear cognates from PUS and Seko makes this risky. Note that in Duri, unlike all the other languages, *R and *-s have merged, both > /h/.

d. The neutralization of final contrasts took place, I believe, via several stages, and by implication, over a considerable period of time. My hypothesis is that it involved the shift of final stops, via an "unreleased" stage (as in BI), to /q/; it began by affecting *-p only,

---

([12]) This variation is found in my Balanipa data; according to the informant, it is the lingua franca of the area. Pelenkahu's Majene data show /q/ > Ø before both suffixes ; it is noteworthy that his dial. thus shows consistent development in just the two areas of variation in my data (the other being the final n/ng variation).

and happened first in absolute final position, then was extended to suffixed forms. Thus, phonemic /p/ with allophones [q] finally, [p] elsewhere, shifted to phonemic /q/. Just at that point, this new phonemic /q/ was reinterpreted as an allophone of *-t, which had also begun to shift toward [q] in final position. This is the stage where Duri is today: e.g. rakaq 'embrace', sirakatan 'embrace one another' < PSS *(dr)akĕ(p?), PAN *ḍakĕp; or tukaq, tukatan 'go up' climb' < PSS *tukat, PAN *tukat. Bug. and Sad. have gone a step further: first, /-r/ and /-s/ became [-q] (but remained [r] and [s] before a suffix); meanwhile the -q/-t- (that is, still /t/) became -q/-q- (now /q/), at which point it was reassigned, arbitrarily either to the -q/-r- or the -q/-s- alternation. Thus in modern Sad., words with /-q/ have suffixed forms with either -r- or -s-: e.g. iruq 'drink', irusan 'drinking place' <PSS *iru(p?), PAN *hiRup; or tukaq 'go up', tukaran 'ladder.' Likewise in Bug., but with the additional shift of /-k/ > /q/, resulting in three possible alternants /r s k/ for /-q/. Mak. and Mdr. have undergone much the same development, but since the continuants /r l s/ were preserved, there were never -q/-r- or -q/-s- alternations to which /-q/ could be reassigned. Assuming unimpeded development in Mdr., we could predict that at some future time /-q/ will be lost entirely. I know of no better illustration of Sapir's concept of *drift* in linguistic change; the tendency to reduce final consonants toward Ø is certainly a hallmark of SSul languages, and indeed, seems also to characterize AN phonology generally. Thus we can conceive of the AN languages as a continuum, beginning with the Philippines, where voiced and voiceless stops, nasals and continuants all occur finally, and ending with, the Oceanic languages where all final consonants have been lost (but reappear in suffixed forms — it seems that the finals are rarely lost *in toto*). For a full discussion of this question, see my paper (Mills 1974).

The principal difficulty with the reconstruction of PSS finals is, of course, that it is largely dependent on outside evidence from Ml., Jav. or the Philippines, where the PAN finals are most faithfully preserved. Where such evidence is lacking, it is possible only to reconstruct an unspecified final, symbolized *(C).

6. *Reduplicated monosyllabic roots.* This is a class of words reconstructed for PAN, and usually retained in their full form only in the Philippine languages, e.g. Tag. bukbok 'woodborer' (BI bubuk) or Tag. dingding 'wall' (BI dinding). In the SSul languages, the medial clusters in such words have undergone changes: the syllable-final nasals become homorganic with the following consonant (and then develop like any other nasal cluster, see §2h-j), otherwise the

first consonant > /q/ with the result that many of the clusters are phonetically indistinguishable from geminates. Examples: Sad. bumbun, Bug. wumpung 'pile up'; Sad. buqbuk, Bug. bĕqbuq (phonemically /bĕbbuq/); or Bug. mimmiq, Mak míqmisiq 'suck'. These could reflect PSS and PAN *bunbun, *bukbuk, *mismis, but I have chosen to assume that the medials clusters were *not* preserved intact into PSS, and so reconstruct PSS *bumbun, *buqbuk, *miqmis. This is the only environment for *q in PSS. In two examples, I reconstruct *h in the same environment, to account for an otherwise irregular correspondence (see *(dr)ehde and *bahba in the Appendix). Should additional examples of this sort be found, *h will be justified; but for the present it remains rather hypothetical, and other analysts might prefer to dismiss the examples as "unexplained irregularities."

7. *Comparative morphology.* Research into this aspect of the language family was slighted, unfortunately, due to lack of time and to my desire to collect fairly large amounts of vocabulary in some of the previously unrecorded languages. Certain affixes are found, indeed, in all the languages, they are not evidence *per se* of any close relationship. However, combined with the phonological evidence, and two important innovations, they do support the hypothesis of a SSul language family. Among others we find:

a. *Verbal prefixes.* (i) *mang- is attested in every language, marking active, generally transitive, verbs. (ii) *maR- and *mi- are in near complementary distribution: *maR- > Bug. Mak. /maq-/ (sometimes /maqr-/ before vowels), /maq-/ also occasionally in Mdr. and Sad. *mi- > mi-/ or /me-/ in Mdr. Sad. Mass. PUS Seko. It is found infrequently in Bug. as /me-/, not at all in Mak., though the parallel prefix /pi-/ (< *pi-) 'agentive; causative' does occur in Mak. In function, *maR-/*mi-, are equivalent to BI ber-. (iii)*taR- and *ti- (comparable in function to BI ter-) have the same distribution as *maR-/*mi-, thus *taR > Bug. Mak. /taq-/ (sometimes /taqr-/ before vowels; frequently too Bug. /caq-/), *ti- > Mdr. Sad. Mass PUS Seko /ti-/. Both *ti- and *mi- may be signs of influence, perhaps as a substratum, from the Toraja languages now spoken in Central Sulawesi, since the /me-/ and /te-/ prefixes in those languages can be derived regularly from PAN *maR- and *taR-. (iv) *ma- > /ma-/ in all languages (but rare in Mak.) as a marker for adjectives. (v) *pa- > /pa-/ in all, forming causative verbs.

b. *Verbal suffixes.* (i) *-i > /-i/ in all the languages, corresponding historically and functionally to BI -i, a locative *cum* transitive suffix. (ii) *-an > Sad. Mass. /-an/, Mak. /-ang/, Bug. /-ang/,

/-ĕng/([13]), Mdr. /-an/ or /-ang/ depending on dialect. It marks objective or dative focus: compare Madurese -an, or some uses of -kan in BI. Seko has /-ing/ in this function (of doubtful cognacy), recalling rather strikingly Balinese -in.

c. *Nominal affixes.* (i) *-an > Bug. /-ang/ and /-ĕng/ (same conditioning as for verbal *-an), Mak. Mdr. Seko /-ang/, all others /-an/, forming abstract or locative nouns. (ii) *ka-...-an > Bug. /a-...-ang,-ĕng/, Mdr. /a-...-ang/, Mak. /ka-...-ang/, Sad. Mass /ka-...-an/, no examples for PUS or Seko; a noun formative comparable to BI ke-...-an.

d. Two innovations worth comment are: (i) the words for *eight* and *nine*. In all the languages, except Mak., these are formed with a prefix *ka- plus the word for 'two' and 'one' respectively — i.e. *eight* can be conceived of as "the second (from ten)", *nine* as "the first (from ten)". Here the SSul languages parallel Ml. and Sundanese, which have comparable constructions even though they are based on other lexical items. (Mak. has borrowed and/or neologized its words for *seven, eight* and *nine).* All the neighboring languages on Sulawesi retain reflexes of PAN *walu 'eight' and *siwa 'nine'. (ii) A form of verbal "conjugation" employing affixes to denote both the subject and the object of a verb — e.g. Bug. Mdr. /uítai/ 'I see him/it/them' (< /u-/ 'I', /ita/ 'see', /-i/ '3d person'). Data on this point are lacking in PUS and Seko, but the same constructions almost surely occur. The immediately contiguous languages of Rampi, Bada and Leboni (but not Bare'e) share this feature, perhaps due to influence from SSul ([14]). We can reconstruct: prefixes *ku- 'I', *mu- 'you' (Mak /nu-/ is not directly relatable to this, but reflects a variation mu/nu that goes back at least to Proto Indonesian, if not PAN), *na- '3rd person', *ta- 'we, you (plural)'. These prefixes indicate only the subjects or agent. To indicate the object, or the subject of an intransitive verb, the suffixes: *-a(C) 'I/me', *-ko 'you', *-i '3d person', *-ki(C) 'we/us; you (plural)', *-ka(m?) 'we/us (exclusive)' — this latter obsolete in most languages. All of these clearly derive from the PAN pronouns *aku, *kaw, *ia or *-na, *kita, *kami, but the reasons for the vowel losses are obscure.

([13]) Their occurrence is phonologically conditioned: /-ang/ after vowels, /-ĕng/ after /q/ and /ng/.

([14]) My impression is that a direct genetic relationship between these languages and SSul is unlikely; but the hypothesis cannot even be tested until much more data are available from the Toraja area.

8. *Subgrouping.* Much more information on dialects, and on the grammar of individual languages, is needed before a really accurate grouping of the SSul languages can be determined. What I propose here is based primarily on phonological developments, together with an informal lexicostatistical analysis (based on the Swadesh 200-item wordlist which, happily, coincides with the phonologically-based classification. We can also tentatively posit a "homeland" area from which the dispersion of linguistic groups can be tied in with geographical factors: to wit, the lower course of the Sa'dan river. This is, of course, the area of Sulawesi to which the Sa'dan Toraja traditionally ascribe their origin. ([15])

Probably the first to break off were the linguistic ancestors of the Makassarese. Their move must have taken place before the shift Nasal + voiceless stop > geminate vl. stop, which is found, in greater or lesser degree, in all the other languages. They moved into the flatter land to the south, but must also have spread east to the shore of Teluk Bone — witness the river name Jenemaeja (Mak. jeqne 'water', maeja 'red') south of Palopo. Lexicostatistically, Mak. shows very low percentages with all the other languages, implying a long separation.

One of the next groups to move must have been the Mandars; since the desirable land to the south was already occupied, they moved north along the coast into their present area extending roughly from Polewali to Cendana, but not, apparently, very far inland. Meanwhile, probably from a more easterly part of the "homeland", the Bug. must have moved up the Sa'dan valley, then across to the coast from the Rantepao area (like today's highway) to the Palopo area — the easiest route, avoiding the high mountains around Latimojong. Subsequently, from the Luwu' area, the Bug. have expanded to the south, within the Christian era if the lontara' can be trusted. Their movement has been, obviously, at the expense of the Makassarese, who must have been pushed steadily south by the Buginese, until the present distribution was achieved sometime prior to European contact (early XVI Century). It would be tempting to view the Bug. expansion as a concomitant of rapid population growth (in turn due to an increased food supply) together with more efficient political organization. Both factors suggest influence from outside: the introduction of (1) wet-rice culture and (2) Indianized political concepts.

The next group to move can be called Toraja, in the broadest

([15]) They claim to have come to Sulawesi from "an island in the south". See van Lijf 1947.

sense, including PUS-Seko and Sa'dan. The PUS-Seko group might have followed the Mamasa river to the west, suggesting that the Sa'-dan people blocked the way up the Sa'dan river. Once in the west, Seko in turn broke off, or was pushed, first toward the coast, then north and finally inland, via the Karama river, to its present location.[16])

The upper Sa'dan valley (Makale-Rantepao area), according to van Lij (1947: 519), is the point from which the Mamasa, Mangki and Rongkong groups have spread; the Mangki (Kalumpang area) are said to have reached their area only around the XVI Century. The brief discussion in Van der Veen (1929) shows that the outlying dialects differ from Makale-Rantepao mainly in the retention of palatals and intervocalic /w/, and some vocabulary.

This leaves — still in the hypothesized homeland area — the languages and dialects of Massenrempulu'. Those for which I have data all measure over 80% lexicostatistically with each other; their next highest percentage is with Sa'dan (71%), and I suspect that a detailed, village-by-village survey would reveal a continuum between Mass. and Sad. comparable to, say, the North German-Dutch border area. The same might be true, too, for the area between Mamasa and its neighbors seems to be due mainly to political and religious factors.

9. A problem that can only be hinted at here, involves the putative presence of a substratum in the SSul languages. If there is even a grain of truth in the old concept of Proto- and Deutero-Indonesian migrations, and if the SSul peoples were among the latter, then it follows that they found the land already inhabited; and the most obvious candidates for the indigenous population are (1) Bare'e speakers, (2) West Torajas, e.g. Rampi, Leboni and Bada' speakers, (3) Bungku-Mori speakers. (From the little data available, it appears likely that these three ultimately form a single group). As the SSul peoples spread north and south, the indigenes were absorbed, enslaved, pushed out, and so forth. This will account for the large number of "irregular" cognates in SSul languages— usually showing open syllable structure where the known PAN etymon shows a final consonant. Many such items, in Matthes' Bug. dictionary, are marked as "Old Bug." or "Basa Bissu", and can easily be found in e.g. Bare'e. Just two examples: "Basa Bissu" taliyu 'back', cf. Bare'e taliku, Napu taleu (Tag. talikud); or "Old Bug." eyo 'sun' Bare'e eo. According to Esser (cf. Noorduyn 1963: 356ff), the language of Wotu retained many of these in current use; a detailed study of that language should be of the highest priority,

---

(16) This route was related to me by one of my Seko informants, who asserted it was "bukan dongeng" — not a legend or myth.

since Wotu lies precisely at the point where Bug., Bare'e and Mori all converge. I must leave it to the anthropologists and archaeologists, however, to determine the validity of my substratum hypothesis.

10. The reconstruction of PSS is, I believe, an important step. It now becomes possible to compare PSS with other proposed subroupings within AN, rather than with PAN directly. It enables us to speculate somewhat more securely about the ultimate origins of the SSul peoples— viz., if they are indeed relative late-comers to the island, then their nearest linguistic relatives can be sought elsewhere. From a methodological point of view, one could wish that the languages showed fewer irregularities. But I believe it is possible to account for most of these in a principled way, under the assumption that we are dealing with a language family whose members (unlike the Indo-European languages) have been in close and constant contact throughout their history. That phenomenon, in itself, should be of interest and relevance to theories of linguistic change.

REFERENCES CITED

Adriani, N. and A.C. Kruijt (1914) : *De Bare'e-Sprekende Toradjas van Midden-Celebes : 3de Deel : Taal- en letterkundige schets der Bare'e taal en overzicht van het taalgebied Celebes-Zuid Halmahera.* Landsdrukkerij, Batavia.

Dempwolff, C. (1934-37):*Vergleichende Lautlehre des Austronesischen Wortschatzes.* Zeitsch. f. Eingeborenen-Sprachen, Beihefte 15, 17, 19.

Lijf, J.M. van (1947) : "Kentrekken en problemen van de geschiedenis der Sa'dan-Toradja-landen." Indonesië I : 518-535.

Matthes, B.F. (1859): *Makassaarsch-Hollandsch Woordenboek.*

—— (1874): *Boegineesch-Hollandsch Woordenboek.* M. Nijhoff, 's Gravenhage.

Mills, R.F. (1974): "Proto South Sulawesi and Proto Austronesian : A possible criterion for subgrouping". Paper presented to the lst International Conference on Comparative Austronesian Linguistics, Honoludu, Hawaii, January 1974.

—— (in progress): *Proto South Sulawesi and Proto Austronesian Phonology.* PhD dissertation, Univ. of Michigan.

Noorduyn, J. (1963): "Mededelingen uit de verslagen van Dr. S.J. Esser". Bijdragen 119 : 329-370.

Polenkahu, R.A. (1967): *Gambaran Sepintas Lalu tentang Dialek2 Mandar*. Unpublished *Sarjana* thesis, IKIP-Makassar.

——, Abd. Muthalib, Muh. Amin and W.M. Manala M., eds. (1972): *Bahasa di Lima Massenrempulu*. Lembaga Bahasa Nasional, Ujung Pandang (Makassar).

Ras J.J. (1968-70): "Lange consonanten in enige Indonesische talen". Bijdragen 124 : 521-541, 126 : 429-447.

Veen, H. van der (1929): "Nota betreffende de grenzen van de Sa'dansche taalgroep en het haar aanverwante taalgebied." Tijdschrift LXIX : 58-96.

—— (1940) : *Tae' (Zuid-Toradjasche)—Nederlandsch Woordenboek*. M. Nijhoff, 's Gravenhage.

APPENDIX : PSS RECONSTRUCTIONS

(NB. Items marked MWB, BWB, SWB, A&K, V (= van der Veen 1929), were not collected by me, but come from a written source. A&K in particular do not indicate geminates or final /q/, hence their forms must be understood as "sic" if the reflexes appear to be irregular. Likewise, Seko forms are all "sic" unless specified "(heard)"; the informants who filled out my word-list were careless (?) about indicating final /q/, and in a few cases I heard a geminate consonant where the informants wrote a single one. Thus while my Seko materials is somewhat flawed, to my knowledge there is no other available, and I consider it important enough to use. I would like to hear from Dutch scholars who may know of Seko word-lists etc. that may exist, for example, in the archives of the Netherlands Bible Society.)

1. *pitu 'seven' : Bug. Sad. Mdr. Duri PUS Seko pitu.
2. *kapas 'plant fibre (BI kapas)' : Bug. apĕq, Mak. kápasaq, Sad. kapaq, DuriC kapah, DurikK, Seko kapa.
3. *tai 'excrement' : Bug. Mak. Sad. Mdr. Mmj. Duri Seko tai.
4. *lima 'five ; hand' : Bug. Mak. Sad. Mass. Mdr. PUS Seko lima.
5. *sulengka 'sit with crossed legs' : Bug. Mdr. DuriK sulekka, Mak. Sad. sulengka.
6. *tama 'enter' : Bug mutama, Mak. antama, Sad. Duri Mdr. Seko tama.
7. *tunu 'burn' : Bug. Mak. Sad. Mass. Mdr. PUS Seko tunu.
8. *tongAn 'true' : Bug. tongĕng, Sad. Mdr. Mass. PUS tongan, Seko tongang. ("*A*" means that either *a or *ĕ is possible.)

9. *sĕbbu 'thousand' : Bug. sebbu, Mak. Sad. Mass. PUS saqbu, Seko sabu.

10. *puso 'heart', : Bug. uso, Sad. puso 'banana blossom' (cf. BI jantung pisang !); Mass SekoP puso 'heart'.

11. *gĕnnĕ(p?) 'enough, complete' : Bug. gĕnnĕq, Mak. Sad. Duri gannaq, Seko kanna, kanaq.

12. *sugi(C) 'rich' : Bug. Mdr. Sad. sugiq, Mak. DuriK sugi, Mak. (Konjo dial.) súgiriq. Seko sukiq. (These point to *sugiR, but cf. Ml./Jav. sugih.

13. *kayu 'wood, tree' : Bug. aju, Mak. Sad. Seko kayu, Mak. (dial.) Mass. PUS kaju, Mdr. ayu.

14. *gayang 'kris ; to stab' : Bug. Duri gajang, Mak. Mdr. gayang, Sad. (dials.) gayang, gajang, gaang.

15. *bete 'fish' : Bug. Mak bete2, Sad. bete 'names of various fish spp.', Seko bete 'fish (in general)'.

16. *bintuin 'star' : Bug. wittoeng, wittoing, Mak bintoeng, Mdr. Seko bittoeng, Sad. bintoen, Duri bintuin, bentuin, PUS bittoen.

17. *lompo 'big ; fat' : Bug. loppo, Mak. Sad. Duri PUS Seko lompo ; note also Mass. loppo 'fat ; fertile (land)'.

18. *tambing 'house, hut' : Bug. tamping, Mak. Sad. tambing 'addition to a house' (e.g. in Sad., the slaves' quarters); SekoP tamming 'house'.

19. *tanduk 'horn' : Bug. Mak. tanruq, Mdr. PUS tanduq, Sad. Duri tanduk, SekoP tannuku (=tánnuk(u)).

20. *tanggung 'carry' : Mak. tanggong, Sad. tanggung Duri tanggungan, SekoP tangkung. Though possibly < Ml./BI, it still shows the expected -ngk- in Seko.

21. *salangga 'shoulder' : Bug. salangka, Mak. Sad. salangga. Presumed borrowing < Bug. : Duri salangka 'joint', Mdr. salakka 'mid-back (kidney area)'. Cf. also Sad. sarangga 'breastbone (of animals)' ; the item shows the semantic "wandering" so typical of body-part names.

22. *insAn 'to know' : Bug. issĕng, Mak. asseng (metathesis?), isseng (the /e/ is irregular), Sad. Mdr. PUS issan, PUS(V) insan, Duri issen (irreg. /e/), Seko issang, isang.

23. *tĕllu 'three' : Bug. tĕllu, Mak. Sad. Mass. Mdr. PUS tallu, Seko italu (but I heard itallu).

24. *pĕnno 'full' : Bug. pĕnno, Mak. (dial.), Mdr. panno, Sad. Mass. Mmj. ponno, SekoP puno.

25. *bĕqbĕ(t) 'wind up' : Bug. bĕbbĕq, balĕbbĕq, Mak. Mdr. Sad Duri Mmj. baqbaq, Seko mambabai. Note also Sad. bigbiq 'to wrap, bind'. (*t assumed on the basis of PAN *bĕgbĕg, Ml. bebat.)

26. *sĕdding 'to feel, experience ; to hear' : Bug. sĕdding, Mak. saqring, Sad. Mdr. Duri saqding ; SekoP pisadingi 'hear'.

27. *raki 'dirt on the skin' : Bug. SekoP rai, Mak. raki.

28. *ciñcin 'ring' : Bug. ciccing, Mak ciñcing, Duri ciñcin, Sad. sissin.

29. *jaru(m?) 'needle' : Bug. Mak. Duri jarung, Sad. darun, Mdr. rarung. PUS darung. SekoP has jarung (probably borrowed) beside mandaung 'to sew' (probably native). The Mdr. /r-/ may also reflect a borrowing.

30. *jĕlluk 'to point' : Bug. jĕlloq, Sad. dullok, dollok, dullok, Duri mangjulluq, Mmj. meñjolloq ; Mdr. pañjolloq, Mdr. (A&K) p'dolok 'index finger'.

31. *ñawa 'soul ; breath' : Bug. Mak. Mdr. Mmj. Duri ñawa, Mdr. Mmj. also nawa, PUS Seko inaha, Sad. penaa. Cf. also Bug. Mak. nawa 'to think, consider'.

32. ñañi 'sing' : Duri ñañi, Mmj, menani ; Sad. Seko menani 'a ritual with song and dance'.

33. *bulan 'moon, month': Bug. wulĕng, ulĕng, Mak. Mdr. bulang, Sad. Mass, bulan, Seko hulang.

34. *balawo 'rat, mouse' : Bug. Mak. Sad. (Rongkong dial.) balawo, Sad. Mdr. Duri balao, PUS, Mmj, balaho, PUS(V) (Bamban dial.) balabo.

35. *(dr)ara 'blood' : Bug. Mmj. dara, "Old Bug." Sad. PUS rara, Mass. dara, rara, SekoL raqa (heard), SekoP raa, PUS(V) lara, "chacha" (=haha?), Mmj. (A&K) haha.

36. *(dr)alan 'path, road' : Bug. lalĕng, Mak. lalang, Sad. Mass. Mdr. PUS lalan, Seko dalang. (The sequence #rVlV... is not permitted in any SSul language ; further, Sad. and perhaps Mdr. and Seko, also forbid #lVrV..., which assimilates to #rVrV...).

37. *ta(dr)ĕ(m?) 'sharp' : Bug. tarĕng, Mak. tarang, Mdr. Mmj. tadang, Sad Mass. PUS taran, Seko tarung. Note also Sad. tinaran, PUS(V) tinarun (< *t-in-a(dr)ĕ(m)) 'blow-gun dart'.

38. *ramba 'chase' : Sad. Mass. SekoL ramba, Mmj. hamba. Perhaps Bug. rampa2 'move a little bit' and Mdr. rimba 'chase'?

39. *ĕrrĕ(t) 'tight' : Bug. (BWB) mĕrrĕq, Bug. (informants), Mak. marreq, Sad. marraq ; SekoP paraq 'to tighten'. Cf. Ml. erat.

40. *ura(t) 'vein, tendon' : Bug. urĕq, Mak. Sad. Mass. Mdr. SekoP uraq ; PUS(V) "uchaq" (uhaq?), Mmj. (A&K) ura, uha, Mdr. (A&K) ue (= ueq?).

41. *pĕzzu 'gall ; liver' : Bug. (BWB) ĕssuq (informants ĕssung), Sad. Duri PUS paqdu, Mmj. (dial.) paqde, Mdr. paqdung, SekoP puru.

42. *(b)uzing 'charcoal' : Bug. os'ng, Mdr. boring, Mmj. ohing, SekoP oring. Presumed borrowings < Bug. : Sad. PUS, Mmj. SekoL os'ng.

43. *atĕ(p?) 'roof' : Bug. atĕkki 'to roof' : Mak. ataq, Mdr. ateq, Mmj. Seko atuq.

44. *(dr)akĕ(p?) 'to embrace' : Mak. Sad. Duri rakaq (note Duri sirakatan 'embrace each other'), Mmj. sihakaq, Mdr. siraetti (base raetti < *(dr)akep + i); perhaps also Sad. rakuq 'graps with the hand'. Bug. rao, kadao is ultimately cognate, but loss of the final is irregular. PAN *dakĕp.

45. *pala(t) 'palm of the hand' : Bug. palĕq, Mak. Sad. Duri PUS Seko palaq, Mdr. paleq. PAN *palag.

46. *ulĕ(t) 'worm' : Bug. ulĕq, Mak. oloq, Sad. Mdr. Duri ulliq, PUS(V) olliq, illiq. The geminate /ll/ is inexplicable.

47. *ĕnnĕm 'six' : Bug. enneng, Mak. Mdr. PUS annang, Sad. Duri PUS annan, Mmj. unnung, annang, Seko unung. PUS(V) unung, Mmj. (A&K) unum.

48. *taněm 'to plant' : Bug. taněng, Mak. Mdr. Mmj. Mass. tanang, Sad. Mass. tanan, Mmj. (A&K) mantanam.

49. *wai 'water' : Bug. wae, uwae, Sad. wai, uai, Duri wai, uai, Mdr. uai, Seko PUS uhai,Mmj. wai.

50. *wase 'axe' : Bug. uwase, Sad. Mass. Mrd. wase, uase, Mmj (A&K) uhase.

51. *so(m)bAl 'sail' : Bug. sompĕq, Mak. sómbalaq, Sad. sombaq, Mdr. sobal ; PUS sumombal, Mmj. sumobal 'to sail'.

52. *ukir 'to write': Bug. Sad. Duri Mmj. ukiq, Mak. úkiriq, Mdr. ukir.

53. *ipaR 'brother-in-law': Bug. Sad. SekoL ipaq, Mak. íparaq, Mdr. ipar, DuriC ipah (DuriK ipaq has irreg. /q/) SekoP ipa.

54. *sangiR 'to sharpen': Bug. sangiq, Mak. sángiriq, Sad. sangeq, DuriC sangeh, DuriK sange.

55. *nipis 'thin': Bug. Sad. manipiq, Mak. nípisiq, Mdr. nipis, DuriC nipih, DuriK, PUS nipi. Seko manipa is irreg.

56. *panas 'hot': Bug. panĕq, Mak. pánasaq 'a plant sp. used in making *acar*'; Sad. panaq 'pepper; ginger; spicy hot'; DuriC panah, DuriK pana 'ginger'; PUS mapanaq 'hot'; Mmj. panaq, panas, SekoP mapanaq 'sick'.

57. *(dr)inding 'wall': Bug. rĕnring, Mak. rinring, Sad. Duri Mdr. PUS Mmj. rinding, Mmj. (A&K) hinding, Seko rinning. PAN *dingding.

58. *duqdu(C) 'sleepy': Bug. caqkaruqduq, Mak. doqdoq, Sad. tikkaruqduq, Mdr. menduqduq.

59. *(dr)ehde 'to boil': Bug. rede, Mak. rere, Sad. Mdr. DuriK reqde, Mmj. (dials.) dede, rere, hehe. Cf. Ml./BI didih; presumed < PAN *ḍihḍih.

60. *bahba 'door, opening': Bug. babang (< *baba + an?), Sad. baqba 'door'; Bug. baba, Mdr. baqba 'river-mouth'; Mak. bawa, (dial. baba) 'mouth'. Dempwolff reconstructed *babah, but Philippine languages, and SSul, point to *bahbah.

Légendes de la planche ci-contre p. 134.