ON THE EVIDENTIAL BASIS FOR THE SOUTH SULAWESI LANGUAGE GROUP

Ülo Sirk
Institute of Oriental Studies,
U.S.S.R. Academy of Sciences, Moscow

In this paper the author analyses a number of phonetic, morphological and lexical features (among them nearly 30 lexical roots and affixal derivates) which are characteristic of the South Sulawesi languages. Some of these features seem to be exclusively shared innovations. The results of three statistical counts are also presented. While the functor statistics and, in general outline, also the etymostatistics agree with the notion of a South Sulawesi language group, the lexicostatistical count sharply contradicts the latter. Nevertheless it is not possible either to abandon or to radically change the notion of the South Sulawesi group, which is supported by a considerable body of qualitative evidence.

1 INTRODUCTION

The notion of South Sulawesi (henceforth: SSul) language group introduced by Esser (1938) has survived remarkably well. For the same unit, Mills (1975a, 1975b, 1981) worked out his proto-language reconstruction. The SSul stock considered in the Grimeses' book (1987) and by Friberg and Laskowske (this volume) matches Esser's SSul group quite neatly as to the general (map) contours, but the language lists given in the two last-mentioned works are much more detailed than Esser's list.[1]

The SSul group also appears in the Language Atlas of the Pacific Area (Wurm, Hattori 1983).

Esser's classification was presented in 1938 without any argument whatsoever. Mills, of course, recognized that the SSul group should be substantiated. In his paper published in Archipel (1975a) he noted two innovations said to be "worth comment",[2] but his discussion concerning them (p.216) remained entirely inadequate. In his dissertation (1975b: the text was probably written subsequently to the paper), Mills considered these features without calling them innovations. The latter term occurs in his dissertation but rarely. Nonetheless throughout this work the author makes numerous remarks like "characteristic of the SSul languages", "distinguishes the SSul languages from the Torajan ones" etc. These remarks concern various structural features, phonetic development, certain lexical roots etc., but they are quite sporadic, and no systematic study is carried out. On what basis the SSul languages are grouped together -- such a question is not even raised expressis verbis.

Mill noticed that the boundaries of the SSul group as drawn traditionally since Esser -- and by himself too -- were not confirmed by lexicostatistics. He[3] counted 100 word Swadesh list percentages among seven SSul languages (Bug, Mak, Mdr, Sad, PUS, Mmj, Seko) and four languages not belonging to this group but spread nearby (Bar, Uma, Le, Mori; all of them are labeled by Mills as Toraja). The results of this count are shown in Mills 1975b, Chart 7 (p.492) and Chart 8 (p.515). Of almost all SSul languages drawn into the count, some of them (i.e., Mak, Seko) score lower than some of the Toraja languages. From this discovery Mills did not draw any far-reaching conclusions. In his opinion, the 'puzzling results' of his lexicostatistical comparison may have arisen, to a high degree, from the poor quality of the lists used (which had been compiled either from dictionaries or with the assistance of informants fluent in Bahasa Indonesia) (Mills 1975b:514-517). However, parts of the contrasts discovered by Mills are clearly too sharp to explain in such a way -- in essence, simply disregarded. Compare, e.g., Sad-Mak 33% and Sad-Uma 46.8%. My counts

(see further) are also suggestive of a rather serious problem arising here.

It is well known even if not so often emphasized that it is possible to reconstruct an imaginary proto-language for any sample of languages which contain a sufficient body of comparable material. To meet this requirement, the languages must be interrelated, but it is not necessary that they should belong to one subgroup. A certain danger lies here. Intuitively, we might expect such an imaginary proto-language to show a larger number of inexplicable irregularities than does one reconstructed on the basis of a real subgroup. But what degree of irregularity should be regarded as critical here? Mills' PSS reconstruction actually abounds with unexplained and incompletely (or unconvincingly) explained exceptions and cases of idiosyncratic development. Here are Mills' own words: "From a methodological point of view, one could wish that the languages showed fewer irregularities" (1975a:219).

Hence it is quite an urgent task to elucidate the basis on which the SSul languages are grouped together. In this paper, I attempt to take a few initial steps in this direction.

2 METHODOLOGICAL PRELIMINARIES

What is indicative of a subgroup? The answer usually given is that the criterion should be formed by exclusively shared innovations (henceforth: ESIs). However, the border-line between innovations and retentions does not appear absolutely clear-cut even if research problems (the impracticability of an exhaustive search for cognates in related languages, our insufficient knowledge of many of the latter, difficulties with the task of cognate decision, etc.) are eliminated from consideration.

Frequently a retention includes an innovative element which may be more or less clearly detachable. The so-called semantic innovation appears to be the simplest case. But the innovative element may also be made up by the fact that an inherited phonetic segment or grammatical marker has entered new environmental and/or structural relations (cf. the ligature -ŋ- before possessive morphs in SSul languages; see below). As for their evidential value, such additionally complicated retentions are not much inferior to innovations.

In view of this, and naturally of research problems as well, I hold that the evidence for a language group cannot be confined to ESIs only.

First of all, however, let us make a few remarks concerning the notion of ESI.

I conceive the expression 'exclusively shared' as 'shared by at least a part of the languages of the group and not found outside the latter (apart from the cases identifiable as borrowings)'. This expression does not presuppose the observability of the innovation in all the languages of the given group (cf. Zorc 1986:155).[4] There may be two reasons why an innovation fails to emerge in part of the languages of the group:

(i) Part of the languages may have lost the given item during their subsequent evolution.

(ii) What is called the proto-language was usually a dialect chain (network of dialects). Quite naturally, an innovation came into existence in a certain dialect of the chain and spread to others. If the dialect chain (i.e. proto-language) had already begun to disintegrate, the innovation may have failed to intrude into some of the dialects.

In view of these points, we must take overlapping innovations into account. In this paper, however, I consciously deviate from this principle in two cases: firstly, in view of the somewhat problematical character of the position of the Makasaric subgroup (Mak, Ko, Sel) I insist that the features to be ascribed to the SSul group should be reflected in Makasaric, and, secondly, in view of prolonged contacts, I disregard features whose distribution in the SSul languages is limited, on the available evidence, to Makasaric and Bug.

Of course, the ESI is a hypothetical category, as is the innovation in general. ESIs are hard to find, and their reliability is particularly low in large and insufficiently studied families. Nevertheless at least a small

number of putative ESIs are needed for establishing a language group (without them, we might take the sum of a few remotely related groups for a single one). This is why I call ESIs the main evidence for the group, whereas all the remaining evidence is labeled additional (although such additional evidence is far from useless.)

The important types of additional evidence are the following:

(i) Non-exclusive innovations which also occur in some remotely related languages (e.g., the prefixal use of the morpheme um-, see further).

(ii) Those non-trivial retentions which clearly are not really conditioned (are not bound to certain historical regions, to certain river basins, to almost inaccessible hinterland areas, or the like). By the term 'non-trivial retentions' I mean the features which, lacking any innovative constituent, neither reflect highly retentive proto-forms nor result from widely attested ways of development.

However, if an isogloss separates the group under consideration from some other group(s) the relations with which seem not quite clear at the given stage of research, then that isogloss should be treated as reflecting a non-trivial retention even if it covers a large number of related languages outside the group. In this case, however, the evidential value of the isogloss is rather low. Retentions which appear to be simply trivial (i.e., highly retentive) or areally conditioned are of no relevance to language subgrouping.

(iii) As additional evidence for a language group, we can treat results of statistical counts not contradicting the qualitative evidence.

The scheme suggested is clearly of an 'ideal' nature; the reality seems to be still more complex. In reality, it is often difficult to distinguish between the subtypes of retentions, and sometimes even between a retention and an innovation. Therefore I don't follow this scheme in arranging the qualitative evidence to be presented.

3 QUALITATIVE EVIDENCE

3.1 Phonetics and linked items

3.1.1 Development of final *-q

Pre-PSS final *-q has been lost in all the SSul languages for which material is available to me, whereas the high vowels that had preceded it have been lowered, and everywhere but in the Seko language the reflexes of the pre-PSS final segments *-iq an *-uq have merged with those of *-ay (still earlier merged with *-əy) and *-aw, respectively. In Seko, judging from the available data (from Grimes and Grimes 1987, and from Martens, to appear), the final segment *-Vq has become a long vowel. Compare:

PrePSS	Most SSul languages	Seko
*-aq	-a	-a:
*-iq	-e	-e:
*-ay	-e	-e
*-uq	-o	-o:
*-aw	-o	-o

True, my data is too scant to find out how regular the vowel lengthening in Seko is.

The lowering of high vowels before final -q unites the SSul languages with Badaic while separating them from the body of the Kaili-Pamona (henceforth: KP) group as well as from Wol and apparently (judging from poor data) from Lay and Wotu. In Mori, Bungku, Toolaki and Muna the situation seems too complex to be elucidated on the scarce data available so far.

If Seko is indeed a SSul language, it would allow us to explain the process of the -Vq > -V development in this group, that is, through V: (this way would prove to be characteristic of SSul, though not necessarily exclusive, as several non-SSul languages might have similarly gone through the -V: stage).

3.1.2 Vestiges of the nasal ligature before possessive pronominal enclitics

This phenomenon, which is of great interest for defining the SSul group, is discussed in some detail in Sirk 1988. Here I will make a few remarks only.

In the majority of the better-known SSul languages the possessive pronominal enclitics, or at any rate a large proportion of them, have pairs of allomorphs like Sad:

-ku	~	-ŋku	'my'
-mu	~	-mmu	'your (sing.)'
-na	~	-nna	'his, her, its, their'
-ta	~	-nta	'our (incl.); your (polite)'
-ki	~	-ŋki	'our (excl.)'
-mi	~	-mmi	'your (pl.)'

Schematically:

Set 1	Set 2
-CV[...]	-N_hCV[...]

In some languages, in compliance with their rules, NC clusters shift to geminates; so Sad -nta finds its counterpart in Bug, Mdr -tta.

In the use of possessive enclitics after vowel-final words, a number of languages -- most regularly Sad -- follow the distributional scheme:

after words ending in e or o	--	set 1
after words ending in i or u	--	set 2
after words ending in a	--	either set 1 or set 2

The Makasaric languages generally do not obey this scheme, but as far as certain individual words are concerned, its influence is often felt in Makasaric.

The typical pre-PSS sources of SSul final vowels are the following:[5] s

-e	<	*-ay (+ -əy), *-iq
-o	<	*-aw, *-uq
-i	<	*-i (in few cases, < *-uy or *-iw)
-u	<	*-u
-a	<	*-a, *-aq

After words ending in -a, we can note the tendency toward the historically conditioned set choice according to the following rules:

after	-a	< *-aq	---	set 1
after	-a	< *-a	---	set 2

This tendency is strongly pronounced in Sad and Bug, while in Mak a number of frequently used a-final words follow it. Compare, for example:

Sad, Bug, Mak lila-na 'his tongue' (pre-PSS *lilaq, PAN *dilaq 'tongue'), and
Sad, Bud, Mak lima-nna 'his hand' (pre-PSS, PAN *lima 'hand').

Evidently the two-set allomorphic structure of possessive enclitics is inherited from some pre-PSS stage. Both the shape of the set 2 allomorphs and the distribution of the sets can be explained by the conjecture that at that stage there existed the nasal ligature which appeared in the postvocalic position only (such a rule is observable in various AN languages). The word + possessive compounds typical of the stage under discussion can be schematized as follows:

	Word				Possessive enclitic
(1)	V	-	$-N_h-$	-	-CV[...]
(2)	C	-		-	-CV[...]

(Note that y, w and *q are classed as consonants.)

Later final -y, -w and -q were lost, and the distribution of the CV and N_h-CV segments (the latter became set 2) was transferred to a new basis which may be called 'essentially historical'. This transfer of the distribution basis makes up the innovative constituent of the observable complex feature (the distribution pattern of the allomorphs of possessive enclitics), which obviously is a retention in essence.

Possessive allomorph pairs like -CV[...] ~ $-N_h$CV[...] are certainly no rarity in the AN family (in the Sulawesi region, besides the SSul languages, we might refer to the Badaic ones, as well as to Rampi, Bungku and Mori). However, to my knowledge, the principle according to which the choice between these allomorphs is governed by the preceding vowel -- a point leading to important conclusions about historical development -- has not been observed anywhere but in SSul languages (even the Badaic languages differ here sharply).

3.1.3 r from PAN *R

The SSul languages, except a few PUS dialects (which have changed r into h), generally reflect PAN *R as r in most positions. While typical of the Batak, Malayic and Chamic groups, of Mad, Maloh, of several languages of the Moluccas, etc., this reflex of *R appears to be virtually absent from the Sulawesi region -- not counting, of course, the SSul languages. (Balantak also seems to be exceptional for the region in that it reflects *R as r. We also see r < *R in the far north, in a dialect of Sangirese.) The r-reflex of *R sharply contrasts the SSul languages with the Badaic ones, and with the KP languages in general, as well as with Mori, Bungku, Toolaki, Lay, Wol, Muna, Bng, Gorontalo, Mong, Minahasa languages, etc.

3.1.4 Development of reduplicated monosyllabic roots

As shown by Mills (1975a:214-215; 1975b:370-386), the sound correspondences between the SSul reflexes of earlier reduplicated monosyllables (henceforth: RMs) of the shape $*C_1VC_2C_1VC_2$, where C_2 was a stop or a non-nasal continuant (glides excluded), allow us to reconstruct for PSS the shape $*C_1VqC_1VC_2$ (if C_2 was nasal it became homorganic with C_1 in the medial cluster, which resulted in PSS $*C_1VN_hC_1VN_2$). In Mills' opinion, "the change of $*C_1C_2$ > PSS *qC is one of the changes which serve to characterize PSS as a separate entity, and to set the SSul languages apart as a distinct subgroup" (1975b:373). I would like to emphasize that the decisive value of this change for language subgrouping is severely reduced by the possibility that some other groups (e.g. the KP group) may have passed through the same stage of evolution, and also by our inadequate degree of understanding of the development of consonant clusters in the SSul as well as in many other AN languages (the last problem cannot be discussed here). Nevertheless the change of $*-C_2C_1-$ in RMs into PSS $*-qC_1-$ may perhaps serve as an additional argument for the SSul group.

3.2 Affixal inventory

Under this heading, I am ready to discuss only one feature, and even its credibility seems rather doubtful so far.

3.2.1 Verbal prefix putatively reducible to *uN- in preconsonantal position

This prefix is represented by Sad uN-, the final nasal of which is changed according to the following scheme:

Stem-initial phoneme	-N >
Stop	homorganic nasal (no deletions takes place); e.g. un-toe 'hold'
nasal, l, r or s	completely assimilated; e.g. ul-lambi 'to reach'
vowel	-nn- ; eg. unn-alli 'to buy'

Similar prefixes occur in a few languages and dialects close to Sad, but before a stem-initial vowel one usually encounters m- or um-.

From Mdr I can cite only one form with u + nasal before a stem-initial consonant: um-baba 'bring' (LM). Before vowel-initial stems Mdr shows umm- or um-.

It is difficult to determine whether we are dealing with the widespread affix um which is, quite idiosyncratically, used as a prefix in preconsonantal position[6] or with an originally different morpheme whose prevocalic allophone has merged with the allophone of um in most languages and dialects of the Sa'dan - Mandar area but not in Sad. Suffice it to say here that each hypothesis can be supported by some (but rather scarce) extra-SSul evidence.

The Bug prefix manifested by the change of w- into mp-, r- into nr-, by gemination or preglottalization of other stem-initial consonants and by mm- (or m-)[7] before stem-initial vowels is probably a cognate of the Sad prefix uN-.

But what about the Makasaric languages?

In Mak and Ko, there exist side by side two aN-shaped verbal prefixes which are both used in transitive verbs: one of them deletes the stem-initial voiceless stop while the other does not. Let us label these prefixes, respectively, aN-1 and aN-2. Cf. the Mak derivatives from the stem -ta'gala':

anna'gala' 'to hold (something)'
anta'galak-i (-i is the third-person pronominal morpheme)
'to hold (it, or a concrete thing mentioned within the same phrase)'[8]

Friberg and Friberg (to appear) call Ko aN-2 transitive definite.

It is conceivable that the prefix aN-2 observable in Mak and Ko is cognate with the Sad prefix under consideration. The vocalism -- namely a for expected u -- is hardly an obstacle for this comparison. The change of prepenultimate (pre-stress) vowels into schwa (which later, in several groups, give o or a, or some other vowel) is a widespread phenomenon in the AN languages. In Mak and Ko, a parallel can be found in the prevocalic prefix amm-, which undoubtedly has developed from um- through the intermediate stages m- > əmm- (the gemination taking place in the position after schwa).

It is certainly of relevance for our argument that uN-like preconsonantal verbal prefixes -- be they ultimately deduced from the widespread infix um or not -- occur rarely in the AN family. No examples are known from the non-SSul languages of the Sulawesi region. Hence -- given that we are not in error regarding the Makasaric languages -- the observations just carried out may be regarded as offering an additional argument for the SSul group.

3.3 Lexicon

3.3.1 The numerals 'eight' and 'nine'

The main body of evidence to be taken into consideration can be presented as follows:

Language (group)		'eight'	'nine'
Sad (V) [9]		karua (two - dua)	kasera (one - misa(ʾ))
PUS dialects	Mbi (G)	karua (two - dua)	kamesa (one - mesa)
	Bmb (G)	kahua (two - dua)	kaseha (one - mesa)
Seko		karoʾa (M) (two - (i)dua)	kamesaʾa (M, G) (one - mesaʾ (G))
Mdr (M)		aruá (two - dua)	amesá (one - mesa)
Mdr (Mu)		arruá (two - dua)	amessá (one - mesa)
Bug		arua (two - dua)	asera (one - seua)
Ko (SK, DK)		karua (two - rua)	---- (nothing comparable)
Sel (C, H)		karua (two - rua)	kaassa (one - assa (C))

Evidently we are dealing with affixal derivatives here. The roots are easily recognizable in most cases. As for -sera, Mills supposes that it may be an otherwise unattested word for 'one' (1975b:826). I would like to emphasize that the word sera 'one' is really present in Maloh, a poorly known language of central Kalimantan which seems to be somehow related to the SSul group.[10] The root of the word for 'eight' is undoubtedly 'two' (PSS *dua; the change of the intervocalic -d- into -r- may be explained by the considerable antiquity of the derivate). So we have for 'nine' the root 'one'; for 'eight' the root 'two'.

In both cases the prefix is identified as PSS *ka-. In view of the Bug and Mdr (Mu) final stress and the segment -a present in the Seko cognates we can suppose that at the PSS stage the derivatives under discussion also contained the suffix -a. That -a might have been the irrealis marker. No doubt such a marker is quite 'logical' in words which may be analysed from the semantic view as 'when (ka-) 2 or 1 would (-a) remain

(or be deducted, or sim.)'.

Although the method of forming the numerals 'eight' and 'nine' by deduction is used by some languages of Eastern Indonesia (Buru, Soboyo, Sula-Sanana), the affixes used there are not cognate either with PSS *ka- or with *ka-...-a. (it is of course another matter that some languages of Eastern Indonesia, particularly those of Aru Islands, have borrowed Bug or Sel numerals). So we may treat the pattern "*ka- + 'one'/'two' (+ *-a)" of forming the numerals 'nine' and 'eight' as an innovation exclusively shared by the SSul languages.

The PSS lexical item *karua(a) 'eight' is an innovation of this kind.

The PSS word for 'nine' poses more problems. In my opinion, it is likely that at the PSS stage there existed *kasera(a) 'nine' besides *sera 'one', but somewhat later, when *sera got lost and the composition of the reflex of *kasera(a) became unclear to speakers, it was discarded in some dialects/languages in favour of a new derivative from 'one': kamesa'a, kamesa, kaassa, or the like. Such derivatives may have been formed by analogy on the model of the word for 'eight', which had remained analysable in most dialects/languages (true, the Seko word for 'eight' may cast doubt on such an explanation). The other possibility, which apparently conforms better to Seko data, is that *kasera(a) and *kamesa(a) already existed side by side in different dialects at the PSS stage.

3.3.2 Non-numeral lexical items

Besides the numerals just examined, we can also find other lexical peculiarities which give evidence for the SSul language group, to wit: (a) characteristic lexical items (lexical roots or derived words), i.e. those not known to have cognates in other languages of the Sulawesi region and at the same time not widely spread in the AN family, and (b) lexical items which certainly are widely spread in the family, the languages of the Sulawesi region not necessarily excluded, but display characteristic ways of formal (phonetic or morphologic) or semantic development in the SSul languages (while the notion of 'characteristic' should be conceived in the same way as in connection with category (a)).

In reality, border-line cases occur where it is hard to say whether we are dealing with two genetically independent roots or with two different developments of the same root. Moreover, further study will certainly compel us to remove a few items from category (a) to category (b) -- given these items are not completely disproved, of course.

In view of these circumstances, I do not base my presentation of language material here on these categories. Instead, I will proceed from the probable evidential weight (eloquence) of items and present them in two loosely demarcated groups:

Group One -- items which appear, on the available evidence, to be limited to the SSul languages, even if more or less serious doubts may arise (and really arise in most cases);

Group Two -- items which certainly are not exclusive for the SSul languages but nevertheless seem worth considering here since they appear not to occur in non-SSul languages of the Sulawesi region (not counting the languages immediately contiguous with the area of the SSul group), and are not widely spread in the AN family. In this group I also include a few items that seem to be problematic in one way or another.

Each group contains both characteristic lexical items (category (a) above) and items displaying characteristic ways of development (category (b) above).

In placing the material within either group, I adhere to the following loose (and sometimes conflicting) rules:

(i) the general sequence should be 'from more eloquent to less eloquent', and

(ii) items of category (b) generally follow those of category (a) and those not readily classifiable in the given terms (a, b).

3.2.2.1 Group One

3.2.2.1.1 *salib/wuT 'fog' (M Diss *sali(b)u(C)): Mak (C) saliu›; Ko (F), Sel (C) salihu›; Bug saliwu›; Sad (V) salebu›; Duri (M) saleu›; Mmj (M), ArT (Kou) salehu›. Cf. the lexical items of the Sangiric languages (San hibu› 'mist, fog', etc.), Sneddon's PSan reconstruction *Ribu[›] (1984:82), and Snd halimun 'mist', OJav limut 'mist, darkness', Bal libut, ributˈ 'pitchdark', Lio lébu 'dense, e.g. of clouds', Kankanay livu:›u 'cloud'. However, no form containing the initial segment sa- is known outside the SSul group. Further, PSS *l does not correspond to PSan *R, neither does PPS *-b- or *-w- correspond to Snd or OJav -m-.

3.2.2.1.2 *cəkke or *cəŋkeX (*cəŋkeT ?) 'cold (feeling cold?)' (M Diss: *cɨkke(C?) (p.266), *(cs)ɨkke(C) (p.668)): Bug macəkke›, Duri, Pti (BLM) macakke›, Maiw (BLM), Pal (G) macakke 'cold'; Sad (V) (ma)sakke, masakka› 'cold', sakke› 'cold water'; Meh (G) masakke›, RA (G) masakka, RB (G) masaŋkaŋ 'cold'; Mak (C) sakke 'feeling of cold, shivering'; Mak (Selayar) (H) <sakke> 'cold (persons)'. Numerous unclear correspondences and doubtless non-phonetic lines of development impede reconstruction. In my opinion, at some time, possibly at the PSS stage, there existed two forms side by side: *cəkke and *cəŋkeX (*cəŋkeT ?), the latter may have been an intensive form. Mak sakke can be derived from *cəkke assuming the sporadic (but not rare) change of *c > s. No cognates are known outside the SSul group.

3.2.2.1.3 *i<r>baX (iqbaX ?) 'to slice, cut' (M Diss: *i-(rl)-baC (p.289), *i-r?-bak (p.708) : Ko (F) -e›ba› 'to slice', Mak (C) e›ba› 'slice'; Mak (C) -e›ba›, Bug -ebba› 'cut into, make an incision'; Sad (V) i›ba›, i›baran 'cut into pieces (a fruit)', i›bak 'to open using force'. Mills relates this PSS etymon to Dempwolff's *bak 'split' (Mills 1975b:708). However, final -k would have been preserved in PSS, and Sad i›ba› and i›baran testify against this final (Sad i›bak -- if cognate -- may be an intensive form). On the evidence presented above, it is difficult to specify *-X. It could hardly be *-q since no clear cases of final -q are known from PSS. Hence the cognacy with Tag hi:wa› 'slice' and its cognates in various other languages of the Philippines seems unlikely. Even if the PSS etymon under consideration should prove to be ultimately cognate with those words, in any case we can claim a specific way of development in PSS, most probably the infixation.

3.2.2.1.4 *luqluR 'to flow' or 'fluid' (M Diss, p.773: "*loqlo(ɣ) or *luqlu(ɣ)"): Mak (C) -lo›loro› 'to flow', Ko (Kajang dialect; DK) < aqloloroq > 'to flow' (in the given text, used for palm juice); Sad (V) lollo› 'dissolved (e.g. sugar); soft (overripe fruit)';[11] Duri (M) lolloh 'to melt'. I agree with Mills' opinion (1975b:773) that the present item reflects reduplication of the monosyllabic root *luR. This reduplication is so far not known beyond the SSul languages.

3.2.2.1.5 *leŋuR (< lə<R>ŋuR ?) 'to break, get out of joint' (M Diss *lengo(ɣ)): Mak (C) le›ŋoro› 'out of joint'; Bug (I) leŋo› 'to break (patah, pecah)'; Sad (V) leŋo› 'to hit heavily so that bones hurt'. Cf. Mong (D) loŋu 'breken door iets heen en weer te buigen', lolоŋu'on 'joint'. The cognacy seems somewhat dubious in view of the final glottal stop in the Mong rootword (-g would be expected). If, however, Mong loŋu' should prove really cognate, we face the task of explaining the vowel e appearing (in place of schwa) in the SSul words. In some cases the schwa adjacent to *R became e in PSS. Hence we might think of the pre-PSS infix -R-, whose other vestige might be found in the glottal stop present in the Mak cognate. If so, however, we must admit that Bug and Sad have not preserved any segments traceable to the putative infix.

3.2.2.1.6 *lampuR 'heap, esp. of harvested rice' (M Diss *lampo(γ); cf. Mills 1981:68): Mak (V) lamporo› 'untidy heap'; Bug lappo› 'heap, esp. of rice bundles'; Sad (V) lampo› meaning dialectally 'heap of rice bundles' or 'certain palm-leaf bag for offered rice'; Duri (M Diss, unglossed) lampoh (-r-). Possibly relatable (but not always without any phonetic problems arising) to lexical items referring to something numerous, excessive, which are found in various AN languages (see Mills 1981),[12] but the meaning 'heap of harvested rice bundles', which the SSul material suggests, is not known to occur outside this language group.

3.2.2.1.7 *timu 'mouth' (M Diss id.): Sel (?) (C)[13] timu 'mouth'; Mak (C) timuŋaŋ 'doorway'; Camp (G) timu(ŋ) 'mouth', timun 'lip'; "To-sada" (A) < timoe > 'mouth'; Duri (M), Pti (BLM) timu 'mouth'; Enr (G) timu 'lip'. Superficially more or less similar words are found in various languages of Eastern Indonesia, the Philippines and even Formosa, but comparison with them usually raises problems which either impede cognate decision or reveal the sufficient idiosyncrasy of the PSS etymon under consideration. Here a few examples are given. Buli smo 'mouth, teeth' would be relatable on the assumption that i was unaccented in the common proto-form, that is, it had the shape of timŭ or timúqu, but I do not know any South Sulawesian evidence supporting such an assumption.[14] Both phonetic and semantic problems arise when we try to compare PSS *timu with Ngadha temo 'thrust (shove) into one's mouth' or with Ttb -timu› 'strike one's lips, or mouth, against something'. (Regarding the phonetic difficulties: the failure of the SSul languages to reflect the segment reflected by Ttb as final glottal stop seems strange. One might guess that the reconstructed PSS form *timu had been assimilated from somewhat earlier *timo which may have derived from *timuq; the Ttb word, too, seems to be derivable from the last-named proto-form. The expected Ngadha reflex of PAN or PMP *timuq would have been **timu; here we might think of the sporadic fluctuation of high and mid vowels. However, I am aware of no internal evidence for putative early PSS *timo, and this conjecture can hardly be convincing. Even if it is correct, we would be left with the semantic problem.) At the same time, PSS *timu is not readily derivable from the etymon reconstructed by Blust (1980) as *timij (PAN) or *timid (PMP), meaning 'chin, jaw', nor is it relatable to PMin *semud, with reflexes referring to 'snout, muzzle' or 'mouth' (Sneddon 1978).

3.2.2.1.8 *asan/ŋ 'all' (M Diss *asang): Mak (C) (ŋ)aseŋ; Ko (G), Sel (G) yaŋase'; Sad (V) nasaŋ, ŋasaŋ; Duri (M) ianasaŋ, solanasaŋ; RA (G) yaŋasaŋ; RB (G) ŋasan; Pal (G) iyaŋasaŋ; Kal (G) iasan; ArT (G) diya'asaŋ; Rnt (G) iya'asan; Bmb (G) iya'asam; Mdr (Mu) nasaŋ; Mmj dialects (G) yanasaŋ, (d)iyanasaŋ; Seko dialects (G, M) nasaŋ, nasanna. Contiguous non-SSul languages: Lem (G) iasana 'all'; cf. Lay (H) < anoena ngase > 'their', Wotu (G) iyamassa 'all'. The final nasal is open to dispute. One might advance some arguments either for -n or for -ŋ. For -n: the RB, Kal and Rnt reflexes; the fact that in Mak the PSS final segment *-an sporadically becomes -eŋ (e.g. Mak areŋ 'name' from PSS *azan), while no clear cases of such development of PSS *-aŋ are known from that language. For -ŋ: PSS *-n should have been maintained intact in Sad, RA and Duri. In fact, these arguments are all weakened due to the fact that in SSul languages the word 'all' is frequently bound with pronominal morphemes following it. A few examples (with reconstructed morphemes): *asán-na 'all of them' or 'all of ...'; *...-ásaŋ-kan '...(e.g. give) everything to us'; *...-asán-ta 'because (or: when) ... (e.g. he gave) everything to us'. Such compounds give ample opportunities for analogical changes in either direction.

The SSul lexical item under consideration may have been derived from a word meaning 'one', most probably from *asa (this etymology is sounder if the debatable final nasal is assumed to have been *-n). Cf. Tonsawang pahasa, Lolak (S) ko'insa, Maloh kausan, Kalinga losan 'all'. However, the form of the basic numeral 'one' -- asa, not * sa, *isa or *usa -- coupled with the absence of any prefixal or infixal elements from the word 'all', would apparently be distinctive of the SSul (and contiguous) languages.

If we should favour the form *asaŋ, then questions arise regarding Aneityum asŋa 'all' (already compared with Mak ŋaseŋ by H.Kern, 1916). What is known about Aneityum sound correspondences does not preclude this relation. However, in my opinion, the Aneityum word must be first investigated within the framework of Oceanic languages, and there we note Fij kaceŋa 'all', which is traced to PEF *kece 'all' (Wurm, Wilson 1975: 4; -ŋa seems to be of suffixal origin). In view of a number of phonetic discrepancies, the last-named PEF etymon can hardly be related to PSS *asan/ŋ. · Therefore I also question the cognacy of Aneityum asŋa with the PSS item under discussion.

3.2.2.1.9 *manan 'all (together)' (M Diss *manAn, Mills 1981 *mane(mn)): Mak (C) manaŋ 'together'; Bug -manəŋ, Camp (G) yamanaŋ, Sad (V) manan, End (BLM), Pti (BLM) iamanan, Maiw (BLM) iamana , Pta (G) yamanaŋ, 'all'. Comparison with semantically corresponding words of various Philippine languages (Kalamansig-Cotabato Manobo ʾanan, Bisaya languages, Mamanwa and Aborlan Tagbanwa tanan, Yogad ʾatanan, Casiguran Dumagat ʾətanan 'all') gives me grounds to reconstruct the PSS form as *manan. If I am right here, the cognacy with To mano 'myriad, ten thousand' (cf. Mills 1975b: 780, 1981:74) is to be considered unlikely. For the distinctive feature of the SSul languages, we can take the initial m- (of infixal origin?). Bima mena 'together', however, may challenge this suggestion. Andio manaŋ 'all' (given in BBS as the second synonym) is undoubtedly a Bug loan.

3.2.2.1.10 *ləqbir/R 'magnificent' or 'to honour' (M Diss ɨbbi(ɣ)): Mak (C) laʾbiriʾ 'magnificent, noble, honoured'; Ko (F) laʾbiri 'to respect, esteem'; Bug ləbbiʾ (-r-) 'magnificent'; Sad (V) laʾbiʾ 'better, best; noble', laʾbiran 'prefer'; Duri (M) pakalaʾbih 'to honour'; Mdr (Mu) (ma)laʾbiʾ 'better, of superior quality'. Perhaps related to PSS * ləqbi 'more' (I agree with Mills). The relation with Bar -labiti (melabiti 'consider oneself as superior', ma(ŋa)labiti 'to swindle, despise') is not clear.

3.2.2.1.11 *b/wəqdin (< *b/wa<r>lin ?) 'can, may' (M Diss *(b)ɨddin): Mak (C) maʾrin; Bug wəddiŋ; Sad (V) maʾdin, waʾdin; Duri (BLM) <waaqdiŋ>; Pti (BLM) <wadiŋ>; Maiw (BLM), End (BLM) waʾdiŋ; Seko (M) madiŋ. Perhaps traceable to a root widely reflected in the Sulawesi region and the Philippines: e.g. Bada (me)waliʾ 'become, take place', Uma wali 'become', Wol membali 'become, come about', Mong (mo)baliʾ 'can, become, come about', Iloko and some other languages of Luzon (ma)balin 'can', etc. If so, we may hypothesize the development through the infixation of *-r- (earlier *-R-), which would constitute the characteristic feature of the SSul group.

3.2.2.12 *ínan (< *ina + -an ?) '(sleeping) place' (M Diss *nginan): Mak (C) -énaŋ 'to lie (esp.: in bed)'; Sad (V) ínan 'place; (dial.) bedroom', aŋénan, laŋénan (both dial.) 'bedroom', naŋénan (dial.) 'children's bedroom'; Duri (dialects) ŋenan mamma' 'tempat tidur' (BLM), ŋenan katindoan 'bed(room)' (M); "Toraja Binuwang" (H) <angenan> 'bedroom'. Possibly we have to do with a derivate with the locative suffix -an. If this is the case, one may suspect it of having arisen as a result of borrowing from some languages where ina (or inaʾ, or inah) had been the regular reflex of the widespread AN root *qinəp; cf., e.g., Bar (Puʾu-Mboto dialect) meina 'to sleep'. It is unlikely that -p- had not left any trace in the source language while -n had been maintained intact there. Consequently, what was borrowed may have been either (1) a suffixal derivative where -p- had been changed into a consonant -- most probably glottal stop or h -- which was subsequently omitted in the borrowing language, or (2) the rootword **ina (**inaʾ, **inah). In the second case, the suffix -an must have been added in the borrowing language. In both cases the reflexes in the modern SSul languages should expectedly be

stressed on the ultima. The penult stress observable in them and, on the available evidence, reconstructible for PSS *inan, leaves us with two possibilities: (1) this reconstruct has gone through a specific line of development; (2) our idea of suffixal derivation was wrong, and PSS *inan is a root morpheme. Concerning the second alternative: the putative root in question may be exclusively South Sulawesian, but this supposition is hard to prove because many languages (of the Sulawesi region, the Moluccas etc.) have words which are ambiguous for *ínan and *qinəp.

3.2.2.1.13 *tedoŋ 'water-buffalo' (M Diss *tedo(N)): Mak (C), Ko (dial.; DK), Bug, Sad (V), Mas (M), Mdr (Mu) tedoŋ. Possible cognates are found in some of the Philippine languages: cf. Mar tidoŋ 'rude', Ceb tiru 'a supernatural being', Tag tíro 'one who makes himself look like a powerful man and bullying others'. However, the meaning 'water-buffalo' (euphemistic?) is not known to occur in languages other than the SSul ones.

3.2.2.1.14 *libuka/ən 'island' (M Diss *libu(k)+an): Mak (C) liukaŋ 'island'; Ko (F) likukaŋ 'islet'; "Salayar" (H) <lihoekang>, Bug libukəŋ, Sad (V) lebukan, "Majene, Mandar" (H) <liboekang>, SekoP (M) libukaŋ 'island'. Contiguous languages: Lay (H) <lehoekang> 'island' (synonym). The SSul words are suffixal derivatives from -libuk, and this morpheme as well as the use of its derivatives for the notion of 'island' seem to go back to the PSS stage. Outside the SSul group -- not counting the Lay language -- the root morpheme -libuk is not known. However, it is a variant of the much more widely spread morpheme *libut 'surround, encircle' (Blust 1980:106). Various non-SSul languages of the Sulawesi region designate the notion of 'island' by means of lexicalized suffixal derivatives of *libut which retain -t- (Wol liwuto, Kul lewuto; Uma lewuto, Mong libuton, etc). So the SSul group is in fact characterized by -k- in place of the -t- of non-SSul languages.

3.2.2.1.15 *tamparan (M Diss id., see p.850): Mak (C), Ko (F), Sel (G) tamparaŋ 'sea' (Mak tamparal-la'ba, lit. freshwater-sea, 'lake'); Bug tapparəŋ, Sad (V) tapparan, Mdr (H) <tapparang> 'lake'; Duri tapparan (M, unglossed); (cognate?) Mmj tampaŋ (M, unglossed). This etymon goes back to the suffixal derivative tampar + -an. The base tampar obviously had the meaning of 'shore' or 'edge' (cf. Buli (m)topo 'shore', Sumbawa tampar 'shore (from the water's side)', etc., and the PPN reconstruction *tapa 'edge'). PSS *tamparan had the meaning of 'sea' or 'large lake', or possibly was used toponymically to designate the deeply incised bay, or even strait, the last renmants of which are Lake Tempe and the near-lying lakes. Cf. Sas tamparan 'surf'. The use of the suffixal derivative under discussion for some aquatic element of landscape (sea, lake, bay, channel, etc.) seems to be a SSul innovation.

3.2.2.1.16 *bintuen (*bintuin ?) 'star' (M Diss *bintuin): Mak (C), Ko (G), Sel (G) bintoeŋ; Bug wittoeŋ, wettoiŋ; Sad (V) bintoen; Duri (BLM) bentuin; RA (G), RB (G) bentoen; Mam (G) bintoen; Kal (G) bittoin; Pal (G) bintueŋ; Mbi (G), Mdr (Mu) bittoeŋ; Mmj dialects (G) bintoeŋ, bittoeŋ; Pta (G), SekoT (G) bittoeŋ (all 'star'). Contiguous languages: Lay (G) bintoeŋ 'star'; Lem (G) bintoi 'moon'. The source is PAN *biCuqən. The SSul languages (as well as Lay and Lem) reflect two irregular changes: (1) *ə > e (or > i [15]), and (2) insertion of the medial nasal accretion. Change (1) is areal; its area also includes the KP (incl. Badaic) languages (e.g. Bar, Bada betu'e, Kul betue 'star') and the Tambee dialect of the Mori language (betue id.). The medial nasal accretion in the given root occurs in a few languages of Formosa (e.g. Bunun bintohan 'star'), in Simalur, etc. For the two changes reflected jointly, however, I cannot cite any trustworthy evidence from other languages than the SSul ones and two languages contiguous with them.[16]

3.2.2.2. Group Two

3.2.2.2.1 *bela 'far' (M Dis *bela): Mak (C) bella; Bug mabela (dial., poet. mawela);[17] Sad (V), Mam (G), RA(G), RB(G), Pal (G), Bmb (G), Meh (G) mambela; Duri (BLM), Maiw (BLM), Pti (G), Pta (G), Mdr (M) mabela;[18] End (BLM) bela; Enr (G) mambela; Kal (G) mambellʌ; "Mamuju, Toraja, Galumpang" (H) <mabelah>, "Toraja Balanipa" (H) <mambela>. Cognacy with Rampi mo'ela 'far' and Bada (ritual language) moela id. is doubtful due to the failure of these items to reflect the initial *b-, which is convincingly reflected by the SSul languages. PSS *bela may have arisen as a result of metathesis of some earlier *labiC into *biCla. This way of development can account for the geminate ll visible in the Mak and Kal reflexes. True, no semantically associable root *labiC has been reconstructed to my knowledge. However, Blust (1973) suggests the PAN reconstruction *alawid/D/j 'far'; among its reflexes he gives Mny lawit 'far' and Mlg lavitra id., which would both fit still better as reflexes of *labiC.

Sumba mbila 'vast, extensive' may challenge the exclusively South Sulawesian nature of putative *biCla.

3.2.2.2.2 *ləntuk 'arrive' (M Diss *lɨ(n)tu(C)): Ko (DK) lantuqmo 'sampailah'; Bug ləttu', lattu' (I: <lettuk>, <lattuk>) 'come'; Sad (V) lattu' 'directly come to...' (probably borrowed); Duri (BLM) dipalantukan 'disampaikan'. See Mills' discussion of this etymon (1975b:760). I would like to add that evidence for final *-k comes not only from Duri, but also from the Bug language of Ide Said's dictionary. If the etymon was *ləntuk, then Sad lattu' (ending in glottal stop instead of the expected final -k) appears to be a loan (indeed Veen, 1940, supplies this word with the remark: "From Bug lĕttu' "). At the same time, I remove Mdr pi/lattowaŋ 'to turn up, show up' given in Mills' Dissertation within the cognate set upon which his reconstruction *lɨ(n)tu(C) is based, from this connection. The Mdr word just cited is a derivate of latto 'nampak, kelihatan', which originates from the early pre-PSS etymon *ləntaw 'float; still water' (for its reflexes, see Mills 1981:74).

Extra-South Sulawesian cognates of PSS *ləntuk apparently occur but rarely. Cf. Simalur ləntug and Sichule löntu', both meaning 'arrive'. Bar (me)letu 'to attack (several people)' may be a cognate reflecting the infixed form *lə<R>tuk (the R-infix usually expresses frequentativity or plurality); in this case, the development must have been: -ə/aR- > -ə/ay- > -e-.

3.2.2.2.3 *kale (< *kalawe ?) 'body, -self' (M Diss: *kalawe and/or *kale): Mak (C), Ko (DK), Sad (V), Mas (M) kale, Bug ale '-self'; Camp (G) ale- in aleu 'I', alemu 'you', etc.; Mdr (Mu) alabe 'sendiri, diri, badan, tubuh'; SekoP kalae (M; unglossed). Mills (1975b:726) explains, rather hesitantly, *kale from *kalawe. But the form kale is also present in Maloh ('-self, body', according to Ad).[19] It seems probable that the forms *kale and *kalawe, whatever their mutual relation, existed side by side in the PSS dialect chain.

3.2.2.2.4 *taya, *tayan (perhaps two grammatical forms) 'to wait' (not treated by Mills): Mak (C) -tayaŋ. "Makasar" (Salayar) (H) <attadjang>, Bug -tajəŋ 'to wait'; Sad (V) taan, (dial.) taya 'to watch over'; "Kada Toraja" (H) <tajanni> 'to wait'; Mdr (Mu) taja 'sedia, siap', tajai 'tunggu'. Two explanations seem available:

(1) PSS *taya and *tayan are suffixal derivatives of a certain AN root *(...)tay (cf. Dempwolff's reconstruction *hantaj, Dyen's *q/hint/Tay, Zorc's PPh *baNtay and *aNtay). It cannot be determined whether the base of the PSS derivates under consideration was disyllabic (like the reconstructed items given above in parentheses; if this was the case, the first syllable must have been lost subsequently, e.g. through reanalysis), or it was the monosyllabic *tay that served as the base.

(2) the PSS derivatives under discussion originate from the early AN

root reconstructed by Tsuchida (1976:232) as *t/TaRaH$_2$ 'to wait'. This assumption presupposes an early borrowing. The modern languages of Central Sulawesi have amalgamated the *-y- which arose from *-R- with the preceding vowel. The borrowing should have taken place in the intermediate stage of the development: -aR- > -ay- > -e-.

Wol ntaa 'to wait' lends itself to either of these explanations, not requiring any borrowing hypotheses.

3.2.2.2.5 *tanete 'hill' or 'slope' (M Diss *tane(m?)te): Mak (C) tanete '(gentle) slope'; Ko (F) tanete 'pasture, meadow'; Bug, Sad (V) tanete 'hill'; Mam (G), Kal (G), Bmb (G), Meh (G), Mmj dialects (G), SekoT (G) tanete 'mountain'. Assuming that this PSS lexical item has arisen from earlier *tani/etay, we can relate it to the words of various languages of the Buru - Ceram area for which Stresemann reconstructed Proto-Ambon *tĕnita 'mountain' (e.g. to Buru tanita 'hill'). Mlg tanety 'hill; land' (for which, true, another etymology has been presented) might also be cognate with the PSS etymon under consideration. Further, cf. Roti lètek (dial. lète) 'mountain' and its apparent cognates in several other languages of Eastern Indonesia which mean 'above' (Tetum lèten, Masiwang leta, etc.). In the Sulawesi region, we should pay attention to San (Sasahara) tadete 'mountain; high sea' (the correspondence SSul n: San d is unexpected, but Sasahara's ritual character should be taken into account). On the contrary, phonetics do not preclude Bar tonete 'k.o. shelf fastened to the house wall above' from entering the same cognate set, but semantic problems arise. One might try to divide this Bar word into to + nete and explain it as 'what is above' (the relativum to is found in a number of KP languages). However, it is difficult to expand this explanation to the SSul words because no relativum ta is known. To is used in this function in archaic Bug, and probably it can be reconstructed for PSS too. Should PSS *tanete really originate from the compound to + nete, it would be characterized by an irregular vowel change.

3.2.2.2.6 *lambi/əR 'long (object)' (M Diss *lambiɣ; Mills 1981: PSS *lamb(eĕ)R): Ko (SK), Sel (G) lambere 'long'; Mak (C) lambere› 'thick and long'; Bug, Pti (BLM), Pta (G) malampe› 'long'; Sad (V), Maiw (BLM), Kal (G), Sum (G), Tap (G), Bot (G) malambe› 'long'; Mdr (Mu) malamber 'long'. On the assumption that the final syllable vowel was i, one should compare the PSS etymon under discussion with Sas (dial.), Bal lambih 'too long', Bar (ritual) mayambi 'long (time)', and probably with Bikol halawig 'long (time)'; Iban lambiŋ 'pendulous' and Mong molambiŋ 'too long' might also be of some interest in view of the seeming tendency of *R to interchange with nasals in various positions. The alternative reconstruction of schwa would push West Bisayan (Zorc) *labəg 'long (object)' into the foreground. Cf. Mills 1981: Note 66.

3.2.2.2.7 *lame 'tuber plant' (M Diss id.): Mak (C) Ko (F), Bug lame 'tuber plant'; Sad (V) lame 'yam'; Mdr (LM) lame 'ubi'; Mdr (Mu) lame aju 'ubi kayu'; "Tomadio-Campalagian" (H) <lame> 'species of tuber'. This SSul word seems fairly isolated in the Sulawesi region. In any case, I entirely lack any comparable, even hypothetically comparable, material from the KP (incl. Badaic) and the Bungku-Laki (sensu Esser) languages. As for Wol lamelame 'potato' and Muna (H) <lame dawa> 'sweet potato', these are borrowings from Bug or Mak (where semantically corresponding compounds lame-lame and lame-jawa are well known).

Phonetically similar words are found in several AN languages outside the region.

In particular, Tag lamí (variant: namí) 'yam sp., according to J.V. Panganiban Dioscorea hispida' may be cognate. If this is the case, the final -e of the SSul word must have arisen from the earlier (pre-PSS) segment *-iq, not from *-ay. This disambiguation, however, is hardly decisive if lame of the SSul languages is compared with Leti, Roma nama 'tuber', or with Roti (dial.) nama 'to twine' (said about plants), or with

Nias (ritual) lama 'tuber'. Whatever the source of the SSul final -e, the correspondence with Leti, Roti or Nias -a appears to be irregular.

Mota nam, Mosina nom and Tangoa ram 'yam' should be first compared with ndamu id. found in a number of languages of Vanuatu, e.g. in Raga. The cognacy of the last-named word with SSul lame seems questionable (in view of final -u and initial nd-). Nevertheless the very possibility of this cognacy can hardly be flatly denied. The great relevance of tuber plants for the early inhabitants of the Malay Archipelago and Oceania should be borne in mind. One may think that this situation favoured irregular modification of terms like 'yam'.

3.2.2.2.8 *ləqte 'thunder' or 'lightning' (not treated by Mills): Ko (F) latte 'lightning strike'; Sel (MW, H) latte 'thunder'; Bug (arch. and dial.) i tte 'thunder' or 'lightning and thunder simultaneously'; Sad (poet.; V) la'te 'lightning'. Contiguous languages: Lay (H) <latte> 'thunder'. This is a derivative of the WMP root *lətiq 'thunder and lightning together' (Blust 1984:76). Cognates (part of them reflecting irregular changes) are fairly numerous in the languages of Central and Southern Philippines (see Zorc 1977:236); aside from them, Blust notes a cognate in Long Terawan. Cf. also Simalur (K) lumāti' 'schwach leuchten, glimmern, glanzen, glitzern'.

3.2.2.2.9 *tinduR 'follow' or 'pursue' (M Diss *tindo(r)): Mak (C) -tindoro' 'go following somebody', -tinruru 'follow'; Mak (Bantaeng dialect; H) <tindóró' asoe> 'hunting with dogs'; Bug -tinro' (-s-) 'follow'; Sad (M) tindo' 'follow in a regular order, in line'[20]; Mdr (Mu) -tindor: sitindor 'beriringan' and other derivatives. The root *tinduR is undoubtedly a doublet of *tunduR, for which we should compare the PPhil reconstruction (Zorc) *t+unduR 'follow', PMin (S) *tunduh id., Bar túndugi 'war ally', Kul, Li, Le natundu 'obedient', Bng -tundu 'pursue', Sumba tundu 'follow'. Reflexes of *tinduR have so far not been noted outside the SSul group, not counting Mang (dial.; Ve) tindu 'mengepung, keroyok, menangkap bersama-sama (kerbau dsb)'. At the same time, *tunduR is represented in Sad (V: tundo' 'ordenen en in rijen plaatsen').

3.2.2.2.10 *tindo 'to sleep' (M Diss id.): Mak (C), Ko (G), Sel (G) tinro, Bug matinro, Camp (G) matindro 'to sleep'; Sad (V) tindo 'dream', matindo 'to lie'; RB (G), Pal (G), End (G) matindo 'to sleep'; Enr (BLM) mattindo-tindo 'to lie'; Maiw (BLM), Pti (BLM), Pta (G) matindo 'to sleep'; Mbi (G) matindo, Bmb (G) metindo 'to lie down'; Rnt (G), ArT (G) metindo 'to sleep'; Meh (G) metindo 'to sleep; to lie down'; Mdr (Mu), Mmj dialects (G) matindo 'to sleep'. Outside the SSul group, regular-shaped cognates are found in Maloh and Sas (both tindo' 'to sleep'). Iban tinduk id. is probably cognate, but the final -k seems strange (glottal stop would have been expected). The SSul, Sas and Maloh items may all derive from PAN *tiDuR (essentially as Mills supposes, 1975b:859), but this etymology presupposes two irregular phonetic changes: (1) final *-R was lost in the SSul languages (which might be considered a distinctive feature of this group), while it changed into glottal stop in Maloh and Sas (also an irregular development), and (2) the medial nasal accretion was inserted. But this etymology is far from being doubtless. Alternatively we may compare the SSul, Sas and Maloh items with Karo tunduh 'to sleep', Gayo and Sund tunduh 'be drowsy', which appear to imply earlier *tund/Duq. In this case, only one irregular change must be posited (i - u fluctuation; cf. the item *tinduR). Finally, it is also possible that the hypothetical *tund/Duq is etymologically related to *tiDuR in some way.

3.2.2.2.11 *kulle 'be able' (M Diss: *ule (? ka/ule)): Mak (C) (ak)kulle 'be able'; Ko (DK) takkulle 'be unable'; "Salayar" (H) <koelle> 'be able'; Bug ulle, makkulle (rarely mulle) 'be able'; Sad (V), Duri (M) kulle 'be able'; Mdr (Mu) ulle 'kuat, sanggup', ma'ulle 'sanggup'. Outside the SSul

group: Maloh (Ad) kulle 'be able, be healthy'. PSS *kulle certainly originates in the widely spread AN root *uliq, but is characterized by two features which are also visible in Maloh: initial k- and geminate l. The latter feature has also been noted for the Mad cognate olle 'be able'. This gemination, however, is not very eloquent because it may have been lost in numerous non-SSul languages which have no geminates. Initial k- (which may be the vestige of a merged prefix) seems to be fairly characteristic of the SSul languages. Beyond this group, it has been noted in several Western KP languages: Uma (Ma) kule', Kul, Li -pakule, Le nompakule 'be able'. The finals -e' and -e occurring respectively, in place of the expected -i' and -i, qualify these words as borrowed. But from where could they come? The final glottal stop of the Uma word testifies against the SSul source, but can be taken as a piece of evidence for the Badaic one. I am not aware of any reflexes of *uliq in Badaic, but it should be borne in mind that no dictionary of those languages exists, and moreover the borrowing may have taken place at some early date.

The problem of Maloh kulie still remains.

3.2.2.2.12 *la e 'to swim' (M Diss *langi): Mak (C), Ko (G), Sel (G) -laŋe, Bug -naŋe, Maiw (BLM) annaŋe, Kal (G), Tap (G) molaŋe, PUS (M) malaŋe. The PAN root was *laŋuy, which should have given *laŋi in PSS. In view of the material just presented we may claim the irregular outcome *laŋe in PSS (true, *laŋi, reflected by a few Mmj dialects, may also have existed in the PSS dialect chain). Interestingly, this root shows an unexpected final -e in some widely separated non-SSul languages of Formosa (Maga dialect of Rukai), Sumatra (Toba), Nusa Tenggara (Roti, Solor, the western dialects of Manggarai), etc. However, this (pre-)PSS development seems to be virtually unique in the Sulawesi region where the only non-SSul example found so far is Menui (BBS) mumaŋe 'to swim'.

4 QUANTITATIVE EVIDENCE

In my study of the problem of defining the SSul language group, I resorted to three statistical methods: lexicostatistics, etymostatistics and functor statistics.

4.1 Lexicostatistical evidence

The lexicostatistical analysis was carried out on the basis of a 100-word list. The list used is the standard Swadesh list from which four items, namely "bark", "round", "seed" and "fat", had been deleted and replaced by "far (away)" (BI jauh), "sea" (BI laut, not lautan), "wind" (BI angin) and "to take" (BI mengambil, not menerima or mendapat). I compared a number of languages and dialects of the southern and central parts of the Sulawesi region and also Mong and Maloh. My lists were compiled from dictionaries and published wordlists including those given in Pelenkahu et al. 1972, Sjahruddin Kaseng 1978, Barr, Barr and Salombe 1979, and the Grimeses' monograph (1987). For Uma, I made use of the data sent by Michael Martens (in personal correspondence).

In the analysis of the wordlists, I tried to use the etymological (comparative) method as far as possible (cf. Grimes and Grimes, 1987). Both items unclear for the "cognate -- non-cognate" dichotomy and doubtless borrowings were eliminated, while the base for percentage computing was correspondingly reduced for every pair of lists.

Table 1 gives the most interesting results of my lexicostatistical comparison. (The scores of a number of poorly studied SSul dialects, namely the Mmj, PUS and Mas ones, and also those of Camp and Pal, are not shown. The table does not include such non-SSul languages as Lem, Rampi, Wolio, Bng, Mong and Maloh.)

Table 1: Lexicostatistical comparisons

Mak (stand.)														
74	Ko													
69	79	Sel	Average Makasaric											
45	53	53	50	Bug										
42	44	49	45	60	Sad									
37	43	45	42	54	61	Mdr								
26	31	30	29	39	46	46	SekoT							
32	32	36	33	45	51	49	52	Bada						
29	30	33	31	39	44	44	48	72	Napu					
28	30	35	31	41	44	43	40	53	60	Kul				
32	30	36	33	40	40	41	44	56	55	71	Uma			
28	26	32	29	43	44	38	40	55	50	57	59	Bar		
24	26	33	28	38	36	35	34	38	36	39	39	48	Mori Atas	
33	38	41	37	47	47	45	40	42	39	44	46	47	43	Wotu
41	47	54	47	42	51	50	38	44	45	46	50	45	38	57	Lay (Barang-Barang)

4.2 Etymostatistical evidence

Etymostatistics is a new method developed by Starostin (1989). In essence, this is a count of the etyma whose reflexes are met with in texts (in principle, it does not matter what kind of texts they are, even dictionary pages can be used). Comparing language A (which may be called 'text language') with language B, we analyze a text portion -- more precisely, a theoretically unlimited number of text portions -- in language A. Omitting loanwords, archaic words and proper names, we write out the root morphemes.[21] From a single text portion, no morpheme is written out more than once (if a morpheme is met with for the second, third, and so on, time it is simply omitted). The procedure of writing out morphemes is stopped when the obtained list reaches a certain predetermined length (in principle, it is conventional, but I, like Starostin (1989), stopped at 100 items). The list obtained in this way is then compared with the vocabulary of language B ('vocabulary language'). Language B is looked at in the entirety of its lexicon. At least one adequate dictionary is needed for this purpose. This is why I was forced to limit myself to four 'vocabulary languages' only.

The etymostatistical method evidently has the advantage that any pair of languages can be tested several times, which permits us to apply methods of statistics, in particular, to compute average percentages, etc. The disadvantages of the etymostatistical method are its labour-consuming nature (if compared with lexicostatistics), the need for adequate vocabularies and the requirement that the historical development of the languages to which it is applied should be understood at least in the most general outline.

The results of my etymostatistical comparison are given in Table 2, which shows the arithmetical means of counts (for any 'text language -- vocabulary language' pair, 2 - 5 counts were carried out).

Table 2: Etymostatistical comparison

'Text language'	'Vocabulary language'			
	Mak	Bug	Sad	Bar
Mak		67	60	45
Bug	70		77	51
Sad	60	68		53
Mdr	64	77	77	53
Bada		53	62	77

4.3 Functor statistical evidence

The third quantitative method employed is functor statistics. For this method, see Zorc 1977:185ff. I used the 100-item list of the following composition: personal pronouns and pronominal clitics (4 sets, totalling 28 items); deictic pronouns (3 items); deictic locational words (4 items); non-deictic locational words (7 items); interrogatives (6 items); numerals

and quantifiers (16 items); negatives (3 items); conjunctions (4 items); affixes (8 items); varia (11 items). It was possible to apply the functor statistics only to a few better-known languages for which enough grammatical information is available.

The functor list was scored, in general, like the lexicostatistical one, but with the following adjustment: if one of the two languages compared lacks an individual item (and there are reasons to believe that this lack is imaginary, due to our insufficient knowledge), then the given pair is scored minus; but if one of the languages lacks a whole set it is eliminated and the base for computing percentages is correspondingly reduced (as with Bar, with its three sets of personal pronominal words/morphemes, with any other language which displays four sets of this kind). Table 3 gives the results of my functor statistical comparison.

Table 3: Functor statistical comparison

Mak (stand.)					
65	Bug				
60	66	Sad			
55	65	66	Mdr		
33	42	51	39	Bada	
29	38	41	40	53	Bar

5 COMPARING DIFFERENT KINDS OF EVIDENCE

The three quantitative methods employed in this study prove to differ in their degree of agreement with the qualitative evidence.

5.1 Functor statistics. As far as better-known SSul languages and Bada are concerned (unfortunately I had no data for such possibly critical languages as Seko, Lay, Wotu and Lem at my disposal), the configuration of the percentages obtained by functor statistics does not conflict with the qualitative evidence.

5.2 Etymostatistics. The etymostatistical percentages obtained clearly testify for the inclusion of Mak into the SSul language group. At the same time, the position of Bada remains somewhat vague. As was to be expected, the elements written out from Bada texts can be more frequently identified in the Bar vocabulary than in the vocabularies of Sad and Bug. Against expectation, however, such elements have proved more numerous in the Sad vocabulary than those coming from Mak texts. Unfortunately there exists no Bada dictionary which might allow us to carry out a cross-test.

5.3 Lexicostatistics. The disagreement between the configuration of lexicostatistical percentages and the qualitative evidence is quite striking. The main point is that the SSul languages located to the north

of the Bug area (Sad, Duri, Mdr etc.) score more highly with certain languages of Central Sulawesi which are so far considered as non-SSul than with the Makasaric languages. Schematically we might delineate two major groupings of languages. One of these groupings would embrace: i. the languages of the northern half of the SSul area (Sad, Duri, Mdr, Mmj, ArT, Seko, etc.); ii. the Badaic languages (Bada, Besoa, Napu, Lem); iii. Uma, Kul, Le (and peripherally also Bar); iv. Wotu; v. Lay. The second grouping would be made up of the Makasaric languages (Mak, Ko, Sel). At the same time, Bug and Camp may be looked upon as forming a link between the two major groupings.

This picture can be interpreted in the following way. Suppose there is a genetically determined group of languages -- let us call it SSul-2 -- that should include the overwhelming majority of the SSul languages with the exception of the Makasaric ones. The relatively high scores which Bug and Camp share with Makasaric can be explained by geographic proximity. If so, Bug and Camp can be included in SSul-2. However, the hypothetical SSul-2 group can hardly be simply equal to the result of the operation: SSul minus Makasaric. It should also include some non-SSul languages, at least the Badaic ones. The main argument would be that the Bada language, which has high percentages with the languages of the northern half of the SSul area, incl. Mdr, is not areally contiguous with any SSul language. The argument is also available for inclusion of a few non-Badaic western KP languages, at least Kul, and for inclusion of Lay. Wotu is contiguous with Bug at present, but does not score especially highly with it, e.g., the Wotu - Mdr percentage is not much lower than the Wotu - Bug one. This may be taken as an argument for including Wotu in SSul-2.

5.3.1 The question is now in order as to how the hypothetical SSul-2 group meets the qualitative evidence.

The qualitative evidence presented in the central part of this paper apparently contradicts the exclusion of Makasaric from the SSul group. Nevertheless we ought not lose sight of the geographic closeness. So what if the common features have been brought into existence by spread of isoglosses? If this is the case, however, there should also exist other features not extending to Makasaric.

When analysing this problem we must distinguish between the features attributable to SSul-2 (i.e. including Badaic and possibly some other KP languages, as well as Wotu and Lay), and those attributable to 'SSul minus Makasaric' (where the term SSul is taken in its traditional sense).

5.3.1.1 SSul-2. So far I have found only one piece of qualitative evidence which appears at first glance to testify in favour of the possible SSul-2 group. This is the lexical item reconstructible as *beluak '(head) hair' (in non-Makasaric SSul languages beluak, welua' and similar; Bada, Besoa (both Ma) welua'; Napu (Ma) welua; Lem (G) velua). But this item is hardly eloquent. The initial segment of *beluak is probably identical with the widely spread rootword bulu 'hair, feather', while the whole item may originate from a compound like *bulu ubak, literally 'hair of head' (cf. Loinan ubak 'head', and also 'a'ba' id. in several Mmj dialects, SekoT ba'ak, Woto baa id., etc.). The vocalism of the first syllable has gone through a certain development including such stages as the shift to schwa (expected in pre-accent syllables; hence, in particular, Lay (G) bɨlua, Wotu (G) baluwa 'hair, rambut'), and shift to e. The latter stage may have been brought about by infixation: -ə<R>- > -əy- > -e. In this case, the word beluak must have been borrowed from some Central Sulawesi (KP?) dialect into early SSul dialects.

5.3.1.2 SSul minus Makasaric. This variant is of less importance because it cannot give any clear-cut solution: even if Makasaric is excluded from the SSul language group, there remain languages (viz Mmj, ArT, Rnt, Seko) which score higher with Bada, and sometimes also with other non-SSul languages of Central Sulawesi, than with Bug (and sometimes also with Enr,

Mdr, and so on). Nevertheless some features widely observable in the non-Makasaric SSul languages seem to be worth comment. From them, I would like to note the following ones:

i. The way of forming the numeral 'four hundred': from *əqpat ratus, instead of *pata( )ratus, which would be expected in-view of PSS *patampulo 'forty' and *patansəqbu 'four thousand'. The compound *pata( )Ratus is reflected by various languages of the region, e.g. by Bar and Mori. The Makasaric languages have lost the root *ratus (at the PSS stage, from the earlier *Ratus).

ii. A few lexical roots, among them:
*bale 'fish' (attested in the Maros dialect of Mak, but there it may be a Bug loan);
*saliwa/ən 'outside' (also in Lem and Wotu);
*taru 'deaf' or 'deafened' (also in Wotu).

5.3.2 There is also another, cardinally different way to improve the agreement of the subgrouping scheme with lexicostatistical percentages: without trying to exclude Makasaric we may make an attempt to include Badaic. Such a scheme is apparently supported by the accentuation rules which unite the SSul (incl. Makasaric) languages with the Badaic ones. As far as known, these rules also extend to Rampi and Le, but such non-Badaic KP languages as Uma, Kul, Li and Bar follow sharply different rules. The rules of Wol and Mori are also different. Unfortunately I have no information on the accentuation patterns of Lay and Wotu.

Another point that might be looked upon as a piece of evidence in favour of the subgrouping scheme under consideration is the lowering of high vowels before the early final *-q. Some other facts of phonetic history, however, militate against this scheme (e.g. the development of final *-q and that of *R sharply differ in the SSul languages and the Badaic ones). It appears to be likely that the two positive features just mentioned, namely the SSul - Badaic accentuation pattern and the lowering of high vowels before *-q, are areal. It is worth noting that there really seem to exist some areal features essentially confined to Badaic and SSul. In particular, the pattern of forming compound numerals expressing multiples of ten, hundred etc. representable as "unit + (no trace of an earlier ligature!) ten, hundred etc. + -na"[22] demonstrates the gradual 'fading out' through the area: from the north (Seko, Napu) where it is used in all or almost all the numerals of the structure "unit + ten, hundred etc." to the south. In Ko (F) it is discernible only in the numerals 700, 800 and 900. Various intermediate stages can be found in Sad, Mdr and Bug.

5.4 So it is hard to escape the conclusion that although the lexicostatistic evidence is at variance with the notion of the SSul language group which is based on the qualitative evidence, nevertheless there is no sufficient reason to abandon this notion or to change it radically (e.g. into SSul-2). True, some correction of the boundaries of the SSul group in the future should come as no surprise, but such a correction would probably concern the position of some languages which are poorly known at present (like Seko, Mnj, ArT), and removal of the Makasaric languages (or part of them) from the SSul group or introduction of, say, the Badaic ones is hardly likely.

NOTES

My study into SSul languages would have been impossible without the assistance of my numerous colleagues who sent me their publications, preprints, and other literature, as well as unpublished materials collected by them, answered my queries in private letters or in personal communication, and so on. I am deeply obliged to Sander Adelaar, René van den Berg, Robert Blust, Michael Chlenov, Otto Christian Dahl, Timothy Friberg,

Victor King, Michael Martens, Roger Mills, Jacobus Noorduyn, James Sneddon, Wim Stokhof, Shigeru Tsuchida, Stephen Wurm, Masao Yamaguchi, and many other Austronesianists. All errors, of course, remain mine.

1 The Grimeses (1987) as well as Friberg and Laskowske (in this volume) assign the Lemolang language to the SSul Stock, but Esser viewed this language as belonging to the Torajan group (Noorduyn 1963). Lemolang is not shown on Esser's map (1938).

2 Only one of the features noted by Mills -- namely the first, concerning the numerals 'eight' and 'nine' -- can be treated as a SSul innovation. The other ("/a/ form of verbal 'conjugation' employing affixes to denote both the subject and the object of a verb") is an areal feature widely spread in the southern and central part of the Sulawesi region (it is also found in various languages of Nusa Tenggara, the Moluccas, the Barrier Islands, some parts of Oceania, etc.).

3 See 'Abbrevations'.

4 The categories of widespread innovation and selective innovation introduced by Zorc (1986:153) are, as a matter of fact, overlapping; the difference between them is made up by their frequency.

5 Note that we are discussing pre-PSS sources of final vowels. Some varieties of Bug (at least in Soppeng) have lost the final glottal stops which had arisen due to certain sound shifts subsequent to the PSS stage, but this recent loss may not have changed the cooccurrence range of individual words. Compare:
'fire': PSS *api, Early Bug *api; 'his fire': Bone Bug api-nna, Soppeng Bug api-nna; 'skin': PSS *kulit, Early Bug *uli'; 'his skin': Bone Bug uli›-na (∿ ulinna); Soppeng Bug uli-na.

6 In many AN languages consonant-initial stems take um as infix. Such derivates also occur in SSul language, e.g. Sad tumampa 'to forge, create' besides tampa id. (also 'form, mould').

7 The possibility that this Bug prefix goes back to um was first noted by Noorduyn (1955:14, note).

8 The translation is based on the example sentences found in Cense's dictionary (1979).

9 Here and further, when presenting data of SSul languages (except Bug) I usually point to my source of information; see 'Abbreviations'. For non-SSul languages, the source is noted occasionally.

10 Some similar traits of the Maloh (Embaloh, Pari) language with the languages of South Sulawesi (Mak and Bug) were noted as far ago as in the middle of the 19th century. The problem was raised anew by Hudson (1978) who pointed to a number of similarities between the so-called Tamanic group (consisting of Maloh dialects) and the SSul languages, esp. Bug. At present it is being investigated by Adelaar. The preliminary results of Adelaar's research into the Maloh - SSul relations were presented in his talk 'Maloh, a South Sulawesi language of Central Kalimantan' (in Leiden, May 1988, not yet published).

11 Mills introduces this Sad word in another cognate set.

12 In addition to the examples presented by Mills, compare Karo ləmpur 'excessive'. The correspondence "PSS *a: Karo ə" should be unexpected in the penult.

13 In Cense's Mak dictionary (1979) timu is identified as "dial. Sal. (= Saleierees) en Bug". The word is not found in any of the Sel

wordlists accessible to me (G, SK, H).

14 The evidence of some languages of the Moluccas (Weda, Gane, Kayoa, East Makian etc. sumo 'mouth', Fordata sumar id.) suggests that the first syllable of the word was not *ti.

15 For the PSS level, it is difficult to choose between *e and *i. In this reconstruction I prefer *e in view of the areal nature of the given development, which forces us to take the non-SSul languages of Central Sulawesi into account.

16 The published Holle list (vol. 7/1, p.208) gives <bintoëng> for 'star' in the list of the language named Tontoli. This "Tontoli" is undoubtedly very close to the Tolitoli language whose 100-item Swadesh list is given by BBS. According to the latter list, the Tolitoli word for 'star' is kikilo'. The presence of a SSul word of Mak rather than Bug appearance in "Tontoli" seems somewhat unexpected.

17 Bug -welai 'to leave (go away from)' may be a derivative of this base.

18 I know Mdr mabela only from Mills' works.

19 Petapa <ale e> 'body' noted by Adriani may be a Bug loan.

20 Veen's Tae' (= Sad) dictionary (1940) does not include any 'general Sad' tindo' in this meaning, but it gives tindo' of Baruppu' dialect as the synonym of tundo' (other dialects or 'general Sad'?); the former is glossed 'in regelmatige volgorde plaatsen'.

21 By this procedure, the test is looked upon as a chain of morphemes. Turning them over one by one, we must decide whether the morpheme under consideration is to be written out or omitted. The problems arising here may differ in different language groups. In the SSul languages, the basic rule "write out root morphemes (i.e., both root words and the root morphemes of derived or compound words), omit affixes" evidently needs a few adjustments. Those adopted by me were, briefly stated, as follows. Regarding non-affixal bound morphemes, I took their relation to the word accent into account. As a general rule, I omitted the morphemes that never bear the accent, which implied omission of proclitic and enclitic pronominal morphemes, various monosyllabic (eventually asyllabic) modal and aspect-marking enclitics, etc. Exception was made for a few proclitics evidently belonging to sets most members of which are not clitic (like the Bug proclitic negative təŋ- co-existing side by side with the non-clitic de'e and aja'). The bound morphemes that regularly bear the accent, among them all disyllabic morphemes and several of the monosyllabic ones, were written out. At the same time, all the non-affixal bound morphemes of the 'vocabulary language' were taken into account, irrespective of their relation to the accent.

My rules concerning non-affixal bound morphemes are of course rather arbitrary and open to criticism (e.g., why is the accent 'drawn in', aod if so then why only in the 'text language'?). It is impossible to present any detailed justification here. I can only say that my rules are seemingly in fairly good agreement both with the structure and with the history of development of the SSul languages. Moreover, any modification of these rules would entail a change of percentages by no more than two or three points, not enough to change the overall picture.

22 Compare, for example:

Bug	arua-pulona 'eighty'	duappulo 'twenty'
Ko (F)	karua-bila anna 'eight hundred'	ruambila a 'two hundred'

Napu talu pulona 'thirty' rompulo 'twenty'

All the numerals given in the first column include the final morpheme -na (of pronominal origin) and lack any trace of the nasal ligature; the situation is exactly contrary in the second column.

ABBREVIATIONS

(a) Names of languages, dialects, language groups, proto-languages.

AN - Austronesian
ArT - Aralle-Tabulahan
Bal - Balinese
Bar - Bare'e (Pamona)
BI - Bahasa Indonesia
Bmb - Bambang (Bambam)
Bng - Banggai
Bot - Botteng
Bug - Buginese (Bugis)
Camp - Campalagian
Ceb - Cebuano
End - Endekan (according to Pelenkahu et al. 1972)
Enr - Enrekang (according to Grimes and Grimes 1987)
Fij - Fijian
Kal - Kalumpang
Ko - Konjo (Highland Konjo and Lowland Konjo are not differentiated)
KP - Kaili-Pamona
Kul - Kulawi
Lay - Laiyolo
Le - Ledo
Lem - Lemolang
Li - Lindu
Mad - Madurese
Maiw - Maiwan
Mak - Makasar(ese)
Mam - Mamasa
Mang - Manggarai
Mar - Maranao
Mas - Mas(s)enrempulu
Mbi - Mambi
Mdr - Mandar
Meh - Mehalaan
Mlg - Malagasy
Mmj - Mamuju
Mny - Maanyan
Mong - Mongondow
OJav - Old Javanese
Pal - Palili
PAN - Proto-Austronesian
PEF - Proto-East-Fijian
PMin - Proto-Minahasan
PMP - Proto-Malayo-Polynesian (in Blust's sense; see, e.g., Blust 1980)
PPh - Proto-Philippine
PSan - Proto Sangiric
PSS - Proto-South-Sulawesi
Pta - Pattae
Pti - Pattinjo
PUS - Pitu-Ulunna-Salu

RA - Rongkong Atas
RB - Rongkong Bawah
Rnt - Rantebulahan
Sad - Sa'dan(ese)
San - Sangir
Sas - Sasak
SekoP - Seko Padang
SekoT - Seko Tengah
Sel - Selayar
Snd - Sundanese
SSul - South Sulawesi
Sum - Sumare
Tag - Tagalog
Tap - Tappalang
To - Tonga
Ttb - Tontemboan
WMP - West Malayo-Polynesian (in Blust's sense;see, e.g., Blust 1980)

(b) Sources of information

A - Adriani 1898
Ad - K.A. Adelaar (unpublished talk, 1988, and private correspondence)
BBS - Barr, Barr and Salombe 1979
BLM - Pelenkahu et al. 1972
C - Cense 1979
D - Dunnebier 1951
DK - Pelenkahu et al. 1971
F - T. Friberg (private correspondence)
G - Grimes and Grimes 1987
H - the Holle lists (published in Pacific Linguistics (Canberra), series D)
I - Ide Said 1977
K - Kähler 1961
Kou - Koubi 1972
LM - Tenriadji and Wolhoff 1955
M - R.F. Mills (various works, mainly the Dissertation (1975b))
Ma - M.P. Martens (pre-prints, and private correspondence)
M Diss - Mills 1975b
Mu - Muthalib 1977
MW - Matthes 1885
S - J.N. Sneddon (various works)
SK - Sjahruddin Kaseng 1978
V - Veen 1940
Ve - Verheijen 1967

(c) Linguistic terms

C - consonant (Not to be confused with C - a proto-phoneme reconstructed for PAN)
dial.- dialectal

ESI - exclusively shared innovation
N - nasal
N_h - nasal homorganic with the following consonant
RM - reduplicated monosyllable
T - voiceless stop
V - vowel
X - non-nasal consonant

BIBLIOGRAPHY

Adriani, N.
1898 Iets over de talen der To Sada en der To Wadu. Mededeelingen van wege der Nederlandsch Zendeling-Genootschap XLII:111-150.

Barr, D.F., Sh.G. Barr, in cooperation with Dr. C. Salombe
1979 Languages of Central Sulawesi: Checklist, Preliminary Classification, Language Maps, Wordlists. Ujung Pandang.

Blust, R.
1973 Additions to 'Proto-Austronesian Addenda' and 'Proto-Oceanic Addenda with Cognates in non-Oceanic Austronesian Languages', No. II. Working Papers in Linguistics 5.3:33-61. Honolulu: Department of Linguistics, University of Hawaii.

1980 Austronesian Etymologies. Oceanic Linguistics XIX, Nos. 1 and 2.

1984 Austronesian Etymologies - II. Oceanic Linguistics XXII-XXIII, Nos. 1-2.

Cense, A.A., in Samenwerking met Abdoerrahim.
1979 Makassaars-Nederlands woordenboek. 's-Gravenhage: M. Nijhoff.

Dunnebier, W.
1951 Bolaang Mongondowsch-Noderlandsch Woordenboek. 's-Gravenhage: M. Nijhoff.

Esser, S.J.
1938 Talen. Sheets 9 and 9b in: Atlas van Tropisch Nederland. 's-Gravenhage: M. Nijhoff.

Friberg, T., and B. Friberg
Notes on Konjo Phonology. To appear.

Friberg, T., and T.V. Laskowske
South Sulawesi Languages. In this volume.

Grimes, Charles E., and Barbara D. Grimes
1987 Languages of South Sulawesi. Pacific Linguistics, Series D, No. 78.

Hudson, A.B.
1978 Linguistic Relations among Bornean Peoples with Special Reference to Sarawak: An Interim Report. In: M.D. Zamora et al. (eds). Sarawak: Linguistics and Development Problems, pp.1-44. Charlottesville: University of Virginia.

Ide Said
1977 Kamus bahasa Bugis-Indonesia. Jakarta.

Kähler, H.
1961 Simalur-Deutsches Wörterbuch. Berlin.

Kern, H.
1916 Taalvergelijkende verhandeling over het Aneityumsch. In: H. Kern, Verspreide Geschriften, vol. 5, 's-Gravenhage. (Originally published in 1906 in Verhandelingen van de Koninklijke Nederlandsche Akademie van Wetenschappen, afdeeling Letterkunde, Nieuwe Reeks 8:1-146)

Koubi, J.
1972 Rapport de mission en Pays Toradja (Août 71 - Janvier 72). Asie du Sud-Est et Monde Insulindien III.3:77-100.

Matthes, B.F.
1885 Makassaarsch-Hollandsch Woordenboek, 2nd ed. 's-Gravenhage.

Mills, R.F.
1975a The Reconstruction of Proto-South-Sulawesi. Archipel 10:205-224.

1975b Proto South Sulawesi and Proto Austronesian Phonology. (Dissertation, University of Michigan) University Microfilms, Ann Arbor.

1981 Additional addenda. In: R.A. Blust (ed). Historical Linguistics in Indonesia, Part I (= NUSA 10), pp. 59-82.

Muthalib, Abdul
1977 Kamus bahasa Mandar-Indonesia. Jakarta.

Noorduyn, J.
1955 Een achtiende-eeuwse kroniek van Wadjo; Buginese historiografie. 's-Gravenhage.

1963 Mededelingen uit de verslagen van Dr. S.J. Esser, taalambtenaar voor Celebes 1928-1944. Bijdragen tot de Taal-, Land- en Volkenkunde 119:329-370.

Palenkahu, R.A., et al.
1971 Dialek Kondjo di Sulawesi Selatan. Udjung Pandang: Lembaga Bahasa Nasional Tjabang III.

1972 Bahasa di Lima Massenrempulu. Udjung Pandang: Lembaga Bahasa Nasional Tjabang III.

Sirk, Ü.
1988 Towards the historical grammar of the South Sulawesi languages: Possessive enclitics in the postvocalic position. In Hein Steinhauer, ed. Papers in Western Austronesian Linguistics No. 4. p. 283-302. Pacific Linguistics Series A, No. 79. Canberra.

Sjahruddin Kaseng
1978 Kedudukan dan fungsi bahasa Makassar di Sulawesi Selatan. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

Sneddon, J.N.
1978 Proto-Minahasan: Phonology, Morphology and Wordlist. Pacific Linguistics, Series B, No. 54.

1984 Proto-Sangiric and the Sangiric Languages. Pacific Linguistics, Series B, No. 91.

Starostin, S.A.
1989 Sravnitel'no-istoričeskoe jazykoznanie i leksikostatistika /Comparative-historical linguistics and lexicostatistics/. In: Lingvističeskaja rekonstrukcija i drevnejšaja istorija Vostoka; materialy k diskussijam na Meždunarodnoj konferencii (Moskva... 1989 g.) /Linguistic reconstruction and the earliest history of

the Orient; materials of the International conference, Moscow 1989/, vol. 1, pp. 3-39. Moscow.

Tenriadji, A., dan G.J. Wolhoff
1955 Lontar Mandar. Bahasa dan Budaja III.3:7-31; III.4:25-38.

Tsuchida, Shigeru
1976 Reconstruction of Proto-Tsouic Phonology. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Veen, H. van der
1940 Tae' (Zuid-Toradjasch) - Nederlandsch Woordenboek. 's-Gravenhage: M. Nijhoff.

Verheijen, J.A.J.
1967 Kamus Manggarai I: Manggarai-Indonesia. 's-Gravenhage: M. Nijhoff.

Wurm, S.A. and Shiro Hattori (eds)
1983 Language Atlas of the Pacific Area, Pt 2: Japan Area, Philippines and Formosa, Mainland and Insular Southeast Asia. Pacific Linguistics, Series C, No. 67.

Wurm, S.A.and B. Wilson
1975 English Finderlist of Reconstructions in Austronesian Languages (post-Brandstetter). Pacific Linguistics, Series C, No. 33.

Zorc, R.D.P.
1977 The Bisayan Dialects of the Philippines: Subgrouping and Reconstruction. Pacific Linguistics, Series C, No. 44.

1986 The Genetic Relationships of Philippine Languages. In: P. Geraghty et al. (eds). FOCAL II: Papers from the Fourth International Conference on Austronesian Linguistics. Pacific Linguistics, Series C, No. 94. Pp. 147-173.