1. Proto-Austronesian (PAN) *s appears to have been a sibilant. Many Austronesian languages have a phoneme /s/, but in general this derives from *s. A comparative survey shows that where they differ, the reflexes of *s preponderantly exhibit greater occlusion than those of *S:

<table>
<thead>
<tr>
<th>Reflexes of PAN *s and *S in representative Austronesian languages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FORMOSA</strong></td>
</tr>
<tr>
<td>Thao</td>
</tr>
<tr>
<td>Ami</td>
</tr>
<tr>
<td>Puyuma</td>
</tr>
<tr>
<td>Paiwan</td>
</tr>
<tr>
<td><strong>PHILIPPINES AND WESTERN MICRONESIA</strong></td>
</tr>
<tr>
<td>Itbayaten</td>
</tr>
<tr>
<td>Agta</td>
</tr>
<tr>
<td>Ilocano</td>
</tr>
<tr>
<td>Tagalog</td>
</tr>
<tr>
<td>Kalamian</td>
</tr>
<tr>
<td>Tagbanwa</td>
</tr>
<tr>
<td>Western BUKIDNON</td>
</tr>
<tr>
<td>Manobo</td>
</tr>
<tr>
<td>Maranao</td>
</tr>
<tr>
<td>Tausug</td>
</tr>
<tr>
<td>Chamorro</td>
</tr>
<tr>
<td>Palauan</td>
</tr>
<tr>
<td><strong>MADAGASCAR</strong></td>
</tr>
<tr>
<td>Merina</td>
</tr>
<tr>
<td><strong>WESTERN INDONESIA AND MAINLAND SOUTHEAST ASIA</strong></td>
</tr>
<tr>
<td>Kadazan</td>
</tr>
<tr>
<td>Long Wat</td>
</tr>
<tr>
<td>Uma Juman</td>
</tr>
<tr>
<td>Iban</td>
</tr>
<tr>
<td>Malay</td>
</tr>
<tr>
<td>Toba Batak</td>
</tr>
<tr>
<td>Javanese</td>
</tr>
<tr>
<td>Jarai</td>
</tr>
<tr>
<td><strong>CENTRAL AND EASTERN INDONESIA</strong></td>
</tr>
<tr>
<td>Sangir</td>
</tr>
<tr>
<td>Proto-Minahasian *s</td>
</tr>
<tr>
<td>Gorontalo</td>
</tr>
<tr>
<td>Uma</td>
</tr>
<tr>
<td>Proto-South</td>
</tr>
<tr>
<td>Sulawesi</td>
</tr>
<tr>
<td>Bima</td>
</tr>
<tr>
<td>Manggarai</td>
</tr>
</tbody>
</table>

*What is most striking about the reflexes of *S is their geographical distribution: within Formosa *S generally appears as a sibilant, but elsewhere it almost invariably became /h/, /*/ or zero. Moreover, among non-Formosan languages segmental reflexes are attested (to date) only from the Philippines, Sulawesi, Borneo, Malaya-eastern Sumatra and mainland Southeast Asia. In the Lesser Sunda and Moluccan islands, and in the great archipelago of the Pacific basin the only reflex that has ever been reported is zero.*

To a large extent - but not entirely - the reflexes of *S follow the general contours of probable linguistic subgroups. Thus, arguments have been presented elsewhere (Blust 1977) for a genetic division of the Austronesian languages as follow:

**Figure 1**

A subgrouping of the Austronesian languages based in part on the reconstruction of the PAN pronouns (after Blust 1977)
guages of the South Halmahera-West New Guinea and Oceanic groups)

SHWNG: South Halmahera-West New Guinea (the MP languages of Halmahera, and of the adjacent north coast of New Guinea as far as the Mamberamo river, including the islands of the Raja Ampat group)

OC: Oceanic (the MP languages of Melanesia, Micronesia and Polynesia except as stated elsewhere)

Note: CMP and EMP may share a common node below MP. The position of Yapese within MP is unclear.

Because the indigenous languages of Formosa were only belatedly incorporated into the Austronesian comparative picture *S was earlier (Dyen 1953) reconstructed as *h. More recently, the discovery of Bornean reflexes such as Kiput s < *bV/s / Vi (awš < *bušuk 'head hair', awš < *bašaq 'floodwater', etc.) not only has increased the likelihood that *S was a sibilant in Proto-Austronesian, but strongly suggests that *S remained a sibilant in Proto-Malayo-Polynesian. The characteristic lenition and subsequent loss of this segment in Malayo-Polynesian (but not in Atayalic, Tsouic or Paiwanic) languages would thus seem to indicate widespread parallel development among at least the Western Malayo-Polynesian languages. By contrast, it has herefore been possible to assume that the loss of *S in Central and Eastern Malayo-Polynesian languages is the result of small number of changes in a few ancestral speech communities (as Proto-Central Malayo-Polynesian and Proto-Eastern Malayo-Polynesian) rather than of a larger number of independent changes. However, it now appears that this assumption can no longer be maintained.

2. The discovery of a non-zero reflex of *S in a WMP language in which it had not previously been reported would come as no great surprise, since a proto-segment must be assumed for Proto-Western Malayo-Polynesian and various interstage languages (Proto-Philippines, Proto-North Sarawak, Proto-Malayic, etc.) in any case. The discovery of such a reflex in a Central or Eastern Malayo-Polynesian language, on the other hand, would compel use to alter our notions of the phonology of Proto-Central Malayo-Polynesian or Proto-Eastern Malayo-Polynesian, and to recognize parallel development as the most probable explanation of the general loss of *S among its descendants.

2.1. Evidence of the type just described is now to hand for Soboyo, A CMP language in the westernmost part of the island of Taliahu in the Sula Archipelago. Although the linguistic situation in the Sula Archipelago is the most poorly known of any discrete island group in Indonesia, Fortgens (1921) has provided a sketch of some 111 pages (including a vocabulary of 69 pages) for Soboyo, together with comparative observations on its close relative, Kadai (eastern Taliahu). Soboyo as described by Fortgens appears to have the following system of phonological contrasts:

### Consonants

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>c</th>
<th>k</th>
<th>ʔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>d</td>
<td>j</td>
<td>g</td>
<td>ʔ</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td>ɾ</td>
<td>ɾ</td>
</tr>
<tr>
<td>s</td>
<td>ɾ</td>
<td>ɾ</td>
<td>ɾ</td>
<td>ɾ</td>
</tr>
</tbody>
</table>

### Vowels

<table>
<thead>
<tr>
<th>i</th>
<th>u</th>
<th>e</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>o</td>
<td>a</td>
<td>o</td>
</tr>
</tbody>
</table>

All stops (exclusive of /ʔ/), /s/ and /ɾ/ may be prenasalized in initial and medial positions. Although some prenasalized initials result from prepousing of the genitive marker (*kali 'skin', *kali 'skin off'; gge 'armpit', gge 'arm pit of'; ndom 'body hair', ndom 'body hair of'); bol 'eye', bol 'eye of'; others appear to belong to the root; mbob 'pig', mbob 'voice', grum 'whale', etc. The phonological status of non derived prenasalized initials is unclear from Fortgens' statements, but general considerations of morpheme structure suggest that they are unit phonemes. In directly inherited words only a, ə, ɪ, ʊ (which probably are allophones of /i/, /ŋ/, /k/ respectively) and ə occur in final position, though other final consonants are found in recent loanwords from Malay (hakim 'judge', curit 'letter').

Fortgens maintains (2-3) that the Soboyo vowels have open and close allophones that are dependent on the placement of stress (penultimate in the word). Their phonetic nature and even their distribution, however, is difficult to determine from the information given.

Finally, a large number of rising diphthongs (ae, ai, ao, au, ei, eu, oi, ou, ui) are recorded in final position. Stress placement (10) shows that these are phonological units: [mese?] 'cook', [matatu] 'person, human being', [tuoï] 'sit', [sosolü] 'ring', etc.

3. An interesting historical question concerns the sources of Soboyo /h/. Even a cursory inspection of the word-list provided by Fortgens shows that /h/ in Soboyo reflects more than one Proto-Austronesian phoneme. Thus, on the first page of the vocabulary we find ahit < *pari 'stingray'. That /h/ is a recurrent reflex of *h is clear from the following additional material:

1) *R > h
   *baRu > baha 'a plant: Hibiscus tiliaceus'
   *d)Uri-an > lhan 'durian'
   *laRiw > lahi 'run, flee'
   *(n)ARa > kana-naha 'a tree: Pterocarpus indicus'
   *naRa > naha 'brother (woman speaking)'
   *Ra(g)kit > ka-haki 'raft'
   *Raya > kaya 'large, great'
   *taRuq > tahtu 'put, place, set'
   *wRej 'linaa' > waho 'rattan'

But from the limited number of relevant instances it appears that Soboyo /h/ also reflects *r:

2) *R > h
   *burit > buhi 'hind-part, rear-end'
   *rebug > habug 'bamboo shoot'
A third source of /h/ is *d/D:

3) *d/D > h

*dakèp 'embrace' > hakò 'hold, cling to'
*dašak > ka-haa? 'clear the throat'
*daYa > hoyo 'toward the interior, landward'
*DuSa > hoyo 'two'
*ke(d)eq > kohoin 'stand'
*m-uDeSi > sa-muhi 'afterwards, behind'

Moreover, /h/ may reflect *p:

4) *p > h

*pa-našik > hana? 'mount, climb'
*pipa > hila 'how much, how many?'
*pitu > hitu '7'
*puluq > hulu '10'
*u(m)pu > uhu 'grandchild'

And in one known example it reflects *b:

5) *b > h

*baŋun > haungin 'rouse, wake up'

Although no instances of the change PAN *l > Soboyo /h/ are known, the correspondence of Proto-Ambon *l to Soboyo /h/ is attested in *labu 'mat': habu 'sleeping mat'. A similar change is found in Wallace's (1869) word-list from the Sula Archipelago which, as Portgens has shown (106-11) almost certainly represents a dialect spoken in northern Sanana and the adjacent coast of Mangole: puluq > foha '10' (apparently also *bulu > ni-foa 'feather', with subsequent loss of the intervocalic consonant). It is thus possible that Soboyo habu 'sleeping mat' is a loan from a dialect in which the change *l > /h/ is recurrent.

4. In addition to the above material /h/ appears in many other Soboyo words. For some of these no etymology is known. For a small number, however, a reconstructed form can be identified, and in these the source of /h/ is PAN *S. Thus, to cite the least problematic instances first, we find:

6) *S > h

a. *Sajek > hao 'sniff, smell'
   (Tagalog halık 'kiss', Cebuano Bisayan halık 'kiss, sniff, put the nose next to')

b. *Samuk > hamo 'run amuck'
   (Tagalog hónok 'hand-to-hand fight')

c. *Sasaq > haa, maqa-haa 'whet, sharpen'
   , paqa-haa 'whetstone'
   (Tagalog hasá? 'whet, sharpen')

d. *Sasek 'dibble' > ka-haa? 'dibble stick'
   (Tagalog hasík 'sowing', Cebuano Bisayan hasták 'make a hole to sow seeds in; dibble stick')

4.1. Various objections might be raised against these comparisons. First, we might assume that they can be explained as chance resemblances. But in each case the phonological correspondences manifest a demonstrably recurrent class. Thus, in 6b, although *u normally yields Soboyo /u/ in the ultima, it yields /o/ in a few other words that appear to be directly inherited: *kamu > kemo '2nd pl.', *pikuv > peko-in 'bent, crooked', *(CT)ku 'bend, curve' > tiko-in 'bay', *tinequ > tino 'weave'.

As noted earlier, in directly inherited words Soboyo permits only four final consonants: R, n, ? and a. The first of these generally derives from *ñ9 (‘na'jin 'wind', ‘na'jin 'sea current', ‘na'jen 'eat', ‘na'lunun 'roll up (of mats, sails)', ‘teken 'punting pole' > tokoin 'poling, pushing off of a boat'), the second from the merger of *m and *? (‘inum > inun 'drink', ‘senem > noo '6', ‘bata > fam 'stem, stalk, trunk', ‘ma-niño 'ma-nimó 'clear, of water') and the third from *k (‘na'nak > ana 'child', ‘manuk > anu 'chicken'). Given only this much information the assumed loss of the terminal nasal in huyu < *Suyuu would appear to weigh against the proposed derivation 6e. But in fact final /u/ is dropped in some probable loanwords, as bawa (Malay bawà) 'onion', pédà 'parang, machete' (Malay pédà 'sword'), and closer inspection reveals that the change PAN *-u > Soboyo zero is attested in nearly as many forms as the change *-g > q: *ba(s)u > basu 'k.o basket', *buku lâli > fuku lâli 'ankle', *ruq(q)u > ba-rugu (‘r > r unexpl. 'thunder', *tuRañ > fali-tuha 'friend, companion, guest', *beCeq 'millet sp.' > bete 'Colocasia antiquorum Schott' (?)). The development of the original final consonant in 6a-e is thus paralleled in other forms.

Soboyo inextricably exhibits more than one reflex of PAN *j. In one known case *j has disappeared:

1. *j > s
   *najan > gaain 'name'

In three others it appears as /l/:

2. *j > l
   *pipa > hila 'how much, how many?'
   *suja sulà-in 'pitfall in which short bamboo stakes are placed'
   *Suaji > ali-? 'younger sibling of like sex'

In view of these comparisons we might be well advised to treat comparison 6a with some
distrust. Yet careful searching shows that
the change *j > y/ not only is found in other
comparisons, but appears to be the most fre-
quent development in the available material:

3. *j > y
   *bajaq 'announce, convey news' > *baya
   ~ *faya 'say'
   *maja 'dry up, evaporate' > *moyo-q
   'dry'
   *qajeq > *agoq 'charcoal'
   *gapeju(s)u(n) > *n-oqu-in11 'gall, gall
   bladder'

Finally, the regularity of the change
*e > o is attested by several examples already
given, and by such additional comparisons as
*gatep > *ato 'thatch', *be(t)aw > *fato 'sister
(man speaking)', *tanem > *danog ~ *tanoq 'to
plant', etc.

4.2. If cps. 6a-6e cannot plausibly be
dismissed as chance resemblances, they might
nonetheless be regarded as loans from some
language in which *S regularly yielded /h/.
Apart from the non-Austronesian language
of Ternate, the most probable source of pre-
European loanwords in Soboyo is Malay, particu-
larly Moluccan Malay.

Dyen (1953:32) noted that PAN *S (then
written *h) became early Malay initial *h, but
as with *h from *q 'there is often competition
both in orthography and speech between syno-
nymers differing in the presence or absence of
an initial *h'. Dyen goes on to quote Malay
amo? 'fit of rage' and asah 'whet', but does
not cite material in support of reconstructions
6a, d and e. Wilkinson, who consistently cross-
references doublets differing in *h ~ *s gives
Malay amok 'furious attack; charge; amuck', asah
'grinding down; whetting, tooth-filing', asak
~ hanak 'stuffing, insertion by force or pres-
sure', uuq ~ huhuq 'rocking, reeling, swaying'.
To sum up, Standard Malay exhibits an initial
vowel in cps. 6b and 6c, but lacks a cognate of
6a. Similarly, although it has probable cognates
of 6d and 6e with initial *h, the
meanings of the Malay and Soboyo forms diverge
significantly from one another. Finally, the
last vowel of Soboyo *ka-hasa is incompatible
with a hypothesis of borrowing from a language
in which *e and *a had merged in the ultima.
It is difficult, then, to argue that any of
cps. 6a-e are loans from Standard Malay.

Given its commercial importance and great-
er proximity, a more likely source of lexical
loans in Soboyo, of course, is Moluccan Malay.
But Moluccan Malay appears to be the last freq-
(1876) and van Hoëvell (1876) is no more con-
vincing a loan-source for the material in
question that is Standard Malay.

Finally, we might consider the possibility
of borrowing from some language other than
Malay. Javanese is a conceivable candidate,
but as Dyen noted some years ago (1953:32) PAN
*S disappeared in Javanese, and is not attested
even in the oldest inscriptive materials.

The next nearest candidates for leading
languages appear to be Gorontalo of northeast-
ern Sulawesi, and Tausug of the Sulu Archipel-
ago in the southern Philippines. But so far
as is known Gorontalo lacks cognates of items
6a-e, and although the available material is
insufficient to determine whether Tausug has
cognates of any of these items, the recurrent
developments *j > j (tśajen > tjan 'name',
*pajey > paej 'riceplant, rice in the field')
and *e > u (*devør *maqu-aŋ 'hear', *ma-
tesas > ma-tuqas 'hard') in Tausug argue
against the hypothesis that this or a similar
language could have served as a loan-source
for the Soboyo items. After considerable
searching, then, we are unable to find a moti-
vated basis for the assumption that the Soboyo
members of cps. 6a-e are loans.

4.3. If cps. 6a-e cannot plausibly be
explained away as chance resemblances or as
loans from a language in which *S regularly
yielded /h/, it is nonetheless possible that
the appearance of /h/ in these forms is due to
a regular phonological development. But if
/h/ is epenthetic in these comparisons, it is
difficult to explain its absence in *aku > aku
'1st sg.', *alap/alqap > *aia 'get, fetch',
*anak > *anq? 'child', *aquin 'wind' > *aquin 'sea
current', *u(m)pu > *uku 'grandchild', *qu^an >
u^ain 'rain', etc. In short, PAN *a~*qa- and
*u~*u- regularly appear as Soboyo initial
*a/, /u/, but PAN *Sa- and *Su- regularly ap-
pear as ha-, hu-.

4.4. Two seemingly discordant comparisons
are known:

1. *SabaŋRa 'northwest monsoon' >
   barat 'north'
2. *Suaji > *ui-' si- 'younger sibling of
   like sex'

Comparison 1 exhibits atypical reflexes
both of *b (normally *f) and of *R (normally *h),
and so may be a loan. However, since the Malay
cognate differs in meaning (barat 'west') Malay
is not a likely loan-source, nor can other po-
tential sources be identified. Moreover, the
semantic development of *timuR 'southeast
monsoon' > timu 'south' (cp. Malay timur 'east')
strengthens the assumption that barat is
directly inherited. As in many Austronesian
languages, prepenultimate *a evidently centralized
in the history of Soboyo (*baqeqRu > beqeqRu
> *bok 'new'), and where it came to be
initial it deleted (*gapeju(s)u > epeju > peju
> eju > n-oqu-in 'gall'). It is possible that
the deletion of prepenultimate shwa also oc-
curred after /h/, thereby leading to loss of
the entire syllable. A similar explanation,
however, evidently is not available for the
loss of *a in cp. 2.

5. Comparisons 6a-e all involve instances
of PAN initial *S. By contrast, a number of
examples indicate that *S unconditionally dis-
appeared in intervocalic position:

7) *S~ > *a
   a. *aRuSu > kaw ahu 'a tree:Casuarina
   rumphiana Miq.'
   b. *daSaK > *ka-haa? 'clear the throat'
   c. *DuSa > *kowo 'two'
   d. *Kali-SepanL2 > *kalipain 'centipede'
   e. *KaSiu > *kawuy 'wood'
   f. *ma-Silu 'poisonous' > *miLu 'poison'
   g. *ma-Suah > *bala-moa 'yawn'
h. *m-ueDeSi > sa-muh'i 'afterwards, behind', ba-sa-muh'i 'later'
i. *pa-naSik > kana? 'mount, climb'
j. *qapejuS(u) > n-oyu-in 'gall'
k. *(t)uSud > tu-in 'knee'
l. *waSLR > wa-jo 'fresh water'

5.1. Two apparent exceptions to the loss of intervocalic *S have been noted:

8) *-S > h

a. *kuSít > kuhi 'knock down with an upward blow, as fruit from a tree, etc.' (Cebuano Bisayan kahi 'poke, jab, touch something with the fingers or something long to remove or obtain it; stick used to get something', Maranao kait 'touch lightly; pick; stick used to pick (as fruit')

b. *ma-Siaq > mahti 'shy, ashamed, embarrassed' (Paiwan masdad, Tagalog ka'yà, Kanowit ka'yà? idem)

Comparison 8a appears to be formally and semantically irreproachable, and so is not easy to dismiss as a product of chance. Comparison 8b is somewhat more problematic. In several known examples an internal sequence of unlike vowels in an original trisyllabic contracted in Soboyo: *buyaq > fòya 'crocodile', *Suajo > uli? 'younger sibling of like sex', *pa-naSik (> panaik) > kana? 'mount, climb'. As clear reflexes of such a reconstructed sequence in word-final position are unknown, it is impossible to determine whether the development in mahti is a special case of the more general change seen in *buayqa > fòya, etc. The latter, however, can perhaps be decided by another kind of evidence: since a closely similar form (mahe) is found in the locally influential language of Ternate, it is best to regard this item as a Ternate loan.

5.2. The preceding material would seem to indicate that while PAN *S-became h-, PAN *-S- disappeared in Soboyo. Comparison 8a, however, raises the possibility that *-S- sometimes became h-

A second apparent instance of Soboyo intervocalic h from *S is:

*Sins > niku (met.) 'snot, nasal mucus'. (Cebuano Bisayan hìyu 'sniffle, draw the snout (sic) up into the nose', Malay tiu 'nghu 'snot, nasal mucus')

This form is of interest for certain reasons other than those already considered. First, since Devin (1969) lists Burusee niku with an identical irregularity, and since other evidence points to the existence of a Sula-Buru group (see Collins, this volume), it is simplest to assume that the prima facie metathesis in this item preceded the separation of these two languages from a common ancestor. The discovery of metathesized or unmethathesized forms in other languages would serve to define the scope of a probable subgroup for which this change can be taken as one piece of evidence.

Second, if Soboyo kahit is taken as evi-

dence that some instances of intervocalic *S persisted as Proto-Sula-Buru *h-, the meta-
thesis in Soboyo, Buru *niku could have pre-
ceded the creation of intervocalic h from other sources. If, on the other hand, compari-
son 8a is regarded as a product of chance the metathesis in *niku would almost certainly have
followed the creation of intervocalic h from other sources, as it would otherwise have violated
an established morpheme structure constraint. A derivation *Sins > niku thus raises the possi-
bility of inferring a relative chronology for certain regular sound changes as well as for the sporadic change in this root.

Unfortunately, however, evidence from other languages of the central Moluccas suggests a
different etymology for Soboyo, Buru *niku (Manipa nìruk, Alang nìrâ, Banda Eli nìrak, Banda Elat runur 'snot', Asilulu nìku, Kaitetu niku 'breathe one's nose'), and it therefore appears preferable to abandon the historical derivation from a form with *S.

Burusee lacks reported cognates of the
Soboyo items in 6a-e, but the occurrence of
as-n < *Sasaq 'gills' and as-n < *Sawak
'waist' suggests that initial *S has dis-
appeared. Since PAN *R yields Buru /h/, the
reflexes of PAN initial *R and *S must have
been distinguished in Proto-Sula-Buru:

PAN *Raya 'large' *Sasaq 'whet' *Sasaq 'gills'
Psb *raya *hasa *hasa
Scb haya hasa - - -
Bur ha-t asa-n

To account for the loss of /h/ < *h, but
not of /h/ < *r in Burusee, then, we evidently
must assume the following order of changes:

Psb Soboyo Burusee
1. *h- > h
2. *r- > h

In short, despite the initial plausibility of
a derivation *Sins > Soboyo, Buru *niku
(met.) we are forced to conclude that Burusee,
unlike Soboyo, exhibits no clear reflex of *S
other than zero.13 Although few relevant data
are to hand, Alang nasa-na, Kairatu nasa-i,
etc. *Sasaq 'gills' nonetheless suggest that
Soboyo is not the only CMP language to preserve
a non-zero reflex of *S.

6. It has been shown that PAN *S regular-
ly appears as Soboyo /h/ in initial position,
and is perhaps occasionally reflected as /h/
in intervocalic position. Moreover, although
the evidence is far less clear-cut, *S may have become /h/ in some other languages of the cen-
tral Moluccas.

In the foregoing section we have assumed that Soboyo and Burusee share a close subgroup-
ing relationship, and that at a slightly higher level of inclusion these languages probably
belong to Stresemann's (1927) 'Ambo' group. Stresemann himself assumed a tripartite sub-
grouping of the languages of the central Moluccas, as follows:14

[Diagram: A -> B -> C]

PAM
As noted in section 2, the discovery of a previously unreported non-zero reflex of PAN *S in a Western Malayo-Polynesian language would occasion little surprise, but the discovery of such a development in a Central- or Eastern Malayo-Polynesian language would demand a basic revision of our notions about the phonological history of these groups. The significance of the development PAN *S > Soboyo /h/ thus depends crucially on the linguistic position of Soboyo.

Space does not permit a full presentation of the evidence and arguments here, but there are some qualitative indications that all of the Austronesian languages of the Lesser Sundas islands east of Sumbawa-rese (western Sumbawa), and all of the languages of the southern and central Moluccas including the languages of the Aru islands and the Sula archipelago belong to a subgroup apart from all other Austronesian languages. The collection of languages so defined corresponds roughly to Eser's (1938) Ambon-Timor group, but also includes the languages of his so-called Bima-Sumba and Sula-Batjan groups, except that Batjan is excluded. Following our earlier usage we shall call this proposed genetic unit 'Central Malayo-Polynesian' (CMP). One of the more striking innovations which appears to be shared exclusively by CMP languages is the merger of PAN *-mb- and *-mp- as PCMP *mb, and of PAN *-nd/D- and *-nt/T- as PCMP *nd.15 As a result of the syncopation of certain prefixal vowels this change is reflected also in initial position:

a) -mb-
   PAN *tumbak 'spear, lance'
   PAM *tubag 'spear'
   Soboyo tobo 'dagger'
   Endeh, Palue tumba 'spear'

c) ma-b
   PAN *ma-ber(e)qat 'heavy'
   PAM *bera 'heavy'
   Adonara ba'at, Riang berat 'heavy'
   Kambera mbotua 'heavy'

e) -nd/D-
   PAN *ma-ndi(b) 'cold'
   PAM *didi 'cold'
   Buru (eb)-ridim 'cold'

b) -mp-
   PAN *kampus 'stomach (of animal)'
   PAM *kabu- 'guts, belly'
   Soboyo kabu-i 'carry in front of the stomach'
   Tetum kabu-n, Kambera kambu 'belly'

d) ma-p
   PAN *ma-bensu 'full'
   PAM *benu 'full'
   Sika benu, Kambera mbina
   Kori mbanuka 'full'

e) -nt/T-
   PAN *pant(t)i 'banana'
   PAM *pudi 'banana'
   Atoni wki, Roti hanti 'banana'

f) -nt/T-
   PAN *gunaq 'intestines', but PAN *guna : PAN *anu 'float on a current,

Several remarks are appropriate at this point. First, as a result of gaps in the evidence relatively full cognate sets that include languages of the central Moluccas and languages of the Lesser Sundas are difficult to find. The otherwise unexplained change *p > b, or *t > d particularly if it occurs in initial position - is, however, a clear indication that the language in question has participated in this distinctive innovation. Soboyo forms such as bia (PAM *piia) 'good, beautiful', danan toen (PAM *tanem, PAM *danak 'tanak) 'to plant' thus constitute persuasive evidence for the inclusion of Soboyo in the Central Malayo-Polynesian group.17

Second, the CMP affiliation of Soboyo seems certain in any case from the number of phonological and lexical isoglosses which in teto appear to be shared exclusively with the languages of Buru. Shared phonological developments from Proto-Central Malayo-Polynesian include *y18 > s (Vor Bu > Sob foku, Bur foku-t 'new'), *R > h (the preceding and *paRa > Sob ahi, Bur pahi 'stingray'), the sporadic meta-thesis in *isa > Sob, Bur sila 'one'19 and the lexically specific split of y in the following items:

1. *-j- > p: *gaqan > Sob gaain, Bur ga-t 'name' (PAM *gala)
2. *-j- > y: *paqju > Sob n-oyu-in, Bur peu-n (= peyu-n?) 'gall, gall bladder' (PAM *pelu)
3. *-j- > l: *pija > Sob hila, Bur pila 'how much, how many?' (PAM *pila)

7. Stresemann (1927:6) summarizes the grammatical and phonological developments from Proto-Austronesian to Proto-Ambon in ten statements, of which the following can profitably be considered:

1. PAN *l, *r * became PAN *l
   *d/D, *z/z

2. PAN *d disappeared without a trace
   (*bis auf Spuren verloren')

3. In nouns PAN final *p and *k merged as glottal stop

4. The PAN final diphthongs *-aw (including *-ew), *-ay (including *-ey) and
   *-uy were almost universally monoph- thongized to /a/, /ua/

7.1. As a concluding contribution it will perhaps be useful to show that once the development PAN *S > Soboyo /h/ is recognized, the integration of this observation into the reconstruction of Proto-Ambon forces us to admit certain readjustments in the representation of other correspondences. In the process we will also be compelled to look more critically at other, unrelated features of Stresemann's reconstruction.

Stresemann (1927:74) indicates that the reflexes of 'Malayo-Polynesian h' (in these instances Dyen's *q) are inconsistent: PAN *qacem 'liver': PAN *qata(y) 'intestines', but PAN *qahud : PAN *anu 'float on a current,
drift', etc. Proto-Ambon *h is reconstructed only in two items (*h)ata(y) 'intestines', *h)ilu 'nose'), in both of which it appears to derive from PAN *q.20 But if *h is chosen to represent the Proto-Ambon reflex of PAN *S, it is evident that some other symbol is needed for Stresemann's PAN *h. Given the symbol used for the Proto-Austronesian distinction, *q would appear to be a reasonable choice. Hence the following revisions:

Stresemann revised meaning
reconstruction

*aku *aku 1st sg.
*ama *ama father
*asa *hasa whet, sharpen
*awa- *hawa waist
*(h)ata(y) *qatay liver
*(h)ilu *gliuqo nose

Our choice of *q (generally reflected as glottal stop in Western Malayo-Polynesian languages) in the Proto-Ambon words for 'liver' and 'nose', however, immediately raises a question about its identity or non-identity with the final segment in e.g. Stresemann's *ana? 'child' or *manu? 'bird'. In short, by a process of serial displacement the reconstruction of PAN *h to account for the Soboyo evidence sets off a chain-reaction reanalysis of Proto-Ambon phonology. Stresemann's own orthography suggests that the initial consonant in the words for 'liver' and 'nose' was distinct from the final consonant in the words for 'child' and 'bird'. Once again, a reasonable choice for the latter is dictated by the Proto-Austronesian symbol; we therefore identify Stresemann's *a? (reconstructed only in final position) with his *k (reconstructed only in non-final position) and write PAN *anak 'child', *manuk 'bird', etc. It is noteworthy that Stresemann's assumption 3 would have to be abandoned in any case because it fails to account for the fact that PAN *p disappeared in Soboyo, whereas *k became glottal stop: *dakep 'seize, embrace' > hako 'grasp, hold tightly', *qatep > aso 'thatch', *aNak > ana? 'child', *manuk > mana? 'bird'.

Our primary purpose here is not to attempt a thorough revision of Stresemann's reconstruction, as such a revision is to be expected on the basis of superior materials from the work of J. Collins, who has recently conducted a historically-oriented linguistic survey of the central Moluccas. Once Soboyo is recognized as a subgroup relative of the languages of Buru, however, certain additional inaccuracies in Stresemann's reconstruction leap to the fore. Among the more important of these are:

1) PAN *1, *j and *r could not have merged as PAN *1. Proto-Austronesian non-final *1 invariably yields Soboyo /l/ (*1aqIC > laq 'sky', *layAR > laya 'sail', *ala > aha 'fetch', *baju > bulu 'wood pigeon', etc.) while, as noted earlier, *1 normally became /y/, and *r is reflected as /h/. Three distinct Proto-Ambon phonemes thus appear to be required: *1, *j, *r. In a continuing chain-reaction, the recognition of PAN *r < PAN *r further requires a distinct symbol for Stresemann's PAN *r (< PAN *R); hence PAN *r.

2) Similarly, although PAN *z/2 sometimes produced Soboyo /d/ (*qa(z)2ay > aza 'chin, jaw') this Proto-Austronesian phoneme or phoneme-pair also yields Soboyo /y/ (*qubits > uyain 'rain'). It is possible that the explanation for this divergent development is to be sought in Stresemann's reconstruction of a PAN *d (< PAN *z/2, etc.), *d (< PAN *nz/nz, etc.) distinction, but Stresemann reconstructs both Proto-Ambon prototypes with *d (*ada-, *udan).

3) Stresemann himself recognized in the parenthetical equivocation of his orthography that the truncation of original final diphthongs so strikingly characteristic of many Central Moluccan languages is not found in all languages that he wished on other grounds to assign to an Ambon group. The retention of final diphthongs in Proto-Ambon must, in fact, be assumed if we are to account for the quality of the last vowel in such Soboyo forms as ase (PAN *gaCey) 'liver', ada (*qa(z)2ay 'chin, jaw', talaq (*t)ali(s)ay) 'a shore tree: Terminalia catappa', foto (*be(t)aw) 'sister (man speaking)', kaso (*kasaw) 'rafter', or paka-nako (*t)aka(0)kaw) 'steal'.

The revisions that we propose are summarized in Table 2:

Table 2

<table>
<thead>
<tr>
<th>PAN</th>
<th>(Stresemann) (revised)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*S</td>
<td>h</td>
</tr>
<tr>
<td>*q</td>
<td>h-</td>
</tr>
<tr>
<td>*k</td>
<td>-?</td>
</tr>
<tr>
<td>*j</td>
<td>1</td>
</tr>
<tr>
<td>*r</td>
<td>1</td>
</tr>
<tr>
<td>*R</td>
<td>r</td>
</tr>
<tr>
<td>*z/2</td>
<td>d</td>
</tr>
</tbody>
</table>
| *-ay, etc. | a(y), etc. | z (?)

8. The net result of our revisions in the preceding section is that we posit a Proto-Ambon phonological system that looks considerably more like that of Proto-Austronesian than does Stresemann's reconstruction. Insofar as some (though not all) of the changes listed in Table 2 result directly from a consideration of the Soboyo reflexes of PAN *S, our study of a seemingly isolated set of observations can be seen to have systematic consequences that reach well beyond the limited problem for which a solution was initially proposed. The clear implication, once again, is that the change *S = h is only one of a number of features attributed by Stresemann to Proto-Ambon that are better explained as the result of innovations in individual daughter languages. If
there is a general theoretical lesson to be learned from particular studies such as the present one, then, it is that many — perhaps most — linguistic changes acquire their attested distribution by diffusion (or, in some cases, drift). A corollary of this conclusion is that preliminary, superficial comparative investigations within a well-defined subgroup will tend to show greater change from family ancestor to subgroup ancestor than in fact took place.

FOOTNOTES

*The original version of this paper was written in the Fall of 1978, and sent to J. Collins, who was then conducting a historically-oriented linguistic survey of the central and southern Moluccas. I am much indebted to him for his comments and additional data, which have led to substantial improvements in the present version.

1. To account for (generally minor) discrepancies in the correspondences holding between the Austronesian languages of Formosa, Dyen (1965) recognized six types of *S and several other segments which exhibit a similar pattern of reflexes either within Formosa (*x₁, *x₂, *x) or in non-Formosan Austronesian languages (*H). Since this paper is not concerned directly with Formosan evidence, *S will be used as a cover term for Dyen’s *S₁–*S₆ and *x. Apart from this deviation PAN reconstructions follow Dempwolff (1934-38) as modified by Dyen (esp. 1953, 1965). Reflexes of *S in Table 1 exclude final *S, which was lost or obscured through analogical reworking.

2. Many of the languages of western and northern Borneo exhibit striking anomalies in the reflexes of PAN voiced obstruents: e.g. Bario Kelabit abâk, Kiput sâk, Long Anap puôk, Bintulu buuk < *bus Suk ‘head hair’, next to buuk, buuk, buuk and buuk < *bulu ‘body hair, feathers’ in the same languages. To explain these observations it has been argued (Blust 1969, 1973, 1974a) that the first of like vowels, or of unlike vowels one of which was shwa, deleted following a voiced obstruent and preceding *S in a language that was ancestral to at least those languages which show evidence of this phonemic split. Because the results of vowel deletion involve fusion with an adjacent consonant the reflex of *S in these languages cannot be stated in purely segmental terms. In non-deleting intervocalic environments *S evidently became glottal stop (*bus Suk > Long Wat ba’u ‘smell, odor’, *baSun > Bario Kelabit di’nun, Long Wat la’uN ‘leaf’).

3. As noted elsewhere (Blust 1973), in conformity with an areal norm penultimate vowels centralized and shortened in Jarai (and other languages of the mainland southeast Asian Chamic group), subsequently deleting where the resulting cluster or phonetically complex segment was areally appropriate. Although *S disappeared in initial and final positions (*Sasaq > aak ‘whet’, *talqS > telay ‘rope’), it became intervocalic glottal stop if preceded by any consonant other than a voiced obstruent (*tuSud > Jarai to?uN ‘knee’); preceding voiced obstruents, on the other hand, were pre-glottalized (*bus Suk > buk ‘head hair’, *baSuq > bu ‘smell, odor’).

4. Proto-Minahasan *bu’uk ‘hair’ < PAN *buSuk ‘head hair’ suggests that PAN *S sometimes became PM *S. However, PM *bu < *baSuq ‘smell, odor’, *dua < *DuSa ‘two’ are best explained on the assumption that *S disappeared in Proto-Minahasan. Given the latter observation the medial consonant in *bu’uk can be seen as a secondary development between derived sequences of like vowels.

5. Fortgens writes * as ‚ as ng and the palatal series j, ɟ, y as dj, s, j respectively.

6. Written -G, as in polot ‘blood’, and -in, as in fulun ‘body hair’. I am indebted to J. Collins for clarifying Fortgens’ orthographic conventions.

7. Although Dempwolff (1934-38) reconstructed *tamburi, Oceanic reflexes such as Nauna (Admiralty) tawau and Fiji Gawai instead indicate *tamburi ‘conch trumpet’.

8. The apparent change *u > o in *laSud > low ‘toward the sea’, *SaSuk > yawo ‘far’ and *niu > nivo ‘cross’ on the other hand, evidently arose through breaking (to earlier *lawed, *javew, *niweR), followed by the regular change *e > o, as with *waSiR > wair > wayeR > wayo ‘fresh water’.

9. But optionally from *-e: *ke(d)e > kokont ‘stand’.

10. For another instance of the change *a > o < o cf. Daya > hoyo ‘toward the interior’; for another instance of the addition of -n cf. *isi > isi-n ‘flesh, contents’.

11. With secondary genitive prefix n- (cf. adem-ade ‘chin, jaw’, uluo-uluo ‘head hair’) and fossilized genitive suffix -in, as in fulun > *bulu ‘body hair, feathers’.

12. Most Austro-Asiatic reflexes *gal- and *kali- are reconstructible with the names of various insects, crustaceans, etc. Soboyo kalipain presumably reflects a form with the latter prefix.

13. Other initially appealing etymologies, as Burmase ba’u-k ‘give forth odor’ < *baSuq ‘smell, odor’ also prove on closer inspection to be illusory. Thus *b normally yields Buru /f/* (SabaRat > fahat ‘west monsoon’, *be(t)a > feta-n ‘sister (man speaking)’, *banua > fen ‘tribe, tribal village’), and for this reason among others fa-n ‘smell, odor’ is a more likely continuation of *baSuq.

14. Stresemann called the proto-languages of his hypothesized subgroups Sub-Ambon, Sub-Buru and Sub-Seran (= Ceram) respectively. Although he does not refer to Proto-Ambon Buru reconstructions, he clearly maintains (p. 8) that the division between A and C is greater than that between A and B.

15. Prenasalization of the bilabial stop may have been redundant in PCMP, but *d and *nd apparently contrasted. If so, in various daughter languages the change *d > /r/ led to reanalysis of /mb/, /nd/ as /b/ (= [mb]), /d/
Bimanese evidently did not participate in this merger (*tampak > dampa 'blunt', *gali-mCag > linta 'water leech', *pandDan > fanda 'pandanus', *DeSem > rindi 'dark'), but can be classified as Central Malayo-Polynesian on other grounds.

16. To account for the comparison Javanese kempun 'ventral side', Malay kempun-an 'bladder', Putunana kapu 'fish stomach' DempoWolff (1934-38) reconstructed *kempun 'ventral side'. The present variant can be posited on the basis of reflexes in CMP languages, and in some Oceanic languages, as Baluan (Admiralties) kapu-n, Penchel (Admiralties) kahu-n 'stomach of an animal', and perhaps Lakalai (New Britain) kapu-tu 'stomach'.

17. Although voicing distinctions are neutralized following a nasal in various languages of northern Sarawak (Blust 1974b: note 13), this change never affects obstruents in morpheme-initial position. Similarly, although PAN *b and *p merged in Proto-Oceanic they did so unconditionally, and this development contrasts with that of *d and *t, *nd/D and *nt/T, which remained distinct.

18. Probably [ŋ].

19. A similar (presumably independent) metathesis is reported in Kemak (Timor) sia 'one'.

20. Dyen (1953 and subsequent works) writes *çaCey, *iJiy. The latter reconstruction (with initial vowel) is supported by such reflexes as Tagabili ToFit, Tongan thi, but *iJiy evidently is needed to account for the initial consonant of Malay hitoq in conjunction with the Proto-Ambon reconstruction, and perhaps for Paluan *Iis (*-ŋ > ŋ unexpl.) 'nose'.

21. Fortgens lists two formally and semantically similar items in which *k has disappeared: *serek > sook 'insert, penetrate', *(t)usuk > tusuk 'penetration of a nail, thorn, etc.', but these clearly are exceptional.

REFERENCES


---------. n.d. Fieldnotes on languages of Borneo, the Lesser Sundas islands and western Melanesia.


***