

THE SOBOYO REFLEXES OF PROTO-AUSTRONESIAN *S

Robert Blust
University of Leiden

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:¹

Table 1

Reflexes of PAN *s and *S in representative Austronesian languages

	*s	*S
FORMOSA		
Thao	t	s
Ami	ts	s
Puyuma	s	s, ø
Paiwan	t	s
PHILIPPINES AND WESTERN MICRONESIA		
Itbayaten	s	h
Agta	t	ø
Ilocano	s	ø
Tagalog	s	h
Kalamian		
Tagbanwa	s	ø
Western		
Bukidnon		
Manobo	s	h
Maranao	s	?, ø
Tausug	s	h
Chamorro	s	ø
Palauan	t	ø
MADAGASCAR		
Merina	s	ø
WESTERN INDONESIA AND MAINLAND SOUTHEAST ASIA		
Kadazan	s	vowel deletion ²
Long Wat	s	?, vowel deletion ²
Uma Juman	s, h	h, ø
Iban	s	ø
Malay	s	h, ø
Toba Batak	s	ø
Javanese	s	ø
Jarai	s	ø, ?, preglottalization ³
CENTRAL AND EASTERN INDONESIA		
Sangir	s	ø
Proto-		
Minahasan	*s	ø ⁴
Gorontalo	t	h, ø
Uma	s, h	ø
Proto-South		
Sulawesi	*s	ø
Bima	s	ø
Manggarai	c	ø

	*s	*S
Roti	s	ø
Leti	s	ø
Fordat	s	ø
Kei	s	ø
Proto-Ambon	*s	ø

WESTERN NEW GUINEA AND THE PACIFIC

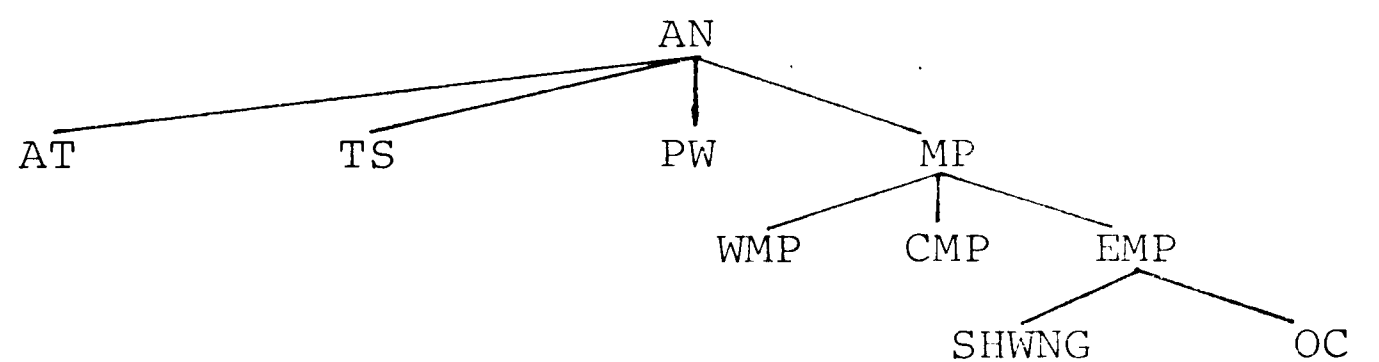
	s	ø
Buli	s	ø
Numfor	s	ø
Proto-		
Oceanic	*s, ns	ø

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/, /ʔ/ 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 archipelagoes 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)



AT: Atayalic (Formosa)

TS: Tsouic (Formosa)

PW: Paiwanic (Formosa)

MP: Malayo-Polynesian (all AN languages outside Formosa)

WMP: Western Malayo-Polynesian (the MP languages of the Philippines and western Indonesia, including Chamorro, Palauan, Chamic and Malagasy)

CMP: Central Malayo-Polynesian (the MP languages of the Lesser Sunda islands east of Sumbawarese, and of the southern and central Moluccas)

EMP: Eastern Malayo-Polynesian (the lan-

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* < *bV₁S / V₁ (*suəʔ* < *buSuk 'head hair', *səiʔ* < *baSaŋ '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 heretofore 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 Taliabu in the Sula Archipelago. Although the linguistic situation in the Sula Archipelago is among 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 Taliabu). Soboyo as described by Fortgens appears to have the following system of phonological contrasts:

Consonants⁵

p	t	c	k	ʔ
b	d	j	g	
m	n	ɲ	ŋ	
f	s	ʃ		h
	l			
	r			
w		y		

Vowels

i	u
e	o
a	

All stops (exclusive of /ʔ/). /s/ and /ʃ/ may be prenasalized in initial and medial positions. Although some prenasalized initials result from preposing of the genitive marker (*kuli* 'skin', *ŋkuli* 'skin of'; *gege* 'armpit', *ŋgege* 'armpit of'; *fuluin* 'body hair', *mfuluin* 'body hair of', *belen* 'eye', *mbelen* 'eye of'), others appear to belong to the root: *mboa* 'pig', *ndio* 'voice', *ngawuni* '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 *e*, *ɲ*,⁶ *ʔ* (which probably are allophones of /t/, /n/, /k/ respectively) and *ŋ* occur in final position, though other final consonants are found in recent loanwords from Malay (*hakim* 'judge', *sarat* '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: [méseʔ] 'cook', [matatúu] 'person, human being', [túoi] 'sit', [sosólui] '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 *ahi* < *paRi 'stingray'. That /h/ is a recurrent reflex of *R is clear from the following additional material:

- 1) *R > h
 *bageRu > *foku* 'new'
 *baRu > *bahu* 'a plant: Hibiscus tiliaceus'
 *(d)uRi-an > *dihain* 'durian'
 *laRiw > *lahi* 'run, flee'
 *(n)aRa > *kana-naha* 'a tree: Pterocarpus indicus'
 *ñaRa > *naha* 'brother (woman speaking)'
 *Ra(ŋ)kit > *ka-haŋki* 'raft'
 *Raya > *haya* 'large, great'
 *taRuq > *tahu* 'put, place, set'
 *waRej 'liana' > *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'
 *rebuŋ > *hobuŋ* 'bamboo shoot'

*tamburi⁷ > *tafuhi* 'conch shell
trumpet'

A third source of /h/ is *d/D:

3) *d/D > h

*dakep 'embrace' > *hako* 'hold, cling
to'
*daSak > *ka-haa?* 'clear the throat'
*Daya > *hoyo* 'toward the interior,
landward'
*DuSa > *howo* 'two'
*ke(d)eq > *kohoin* 'stand'
*m-uDeSi > *sa-muhi* 'afterwards,
behind'

Moreover, /h/ may reflect *p:

4) *p > h

*pa-naSik > *hana?* 'mount, climb'
*pija > *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

*baṇun > *haṇuin* '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 Fortgens 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 > *hayo?* 'sniff, smell'
(Tagalog *halík* 'kiss',
Cebuano Bisayan *halúk*
'kiss, sniff, put the
nose next to')
- b. *Samuk > *hamo?* 'run amuck'
(Tagalog *hámok* 'hand-to-
hand fight')
- c. *Sasaq > *hasa, maṇa-hasa* 'whet, sharp-
en', *paṇa-hasa* 'whetstone'
(Tagalog *hasá?* 'whet, sharpen')
- d. *Sasek 'dibble' > *ka-haso?* 'dibble
stick' (Tagalog *hasík*

'sowing', Cebuano Bisayan
hasúk 'make a hole to sow
seeds in; dibble stick')

- e. *Suyun 'shake, sway, stagger' >
huyu 'shake, wake someone by
shaking', *mo-huyu* 'earthquake'
(Western Bukidnon Manobo *huyun*
'shake something back and
forth, as a tree to make the
fruit fall, or as an earthquake
shakes a house', Malay *huyon*
'rocking, reeling, swaying')

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.', *pikuq > *penko*-in 'bent, crooked', *(CtT)iku 'bend, curve' > *tiko*-in 'bay', *tinegun > *tino* 'weave'.⁸

As noted earlier, in directly inherited words Soboyo permits only four final consonants: *ñ*, *n*, *ʔ* and *ɕ*. The first of these generally derives from *n⁹ (*aṇin 'wind' > *aṇín* 'sea current', *kaṇen > *kain* 'eat', *lunlun > *luluin* 'roll up (of mats, sails)', *teken 'punting pole' > *tokoin* 'poling, pushing off of a boat'), the second from the merger of *m and *ṇ (*inum > *inuṇ* 'drink', *ʔenem > *noṇ* '6', *bataṇ > *fataṇ* 'stem, stalk, trunk', *ma-niṇniṇ > *ma-niniṇ* 'clear, of water') and the third from *k (*aNak > *ana?* 'child', *manuk > *manu?* 'chicken'). Given only this much information the assumed loss of the terminal nasal in *huyu* < *Suyun would appear to weigh against the proposed derivation 6e. But in fact final /ṇ/ is dropped in some probable loanwords, as *bawa* (Malay *bawaṇ*) 'onion', *peda* 'parang, machete' (Malay *pědaṇ* 'sword'), and closer inspection reveals that the change PAN *-ṇ > Soboyo zero is attested in nearly as many forms as the change *-ṇ > ṇ: *ba(s)uṇ > *basu* 'k.o.basket', *buku laliṇ > *fuku lali* 'ankle', *ru(ṇ)guṇ > *ba-runṇu* (*r > r unexpl.) 'thunder', *tuRaṇ > *fali-tuha* 'friend, companion, guest', *beCeṇ 'millet sp.' > *bete* 'Colocasia antiquorum Schott' (?). The development of the original final consonant in 6a-e is thus paralleled in other forms.

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

1. *j > ø
*ṇajan > *ṇaain* 'name'

In three others it appears as /l/:

2. *j > l
*pija > *hila* 'how much, how many?'
*suja *sula*-in 'pitfall in which
short bamboo stakes are placed'
*Suaji > *uli-?* '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 frequent development in the available material:

3. *j > y

- *bajaq 'announce, convey news' > *baya*
~ *faya* 'say'
- *maja 'dry up, evaporate' > *moyo-ŋ*
'dry'10
- *qajen > *ayon* 'charcoal'
- *qapejuS(u) > *n-oyu-inl1* '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 > *foto* 'sister (man speaking)', *tanem > *danoŋ* ~ *tanon* '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 became /h/. Apart from the non-Austronesian language of Ternate, the most probable source of pre-European loanwords in Soboyo is Malay, particularly 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 synonyms 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* ~ \emptyset gives Malay *amok* 'furious attack; charge; amuck', *asah* 'grinding down; whetting, tooth-filing', *asak* ~ *hasak* 'stuffing, insertion by force or pressure', *uyon* ~ *huyon* '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-haso?* 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 greater proximity, a more likely source of lexical loans in Soboyo, of course, is Moluccan Malay. But Moluccan Malay as described by de Clercq (1876) and van Hoëvell (1876) is no more convincing a loan-source for the material in question than 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 inscriptional materials.

The next nearest candidates for lending languages appear to be Gorontalo of northeastern Sulawesi, and Tausug of the Sulu Archipelago 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 > \emptyset (+*ŋajan* > *ŋan* 'name', **pajey* > *pay* 'riceplant, rice in the field') and *e > u (**deneR* mag-*dunug* 'hear', **ma-teRas* > *ma-tugas* '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 motivated 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/alaq > *ala* 'get, fetch', *aNak > *ana?* 'child', *aŋin 'wind' > *aŋin* 'sea current', *u(m)pu > *uhu* 'grandchild', *quZaŋ > *uyain* 'rain', etc. In short, PAN *a-/*qa- and *u-/*qu- regularly appear as Soboyo initial /a/, /u/, but PAN *Sa- and *Su- regularly appear as *ha-*, *hu-*.

4.4. Two seemingly discordant comparisons are known:

1. *SabaRat 'northwest monsoon' > *barat* 'north'
2. *Suaji > *uli-?* '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 potential 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 (**bageRu* > *begeRu* > *beRu* > *fohu* 'new'), and where it came to be initial it deleted (**qapejuS(u)* > *epeju* > *peju* > *eju* > *n-oyu-in* 'gall'). It is possible that the deletion of prepenultimate shwa also occurred 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 disappeared in intervocalic position:

7) *-S- > \emptyset

- a. *aRuSu > *kaw ahu* 'a tree:Casuarina rumphiana Miq.'
- b. *daSak > *ka-haa?* 'clear the throat'
- c. *DuSa > *howo* 'two'
- d. *kali-Sepan12 > *kalipain* 'centipede'
- e. *kaSiw > *kayu* 'wood'
- f. *ma-Silu 'poisonous' > *milu* 'poison'
- g. *ma-Suab > *bala-moa* 'yawn'

- h. *m-uDeSi > sa-muhi 'afterwards, behind', ba-sa-muhi 'later'
- i. *pa-naSik > hana? 'mount, climb'
- j. *qapejuS(u) > n-oyu-in 'gall'
- k. *(t)uSud > tu-in 'knee'
- l. *waSiR > wayo '(fresh) water'

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

8) *-S- > h

- a. *kuSit > kuhi 'knock down with an upward blow, as fruit from a tree, etc.' (Cebuano Bisayan *kúhit* 'poke, jab, touch something with the fingers or something long to remove or obtain it; stick used to get something', Maranao *koit* 'touch lightly; pick; stick used to pick (as fruit)')
- b. *ma-Siaq > mahi 'shy, ashamed, embarrassed' (Paiwan *masiaq*, Tagalog *hiyá?*, Kanowit *hiyá?* 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 trisyllable contracted in Soboyo: *bugaya > foya 'crocodile', *Suaji > uli-? 'younger sibling of like sex', *pa-naSik (> panaik) > hana? '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 *mahi* is a special case of the more general change seen in *bugaya > foya, etc. The issue, 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:

*Sigus > ŋihu (met.) 'snot, nasal mucus'. (Cebuano Bisayan *hiŋus* 'sniffle, draw the snout (sic) up into the nose', Malay *iŋus* ~ *hiŋus* 'snot, nasal mucus')

This form is of interest for certain reasons other than those already considered. First, since Devin (1969) lists Buruese *ŋihu* 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 unmetathesized 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 *kuhi* is taken as evi-

dence that some instances of intervocalic *S persisted as Proto-Sula-Buru *-h-, the metathesis in Soboyo, Buru *ŋihu* could have preceded the creation of intervocalic h from other sources. If, on the other hand, comparison 8a is regarded as a product of chance the metathesis in *ŋihu* 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 *Sigus > ŋihu thus raises the possibility 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 *ŋihu* (Manipa *niruk*, Alang *nirut*, Banda Eli *niruto*, Banda Elat *nurut* 'snot', Asilulu *niru*, Kaitetu *niku* 'blow one's nose'), and it therefore appears preferable to abandon the historical derivation from a form with *S.

Buruese lacks reported cognates of the Soboyo items in 6a-e, but the occurrence of *asa-n* < *Sasaŋ 'gills' and *awa-n* < *Sawak 'waist' suggests that initial *S has disappeared. 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'	*Sasaŋ 'gills'
PSB	*raya	*hasa	*hasan
Sob	haya	hasa	- - -
Bur	ha-t	- - -	asa-n

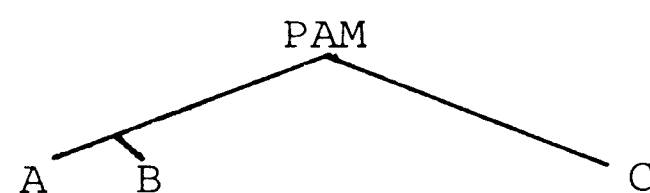
To account for the loss of /h/ < *h, but not of /h/ < *r in Buruese, then, we evidently must assume the following order of changes:

	PSB	Soboyo	Buruese
1.	*h- >	h	ø
2.	*r- >	h	h

In short, despite the initial plausibility of a derivation *Sigus > Soboyo, Buru *ŋihu* (met.) we are forced to conclude that Buruese, unlike Soboyo, exhibits no clear reflex of *S other than zero.¹³ Although few relevant data are to hand, Alang *hasa-na*, Kairatu *hasa-i*, etc. < *Sasaŋ '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 regularly 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 central Moluccas.

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



A: Proto-Ambon
B: Proto-Buru
C: Proto-Ceram

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 Sunda islands east of Sumbawarese (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 Esser'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-	b) -mp-
PAN *tumbak 'spear, lance'	PAN *kampuŋ ¹⁶ 'stomach (of animal)'
PAM *tubaq spear'	PAM *kabu- 'guts, belly'
Soboyo <i>tobo?</i> 'dagger'	Soboyo <i>kabu-i</i> 'carry in front of the stomach'
Endeh, Palue <i>tumba</i> 'spear'	Tetum <i>kabu-n</i> , Kambera <i>kambu</i> 'belly'
c) ma-b	d) ma-p
PAN *ma-beR(e)qat 'heavy'	PAN *ma-penuq 'full'
PAM *bera 'heavy'	PAM *benu 'full'
Adonara <i>ba?at</i> , Riung <i>berat</i> ,	Soboyo <i>bonu</i> 'full, deep'
Kambera <i>mbotu</i> 'heavy'	Sika <i>benu</i> , Kambera <i>mbinu</i>
	Kodi <i>mbanuka</i> 'full'
e) -nd/D-	f) -nt/T-
PAN *ma-(d)in(d)in 'cold'	PAN *pun(t)i 'banana'
PAM *didi 'cold'	PAM *pudi 'banana'
Buru (eb)- <i>ridin</i> 'cold'	Atoni <i>uki</i> , Roti <i>huni</i> 'banana'

e) -nd/D-

f) -nt/T-

Atoni ma-*nikin*

Roti maka-*lini*

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* (PAN *pia) 'good, beautiful', *danon* ~ *tanon* (PAN *tanem, PAM *daně ~ *taně) '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 toto* appear to be shared exclusively with the languages of Buru. Shared phonological developments from Proto-Central Malayo-Polynesian include *vl8 > f (*vəRu > Sob *fohu*, Bur *fehu*-t 'new'), *R > h (the preceding and *paRi > Sob *ahi*, Bur *pahi* 'stingray'), the sporadic metathesis in *isa > Sob, Bur *sia* 'one'¹⁹ and the lexically specific split of *j in the following items:

1. *-j- > ø: *ŋajan > Sob *ŋaain*, Bur *ŋa-t* 'name' (PAM *ŋala)
2. *-j- > y: *pəju > 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, j, r became PAM *l
*d/D, z/Z
*n, ñ
2. PAN *S 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 monophthongized to /a/, /u/

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 *qaCey 'liver' : PAM *(h)ata(y) 'intestines', but PAN *qañud : PAM *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 PAM *h. Given the symbol used for the Proto-Austronesian distinction, *q would appear to be a reasonable choice. Hence the following revisions:

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

Our choice of *q (generally reflected as glotal 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 PAM *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 *? (reconstructed only in final position) with his *k (reconstructed only in non-final position) and write PAM *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 > *ato* 'thatch', *aNak > *ana?* 'child', *manuk > *manu?* 'bird'.²¹

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 inadequacies in Stresemann's reconstruction leap to the fore. Among the more important of these are:

- 1) PAN *l, *j and *r could not have merged as PAM *l. Proto-Austronesian non-final *l invariably yields Soboyo /l/ (*lanic > *lani* 'sky', *layaR > *laya* 'sail', *alap/alaq > *ala* 'fetch', *baluj > *bule-falu* 'wood pigeon', etc.) while, as noted earlier, *j normally became /y/, and *r is reflected as /h/. Three distinct Proto-Ambon phonemes thus appear to be required: *l, *j, *r. In a continuing chain-reaction, the

recognition of PAM *r < PAN *r further requires a distinct symbol for Stresemann's PAM *r (< PAN *R); hence PAM *R.

- 2) Similarly, although PAN *z/Z sometimes produced Soboyo /d/ (*qa(zZ)ay > *ade* 'chin, jaw') this Proto-Austronesian phoneme or phoneme-pair also yields Soboyo /y/ (*quZaN > *uyain* 'rain'). It is possible that the explanation for this divergent development is to be sought in Stresemann's reconstruction of a PAM *d (< PAN *z/Z, 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 *ate* (PAN *qaCey) 'liver', *ade* (*qa(zZ)ay 'chin, jaw', *talise* (*(t)ali(s)ay) 'a shore tree: Terminalia catappa', *foto* (*be(t)aw) 'sister (man speaking)', *kaso* (*kasaw) 'rafter', or *paka-nako* (*(t)a(n)kaw) 'steal'.

The revisions that we propose are summarized in Table 2:

Table 2

Proposed revisions of Stresemann's Proto-Ambon phonology based on the inclusion of Soboyo reflexes

PAN	PAM (Stresemann)	PAM (revised)
*S	∅	h
*q	h-	q-
*k	-?	-k
*j	l	j
*r	l	r
*R	r	R
*z/Z	d	z (?)
*-ay, etc.	a(y), etc.	ay, etc.

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 > ∅ 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 *əb^huk*, Kiput *suəʔ*, Long Anap *puk*, Bintulu *buk* < *buSuk 'head hair', next to *buluh*, *buləw*, *bulu*, *buləw* ~ *vuləw* < *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 (*baSuq > Long Wat *baʔuʔ* 'smell, odor', *daSun > Bario Kelabit *daʔun*, Long Wat *laʔun* '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 > *asah* 'whet', *taliS > *tələy* 'rope'), it became intervocalic glottal

stop if preceded by any consonant other than a voiced obstruent (*tuSud > Jarai *təʔut* 'knee'); preceding voiced obstruents, on the other hand, were preglottalized (*buSuk > *ʔbuk* 'head hair', *baSuq > *ʔbəw* 'smell, odor').

4. Proto-Minahasan *buʔuk 'hair' < PAN *buSuk 'head hair' suggests that PAN *S sometimes became PM *ʔ. However, PM *bow < *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 *buʔuk can be seen as a secondary development between derived sequences of like vowels.

5. Fortgens writes ʔ as ' , ŋ as ng and the palatal series j, š, y as dj, s, j respectively.

6. Written ^{-C}, as in *poloic* 'blood', and -in, as in *fuluin* '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 (Admiralties) *tahuy*, Fijian *davui* instead indicate *tambuRi 'conch trumpet'.

8. The apparent change *u > o in *laSud > *lowo* 'toward the sea', *Za(S)uq > *yawo* 'far' and *niuR > *nuo* 'coconut', on the other hand, evidently arose through breaking (to earlier *lawed, *jaweq, *niweR), followed by the regular change *e > o, as with *waSiR > *wair* > *wayeR* > *wayo* 'fresh water'.

9. But occasionally from *-ŋ: *ke(d)eŋ > *kohoin* 'stand'.

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

11. With secondary genitive prefix n- (cf. *ade:n-ade* 'chin, jaw', *ulu:n-ulu* 'head hair') and fossilized genitive suffix -in, as in *fulu-in* < *bulu 'body hair, feathers'.

12. Most Austronesian languages reflect *qali-Sepan, but both *qali- 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 Buruese *bahu-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)aw > *feta-n* 'sister (man speaking)', *banua > *fena* '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/

(= [nd]). Bimanese evidently did not participate in this merger (*tampak > dampa 'blunt', *qali-meCag > linta 'water leech', *panDan > fanda 'pandanus', *DeSem 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 *kěmpun*-an 'bladder', Futunan *kopu* 'fish stomach' Dempwolff (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, Penchal (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 [β]

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

20. Dyen (1953 and subsequent works) writes *qaCey, *ijun. The latter reconstruction (with initial vowel) is supported by such reflexes as Tagabili *qilun*, Tongan *ihu*, but *qijun evidently is needed to account for the initial consonant of Malay *hidun* in conjunction with the Proto-Ambon reconstruction, and perhaps for Palauan *ʔis* (*-ŋ > ø unexpl.) 'nose'.

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

SOURCES OF MATERIAL

Adonara : Fox n.d.	Lakalai : Chowning
Ana Kalang : Fox n.d.	(1973)
Atoni : Fox n.d.,	Long Anap : Blust n.d.
Blust n.d.	Long Wat : Blust n.d.
Baluan : Blust n.d.	Malay : Wilkinson (1959)
Bario Kelabit :	Maranao : McKaughan and
Blust n.d.	Macaraya (1967)
Bima : Blust n.d.	Nauna : Blust n.d.
Bintulu : Blust n.d.	Paiwan : Ferrell (1978)
Buruese : Devin (1969)	Palauan : McManus (1977)
Cebuano Bisayan :	Palue : Fox n.d.
Wolff (1972)	Penchal : Blust n.d.
Endeh : Fox n.d.	Proto-Ambon :
Fijian : Capell (1968)	Stresemann (1927)
Jarai : Blust n.d.	Proto-Minahasan:
Kambera : Fox n.d.	Sneddon (1976)
Kanowit : Blust n.d.	Proto-Oceanic :
Kemak : Fox n.d.	Grace (1969)
Kiput : Blust n.d.	Proto-South Sulawesi
Kodi : Fox n.d.	(Mills 1975)

Riung : Fox n.d.	Ternate : Fortgens
Roti : Fox n.d.	(1921)
Sika : Fox n.d.	Tetum : Fox n.d.
Tagabili : Reid (1971)	Tongan : Churchward
Tagalog :	(1959)
Panganiban (1966)	Western Bukidnon Manobo:
Tausug : Reid (1971)	Elkins (1968)

REFERENCES

- BLUST, Robert A. 1969. 'Some new Proto-Austronesian trisyllables'. *Oceanic Linguistics* 8:85-104.
- 1973. 'The origins of Bintulu ʔ, ʔ'. *BSOAS* 36:603-20.
- 1974a. *The Proto-North Sarawak vowel deletion hypothesis*. Unpublished PhD dissertation. Honolulu, University of Hawaii.
- 1974b. 'A Murik Vocabulary, with a note on the linguistic position of Murik.' *Sarawak Museum Journal* 43 (N.S.):153-89.
- 1977. 'The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report'. *Working Papers in Linguistics* 9.2:1-15. Honolulu, University of Hawaii.
- n.d. Fieldnotes on languages of Borneo, the Lesser Sunda islands and western Melanesia.
- CAPELL, Arthur. 1968. *A new Fijian dictionary (third edition)*. Suva, Government Printer.
- CHOWNING, Ann. 1973. Milke's 'New Guinea Cluster': the evidence from northwest New Britain. *Oceanic Linguistics* 12:189-243.
- CHURCHWARD, C.M. 1959. *Tongan dictionary*. London, Oxford University Press.
- CLERCQ, F.S.A. de. 1876. *Het Maleisch der Molukken*. Batavia.
- DEMPWOLFF, Otto. 1934-38. *Vergleichende Lautlehre des austronesischen Wortschatzes. Zeitschrift für Eingeborenen-Sprachen*, supplements 15 (1934), 17 (1937), 19 (1938).
- DEVIN, Chaumont. 1969. *Buruese-English dictionary*. Typescript.
- DYEN, Isidore. 1953. *The Proto-Malayo-Polynesian laryngeals*. Btimore, Linguistic Society of America.
- 1965. 'Formosan evidence for some new Proto-Austronesian phonemes'. *Lingua* 14:285-305.
- ELKINS, Richard E. 1968. 'Manobo-English dictionary'. *Oceanic Linguistics Special Publication no. 3*. Honolulu, University of Hawaii Press.
- ESSER, S.J. 1938. 'Talen'. *Atlas van Tropisch Nederland*. Amsterdam, Koninklijk Nederlandsch Aardrijkskundig Genootschap.
- FERRELL, Raleigh. 1978. *Paiwan dictionary* (unpublished English draft translation of French language edition).
- FORTGENS, J. 1921. *Bijdrage tot de kennis van het Soboyo (eiland Taliabo, Soelagroep)*. The Hague, Martinus Nijhoff.
- FOX, James J. n.d. Comparative vocabularies of languages of the Lesser Sunda islands. Ms.

- GRACE, George W. 1969. 'A Proto-Oceanic finder list'. *Working Papers in Linguistics* 1.2: 39-84.
- HOËVELL, G.W.W.C. van. 1876. *Vocabularium van vreemde woorden voorkomende in het Ambonsch-Maleisch*. Dordrecht, Blussé & van Braam.
- MCKAUGHAN, Howard P. and Batua A. Macaraya. 1967. *A Maranao dictionary*. Honolulu, University of Hawaii Press.
- MCMANUS, Edwin G. 1977. *Palauan-English dictionary*. Edited and expanded by Lewis S. Josephs, with the assistance of Masa-aki Emesiochel. Honolulu, University Press of Hawaii.
- MILLS, Roger F. 1975. *Proto-South Sulawesi and Proto-Austronesian phonology*. Unpublished PhD dissertation. Ann Arbor, University of Michigan.
- PANGANIBAN, Jose Villa. 1966. *Talahuluganang Philipino-Ingles*. Manila, Government Printing Office.
- REID, Lawrence A., ed. 1971. 'Philippine minor languages: word lists and phonologies'. *Oceanic Linguistics Special Publication* no. 8. Honolulu, University of Hawaii Press.
- SNEDDON, J.N. 1976. *Proto-Minahasan: phonology, morphology and word list*. Typescript.
- STRESEMANN, Erwin. 1927. 'Die Lauterscheidungen in den ambonischen Sprachen'. *Zeitschrift für Eingeborenen-Sprachen*, supplement 10. Berlin, Dietrich Reimer.
- UHLENBECK, E.M. 1971. 'Indonesia and Malaysia'. Pp. 55-111. Thomas A. Sebeok, ed., *Current Trends in Linguistics*, vol. 8. The Hague, Mouton & Co.
- WALLACE, Alfred R. 1869. *The Malay archipelago*. 2 vols. London, Macmillan.
- WILKINSON, R.J. 1959. *A Malay-English dictionary* (romanised). London, Macmillan & Co.
- WOLFF, John U. 1972. 'A dictionary of Cebuano Visayan'. *Philippine Journal of Linguistics Special Monograph* no. 4. Manila, Linguistic Society of the Philippines.

* * *