THE PEPET IN SAMA-BAJAW

A. KEMP PALLESEN

1. TOPIC OF DISCUSSION

PAN *ə, the so-called 'pepet' vowel of Proto-Austronesian, described by Conant (1912) as the "original indifferent vowel... colorless and indefinite in pronunciation", is notable for the variety of its realisations in the daughter languages. This variety is described for Sama-Bajaw[1], a sub-grouping of Austronesian, where it finds especially rich expression. Some explanatory hypotheses are proposed for the splitting of PAN *ə into three Proto Sama-Bajaw phonemes.

2. SAMA-BAJAW AS A DISTINCT SUB-GROUPING

Sama-Bajaw languages occur in greatest number and diversity along the Sulu Archipelago (Philippines) and up into Sibuguey Gulf on the western end of Mindanao. Other languages of the group are Abaknon (AB)[2] in the

Central Philippines; Sabah Bajau (NB) on the West Coast of Sabah,
Malaysia; and a number of closely related but geographically dispersed
groups (IB) in Indonesia - in Sulawesi, the Moluccas, and the Lesser
Sunda Islands.

The position of the Sama-Bajaw subgrouping within Austronesian has
not yet been satisfactorily determined. It is a Hesperonesian sub-group,
but the data do not clearly support the generally accepted assignment to
the Philippine Hesion[3]. Central Sulu (CS) languages share 31% of
cognates with both Malay and Cebuano on a 372 word list[4], and 38% with
Siocon Subanon, a Meso-Philippine language of western Mindanao. Sabah
Bajau (NB) shares 37% with Malay and 32% with Siocon Subanon. (The 54%
which Central Sulu languages share with Tausug, a displaced Southern
Bisayan language spoken in Central Sulu, is due to considerable linguis-
tic convergence.) The sub-group is atypical, though perhaps not unique
(if Sangir is a member of the Philippine Hesion), in having *h as the
regular reflex of PAN *R.

Within the sub-grouping Sama-Bajaw is divided into six sections on
the basis of cognacity, syntactic typology, and reflexes of PSB phonemes.
The dividing lines of these sections correspond to discontinuities in
chains of mutual intelligibility. Since it is not important for the
topic of this paper, this internal subdivision is not discussed in fur-
ther detail, beyond noting that apart from relationships with (AB),
which shows a considerable degree of convergence with Central Philippine
languages, the lowest internal Sama-Bajaw cognate relationship is be-
tween (CS) and Indonesian Bajaw (IB), with 54%.

3. THE DATA

Accurately recorded data based on the Elkins word list, and consider-
ably augmented in several cases, are available for all sections of Sama-
Bajaw and form the basis of this study. The notation is essentially
phonemic unless otherwise indicated. The Proto Sama-Bajaw forms cited
are from a corpus of about 2500 reconstructions currently in preparation.

4. NORMAL PROTO SAMA-BAJAW REFLEXES OF PAN *ə

Sama-Bajaw languages exemplify Conant's[5] 'law of gemination'. Gemina-
tion in SB is the result of two processes: (1) the gemination of a
single medial consonant, and (2) the assimilation of the first consonant
of a heterorganic cluster formed by the reduplication of a monosyllable.
Non-geminate homorganic clusters are formed by the assimilation of an
infixed *ŋ[6] to the point of articulation of the following consonant.

A following homorganic cluster, geminate or otherwise, is the essen-
tial environment for the reflection of PAN *ə as *ə[7]. Elsewhere, i.e.
where this environmental condition is not met, the reflex of PAN *ə is
*ʉ[8] or *a. PSB is thus like Malagasy in having diverse reflexes of PAN
*ə in penult and ultima, but it does not conform to Conant's[5] observa-
tion that "the Philippines languages, wherever they show this double
vocalism, have a in the penult". The phonemic status of *ə and *ʉ in
PSB is discussed in Section 8.

4.1 The three PSB reflexes of PAN *ə are now presented and exemplified
for the penultimate and ultimate syllables, i.e. for the dissyllabic
words which comprise the great proportion of the Proto-Austronesian
corpus. The proposal of this paper is that the differentiation of PAN
*ə into the PSB reflexes *ə, *ʉ and *a is most satisfactorily explained
by positing contrastive stress in the dialect of PAN[9] from which PSB
most immediately derives, at least in those words which contain PAN *ə.
This stress hypothesis is discussed in Section 6.

 *ə is the reflex of unstressed PAN *ə in the penult; *a is the reflex
of unstressed PAN *ə in the ultima, when the vowel of the penult is
other than PAN *ə; *ʉ is the reflex of PAN *ə elsewhere, i.e. of stres-
sed PAN *ə in penult and ultima, and of unstressed PAN *ə in the ultima
when the vowel of the penult is also PAN *ə:

 PAN *Rəbáh 'to throw down, destroy' > *həbbaq 'to topple'; PAN
*səpsép 'to suck' > *səssʉp; PAN *lípət 'to fold' > *lipat 'to fold in
two'; PAN *léŋah 'kind of plant (sesame)' > (*)lʉŋa[10] 'Sesamum orientale';
PAN *kémkəm 'to keep close' > (*)kʉmkʉm 'to hold in closed fist'; PAN(B)
*SéRə(Ct) 'constriction, to tighten' > (*)hʉgʉt[10] 'to be secure'; PAN
*hajéŋ 'charcoal' > *arʉŋ; PAN *lə(ŋ)béŋ 'grave, ditch' > *ləbbʉŋ 'to
bury'.

4.2 Lack of data makes it impossible to give an adequate description of
the PSB reflexes of PAN *ə in the antepenultimate syllable. The one PAN
example which has a PSB reflex is unsatisfactory because of the ambi-
guity of the Ja. and Ml. evidence for PAN *ə[11]:

 PAN *sərampaŋ 'sprout, tine' (the citations in VLAW suggest 'fish
spear' as a more satisfactory gloss) > *sahapaŋ 'multi-tined fish spear'.
There are numerous comparisons, however, between Malay and Sama-Bajaw
which show that SB regularly has *a where Ml. has ə in the antepenult
(the possible loan status of some of the following forms does not affect
the point of the relationship between the vocalisms):

 PSB *jambatan 'bridge, pier' : Ml. jəmbatan, jambatan 'bridge';
*kəranjaŋ 'open-weave basket' < PAN *karanzaŋ : Ml. kəranjaŋ 'basket';

(*)salassay < *(saləssay) *'to settle a dispute'* : Ml. saləsay *'to solve, settle, complete'*. It seems probable that the correspondence of PSB *a and Ml. ə in the antepenult reflects PAN *a.

Although the PAN > PSB evidence is sparse, the internal evidence of Sama-Bajaw supports the reconstruction of five vowels in the antepenultimate syllable: *a, *i, *u, *ə, *ʉ. Three of these are of interest here[12]; *a is included because of the partial overlap of its reflexes with those of *ə and *ʉ.

4.21 *a is posited for the correspondence: e (occasionally a in (YK), ə (occasionally a) in (NB)[13], a in other languages:

   (*)bagunbun *'dust'*. (ZB, NS.2-3, CS.5, SS.4) bagumbun, (YK) begumbun, (NS.4-9, WS, CS.2, SS.1-4, Tsg.) bagunbun.

   *saruduŋ *'roofed area'*. (YK) seuduŋ *'extension of roof overhang'*, (CS.2) sauduŋ *'roofed, unwalled part of house'*, (NB.1) səruduŋ *'cooking area'*.

4.22 *ə is posited for the correspondence: a in (AB), i (with a few exceptions) in (ZB) and (NS-SS), e or i in (YK)[14], ə in (NB), and ə or i in (IB)[15]:

   *dəbuhiq *'last night'*. (AB) dabuqi, (YK) debuhiq, (NS-SS) dibuhiq, (NB) dəbuiq, (IB) dabuiq.

   *ənsəllan *'oil'*. (YK) isellan, insellan, (CS.2,5) ənsəllan, (CS.3, JM) insəllan, (NB) ənsəlan, [IB.9] ansəlan.

4.23 *ʉ is posited for the correspondence: a in (AB), u in (ZB, YK, NS-SS and JM), u or ə in (NB), a or i in (IB). Only the (AB) and (IB) reflexes distinguish this set from *u.

   *kʉhapoq *'grouper, rock cod'*. (YK, NS.3, WS.1, CS.2) kuhapoq, (SS.4) kuhapuq, (IB.1) kiapuq.

   *tʉqolan *'bone'*. (AB) taqulaŋ, (ZB.3, CS.2-3, SS) toqolaŋ (*tuqolaŋ).

The internal PSB evidence thus suggests that PAN *ə in the antepenult split into the reflexes *ə and *ʉ, a split analogous to that described for the penult. The primary conditioning factor in the antepenult, however, does not appear to be contrastive stress, as is posited for the penult (in Section 6). The small number of examples permits only the following tentative conclusion at this stage:

   In the antepenultimate syllable *ʉ is the reflex of PAN *ə before the consonants *h, *w and *l; *ə is the reflex of PAN *ə otherwise.

   *h and *w are consonants which do not geminate, and this limitation inhibits the reflex of PAN *ə as *ə in the penult (discussed in

Section 6.1). It is reasonable to suppose that the same inhibiting function is exercised in the antepenult also. *l follows *u in two of the six examples where *u is reconstructible in the antepenult, but there is no equivalent restriction on the gemination of *l. The inclusion of *l in the above distribution statement remains unexplained at present.

No gemination takes place following *ə in the antepenult, contrary to what might be expected from the constraints on the realisation of PAN *ə as *ə in the penult. There is apparently an overriding constraint which may be informally stated as "no word may have two sets of geminates". Such a constraint is current in (CS.2), a language which reflects penultimate *ə as ə: buhat *to serve refreshments* becomes (with stative prefix mag- and initial CV reduplication) magbubuhat *one who serves refreshments*. təttaq *to chop firewood* becomes, by the same affixation process, magtatəttaq *one who chops firewood*. Where a homorganic cluster (nasal plus stop) follows *ə in the antepenult the reflex ə does occur in (NS-SS), but not otherwise:

   *əntəllo *egg*. (AB) antallu, (NS.1-2, CS.2-3,6, SS.1,4-5) əntəllo, (IB.1-2,4) antillo.

   *dəqilaw *yesterday*. (AB) daqillaw, (ZB.1,3, NS, WS.2, CS, SS.1,3-4) diqilaw, (IB.1,5) dilaw.

5. IRREGULAR REFLEXES

Three hypotheses are considered as possible explanations for reflexes of proto-phonemes which are aberrant insofar as they do not conform to the rules stated in Section 4[15]: (1) that the words in question are loans, (2) that the reconstructed proto-forms are incorrect, or incomplete insofar as they fail to indicate phonological variants of the proto-form (doublets), and (3) that there are further developments within PSB, either regular processes of sound changes for which the conditioning factors must be discovered and described, or innovations (apparently unconditioned variants of protoforms). Each of these hypotheses has its legitimate function[16], but of the three borrowing is the least rigorous and the most open to misuse. Dyen's cautionary word (Dyen 1953) in this respect is worth keeping in mind:

> "If phonetically similar phonemes occur in the words of
> related languages in contact, and their correspondence does
> not agree with any formula that has been otherwise estab-
> lished, it is not safe to conclude that the words are
> related only by borrowing. ...it may be discovered that
> this (borrowing) explanation is insufficient"

5.1 Almost all the exceptions to the rules given in Section 4 are those
where PAN *ə > *a, a situation similar to that noted by Llamzon (1966)
for Tagalog. In this section the general problem of *a reflexes of PAN
*ə is considered, for ultima as well as penult, since there are plainly
possible alternatives to the interpretation that *a in the ultima is the
reflex of unstressed PAN *ə.

5.11 Malay is an obvious possible source for SB loans; it has exerted
considerable influence on Sulu languages for many centuries, and it
reflects PAN *ə as a in the penult. There is, furthermore, evidence
that other Philippine languages replace the penultimate or antepenultimate
ə of Ml. loans with a (a replacement noted by Dyen (1953) for Ngadju
Dayak loans from the same source)[17]. In the following examples from
non-Sama-Bajaw languages the supposition that we are dealing with loan
words is supported by the occurrence of non-inherited features in addi-
tion to the replacement of PAN *ə with a, i.e. by the presence of a
Malay prefix or of a second irregular prefix of a PAN phoneme:
    Mar. parampuan *woman* : Ml. pərəmpuan; Tsg. jarat *noose* : Ml.
jərat (the expected Tsg. reflex of PAN *zərat is (dj)ulat). SB lan-
guages show the same pattern: ·
    CS.3 balbaŋsa *of noble birth* : Ml. bərbaŋsa.
    The foregoing evidence justifies consideration of the possibility of
borrowing from Malay as an explanation for PAN *ə > *a irregularities.
It also suggests an alternative source for Tg. katám *carpenter's plane*
(if it is in fact a borrowing), for which Llamzon[18] proposed Kapampangan
as a plausible source[19]. A Malay source would also explain the presence
of katam in Sulu languages.
    Such a hypothesis would be strengthened if a Philippine language were
to exhibit the expected regular reflex of PAN *ə in PAN *kətəm (e.g. Tg.
kitím, CB kutum or PSB (*kutum/kəttum)). This is the case with PSB
(*)ganap and *gənnap *to become complete* (Tg. ganap and Ml. gənap) for
which VLAW gives only PAN *gənəp as a possible source. Alongside the
PSB doublet, which may be a borrowing, we have the CS.2 form kənnop *to
add to* < *(kənnup). Apart from the devoicing of the initial consonant
this is a regular reflex of PAN *gənép.
    There is, however, no *proof* of the borrowing hypothesis for such
problem cases. The apparently irregular form may be a regularly re-
flected doublet of the proto-language or of some meso-language. As more
data show the same 'irregular' reflex occurring in languages not in
contact, and as such forms prove to be unambiguously reconstructible for
the proto-languages of distinct sub-groupings, a doublet hypothesis
gains strength. A borrowing hypothesis, while not ruled out, requires
more substantive evidence – or greater ingenuity.

5.12 The problem of PAN *ə > *a is one for which alternatives (2) and
(3) must be thoroughly investigated before a borrowing explanation is
entertained. Alternative (2) - doublets in a proto-language - is
supported by a significant number of PAN doublets which show an alterna-
tion between PAN *ə and PAN *a in both penult and ultima:

PAN *basuR/*bəsuR 'satisfied'; PAN *basih/*bəsih 'iron'; PAN *dampul/
*dəmpul 'adhesive'; PAN *kaRat 'to bite' / *kəRət 'to cut off'; PAN
*laŋkəp 'to be complete' / *raŋkap 'to be combined'; PAN *tanaq/*tanəq
'ground, earth'; PAN *zagzag/*zəgzəg 'to gain a firm footing'.

Of the corpus of PSB words which reflect a PAN form with an infixed
*ŋ a remarkably high proportion - over half - show the reflex *a for PAN
*ə in the penult. Among them are:

PAN *gəndaŋ 'kettledrum' > *gandaŋ 'drum; to beat time'; PAN *gəntih
'to stop, finish' > (*)hantiq 'to break a journey' (Ml. mənanti 'to wait
for', hənti 'to stop'); PAN *gənDit 'a girdle' > *kan(dj)it 'scarf;
sash'; PAN(B) *ləŋTik 'graceful curve' > (*)lantik (Tg., Mar. lantik);
PAN *təmpət[20] 'domicile' > (*)tampat 'site where psychic power is con-
centrated' (Mar. tampat 'tomb'); PAN *tə(n)tu 'certainly' > *tantu (Mar.
tanto, CB tantu 'to an extreme degree'; (cf. also PKr. tantu 'certainly'
and mutu tantu 'too, overly', and the suggested Port. origin for this
disputed reconstruction).

It can hardly be the case that there is any greater inherent probab-
ility that such forms will be borrowed from Malay; rather it is evidence
of the facility with which *ə and *a alternate in this environment, at
some stage in the development of PSB.

The *ə/*a alternation before an infixed *ŋ was evidently an estab-
lished pattern at the PAN level. Examples given below are those which
have a Sama-Bajaw reflex:

PAN *lamba(r)/*ləmba(r) 'material, "piece" as a numerical classifier'
> *lamba 'strand, piece'; PAN *ma(n)tah/*məntah 'unripe, raw' > *mataq;
PAN *sambah 'respect' / *səmbah 'reverence' > (CS.2) səmba 'to revere'.

5.13 One further set of data remains to be considered, where PAN *ə > *a
in the final syllable:

PAN *li(ŋ)kər 'to circle, wind' > (*)(ka)liŋkar 'to surround'; PAN
*lipət 'to fold' > *lipat 'to fold in two'; PAN *rapət 'joined together'
> (*)(lr)apat 'to be joined without any gap' (Mar. lapat 'airtight', Tg.
lapat 'fitted in'); PAN *saŋət 'excessive' > *saŋat 'intensity; sever-
ity'; PAN *tabəq 'to be fat' > (*)tabaq 'fat which adheres to flesh'
(CB tabaq 'pork fat', Mwa tabaq, Tg. tabáq 'to be fat').

It is proposed in Section **6** that PSB *a, when it corresponds to PAN *ə
in the ultima, is the regular reflex of PAN [ʌ], the unstressed allophone

of PAN *ə in the ultima. Such a hypothesis must contend with explana-
tions in terms of borrowing and doublets (the alternatives (1) and (2)
above). For the borrowing hypothesis it may be argued that the SB forms
in the above examples are exactly what would be expected if they were in
fact loans from Malay. One reflex at least, however, cannot be ex-
plained as a borrowing from Ml. (nor from a Philippine language which
reflects PAN *R as g):

    PAN *bárəq *ulcer, abscess* > *bahaq *swelling, a boil*.

As has been shown in 5.11 there is ample precedent from PAN for
positing the existence of doublets which exhibit the a/ə alternation in
the ultima. Examples like PAN *rapət above, for which diverse Philip-
pine languages reflect a in the ultima, also suggest the possibility of
doublets, at least at the level of a Proto-Philippine meso-language.

The conclusions indicated at this stage are: (a) that the presence of
undoubted Malay loans in SB and in other Philippine languages, and the
pattern of ə to a replacement in such words, support the possibility of
borrowing as an explanation; (b) that there is evidence from PAN in sup-
port of the hypothesis that some PAN *ə > *a forms are the regular re-
flex of previously unidentified PAN doublets; (c) that in view of the
high frequency of forms showing the a/ə alternation before a homorganic
nasal-stop cluster some at least of these must be inherited from a lan-
guage ancestral to PSB, and (d) that the data suggest that some of the a
forms are either PSB innovations, or the result of regular phonological
processes in the ultima. The arguments for the latter proposal are
presented in Section 6.3 and are not repeated here.

The choice between the competing hypotheses is not resolved. Each
case must be considered separately; but any borrowing hypothesis can be
sustained only by good evidence.

5.2 Irregularities other than PAN *ə > *a make up a small residue, all
of which is now considered. In the following material the expected re-
flex, indicated by [...], follows the actual PSB or daughter language
form:

    (*)bəkkʉs *[bak(aʉ)s], (CS.2) bekos [bak(ao)s] *to be marked by
transverse bands* : PAN(B) *bakəs *belt*. This is probably a PSB in-
novation, with *ə replacing the vowel of the penult[21]. The (CS.2) form
is a secondary development, perhaps via *(bekʉs).

    (C) (YK) dellet *[d(ei)llet] *noose* : PAN *dirət *rope* (cf. PAN *zərat
loop). Probably a PSB innovation with vowel replacement and stress
shift, from *(dərrʉt).

    *hadʉn *[həddan/hʉdan] *ladder* : PAN *həzan. A PSB innovation by
vowel metathesis.

*intaq *[entaq] 'to eat raw' : PAN *qə(n)taq 'to be raw'. A PSB in-
novation which is unusual with regard to the replacement of PAN *ə by
*i.

*kehet *[kuhut] 'to cut up; to slice' : PAN *kəRət 'to cut off'.
Probably a PSB innovation, without parallel other than the next example.

*keket *[kukut/kəkkut] 'to bite' : PAN *kəkət 'to hold fast'. Ho.
hehitrah (VLAW) 'grip, bite' confirms the semantic connection. This and
the preceding example display the regular (YK) reflex of the posited
*(kuhut) and *(kukut), but it is unlikely that the (YK) form was the
loan source for other PSB languages, in view of the wide spread of un-
ambiguous reflexes of the reconstructions given.

*kijut *[kudut/kəddut] 'small muscular movement; twitch' : PAN *kədut
'to smart' (VLAW gives Ja. kədut 'muscular jerkings'). It is not clear
that the PAN and PSB forms are directly related, though a doublet may
be involved. Cf. PSB *kuddət/kəddut 'startled movement', and PAN(B)
*kə(zZ)u(Ct) 'pinch' (with Singhi kujet noted as metath.).

    (NS, CS) leŋog [doŋog/dəŋŋog] 'noisy, disturbed' : PAN *dəŋəg 'to
hear'. Probably not inherited from this PAN form. The CB data,
liŋug-liŋug 'to ignore by pretending not to hear' and duŋug 'to hear',
suggest a doublet, PAN *(liŋəg).

    (CS.2) lipaq [lopaq/ləppaq] 'to apply cosmetic paste' : PAN *ləpah
'paste'. Probably a reflex of a doublet with PAN *(lipaq) which VLAW
suggests somewhat uncertainly (cf. Ml. lepa 'plaster, mortar').

    (YK) luguŋ [leguŋ/legguŋ] 'thunder' : PAN(B) *ləguŋ 'booming sound'.
Probably the reflex of a doublet *(luguŋ). Cf. Mar. ruguŋ and WBM ruguŋ
'thunder', and the same vocalism in CB huguŋ 'produce a steady humming
sound' and in Tg. hugoŋ 'noise like thunder'.

*pəssil *[pusul/pəssul] 'to apply pressure with the fingers; to mas-
sage' : PAN(B) *pəcəl 'to squeeze in the hands'. Possibly a PSB innova-
tion, but cf. the final vowel of CB pisíl, písil 'to squeeze s.t. with
the tips of the fingers wide apart'.

    (*)sande(lr) *[sand(au)] 'to derive support from' : PAN *sanDəR 'lean
against'. Probably the reflex of a doublet. Cf. Batak sandir 'to lean
against a wall', and CB sandig 'to lean against'.

    (*)sumbaliq *[səmbəlliq] 'to slaughter by draining blood' : PAN
*səmbəlih 'to slaughter according to Moslem rites'. Probably a Philip-
pines innovation. Cf. Mar. sombaliq, WBM sumbaliq.

    (CS.2) tanog [tonog/tənnog] 'resonant' : PAN *tənəR 'voice'. Probably
a PSB or CS innovation. Cf. CB tunug 'widely known, resonant'; Tg.
tinig 'voice'; which are regular reflexes of the PAN form.

*timbak *[təmbak] 'to shoot a firearm' : PAN *təmbak. Either a bor-
rowing from Tg. (timbak) (unrecorded), or the reflex of a Philippines
innovation. Cf. Mar. timbak.

*tiŋʉr *[tʉŋʉr/təŋŋʉr] *'tanbark mangrove (Ceriops sp.)'* : PAN *təŋər. Probably a PSB innovation.

(*)tuman *[tʉman/təmman] *'certain; to fulfil'* : PAN *təman *'to be accustomed'*. Possibly a borrowing from CB tuman *'extreme'*, or from Mar. toman *'true; to fulfil'* (in which the expected Mar. reflex would be (təman)). Possibly also the reflex of a doublet *(tuman). Cf. the Ja. doublet: təmən *'to be faithful'*; tuman *'accustomed'*.

6. CONDITIONING FACTORS FOR THE PSB REFLEXES OF PAN *ə

Dempwolff[22] noted the occurrence of stress in the three primary languages of his study, but concluded that it was not possible to compare stress in a useful way. Zorc (1972), in a paper entitled 'Current and Proto Tagalic Stress', raises the question:

> "If... one posits stress as predictable and hence non-phonemic for Proto Malayo-Polynesian, how is it that it does become phonemic in so many daughter languages in the Philippines?"

In a concluding paragraph, having noted the need to posit both penultimate and ultimate stress for Proto Tagalic, and the existence of at least one minimal pair, he raises the further question:

> "Did Proto Tagalic inherit or innovate its ultimate stresses. A careful investigation of the lexicons of Tagalic and other Philippine languages is called for. As time progresses and areas of research increase we should be able to discover just how far back in time we can postulate phonemic stress."

It is here proposed that differential stress in PAN[9] was the primary conditioning factor for the splitting of PAN *ə into the PSB reflexes *ə, *ʉ and *a. According to this hypothesis, allophonic variations of PAN *ə existed as a result of contrastive stress, such that unstressed PAN *ə in the penult was manifested by the normative allophone [ə]; unstressed PAN *ə in the ultima was manifested by the lowered allophone [ʌ] when the vowel of the penult was other than PAN *ə, and by the raised allophone [ɨ][23] when the vowel of the penult was also PAN *ə. Stressed PAN *ə in both penult and ultima was manifested by the raised allophone [ɨ]. The distribution of the posited allophones corresponds to the distribution of the PSB reflexes as stated in Section 4.1, so that PAN [ə] > *ə; PAN [ʌ] > *a, and PAN [ɨ] > *ʉ, as in the following:

PAN		PSB	
*léŋah [lɨŋah] *'sesame'*	>	*lʉŋa	*'sesame'*
*gətáh [gətah] *'sap'*	>	*gəttaq	*'sticky sap'*
*téRəb [tɨRɨb] *'quantity, mass'*	>	(*)tʊgʊb[24]	*'heavily loaded'*
*ləŋén [ləŋɨn] *'forearm'*	>	*ləŋŋʉn	*'upper arm'*
*pútər [putʌr] *'to turn around'*	>	*putar	*'to revolve; to unwind'*

6.1 Stress is indeterminate for words where other factors supercede
stress as the primary conditioning factor. Two such cases are suggested
by the data.

6.11 Stress is indeterminate when PAN *ə in the penult is followed by a
consonant which does not geminate in PSB, e.g. PSB *h. The reflex of
PAN *ə as *ʉ is inhibited, since a following homorganic cluster is an
essential condition for this reflex in the penult:

      PAN *bəRas   'husked rice' > *bʉhas
      PAN *bəRsay[25] 'canoe paddle' > *(bʉhsay) > *busay
There are thus two processes by which *ʉ could have been derived from
PAN *ə in the above examples. If stress occurred on the penult the
posited PAN [ɨ] allophone of PAN *ə would have been reflected as *ʉ;
had stress occurred on the ultima the unstressed PAN [ə] allophone would
likewise have been reflected as *ʉ, due to the limitation just de-
scribed.

6.12 Stress is also indeterminate for forms with an infixed PAN *ŋ
following penultimate PAN *ə. The present data provides no evidence of
contrast between *ə and *ʉ in this environment. This appears to be more
than a data limitation, and the possible phonetic explanation is sug-
gested in Section 7.2. No stress has been assigned to examples such as
those given in 6.11 above, or to the following:

      PAN *kə(m)bar 'twin' > *kəmbar, (*)kambar; PAN *gəntiŋ 'thin' >
*gəntiŋ 'slim-waisted'.

6.2 The differentiation of stress carries a low semantic load:

PAN
*təŋáh	'middle'	> *təŋŋaq	'middle (range)'	
*téŋah	"	> *tʉŋaq	'middle (point)'	
*lí(ŋ)kər	'to circle, wind'	> (*)(ka)liŋkal	'to surround'	
*li(ŋ)kér	" " "	> *(liŋkʉl)	(CS.2) leŋkol 'halo about sun'	
*bə(n)téŋ	'stomach'	> *bəttʉŋ	'stomach, bulge, pregnant'	
*bé(n)təŋ	"	> (*)bʉtʉŋ[26]	'mature, green coconut'	
*ləpás	'loose, free'	> *ləppas	'emerge; pierce through'	
*lépas	" "	> (*)lʉpas	'to be lost'	
*(təgél)	('persistent')	> *təggʉl	'long period of time'	
*(tégəl)	(")	> *(tʉgʉl)	(CS.2) togol 'persist'	

      In some cases the reflexes of stress-differentiated PAN pairs show
no current semantic contrast in SB languages:

*kəbél	*invulnerable*	> *(kəbbʉl)	(YK)	kebbel
*kə́bəl	"	> *(kʉbbʉl)	(CS.2)	kobol
*huléj	*maggot*	> *(ulʉt)	(YK)	olet
*hú́ləj	"	> *(ulat)	(NS, CS)	ulat

6.3 As has been mentioned in Section 5.13 there are competing hypo-
theses for the proposal that PAN *ə in the ultima had stressed and un-
stressed allophones [ɨ] and [ʌ] which were reflected in PSB as *ʉ and
*a respectively. The following arguments are not conclusive, but they
warrant serious consideration of the proposal.

6.31 The parallelism between the split in the penult and the posited
split in the ultima is phonetically plausible. In both syllables the
phonetically higher reflex is that of the posited stressed allophone of
PAN *ə. The unstressed allophones and their PSB reflexes differ in
phonetic height, with the lower occurring in the ultimate syllable.

6.32 If two languages, not closely related and not in contact, share a
distinctive phonological phenomenon there is reason to seek an explana-
tion in similar phonological processes. Tagalog, for which the regular
reflex of PAN *ə is i, and Sama-Bajaw both show significant numbers of
words which exhibit ultimate a instead of the expected reflex of PAN *ə.
Llamzon[27] cites some doublets which parallel the SB data:
    batís, batás *to ford a creek* < PAN *batəs *boundary*; hilís, hilás
cut; lapit[28] *close, draw near*, lapat *fitted in* < PAN *rapət
joined together; sapit *arrived*, sapat *attached*.

One would expect, if doublets are the correct explanation for these a
reflexes in both languages, to find evidence of the same doublets in
other languages of the Central and Southern Philippines. This evidence
does not appear to be abundant in Cebuano Bisayan, but more data are
needed, both from CB and other languages such as Bikol, Maranao, Subanon
and Sangir.
    If *(baRaq) *ulcer, abscess* is reconstructed as a doublet (alongside
PAN *baRəq) on the evidence of Tg. bagáq *abscess* and PSB *bahaq *swel-
ling; boil*, we must suppose the persistence of both members of the
doublet down to the Proto Tagalic stage at least. Cf. CB baguq *disease
characterised by... swelling in the region of the stomach* (and also
CDu bagəq *to be swollen*).
    The persistence of such doublets is not unprecedented, to be sure,
but it is significant that the two languages which justify the posited
doublet are languages which exhibit similar alternations between a and
the regular reflex of PAN *ə. It thus seems reasonable to consider the

possibility that we are not dealing here with a doublet but with a
distinctive, shared phonological development.

6.33 The hypothesis of contrastive stress in the ultima is consistent
with the alternations displayed within PSB, not all of which have an
established PAN source. In the first example below the shift of stress
to the ultima is accompanied by reduction of original PAN *i in the
penult to *ə[29]:

PAN *lípət   *'to fold'   > *lipat   > *'to fold in two'*
     *(ləpét)[30]   > *ləppʉt > *'to fold up'*
     *(táhəp)   *'to winnow'* > *tahap
     *(tahép)   > *(tahʉp) > (AB) tahup

7. THE DEVELOPMENT OF GEMINATION IN PSB

   Universal penultimate stress is reconstructible for PSB. It is
therefore proposed that the posited contrastive stress of PAN became
non-contrastive in PSB. In all the PSB daughter languages which re-
flect it ə is a short vowel. It does not occur with the vowel lengthen-
ing which is the usual concomitant of stress in the daughter languages
(and presumably reconstructible as a feature of stress in PSB also).
It has been proposed in Section 6 that PSB *ə only occurred as the re-
flex of unstressed PAN *ə in the penult. It is further proposed that
with the shift of stress from the ultima to the penult the short vowel
was retained and that compensatory length was carried by the following
consonant, thus causing gemination. Thus PAN *bəkás *'inheritance'* be-
comes, with the regularisation of stress, PSB *bə́k:as *'trace of former
occupancy'*.

7.1 In (NB), the only languages which do not have universal penultimate
stress[31], stress falls on the ultima only when the vowel of the penult
is ə. In (NB), furthermore, the length of the medial geminate is sub-
stantially reduced, so that it no longer carries the length feature of
the stressed syllable. Stress thus reverts to the ultima. In the fol-
lowing examples the raised period indicates a very slight lengthening
of the consonant; rather more when the ə is initial:

PSB *əmmaq *'father'* > (NB) əmaq[32] [m·áq] and > (CS) əmmaq [m̂maq]; PAN
*bə(n)tis *'shin-bone'* > *bəttis *'leg'* > (NB) bətis [bət·ís] and > (CS)
bəttis [béttis].

(CS) languages display an extreme degree of transference of syllabicity
from stressed ə to the following consonant. The phonetics below are
simplified with respect to the allophones of the other segments:

PSB *ənnʉm *'six'* > (CS.2) ənnom [ńnom]; PSB *əmpit *'to hold a child (on the lap)'* > (CS) əmpit [ḿpit]; PSB (*)əggok *'to make a gulping motion'* > (CS.2) əgguk [ˈǵguk] (the initial g is slightly imploded); PSB (*)əttus *'disease which pocks the skin'* > (CS) əttus [ˈttus] (tongue tip is held in the stop position for the normal duration of a stressed vowel).

7.2 Where PSB reflects a PAN word in which PAN *ə in the penult oc-curred before an infixed *ŋ, only *ə occurred as the reflex of PAN *ə. This is not surprising; the nasal carried the compensatory length when stress occurred on the penult, and the normative vocalism of PAN *ə was maintained. This environment, it is presumed, conditioned the allo-phones of PAN *ə so that [ə] (rather than [ɨ]) was the allophone of stressed PAN *ə in the penult before a nasal-stop cluster. Since both stressed and unstressed PAN *ə in this environment have the reflex *ə it is not possible to determine stress.

8. PHONEMIC STATUS OF THE PSB REFLEXES *ə AND *ʉ

We now consider the question of the phonemic status in PSB of the PSB reflexes *ə and *ʉ of PAN *ə. It has been proposed in the previous section that contrastive PAN stress became non-contrastive in PSB. The contrastive function of stress in PAN was thus taken over by phonemic-ally contrastive vowels, as witnessed by the phonemic status of the reflexes of *ə and *ʉ in all daughter languages except (YK)[33]. PSB thus represents the completion of a transitional period during which contrastive stress plus sub-phonemic vocalic contrast gave place to non-contrastive stress and phonemic vowel contrast. The gemination of the medial consonants was a by-product of the change in the pattern of stress placement, and played only a secondary role in marking phonemic contrast.

9. PROCESSES BY WHICH ə OR ʉ IS INNOVATED IN SB

There are four processes by which ə or ʉ is developed from vowels other than PAN *ə. In some cases the development is reconstructible for PSB; in others it appears to have taken place subsequent to the sub-division of PSB.

9.1 The first of these innovations is based on the analogy of stress-conditioned reflexes of PAN *ə in a reduplicated monosyllable. The effect of contrastive stress on the PSB reflexes is displayed in the

following examples; with ultimate stress the heterorganic cluster of
reduplicated monosyllables is assimilated to a geminate and the vowel of
the penult is reflected as *ə:

PAN *kə́mkəm 'to keep close' > (*)kʉmkʉm 'to hold in closed fist';
PAN *səpsə́p 'to suck' > *səssʉp; PAN *kə́pkəp 'to grasp' > (*)kʉpkʉp,
but PAN *kəpkə́p > (*)kəkkʉp.

A similar process based on this pattern may operate on a reduplicated
monosyllable with a vowel other than *ə. The medial cluster becomes a
geminate, and the vowel of the penult is sometimes innovated as ə:

PSB *kabkab 'to fan' > *(kəkkab) > (YK) kekkab; (*)kutkut 'to
scratch' > *(kəkkut) > (ZB.3) kukkut, and (CS.2) kakkut (where the
vowel of the penult in both cases is a secondary development); *səssʉk
'narrow gap' and (CS.2) soksok < *(sʉksʉk) suggest a back formation on
the same analogy, although the (CS.2) form may be original, lost in
other languages to the innovating modification.

Assimilation of a heterorganic cluster does not, however, obligator-
ily require this vowel replacement:

(*)bugbug 'starch food cooked till mushy' > (YK) bʉbbug 'rice gruel'.

The alternation also occurs within a single daughter language, mar-
king minor semantic modification:

(CS.2) loklok 'to withdraw into a hole in order to avoid detection';
(CS.2) ləllok 'to be just below surface, as a nail driven into wood'.
(The CS.2 vowel o is a merger of two PSB proto-vowels, *ʉ and *o, so
that the above forms may be reflections of PSB *(lʉklʉk).)

9.2   The second innovation is based on the alternation between *ə and
some other vowel (not a) in the penult, with the regular gemination of
the medial consonant after the innovated ə. There are precedents for
this pattern in PAN:

PAN *kumuR, *kəmuR 'to move to and fro in the mouth'.
In some cases the innovation is of PSB provenance:

(*)bəttad 'to stretch out; to set s.t. out' and (NS, CS) bitad 'to
pull taut'. Cf. PAN(B) *bi(ŋ)CtT)ad 'spread out'.

(*)bulud 'mountain' (in NS.7-10, WS, CS, SS.1-3) and *bəllud 'moun-
tain' (in ZB, NS.2-3, SS.4-5, NB, IB).

The alternation is also exploited to mark semantic modifications
within a single daughter language, apparently as a fairly recent innova-
tion. A few examples are presented from CS.2, where the alternation is
abundant although no longer an active process:

babas 'to abate, as wind'   bəbbas 'to abate, of swelling'
bakat 'cut in flesh'   bəkkat 'to break a line'
kilat 'lightning; to flash'   kəllat 'to show light, of sun'

kulit *'outer layer, skin'*   kəllit *'leathery'*
tigaŋ *'dried out'*   təggaŋ *'barren womb'*
tubud *'to well up, as water'*   təbbud *'to billow up, as smoke'*

9.3 PAN *a in the ultima before final PAN *h is regularly reflected as
PSB *ʉ[34]:

   PAN *kimah *'giant-shell'* > *kimʉ *'giant clam'*; PAN *duwah *'two'* >
*duwʉ; PAN *taliŋah *'ear'* > *taliŋʉ.

9.4 PAN *a in the penult sometimes becomes *ʉ when the vowel of the
ultima is the *ʉ reflex of PAN *ə, and the medial consonant is *l or
*r:

   PAN *balə́s *'to retaliate'* > *balʉs or *bʉlʉs *'to reciprocate an ac-
tion'*; PAN *tazəm *'sharp'* > *tarʉm or *tʉrʉm.

10. REFLEXES OF PSB *ə AND *ʉ IN DAUGHTER LANGUAGES

   PSB had seven phonemic vowels after the regularisation of stress
placement on the penultimate syllable:
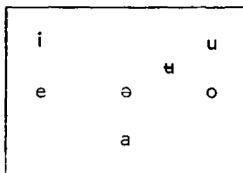
Two daughter languages, (ZB.1, WS.1) have retained this seven-vowel
system. (ZB.2-3, NS, WS.2, CS, SS, JM, NB, IB) have all reduced the
PSB proto-system to a six-vowel system by merging *ʉ with another
phonemic vowel (with either *a, *o or *u). Some (NS) and (IB) languages
are in a stage of transition between six and five vowels, as ə, a
phonetically unstable segment, merges diversely with other phonemic
vowels. (YK) has a five-vowel system resulting from the merger of both
*ə and *ʉ with *e, and (AB) has a three-vowel system resulting from the
merger of *o and *ʉ with *u, of *ə with *a, and of *e with *i.

10.1 The reflexes of PSB *ə and *ʉ are now presented for each of the
daughter languages (V_1 refers to the vowel of the penult; V_2 to the
vowel of the ultima):

10.11 In the antepenult *ə > a in (AB); > i in (ZB, NS-SS, JM),
usually; > e or i in (YK); > u or ə in (NB); > a or i in (IB):

*təmboroq *'finger'*. (AB) tamburuq *'hand'*, (ZB) tomboq < *(tomboloq)
< *(təmboloq), (YK) timbooq *'index finger'*, (NB) təmboro *'index finger'*.

10.12 In the penult *ə > a in (AB); > ə in (ZB, NS-SS, JM, NB, IB),
with secondary modifications as indicated for (ZB, NS, IB); > e in
(YK):

*bəttʉŋ *'stomach; pregnant'*. (AB) battuŋ, (ZB.3, NS.2, CS, SS.2,
JM) bəttoŋ, (WS.1) bəttʉŋ, (WS.2) bəttuŋ, (NB.1) bətoŋ, (IB.8) bətah.

*əmmaq *'father'*. (ZB.1, NS-SS, JM) əmmaq, (NB) əmaq, (IB.1) əmmaq
'mother'.

10.121 In (ZB.2, NS, IB) ə is an unstable segment, and secondary pro-
cesses are still continuing, by which ə is replaced or merged with an
existing phoneme. In (ZB.2) and (NS) two processes operate and compete
with each other and with the inherited reflex ə. ə usually persists
when it is initial in a word; elsewhere ə is sometimes (apparently un-
predictably) replaced by e (especially in NS.6), or ə assimilates to
V_2 (especially in ZB.2, NS.1,9-10):

*ləŋŋʉn *'upper arm'*. (ZB.1, WS.1) ləŋŋʉn, (ZB.2, NS.2-3,7-10, CS.4,
SS.4) loŋŋon, (ZB.3, CS.1-3,6, SS.1,3) ləŋŋon, (YK) leŋŋen, (WS.2)
ləŋŋun.

*bəlla *'to cook'*. (ZB.1, NS.3,6-7, WS.2, CS.2-6, SS.1,4, JM) bəlla,
(ZB.2-3, NS.10) balla, (YK, NS.3,5,8) bella.

10.122 In (IB) ə usually persists when it is initial in a word; else-
where it is usually replaced by i when V_2 is *ʉ or *a, and sometimes
replaced by i or u in other environments:

*bəlla *'to cook'*. (IB.1) billa.

*bəlli *'to buy'*. (IB.1) bəlli, (IB.2) billi.

*bənnʉd *'muscular cramp'*. (IB.1) binnaq.

*bərrat *'heavy'*. (IB.1) birraq, (IB.4) bərraq.

*dəppʉ *'fathom'*. (IB.1) duppa, [IB.9] dəpe.

*əllum *'to be alive'*. (IB.1) əllumaŋ, əlloŋ.

*kəlloŋ *'neck'*. (IB.1) killoh, (IB.2) kulloq, [IB.9] kəlo.

10.21 In the antepenult *ʉ > a in (AB); > u in (ZB, YK, NS-SS, JM); >
u or ə in (NB); > a or i in (IB):

*dʉwaqi *'to descend'*. (AB) dawaqi, duwaqi, (ZB, YK, NS-SS, JM, NB)
duaqi.

*kʉhapoq *'grouper; rock cod'*. (YK, NS.3, CS.2) kuhapoq, (IB.1)
kiapuq.

10.22 In the penult *ʉ > u in (AB, WS.2); > ʉ in (ZB.1, WS.1); > o in (ZB.2-3, NS, CS, SS, JM, NB); > e in (YK); > o in (IB) when both V_1 and V_2 are *ʉ (there are no data for *ʉ in penult and a different vowel in the ultima):

   *bʉtaŋ *'to place; to install'*. (AB) butaŋ, (WS.1) bʉtaŋ, (CS.2-3) botaŋ.

   *dʉhʉŋ *'to desist from doing s.t.'*. (YK) deheŋ, (NS.1, CS) dohoŋ, (IB.1) doqoh *'to abate, of a storm'*.

   *bʉlʉs *'to reciprocate an action'*. (AB) bulus *'to take revenge'*, (NS.1, NB.1) boos *'to borrow'*, (WS.1) bʉʉs *'to borrow'*. (NB.4) bohos < *(boos) *'to take revenge'*.

10.23 In the ultima, in an open syllable, *ʉ > a in (AB, ZB, NS-SS, JM, IB); > e in (YK); > o in (NB):

   *hayʉ *'big'*. (AB, ZB.1-2) haya, (YK) haje, (CS, SS) hea, (NB) oyo.

   *sawʉ *'snake'*. (AB) sawa, (ZB, NS-SS, IB.1) soa, (YK) sawe, (JM, NB) soo.

   In the ultima, in a closed syllable, *ʉ > u in (AB, WS.2); > ʉ in (ZB.1, WS.1); > o in (ZB.2-3, NS-SS, JM); > e in (YK); > a in (IB):

   *anʉm *'to weave a mat'*. (ZB.1) anʉm, (ZB.3, NS.3, CS.2-3, SS.2-4) anom, (ZB.2, WS.2) anum, (NB.1-2) ŋonom, (IB.1,4-5) anaŋ.

   *dakʉp *'to seize'*. (CS.3) dakop *'to abduct a woman'*, (NB.1-2) dəkop, dokop, (IB.9) dakap.

   *ənnʉm *'six'*. (AB) annum, (ZB.1, WS.1) ənnʉm, (ZB.2, NS.3,5,10) onnom, (ZB.3, NS.2,7-9, WS.2, CS, SS, JM) ənnom, (NB.1,3-4) ənom, (IB.1,5) ənnaŋ.

10.24 In (NB) a secondary replacement of the o reflex of *ʉ is taking place, such that o > a.[34] There is considerable alternation between a and o in (NB), in both directions, possibly under the influence of cognates with Malay which show a regular correspondence between a and o in the ultima.

   *kəttʉp *'to slice, as meat'*. (NB.1) ŋətap *'to chop small'*.

   *ənnʉm *'six'*. (NB.1,3-4) ənom, (NB.1-2) ənam, Ml. ənam.

   *(h)iŋkʉt *'to tie'*. (NB.2) eŋkot, (NB.1) iŋkaq, əŋkaq, Ml. ikat.

   *sərrʉt *'tightly wedged'*. (NB.1) sorot, (NB.2) sərat, Ml. sərat *'jammed'*.

   *səssʉr *'to compel s.t. to move'*. (NB.1) səsar, səsor *'to drive away'*, Ml. səsar *'to push s.t. aside'*. Cf. also Ml. pərtama *'first'*, borrowed as (NB.1) pərtomo.

11. CONCLUSION

In order to faciliate the description of the Proto Sama-Bajaw re-
flexes of PAN *ǝ - an interesting and rather complex set of data - we
have proposed that contrastive stress was a feature of a language
ancestral to PSB, though not necessarily of Proto-Austronesian. Most of
the SB data are explicable in terms of this hypothesis; a borrowing or
doublet hypothesis is indicated for only a small proportion of the
corpus.

However, satisfactory though the hypothesis appears to be for the
PSB data, it needs to be tested against the data of other PAN sub-
groupings. Of particular value, it seems, are the data of those lan-
guages in which PAN *ǝ splits into two or more phonemes, and those lan-
guages which exhibit secondary developments such as gemination of the
medial consonant after penultimate PAN *ǝ. Another useful set of data
may be provided by languages which, like Sama-Bajaw, have phonemic
contrast between u and o, and between i and e. Such splits may be due
in part to contrastive reflexes of stress-conditioned allophonic varia-
tions of PAN *u and PAN *i.

NOTES

1. Sama is a widespread autonym (from Sulu, Sabah and Indonesia), and
is reconstructible as the proto-form by which speakers of the language
referred to themselves. Bajaw is the Malay designation used throughout
Malaysia and Indonesia, and for one group in the Southern Philippines.
The compound name of the title has been chosen to facilitate identifica-
tion, and at the same time to retain the autonym.

   Other names which occur currently and in the literature are Orang
Laut, Samal (the usual external designation for Sama in the Philippines),
Luwaqan (nomadic Philippines Sama) and "Sea Gypsies".

2. The following conventions have been followed: Proto-Austronesian
reconstructions have been written in Dyen's orthography with the excep-
tion that the pepet is written as *ə, and the symbol e retained to
indicate the mid-front vowel [e] which is phonemic in Sama-Bajaw.
ʉ indicates a high back-central unrounded vocoid; ɨ indicates a high
central unrounded vocoid. I have followed Blust in writing the PAN in-
fixed nasal as *ŋ in his reconstructions, but have not modified the
VLAW material. Malay citations have been modified by writing the non-
syllabic high vocoids as -w and -y, the digraph ng as ŋ, and e as ə
where it represents the mid central vowel.

   Since we are dealing with two levels of reconstruction the convention
is followed of prefixing all Proto-Austronesian forms with the abbrevia-
tion PAN, and only prefixing Proto Sama-Bajaw forms with PSB when there
is a possibility of confusion. PSB forms well-attested in Sulu lan-
guages, but unattested outside them, are indicated by (*), and hypo-
thetical forms by *(...); thus PAN *bélah *to split* > *(bʉlaq) > (CS.2)
bolaq *bamboo lath*. Primary stress is marked by a grave accent over
the vowel of the stressed syllable. [..] enclosing a language code
indicates that there is some uncertainty about the phonetic interpreta-
tion of some part of the following citation. Language citations appear

bold type and are identified by a preceding letter code in the case of
Sama-Bajaw languages, or by an abbreviation.

Sources and abbreviations are as follows:

PAN: Proto-Austronesian (Dempwolff); PAN(B): Proto-Austronesian
(Blust); PSB: Proto Sama-Bajaw; Batak (Rodda); CB: Cebuano Bisayan
(Wolff); CDu: Casiguran Dumagat (Headland & Headland); Ho.: Hova (VLAW);
Ja.: Javanese; Mar.: Maranao (McKaughan and Macaraya); Ml.: Malay,
Indonesian (Echols and Shadily); Mwa.: Mamanwa (Miller and Miller);
Minahasa and Togian Bajaw (Adriani); PKr.: Papia Kristang (Hancock);
Proto East Mindanao - a subgrouping of Southern Bisayan which includes
Mamanwa and Tausug (author's fieldnotes); SB: Sama-Bajaw daughter lan-
guages, collectively considered; Skr.: Sanskrit (Gonda); Tg.: Tagalog
(English); Tsg.: Tausug (Taqusuug) (Ashley); VLAW: *Vergleichende Lautlehre
des Austronesischen Wortschatzes*; WBM: Western Bukidnon Manobo (Elkins).

Sama-Bajaw daughter languages have been assigned letter codes, with
following letters indicating dialect or sub-dialect variations within
each coded section. The expansion of the code is indicated as a
mnemonic aid: (AB-Abaknon): Abaknon or Capulenyo (Capul Island, Central
Philippines, between Masbate and Northern Samar); (ZB-Zamboanga):
Batuan, Lutangan and Sibukuq Sama (Sibuguey Gulf and Zamboanga Pen-
insula, Mindanao); (YK-Yakan): Yakan (Basilan Island); (NS-Northern
Sulu): Baangingiq and related dialects (northeastern Sulu, with migrant
groups in other parts of the Philippines); (WS-Western Sulu): Pangutaran
and Soang Bunaq (islands about 35 miles north-west of Jolo Island in
Sulu); (CS-Central Sulu): Sama Dilaut (nomadic and semi-nomadic, from
eastern Sabah to Zamboanga) and several related dialects spoken by
sedentary groups (central Sulu); (SS-Southern Sulu): Tanduq Baas,
Simunul and Sibutuq (southwestern Sulu and eastern Sabah); (JM-Jama
Mapun): Mapun (Cagayan Sulu Island, north of Sandakan, Sabah); (NB-North
Borneo): West Coast Sabah Bajaw (Labuan Island and the coastal fringe
from Papar to Banggi); (IB-Indonesian Bajaw): Indonesian Bajaw (Indone-
sian Timor, Roti Island, Kajoa in the western Halmaheras, and various
places around the Sulawesi coast). (CS.2) indicates Sama Dilaut of the
Siasi Lagoon area, Siasi, Central Sulu, the language with which the
writer is best acquainted. (NS-SS) includes (NS, WS, CS, SS), a single
section characterised by a chain of mutual intelligibility. Sulu lan-
guages include (ZB, YK, NS-SS, JM).

The following acknowledgements are made of unpublished sources:
James Fox (Roti data (IB.3)), Charles Frake and Carol Molony (Kajoa,
Halmahera data (IB.2)), Gerard Rixhon (Sulawesi data (IB.4-6)), and
various members of the Summer Institute of Linguistics, Philippines,
especially Dietlinde Behrens and Janet Pack (YK), Jeanne and Helen Miller

(Mamanwa), Charles Walton and William Hall (various Sama word lists)
and Siocon Subanon (Hall).

3. Dyen (1965) assigns one language of the sub-group, Yakan, to the
Philippine Hesion as a first order sub-grouping, and indicates the
problematic nature of the assignment by a double asterisk.

4. The word list is a list adapted by Richard E. Elkins from the
Swadesh list for the use of the field workers of the Summer Institute of
Linguistics in the Philippines. A comparison of cognate counts between
the Swadesh and the Elkins list indicates a difference of ± 4%. The
Elkins list forms the basis of *Philippine Minor Languages: Word Lists
and Phonologies* (Reid, ed.).

5. Conant op. cit.

6. Blust's (1970) assumption has been followed that the data are best
explained by positing a single segment *ŋ which was sometimes infixed
preconsonantally in PAN, but Dempwolff's reconstructions have not been
rewritten in accordance with this assumption.

7. A minor exception to this statement occurs in the antepenultimate
syllable, where PSB *ə occurs without a following homorganic cluster.
This case is discussed in Section 4.2.

8. Since the daughter languages show a wide diversity of reflexes of
PSB *ʉ (e, a, o, u) there was no clear phonetic identity for this proto-
form. ʉ was chosen for pragmatic reasons: (1) it is a unique symbol
for the correspondence set, and (2) in the only two languages which
maintain the seven-vowel phonemic system of PSB the current phonetic
realisation of this proto-phoneme is [ʉ], a high back-central unrounded
vocoid. In view of the other reflexes of PSB *ʉ no assumption is made
about the rounding of the proto-phoneme.

9. Since Dempwolff's and Blust's reconstructions are based essentially
on languages of Dyen's Hesperonesian Linkage, the proto-language of the
forms indicated as PAN in this paper are more correctly of Proto-
Hesperonesian provenance. The ancestor of PSB for which contrastive
stress is posited is thus Proto-Hesperonesian or some lower order proto-
language. More work is needed before such mid-level groupings can be
satisfactorily identified; the reconstruction of reliable forms for low
level proto-languages is a step in this direction. The label PAN in
this paper, when it refers to the stress hypothesis, should be read as
shorthand for "PAN or some lower order proto-language of Proto-
Hesperonesian".

10. These examples, because of their relatively limited range of oc-
currence, are susceptible to interpretation as loans from a language
which has ə or ʉ reflexes of PAN *ə in both penult and ultima, i.e. a
language such as Proto East Mindanao and its daughter language Tausug.
Upland Tausug maintains the contrast between ʉ and u, lost in Lowland
Tausug. There are, however, other data showing this vocalism which
have good PSB credentials though without an identified PAN source:
*dʉhʉŋ 'to desist from doing something'.

11. Ml. and Ja. ə in the antepenult may derive from either PAN *ə or
*a, an ambiguity which Blust indicates where appropriate, in his Ad-
denda. Malay regularly, though not invariably, replaces original a with
ə in the antepenult: mərdeka 'free' (Skr. maradeka); kəpala 'head'
(Skr. kapāla 'skull'); fatwa (Ar.) 'instructions' and pətua '(religious)
advice', where the Arabic word has been modified in the second form to
conform with Ml. phonology.

12. *u and *i do not constitute a problem since their reflexes are in-
variable in all PSB languages (except (NB), which has extensive vowel
reduction in the antepenult).

13. (YK) and (NB) are similar to Malay in having a rule by which
phonemic a is modified in the antepenult: (YK) saqi 'mother' and ṣeqinen
'his mother'; (NB.1) papag 'to beat' and pinəpagan 'beaten'. Cf. (CS.2)
saqina 'his mother' and binabagan 'provided with a cross-beam' (from
babag 'cross-beam'). The above illustrations reflect only part of a
more complex process in both languages. There are other constraints
for (YK) on the reduction of a to e, and for (NB) other vowels and other
environments are implicated.

14. (YK) reflexes are about evenly divided between e and i, and (IB)
reflexes between a and i. There is, however, no correlation between
these sets of variations, and insufficient data in either case to sug-
gest any conditioning hypothesis. It should be noted that in (IB) *ə
has multiple and largely unpredictable reflex potential in the penult
as well. In some cases there are currently competing forms: (IB.1)
limbaq, ləmbaq, lumbaq 'to repeat'.

15. I have not included analogic replacement of phonemes since, except
for those cases where innovations can be shown to have spread by some
statable process (Zorc's (1972) morphological regularisation of stress
to indicate morphemic contrast between segmentally identical forms is a
good example), this explanation is frequently a less satisfactory treat-
ment of the data than simply noting that the problem is unexplained.

16. The correct solution may, of course, require a combination of these hypotheses. Doublets may not be of proto-language provenance, but may be the result of early inter-dialect or inter-language borrowing. Once borrowed they become inherited forms of the daughter languages, subject to native laws of sound change. Again, since sound change processes lose their impetus or change phonological direction, the phoneme X of a loan word may undergo different (though regular and describable) processes from those which operated earlier on a phonemically identical phoneme in an inherited word.

17. An alternative explanation is that the Malay source was a dialect of Ml. which realised standard Malay ə as a, although this is not the case for current contacts between Sulu languages and Malay.

18. Llamzon, op. cit.

19. It is not necessary to suppose that a source language form was borrowed in the form in which it currently appears. Lowland Tausug, which reflects PAN *ə as u in all positions, and once assimilated the ə of borrowed Sama forms in the same way (Tsg. kuddum *to frown* probably from PSB *kəddum *to close the eyes*), now assimilates borrowings from Sama languages in a different way, generally by harmonising Sama ə with the final vowel of the word:
    Sama dənda *woman* becomes danda; Sama pəssi *fishhook* becomes pissi; Sama əmboq *ancestor* becomes umbuq when spoken by Tausug who do not control the phonology of Sama.

20. PAN *təmpət *domicile* is poorly attested in VLAW.

21. Innovated ə is discussed in Section 9. This is a fairly typical example of the process described in 9.2.

22. Dempwolff (1934-8), op. cit.

23. The harmonising of the vowels in this context may have been a PSB process; the evidence is ambiguous.

24. This word is possibly a loan from a Philippine language in view of the *g reflex instead of the expected *h reflex of PAN *R. Medial PAN *R is reflected as PSB *g more frequently than initial and final PAN *R, and there may be conditioning factors. Cf. also such forms as *pəggaq < PAN *pəRaq *to squeeze out juice*, and the doublet *taguq/ ta(h)uq *to keep*, both well attested for PSB. The doublet probably reflects, respectively, PAN(B) *(CtT)aguq *hidden* and PAN *taRuq *to set down*.

25. Dempwolff (1925) cites this form, but only *bəsay is given in
VLAW. The *R infix is supported by such Philippine reflexes as Proto
East Mindanao *bəgsay 'canoe paddle'.

26. Both PSB forms appear to derive from PAN *batəŋ. The form (*)butuŋ
'mature, green coconut' though an innovation with this meaning, may be
an example of the exploitation of stress contrast. The semantic connec-
tion is at least plausible, with a shared component of swelling and
fruition.

27. Llamzon, op. cit.

28. Blust (1970) proposes an alternative etymology for Tg. lapit:
PAN(B) *ra(ŋ)pi(Ct) 'approach'.

29. See Section 9.2 for a discussion of this innovation pattern.

30. The change in stress here may be an innovation of early PSB
provenance, and the form in parentheses here is included only to facili-
tate comparison with the established PAN form. No claim for PAN status
is intended.

31. I have no direct data for (AB) stress, though the material was
recorded by an experienced field worker who would, I expect, have
indicated contrastive stress. In any case there is no evidence in the
corpus of some six hundred (AB) words of homophonous pairs for which
one might expect stress to be the contrastive feature, and it is as-
sumed with reasonable confidence that the above statement is equally
applicable to (AB).

32. The phonemic status of the single versus geminate consonant con-
trast in (NB) is uncertain. It carries a low contrastive load in any
case, even when the vowel of the penult is other than ə.

33. (YK) reflects PSB *ə and *u as e in all environments (merging
these both with *e), and *a as e under predictable, stress-governed
conditions.

34. The PSB reconstruction of this phoneme is based on the correspon-
dence set: -e in (YK), -o in (NB), -a elsewhere. The (YK) form is cru-
cial for the identification of this set, since the (NB) evidence is
somewhat ambiguous. In (NB) there is a secondary development whereby
a > o in an open final syllable, and penultimate a > o when the final
vowel is o. Thus:
     *mata 'eye' > (NB) moto; *hadun 'ladder' > (NB) odon.

A similar centralising process takes place with PAN *u and PAN *i
before final PAN *h or PAN *q, with these variants reflected in PSB as
*o and *e.

REFERENCES

ADRIANI, N.
  1900   *Tijdschrift voor Indische Taal-, Land- en Volkenkunde.*
         Deel XII/460-90. Batavia, Albrecht and Co.; 's Hage: M.
         Nijhoff.

ASHLEY, S.
  1975   *Tausug-English Dictionary.* In press.

BLUST, Robert A.
  1970   Proto-Austronesian Addenda. *Oceanic Linguistics* 9(2):104-62.

CONANT, Carlos Everett
  1912   'The pepet law in Philippine languages'. *Anthropos* 5:920-47.

DEMPWOLFF, Otto
  1925   'Die L-, R- und D-Laute in austronesischen Sprachen'. *Zeit-
         schrift für Eingeborenen-Sprachen* 15:19-50, 116-38, 223-38,
         273-319. Berlin: Dietrich Reimer.

  1934- *Vergleichende Lautlehre des Austronesischen Wortschatzes.*
  1938   Band I-III. Berlin: Dietrich Reimer.

DYEN, Isidore A.
  1953   'Dempwolff's *R'. *Language* 29:359-66.

  1965   *A Lexicostatistical Classification of the Austronesian Lan-
         guages.* Indiana University Publications in Anthropology and
         Linguistics, Memoir 19 of the *International Journal of American
         Linguistics.* Baltimore: Waverly Press, Inc.

ECHOLS, John M., and Hassan SHADILY
  1961   *An Indonesian-English Dictionary.* New York: Cornell Uni-
         versity Press.

ELKINS, Richard E.
    1968   *Manobo-English Dictionary*. *Oceanic Linguistics Special*
            Publication No.3. Honolulu: University of Hawaii Press.

ENGLISH, Leo James
    1965   *English-Tagalog Dictionary*. Department of Education,
            Republic of the Philippines. Clayton, Victoria: Wilke and
            Co., Ltd.

GONDA, J.
    1952   *Sanskrit in Indonesia*. Nagpur: The International Academy of
            Indian Culture.

HANCOCK, Ian F.
    1973   'Malacca Creole Portuguese: a brief transformational outline'.
            Te Reo 16:23-44.

HEADLAND, Thomas N., and Janet D. HEADLAND
    1974   *A Dumagat (Casiguran) - English Dictionary*. PL, C28. The
            Australian National University.

LLAMZON, Teodoro A.
    1966   'Tagalog reflexes of PMP *ə'. *Anthropological Linguistics*
            8/3:13-23.

McKAUGHAN, Howard P., and Batua A. MACARAYA
    1967   *A Maranao Dictionary*. Honolulu: University of Hawaii Press.

MILLER, Jeanne, and Helen MILLER
    n.d.   Mamanwa-English Dictionary. Summer Institute of Linguistics.
            Typescript.

REID, Lawrence A., ed.
    1971   *Philippine Minor Languages: Word Lists and Phonologies*.
            Oceanic Linguistics Special Publication No.8. Honolulu:
            University of Hawaii Press.

RODDA, Rosemary, and Audrey MAYER
    n.d.   Batak-English Dictionary. Summer Institute of Linguistics.
            Typescript.

ZORC, David R.
    1972   'Current and Proto Tagalic Stress'. *Philippine Journal of*
            Linguistics 3/1:43-57.