STUDIES IN PHILIPPINE LINGUISTICS

Volume 2 Number 1 1978

Casilda Edrial-Luzares and Austin Hale, series eds.

ISSN: 0119-6456

Sample Citation Format

Ma. Lourdes S. Bautista. 1977. "The noun phrase in Tagalog—English code switching". Studies in Philippine Linguistics 1:1, 1–16. Online. URL: http://www.sil.org/asia/philippines/book_store.html#culture [etc.] + access date.

PRE-SANGIR *l, *d, *r AND ASSOCIATED PHONEMES

Kenneth R. Maryott
Summer Institute of Linguistics
Philippine Branch

Part I

The differences that have developed between Sangihé and Sangiré (otherwise known as Sangir and Sangil respectively) have made these languages particularly suggestive of processes that may account for some problematic correspondences in historical studies.

Sangihé and Sangiré reputedly had a common origin two or three hundred years ago in the Sangir Islands, Northern Sulawesi, Indonesia, where in fact the far greater proportion of the Sangihé speakers still live. The Sangiré have long since emigrated to the southeastern Philippines, where they have been followed in recent years by a small contingent (about 5000) of the Sangihé. The Sangihé have largely accepted the Christian religion and otherwise have been influenced by the more central culture of Indonesia. The Sangiré have embraced Islam and have become oriented rather toward the Philippines.

No thoroughgoing comparison of these similar but distinctive languages has yet been made. However, the present paper reports the results of an initial study that reconstructs a significant set of phonemes featuring those I am calling Pre-Sangir (PS) *l, *d, and *r.

Sounds of this general kind have long held the attention of serious students of Austronesian languages, due mainly to the unusual tendency of these sounds to interchange among themselves.[1] But in Sangihé and Sangiré it is the extent and the complexity of such interchange that is so arresting, both within and between these two languages.

To illustrate, the two lateral phonemes of Sangihé (Sh), simple (l) and retroflexed (ll), and the two comparable laterals of Sangiré (Sr) plus a zero element (l, ll, Ø) all are conditioned replacements of the original phoneme here designated PS*l (e.g. Sh 'langi' / Sr 'langi' 'sky', Sh 'béli / Sr bé'lli 'buy', Sh 'pullo / Sr 'puo 'ten'; note that citations from the two languages will always be given in this order, Sangihé followed by Sangiré). PS*l, however, is not the only phoneme giving rise to Sangihé l. A proto-phoneme which becomes l for Sangihé but for Sangiré y (e.g. bu'ala / bu'aya 'crocodile') I am calling PS*y.

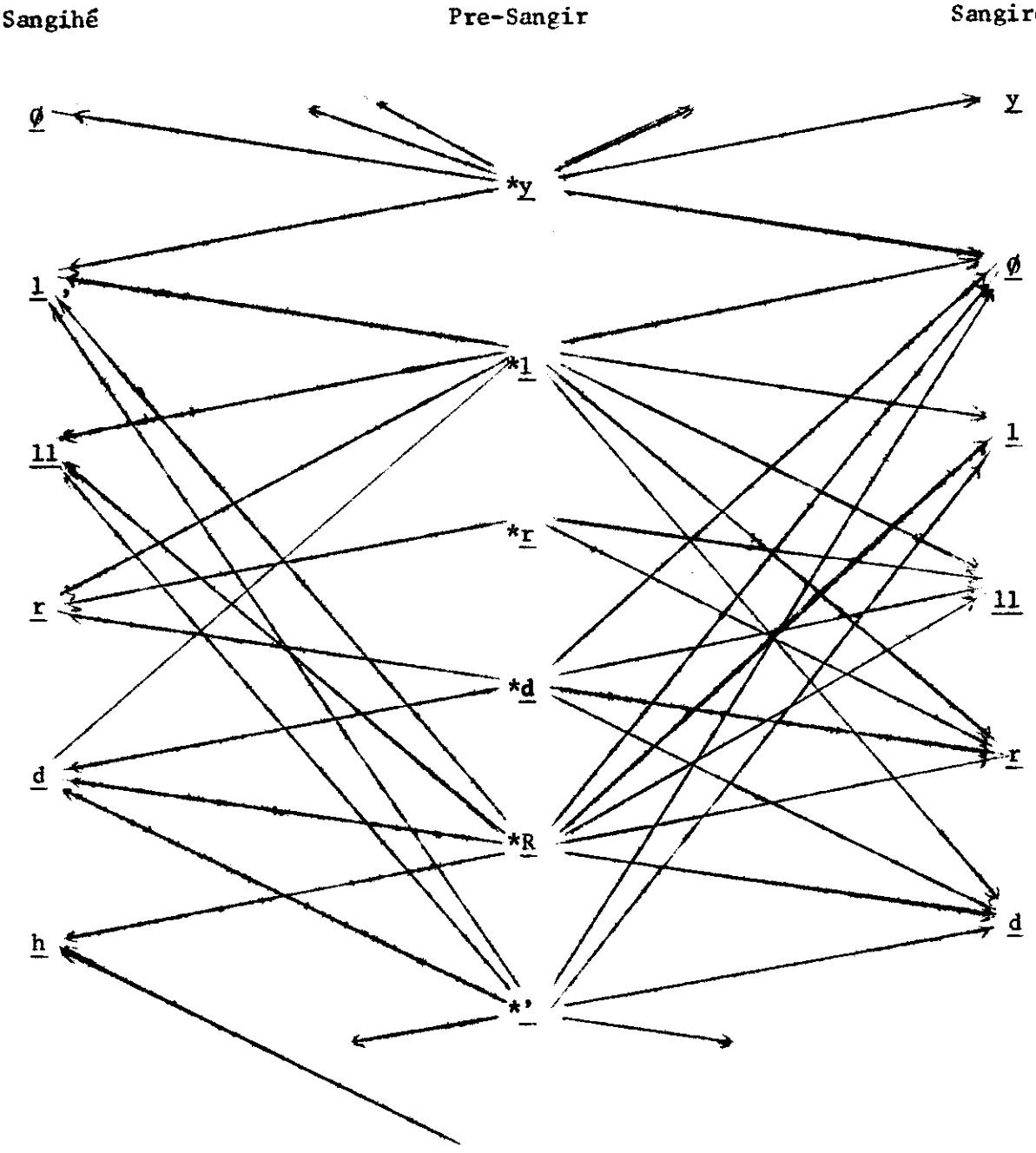
But another merger involves the laterals. The phoneme PS*r, which commonly derives in Sangihé as the flap r, in Sangiré becomes ll, the same phoneme that also derives from PS*l (e.g. 'barisé' / 'ballisé' 'line up'). Thus, it is quite possible for Sangiré to neutralize a contrast retained by Sangihé between forms containing *r and *l respectively (e.g. 'barisé' 'line up', 'ballisé' 'reciprocate' /

'ballisé' 'line up' and 'reciprocate'). Not only so, but in the Mindanao dialect of Sangiré (MinSr), as opposed to the Sarangani dialect (SarSr) which we have been discussing, there is no retroflexed lateral and SarSr ll regularly appears as MinSr r (e.g. SarSr 'ballisé' / MinSr 'barisé' 'line up' and 'reciprocate'). But another proto-phoneme, the PS*R, becomes the flap r in both dialects (and h in Sangihé). Therefore the Mindanao dialect undergoes a three-way neutralization of contrast between PS*r, *l, and *R (e.g. Sh 'barisé' 'line up', 'ballisé' 'reciprocate', and 'bahisé' 'salt fish' / SarSr 'ballisé' 'line up' and 'reciprocate', 'barisé' 'salt fish' / MinSr 'barisé' 'line up' and 'reciprocate' and 'salt fish').

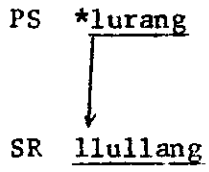
Further, under other stateable conditions, a still different proto-phoneme, PS*d, becomes Sh r but Sr ll, while elsewhere *l becomes d, *d becomes Ø, and *R or even *' becomes l in one or both languages (secs. 2.2, 2.1, 1.1.3, 2.3 respectively).

It is not my purpose at this point to confuse, but merely to underscore the complexity in these languages of the genetic relationships between phonemes of the l-d-r type. It may be appropriate to portray this interrelatedness graphically, at the same time assuring the reader that it will be the primary aim in the remainder of the paper to try to resolve the complexity and state as simply as possible the interesting conditions that occasion it. The diagram that follows indicates only the regular sound shifts of the specified phonemes, the nature and the conditions of these changes to be described later. The shifts of Mindanao Sangiré are not represented, since for want of additional data I will be forced to eliminate this dialect from further consideration.

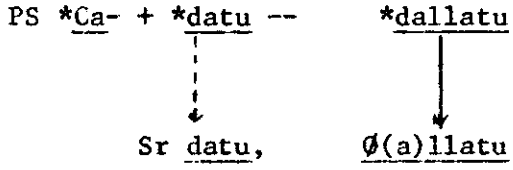
In order to reduce the problem to more manageable proportions, we may observe certain distinctions between diachronically related morphemes characterized by Hoenigswald as "degree of phonemic affinity" (1960:48ff), noting the kinds of conditions that determine this affinity. Besides the absence of any correspondence at all, Hoenigswald cites (1) phoneme-by-phoneme correspondence, or identity of morpheme under sound change, the replacement here presumably being regular and general throughout the language (e.g. PS *lurang > Sr 'llullang 'load', where initial ll in this environment is a replacement of *l; sec. 3), (2) morphophonemically related replacement, where the affinity between replaced and replacing morphemes is mediated through an allomorph of an earlier morphophonemic process (e.g. PS *Ca- + *datu --> *dallatu > Sr 'llatu 'high chieftan', where ll is a replacement of *d; sec. 2.2), and (3) replacement via dialect borrowing, the affinity here being mediated through a borrowing from a cognate language, the indigenous form sometimes but not always being lost in the process (e.g. PAN *rantay > Magindanao rantay > Sr 'llante 'chain', where ll is a replacement of *r; indigenous form (?) 'rotay 'hemp (rope)'). These three types of replacement and their exemplification in Sangiré may be summarized in the following diagram.


Sangihé
Pre-Sangir
Sangiré
ø
l
ll
r
d
h
*y
*l
*r
*d
*R
*'
y
ø
l
ll
r
d

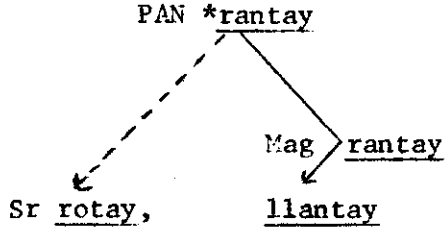
(1) Phonemic Correspondence



(2) Morphophonemic Correspondence



(3) Loan Correspondence



Note that in the third case, dialect borrowing, the correspondence illustrated is to a morpheme not of Pre-Sangir but of Proto-Austronesian. In fact, one is hard pressed to find instances of correspondence to Pre-Sangir through dialect borrowing, at least within the scope of our topic. Replacements of this type, therefore, require no treatment here. Rather, the discussion to follow is organized around just the first two types of

affinity, here redesignated direct correspondence, or that involving exclusively phonological processes, and indirect correspondence, or that implemented through morphophonemic alternations. Direct correspondence is treated herewith in Part I of this study and indirect correspondence, in Part II which will appear in the next number of this series. Both direct and indirect correspondence are dealt with here as instances of primary sound change, or change in sounds while their phonological environments remain constant. Secondary change, during which the environment also changes and in these languages exerts an assimilative influence, will be considered in a Part III of this study.

1. Direct Correspondence

In the languages before us, direct correspondence may involve either conditioned or unconditioned sound change.

1.1 Conditioned Change

The reconstructions that have undergone conditioned, or environmentally diversified change are the proto-phonemes PS *l, *y, and *d.

1.1.1 PS*l

Of these three phonemes, PS*l shows the most complex change. The rules and representative illustrations for this phoneme are presented below. (The spelling of cited forms here and throughout is explained in summary phonologies of Sangihé and Sangiré in Appendix I. Dempwolff's reconstructions for Proto-Austronesian, where known, appear in Dyen's orthography parenthesized after each example. Within his reconstructions diagonals indicate that Dempwolff has only indirect evidence for the phoneme enclosed thereby, and his parentheses indicate an optional phoneme.

(1) As the initial consonant of the word base, PS*l becomes l in both Sangihé and Sangiré (in this environment only silence or the high central vowel é can precede, all other vowels being preposed by another consonant, if only glottal stop).

'labo' / 'labo' (*labaq) 'many, much'
'laku' / 'laku' 'loose-hanging garment'
'lama' / 'lama' 'plate, dish'
'lana / 'lana 'oil'
'langi' / 'langi' (*laNit) 'sky'
'laudé' / 'laudé' (*lahud) 'open sea'
'lese' / 'lesé' 'wound, sore'
'lét:ng / 'lé'bing (*le(m)beN) 'grave'
'likudé' / 'likudé' (*likuD) 'back (anatomic)'
'lima / 'lima (*limah) 'hand, five'
'limasé' / 'limasé' (*limas) 'bilge water'
'lipang / 'lipang (*l'pan) 'centipede'
'liu / 'liu (*laluh) 'across'
'longi' / 'longi' 'gums'
'lua' / 'lua' 'breaking waves'
'lukadé' / 'lukadé' 'watch for'

In the following examples, the Sangiré high central vowel in base-initial position is weak and is not retained unless preposed by certain affixes (e.g. mangé'lu' 'swallow (s.t.)' but 'lu' 'Swallow!'). The Sangiré examples below are all assumed to pattern similarly, though for the second and third, prefixed forms that would verify the recoverability of initial é are not actually attested. In this group of correspondences the vowel is considered always to be the base-initial phoneme, and in Sangiré its loss when initial is handled in the morphophonemics.

'élang / é'lang 'slave'
'éli' / é'li' 'seabottom'
'élo / é'law (*a(n)daw) 'day'
'élu' / é'lu' (*luduq) 'swallow'
'élung / é'lung (*guluN) 'roll up'

(2) As the non-initial consonant of the base:
(2a) after a front vowel (i, e), PS*l becomes l in both Sangihé and Sangiré.

'baelé' / 'baelé' 'field'
'bele' / 'belé' (*belut) 'hold'
'bilang / 'bilang (*bilaN) 'count'
'biling / biling (*biliN) 'turn'
''ilang / ''ilang (*qilaN) 'lose'
''ilu / ''ilu 'poison'
'kila' / 'kila' (*kilat) 'lightening'
'lila / 'dila (*dilaq) 'tongue'
'pakelé' / 'pakelé' (paNkal?) 'heel'
'pile / 'pile (*piliq) 'choose'
'sili / 'sili (*siliq) 'defer'
'silo / 'silaw (*hilaw?) 'see'

(2b) after the high-central vowel (é), and:
(2b1) before a front vowel (i, e), PS*l becomes l in Sangihé and ll in Sangiré.

'bélisé' / bé'llisé' 'rancid'
'béli / bé'lli (*belih) 'buy'
'déli' / dé'lli' 'flame'
'déling / dé'lling 'betel leaf'
'géli' / gé'lli' ~ 'gi' (*beRay) 'give'
'péli 'taboo' / pé'lli 'incest'

(2b2) before a back or low vowel (u, o, a), PS*l becomes l in Sangihé and ø in Sangiré.[2]

ba'bélo / babé'aw (*babaq *a(n)daw) 'afternoon'
'béla' / bé'a' (*belaq) 'middle'
'béladé' / bé'adé' (*belaj) 'sea swell'
'délahé / dé'aré 'yard'
'délang / dé'ang 'storm cloud'
'délu' / dé'u' 'thunder'
lé'laba / lé'aba (*lawah) 'cobweb'
sé'laeng / sé'aeng 'beach'
'téla' / té'a' (*telah?) 'fly'
'télu / 'taw (*teluh) 'three'
'téluhé' / 'auré' (*/t/eluR) 'egg'

Note that in the last two examples, the Sangiré bases have a preceding ∅ instead of the é which is the condition of this rule. The explanation is, of course, that there was an additional sound change--é > a under certain circumstances--following the time stage at which Rule 2b2 shows exceptionless application. More generally, none of the rules for the change of PS*l, *r, etc. can be stated in terms of the present phonologies. The fact that it is almost possible to do so attests the relative lateness of these rules, or perhaps the stability of the environments.

A more troublesome counterexample of this rule is Sr pé'lò 'put' (Sh 'pélò). There are two possible explanations: (1) the high central é is the reflex of an earlier front vowel, following which PS*l would not be lost (Rule 2a); or (2) according to Collinge's principle of "high frequency, low integration" (1970), pélò, which is the common and much-used term for 'let go!' did not participate in the change that lost PS*l in the same environment in less frequent morphemes.

(2c) after a back or low vowel (u, o, a), and:

(2c1) before a front vowel (i, e), PS*l becomes ll in both Sangihé ang Sangiré.

'balle / 'balle (*balay) 'house'
'balli' / 'balli' (*balik) 'reverse, return'
'bolleng / 'bolleng 'pull'
'bulle / 'bulle 'forget'
'dalleng / 'dalleng (*dalan 'path, way'
'dalligé' / 'dalligé' 'buttress root'
én'salli' / én'salli' (*/s/aleR) 'floor'
'palledá' / 'palledé' (*palaj) 'palm of hand'
'pulle' / 'pulle' (*pulaN) 'go home'
sa'lliu / sa'lliu 'exchange'
'talli / 'talli (*talih) 'rope'
talli'mé'do / tallimé'do 'finger'
'tulli / 'tulli (*tuli) 'ear'
'tullidé' / 'tullidé' 'straight'

In Sangihé, 'suling (Sr 'sulling, PA *suliN) 'flute', the non-retroflexed l occurs instead of the retroflexed ll that would normally be expected. The only explanation that suggests itself is that 'suling is really an unassimilated loan from Indonesian (viz. In 'suling 'flute'), as are the names of most Sangihé musical instruments.[3]

(2c2) before a non-front vowel (u, o, a, é), PS*l becomes ll in Sangihé, and Ø in Sangiré (with the coalescence of adjacent vowels if similar and the loss of the first vowel if it is unstressed and the second vowel is the high central).

'alla' / 'a'~ 'aya' (*halap) 'get'
ba'lloto / ba'oto 'eel'
'ballu / 'bau (*baluh) 'widow'
bu'llaeng / bu'aeng 'gold'
'bullang / 'buang (*bulan) 'moon'
'bullude' / 'budé' 'mountain'
'dallung / 'daung (*Dalem) 'interior'
'gallang '(brass) bracelet' / 'gang 'brass'
ko'llano / ko'ano 'king'
'malla / 'ma 'pond, swamp'
'pullo / 'puo (*puluq) 'ten'
'pullu' / 'pu' (*pulut) 'sap'
'salla / 'sa (*salaq) 'mistake'
'sollo / 'so ~'soyo (*suluq) 'lamp'
'sallu' / 'saure' (*saluR) 'river'
'sullung / 'sung 'same'
'tallang / 'tang 'run'
tallo'ara / taw'alla 'middle'
'tolle / 'toay 'tail'
tu'allagé / tu'agé 'light'
''asallé' / ''asé' 'provided that'
bé'basallé' / bé'basé' (*baluh?) 'squash'
'kapallé' / 'kapé' 'ship'
'komollé' 'collect' / 'kumé' 'collection of sayings, lore'
'pundallé' / 'pundé' (*pundul) 'paddle'
'déngkallé' /déng'kaé' 'driftage'
'sékollé' / sé'koé' 'cough
'simbullé' /'bué' (*bukal?) 'spring of water'

The listing here of Sh 'tolle (Sr 'toay) 'tail', even though ll precedes a front rather than a back vowel, underscores the fact that these changes are not the most recent (cf. (2b2)). Actually, the final e in Sangihé represents an earlier PS*ay (which has been retained in Sr 'toay; sec. 1.1.2), and the situation at that point in the chronology is wholly consistent with the terms of this rule as stated.

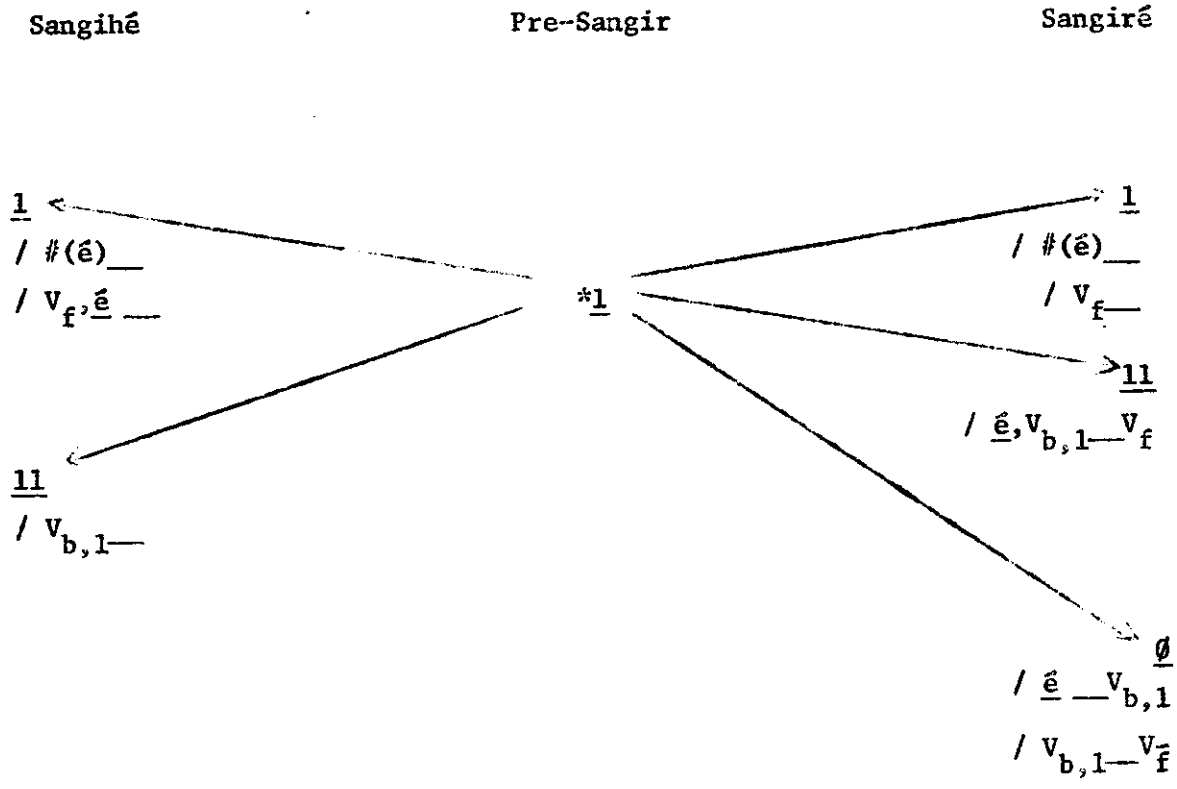
Parenthetically, attention should be called to a fact about the high central vowel that may not be evident from the above. With respect to the conditioning of the changes of PS*l, the high central (é) tends

to pattern with the front vowels in Sangihé but with the back vowels in Sangiré. Thus, following é, PS*l becomes Sangihé l just as it does following the front vowels (e.g. Sh 'téla' 'fly' and 'bilang 'count', vs. Sr. té'a' 'fly' but 'bilang 'count'), whereas PS*l following é becomes Sangiréø just as it does following the back vowels (e.g. Sr té'a' 'fly' and 'puo 'ten', vs. Sh 'téla' 'fly' but 'pullo 'ten').

But again concerning PS*l itself, I was quite puzzled at first about its development into Sr ø and ll regularly (Rules 2c and b) but also into y irregularly; e.g. ''alla' / ''a' ~ ''aya' (*halap) 'get', 'sollo / 'so ~ 'soyo (*suluq) 'lamp', and 'géli' / gélii' ~ 'gi' (undoubtedly via *géyi' with weak é lost and y coalescing with i) 'give', and, with the complete absence of a "regular" alternate, ba'lline / ba'ine (again y coalescing with i) 'other'. Then I discovered Conant's "Indonesian l" article and learned that the phenomenon that troubled me was by no means unique to Sangiré. According to Conant this phoneme, although "one of the most stable of the original consonantal sounds of Austronesia," does change or is lost altogether in some of the daughter languages. The possibilities include regular shifts to r (Sambali, Tahitian, Rapanui, and certain languages of Formosa and Borneo), d (Inibaloi, Batan, and some Malagasi dialects), either x or h under certain conditions in Batan, and q in Ilongot. And, as frequently happens in Sangiré, intervocalic l is lost in Cebuano, Sulu, Tagalog, Bontok, Kankanai, Samal, and Mandaya, and some Formosan, New Guinea and Marquesas Island languages. But more to the point, l becomes y under certain conditions in Tagalog, Bontok, Isinai, and Mandaya, as well as non-Philippine Bare'e and Palau.

Concerning the situation in Mandaya, Conant states: "Final l regularly becomes i in Mandaya... Furthermore, this tendency to palatalize l to i or y is seen even in intervocalic position where in some words l may be either lost or changed to y [italics mine], e.g. Mandaya sâup or sayup: Bisaya sâlop 'set (of heavenly bodies)'" (p. 186). It would seem from this that the Sangiré alternation of ø or ll with y for PS*l is only part of a process which is much more inclusive geographically and linguistically, and whatever the explanation, it properly falls outside the scope of this study on Pre-Sangir.

To summarize sec. 1.1.1, the regular sound changes of PS*l may be displayed as in the following diagram. The symbols used in the notational representation of phoneme environments are the ones conventionally used and should be clear to the reader upon comparison of notational with discursive forms of the rules.


Sangihé
Pre-Sangir
Sangiré
l
/ #(ẽ)__
/ V_f,ẽ __
ll
/ V_b,l—
*l
l
/ #(ẽ)__
/ V_f—
ll
/ ẽ,V_b,l—V_f
ø
/ ẽ —V_b,l
/ V_b,l—V_f̄

1.1.2 PS*y

Implied above is the fact that where Sh l follows back or low vowels it cannot be a reflex of PS*l; in these environments it is always a reflex of PS*y. But by virtue of the l merger of *y and *l reflexes, and also of the non-contrastive environments where these reflexes may occur, Sangihé exhibits such lateral pairs as 'dalung (PAN* Dayung) 'oar' and 'dallung (PAN*Dalem) 'interior'. Sangiré also shows lateral contrasts but not in connection with PS*y (see sec. 1.2.1).

For PS*y itself, the rules and their exemplification are quite straightforward.

(1) As the initial consonant of the base, PS*y becomes i preceded by ' in Sangihé and y in Sangiré.

'i'akang / 'yakang 'eldest sibling'
'i'upung / 'yupung 'grandparent'
'i'uta' / 'yuta' 'innumerable'

(2) As a medial consonant of the base:

(2a) after a front vowel (i, e), PS*y becomes ∅ in both Sangihé and Sangiré. [4]

'bea' / 'bea' (*biyak) 'excuse oneself'
'heo / 'reo (*hiyak) 'cry, scream'
'pia / 'pia (*pi/y/a/h) 'good'
'sie / 'sie (*hiyah) 'he, she'
'tiu'/ 'tiu' (*/h/iyup) 'blow'

(2b) after a non-front vowel (é, a, u, o), PS*y becomes l in Sangihé and y in Sangiré.

''alo / ''ayo (*kayaw) 'raid'
bu'ala / bu'aya (*b/uq/ayah) 'crocodile'
'dala / 'daya ~'da (*Dayah) 'inland'
'dalung / 'dayung (*Dayung) 'oar'
'dulung / 'duyung (*DuyuN) 'sea cow'
'kalu / 'kayu (*kayuh) 'tree, wood'
ké'laeng / ki'yaeng 'type of bamboo'
'nalang / 'nayang ~'nang (*hayam) 'toy, play'
'palung / 'payung (*payuN) 'umbrella'
'pulu 'mad, desirous' / 'puyu 'mad' ~ 'pu 'desirous' (*puyuh)
'pulung / 'puyung 'grandchild'
'tala / 'taya ~'ta 'no'

Several forms require comment. Sr 'daya ~ 'da, 'taya ~ 'ta and 'nayang ~ nang are evidently subject to the same forces that produce ''a' ~ ''aya' and 'so ~ 'soyo from etymons with PS*l. This fact seems to confirm our conclusion in the preceding section that whatever process is at work, it is more extensive than one producing merely a y variant of ∅ from original *l or, we might add here, a ∅ variant of y from original *y. There are, in fact, several Sangiré forms for which there is not even the regular y correspondence to PS*y, but only the irregular ∅ option:

'bala / 'ba (*baya/q/) 'disregard', 'balaré' / 'badé' (*bayaD) 'pay for', 'dolohé' / 'doré' 'send person', ho'loéng / 'rung 'wet', and mahu'ala / maru'a 'maiden'. [5]

In the case of Sr ki'yaeng, the assumption that there was an original é that shifted to i before y is not hazardous, since é is in any case a weak and readily altered vowel.[6]

(3) As the final consonant of the base:

(3a) after a high or back vowel (i, é, u, o) PS*y together with preceding *a becomes e in Sangihé and ay in Sangiré.

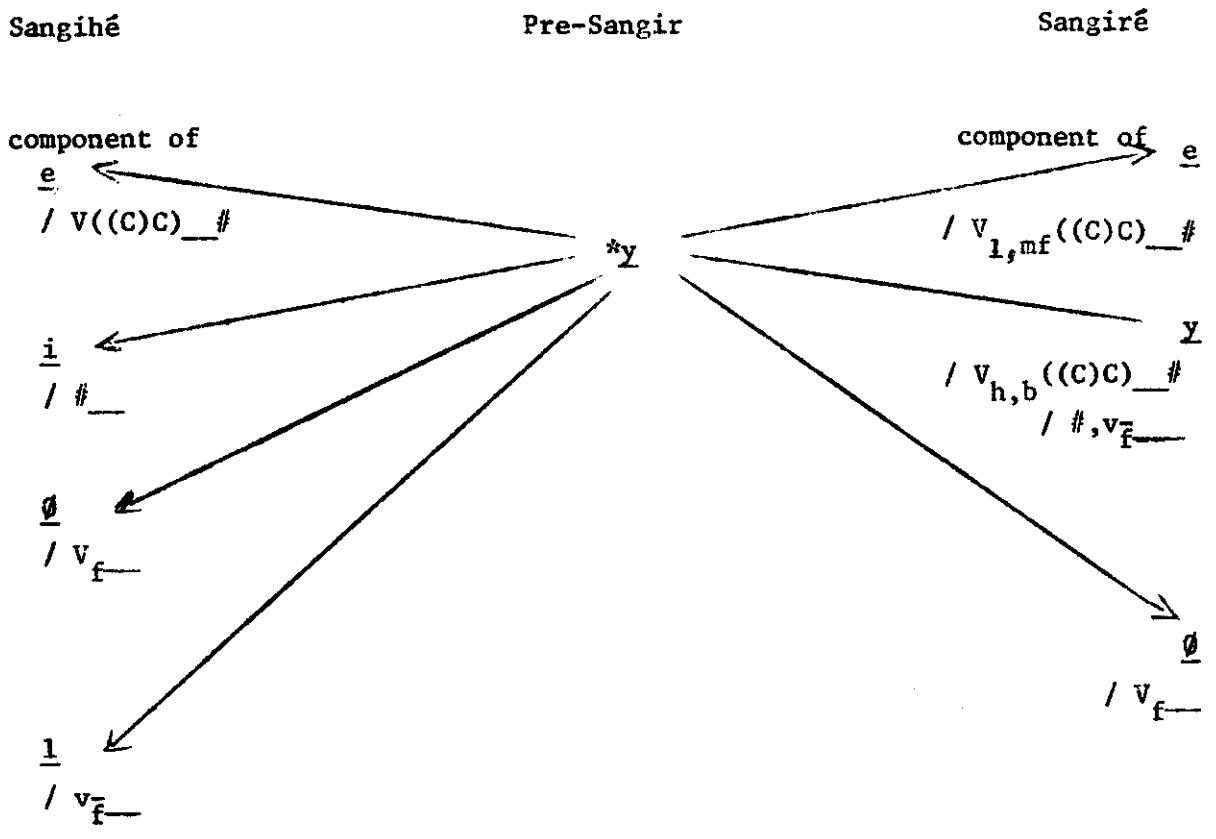
ba'bine / ba'binay (*binay) 'woman'
'hote / 'rotay 'abaca'
'lége / lé'gay (*gelih) 'laugh'
'tumbe / 'tumbay 'brink'
''ue / 'uay (*huway) 'rattan'

(3b) after a low or mid-front vowel (a, e), PS*y together with preceding *a becomes e in both Sangihé and Sangiré.

''ate / ''ate (*hatay) 'liver'
'balle / 'balle (*balay) 'house'
'hengke / 'bengke (*laNkay) 'lift'
'lente / 'lente (*lantay) 'cross (flooring)'
'pande / 'pande (*panday) 'able person'
'sake / 'sake (*sakay) 'embark'

In the absence of any preceding vowel at all, the occurrence of Sangiré e as opposed to ay is unpredictable (cf. 're (Sh 'éhe) 'kind of grass' and 'may (Sh 'éme) 'rice plants').

The history of PS*y is summarized diagrammatically as follows.


Sangihé
Pre-Sangir
Sangiré
component of
e
/ V((C)C)__#
i
/ #__
ø
/ V_f—
l
/ v̄_f—
*y
component of
e
/ V_l,mf((C)C)__#
y
/ V_h,b((C)C)__#
/ #,v̄_f—
ø
/ V_f—

1.1.3 PS*d

PS*d is subject to just two rules, one of which is sharply restricted in its application.

(1) PS*d becomes d in Sangihé, and in Sangiré d except both after silence or a front vowel and before an original *l that is bracketed by similar back or low vowels.

'dalleng / 'dalleng (*dalan) 'road'
'dallung / 'daung (*Dalem) 'inside'
'daung / 'daung (*dahun) 'leaf'
'déka' / déke' (*deket) 'stick'
'diadi / 'diadi (*zadih) 'create, establish'
'diko' / 'diko' (*deket) 'ignite'
'dingihé / 'dingiré' (*deNeR) 'hear'
'dua / 'dua (*Duwah) 'two'
'duku' / 'duku' (*dukut) 'grass, vegetable'
én'dai' / én'dai' (*/h/anDah) 'there'
'éntudé' / én'tudé' (*qa(n)ted) 'bring, take'
'likudé' /'likudé'(*likuD) 'back'
'palledé' / 'palledé' (*palaj) 'palm of hand, sole of foot'
'panda / 'panda (*pandaN) 'look at'
'peda / 'peda (*peDa) 'bolo'
'pédu / pé'du (*pejuh) 'gall, anger'
'tada / 'tada (*/t/i (n)zak) 'tread'
talli'médo / tallimé'do (*tu(n)duq) 'finger'
'tukadé' / 'tukadé' (*/t/ukat) 'ladder, doorway'
''undang / ''undang (*qa(n)daN) 'medicine'

Counterexamples are 'nguda / 'ngulla (*hudah) 'young', and 'lihadé' / 'llehallé' 'chafe, strip away', in which the Sangiré forms do not manifest d as do the Sangihé. What has happened is that for Sangiré, PS*d first shifted to its intervocalic counterpart PS*r and thence to ll in accordance with the PS*r rule. Why this intervocalic change did not occur in the Sangihé forms--or for that matter, in other Sangihé and Sangiré forms (e.g. 'tada / 'tada 'tread')--is something that can be looked into later (Appendix II).

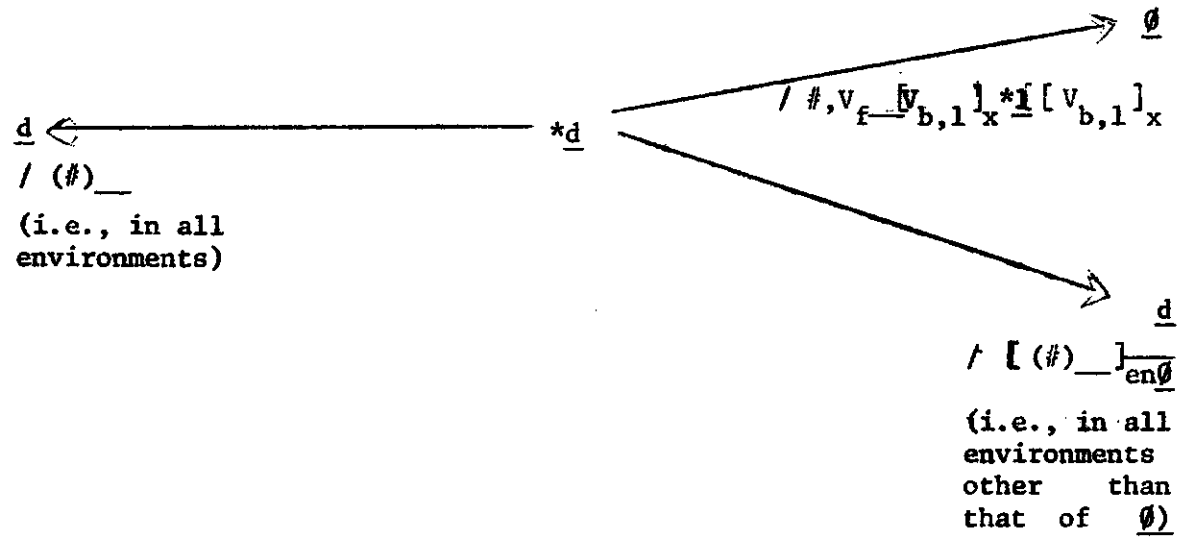
(2) PS*d becomes Ø in Sangiré after silence or a front vowel and before an original *l which is bracketed by similar back or low vowels (the first of which is then lost).

ki'ralla (ki + 'dalla) / ki'la (*lala) 'know, recognize'
dallo'ati / la'wati (*zuluk + *qatay?) 'roundworm'
'i'rollo ('i + 'dollo) / 'i'lo 'rest'
'dollosé' / 'losé' (*sulur) 'interchange, borrow'
do'llosi / 'losi 'k.o. sea fish'
'dolloi / 'loy 'sapling'
'dullu / 'lu (*zuluh) 'upland, wilderness'

Despite the dissimilar vowels bracketing Sh dallo'ati above, the absence of initial d in the cognate Sr la'wati is not irregular. It is quite certain that prior to the loss the Sangiré form had the low central a both before and after the medial lateral. Sr aw is a sister reflex of Sh o in just such environments throughout the language.

A genuine exception to this rule is Sr 'doré' (Sh 'dollohé') 'bald'. But even here there may be conditioning factors, which could even be written into the rule if there were more than just this one instance of the divergence. Notice that the loss of the initial CV - from the reconstructed PS*doloR would have resulted in *lor, a collocation of consonants intolerable to Sangiré (see sec. 3). Therefore the d was retained and the l itself lost, in this case by PS*l Rule 2c2. Implicit here is the assumption that this *d rule normally applied before the *l rules. Otherwise the lateral would have been lost before the environment fulfilled the conditions for the loss of d. [8]

Sangihé | Pre-Sangir | Sangiré



1.2 Unconditioned Change

It is reconstructed PS*r and *R that for each language have reflexes consistently the same in all environments.

1.2.1 PS*r

The fact that PS*r occurs only medially in word bases insures against complicating phenomena at morpheme borders. There is therefore but one rule and that is simply stated. First, however, there is one preliminary to dispose of.

In Sangíe, ll, the reflex of PS*r, occurs in some of the same environments as ll and l, the reflexes of PS*l. It is these correspondences that result in the Sangiré homonymy I referred to in the introduction ('ballisé' 'line up' and 'reciprocate'), as well as in lateral pairs like 'silli (*sijih) 'separate hulled from unhulled grain' and 'silli (*silig) 'defer to'.

PS*r becomes r in Sangihé and ll in Sangiré.

'bera / 'bella (*bajaq) 'word, speak'
'biru' / 'billu' (*bi/r/uh) 'blue'
'dirisé' /'dillisé' 'bandage'
'doro' / 'dollo' 'alight, roost'
''irung / ''illung (*hijuN) 'nose'
'kere / 'kelle 'like, as'
'kire / 'killay 'agree'
'kurang / 'kullang (*kuraN) 'lack'
'kuring / 'kulling (*ku/dD/en) 'cooking pot'
'kurung / 'kullung (*kuruN) 'enclosure'
'mara / 'malla (*maja/h/) 'dry'
'mura / 'mulla (*mudaq) 'easy, inexpensive'
'parasé' / 'pallasé' (*paras) 'smooth'
pa'renta / pa'llenta (*/r/intaq) 'command'
'pira / 'pilla (*pijah) 'how many or much'
'sarang / 'sallang (* ha/r/aq) 'direction'
sa'ria / sa'llia (*besa/r/) 'large, course'
'surate' / 'sullate' (*surat) 'writing'
ta'rima' / ta'llima' (*ra/r/imah) 'accept'
''urang / ''ullang (*qu(n)DaN) 'shrimp'

Though r does not normally occur initially in Sangihé, a small class of loans exists where r is initial (e.g. Sh raha'sia (<In rahasia) 'secret'). And where these or r-initial forms from other languages are taken into Sangiré, the shift is as stated in the rules above (e.g. Sr llaha'sia (<Sh or Mag rahasia) 'secret truth').[9]

Sangihé	Pre-Sangir	Sangiré
r ← / (#)__ (i.e., in all environments)	*r	→ ll / (#)__ (i.e., in all environments)

1.2.2 PS*R

The last reconstruction to be considered in this part of the paper, PS*R is replaced consistently by the same phoneme in all environments.

PS*R becomes h in Sangihé and r in Sangiré.

''ahusé' / ''arusé' (*qaRus) 'succeed, change with'
'bahe' / 'bare' (*beRat) 'northwest monsoon'
'bẽha' / 'bẽ'ra' (*beRat) 'heavy'
'bibihé' / 'bibiré' (*bibiR) 'lips'
'bohasé' / birasé' (*beRas) 'hulled rice'
'boheng / 'boring (*hajeN) 'charcoal'
bo'hẽsi' / boré'si' (*burih, *resi) 'smooth'
'buhu / 'buru (*beRuh) 'new'
'daha / 'dara (*/dD/aRaq) 'blood'
'ẽhe / 're 'kind of tall grass'
'dolohé' / 'doré' 'send person'
'haki' / 'raki' (*Ra(N)kit) 'raft, bind together'
'hasu' / 'rasu' (*Ratus) 'hundred'
'hẽbi' / rẽbi' (*Rabih) night, 24-hour period'
'hibu / 'ribu (*ribu) 'thousand'
hum'bia / rum'bia (*rumbi/y/ah) 'sago palm'
'husu' / 'rusu' (*Rusuk) 'rib'
''iha' / ''ira' (*huRat) 'tendon, vein'
'meha / 'mera (*hiRaq) 'brown'
'puhe / 'puray 'seashell'
'sangihé' / 'sangiré' 'Sangihé, Sangiré
''uhasé' / ''urasé' (*quRas) 'wash'

The only form worthy of notice is Sangihé 'sallu' (Sr 'saurẽ', *saluR) 'river', the derivation of which is regular except that the representation of *R has fiminished almost to vanishing, the only residue being glottal stop. This glottal is customary with such near losses of a final consonant though the losses themselves are relatively infrequent. For as Lopez (n.d.:42-43) puts it, "Sound decay is prevented by means of a supporting vowel in Sangir," viz. the high-central vowel ẽ together with its invariable accompaniment when final, the glottal stop.

It should perhaps be noted here that PS*R is not the only proto-phoneme deriving as h in a daughter language. PS*h too, which has h reflexes in Sangihé and Sangiré as well (e.g. 'harapẽ / 'hallapẽ 'hope, expect'), thus gives evidence of yet another merger in these languages from previous to present time stage. However, since PS*h does not develop into phonemes of the l-d-r variety, it warrants no further attention here.

Sangihé	Pre-Sangir	Sangiré
h	*R	r
/ (#)__		/ (#)__
(i.e., in all environments)		(i.e., in all environments)

Preview: Of the sound correspondences charted on page 117, Part I has now accounted for 19 (nine for Sangihé and ten for Sangiré), or approximately 44% of the total of 43 correspondences. Though less than half of the complexity referred to at the outset has been resolved, the description thus far has perhaps been useful in its own right and certainly is foundational for what follows. Specifically, the next topic will move into an area in which there is the disappearance of base-initial phonemes rendered intervocalic by the juxaposition of an affix at some previous point in time. Such an absence of consonants, the coalescence of their bracketing vowels, and their apparent replacement by other consonants can be exceedingly confusing in reconstructing primitive forms as well as in understanding the current grammar. But more of this in Part II.

APPENDIX

The following is a brief description of the Sangihé and Sangiré phonemes featured in this article. The statements are abstracted (with minor corrections) from three of my own papers (1963, 1968, 1977), but an earlier and very different approach to the Sangihé phonology may be found in N. Adriani (1893:40-52).

To summarize for Sangihé, the phoneme inventory consists of 17 consonants, six vowels, and word stress with an extremely light functional load except in loans. [10] In the following list, phonemic symbols represent their respective phonetic counterparts unless otherwise indicated by a symbol in brackets. Conditioned variants are indicated in parentheses where relevant. The consonants are: b, b̲, d, g, g̲, h, k, l, ll, [ɬ], m, n, ng [ŋ], p, r [ɾ], s, t, ’. The vowels

are: a, e ([ɛ] in closed syllables, [e] elsewhere), ë ([ə] preceding stressed syllables, [ɨ] elsewhere), i ([ɪ] in nasal-final syllables, [i] elsewhere), o ([ɔ] in glottal-final syllables, [o] elsewhere), u ([ʊ] in nasal-final syllables, [u] elsewhere).

Details on the phonemes in this paper are as follows.

l [l] is the voiced alveolar lateral non-vocoid. It occurs only initially and medially, and infrequently immediately following the glottal stop. It is subject to doubling after the high central vowel ë.[11]

ll [ɭ] is the voiced retroflexed flapped lateral non-vocoid. It occurs only medially following back and low central vowels.

r [ɾ] is the voiced alveolar flapped non-vocoid. It occurs only medially.

d [d] is the voiced alveolar stop non-vocoid. It occurs initially and medially, and also immediately following the glottal stop or a homorganic nasal.

h [h] is the voiceless glottal fricative non-vocoid. It occurs initially and medially, and infrequently immediately following the glottal stop.

' ['] is the voiceless glottal stop non-vocoid. It occurs initially, finally, medially rarely, and as the first member of consonant pairs.

The general distribution of the phonemes may be stated in terms of the phonemic syllable as follows: (1) the syllable consists of a single vowel which may be preceded and/or followed by a consonant (formulaically, ± C + V ± C); and (2) the four syllable types (viz. CV, CVC, VC, V) co-occur freely except that vowel-initial syllables do not follow closed syllables.

The specific distribution of the phonemes is distinguished by (1) the non-occurrence utterance-initially of any vowel except ë; (2) the occurrence of the fricatives b and g and the flapped consonants ll and r only intervocalically, but never immediately after ë; (3) the limitation of consonant clusters to nasal plus homorganic stop or sibilant, or to glottal stop plus any consonant except fricative, retroflexive, or glottal stop; and (4) the non-occurrence utterance-finally of the high central vowel ë or of any consonant except glottal stop ' or the velar nasal ng. The velar nasal (-ng) invariably replaces any word-final nasal of proto-forms or loans. The utterance-final sequence high central vowel plus glottal stop (-ë'), a paragogic or echo phenomenon that occurs with high frequence in Sangihé (and Sangiré), is invariably added to proto-forms or loans which retain a final consonant other than a nasal or glottal stop. Though this final sequence is predictable, I transcribe it here to facilitate the reading of cited forms.

The Sangirẹ́ inventory consists of 18 consonants, six vowels, and indigenous word stress with a relatively heavy functional load. The consonants are the same as for Sangihé except that Sangirẹ́ does not have g and does have w and y. Sangirẹ́ vowels are the same as for Sangihé, but have minor differences in allophonic variation.

The featured consonants for Sangirẹ́ are:

y [y] is the voiced high close front unrounded non-syllabic vocoid. It occurs in all word positions. In open unstressed non-final syllables immediately following either a low vowel or an alveolar stop or nasal, it occurs in free variation with allophones of the phoneme i.

l [l] is the voiced alveolar lateral non-vocoid. Its distribution is identical with that of Sangihé l.

ll [ɭ̆] is the voiced retroflexed flapped lateral non-vocoid. It occurs medially following low central and back vowels.

[əɭ̆] is the voiced retroflexed flapped lateral non-vocoid with retroflexed mid central non-syllabic vocoid on-glide. It occurs initially and medially in complementary distribution with its co-allophone.

r [ɾ̆] is the voiced alveolar flapped non-vocoid. It occurs initially and medially.

[r^{g}] is the voiced alveolar flapped non-vocoid with velar friction. It occurs in free variation with the allophone [ɾ̆] of this same phoneme; it is the more frequent.

d [d] is the voiced alveolar stop non-vocoid. Its distribution is the same as that of Sangihé d.

' ['] is the voiceless glottal stop non-vocoid. Its distribution is the same as that for the Sangihé '.

The general and specific distribution of the Sangirẹ́ phonemes is the same as that for the Sangihé except that the high central vowel ẹ́ is never stressed unless there is no other vowel in the word to which stress can shift, the fricative b and the flaps ll and r may occur initially as well as intervocalically, and also after ẹ́, and w and y may occur finally. In addition, the final sequence -ẹ́' as such is non-predictable and thus phonemic, though once the sequence occurs, the glottal stop, being the invariable accompaniment of the high-central vowel in this position, must occur and is itself therefore non-phonemic.

FOOTNOTES

[1]Of the early scholars investigating such phenomena, Conant and Dempwolff are of particular interest to us here. Conant has produced an especially helpful study entitled "Indonesian l in Philippine Languages" (1916), and also the article "The RGH Law in Philippine Languages" (1911). The latter is of direct relevance to PS*R, one of the phonemes treated here as intimately associated withPS*l, *d and *r. Dempwolff has given us an invaluable set of pertinent reconstructions in his monumental "Vergleishende Lautlehre des austronesischen Wortschatzes" (1938), but he has also published a work even more germaine to my reconstructions, "Die L-, R- und D-Laute in austronesischen Sprachen" (1924-25). Bloomfield (1927) reviews this article as an outline of "the sound history of one of the world's largest linguistic stocks" and says of it, "Dempwolff's study amounts to a comparative Austronesian phonology, probably as complete and detailed as human ingenuity could make it with the data that are available... Further advance will be possible, it would seem, only when more material can be used." It is my hope that the data and analysis in this present study will be counted among the many useful responses to Bloomfield's call for additional materials.

[2]Notice that these rules do not provide for an environment in which the high central é not only precedes but also follows the lateral within the word base. The omission is made simply because no reflex of PS*l has been found to occur in that environment.

[3]Of course, Sangiré also exhibits such unassimilated loans, one of many from Visayan being ma'lita 'suitcase'.

[4]An alternate analysis might be that PS*y did not occur in this environment, PAN*y having become *Ø before reaching the Pre-Sangir stage. It seems to me the matter is indeterminable here, and I have just opted for the simplest analysis.

[5]Interestingly enough, almost every *y that is bracketed between similar non-front vowels has the Ø alternate as replacement.

[6]See Maryott (1963) for a brief synchronic study of this vowel in Sangihé, and Conant (1912) for a more general diachronic study.

[7]In Steller and Aebersold's usually impeccable Sangihé dictionary (1959), I take the form "mangirilo" (under "dolo", p. 105) to be a typographical error for "mangirolo', and informants confirm this.

[8]There is just conceivably a way of accounting for the data other than by this rule. Steller and Aebersold (1959) list a number of

Sangihé words like "dolo" 'rest, etc.' as having (presumably) dialectal variants of the form "lolo" 'rest, etc.' i.e., forms in which l replaces d initially. Might it not be plausible to find such variants for each of the Sangihé forms exemplifying this rule, then claim that not the initial consonant but the medial consonant was lost, quite routinely by *l rule 2c2. The l-initial variants do exist. I use one of them to explain a Sangiré form that would otherwise have constituted an exception to this *d rule 2. (Thus, Sr 'luaw 'saliva, spittle' does not pair with Sh 'dullo 'saliva' etc.', with the dissimilar bracketing vowels and the metathesis that this would entail, but rather with 'llullo, a Northern Sangihé variant not listed by Steller and Aebersold but cited by an acquaintance of ours from that area.) But these variant forms are not all that easy to come by. Of the seven examples under Rule 2, Steller and Aebersold list only "dolo" as possessing the variant, and elicitation from the Sangihé themselves proves little better. In any case, the initial-d loss is required in sec. 2.2 and there nothing resembling the l-variant analysis will suffice.

[9]I have no data on borrowings from Sangiré by Sangihé.

[10]This will correct the claim in my 1963 article to the effect that stress is non-phonemic in Sangihé.

[11]It is just such doubling that motivated the Dutch scholars, at that point writing phonetically rather than phonemically, to use double-l, double-m, double-n, etc. immediately following the high central vowel. This is very different from my use of the double-l to symbolize the retroflexed lateral (Dutch "l") in contrast to single-l to represent the non-retroflexed lateral (Dutch "l" or "ll").

REFERENCES

Adriani, N. (1893). Sangireesche Spraakkunst. Leiden: A. H. Adriani.

Bloomfield, Leonard (1927). Review of Otto Dempwolff, "Die L-, R- und D-Laute in austronesischen Sprachen"; Language 3:199.

Collinge, Neville E. (1970). Collectanea Linguistica, essays in general and genetic linguistics. Janua Linguorum, Series Minor 21. The Hague: Mouton.

Conant, Carlos E. (1911). "The RGH Law in the Philippines," Journal of the American Oriental Society 31:70-85.

_____ (1912). "The Pepet Law in Philippine Languages," Anthropos 7:920-947.

_____ (1916). "Indonesian l in Philippine Languages," Journal of the American Oriental Society 36:181-196.

Dempwolf, Otto (1924-25). "Die L-, R- und D-Laute in austronesischen Sprachen," Zeitschrift für Eingeborenen-sprachen 15:19-50, 116-138, 223-238, 273-319.

_____ (1938). "Vergleichende Lautlehre des Austronesischen Wortschatzes" III Band: "Austronesischen Wörterverzeichnis," Zeitschrift für Eingeborenen-sprachen. Berlin: Verlag von Dietrich Reimer (Andrews and Steiner).

Hoenigswald, Henry M. (1960). Language Change and Linguistic Reconstruction. Chicago: Univeristy of Chicago Press.

Lopez, Cecilio (n.d.). Studies on Dempwolff's "Vergleichende Lautlehre des Austronesischen Wortschatzes". Manila: Summer Institute of Linguistics.

Maryott, Kenneth R. (1963). "The Phonology and Morphophonemics of Tabukang Sangir," Philippine Social Sciences and Humanities Review 26:111-126.

_____ (1968). Reference Clauses in Sangir. (Unpublished thesis, Hartford Seminary Foundation).

_____ (1977). "The Phonemes of Sarangani Sangiré," Studies in Philippine Linguistics 1:264-279.

Steller, K. G. F., and W. E. Aebersold (1959). Sangirees-Nederlands Woordenboek. 'S-Gravenhage: Martinus Nijhoff.

Wojowasito, S., and W. J. S. Poerwadarminta (1961). Kamus Bahasa Indonesia-Inggris. Jakarta: Penerbit Tiara.