The Philippine Languages and the Determination of PAN Syllable Structure

John U. Wolff
Cornell University

Professor Lawrence Reid has made important contributions to our understanding of the history of the languages of the Philippines. It is fitting for a paper which points out the role of the Philippine languages in establishing the Proto-Austronesian (PAN) phonology to be included in a group of papers dedicated to him. In this paper I propose to reexamine forms, many of which have been discussed previously in the literature, and show how the Philippine languages show reflexes which provide crucial evidence for certain areas of PAN phonology that have heretofore gone unrecognized. Specifically, they provide evidence for the syllable structure of PAN roots and the kinds of vowel sequences that occurred in the PAN root. This matter is discussed in Section 1. A study that reconstructs vowel sequences should at the same time examine consonant sequences, for the question of whether or not consonant clusters existed in PAN also impinges on the issue of the kinds of vowel sequences which occurred. Section 1 discusses the vowel sequences and Section 2 presents the data which indicate that consonant clusters did not exist.

1. Syllables Beginning in a Vowel

Dempwolff and most historians of the AN languages after him reconstructed almost no vowel sequences in PAN roots. However, the Philippine languages provide evidence for the reconstruction of vowel sequences in many roots, even though the majority of AN languages show single vowels. This evidence comes from the development of a glottal stop to separate an original Pre-Philippine vowel sequence. This development of the glottal stop is probably not a Proto-Philippine process. It took place at a later point and continued as an ongoing change over a period of time. It covers a wide

---

1 The term 'Philippine languages' refers to a group which evidently includes languages outside of the Philippines as well as all of those (or at least all but a very few) in the Philippines. The facts here discussed characterize the languages of Northern Sulawesi as well as the Philippines, but they do not characterize all of them (see footnote 3). Nevertheless, in this paper I will talk of the Philippine languages in discussing these facts even though they are not relevant for all the languages of the group. I use the abbreviations: Tg for Tagalog, Cb for Cebuano, Ilk for Iloko, Ml for Malay, AN for Austronesian, and PAN for Proto-Austronesian.

2 The term 'Pre-Philippine' refers a stage between PAN and the contemporary Philippine languages. 'Pre-Philippine' may be contemporary with but is probably later than 'Proto-Philippine'. An innovation is said to characterize 'Proto-Philippine' if it is reflected in all languages of the group. An innovation is 'Pre-Philippine' if it began in Proto-Philippine times but has not spread enough to be reflected in all of the current daughter languages.
Before we continue with the development of the glottal stop, we need to provide the evidence for the existence of vowel sequences, for it is not immediately obvious that these existed and in which kinds of roots these occurred.

1.1 Evidence for the existence of a vowel sequence

There are two kinds of roots in which vowel sequences have been posited: disyllabics and trisyllabics. Disyllabic roots are of the type exemplified by the root ‘eat’, which is widely reconstructed as *kaen, on the theory that forms reflecting *kan result from a contraction of *kaen (Robert Blust, pers. comm.). Even though reflexes of *kan are much more widespread than those of *kaen, forms which have the appearance of reflexes of *kaen are found in languages of Taiwan (specifically Bunun and Amis) and also in the Philippines, that is, in languages which are related only by being descendents of PAN, and not in a later subgroup. The argument is that this shared feature must be characteristic of PAN. However, I argue that the development of *kaen is the product of independent innovations in languages of Taiwan and the Philippines. In the case of *kan and other roots like it, I argue that these were not disyllabic roots in PAN but rather that they were monosyllabic roots in PAN, and that the disyllabic forms are the product of the process of syllable-internal lengthening which took place independently and repeatedly in post-PAN times throughout the range of the AN languages (Wolff 1999). My argument is supported by the fact that the process of creating the disyllabic reflexes of monosyllabic roots is found throughout the range of the AN languages and proceeded in a variety of different ways. In fact, one language may reflect one and the same AN root in a number of ways. Specifically, *kan took many forms, one of which was lengthening of the vowel, and even the process of syllable internal lengthening was different in different languages: the disyllabic reflexes do not uniquely reflect *kaen. Tg reflects káʔin, showing a lengthened root vowel raised in the second mora with a glottal stop intercalated; Tondano and other languages of northern Sulawesi reflect kaʔan, developing an intercalated glottal stop, but not raising the second mora; Bunun and Amis on Taiwan raise the second mora like Tg, and Bunun intercalates a glottal stop kaʔun, whereas Amis simply has hiatus kaen. In short, the process of disyllabization in these languages was independent, as was the process of glottal stop insertion in the languages that undergo this process. In other words there is no evidence for a PAN vowel sequence in the roots which can be reconstructed as monosyllabics.

1.1.1 More on lengthened monosyllabic roots

PAN *kan is not the only monosyllabic root which shows lengthening in the Philippine languages. Here are some other examples:

- *cuk ‘be inside’
  - Cb suʔuk ‘inside space’, Tg pások ‘enter, insert’
- *dañ ‘old’
  - Cb dáʔan ‘old, not new’, Atayal raral ‘old’
- *kaŋ ‘legs apart’
  - Cb kaʔang ‘walking with the legs apart’, MI terkangkang ‘with legs flung wide apart’
- *luk ‘curved’
  - Tg lúʔok ‘bay’, MI teluk ‘bay’, peluk ‘embrace’
- *pan ‘bait’
  - Tg páʔin ‘bait’, Paiwan pan ‘bait’, Proto-Oceanic *pan-i ‘to bait’

The following languages are among those which do not seem to develop a glottal stop in between two vowels: Itbayatan, Ivatan, Dumagat, Pangasinan, and in the south: Kadazan, Gorontalo.
1.2 The glottal stop in the Philippines as the reflex of earlier hiatus

My view that glottal stop insertion is a late process in Tg is further supported by the fact that Tg and other Philippine languages inserted glottal stop in forms which developed a vowel sequence in their post-Proto-Philippine history. This is exemplified by a sporadic reflex of *l with Tg /?l/, where the *l is lost and a glottal stop is inserted to separate the vowel sequence which developed, e.g., *quleg ‘worm’ > ?u?ód ‘worm’. In summary, these monosyllabic roots show that the occurrence of a glottal stop in Philippine languages reflects an earlier hiatus that developed from the addition of a mora to a root vowel.¹

Philippine /?l/ has another source as well: PAN *q becomes /?l/ in languages which develop this intervocalic /?l/. Thus, there are cases where /?l/ reflects *q and cases where it does not reflect *q. I hypothesize that in the cases where the Philippine languages show a glottal stop where other languages do not reflect *q, there was an earlier hiatus (by which I mean a sequence of unlike vowels or a single vowel of two morae).² In addition to the monosyllabic roots which developed vowel sequences by stretching the nucleus cited in §1.1ff. above, there are roots of more than one syllable in which a glottal stop is reflected where no *q can be reconstructed (that is, no reflex of *q is reflected in AN languages which show a reflex of *q other than /?l/). In some cases this [?] occurs between two vowels in the Philippine languages, but in other cases it occurs in a consonant cluster. In no cases of glottal stop insertion is there motivation for the development of hiatus in post-PAN times. That is, no process of syllable lengthening can be documented for any roots other than monosyllabic roots (cf. §1.5). The conclusion is that these Philippine glottal stops reflect a feature of the PAN phonology. Examples are given in §1.3. In cases where the glottal stop appears in Philippine languages intervocalically not from *q, we may hypothesize that the Pre-Philippine etymon contained a vowel sequence and had the same number of syllables as there are in the attested forms. In cases where the glottal stop occurs pre- or post-consonantally and not from *q, we hypothesize that the Pre-Philippine etymon was trisyllabic. That is, we assume an earlier occurrence of a short (unstressed) vowel between the consonant and the [?]. The short vowel was subsequently lost by syncope (a normal development of medial short vowels in trisyllabic roots). The evidence which supports the hypothesis of a trisyllabic root is presented in footnote 7.

The hypothesis of glottal stop insertion between two vowels is given further support by developments in root-initial position: in root initial position in the

---

¹ This sporadic loss of a reflex of *l is probably due to dialect mixture. Further, not only Tg /?l/ is intercalated in vowel sequences but also Tg /h/: e.g. a variant of the reflex of *quleg ‘worm’ /?u?ód/ ‘worm’ is /?u?ód/. Other examples: *balay ‘hall’ > báhay ‘house’. Further, other vowel sequences which develop sporadically get separated by /?l/: *bayi ‘woman’ > babá?e (with reduplication of the first syllable), *taw ‘man’ > tá?ó ‘man’. The intercalation of /h/ does not happen in root-initial position. It is a sporadic process medially and at the end of a root, but in some languages (e.g. Tg) intercalation of [h] is what normally happens to the exclusion of the insertion of [?] at the end of a root before the addition of a vowel-initial suffix.

² The Philippine languages only reflect hiatus. If we compare cognates from languages in Taiwan, we must conclude that a laryngeal was lost in Proto-Philippine (or possibly individually in all the Philippine languages) — cf. §1.4, below.
Philippine languages, the insertion of [ʔ] is automatic. That means, these languages show initial glottal stop in roots which had PAN *q- and also in those roots which began with a vowel. In short, the insertion of intervocalic glottal stop is a further example of the same process which took place in word-initial position.

To summarize, syllable-initial vowel is reconstructed for Pre-Philippine when the Philippine languages manifest [ʔ] corresponding to syllable-initial vowel in languages that reflect *q in contrast with its absence. This means that syllable-initial vowel is reconstructed not only root initially but also in the middle of a root when Philippine languages manifest VʔC corresponding to VC in other languages and also when they manifest CʔV corresponding to CV in other languages.

1.3 Examples of intercalation of glottal stop

The most widespread examples of vowel-initial syllables are at the beginning of roots. Medially, the examples are limited. The following forms are identified as having non-initial syllables beginning in a vowel (that is, internal vowel sequences not separated by a consonant): *bacueq ‘wash’, *beyeat ‘heavy’, *bitieq ‘calfes’, *bueni ‘ringworm’, *yabiñ ‘night’, *isewab ‘yawn’, *kanuec ‘squid’, *kaewit ‘hook’, *kua ‘what’, *lawen ‘long (time)’, *tieyeb ‘belch’, *tinueq ‘weave’ (*tuen + *-in-; *tuen < *tun). 6

As we say above, it is the occurrence of a glottal stop in the Philippine languages as opposed to the absence of any reflex of *q in other languages which enables us to identify these roots. An example of Philippine /ʔ/ from *q as opposed to Philippine /ʔ/ from hiatus are the following two forms, the first of which exemplifies the reconstruction with syllable-initial *q- and another which exemplifies reconstruction with syllable-initial ?: in Cb bagʔu ‘new’ the /ʔ/ reflects *q and we can reconstruct *q because other languages also show reflexes of *q (e.g., Bunun vaqlu ‘new’, Tongan faʔou ‘new’, etc.). However, for Cb bügʔat the cognates in other languages do not reflect *q (e.g., Tongan mam-fa ‘heavy’ without the /ʔ/ which appeared in the word for ‘new’; Old Javanese weat ‘heavy’ — Old Javanese reflects *q with /h/, but here there is no /h/; Moken beat ‘heavy’ — where Moken reflects *q with /k/, etc.). We conclude that the Cb form bügʔat is from an earlier *bugeʔát (by a rule of syncope which applies in the unaccented penults of trisyllabics in Cb and many other languages). 7 This form *bugeʔat in turn is from *beyeʔat which developed from an earlier *beyeat when a [ʔ]

6 *tun is widely reflected. *tinueq is in fact derived from a root *tun which has been disyllabized by lengthening the nucleus (> *tuen to which -in- has been added). Evidence for *tuen is Cb tūʔun ‘be directed toward a focus’. Evidence for *tun is Tg tunton ‘give guidance/direction’, Tondano tonton ‘lower on rope’, Old Javanese tumtun ‘rope for guiding animal’. *tun can be reconstructed to mean ‘lead on rope, thread’ and *tinueq ‘weave’ is derivative of the meaning ‘thing lead on thread’. This etymology was first proposed by Blust (1976) and quoted by Mahdi (1988: 104). However, Mahdi and Blust propose *teun, presumably on the basis of forms like Bunun tinʔun ‘weave’. I propose that *tun became *tuen by lengthening the nucleus, with the off-glide at the end of the nucleus rather than the beginning, for in all other cases of lengthened vowels it is the second mora which slides upward, not the first.

7 Trisyllabics with short penult in PAN had stress on the final syllable. The evidence for this is threefold: (1) various languages located across the board from Taiwan to Oceania manifest apocope of the first syllable, which is inconsistent with stress on first syllable of the root; (2) many Philippine languages have stress on the final syllable in roots derived from trisyllabics with a short medial vowel; (3) evidence in Taiwan languages of the reflex of *t in reduplicated monosyllabics, the details of which cannot be discussed here.
was automatically inserted between two adjacent vowels. The same process of glottal stop insertion and syncope explains the reflexes of *kaewit: Cb kawitan 'hook', ML kait ‘connect’ (where *q would have produced /h/ in ML); and *bueni ‘ringworm’: Cb buní 'ringworm', Tg buni 'shingles', and Proto-Polynesian *pune (where a /h/ would have occurred if there had been a *q in PAN).

In the case of *lawen, there was no syncope as there were only two syllables, but a vowel sequence developed when *we became monophthongized. MI and Muna retain vowel sequences, but the Philippine languages develop /?/ between the two vowels: Cb la?un 'aged, matured', MI laun 'long time', Muna lao 'long (of the dry season)'. The vowel sequence which developed in MI and Muna reflexes of *lawen can be compared with the *q reconstructed in *taquweñ 'year', where the Philippine languages develop /?/ and MI and Muna clearly reflect PAN *q (MI manifests /h/ and Muna manifests /?/): Tg ta?un, MI tahun, Muna tayu ‘year’.

In the case of *bitiic ‘calf’: Cb batíis ‘calf’, and *yabií ‘night’: Cb gabíi ‘night’, a stressed penult is reconstructed, and there was no syncope in Cb (but there was in Samareño and dialectal Tg gabí ‘night’). A glottal stop was inserted between the vowel of the penult and the beginning of the final syllable. There is no evidence of a PAN *q in either of these roots.

In the case of *tinuen ‘weave’, there are no reflexes in the languages of the Philippines which reflect hiatus with an inserted glottal stop. However, some languages of Taiwan which sporadically insert [?] in vowel sequences reflect *tinuen or *tuen, e.g. Bunun ma-ti?un ‘weave’ (from an earlier *tine?un by vowel metathesis and syncope of the medial syllable), Amis mi-ti?un ‘weave’). There are other forms showing this glottal stop insertion discussed in §1.4, below.

1.4 Was there hiatus or were there PAN laryngeal phonemes?

Corresponding to the hiatus Pre-Philippines (or /?/ not reflecting PAN *q) the following languages from Taiwan reflect /h/: Atayalic, Pazeh, Saisiat, Bunun, and Amis, but not between like vowels. Between like vowels there is hiatus in Amis and in Rukai (Budai dialect), and contraction in the other languages. There are two possible interpretations: (1) that the /h/ developed in these languages as a transitional phenomenon; (2) that the /h/ reflects a PAN phoneme *h which was lost in the other AN languages of Taiwan and also in the languages outside of Taiwan (with contraction

8 We reconstruct *be?eat rather than **be?eat, both of which would develop into Cb bug?at. (The contrast between /C/ and /Ci/ is lost in Cb.) I reconstruct *be?eat on the hypothesis that Tg reflects a contrast between /Ci/ and /Ci/. The /?/ is lost in most dialects of Tagalog, but there is compensatory lengthening of the vowel preceding the /?/ when the /?/ is lost. However, there is no compensatory lengthening when a /?/ following a C is lost: e.g. kāmin ‘rice’ (< *ka?men < *ka?menen) as opposed to pusūn ‘lower abdomen’ (< *pusūn < *pusu?un). Thus, Tg bigāt ‘heavy’ indicates earlier *big?at (< *be?e?at) rather than **bi?gat. Similarly, we reconstruct *kaewit, rather than **kawit on the basis of Tg kāwit ‘hook’ (< *ka?wit) and *bueni rather than **bunei on the basis of Tg buni ‘shingles’ (< *bu?ni).

9 I assume a general rule which affects a large number of languages across the board, that the sequences ‘we and *ew in almost all positions become /u/. Except for certain metrical environments (whose nature I have not yet managed to determine) *ew > /u/ almost everywhere except in some of the languages of Northern and Central Taiwan.

10 Most current MI dialects do not retain /h/ between unlike vowels, but this is a recent development, and we are referring to an older stage of MI reflected in the writing system in which the /h/ is still retained. Indeed there are dialects currently spoken which continue to maintain /h/ between unlike vowels.
taking place in most of them, but not in many of the Philippine languages). It is this second explanation which best fits the facts. Namely, there is general evidence (archeological as well as linguistic) of the spread of the AN languages from Taiwan southwards and outward (Pawley 2002, §2ff. 256–8). This evidence indicates not only that the extra-Taiwan languages are in a subgroup but also that Atayalic, Pazeh, Saisiat and Amis themselves are the descendents of primary and separate off-shoots from PAN. If this is so, then it should be deduced that intervocalic \( /h/ \) is a retention and not an innovation.\(^1\)

A second reason for hypothesizing \(*h-\) rather than hiatus is that PA\(^2\)\(*h-\) also must be reconstructed in initial and final position (although the Philippine languages provide no evidence for the reconstruction of this phoneme in those positions).

The following forms exemplify PA\(*h-\):

- \(*bahañi ‘panel’\) Bunun \(banhil ‘cypress’\), Amis \(fahdil ‘cedar’\), Rukai \(baalhi ‘cypress’\), Tg \(páñig ‘panel’\)
- \(*buhet ‘squirrel’\)\(^3\) Atayal \(bhuT, Pazeh bhut, Saisiat kahboet, Amis fohet ‘squirrel’, Tg \(bú’ot ‘rabbit’\)
- \(*láhud ‘seaward’\) Pazeh \(rahut ‘seaward’, Saisiat lahaer ‘downhill’, Cb \(lawud ‘sea, seaward’ (*h is replaced by glide between high and low vowels)\)
- \(*yabihí ‘night’\) Amis \(lafii ‘midnight’, Cb \(gabí?i ‘night’\)

1.5 Could the reconstructed vowel sequences in fact have been single vowels which were later expanded?

In §1.1, I argue that the roots such as \(*kan were monosyllabic and not disyllabic \(*kaen, as others have posited. The analogous question could be asked of the trisyllabic roots, listed in §1.3. These are not attested in languages of Taiwan and provide no evidence for \(*h. Could these roots be reconstructed as \(*bacuq, \(*beyat, \(*buni and so forth; that is, could it be possible to assume that the reconstructed trisyllabic forms which explain the glottal stop in the Philippine languages were in fact developed by a process of lengthening, as in the case of the monosyllabic root? Such a scenario is theoretically possible, but without a phonological motivation for lengthening one of the two syllables (such as in the case of the monosyllabic roots), we must assume that the reconstructed roots indeed had three syllables.\(^4\) This assumption is supported by the fact that we are obliged to reconstruct trisyllabic roots in forms that are not attested in

---

\(^1\) This is essentially the view presented by Zorc (1982). However, Zorc proposed the reconstruction of two laryngeal phonemes. The extra putative phoneme is based on a failure to recognize the PAN reflexes in Saisiat and taking non-contrastive attestations of \(/h/\) and \(/ʔ/\) in Bunun to be contrastive. There is only one set of correspondences involving \(/ʔ/\) in the Philippines and \(/h/\) (or hiatus between like vowels) in these languages of Taiwan.

\(^3\) Tsou \textit{buhetsi ‘squirrel’} is a borrowing from an unknown source, in which PAN \(*-h-\) was reflected as \(/h/\).

\(^4\) Or if they had two syllables, one of the two syllables must have been bimoraic. It is probably correct to reconstruct all of these forms as trisyllabic and not disyllabic with one bimoraic syllable. In Cb and other Philippine languages, the vowel sequences in this kind of root are treated differently from the vowel sequences which developed from monosyllabic roots which became disyllabic by developing a bimoraic root vowel. The best example is the word for ‘calf’, which I reconstruct as \(*beťes (Cb manifests \textit{batí} ‘calf’ where the final \(*e is assimilated to the preceding \(*i, whereas \(*pit is reflected as \textit{piʔut ‘narrow, not afford ing enough space’, developed from a bimoraic \(*piet without assimilation). Not all Philippine languages treat the two types differently. For example, Mongondow manifests \textit{bosiʔot ‘calf} with no assimilation.
current languages with vowel sequences (§2). Inasmuch as roots with /ʔ/ not from *q in the Philippines correspond to /h/ < PAN *h in languages of Taiwan, we should assume that the roots with glottal stop not from *q ultimately derive from roots with *-h- at an earlier stage, even though no attestations in languages from Taiwan have been found.

1.6 Were glides contrastive in vowel sequences consisting of high and low vowels?

The evidence that a vowel sequence containing a high vowel contrasted with a sequence broken by a semi-vowel, *y or *w, is slender. That is, there is only very little evidence for a contrast between *ua and *uwa, *ia, and *iya, etc. The Philippine languages manifest /uwa/ which presumably reflects *uwa, as in the following forms:

*banuwa ‘continent’  
Cb banwa ‘forest’

*binuwà ‘k.o. tree’  
Cb binuwang ‘k.o. tree (Macaranga tanarius)’

*buwaq ‘fruit’  
Cb buwa? ‘endosperm of germinating  
coconut’, Ilk bowa ‘betel palm and nut’, MI  
buah ‘fruit’ (never with medial /h/)

*quway ‘rattan’  
Cb ?ʻuway ‘rattan’, Ilk way ‘rattan’

However, the evidence that a sequence /uʔa/ in the Philippine languages reflects PAN *ua is slender. There is only one form in which Philippine languages manifest /uʔa/ and which are possibly cognate with forms that cannot reflect PAN *q. This correspondence indicates a sequence *ua. There is one other form that may possibly provide further evidence:

*kua ‘say’  
Cb kuʔan ‘filler word’, Old Javanese kwa ‘like  
this’ (without /-h-/ < PAN *q)

*yuang ‘hole, depression’  
Tg guʔang ‘hollow, empty inside’,14 Old  
Javanese rong ‘hole’ (not directly inherited,  
but not reflecting *-q-), MI ruang ‘hollow  
space’ (never with /-h-/)

There is practically no evidence for a similar contrast in other sequences. If there was indeed a contrast between *uwa and *ua then surely the corresponding contrast between *iya and *ia must have existed, but I have found only one form which might possibly provide evidence. This is the following form, where Amis shows that the PAN form did not have *q and Cb shows a glottal stop between the sequence [ia] which developed after metathesis:15

*isewab ‘yawn’  
Cb huyʔab (<*suia, with metathesis of the  
first and second syllables), Amis sowab, Rukai  
ma-sewasewab ‘yawn’

Except for this form the Philippine languages reflect /iʔa/ only where there is clear evidence of PAN */iqa/ and therefore provide no evidence for a contrast. There are forms in the Philippine languages reflecting PAN forms which we reconstruct with *iya,

---

14 However, Tg is not a reliable witness, for the contrast between /-ʔ-/ and its absence was lost at a certain point of time in Tg dialects. Later the contrast obtained again in all dialects, and /-ʔ-/ was introduced in environments in which it etymologically had never occurred.

15 I hypothesize that Amis initial syllables with /ʔV/ from *qV do not get lost. The loss of the initial syllable in Amis indicates that the root had a vowel initial.
but these could as well be reconstructed with *ia.\textsuperscript{16} There are no PAN forms which could be reconstructed with *ai or *ayi, or with *au or *awu.

The Philippine languages provide no evidence for a contrast between *iu and *iyu. Forms which can be reconstructed with one or the other of these sequences reflect *iyu (sometimes syncopated to *yu) in the Philippines. Bunun, Amis, and Puyuma from Taiwan show two reflexes for the forms which reflect *i followed by *u, i.e., indicate that PAN possibly had a contrast between *iyu and *iu, but the evidence is not unequivocal.\textsuperscript{17} Two forms seem to reflect *iyu:

\begin{itemize}
  \item \textit{bayiyus} ‘storm’ Amis \textit{faliyos} ‘typhoon’, Puyuma \textit{yayiw} ‘typhoon’, Bunun \textit{balivus} ‘storm’,\textsuperscript{18} Tg \textit{bagyo} ‘typhoon’ (with syncope)
  \item \textit{qiyut} ‘sexual intercourse’ Bunun \textit{paquit} (with metathesis prior to the development of the /w/ between the *i and *u), Puyuma \textit{maha-hiyut}, Cb \textit{?iyut}, Tg \textit{?iyot}, Ilk \textit{yut} ‘have intercourse’
\end{itemize}

One form seems to reflect *iu:

\begin{itemize}
  \item \textit{lius} ‘turn around’ Tongan \textit{liliu} ‘turn around’, Amis \textit{maliun} ‘revolve’, Puyuma \textit{mulius} ‘turn around’, Cb \textit{liyu} ‘circle around’, \textit{saylu} (sa-\textit{lyu} with syncope and metathesis) ‘pass by’
\end{itemize}

2. Syllables Ending in a Consonant and Trisyllabic Roots with Short Medial Vowels

In PAN root-final position any consonant is free to occur. However, in a root internal syllable no final codas occurred.\textsuperscript{19} This proposition is based on the assumption that PAN had trisyllabic roots with unstressed *e as the vowel of the penult (or had unstressed [u] before [w] and unstressed [i] before [y]). This penultimate vowel in some cases later became lost by syncope. The Philippine languages mostly show syncope when the penult is short (not stressed), and the evidence for the existence of trisyllabic roots in PAN, rather than consonant sequences in between which an epenthetic vowel developed, comes from the fact that there were two ways in which these roots were disyllabized: in some languages these roots were disyllabized by loss of the initial syllable, whereas in others they were disyllabized by syncope of the medial syllable. For example, we reconstruct *qapegu ‘gall, bile’ rather than **qapgu; *tuqeläñ ‘bone’ rather than **tuqlañ. The form *qapegu is disyllabized in Tg by syncope of the

\textsuperscript{16} For example, I reconstruct *iya in the following forms, but these might as well have been with *ia: (i) *iya ‘he, she’: Cb s-iya ‘he, she’, Ml ia ‘he, she’ (never with /-h/-); (ii) *iyak ‘shout’: Tg ?iyak ‘cry’, Ml ter-ia ? ‘shout’ (never with /-h/-); (iii) *tiyañ ‘belly’: Cb tïyañ ‘belly’, Ml tïtun ‘womb’ (never with /-h/-).

\textsuperscript{17} The evidence is only from three forms. Authors transcribe Amis and Puyuma forms some with /iu/ and others with /iyu/, but this may well be an inconsistency in transcription and not reflect a phonemic distinction in these languages. Furthermore, as the data show, not only do we have to assume that there is a contrast between /iu/ and /iyu/ in these languages, but there are other assumptions as well that must be made to explain two different transcriptions.

\textsuperscript{18} I am assuming that PAN *iyu > Bunun /ivu/. In Bunun, post-vocalic *w > /v/.

\textsuperscript{19} Prefixes ending in a consonant are reconstructed. In other words, within a word consonant clusters occurred, but not within a root.
medial syllable - 'apdo ‘gall’; in Toba Batak, with loss of the initial syllable - pogu ‘gall’.

There is further support in the case of *qapegu and other reconstructed roots which are reflected with a consonant cluster in some attested languages, in that not all languages disyllabize the root. Some languages still manifest a trisyllabic root. For example, *qapegu is reflected with MI hamperu, Paiwan qapedu, Western Bukidnon Manobo epem ‘gall’, etc. In the case of *tuqelañ, there are no examples of loss of the first syllable, but there are trisyllabic forms that are reflected in attested languages which otherwise have consonant clusters, e.g., Paiwan tsuqelañ ‘bone’. There are many forms which like *qapegu provide evidence from loss of the first syllable in the occurring reflexes that the PAN form was trisyllabic, and there are cases like *tuqelañ for which the existence of trisyllabic reflexes, even in languages which have consonant clusters, is the only evidence.

Weakening of the first syllable, i.e., centralization of the first vowel of the root (similar to loss, described immediately above), also provides evidence that the attested form reflects a trisyllabic root. For example, in the case of *banuwa ‘land’, a trisyllabic root must be reconstructed to account for MI benua ‘continent’, although the Philippine languages show syncope of the penult, e.g., Tg banwa ‘village’. Another example is found in *baqeyüh ‘new’. Here there is evidence for a trisyllabic root in the occurrence of reflexes with weakened first syllable in many languages, e.g., Tondano weru ‘new’, and also in the occurrence of trisyllabic roots over a wide range of languages, e.g., MI beharu (dialectic), Tongan foqou, Sa’a haalu ‘new’.

However, root-internal nasal clusters, excepting clusters with *ñ, can be reconstructed for PAN. The only form with root internal *ñ which is reflected in an attested language as part of a cluster with a following consonant is the form *qañegaw ‘day, sun’, which cannot be reconstructed as **qañgaw. The initial syllable is lost and the medial vowel is retained in some languages of Indonesia: Bugis eso ‘sun, day’, Tondano édo ‘sun’, Manggarai leso ‘sun, day’, Ratahan law ‘day’ all reflect *negoegaw with loss of the initial *qa-. In other languages this root becomes disyllabic by loss of the medial vowel: e.g., Ngaju Dayak andaw ‘day’, Cb ʔadlaw ‘sun, day’.

In the case of doubled monosyllabic elements (C₁V₁C₂C₁V₁C₂), it is assumed that such forms were actually a case of doubled monosyllabic roots, and in those cases C₂ was any consonant. There is some evidence that doubled monosyllabic roots were in fact trisyllabic; that is, an epenthetic vowel developed between the two monosyllabic elements if they ended in a C. For example, in MI doubled roots with /a/ are reflected

---

20 It is the fact that the first syllable is lost in some languages that constrains us to reconstruct a trisyllabic root, not the occurrence of a trisyllabic root in some of the currently attested languages, for it could well be that the vowel of the middle syllable in those cases was inserted by a process of epenthesis. However, loss of the first syllable implies the existence of a previous trisyllabic root. Loss of the first syllable is sporadic and scattered and in no way defines a subgroup. In short, there must have been a trisyllabic precursor for all cases manifesting weakening or loss of the first syllable. Furthermore, roots which have no reflexes other than those with medial consonant clusters can well have developed from trisyllabic forms, even if no trisyllabic form or form reflecting loss of the first syllable is attested in the current languages. In short, it is not possible to reconstruct consonant clusters for PAN.

21 This is not an especial characteristic of the Philippine languages. In fact, this sort of evidence comes largely from languages outside of the Philippines, and the Philippine languages provide only a few examples.

22 *ñ was very likely not a nasal but rather a lateral. The reason for this is not only that its reflexes are laterals in most languages of Taiwan, but also they are laterals in some environments in languages outside of Taiwan (Wolff 1993).
with an /e/ in the first syllable: *tactac ‘rent, cut through’ > Ml tetas ‘cut through’. The /e/ of the first syllable developed by a rule that the vowel of the antepenult is weakened. The Ml development must have been *tac > *tactac > *tacetác > *tesetás > tetas. Similarly, *dap > *dapidap ‘Erythrina sp.’ > dedap ‘Erythrina sp.’ In fact this is the case of all examples of reduplicated monosyllabic elements with a vowel nucleus in *a in Ml (but not those with high vowel nuclei). The Philippine languages support the assumption of the existence of the epenthetic vowel indirectly. There were two root stress patterns in PAN: final stress and penultimate stress. In some cases the reduplicated root developed penultimate stress. Furthermore, trisyllabic roots with stress on the final syllable (cf. footnote 7) are syncopated in the Philippines, but not those with penultimate stress. Examples are as follows:

*bic ‘sprinkle’
Ilk, Pangasinan, Cb bisibis ‘sprinkle with water’ (< *bicébic) and also Tg balisbisán ‘gutter, rain-trough’ (< *bicabic)

*ŋuc ‘growl’
Bikol ngusngus (< *ŋucẹŋuc), Ilk ngusúngus ‘growl’ (< *ŋucęŋuc)

*sap ‘g Grope’
Paiwan sapsap ‘dig, scratch surface’, Amis sapsap ‘feel with hands’ (< *sapesáp), Samar-Leyte hapūhap ‘look for by probing’, Bikol hapūhap ‘g Rope in dark’, Cb hapūhap ‘stroke’ (< *sapésap)

Thus, the existence of reduplicated monosyllabic roots with a stressed epenthetic vowel between them is further support for a theory that says PAN roots which consisted of a monosyllabic root that had been doubled were made trisyllabic with an epenthetic vowel between the two monosyllables.23

In the case of roots with a medial *ɣ or *w, that is, roots with a consonant preceded by a *ɣ or followed by a *w, usually there is no evidence in the attested reflexes for the reconstruction of a trisyllabic root. Nevertheless, no cluster is reconstructed, i.e., we reconstruct medial *Cuw (rather than medial **Cw) and medial *leC (rather than medial **leC). One reason to support this is that [ɣ] and [w] in articulatory terms do not share a characteristic which no other sound has. But more important, there are cases where evidence for PAN medial *Cuw or medial *leC exists. That is, we must reconstruct *banuwa on the basis of Ml benua ‘continent’ (see above), and we must reconstruct *sayejañ ‘ladder’ on the basis of Ngaju Dayak rejan ‘ladder’ (with apocope of the initial syllable). That means that other attested roots with a medial consonant followed by a *w or a medial *ɣ followed by a consonant must be reflexes of trisyllables.

23 I do not assume the reconstructed proto-language to be an invariant language spoken at a single point in time. I believe it is necessary to assume that the PAN forms which reflect currently attested forms may indeed have occurred as variants (that is, have been part of an on-going change), have come from different points in time, or have existed as dialectal variants in the proto-language. The fact that the first vowel is weakened in Ml in the case of doubled monosyllabic roots with /a/ but not in roots with the high vowels *i and *u indicates that the development of this epenthetic vowel was an on-going process.
even in cases where there is no evidence for that in the attested reflexes.\textsuperscript{24} There are very few examples of forms for which no evidence for a trisyllabic root in PAN appears in one or another attested language. Some examples are *byeçay ‘oar’, *buyenay ‘Antidesima bunius’, *dasuwen ‘leaf’, *qayuwaç ‘k.o. mullet’.

\textsuperscript{24} The alternative would be to assert that there was a contrast between medial sequences (a) \textsuperscript{*}yeC and (b) \textsuperscript{*}yeC, and between (a) \textsuperscript{*}Cuw and (b) \textsuperscript{*}Cw. Such an assertion is contradicted, however, for a medial vowel must be reconstructed in some cases of (a), but nothing requires the reconstruction of (b); there is no evidence against (a) in any root. In fact, there is evidence for a trisyllabic root (a), in the case of the vast majority of roots with medial \textsuperscript{*}ye and \textsuperscript{*}w, but the language or languages which provides the evidence are different for each root. For example, in the case of *byeçek ‘bundle’, Bunun \textit{ma-luku} ‘tie up’ (with loss of initial syllable) provides the evidence for case (a). In the case of *bayçqäj ‘molar’, the evidence comes from Kelabit \textit{bera’ang} ‘molar’, whereas Bunun manifests a syncopated reflex \textit{dalqam} ‘molar’, and so forth.
References


