

PROTO AUSTRONESIAN ACCENT REVISITED

R. DAVID ZORC
Senior Lecturer
School of Australian Linguistics
Darwin Community College, Australia

1. INTRODUCTORY SURVEY

A decade has now passed since my initial attempt (Zorc 1972)¹ to encourage fellow Austronesianists to look into accent in AN languages, and to determine how far back the phenomenon found in diverse Ph languages can be reconstructed. Since that time other researchers have demonstrated that accent plays a role in the historical development of their respective languages. Recently, Dahl has called together a team to look at AN accent (Biggs, Cowan, Dahl, Li, Tsuchida, and Zorc). Hence, a brief survey of recent work and hypotheses is in order as a manifesto for further study.

Zorc (1972) showed that accent needed to be reconstructed for Proto Tagalic (PCP), and sought to explain how accent on some forms may have arisen [e.g. *ə in the penult resulted in oxytonality; certain form classes (pronouns, numerals, interrogatives, deictics, and vocatives) also were oxytone], but that there was a residue of oxytonal forms which could not be explained and appeared to contradict the long-standing assumption that stress or accent fell on the penult in PMP or PAN.

Zorc (1978) again challenged the penult-accent assumption, listing a few (but diverse) AN languages where stress regularly falls on the ultima. The study demonstrated that penult vowel length or shortness must be reconstructed for at least PPH, that geminate consonants might be the result of a preceding short vowel, and that morphological accent (on vocatives and statives, etc.) was at least necessary for PHN. Although data supporting 153 reconstructions were presented, an appendix included over 500 etyma with accent, derived from and justified by cognates cited in Reid (1971) and McFarland (1977).

Zorc (1979)² showed how contrastive accent (vowel length) can develop in a language, and demonstrated that Pangasinan was basically oxytonal, with phonemic length introduced by the loss of the first member of a consonant cluster, parallel to that on some Tagalog forms [observed earlier in Zorc (1972:46)]; thus PPH *baqRuh 'new' > Png ba:lu, Tag ba:go.

Cowan (1974 and in press) considers the reconstruction of vowel quantity essential in deriving synchronic minimal pairs and other forms in Acehnese (see section 7 herein).

Palleesen (1979)² showed how the split of PAN *ə into three Proto Sama-Bajaw phonemes could be accounted for if accent was posited for pre-PSB:³ (1) PAN³ *ə > PSB *u, (2) PAN³ *əCV > PSB *əCCV, (3) PAN³ *ə₂ > PSB *a. His study has demonstrated that yet another AN subgroup has gemination after *ə, but also that *ə could be accented,⁴ and would not then yield gemination; thus, PAN *iəŋaq > PSB *iəŋaq 'middle (range)', but PAN *iəŋaq > PSB *təŋaq 'middle (point)'.

¹I feel that Zorc (1972) was a clumsy first attempt, and I disagree with the shape of many of the reconstructions presented therein. Many of these have been modified or corrected in Zorc (1977: 50f, 216-19), where the term 'Proto Central Philippine' replaces 'Proto Tagalic'. However, the correction of PBS, PCP *? < PAN *q as opposed to PAN *?, was not taken up until Zorc (1981).

²Although published in 1979, each of these articles was written considerably earlier (circa 1975), and therefore preceded Zorc (1978).

³Palleesen (1979: footnote 9): 'The label PAN in this paper, when it refers to the stress hypothesis, should be read as shorthand for "PAN or some lower proto-language of Proto-Hesperonesian".'

⁴I erroneously did not take up this point in Zorc (1978), although it is discussed in Zorc (1972:50 and 1977:218), positing a PCP *bə:təŋ 'to pull' and PCP *bə:təŋ 'young coconut' (i.e., pulled off before it is ripe). I take at least some of the accent pairs for Proto Sama-Bajaw to be indicative of what I have called 'morphological use of accent' (1978:91ff).

Dahl (1982) reviews my 1978 paper, surveys accent in several AN languages, and makes explicit the reconstruction of short vowels where gemination appears.⁵ He also explains more of the Madurese geminates (correctly dismissing nasal accretion),⁶ and introduces Makassarese and other South Sulawesi speech varieties as criterion languages.⁷ However, Dahl reconstructs long vowels on the basis of absence of gemination in Madurese, which does not agree with oxytonality reconstructed for PPH.⁸ It is clear that some of the accent-marking conventions introduced in Zorc (1978) require revision (see section 3 below). I fully agree with Dahl (1982:112) 'that the geminates have developed after short penultimate vowels The quantity feature has, so to speak, moved from the vowel to the consonant'. [See sections 8.2 and 8.3 herein.]

I am pleased that all studies of individual AN languages or subgroups that I have received or seen in the last five years clearly indicate the role of accent. If it predictably falls on a given syllable, this is specified in the introduction; if it is unpredictable, it is indicated on the individual entries. For example, in a study of the Sangiric languages, Sneddon (to appear) discusses word stress in each speech variety, and offers a particularly astute explanation of how stress contrasts evolved in Ratahan: 'stress shifted from a penultimate high vowel to an immediately following vowel in a final closed syllable (*Ríud > Rth iúr 'to pull', tían > Rth tián 'belly' . . .). This stress shift operated before the loss of *h, reflecting PSan *R, in Rth. Subsequent loss of *h resulted in stress contrasts (*níuR > Rth níu 'coconut', *níRu > Rth níu 'winnowing pan', *búat > Rth buá? 'to stand up', *súRat > súa? 'fish poison').⁹ Another fine study is that of Adelaar (1981) on Proto-Batak, wherein the role of accent for each Batak dialect is discussed.

2. SOME NOTES ON FAITH

In a recent review I noted:

The decisiveness of any solution . . . ranges from near universally-accepted changes (metathesis, assimilation, dissimilation, etc.) to the setting up of additional correspondence sets, and from over-reliance on the principle of economy to an act of faith in a particular reconciliation of divergent forms (can one reconstruction do? or should doublets or different reconstructions be made?) . . . But this is precisely where the act of faith comes in: Scholars can offer alternative solutions, and those that seem the most economical *and* reasonable should be the most believable. (1981a:45)

⁵Zorc (1978:97f) suggested that this area should be studied, but did not make explicit the reconstruction of *V when gemination was observed.

⁶Thus, PHN *búká? 'open', PAN *Ca?ás 'above', PAN *Sásí? 'contents' 'all. . . had short penult in PAN' [Dahl (1982:110)]. And I agree.

⁷Based on Mills (1975 – which to date I do not have access to) and his own research, I will not add to the list of PHN, PMP, or PAN reconstructions with penult *ǝ herein, because I consider this to be sufficiently established. However, some of the 'new reconstructions' (beyond those cited in Zorc 1978) would include: PHN *bǝnǝn 'thread' (Akl bunǝn, Ml bǝnǝn, Mkr bannǝn), PMP *pǝñuh [turtle] (Ib pǝñu?, Mkr paññu, Tbl hǝnuh, Sml pǝnno), PHN *qǝlǝt 'space, interval' (Tbl kǝlǝt 'in succession', Ml hǝlat, Md ǝlla?, Mkr alla?, Akl ?uhít). Some new reconstructions based on gemination in Mkr after a vowel other than *ǝ are presented in section 5 (herein).

⁸Dahl's PAN *a:ku [PPH *?akú (#139)], *t'i:da [PPH *sidá (#144)], *ki:t₁a [PPH *kita (#47)], *a:nu [PPH *?anǝh (#148; see Zorc 1981:#P112)] I do not feel that lack of gemination in Madurese proves length in PAN, and oxytonality alone in the Philippines does not prove shortness (see section 3), hence the errors in Zorc (1978) and perhaps in Dahl (1982). I do accept Dahl's PAN *a:lak and *i:naH as possible base forms, with PPH *?aná and *?iná, *?iná? as resultant vocatives.

⁹The importance of Sneddon's analysis cannot be underestimated. The agreement of Rth tián 'belly' with PPH *tián, Rth níu 'coconut' with PPH *níuR, and Rth níu 'winnowing pan' with PPH *ní:Ru is accidental, and clearly not a case for sporadic retention of vowel quantity.

In an article (in press), Blust says:

During the past decade several writers have proposed that PAN phonology be revised in various ways. In my view most of these proposals suffer from serious methodological inadequacies (Dahl 1976, Wolff 1974, Dyen 1978 and — to a lesser extent — Prentice 1974) or basic incompatibilities with the evidence (Zorc 1978).

Whereas Dahl has come to *believe* in the necessity to reconstruct accent, it is clear that Blust has not. In a personal correspondence, Blust sent me data on Kelabit and some other languages that did not conform with Zorc (1978 and 1981: footnote 24). Indeed, there was an incompatibility with the Ph evidence alone, for I sought to account for three phenomena with two:

- | | | | |
|-----|------------------------------------|------------------------------|----------------------------|
| (1) | Ph C \check{V} :CV(C) | [long/accented penult vowel] | < PPH *CV:CV(C) |
| (2) | Ph C \check{V} C \check{V} (C) | [accent on ultima] | < PPH *C \check{V} CV(C) |
| (3) | Ph/Md CVCCV(C) | [consonant gemination] | |

I hope to offer herein more plausible evidence that accent contrasts did obtain in PAN, and that there is a need to reconstruct: (1) vowel length (*V:), based on the evidence of Ph languages which is not contradicted by evidence from other AN languages, (2) vowel shortness (* \check{V}), based on gemination (or other strengthened reflexes) of consonants following such vowels, (3) accent on the ultima (* \check{V}_1), based on the evidence of Ph languages where neither length or shortness in the penult can be established, but oxytone roots appear, and (4) unknown (*V), where evidence of any accent pattern is ambiguous, contradictory, or lacking. Two points should be understood.

Firstly, the basis of my 'faith' in the reconstructability of accent (vowel quantity, and possibly stress) is the overwhelming agreement of Northern and Central Philippine languages that share hundreds of cognate forms and inflections, which cannot plausibly be explained as the result of the loss of a consonant phoneme, borrowing, or convergence (= 'drift'). Ultimately, the question of the subgrouping (and time-separation) of languages like Bisayan and Bontok is at issue. Few scholars today believe in a node 'Proto Philippine'. It is clear that the Sama-Bajaw languages are intrusive, and belong to some 'Indonesian node'. The Bashiic (Yami, Ivatan, Itbayaten) and Bilic (Bilaan, Tboli, Tiruray) languages have been up for grabs to just about any subgrouping hypothesis, while other languages (e.g., Ilongot) have become less controversial. While I still cling to the theory that most of the linguistic groups of the Philippines [including some of Northern Celebes (Mongondow, Minahasan, Gorontalo?) and some Ph intrusives in Sabah (Illanun)]¹⁰ form a single AN subfamily, Reid's most recent subgrouping hypothesis¹¹ includes only Southern Mindoro, Central and Southern Palawan, and Central Philippines within the Malayo-Polynesian node. Northern and Southern Philippine languages merge at an Extra Formosan

¹⁰See Fleischman (1981). The inclusion of Mongondow and Gorontalo is based primarily on the work of Mathew Charles, and of Minahasan on Sneddon. Charles had long ago suggested to me that Kadazan may be a Ph language, but I defer to Blust's hypothesis that includes Sabahan languages in a larger North-Sarawak-Sabahan AN node. While I take issue with Blust's vowel-deletion hypothesis on reconstructional grounds, I do not refute the subgrouping implications of the split of PAN *b, *d, *D, *j into two correspondence sets among Bornean languages. This phenomenon appears to be an innovation of high quality, and hence of considerable power in subgrouping theory. My hypothesis suggests that this split arose as the result of consonant strengthening after (randomly or sporadically retained) short penult vowels [rather than vowel deletion and contact with PAN *S, as Blust proposed]. The fact that this innovation crosses typologically different languages ('Philippine' vs 'Indonesian') may be the object of some concern, but I am not in a position (data- or time-wise) to put forward an alternate subgrouping of the Sabahan languages.

¹¹Personal communication, cited in Reid's letter to Merritt Ruhlen (dated 27 Aug 82), although the subgrouping is dated 21 Nov 81. It is not my intention to be critical of this hypothesis (as I am not aware of the evidence that has motivated it); I merely wish to make the reader aware of alternative interpretations (of Ph subgrouping and, hence, the provenance of Ph accent).

node along with Amis and PMP. If Reid is correct, then CPh and NPh vowel and consonant quantity agreements must be posited as pre-PMP. If, on the other hand, my subgrouping is correct, then Ph accent *may* be an innovation (see Dyen 1971:45) – one of high quality in a subgrouping hypothesis. However, I do not make recourse to such an innovation, as will be seen hereunder, and has been suggested in Zorc (1972, 1977, 1978, 1979).

Secondly, loss or regularization of accent should not give us pause. The antiquity of the preserved penult accents in Malagasy or the recentness of oxytonality in Kuyonon and Inland Bikol dialects demonstrate only that accent is and was a fragile phoneme (akin to PAN *j or *R or *S), and is subject to the shifts or losses or resurrections noted for the less controversial members of the PAN phonemic inventory. Similarly, partial retention (or loss) of accent, while puzzling, must be accepted. Mansakan short vowels (Svelmoe and Abrams 1955, Abrams 1963) and Kalagan long vowels (Dawson 1958, Reid 1971) correspond with the quantity reconstructed for PCP (Zorc 1977). Yet neither Msk or Klq preserves more than a handful of the total PCP accent inventory.¹² The absence of a Msk **māta 'eye' or of a Klq **?a:wid 'grasp' does not invalidate the reconstruction of a PCP (and PPH) *matá 'eye' or *ha:wid 'hold',¹³ even if the loss of the corresponding accent pattern cannot presently be adequately explained.¹⁴ An analogous or parallel development is suggested hereunder for the clearly random retention of short vowels in Makassarese or Madurese (leading to consonant gemination) or North-Sarawak languages (leading to strengthened reflexes).

3. QUANTITY AND STRESS IN AUSTRONESIAN LANGUAGES

In my 1978 paper, I suggested guidelines for classifying languages typologically (and regardless of their genetic subgrouping) on the basis of the role of accent. The can be found as items A → E in Table 1. Further kinds of sub-classification were implicit in the paper, or have come to light since (items F → K in Table 1).

TABLE 1. CLASSIFICATION OF LANGUAGES BASED ON THE ROLE OF ACCENT

(A) Phonemic length and shortness, as *inherited* from PPH (< PHN/PMP/PAN): Bisayan (except Kuyonon and Tausug), Coastal and Pandan Bikol, Balangaw, Bontok, Hanunoo, Ifugao, Ilokano, Isneg, Itneg, Itawit, Kalinga, Kapampangan, Kankanay, Malaweg, Sambal, Tagalog

(B) Phonemic accent (quantity or stress), as *secondarily introduced*, generally due to consonant loss or borrowing:
Casiguran-Dumagat, Ibanag, Pangasinan, Old-Javanese, Ratahan, Malagasy

¹²All of the Mansakan short vowels I have data on are: ?ába 'chest', ?ábug 'dust', ?ágaw 'surely', ?álag 'also', ?arígi 'shrimp-fat', bǎ?ug 'rotten', bága? 'abscess', báyaw 'brother-in-law', béteŋ 'young-coconut', bílas [in-law], bítay [yam], bíyag 'full, satisfied', búkad 'to clean outside of house', búkuŋ 'bent, crooked', búta 'blind', dáyaw 'good', dúnut 'rotten', gǎbi 'night', gǎmut 'root', kúlan 'lie-down', lá?ug (no gloss, but in contrast to la?ug 'to string beads'), lánut 'rice-water', lápas (no gloss, in contrast with lǎpas (no gloss)), láwas 'tree-trunk', sǎrig 'to trust', tú?ud 'knee', túyu? 'to commit suicide'; there is in addition the prefix yamǎ- 'able to' (in contrast with yama- 'accidentally'). All of the Kalagan long vowels I have data on are: ?aba:ga 'shoulder', ?a:kuk 'vagina', ?a:lad 'fence', ?a:lun 'shadow', ?a:tag 'give', ba:ga? 'lungs', balaŋa:wan 'rainbow', ba:lu 'widow', da:gat 'sea', dala:gan 'run', da:lan 'trail', ?ina:si 'wine', ka:n 'eat', la:bun 'cloud', (ma)la:lem 'deep', la:nut 'abaca', la:was 'body', (ma)la:wig 'long', mala:lan 'thousand', ma:wat 'far', minta:lun 'defecate', ŋa:lan 'name', pa:lad 'palm (of hand)', pa:naw 'walk', pa:nda? 'short', pa:nid 'wing', pa:wa? 'swidden', sa:kil 'heel', sa:kit 'pain', ta:kaw 'steal'. Etymological justification for most of these can be found in Zorc (1977 and 1978).

¹³PPH *matá (Zorc 1978:#52), PPH *há:wid (Zorc 1977:footnote 13).

¹⁴It would appear that Mansakan has become a paroxytone language for the most part due to influence from its Manobo neighbors, while Kalagan has become oxytone due to influence from its Bilic neighbors. The vowel quantity thus far preserved is likely to die out in a generation or two.

- (C) Length contrasts in the ultima, resulting from compensation for the loss of a consonant:
Tausug, Butuanon, Kamayo, Cebuano
- (D) Phonemic length, as the result of coalescence (or crasis) of vowels, which does not correspond with stress (pitch accent):
Kuyonon, Tungho-Saisiat
- (E) Phonemic (1) length or (2) shortness, retained sporadically as remnants of a pre-existing system.
Mansaka (shortness), Kalagan (length)
- (F) Consonant length (1: gemination, or 2: strengthening) following a short vowel (generally, but not always *ə):
Bagobo, Ilokano, Kagayanen, Obo, Isneg, Itneg, Malaweg, Tagabawa, Madurese, Buginese, Sama-Bajaw
- (G) Oxytone = accent (with or without secondary vowel lengthening) falling regularly on the ultima:
Acehnese, Bilaan, Javanese, Ivatan, Kerinci, Palau, Puyuma, Tboli, Takituduh-Bunun, Saisiat, Uma-Juman, Yogad
- (H) Paroxytone = accent (with or without secondary vowel lengthening) falling regularly on the penult:
Gaddang, Makassarese, Paiwan, Pazeh-Kahabu, Ishbukun-Bunun, Maanyan
- (I) Proparoxytone = accent falling regularly on a prepenultimate syllable (or on the first syllable of a polysyllabic word):
Saaroa, Mantaوران-Rukai
- (J) PAN *ə influences accent in a different way from the other vowels:
Atayal, Malav, Sarangani-Manobo, Tiruray
- (K) Accent is used inflectionally = morphemic accent:
Chamorro, Kanakanabu, Motu, Toba-Batak, Angkola-Batak, most CPh languages

Such a classification may be more useful on given *words*, rather than on languages as a whole, although in some cases only one or two statements may apply and summarize concisely the role of accent in a language. Problematic cases include: Tagalog (A/B,D,F2, K),¹⁵ Ilokano (AF/B),¹⁶ Cebuano (A/C,F2,K),¹⁷ More clearcut cases would include: Malay (HJ), Chamorro (HK), Malagasy (B/H), Paiwan (H).

I would like to be able to make a comprehensive list of AN languages on this basis, and insure the accuracy of statements in Table 1. I would deeply appreciate your assistance in achieving this goal.

The accent system of Ph languages and the need to reconstruct long and short vowels for the ancestor of CPh and NPh languages ('PPH') has been taken up in my 1978 paper, and need not be reiterated here. Now I would like to discuss: the origins of secondarily lengthened vowels (§4) and consonants (§5), since they are at issue in the ultimate reconstruction of PAN quantity distinctions; secondary vowel shortness (§6);

¹⁵Tagalog basically retains original quantity distinctions, but particular forms exhibit interference from consonant loss (Tag ?a:raw < *qaljāw 'day'), crasis ?a:nim < *?a?ənəm 'six'), strengthened reflexes (?irí 'this' < *?i+di; expected **?ilí, cf: Tg ha-lí: ka 'come here!'), and morphemic accent (hi:hip < PSP *həyap 'to blow'; expected **hiyap, cf: hi:p-an 'blow on/at!')

¹⁶Ilokano has inherited length and shortness distinctions, and geminate consonants after short vowels; but some long high vowels are overridden if followed by an original laryngeal (Ilk rabi:í < PAN *Rabi:ʔiH 'night', Ilk luwa < PAN *lu:Səq 'tear').

¹⁷Cebuano has an accent situation similar to that described for Tagalog (footnote 15), but also has phonemically long final vowels in its *l-losing dialects (Ceb ba:y < PMP *baláy 'house').

