

# *Evidence for a Celebic supergroup*

---

DAVID MEAD

## 1 Introduction

Sulawesi and its offshore islands are home to more than one hundred Austronesian languages. As these languages have become better known, researchers have parcelled them into from four to eleven subgroups. Following the most current scholarship I recognise the following ten subgroups (see map for their distribution).<sup>1</sup> An estimate of the number of languages in each group follows in parentheses.

Sangiric (5)  
Minahasan (5)  
Gorontalo-Mongondow (9)  
Tomini-Tolitoli (11)  
Kaili-Pamona (16)  
Saluan-Banggai (5)  
Bungku-Tolaki (15)  
Muna-Buton (12)  
Wotu-Wolio (5)  
South Sulawesi (29)

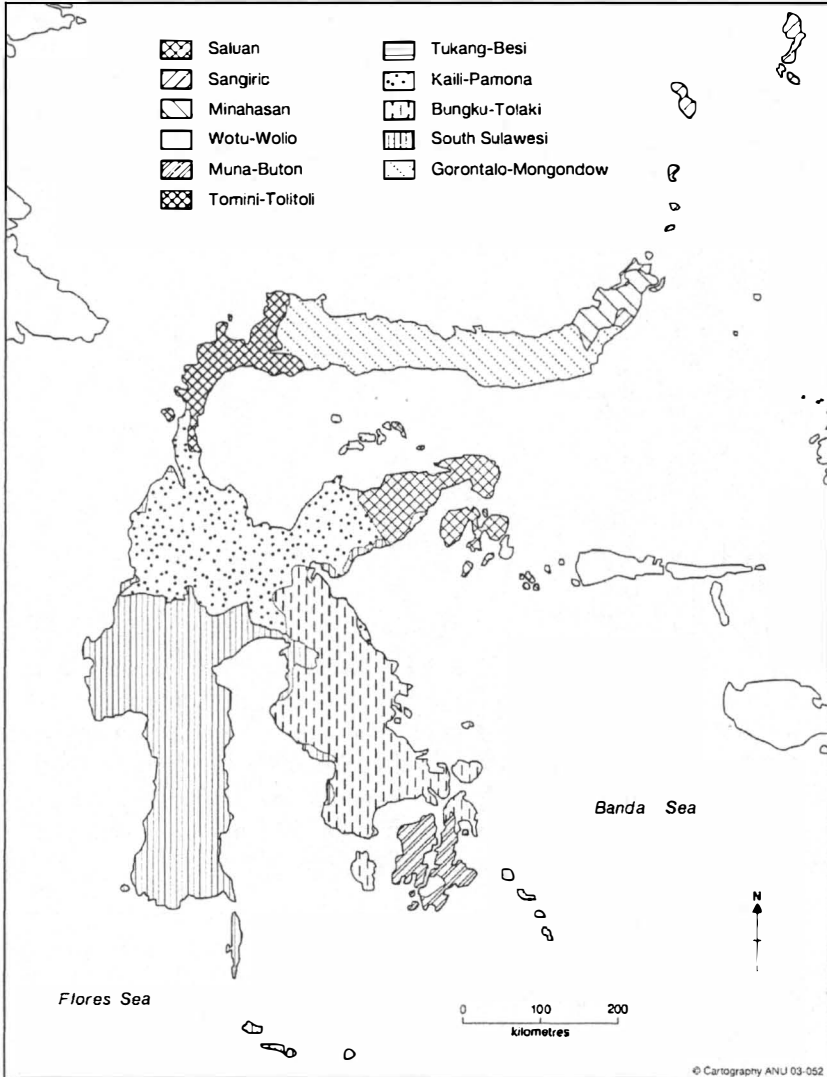
This paper is concerned with six of these subgroups, which collectively cut a broad swath across central and south-eastern Sulawesi — the Tomini-Tolitoli, Kaili-Pamona, Saluan-Banggai, Bungku-Tolaki, Muna-Buton and Wotu-Wolio subgroups. I propose that these groups are genetically related, composing what is here called the Celebic supergroup. My approach is bottom-up: I first show that the Bungku-Tolaki and Muna-Buton languages deserve to be united under a single node, the South-eastern Celebic macrogroup (§3). While it is clear that this macrogrouping includes the Tukang Besi languages, but the position of Tukang Besi within this

---

<sup>1</sup> Apart from the inclusion of Banggai with other Saluan languages (Mead this volume), and the splitting off of the Wotu-Wolio languages from Muna-Buton (Donohue in press), these are the same ten 'established microgroups' outlined in Sneddon (1993). For a summary of some other classification schemes of Sulawesi languages not discussed in this paper, see Mead (1999:179–180).

group requires further discussion (§4). In turn, the South-eastern Celebic languages group closely with the Saluan-Banggai languages under an Eastern Celebic node (§5). Finally, at the most inclusive level I adduce evidence for a Celebic supergroup comprising the newly proposed Eastern Celebic macrogroup along with the Tomini-Tolitoli, Kaili-Pamona and Wotu-Wolio subgroups (§6). The notion of a Celebic supergroup in turn refines our view of the Austronesian settlement of Sulawesi. I comment on this in the conclusions (§7), and list questions for further research.

Before proceeding to subgrouping arguments, I first review the historical and comparative literature concerning Sulawesi languages.



**Map 1: The Celebic supergroup**

## 2 Historical and comparative studies of Sulawesi languages

Rigorous historical and comparative studies of Sulawesi languages have been undertaken only in the past thirty years. Beginning with Mill's (1975) work on Proto South Sulawesi, reconstructions of protolanguages have also appeared for Minahasan (Sneddon 1978), Sangiric (Sneddon 1984), Gorontalo-Mongondow (Usup 1986), Kaili-Pamona (Martens 1989b) and Bungku-Tolaki (Mead 1998), while initial historical and comparative work has also been carried out in Muna-Buton (van den Berg 1991a, 1991b, and this volume), Saluan-Banggai (Mead this volume) and Wotu-Wolio (Donohue in press). All told, then, only one language group of Sulawesi has yet to come under the attention of historical and comparative linguists — the Tomini-Tolitoli languages spoken in northwestern Sulawesi.

These studies have brought to light a considerable amount of data on Sulawesi languages. They have refined our notions of subgroup boundaries and have done much to clarify our understanding of relationships within such lower-level groupings. Concurrent attempts to establish higher-level connections *between* subgroups, however, have largely been unsuccessful.

Starting from his work in South Sulawesi, Roger Mills (1975:517ff.) investigated but could find no reason for macrogrouping South Sulawesi languages with the Kaili-Pamona, Bungku-Tolaki or Muna-Buton groups, which he collectively referred to as 'Toraja' languages (the existence of a Wotu-Wolio group separate from Muna-Buton was unknown to Mills). Later Mills (1981) sketched out four major subgroups across the island of Sulawesi, namely the South Sulawesi languages, the 'Toraja' languages, the North Sulawesi languages, and finally the Saluan languages. His groupings, however, were mostly impressionistic. He speculated, for example, that the Saluan languages including Banggai were ultimately connected with Philippine languages, and might be relatively recent arrivals in Sulawesi. He left the Tomini-Tolitoli languages out of consideration owing to the small amount of material then available. Ülo Sirk (1981) reached the same conclusions as Mills concerning South Sulawesi languages. While Sirk identified several 'old' lexical items which South Sulawesi languages shared with their neighbours, particularly Kaili-Pamona languages, and structural similarities which they shared with languages of South-east Sulawesi, he concluded that such similarities merely pointed to a long period of contact.

Since those early days, other important contributions have been made to our understanding of macrogrouping of Sulawesi languages.

In 1989 James Sneddon published the results of his comparison of North Sulawesi languages from the perspective of both historical sound change and lexical innovations. In this careful study, he found no basis for grouping the Gorontalo-Mongondow, Minahasan or Sangiric languages with each other, nor did he find support for grouping any of these three with other language groups of Sulawesi. He concluded that 'the search for close affinities [of these three microgroups] must be directed northward, to the languages of the Philippines' (Sneddon 1989:103)

Work in that direction was already in progress. In 1991, Blust identified the Gorontalo-Mongondow languages (but not the Minahasan or Sangiric languages), as belonging to his newly proposed Greater Central Philippines macrogroup. In essence, the Gorontalo-Mongondow languages are relatively recent arrivals in Sulawesi and share a closer genetic affiliation with Tagalog and other Philippine languages than they do with any language group on Sulawesi.

Van den Berg (1996b) is notable as the only author to sketch out a possible basis for a Celebic macrogroup, corresponding to Mill's 'Toraja' group, while at the same time he left open the question whether this group might also include the Tomini-Tolitoli and Saluan languages with Banggai. Although the issue of subgrouping was tangential to the thrust of his paper, van den Berg was on track in regard to certain sound changes, namely the loss of consonant clusters, the monophthongisation of PMP final diphthongs, and the shift of PMP \*e (schwa) to a back rounded vowel. Final consonant loss, however — the change which van den Berg gave first in his list of shared phonological innovations — cannot be used for subgrouping. As I have argued elsewhere (Mead 1996), final consonant loss across central and south-eastern Sulawesi must be an areal feature, not a shared innovation, and this weakens van den Berg's overall argument. Below I refine his notion of a Celebic macrogroup, as well as adduce new evidence which allows the Tomini-Tolitoli and Saluan-Banggai languages to be brought into it.

Finally, in his forthcoming article 'The pretenders to Muna-Buton', Donohue demonstrates from historical sound change that five Muna-Buton languages belong in their own subgroup, which he labels the Wotu-Wolio group after the two most prominent of these five languages. Unlike the other language groups of Sulawesi which occupy geographically contiguous areas, the Wotu-Wolio languages are spoken in widely separated enclaves. Wolio and Kamaru are spoken on the island of Buton in South-east Sulawesi, Kalao and Laiyolo (including Barang-Barang) are spoken on and in the vicinity of Selayar Island off the southern coast of South Sulawesi, while the fifth, Wotu, is spoken at the northern tip of Bone Bay. At the same time, Donohue was unwilling to place the *Tukang Besi* languages either within his Wotu-Wolio group or with the remaining Muna-Buton languages, nor has he chosen to comment elsewhere on the classification of *Tukang Besi*.

### 3 South-eastern Celebic

On the heels of my work on the Bungku-Tolaki languages (Mead 1998), I suggested that these languages probably link closely with the Muna-Buton languages. Elsewhere I have referred to this grouping as 'South-eastern Celebic' (Mead 2001, 2002). I would now like to put this grouping on a firmer footing by citing the sound changes which are shared by all the indigenous languages of south-eastern Sulawesi, excluding Wolio and Kamaru.

This grouping includes *Tukang Besi*. The four principle *Tukang Besi* Islands stretch in a south-eastward direction off the southern coast of Buton Island, at the extreme tip of the south-eastern peninsula of Sulawesi. While there is some dialect chaining, following Donohue (2000) it is possible to recognise two *Tukang Besi* languages, *Tukang Besi North* spoken on the islands of Wanci and Kaledupa, and *Tukang Besi South*, spoken on the islands of Tomea and Binongko (and including Bonerate, spoken on islands to the south of Selayar Island in South Sulawesi).

The following twelve changes have been discussed at length by Mead (1998) for Bungku-Tolaki languages and by van den Berg (1991a, b) for Muna and its dialects. Van den Berg (this volume) also mentions most of these changes in regard to *Tukang Besi*. Where known to me, I include data from Kioko by way of exemplifying another Muncic language, and from Cia-Cia for a Butonic language. Bungku-Tolaki data is from Mead (1998). Proto Muna forms are from van den Berg (1991a), otherwise Muna data is from van den Berg (1996a). Kioko data is from my own field notes. Cia-Cia data is from van den Berg (1991c and pers. comm.). *Tukang Besi* data

is from Donohue (1999, 2000). (PBT = Proto Bungku-Tolaki, PM = Proto Muna; TB = all Tukang Besi isolects, otherwise Tukang Besi isolects are referred to by island name. Leftmost forms are Proto Malayo-Polynesian (PMP) unless otherwise noted.)

1. Consonant cluster reduction  $*C_1C_2 > *C_2$ , provided the initial consonant  $C_1$  was not a nasal

---

<i>*sepsep</i> ‘suck’	> PBT <i>*sosoQ</i> , Muna <i>soso</i> , Cia-Cia <i>sosopi</i>
<i>*dutdut</i> ‘pluck’	> PBT <i>*ruruQ</i> , Muna <i>ruru</i>
<i>*tuktuk</i> ‘knock’	> PBT <i>*tutuk-i</i> , Wanci <i>tutu</i> ‘pound, smith’
<i>*qali/kali-petpet</i> ‘firefly’	> PBT <i>*olimpopoQ</i> , Muna, Kioko, Wanci <i>kalipopo</i>

---

In nasal clusters, the nasal assimilated to the point of articulation of the following consonant, compare PMP *\*demdem* ‘dark’ > PBT *\*rondoma*, Muna *rondo*, Wanci *morondo* ‘night’.

2. Loss of PMP  $*h > \emptyset$

---

<i>*hapuy</i> ‘fire’	> PBT <i>*apuy</i> , Muna <i>ifi</i> , Cia-Cia <i>api</i> , Wanci, Kaledupa <i>ahu</i>
<i>*hikan</i> ‘fish’	> PBT <i>*ikaN</i> , Wanci <i>ika</i>
<i>*kahiw</i> ‘wood’	> PBT <i>*kayu</i> , Muna, Cia-Cia <i>sau</i> , TB <i>kau</i> (but Wanci <i>kau</i> ~ <i>ka<sup>ə</sup>u</i> )
<i>*buhuk</i> ‘head hair’	> PBT <i>*wuuQ</i> , Muna, Kioko <i>wuu</i>
<i>*dahun</i> ‘leaf’	> PM <i>*roo</i> , Cia-Cia, TB <i>ro<sup>ə</sup>o</i>

---

Addition of glottal stop between like vowels is a regular feature in Tukang Besi. The Wanci form *ka<sup>ə</sup>u* ‘wood’ is problematic, but is insufficient to maintain that Tukang Besi has retained a non-zero reflex of PMP  $*h$  (see further van den Berg in this volume).

3. Rhotacisation of PMP  $*d > *r$

---

<i>*depa</i> ‘fathom’	> PBT <i>*ropa</i> , Muna <i>rofa</i>
<i>*duRi</i> ‘thorn’	> PBT <i>*rui</i> , PM <i>*ka-rui</i> , Kioko <i>xii</i> , Cia-Cia <i>rui</i> , Wanci, Kaledupa, Binongko <i>ruhi</i> , Tomea <i>rihi</i> , Bonerate <i>rihu</i>
<i>*daRaq</i> ‘blood’	> PBT <i>*raRaq</i> , PM <i>*rea</i> , Kioko <i>xea</i> , Cia-Cia <i>rea</i> , TB <i>raha</i>
<i>*daqan</i> ‘branch’	> Tolaki <i>ra/a</i> , Muna <i>ragha</i> , Kioko <i>kaxa<sup>ə</sup>a</i> , Cia-Cia <i>raha</i>
<i>*dalem</i> ‘inside’	> PBT <i>*laroN</i> (<met.), PM <i>*lalo</i> , Kioko <i>lalo</i> , Wanci, Kaledupa, Bonerate <i>laro</i> (<met.), Tomea, Binongko <i>lalo</i>
PWMP <i>*kiday</i> ‘eyebrow’	> PBT <i>*kire</i> , Muna <i>kire</i> , Kioko <i>kixe</i> , Wanci <i>kire</i>
<i>*qadep</i> ‘front’	> PBT <i>*aroQ</i> , Wanci <i>aropa</i> , Kaledupa, Binongko <i>aro</i>
<i>*tuduR</i> ‘sleep’	> PBT <i>*turuR</i> , Muna <i>tuturu</i> , TB <i>moturu</i>
<i>*qudanj</i> ‘shrimp’	> PBT <i>*uraN</i> , Muna <i>ghura</i>

---

Kioko /x/ is a voiceless velar fricative (phonetically in free variation with [h]). Van den Berg (this volume) postulates a split of PMP  $*d > *d, *r$ , apparently on the basis of a few (clearly minority) cases where  $*d$  remained /d/, for example PMP *\*duha* > PBT *\*rua*, Cia-Cia *rua*, but Muna *dua*, *rua-*, Kioko *xudua*, Wanci *dodua*; also PMP *\*deŋeR* > PBT *\*rongoR*, but Wanci, Kaledupa *rodongo*, Tomea, Binongko, Bonerate *dongo*.

## 4. Monophthongisation of PMP final diphthongs \*-ay and \*-ey &gt; \*e

*qaZay	'chin'	> PBT *ase, Muna <i>ghase</i> , Kioko <i>ase</i> , Cia-Cia <i>hae</i>
*qatey	'liver'	> PBT *ate, PM *qate, Kioko <i>ate</i> , Cia-Cia <i>hate</i>
*quey	'rattan'	> PBT *ue, PM *que, Kioko <i>ue</i> , Wanci, Kaledupa ?ue, other TB <i>ue</i>
*m-atey	'die, dead'	> PBT, PM *mate, Kioko, TB <i>mate</i>

## 5. Monophthongisation of PMP final diphthongs \*-aw and \*-ew &gt; \*o

*qalejaw	'day, sun'	> PBT *oleo, PM *qoleo, Kioko <i>oleo</i> , Cia-Cia <i>holeo</i> , TB ?oloo
*babaw	'over'	> PBT *wawo, Muna, Kioko <i>wawo</i> , TB <i>wawo</i>
*kasaw	'rafter'	> PBT *kaho, Muna <i>saho</i> , Kioko <i>sa?o</i> , Wanci <i>kaso</i> 'ridge pole'
*behew	'odour'	> PBT *woo

## 6. Backing of PMP \*e (schwa) &gt; \*o

*telu	'three'	> PBT, PM *tolu, Kioko, Cia-Cia <i>tolu</i>
*qitem	'black'	> PBT *itoN, Muna <i>ghito</i> ; Kioko <i>ito</i> , Cia-Cia <i>kito</i>
*qatep	'roof'	> PBT *atoQ, PM *qato, Kioko <i>ato</i> , Cia-Cia <i>hato</i>

## 7. Lowering of PMP \*i &gt; \*e preceding final \*q

*putiq	'white'	> PBT *pute, Muna, Kioko, Cia-Cia <i>pute</i> , TB <i>mopute</i>
*binehiq	'seed rice'	> PBT *bine, Muna, Kioko, Cia-Cia, Wanci <i>wine</i>
*uliq	'return'	> Tolaki <i>pule</i> 'return home' TB <i>pule</i> 'repeat'
*piliq	'choose'	> PBT *pile (but Muna <i>pili</i> , from Wolio or Malay?)

The lowering of PMP \*i > \*e preceding final \*q is not recognised by van den Berg (1991a, 1991b), and is presented here for the first time as a regular sound change characterising Muna-Buton and Tukang Besi languages.

## 8. Raising of PMP pretonic \*a &gt; \*o

*qalejaw	'day, sun'	> PBT *oleo, PM *qoleo, Kioko <i>oleo</i> , Cia-Cia <i>holeo</i> , TB ?oloo
*qalipan	'centipede'	> PBT *olipaN, Wanci <i>oliha</i>
*paniki	'bat'	> PBT *poniki, Muna, Kioko <i>ponisi</i> , TB <i>honiki</i>
*baqeRu	'new'	> Tolaki <i>wo?ohu</i> , PM *buqou, Kioko <i>wu?ou</i> , Cia-Cia <i>wukou</i> , TB <i>wo?ou</i>

## 9. Loss of PMP medial \*-w- &gt; ∅

*sawa	'snake, python'	> PBT *saa, Muna <i>saa</i> , Cia-Cia, TB <i>sa?a</i>
*hawak	'waist'	> PBT *aaQ, Muna <i>aa</i>
*tawa	'laugh'	> PBT *taa, Muna <i>futaa</i> , Kioko <i>fotaa</i>
*kawit	'hook'	> PBT *kaiQ, Muna, Wanci <i>kai</i>

## 10. Split of PMP \*s &gt; \*s, \*h

* <i>lasuq</i>	'penis'	> PBT * <i>lahuq</i> , Cia-Cia, Wanci <i>lau</i>
* <i>salaq</i>	'mistake'	> PBT * <i>halaq</i> , Muna <i>hala</i> , Cia-Cia, TB <i>sala</i>
* <i>sabuR</i>	'scatter'	> PBT * <i>hawuR</i> , Muna <i>hewi</i>
* <i>sa-puluq</i>	'ten'	> PBT * <i>hopuluq</i> , Muna, Kioko, Cia-Cia, TB <i>ompulu</i>
* <i>tasik</i>	'sea'	> PBT * <i>tahiQ</i> , Muna <i>tehi</i> , Kioko <i>te?i</i> , Cia-Cia, Tomea, Binongko, Bonerate <i>tai</i>
* <i>tasak</i>	'ripe'	> PBT * <i>tahaq</i> , Muna, Cia-Cia <i>taha</i> , Wanci <i>mota?a</i>
* <i>qasiRa</i>	'salt'	> PBT * <i>ohia</i> , Muna <i>ghohia</i>
* <i>pusej</i>	'navel'	> PBT * <i>puhoy</i> , Muna <i>puhe</i> , Cia-Cia <i>puse</i>
* <i>isi</i>	'contents'	> PBT * <i>ihi</i> , Muna <i>ihi</i> , Cia-Cia <i>isi</i>
* <i>siku</i>	'elbow'	> PBT * <i>hiku</i> , Muna, Cia-Cia, Kioko, Wanci <i>siku</i>
* <i>siwa</i>	'nine'	> PBT * <i>sio</i> , Muna, Kioko, Cia-Cia <i>siua</i> , TB <i>sia</i>
* <i>pisaw</i>	'knife'	> PBT * <i>piso</i> , Muna, Kioko <i>piso</i>
* <i>beRsay</i>	'paddle'	> PBT * <i>bose</i> , Muna, Kioko, Wanci <i>bose</i>
* <i>esa</i>	'one'	> PBT * <i>asa</i> , Muna, Cia-Cia <i>ise</i> , Kioko <i>seise</i> , Wanci <i>sa?asa</i> , Bonerate <i>asa</i> , other TB <i>assa</i>

The split of PMP \*s into both \*s and \*h is one of the most significant changes for subgrouping South-eastern Celebic languages together, yet it is not unproblematic. A notable feature about this change is that there is no ready explanation for the conditioning environment which caused the split, either in terms of word stress, position within the word, or surrounding phonemes. The two major patterns are either that PMP \*s became \*h in Proto Bungku-Tolaki, /h/ in Muna, and zero in Cia-Cia and Tukang Besi, or else it remained /s/ in all four. Nonetheless, sometimes a mixed pattern is observed. Table 1 is a rearranged presentation of some of the above data, particularly where Cia-Cia and/or Tukang Besi reflexes are known. Forms that exhibit weakening of PMP \*s are shown in bold.

Table 1: Weakening of PMP \*s

	PBT	Muna	Tukang Besi	Cia-Cia
* <i>lasuq</i> 'penis'	* <i>lahuq</i>	—	<i>lau</i>	<i>lau</i>
* <i>sa-puluq</i> 'ten'	* <i>ho-puluq</i>	<i>ompulu</i>	<i>ompulu</i>	<i>ompulu</i>
* <i>tasik</i> 'sea'	* <i>tahiQ</i>	<i>tehi</i>	<i>tai</i>	<i>tai</i>
* <i>tasak</i> 'ripe'	* <i>tahaQ</i>	<i>taha</i>	<i>ta/a</i>	<i>taha</i>
* <i>pusej</i> 'navel'	* <i>puhoy</i>	<i>puhe</i>	—	<i>puse</i>
* <i>isi</i> 'contents'	* <i>ihi</i>	<i>ihi</i>	—	<i>isi</i>
* <i>salaq</i> 'mistake'	* <i>halaq</i>	<i>hala</i>	<i>sala</i>	<i>sala</i>
* <i>kasaw</i> 'rafter'	* <i>kaho</i>	<i>saho</i>	<i>kaso</i>	<i>kaso</i>
* <i>siku</i> 'elbow'	* <i>hiku</i>	<i>siku</i>	<i>siku</i>	<i>siku</i>
* <i>esa</i> 'one'	* <i>asa</i>	<i>ise</i>	<i>asa</i>	<i>ise</i>
* <i>siua</i> 'nine'	* <i>sio</i>	<i>siua</i>	<i>sosia</i>	<i>siua</i>
* <i>sawa</i> 'snake'	* <i>saa</i>	<i>saa</i>	<i>sa?a</i>	<i>sa?a</i>
* <i>beRsay</i> 'paddle'	* <i>bose</i>	<i>bose</i>	<i>bose</i>	—

One interpretation of this data is that while the weakening of \*s may have begun in Proto South-eastern Celebic, it continued to diffuse lexically and areally, reaching its fullest extent in Bungku-Tolaki languages, and its least extent in Tukang Besi and Cia-Cia. This account, however, runs into a conundrum. The weakening of \*s, to the extent that it did occur, seems to have largely been completed by Proto Bungku-Tolaki; daughter languages simply fail to exhibit evidence of any lexical or areal diffusion post-Proto Bungku-Tolaki.<sup>2</sup> On the other hand, if diffusion of this change is a recent phenomenon in the Muna-Buton area, then there is no way to account for the high degree of concordance between Muna and Bungku-Tolaki languages. It is also possible, however, that forms such as Cia-Cia *puse* and *isi*, and Cia-Cia and Tukang Besi *sala*, *kaso* and *siku* are later borrowings which have obscured an earlier, more regular state of affairs. In this case a larger proportion of *s*-forms would indicate greater influence from an outside language (for example Wolio) in which PMP \*s never weakened. Until such time as more lexical material becomes available, both from Tukang Besi and other Muna-Buton languages, it may be premature to decide between the two explanations.

### 11. Depalatalisation of PMP \*Z > \*s

---

*Zalan 'road'	> PBT *salan, Muna, Wanci, Kaledupa <i>sala</i>
*qaZay 'chin'	> PBT *ase, Muna <i>ghase</i> , Kioko, Wanci <i>ase</i> (but Cia-Cia <i>hae</i> )
*quZan 'rain'	> PBT *usan, Muna <i>ghuse</i> , Kioko <i>ise</i> , Wanci <i>uselau</i> 'k.o. storm' (but Cia-Cia <i>kia</i> )

---

The Cia-Cia forms, which exhibit PMP \*Z becoming zero, are problematic unless it can be shown that they went through an /s/ stage before being lost. Otherwise, it may be necessary to reconstruct \*Z for Proto Muna-Buton (a difficulty which van den Berg does not address), and hence also for Proto South-eastern Celebic.

### 12. Depalatalisation of PMP \*ñ > n

---

*peñu 'turtle'	> PBT *ponu, Muna, Kioko <i>ponu</i>
*wañi 'bee'	> PBT *hoani, Muna <i>ani</i>
*nia, ña '3S possessive'	> PBT *-no, Muna, TB <i>-no</i>

---

The depalatalisation of PMP \*ñ is best viewed as dependent upon (or going along with) the depalatalisation of \*Z. A parallel case is found in Saluan-Banggai languages, where \*ñ was depalatalised to /n/ only in those languages where \*Z was also depalatalised to /d/ or /s/ (Mead this volume). Cross-linguistically, no language is known to have more nasal stops than oral stops (Ferguson 1963).

A thirteenth and fourteenth sound change could be marshalled in support of a South-eastern Celebic group. The change of PMP \*j > \*y,  $\emptyset$  is exhibited by all South-eastern Celebic languages, as well as other Sulawesi languages. This change is discussed at some length in §6.

---

<sup>2</sup> The only forms where present-day Bungku-Tolaki languages differ is where the PMP form contained two occurrences of \*s. In this case phonotactic constraints may have played a further role. Compare PMP \*susu 'breast' > Moronene, Wawonii, Bungku *susu*, but Tolaki *uhu*, Mori and Padoe *uo*.

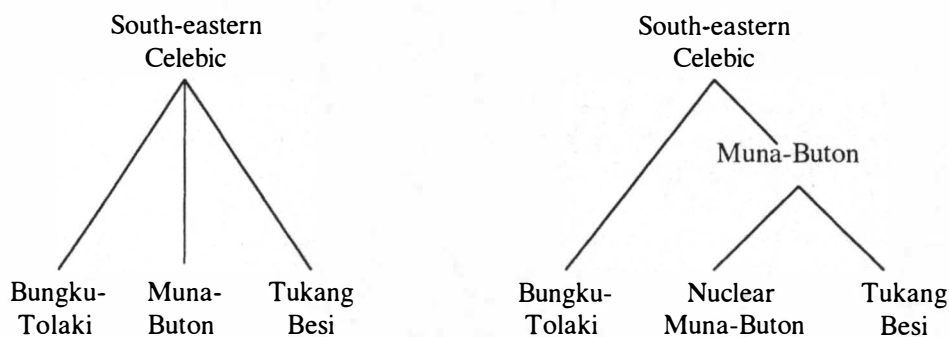


A split of PMP *\*b* into /b/ and /w/ is also exhibited by all South-eastern Celebic languages. Even more than the weakening of *\*s*; however, the weakening of *\*b* has clearly continued to diffuse into many present-day languages, and it is unclear to what extent this change should be attributed to their common ancestor. See further Mead (1998:35–40), van den Berg (1991b:12 and this volume) and Donohue (in press).

#### 4 Relationships of South-eastern Celebic languages to each other

In his important paper clarifying the boundary between Wotu-Wolio and Muna-Buton languages, Donohue (in press) declined to affiliate the Tukang Besi languages with either group, effectively leaving Tukang Besi ‘orphaned’ in any classification scheme for south-eastern Sulawesi languages. Van den Berg considered this to be in error. In his paper (this volume), he dismisses two supposed objections to including Tukang Besi with the other Muna-Buton languages, and adduces phonological changes as well as grammatical and lexical evidence for bringing the Tukang Besi languages back into the fold, so to speak, with the other Muna-Buton languages.

While one can argue, as I have above, that the Tukang Besi languages are South-eastern Celebic languages, the question remains whether they share a further, closer relationship with Muna-Buton languages. Compare the two diagrams of Figure 1.



**Figure 1:** Two views of the classification of south-eastern Celebic languages

In that van den Berg does not intend his Muna-Buton node to also include the Bungku-Tolaki languages, he must support the configuration given on the right. However, in arguing for the position he gives to the Tukang Besi languages, van den Berg generally does *not consider their position also with respect to Bungku-Tolaki languages*. I devote the remainder of this section to a consideration of his claims in this light.

In regard to sound changes which could prove a close relationship, one could hazard that final consonants were lost in the ancestor to (nuclear) Muna-Buton and Tukang Besi. This would necessarily exclude the Bungku-Tolaki languages, which all lost final consonants in a drift-like tendency (reflexes of almost all PMP final consonants are reconstructible for *their* common ancestor; see Mead 1996). Apart from this, there is only one other sound change to my knowledge which supports van den Berg’s grouping, namely the fate of PMP initial *\*w*, which

became zero in Tukang-Besi and the remaining (nuclear) Muna-Buton languages, but became /h/ in Bungku-Tolaki languages (recall from the preceding section that PMP \*-w- became zero in all South-eastern Celebic languages). As can be seen from Table 2, to date evidence that Tukang Besi participated in this change is limited, as far as I know, to two forms of the number 'eight'.

**Table 2:** Fate of PMP initial \*w

		PBT	Muna	Tukang Besi
*wa-walu	'eight' (free)	*hoalu	oalu	oalu
*walu	'eight' (bound)	*halu	alu	alu
*wakat	'root'	*haka	aka	—
*(wa)wañi	'honeybee'	*hoani	ani	—

The picture from sound change is not conclusive. Against this weakly attested change (PMP \*w- > Ø), there are two well-attested changes which could link nuclear Muna-Buton languages instead with Eastern Bungku-Tolaki (!) languages. These changes are the merger of PMP \*R with \*y, and the subsequent loss of \*y (from all sources), usually with fronting of the preceding vowel. See van den Berg (1991b:14–15, this volume) and Mead (1998:114) for details of these changes.

Turning to grammatical changes, van den Berg suggests that Proto Muna-Buton had six pronoun sets. Cognates for five of these pronoun sets are also attributable to Proto Bungku-Tolaki. Only irrealis subject pronouns are not reconstructible for Proto Bungku-Tolaki, though they did develop later as an areal feature in at least four Bungku-Tolaki languages (termed 'future pronouns' in Mead 1998). It would be of greater validity to show that Muna-Buton and Tukang Besi exclusively shared innovations in *reconstructed pronominal forms* — but no Proto Muna-Buton pronoun sets have yet been reconstructed.

Likewise the use of *-um-* to form subject relative clauses is characteristic of Bungku-Tolaki languages (Mead 1998). A requestive prefix *pepe-* with parallel semantics and morphosyntax is also found in Mori Bawah (Esser 1933:326), while in the Kaili-Pamona language Uma, this prefix has the form *pope-* (Martens 1988:184). Only the prefixes \*pa- 'occupation' and \*para- 'iterative' appear not to have cognates in Proto Bungku-Tolaki.<sup>3</sup>

A deictic opposition between *t*-forms meaning 'upwards' and *w*-forms meaning 'downwards' is also found in Tolaki; compare *ikita* 'up there' with *ikua* 'down there' (from earlier \**i-ki-wa*). While the opposition has been lost in present-day Padoe, Esser (1927) recorded (among other forms) Padoe deictic adverbs *tehea* 'up there' and *wehea* 'down there' and deictic determiners *ta?a* 'that upwards' and *wa?a* 'that downwards'. Clearly the \**t* versus \**w* deictic distinction must be a retention from Proto South-eastern Celebic, not an innovation which exclusively links Muna with Tukang Besi.

Therefore while the grammatical evidence does not contradict grouping the Muna-Buton languages with Tukang Besi vis-a-vis the Bungku-Tolaki languages, it cannot be said to argue for such a grouping either. This leaves van den Berg's (rather impressive) list of possible lexical

<sup>3</sup> Esser (1933:301) lists four Mori Bawah nouns formed with the prefix *pa-*, which he considered to be borrowings from Buginese. Van den Berg's Proto Muna-Buton prefix \**para-* may have a cognate in the Uma diffuse prefix *mpara-* (Martens 1988:197).

innovations (including irregular phonological developments in specific lexical items) as the primary basis upon which to posit a close link between the Tukang Besi and other Muna-Buton languages. Like any other initial offering of supposed lexical innovations, van den Berg's list will undergo a process of scholarly refinement, as more scholars become involved. Even though lexical innovations by their nature constitute a weaker kind of evidence, I provisionally accept van den Berg's classification of the Tukang Besi languages. At this point a great deal depends on finding out more about the languages of central and south-eastern Buton, where unfortunately our best data is often still inadequate.

## 5 Eastern Celebic

If we now take the sound changes which South-eastern Celebic languages share in common, and compare them with surrounding language groups, it becomes apparent that South-eastern Celebic languages are most closely related to languages directly to the north. No less than eight of the fourteen sound changes outlined in §3 are also shared by the five Saluan-Banggai languages of eastern Sulawesi. The following data have been excerpted from Mead (this volume).

### 1. Consonant cluster reduction $*C_1C_2 > *C_2$ , where $C_1$ was not a nasal

---

<i>*tuktuk</i>	'forge'	> Banggai, Balantak <i>tutuk</i>
<i>*gisgis</i>	'rub'	> Balantak, Andio, Saluan, Bobongko <i>geges</i>
<i>*sepsep</i>	'suck'	> Banggai, Balantak, Saluan, Bobongko <i>sosop</i>

---

### 2. Loss of PMP $*h > \emptyset$

---

<i>*hapuy</i>	'fire'	> Balantak, Andio, Bobongko <i>apu</i> , Saluan <i>apu ~ apuu</i>
<i>*hasanj</i>	'fish gills'	> Balantak, Andio, Saluan, Bobongko <i>ansang</i>
<i>*buhuk</i>	'head hair'	> Banggai <i>buuk</i> , Balantak <i>wuuk</i>

---

### 3. Rhotacisation of PMP $*d > *r$

---

<i>*duha</i>	'two'	> Banggai <i>lua</i> , Balantak, Andio <i>rua</i> , Saluan <i>ohua?</i>
<i>*duRi</i>	'thorn'	> Balantak <i>ruri?</i> , Andio <i>rii?</i> , Saluan <i>hii?</i>
<i>*dahun</i>	'leaf'	> Banggai <i>loon</i> , Balantak, Andio <i>roon</i> , Saluan <i>hoon</i> , Bobongko <i>ron</i>
<i>*daRaq</i>	'blood'	> Balantak <i>rara?</i> , Andio <i>raa?</i>
<i>*qudaŋ</i>	'shrimp'	> Balantak <i>urang</i> , Saluan <i>uhang</i>
<i>*qadep</i>	'front'	> Balantak <i>arop</i> , Andio <i>aropon</i> , Saluan <i>ahop</i>
<i>*pawed</i>	'sew thatch'	> Banggai <i>paul</i> , Balantak <i>paur</i>
<i>*tuhud</i>	'knee'	> Banggai <i>tuul</i> , Balantak <i>tuur</i> , Andio <i>utur</i> (<met.)

---

Note that  $*r$  subsequently became  $/l/$  in Banggai and  $/h/$  in Saluan.

## 4. Monophthongisation of PMP final diphthongs \*-ay and \*-ey &gt; \*e

---

*qaZay	'chin'	> Banggai <i>ade</i> , Balantak <i>asi</i> , Andio <i>ade</i> , Saluan, Bobongko <i>aje</i>
*m-atey	'die, dead'	> Banggai, Andio, Saluan, Bobongko <i>mate</i>
*qatey	'liver'	> Banggai, Balantak, Andio, Saluan, Bobongko <i>ate</i>

---

## 5. Monophthongisation of PMP final diphthongs \*-aw and \*-ew &gt; \*o

---

*linaw	'clear (water)'	> Banggai, Balantak, Bobongko <i>molino</i>
*babaw	'above'	> Banggai, Andio <i>babo</i> , Balantak <i>wawo</i> , Saluan <i>bawo</i> , Bobongko <i>bafo</i>
*kasaw	'rafter'	> Banggai, Balantak <i>kaso</i> , Andio, Saluan, Bobongko <i>kaso?</i>
*behew	'odour, stink'	> Banggai <i>boo</i> , Balantak <i>woo</i>

---

## 6. Backing of PMP \*e (schwa) &gt; \*o

---

*qatep	'roof'	> Balantak, Andio, Saluan, Bobongko <i>atop</i>
*utek	'brain'	> Balantak, Andio, Saluan, Bobongko <i>utok</i>
PWMP *beŋel	'deaf'	> Banggai, Balantak, Andio, Saluan, Bobongko <i>bongol</i>
*relu	'three'	> Banggai <i>tolu</i> , Balantak, Andio <i>tolu?</i> , Saluan <i>totolu?</i> , Bobongko <i>totolu</i>

---

## 7. Lowering of PMP \*i &gt; \*e preceding final \*q

---

*putiq	'white'	> Banggai <i>moute</i> , Saluan, Bobongko <i>mopute?</i>
*piliq	'choose'	> Banggai <i>ilei</i> , Andio, Saluan, Bobongko <i>pile?i</i>
*binehiq	'seed rice'	> Balantak <i>wine?</i> , Andio, Saluan, Bobongko <i>bine?</i> 'seedling'

---

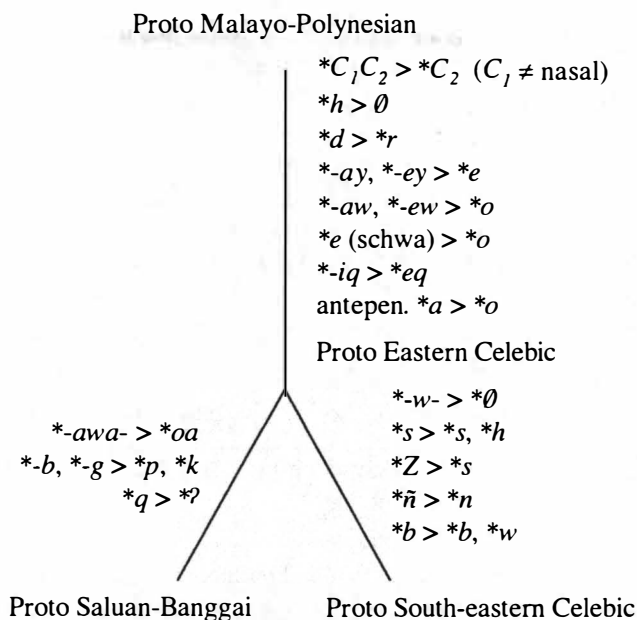
## 8. Raising of PMP pretonic \*a &gt; \*o

---

*laqia	'ginger'	> Balantak, Andio, Saluan <i>loiya?</i>
*paniki	'bat'	> Banggai <i>uniki</i> , Balantak, Saluan <i>poniki?</i>
*qasawa	'spouse'	> Banggai <i>osoaan</i> 'to marry', Saluan, Bobongko <i>osoa</i>

---

Taken together, these eight changes suggest a close relationship between Saluan-Banggai and South-eastern Celebic languages. I unite them in an Eastern Celebic macrogroup as shown in Figure 2.



**Figure 2:** The eastern Celebic macrogroup

The Saluan-Banggai languages are distinguished from Eastern Celebic by three further changes, also detailed in Mead (this volume). Two of these changes are problematic for subgrouping. The change of PMP *\*q* to glottal stop is widespread in Austronesia. It is found, incidentally, in all Bungku-Tolaki languages, in Tukang Besi, and in many Muna-Buton languages.<sup>4</sup> The value of this change for subgrouping is close to nil. Second, devoicing of final consonants could perhaps even be attributed to Proto Eastern Celebic, but since Bungku-Tolaki and Muna-Buton languages lost final consonants, at this point we simply lack the evidence for attributing this change to a higher level.

Despite the meagre remaining evidence, namely the single change PMP *\*-awa-* > *\*oa*, I maintain Saluan-Banggai as a separate subgroup. This change distinguishes the Saluan-Banggai languages from South-eastern Celebic languages where *\*w* was lost word-medially in Bungku-Tolaki, Muna-Buton and the Tukang Besi languages.

South-eastern Celebic languages in turn are distinguished from Eastern Celebic by possibly five further changes, though with caveats discussed in §4.

## 6 Evidence for a Celebic Supergroup

If we are correct in postulating an Eastern Celebic macrogroup, then it is clear that a 'Toraja' group as conjectured by Mills (1975, 1981) or a Celebic group as envisioned by van den Berg (1996b) encompassing the Kaili-Pamona, Bungku-Tolaki and Muna-Buton subgroups can not be supported. The Bungku-Tolaki and Muna-Buton languages (including Tukang Besi) are more

<sup>4</sup> Notably PMP *\*q* remained a uvular fricative in Muna, while in Cia-Cia, Kumbewaha and Kaimbulawa it became /k/ (usually preceding high vowels) or /h/ (usually preceding mid to low vowels).

closely related to the Saluan-Banggai languages on the eastern peninsula of Sulawesi than they are to the Kaili-Pamona languages in the heart of central Sulawesi. Is it possible, though, that the Kaili-Pamona languages could share some relationship, albeit a more distant one, with Eastern Celebic languages?

Of the eight changes listed in §5 which constitute the basis for an Eastern Celebic macrogroup, five are shared by Kaili-Pamona, as well as the Tomini-Tolitoli and Wotu-Wolio language groups. Proto Kaili-Pamona (PKP) reconstructions are from Martens (1989b). Pamona data are from Adriani (1928). Wotu-Wolio data are from Donohue (in press), Laidig and Maingak (1999), Anceaux (1987), and an unpublished Wotu word list collected by Wyn Laidig. Tomini-Tolitoli (TT) data is from Himmelmann (2001). Capital *E* (as in Dondo and Tialo *iʔanE* 'fish') represents a paragogic vowel.

1. Consonant cluster reduction  $*C_1C_2 > *C_2$ , where  $C_1$  was not a nasal

---

* <i>tuktuk</i> 'forge'	> Pamona, Wolio <i>tutu</i> 'pound, crush' Tialo, Dondo, Totoli, Boano <i>tutu</i> 'pound'
* <i>gisgis</i> 'rub'	> Pamona <i>gegesi</i> , Wolio <i>gigisi</i> , Taje <i>gegesi</i> , other TT <i>geges</i>
* <i>sepsep</i> 'suck'	> Wolio <i>sosopi</i> , Dampelas <i>monosop</i> , most other TT <i>sosop</i>
* <i>qali/kali-petpet</i> 'firefly'	> PKP <i>*kalipopoʔ</i> , Dampelas <i>kalipopo</i> , Taje, Boano <i>alipopo</i> , Tajio <i>alipopot</i>
* <i>kitkit</i> 'bite'	> Pamona <i>kiki</i> , Kalao <i>kekiti</i> (but Laiyolo <i>kikkiʔ</i> from South Sulawesi influence?)

---

2. Loss of PMP  $*h > \emptyset$

---

* <i>hapuy</i> 'fire'	> PKP <i>*apu</i> , Wotu, Laiyolo, Kalao, Kamaru <i>apu</i> , Tialo, Dondo <i>apiy</i> , other TT <i>api</i>
* <i>hikan</i> 'fish'	> Lauje <i>iʔang</i> , Tialo, Dondo <i>iʔanE</i> , Boano <i>ikan</i>
* <i>buhuk</i> 'head hair'	> Wolio <i>buu</i> 'nape of neck', Totoli, Boano <i>buok</i>
* <i>kahiw</i> 'wood'	> PKP <i>*kayu</i> , Laiyolo, Kalao <i>kaju</i> , Dampelas, Totoli, Boano <i>kayu</i> , Tajio <i>ayu</i> , other TT <i>ʔayu</i>

---

3. Rhotacisation of PMP  $*d > *r$

---

* <i>duha</i> 'two'	> PKP <i>*ro-</i> , Laiyolo, Kalao, Wolio <i>rua</i> , Taje <i>rorua</i> , Tajio <i>orua</i> (but PKP <i>*dua</i> , Wotu <i>dua-</i> , <i>duango</i> , Balaesang <i>dorua</i> , Pendau <i>doruo</i> , Lauje <i>doluo</i> , Totoli, Boano <i>dóua</i> )
* <i>duRi</i> 'thorn'	> PKP <i>*rui</i> , Wotu, Laiyolo, Kalao, Wolio, Kamaru <i>rui</i> , Lauje, Totoli, Boano <i>lui</i> , Tialo, Dondo <i>lugitE</i> , other TT <i>rui</i>
* <i>daRaQ</i> 'blood'	> PKP <i>*raaʔ</i> , Wotu, Wolio, Kamaru <i>raa</i> , Laiyolo, Kalao <i>raʔa</i> , Lauje, Totoli <i>laa</i> , Boano <i>laʔa</i> , Tialo <i>laga</i> , Dondo <i>laganyo</i> , other TT <i>raa</i>
* <i>dejeR</i> 'hear'	> Laiyolo, Kalao, Wolio <i>rango</i> , Lauje, Tialo, Dondo <i>longo</i> , Boano <i>longa</i> (but Pamona <i>donge</i> )
* <i>danaw</i> 'lake'	> PKP <i>*rano</i> , Wolio <i>rano</i> , most TT <i>rano</i> (but Lauje, Tialo, Boano <i>dano</i> )

---

---

* <i>dalem</i> 'deep'	> PKP * <i>rala</i> , Wotu, Laiyolo, Kalao <i>lara</i> (<met.), Taje, Tajiio <i>raro</i> , Pendau, Lauje, Dondo <i>lalong</i> , Tialo <i>lalongE</i> , Totoli <i>lalom</i> , Boano <i>laom</i>
* <i>dahun</i> 'leaf'	> Uma <i>rau</i> , Balaesang, Dampelas, Taje, Pendau, Tajiio <i>roong</i> , Lauje, Tialo, Dondo <i>loongE</i> , Totoli <i>laeng</i> (but Boano <i>da?un</i> )
* <i>qudaj</i> 'shrimp'	> Pamona <i>ura</i> , Balaesang, Tajiio <i>urang</i> , Boano <i>ulang</i>
* <i>kuden</i> 'cookpot'	> PKP * <i>kura</i> , Laiyolo <i>kuro</i> , Kalao <i>kura</i>

---

Northern Tomini (Lauje, Tialo and Dondo) languages along with Totoli and Boano exhibit the further change of \**r* > /l/.

#### 4. Monophthongisation of PMP final diphthongs \*-ay and \*-ey > \*e

---

* <i>qaZay</i> 'chin'	> PKP * <i>aje</i> , Laiyolo, Kalao, Wolio, Kamaru <i>ade</i> , Balaesang, Totoli <i>ngade</i> , Boano <i>ngade?</i> , Pendau <i>nganje</i> , other TT <i>ngaje</i>
* <i>anay</i> 'termite'	> PKP * <i>ane</i> , Wolio <i>ane</i> , Balaesang, Tajiio <i>ane</i> (also Taje, others <i>siane</i> )
PWMP * <i>kiday</i> 'eyebrow'	> PKP * <i>kire</i> , Muna <i>kire</i> , Kioko <i>kixe</i> , Wanci <i>kire</i> , Balaesang <i>kire</i>
* <i>m-atey</i> 'die'	> PKP * <i>mate</i> , Laiyolo, Kalao, Wolio <i>mate</i> , Balaesang <i>maate</i> , Pendau <i>naate</i> , Lauje <i>mate</i> , all other TT <i>ate</i>
* <i>qatey</i> 'liver'	> PKP * <i>ate</i> , Wotu, Wolio, Kalao <i>ate</i> , Laiyolo <i>ati</i> , all TT <i>ate</i>

---

#### 5. Monophthongisation of PMP final diphthongs \*-aw and \*-ew > \*o

---

* <i>babaw</i> 'above'	> PKP * <i>wawo</i> , Layolo <i>bafo</i> , Kalao <i>bavo</i> , Wolio <i>bawo</i> , Taje <i>wawo</i> , Lauje <i>babo?</i> , Totoli <i>babo</i>
* <i>kasaw</i> 'rafter'	> PKP * <i>kaso</i> , Laiyolo, Kalao <i>kasu</i> , Dampelas, Taje, Lauje, Tialo <i>?aso</i> , Dondo <i>aso</i> , Totoli, Boano <i>kasu</i>
* <i>pisaw</i> 'knife'	> Wotu, Laiyolo, Kalao, Wolio, Lauje, Boano <i>piso</i> , Tialo, Dondo <i>pisoyE</i>
* <i>behew</i> 'odour, stink'	> Pamona <i>boo</i> , Dampelas <i>noboomo</i> , Pendau, Lauje, Dondo <i>boo</i> , Tialo <i>memboo</i>

---

A sixth change could possibly be added to this list, namely PMP \**e* (schwa) > \**o*, which is regular in all these languages apart from a number of exceptions where \**e* > *a* in Kaili-Pamona and Wotu-Wolio languages (Martens 1989b; Donohue in press). Even if we were to take all five (or even six) changes together, however, their value for defining a larger Celebic group remains low, since all five changes occur relatively frequently in the Austronesian world. There is, however, another change by which these languages may be grouped, and which, when added to the five changes above, makes a strong case for macrogrouping. It concerns the fate of PMP \**j*.

PMP \**j* has been reconstructed only in word-medial and -final position. Assigning a phonetic value to PMP \**j* has been problematic. While Dahl (1981:92, 152) takes the position that \**j* was a palatal stop or affricate, other scholars have favoured a velar interpretation. Wolff (1988) considered \**j* (his \**g*) to be a voiced velar stop, Blust (1990:234) considers it to be a palatalised velar stop [gʲ] which had no voiceless counterpart, while Ross (1992) has argued that it was

likely a voiced velar fricative [ɣ, ɣʷ]. Its reflexes in Tagalog are *-l-* (medially) and *-d* (finally), in Toba Batak *-g-* and *-k-*, in Malay *-d-* and *-t-*, in Javanese *r* in all positions and in Madurese *l*.

Consider now the realisation of *\*j* in Sulawesi languages. In the three microgroups of northern Sulawesi, PMP *\*j* became *\*d* in final position; compare Proto Gorontalo-Mongondow *\*pusod*, Proto Minahasan *\*pusəd* and Proto Sangiric *\*pusid* (< PMP *\*pusej* ‘navel’). In medial position *\*j* merged with the reflexes of PMP *\*Z* in Proto Gorontalo-Mongondow (Usup 1986:277–279), and with reflexes of *\*Z*, *\*D* and *\*d* in Proto Sangiric (Sneddon 1984). Among Minahasan languages, medial PMP *\*j* and *\*D* both became Proto Minahasan *\*d* following *\*e* (schwa), otherwise PMP *\*-j-* has the distinct realisation pattern of /l/ in Tonsawang and /r/ in other Minahasan languages (for which correspondence Sneddon reconstructed Proto Minahasan *\*r-*) (Sneddon 1978; 1989:97–98).

Word-medially PMP *\*j* must have been maintained as a separate phoneme into Proto South Sulawesi, as it is realised as *-s-* in Buginese but as *-r-* in other languages (on the basis of which Mills reconstructed PSS *\*z*). In word-final position, however, PMP *\*-j* presumably became *\*d* and thence merged with *\*t* in the general final-stop devoicing which occurred prior to Proto South Sulawesi. From there it is reflected as glottal stop in most present-day South Sulawesi languages (Mills 1975:553, 556).

In contrast to these four language groups—three in the north and one in the south of Sulawesi—across the rest of Sulawesi *\*j* did not merge with *\*Z* or *\*d* in any position. In fact it has only two principal realisations, either as *y* (sometimes resegmented in final position to *i*) or as zero, often accompanied by fronting of the preceding vowel. Because this change is significant for postulating a Celebic macrogroup, I discuss it here in some detail. Table 3 gives reflexes of nine PMP etyma containing *\*j*, both in medial and in final position. Non-cognate forms (lexical replacements) are indicated by underlining, while a dash (—) indicates a lack of data. Forms in bold receive further discussion below.

On the basis of these cognate sets, I reconstruct Proto Celebic *\*qapəyo*, *\*pae* (from earlier *\*paye*), *\*qaləyo*, *\*ipian* (from earlier *\*ipiyan*), *\*ngayam*, *\*quləy*, *\*laləy*, *\*pusəy* and *\*palay*. PMP *\*j* became *\*y*, and further, it would appear, became a transition glide and was subsequently lost between /a/ and a front vowel.<sup>5</sup>

From Table 3, we can also note the occurrence of ‘irregular’ reflexes of *\*j* in border areas from the influence of surrounding languages. Among the Wotu-Wolio languages spoken in the area of Salayar Island off the southern tip of South Sulawesi, Kalao *asə* ‘field rice’ (with *\*j* reflected as /s/) must clearly be ascribed to Buginese influence. Similarly Kalao *pi<sup>?</sup>du* and Laiyolo *pidu* ‘gall’ are also likely borrowings from a South Sulawesi language, particularly as the Kalao form exhibits consonant doubling (compare PSS *\*pizzu*). In Laiyolo *pəllə<sup>?</sup>* ‘palm, sole’, final glottal stop is likewise the typical South Sulawesi reflex of PMP *\*-j*, and so this form must also be ascribed to South Sulawesi influence.

<sup>5</sup> The simpler statement, that *\*j* was lost contiguous to a front vowel, is contraindicated by Padoe *penei* ‘wing’ < Proto Celebic *\*paniy* < PMP *\*panij* ‘wing’ (see Mead 1998:64). Further data — especially a full account of the fate of PMP *\*y* in medial position — may lead this statement to be refined. In languages which have been investigated to date, PMP *\*-j-* and PMP *\*-y-* shared the same fate in Kaili-Pamona (Martens 1989b), Bungku-Tolaki (Mead 1998:67) and Muna (van den Berg 1991b:14–15). Evidence also points in this direction in Saluan-Banggai languages, but (because of a lack of data) is somewhat inconclusive (Mead this volume).



**Table 3:** Etyma containing PMP \*j and reflexes in selected Sulawesi languages

	'gall, bile' *qapeju	'field rice' *pajey	'sun' *qalejaw	'when' <sup>1</sup> *i-pija-n	'name' *ŋajan	'snake' *qulej	'fly' *lalej	'navel' *pusej	'palm' *palaj
TOMINI-TOLITOLI									
Boano	<i>poyu</i>	<i>pae</i>	<i>ondo</i>	<i>pilan</i>	<i>langan</i>	<i>ule?</i>	<i>pikot</i>	<i>pusol</i>	<i>paak</i>
Totoli	<i>peu</i>	<i>bini</i>	<i>ondo</i>	<i>pilan</i>	<i>ngalan</i>	<i>ule</i>	<i>lale</i>	<i>pisol</i>	<i>palak</i>
Dondo	<i>poyu</i>	<i>bo?ung</i>	<i>oloyo</i>	<i>sogaubengi</i>	<i>tope</i>	<i>ule</i>	<i>lale</i>	<i>puse</i>	<i>pale</i>
Lauje	<i>poyu</i>	<i>bo?ung</i>	<i>oleo</i>	<i>sogaumbéng</i>	<i>tope</i>	<i>ule</i>	<i>lale</i>	<i>puse</i>	<i>pale</i>
Balaesang	<i>peit</i>	<i>boas</i>	<i>sekat</i>	<i>mpiang</i>	<i>tope</i>	<i>ule</i>	<i>lale</i>	<i>puse</i>	<i>pale</i>
Pendau	<i>opoyu</i>	<i>pae</i>	<i>eleo</i>	<i>nasae</i>	<i>sango</i>	<i>ule</i>	<i>lale</i>	<i>puse</i>	<i>tanatang</i>
KAILI-PAMONA									
Da'a	<i>mpoyu</i>	<i>pae</i>	<i>eo</i>	<i>nepia</i>	<i>sanga</i>	<i>ule</i>	<i>lale</i>	<i>puse</i>	<i>pale</i>
Pamona	<i>poju</i>	<i>pae</i>	<i>eo</i>	<i>impia</i>	<i>sanga</i>	<i>ule</i>	<i>yale</i>	<i>puse</i>	<i>pale</i>
Uma	<i>poju</i>	<i>pae</i>	<i>eo</i>	<i>nto?uma</i>	<i>hanga?</i>	<i>ule</i>	<i>dali?</i>	<i>puhe</i>	<i>pale</i>
Napu	<i>puru</i>	<i>pare</i>	<i>alo</i>	<i>impira</i>	<i>hanga</i>	<i>ile</i>	<i>dale</i>	—	<i>palanta</i>
Bada	<i>puru</i>	<i>pare</i>	<i>alo</i>	<i>himpira?</i>	<i>hanga?</i>	<i>ile</i>	<i>dali?</i>	—	<i>palanta?</i>
Besoa	<i>puru</i>	<i>pare</i>	<i>alo</i>	<i>impira</i>	<i>hanga?</i>	<i>leloto</i>	<i>dali?</i>	<i>pohi?</i>	<i>palanta?</i>
WOTU-WOLIO									
Wotu	—	<i>bae</i>	<i>iyo</i>	<i>dipia</i>	<i>sanga</i>	<i>ulo</i>	<i>lale</i>	—	—
Wolio	<i>mapai</i>	<i>bae</i>	<i>eo</i>	<i>naipia</i>	<i>saro</i>	<i>ulo</i>	<i>lale</i>	<i>puse</i>	<i>randa</i>
Kalao	<i>pi?du</i>	<i>asə</i>	<i>ajo</i>	<i>laipia</i>	<i>sanga</i>	<i>ulo</i>	<i>lale</i>	—	<i>pele</i>
Laiyolo	<i>pidu</i>	<i>bae</i>	<i>ajo</i>	<i>ripia</i>	<i>sanga</i>	<i>ulo</i>	<i>lale</i>	<i>puse</i>	<i>palla?</i>

<sup>1</sup> PMP \*pija 'how many' has been lexically replaced in a number of Celebic languages, therefore I have chosen instead to cite reflexes of \*i-pija-n 'when' (which strictly may not be reconstructible to PMP; see Mead 2001:170). The pattern of realisation of \*j is the same, but more obvious from reflexes of the latter.

	'gall, bile' <i>*qapeju</i>	'field rice' <i>*pajey</i>	'sun' <i>*qalejaw</i>	'when' <i>*i-pi ja-n</i>	'name' <i>*ηajan</i>	'snake' <i>*qulej</i>	'fly' <i>*lalej</i>	'navel' <i>*pusej</i>	'palm' <i>*palaj</i>
SALUAN-BANGGAI									
Banggai	<i>sopot</i>	<i>labue</i>	<i>oloyo</i>	<i>noian</i>	<i>sambu</i>	<i>uloy</i>	<i>poos</i>	<i>pusoy</i>	<i>palalap</i>
Balantak	<i>opoyu?</i>	<i>pae</i>	<i>ilio</i>	<i>ipi</i>	<i>ngaan</i>	<i>ule</i>	<i>laale</i>	<i>puse</i>	<i>palaa</i>
Andio	<i>poyu?</i>	<i>pae</i>	<i>sina</i>	<i>ipian</i>	<i>ngaan</i>	<i>ulo</i>	<i>laalo?</i>	<i>puse</i>	<i>pala</i>
Saluan	<i>pou?</i>	<i>pae</i>	<i>sina</i>	<i>hipian</i>	<i>sanggo</i>	<i>ulo, uloo</i>	<i>lalo, laloo</i>	<i>pusoo</i>	<i>palaa</i>
Bobongko	<i>opou?</i>	<i>pae</i>	<i>dolag</i>	<i>torikuka</i>	<i>sanggor</i>	<i>bintana?</i>	<i>lalo</i>	<i>puso</i>	<i>pala</i>
BUNGKU-TOLAKI									
Kulisusu	<i>upeu</i>	<i>pae</i>	<i>oleo</i>	<i>impia</i>	<i>ngee</i>	<i>ule</i>	<i>lale</i>	<i>puhe</i>	<i>pele</i>
Mori Bawah	<i>upeu</i>	<i>pae</i>	<i>oleo</i>	<i>te?ipia</i>	<i>ngee</i>	<i>ule</i>	<i>lale</i>	<i>puhe</i>	<i>pele</i>
Padoe	<i>upeu</i>	<i>pae</i>	<i>olo</i>	<i>te?epie</i>	<i>nee</i>	<i>ule</i>	<i>laloi</i>	<i>puhoi</i>	<i>palai</i>
Tolaki	<i>posu</i>	<i>pae</i>	<i>oleo</i>	<i>te?ipia</i>	<i>tamo</i>	<i>ule</i>	<i>lale</i>	<i>puhe</i>	<i>pele</i>
MUNA-BUTON									
Muna	<i>ghufei</i>	<i>pae</i>	<i>gholeo</i>	<i>naefie</i>	<i>nea</i>	<i>ghule</i>	<i>pepi</i>	<i>puhe</i>	<i>randa</i>
Kioko	<i>piu</i>	<i>bae</i>	<i>oleo</i>	<i>naifie</i>	<i>kona</i>	<i>ule</i>	<i>buhoto</i>	—	<i>handa</i>
Cia-Cia	<i>hopiu</i>	<i>bae</i>	<i>holeo</i>	—	<i>ngea</i>	<i>kule</i>	<i>bei</i>	<i>puse</i>	<i>randa</i>
Tukang Besi	<i>ho?ou</i>	<i>bae</i>	<i>?oloo</i>	<i>ehia</i>	<i>ngaa</i>	<i>sa?a</i>	<i>lalo</i>	—	—

Non-*y* reflexes of PMP \**j* are also found among the Badaic languages of Central Sulawesi; compare Napu, Bada, Besoa *puru* ‘gall’, *pare* ‘field rice’, Napu, Besoa *impira*, Bada *himpira* ‘when?’, Besoa *pohi?* ‘navel’. As delineated above, /*r*/ (medially) and glottal stop (finally) are typical South Sulawesi reflexes of PMP \**j*, compare Seko *puru* ‘gall’, Sa’dan Toraja *pare* ‘field rice’, Seko Tengah *napiranga* ‘when?’ and Seko Lemo *posi?* ‘navel’. Likewise Uma, Bada and Besoa *dali?* and Napu *dale* ‘fly’ exhibit similarities with South Sulawesi languages. The irregular dissimilation of \**l* > /*d*/ observed in these forms is also found across the border in Mamuju and Seko (Mills 1975:748), but since in these South Sulawesi languages PMP \**-ej* > PSS \**-it* > \**-it* > /*i?*/ is regular, there can be no doubt about the direction of borrowing. Martens (1989a) considers such cases to reflect a long period of contact and borrowing between the Badaic languages and South Sulawesi languages, particularly Seko. He also considers it probable that Napu, Bada, and Besoa *alo* ‘sun, day’ was likewise borrowed from South Sulawesi — compare Seko *alo* ‘day’ (< PSS \**ilzo*) — whereas other Kaili-Pamona languages consistently have *eo*.

This large number of supposed borrowings, however, leaves Napu and Besoa *ile* ‘snake’ as the only form known to me where it could be proposed that PMP \**j* passed through a *y*-stage. This form, however, has unexplained \**u* > /*i*/ in the initial syllable. Furthermore, this irregularity is also found in the northern South Sulawesi language area; compare Seko Lemo, Mamuju and Bambam *ile* (Mills 1975:877). While it is possible to account for the presence of non-*y* reflexes of \**j* in Badaic languages through borrowing, the sheer number of exceptions suggests an alternative hypothesis: Badaic languages are genetically South Sulawesi languages.<sup>6</sup>

An inspection of Table 3 also reveals that in several cases the Boano and Totoli forms also exhibit unusual (non-*y*) reflexes of \**j*. Both of these languages are located in the northern area where Celebic languages border the Gorontalic subgroup, and these forms may reflect influence from Gorontalo. However, until a study of historical sound change in the Tomini-Tolitoli languages is undertaken, the identification of loan words remains problematic. Indeed, the question remains open whether Boano and Totoli should even be subgrouped with the Tomini languages (see Himmelman 2001:19–20).

<sup>6</sup> Besides the fate of PMP \**j*, other areas of difference to explore include PMP \**R* (reflected as zero in other Kaili-Pamona languages but as /*r*/ in South Sulawesi languages), PMP \**-uy* (reflected as /*u*/ in other Kaili-Pamona languages, but as /*i*/ in South Sulawesi languages), and PMP \**-uq* and \**-iq* (vowels were lowered in Proto South Sulawesi, but not in Proto Kaili-Pamona).

The claim that Badaic languages reflect PMP \**q*, while this phoneme was lost in Proto South Sulawesi (Mills 1975:518), is a red herring. Unknown to Mills, Seko Padang reflects PMP \**q* as length on the vowel, either when \**q* occurred in final position or contiguous to a final vowel (in the latter case, presumably via intermediate metathesis/float to final position). Therefore reflexes of PMP \**q* must be reconstructed for Proto South Sulawesi regardless. Compare the following data (Laskoswke 1995 and pers. comm.).

PMP		Seko Padang
* <i>puluq</i> ‘ten’	>	<i>pulo:</i>
* <i>dilaq</i> ‘tongue’	>	<i>lila:</i>
* <i>tuqah</i> ‘old’	>	* <i>tuq</i> > <i>tua:</i>
* <i>taqi</i> ‘feces’	>	* <i>taiq</i> > <i>tai:</i>

Apart from these exceptions, only a few other forms would appear to have non-y reflexes of PMP \*j. The Uma and Pamona form for 'gall', *poju* [podzu], and the Tolaki form *posu*, however, are not borrowed, but rather result independently from fortition of \*y, which is regular in defined contexts in Uma and Pamona (Martens 1989b) and also occurred sporadically in Tolaki (Mead 1998:115). The Kalao and Laiyolo form for 'sun', *ajo* [adzo], is unexplained in the present analysis. As far as I can tell it has not been borrowed from any present-day South Sulawesi language, and perhaps is an inherited form (from earlier \*alyo?).

Finally, while PMP \*j appears to have become zero or merged with PMP \*y in medial position, on the basis of counter-additive reasoning the change of PMP \*-j > Proto Celebic \*y must have occurred *after* the monophthongisation of PMP \*-ay and \*-ey to \*e. Among Celebic languages PMP \*-ay and \*-ey are universally reflected as /e/. And while in many cases PMP \*-aj and \*-ej are also reflected as /e/, enough languages reflect these otherwise to indicate that a merger of PMP \*-ay, \*-ey, \*-aj and \*-ej did not occur.

I would also like to make an initial proffering of two lexical innovations which support the newly proposed Celebic macrogroup. These forms are:

PCel \*panianan 'parent-in-law' > Totoli *poneanan* 'parent-in-law' (also 'child-in-law?'), Boano *ponianan* 'parent-in-law' (also 'child-in-law?'), Uma *piniana* 'parent-in-law', Kulawi *paniana* 'parent-in-law', Kulisusu *poniana* 'parent-in-law', Moronene, Mori Bawah *poni* 'parent/child-in-law' (reciprocal term). Compare also Moronene, Muna *samponi* 'child's spouse's parent' (with prefix *sa-* 'one').

PCel \*manian 'child-in-law' > Dampelas *maniang*, Lauje *meniang*, Tialo, Dondo *monianE* 'parent-in-law' (also 'child-in-law?'), Pendau *meniang* 'parent-in-law' (*not* reciprocal), Da'a, Kulawi, Lindu, Pamona (Ampana dialect) *mania* 'child-in-law', Uma *minia* 'child-in-law', standard Pamona (Adriani 1928:s.v.), Wolio *mania* 'child/parent-in-law' (reciprocal term), Balantak *monian* 'child/parent-in-law' (reciprocal term), Kulisusu *ana monia* 'child-in-law'.

Only three languages known to me — Kulawi and Uma in the Kaili-Pamona area, and Kulisusu in the Bungku-Tolaki area — have retained both forms (though further lexical research may uncover other languages where this is so). Nonetheless, in that these words constitute a derivationally related pair, their reconstruction is mutually supported. That a number of daughter languages reflect only one member of the pair must be accounted for by semantic shift to a reciprocal term, lexical replacement, or both (compare for example Pendau which now has *meniang* 'parent-in-law' and unrelated *tomodait* 'child-in-law'). For such reasons it is also untenable that present-day forms could have obtained their distribution through borrowing.

## 7 Conclusions

The principal results of this study are summarised in Figure 3. This figure shows the branches and nodes which have been argued for in the preceding section. As noted there and again below, some branches have a tentative status. At the highest level I propose a group which encompasses almost all the languages of central, eastern and south-eastern Sulawesi as well as a handful of languages located in South Sulawesi. Following van den Berg (1996b), I label this the Celebic supergroup. Notably, van den Berg is the only other author to have speculated about a

macrogroup comprising these same languages. However, he did not define his reasons for including the Saluan-Banggai or Tomini-Tolitoli languages in his macrogroup, and presented a different view of subgrouping within it.

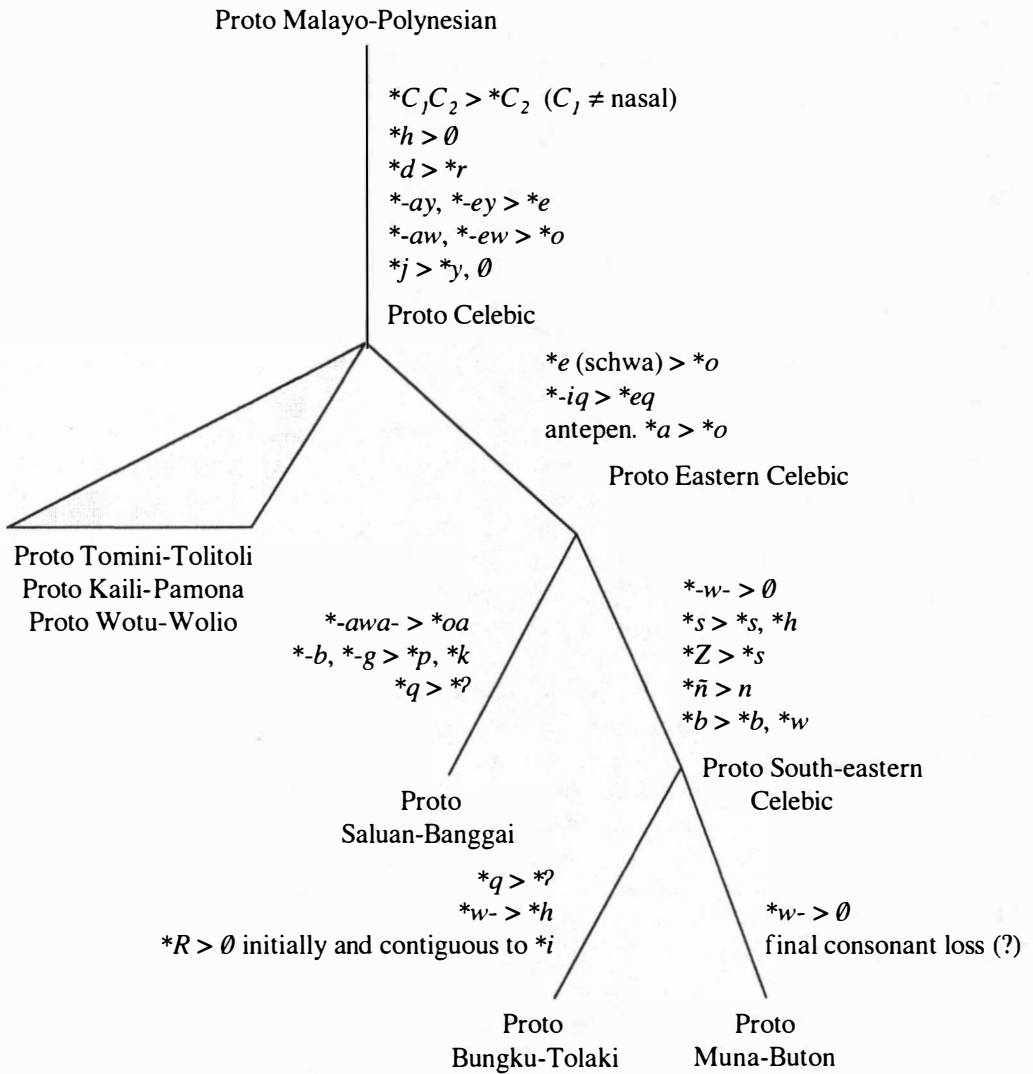


Figure 3: The Celebic supergroup

At present it is unclear how many primary branches should be posited under the Celebic node. The shaded triangle in Figure 3 indicates that while there is evidence for bringing the Tomini-Tolitoli, Kaili-Pamona and Wotu-Wolio subgroups into a Celebic Supergroup, I make no claim as to how these groups may be related to each other. On the other hand, there are three sound changes which would apparently allow us to distinguish Eastern Celebic as a separate branch.

The change of PMP *\*e* (schwa) > *\*o* is, however, also broadly characteristic of all Celebic languages. But because some exceptional cases where *\*e* > *\*a* are known from Kaili-Pamona and Wotu-Wolio languages, I have not assigned this change to a higher level. The raising of PMP antepenultimate *\*a* > *\*o* is also a change of a general nature. This leaves the lowering of PMP *\*i* > *\*e* before final *\*q* as the most significant change for subgrouping Eastern Celebic languages together. To be sure, this same lowering has been noted as an independent change elsewhere, even in other parts of Sulawesi including Proto Sangiric (Sneddon 1989:90), Proto South Sulawesi (Mills 1975:545), and, perhaps not surprisingly, in the Badaic languages (Martens, cited in Sneddon 1989:90). In these other cases, however, the lowering of *\*i* > *\*e* before *\*-q* was accompanied by a parallel lowering of *\*u* > *\*o* in the same environment. Only in Eastern Celebic do we find lowering of *\*i*, *unmatched* by a parallel lowering of *\*u*.<sup>7</sup>

Despite the encompassing nature of the present study, important questions remain. Here I mention areas for further research in the area of Sulawesi historical linguistics. Several of these issues have been mentioned in the preceding sections, but it seems beneficial to bring them together here. Many of the following comments can be considered additional commentary on Figure 3.

*What is the position of Boano and Totoli?* Totoli is spoken by 25,000 speakers in the border area where Celebic languages meet Gorontalic languages, and Boano in the same area by about a tenth that many speakers. These two languages are not Gorontalo-Mongondow languages, but if reflexes of PMP *\*j* are an indication (Table 3 above), they may not fit comfortably into a Celebic group as defined here either. Himmelmann, who has provided a wealth of new data on Tomini-Tolitoli languages, simultaneously cautions that the genetic unity of this group has never been established. In particular he singles out Totoli and Boano for their divergent phonology, lexicon and grammar (Himmelmann 2001:20).

*What is the relationship between the Kaili-Pamona and Wotu-Wolio languages?* A number of tantalising leads have been proposed, but none have been followed up. The Dutch linguist J.C. Anceaux recognised striking structural similarities between Wolio and Pamona, but never published on this topic (René van den Berg pers. comm.). Donohue (in press) suggests the change of PMP *\*e* (schwa) > *\*a* in a number of lexical items (otherwise regularly PMP *\*e* > /o/) could be used to link the Kaili-Pamona and Wotu-Wolio languages. Even a consideration of geographical location suggests that the Wotu-Wolio languages are more likely to share a closer relationship to Kaili-Pamona than to Tomini-Tolitoli languages. Unfortunately, only Wolio has been well documented, and other Wotu-Wolio languages have been influenced by South Sulawesi languages. A careful study will be needed to pull out the traces of historical connection. To date, not even an internal classification of the five Wotu-Wolio languages has emerged.

*What is the position of the Badaic languages?* Culturally, speakers of Napu, Bada and Besoa (collectively referred to as the Badaic languages) identify with their Kaili-Pamona neighbours, yet their word stock bears affinities to South Sulawesi languages. Martens (1989) investigated this situation, and concluded that Badaic languages were genetically Kaili-Pamona languages that had borrowed lexically from South Sulawesi languages, particularly Seko. In that Badaic languages overwhelmingly reflect *\*j* as /r/ (medially) or glottal stop (finally) — the prototypical South Sulawesi reflexes — this study suggests the opposite. That is to say, the Badaic languages

<sup>7</sup> The lowering of *\*u* occurred as a further change in two sub-branches of Eastern Celebic, namely Eastern Saluan-Banggai (Mead this volume) and Western Bungku-Tolaki (Mead 1996:80ff.).

may genetically be South Sulawesi languages which have been influenced lexically by Kaili-Pamona languages.

*What is the internal classification of South Sulawesi languages?* If Badaic languages are genetically South Sulawesi languages, then clearly Mills' (1975:490ff.) internal classification of South Sulawesi languages needs to be reworked. In fact such a re-evaluation has been needed on other grounds. In particular, Mills' internal groupings were based heavily (but not exclusively) on what happened to consonants in word-final position. Experience elsewhere in Sulawesi, however, has shown that processes of final consonant weakening, merger and loss are likely to exhibit areal diffusion, and thus not to be valid indicators of genetic affiliation (Sneddon 1993; Mead 1996). Since Adelaar (1994), it has been clear that the Tamanic languages of Borneo are genetically South Sulawesi languages, most closely related to Buginese. These languages, which have been more conservative with regard to final consonants, are likely to help in sorting out what changes can (or cannot) be attributed to higher genetic levels within South Sulawesi.

*Are Muna-Buton and Bungku-Tolaki valid subgroups?* Evidence from sound change alone is — and will probably remain — an insufficient basis for establishing a Muna-Buton subgroup. As discussed in §3, a Muna-Buton grouping is provisionally accepted based on the probable lexical innovations set forth in van den Berg (this volume). I have not touched on the genetic unity of the Bungku-Tolaki languages. Although the loss of PMP \*R initially and contiguous to \*i can be attributed to Proto Bungku-Tolaki (not mentioned previously in this paper, but detailed in Mead 1998:58–60), this change will need to be taken in hand with the fate of PMP \*R in nuclear Muna-Buton languages. The remaining changes attributed in Figure 3 to Proto Bungku-Tolaki (PMP \*w- > \*h, PMP \*q > glottal stop) are less consequential for subgrouping. Lexical innovations in support of a Bungku-Tolaki subgroup can be found in Mead (1998:86–87). Within the framework provided by this study, can additional evidence be adduced in support (or refutation) of these two subgroups?

Despite these unknowns, by stages we are improving our understanding of historical relationships among Sulawesi languages. Instead of the ten subgroups listed at the beginning of this paper, following the results of this study we need recognise only five genetic groupings across the island of Sulawesi (though the position of Totoli and Boano, tentatively included as Celebic languages, remains suspect):

Sangiric  
 Minahasan  
 Gorontalo-Mongondow  
 Celebic  
 South Sulawesi

Of these five groups, Gorontalo-Mongondow could be considered exo-Sulawesi in that these languages originated at a later date from the Philippines, Proto Gorontalo-Mongondow speakers supposedly arriving in northern Sulawesi around 500BC (Blust 1991:103–104). Whether the other four groups are truly indigenous languages of Sulawesi — directly descended from their Proto Malayo-Polynesian ancestor — or whether they too share higher-level genetic relationships to languages outside of Sulawesi, remains to be seen. Sangiric and Minahasan languages have long been considered 'Philippine' languages on typological grounds, but proving a genetic connection to particular (or all) Philippine languages has proved more elusive. Initial evidence for their inclusion in a larger Philippines group has been accumulated by Zorc (1986),

but consists of lexical innovations only. Among his ninety-eight Proto Philippine lexical innovations, sixteen have reflexes in Proto Minahasan or a Minahasan language, while fifteen have reflexes in Proto Sangiric or a Sangiric language. This evidence has not been critically evaluated.

A link between South Sulawesi languages and languages of central and south-eastern Sulawesi has been disparaged. Adriani commented, 'Van het Boegineesch onderscheiden zich de Toradjasche talen in vele opzichten' (The Torajan [Kaili-Pamona] languages distinguish themselves from Buginese in many respects), then proceeded to marshal two pages worth of evidence (Adriani & Kruyt 1914:91–93). In similar fashion Mills noted that South Sulawesi languages differed from languages of central and south-eastern Sulawesi in respect to both historical sound change and verb morphology, and was inclined to note instead, with few specifics, affinities between South Sulawesi languages, Malay, and Madurese (Mills 1975:499, 517–519, 1981:60). Ross, however, has suggested that South Sulawesi may not be as distinct from Celebic languages as is often supposed, at least typologically. He partially reconstructs a 'Proto Sulawesi' system of verb morphology, from which South Sulawesi systems could also be derived (Ross 2002:462–464).

Finally, it is interesting to note how much the view of macrogrouping of Sulawesi languages presented above comes back around to a view presented ninety years ago by the pioneer of Sulawesi linguistic studies, the missionary Nicolas Adriani.

Door zijne verwantschap met het Bobongkosch (op de Togian-eilanden) en het Gorontaleesch, wijst het Loinansch op eene strooming der bevolking van de N. helft van het Noordelijk schiereiland naar het Z. toe, die voorbij het gebeid van het Gorontaleesch zich heeft verdeeld, waarbij de Loinansche tak over de Togian-eilanden naar den vasten wal ten O. van Tandj. Api is gegaan, om zich daarop naar het O. (Balantaksch) en naar het Z. (Boengkoesch) te verbieden, terwijl een andere tak naar het W., daarop naar het Z. is gegaan, en ten Z. van den Evenaar weder een tak in O.lijke richting heeft afgescheiden. Zoo mag men dus aannemen dat het Bare'e, als meest O.lijke uitlooper van dezen laatstgenoemden zijstroom, bij Tandj. Api weder op het Loinansch is gesluit.

[By its relationship with Bobongko (on the Togian Islands) and Gorontalo, the Loinan language points to a southward migration of the inhabitants of the north half of the northern peninsula, which divided near present-day Gorontalo: the Loinan [Saluan] branch proceeded via the Togian islands to the further shore east of Tanjung Api, spreading therefrom to the east (Balantak) and to the south (Bungku), while another branch proceeded to the west and then to the south, and then south of the equator a further branch separated back in an easterly direction. So one may consider that Bare'e [Pamona], as the most eastern extension of this last-named side flow, was arrested up against Loinan.] (Adriani & Kruyt 1914:89)

Given that Adriani developed his picture of migration at a time when typological concerns played an equal role with sound change in determining language relationships, it is remarkable how little needs to be changed. A clear amendment to Adriani's hypothesis is that we can no longer maintain that Celebic languages are descended from Gorontalic languages. Adriani's Loinan branch, however, has clear parallels to the Eastern Celebic branch proposed in this study, though here we must add that this 'southward flow' did not stop at present-day Bungku, but eventually encompassed all of peninsular south-eastern Sulawesi, including Muna, Buton, and the Tukang Besi Islands. Whether the Tomini-Tolitoli and Kaili-Pamona (and the now recognised Wotu-Wolio) languages together constitute a second, separate branch of Celebic, as



Adriani also supposed, remains to be seen. Once that issue has been settled, we can begin (again) to theorise about a Proto Celebic homeland.

Considering that Adriani had far less — and less reliable — data than were available to this author, the present work stands as a tribute to his genius and prescience.

## References

- Adelaar, K. Alexander, 1994, The classification of the Tamanic languages. In Tom Dutton and Darrell T. Tryon, eds *Language contact and change in the Austronesian world* 1–42. Trends in Linguistics, Studies and Monographs 77. Berlin: Mouton de Gruyter.
- Adriani, N., 1928, *Bare'e-Nederlandsch woordenboek met Nederlandsch-Bare'e register*. Leiden: E.J. Brill.
- Adriani, N. and A.C. Kruyt, 1914, *De Bare'e-sprekende Toradja's van Midden Celebes*, vol. 3: *Taal- en letterkundige schets der Bare'e-taal en overzicht van het taalgebied Celebes-Zuid-Halmahera*. Batavia: Landsdrukkerij.
- Anceaux, J.C., 1987, *Wolio dictionary (Wolio-English-Indonesian)*. Dordrecht: Foris.
- Berg, René van den, 1991a, Muna dialects and Munc languages: towards a reconstruction. In Ray Harlow, ed. *VICAL 2: Western Austronesian and contact languages: papers from the Fifth International Conference on Austronesian Linguistics*, 21–51. Auckland: Linguistic Society of New Zealand.
- 1991b, Muna historical phonology. In James N. Sneddon, ed. *Studies in Sulawesi linguistics*, part 2, 1–28. NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, vol. 33. Jakarta: Universitas Katolik Indonesia Atma Jaya.
- 1991c, Preliminary notes on the Cia-Cia language (South Buton). In Harry A. Poeze and Pim Schoorl, eds *Excursions in Celebes: een bundel bijdragen bij het afscheid van J. Noorduyt als directeur-secretaris van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde* Verhandelingen van het KITLV 147, 305–324. Leiden: KITLV Press.
- 1996a, *Muna-English dictionary*. Leiden: KITLV Press.
- 1996b, The demise of focus and the spread of conjugated verbs in Sulawesi. In H. Steinhauer, ed. *Papers in Austronesian linguistics*, No. 3, 89–114. Canberra: Pacific Linguistics.
- Blust, Robert, 1990, Summary report: linguistic change and reconstruction methodology in the Austronesian language family. In Philip Baldi, ed. *Linguistic change and reconstruction methodology*, 133–153. Trends in Linguistics, Studies and Monographs 45. Berlin: Mouton de Gruyter.
- 1991, The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–129.
- Dahl, Otto Christian, 1981, *Early phonetic and phonemic changes in Austronesian*. Oslo: Universitetsforlaget.
- Donohue, Mark, 1999, *A grammar of Tukang Besi*. Mouton Grammar Library 20. Berlin: Mouton de Gruyter.
- 2000, Tukang Besi dialectology. In Charles E. Grimes, ed. *Spices from the east: papers in languages of eastern Indonesia*, 55–72. Canberra: Pacific Linguistics.
- in press, The pretenders to Muna Buton. In John Bowden and Nikolaus Himmelmann, eds *Papers in Austronesian dialectology and subgrouping*. Canberra: Pacific Linguistics.

- Esser, S.J., 1927, *Klank- en vormleer van het Morisch*, part 1. *Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen*, vol. 67/3. Leiden: Vros.
- 1933, *Klank- en vormleer van het Morisch*, part 2. *Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen*, vol. 67/4. Bandoeng: Nix.
- Ferguson, Charles A., 1963, Some assumptions about nasals: a sample study in phonological universals. In Joseph H. Greenberg, ed. *Universals of language*, 2nd edn, 53–60. Cambridge: MIT Press.
- Himmelmann, Nikolaus, 2001, *Sourcebook on Tomini-Tolitoli languages: general information and word lists*. Canberra: Pacific Linguistics.
- Laidig, Wyn D. and Sahabu Dg. Maingak, 1999, Barang-barang phonology: a preliminary description. In Wyn D. Laidig, ed. *Studies in Sulawesi linguistics*, part 7, vol. 46, 46–83. NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia. Jakarta: Universitas Katolik Indonesia Atma Jaya.
- Laskowske, Thomas, 1995, A grammatical sketch of Seko Padang. Unpublished typescript.
- Martens, Michael, 1988, Notes on Uma verbs. In H. Steinhauer, ed. *Papers in Western Austronesian linguistics*, No.4, 167–237. Canberra: Pacific Linguistics.
- 1989a, The Badaic languages. In Sneddon, ed. 1989:19–53.
- 1989b, Proto Kaili-Pamona: reconstruction of the protolanguage of a language subgroup in Sulawesi. Unpublished typescript, Hasanuddin University and SIL.
- Mead, David, 1996, The evidence for final consonants in Proto-Bungku-Tolaki. *Oceanic Linguistics* 35:180–194.
- 1998, Proto-Bungku-Tolaki: reconstruction of its phonology and aspects of its morphosyntax. PhD dissertation, Rice University.
- 1999, *The Bungku-Tolaki languages of south-eastern Sulawesi, Indonesia*. Canberra: Pacific Linguistics.
- 2001, The numeral confix *\*i- -(e)n*. *Oceanic Linguistics* 40:167–176.
- 2002, Proto Celebic focus revisited. In Wouk and Ross, eds 2002:143–177.
- Mills, Roger F., 1975, Proto South Sulawesi and Proto Austronesian Phonology. PhD dissertation, University of Michigan.
- 1981, Additional addenda. In Robert A. Blust, ed. *Historical linguistics in Indonesia*, part 1, vol. 10, 59–81. NUSA Linguistic Studies in Indonesian and Languages in Indonesia. Jakarta: Universitas Atma Jaya.
- Ross, Malcolm D., 1992, The sound of Proto-Austronesian: an outsider's view of the Formosan evidence. *Oceanic Linguistics* 31:23–64.
- 2002, Final words: research themes in the history and typology of western Austronesian languages. In Wouk and Ross, eds 2002:451–474.
- Sirk, U.K., 1981, The South Sulawesi group and neighboring languages. *Indonesia Circle* 25:29–36.
- Sneddon, James N., 1978, *Proto-Minahasan: phonology, morphology and wordlist*. Canberra: Pacific Linguistics.
- 1984, *Proto-Sangiric and the Sangiric languages*. Canberra: Pacific Linguistics.

- 1989, The North Sulawesi microgroups: in search of higher level connections. In Sneddon, ed. 1989:83–107.
- 1993, The drift towards final open syllables in Sulawesi languages. *Oceanic Linguistics* 32:1–44.
- Sneddon, James N., ed., 1989, *Studies in Sulawesi linguistics, part 1* NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, vol. 31. Jakarta: Universitas Katolik Indonesia Atma Jaya.
- Usup, Hunggu Tadjuddin, 1986, Rekonstruksi proto-bahasa Gorontalo-Mongondow. PhD dissertation, Universitas Indonesia.
- Wouk, Fay and Malcolm Ross, eds, 2002, *The history and typology of Western Austronesian voice systems*. Canberra: Pacific Linguistics.
- Wolff, John U., 1988, The PAN consonant system. In Richard McGinn, ed. *Studies in Austronesian linguistics*, 125–147. Ohio University Monographs in International Studies, Southeast Asia Series 76. Athens, OH: Center for International Studies.
- Zorc, R. David, 1986, The genetic relationships of Philippine languages. In Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 147–173. Canberra: Pacific Linguistics.

