

# Problems with Blust's Phonology for Protoaustronesian

By Connor O'Donnell in 2025

[www.parallax.net.nz](http://www.parallax.net.nz)

Blust has done a truly excellent job at reconstructing the Protoaustronesian language. However, I feel that there are numerous extremely obvious discrepancies in his reconstructed consonant inventory, which is displayed below in a condensed format.

*/p/, */b/	*/t/, */d/	*/D/	*/k/, */g/, */j/	*/q/
*/m/	*/n/	*/ñ/	*/ŋ/	
	*/r/			*/R/
	*/S/	*/s/		*/h/
	*/C/	*/c/, */z/		
*/w/		*/y/		
	*/l/	*/N/		

When converted into the standard international phonetic alphabet, it becomes the following.

*/p/, */b/	*/t/, */d/	*/d/	*/k/, */g/, */g <sup>j</sup> /	*/q/
*/m/	*/n/	*/ɲ/	*/ŋ/	
	*/r/			*/R/
	*/s/	*/ç/		*/h/
	*/ts/	*/çç/, */j̄j̄/		
*/w/		*/j/		
	*/l/	*/ʎ/		

To avoid confusing people who are already familiar with the subject, I will use Blust's transcription system instead of the international phonetic alphabet when referencing his reconstruction.

## \*/g/ and \*/j/

These have mutually exclusive distributions, and the only difference in their realisations is that one is palatalised. Therefore, they're allophones. I would reconstruct them as \*/g/ which is allophonically realised as \*[j] intervocally and in coda position.

## \*/ñ/

In Blust's dictionary, there are only ten individual words that contain this phoneme.

### Group 1

- \*/ñamñam/, \*/tañam/
- \*/ñamuR/
- \*/qañud/

## Group 2

- *\*/ñawñaw/, \*/bañaw/, \*/Səñaw/, \*/Siñaw/*
- *\*/pəñu/*
- *\*/təñəb/*

The words in group 1 appear to have the same outcomes in Formosan languages as they would have if they contained *\*/N/* instead. For instance, Thao yields /z/ in its reflexes of both *\*/ñamñam/* and *\*/naNaq/*, Atayal yields /l/ in its reflexes of both *\*/tañam/* and *\*/RiNuk/*, and Saisiyat yields /l/ in its reflexes of both *\*/qañud/* and *\*/baNaS/*.

The words in group 2 appear to have the same outcomes in Formosan languages as they would have if they contained *\*/n/* instead. They appear to be reconstructed with *\*/ñ/* in Protoaustronesian on the basis that there are a few Malayopolynesian cognates that ended up with /ŋ/.

As we know, both *\*/n/* and *\*/N/* merged into *\*/n/* in Protomalayopolynesian. I suspect these ten words were part of a small sporadic shift from *\*/n/* to *\*/ñ/* during the Protomalayopolynesian era, not before. All other instances of *\*/ñ/* in Protomalayopolynesian are in words that are not found in Formosan languages.

Therefore, based on the extremely low distribution, bivariability in Formosan reflexes, and the ability to produce the same outcomes with other existing phonemes, I would reconstruct it as either *\*/n/* or *\*/N/*, depending on whether the word is in group 1 or group 2.

### *\*/N/*

In most Formosan languages, this became either /n/ or a type of /l/, in which case, it had a slight tendency to be palatalised. In Protomalayopolynesian, it simply became *\*/n/*. To me, it seems that it could easily be reconstructed as either *\*/ʎ/* or *\*/ŋ/* in Protoaustronesian. I prefer to use *\*/ŋ/* because it completes the row of nasals in the case that Blust's *\*/ñ/* did not exist.

## **\*/s/ and \*/c/**

<b>PAN *susu<sub>1</sub> female breast, udder</b> 		
<b>Form.</b>		
Saisiyat	hoehœʔ	female breast
	ka-hœhœʔ-an	nipple of the breast
Hoanya	tutu	female breast
Sakizaya	cucuʔ	female breast
Thao	pa-tutu-an	nurse him! (e.g. crying child)
	tutu	breast
Bunun	susuʔ	female breast
	susu	breast
Bunun (Isbukun)	susuʔ	female breast
Saaroa	ʔu-susu	female breast
Amis	coco	breast of a female; udder; teat
	mi-coco	to suck or nurse; give the breast to
Puyuma	ʃuʃu	breasts
	ra-ʃuʃu-an	having big breasts
Paiwan	tutu	breasts

As seen in the image, in most Formosan languages, \*/s/ became either /s/, /t/, or /tʃ/. In Protomalayopolynesian, it became /s/. The only evidence of palatalisation at this stage is in Puyuma as far as I can see. Based on those outcomes, I would reconstruct that it was originally \*/tʃ/.

In Blust's dictionary, there are only four words that contain \*/c/.

- \*/lœcɪk/
- \*/pœcək/
- \*/pœcœq/
- \*/qapucuk/

All of these appear to have the same outcomes in Formosan languages as they would have if they contained Blust's \*/s/ instead. They appear to be reconstructed with \*/c/ in Protoaustronesian on the basis that several Western Malayopolynesian cognates ended up with /tʃ~c/. However, much like \*/ñ/, I suspect that these four words were affected by a small sporadic shift from \*/s/ to \*/c/ during the Protowestern Malayopolynesian era. All other instances of \*/c/ in Protowestern Malayopolynesian are in words that are not found in any

Formosan language, and don't contain /tʃ~c/ in any other branch of Malayopolynesian as far as I can tell.

Therefore, I would reconstruct both \*/s/ and \*/c/ as \*/tʃ/. Separating them in Protoaustronesian is like trying to reconstruct an underlying reason in Old English for the trap bath split to happen centuries later in a few dialects.

**\*/C/**

PAN *Caliŋa ear 𑖀		
Form.		
Bunun	taiŋa	ear
Saaroa	caliŋa	ear
Rukai (Budai)	caliŋa	ear
Rukai	caliŋa	ear
Rukai (Mantauran)	caliŋa	ear
Puyuma	Taŋila-ŋila-yan	the Jew's ear fungus, Auricularia-Auricula Judae
Paiwan	tsaliŋa	ear
	tsaliŋa-liŋa	Coleus scutellarioides; Iresine herbstii
	tsaliŋa nua djaum	eye of a needle
Taokas	sanina	ear
Favorlang/Babuza	charina	ear

As seen in the image, in most Formosan languages, this became either /s/, /t/, or /tʃ/. In Protomalayopolynesian, it became \*/t/. The range of its outcomes are pretty much exactly the same as those of \*/s/, just in different places, so it was clearly a separate phoneme from \*/s/.

I would choose to reconstruct it as \*/c/, because it mirrors the outcomes of Blust's \*/N/. In Protomalayopolynesian, \*/N/ merged into \*/n/, and \*/C/ merged into \*/t/. Those two defining shifts would be in harmony if \*/N/ and \*/C/ shared the same place of articulation and each had the same manner of articulation as the phonemes they became. It effectively turns it into one simple shift, which can be parsimoniously defined as "The alveolar palatal distinction was lost", instead of trying to explain how an affricate turned into a plosive and at the same time a palatal approximant turned into an alveolar nasal.

## \*/z/ and \*/D/

PAN *zalan path, made by a human as opposed to an animal; way or means to do something <span style="float: right;">↑</span>		
Form.		
Basai	tsatsan	path, road
Trobiawan	tsatsan	road
Kavalan	mani-razan	to take the same way
	nan-razan	same way of life
	razan	road, path, trail
	sa-razan	to pass through the bush
Saisiyat	raLan	road
Hoanya	dzalan	road
Sakizaya	daðan	path, road
Thao	saran	path, road
	mu-saran	go out onto the road; take out onto the road
Pazeh	darán	road, trail
Bunun	daan	road
	daan	road, path
Bunun (Isbukun)	daan	road
Kanakanabu	caánə	road, path
Tsou	cronə	road
Amis	lalan	path, road, street, highway; direction figuratively; means to do something
Siraya	darəŋ	road, path
Puyuma	daɭan	road
	mi-a-daɭan	spirits of the road
	p<in>u-daɭan-an	place where a road has been made
Paiwan	djalan	road, trail, path
	ta-djalan a tsautsau	same kind of people
	pu-djalan	to have a road; make a road
	ta-djalan	same, same kind of ("one road")
Favorlang/Babuza	tarran	road, path

As seen in the image, in most Formosan languages, \*/z/ became either /d/, /t͡s/, or /r/. In Protomalayopolynesian, it became \*/z/. The only evidence of palatalisation as far as I can see is in Paiwan \*/j/. Based on those outcomes, I would reconstruct that it was originally realised as \*[dz] with a slight tendency to go in the postalveolar direction.

In Blust's dictionary, there are only three words that contain \*/D/.

- \*/ikuD/
- \*/NataD/
- \*/SatəD/

Even in Protomalayopolynesian, there are only eight. Additionally, \*/z/ and \*/D/ have mutually exclusive distributions. Therefore, they're allophones, and the small distribution can be explained by the fact that \*/D/ is only one environment of a more widespread phoneme.

### PAN \*NataD outside ↑

#### Form.

Saisyat	latar	outside
Rukai	látəqə	outside
Rukai (Budai)	látəqə	outside
Rukai (Tanan)	latátəqə	outside
Rukai	latátəqə	outside
Proto-Rukai	*lataqə	outside

### PAN \*SateD accompany, escort; send, return ↑

#### Form.

Puyuma	ʔatəq	send, deliver
Paiwan	satjez	return something; object which is being returned

As seen in the images, in Formosan languages, \*/D/ became either /q/, /z/, or /r/. In most Western Malayopolynesian languages, it became either /d/ or /r/. From those outcomes, I find it hard to say if this phoneme was originally pronounced any differently in the \*/D/ environment, compared to the \*/z/ environment. So I would simply reconstruct them both as \*/d̪z/.

### \*/R/

Within the Formosan languages, this has a huge variety of outcomes, ranging from /r/ (Saaroa, Tsou) to /l/ (Amis, Bunun) to /g/ (Atayal) to /x/ (Pazeh). What's more is that all of those can be found in Malayopolynesian languages as well!

This phoneme was certainly a liquid of some kind. In order for /g/ and /x/ to be connected to the others, they would have had to be /R/ at some point. I would reconstruct the original phoneme in both Protoaustronesian and Protomalayopolynesian as \*/r/. \*/r/ to /l/ and \*/r/ to /R/ are both very commonly attested sound changes cross linguistically. However, \*/R/ to /r/ and \*/R/ to /l/ are not. And it couldn't have originally been \*/l/ because there is already a \*/l/.

**\*/r/**

This had much less variety in its outcomes than \*/R/, mostly sticking to alveolar liquids and the occasional /d/. I would reconstruct it as \*/r/.

### My Resulting Reconstructed Consonant Inventory

*/p/, */b/	*/t/, */d/	*/c/	*/k/, */g/	*/q/
*/m/	*/n/	*/ɲ/	*/ŋ/	
	*/r/			
	*/ɾ/			
	*/s/			*/h/
	*/ʃs/, */d͡z/			
*/w/	*/l/	*/j/		

### Phonology Conversion Matrix

Blust's	Mine
*/g/	*/g/
*/j/	
*/n/	*/n/
*/ñ/	*/ɲ/
*/N/	
*/S/	*/s/
*/s/	*/ʃs/
*/c/	
*/C/	*/c/
*/z/	*/d͡z/
*/D/	
*/R/	*/r/
*/r/	*/ɾ/
*/y/	*/j/

### Resources

**Blust's Dictionary**  
<https://acd.clld.org/>