

**WORKING PAPERS**

**IN**

**LINGUISTICS**

The notes and articles in this series are progress reports on work being carried on by students and faculty in the Department. Because these papers are not finished products, readers are asked not to cite from them without noting their preliminary nature. The authors welcome any comments and suggestions that readers might offer.

Volume 46(8)  
December  
2015

DEPARTMENT OF LINGUISTICS  
UNIVERSITY OF HAWAI'I AT MĀNOA  
HONOLULU 96822

An Equal Opportunity/Affirmative Action Institution

DEPARTMENT OF LINGUISTICS FACULTY  
2015

Victoria B. Anderson  
Andrea Berez-Kroeker  
Derek Bickerton (Emeritus)  
Robert A. Blust  
Lyle Campbell  
Kenneth W. Cook (Adjunct)  
Kamil Deen  
Patricia J. Donegan (Chair)  
Katie K. Drager  
Emanuel J. Drechsel (Adjunct)  
Michael L. Forman (Emeritus)  
Gary Holton  
Roderick A. Jacobs (Emeritus)  
William O’Grady  
Yuko Otsuka  
Ann Marie Peters (Emeritus)  
Kenneth L. Rehg (Adjunct)  
Lawrence A. Reid (Emeritus)  
Amy J. Schafer (Acting Graduate Chair)  
Albert J. Schütz, (Emeritus, Editor)  
Jacob Terrell  
James Woodward Jr. (Adjunct)

# TOWARD A PHONOLOGICAL RECONSTRUCTION OF PROTO-SULA<sup>1</sup>

TOBIAS BLOYD

**ABSTRACT.** This paper describes the primary dialect division in Sula, an under-documented language of eastern Indonesia. It uses the Comparative Method with new primary data to describe the protolanguage and its transformations, and in the process helps to narrow the regional literature gap.

Collins (1981) placed Sula within a Buru–Sula–Taliabo subgroup of Proto–West–Central Maluku within the Austronesian family. There have since been challenges to that subgrouping, but to my knowledge none of them have included new primary data, and the work of Collins remains the most thorough historical evaluation of these languages. The current paper does not consider higher branches of Austronesian, but it is intended to provide data with which higher-level subgroups can be measured and refined.

Collins (1982:82) suggested that there is one primary dialect division separating Mangon and Sanana, and my own fieldwork corroborates this interpretation.<sup>2</sup> However, his analysis is mostly limited to consonants, and the data is mainly from the Mangon dialect. He writes (1981:35) that our knowledge about the languages of Sula is in an elementary state, and that this limits our ability to analyze the reflexes of PAN vowels. Collins (1981:41) also notes that the Sanana dialect has “undergone a number of obfuscating innovations.” The present paper compares the Sanana and Mangon dialects in order to identify these “obfuscating innovations,” and thereby help to make sense of the vowel correspondences within the Sula language. These findings should help future research into the position of Sula within the Austronesian family by providing firmly established Proto-Sula reconstructions.

**1. LANGUAGE BACKGROUND AND PREVIOUS LITERATURE.** Sula is an under-documented Austronesian language of the Maluku Utara province of Eastern Indonesia. The language is spoken in the Sula Archipelago, which includes the islands of Sanana, Taliabo, and Mangon.<sup>3</sup> This archipelago has been frequently listed among the least studied regions in Indonesia (Collins 1981, 1982). Most of the work published on the language is built on wordlists collected by early explorers and missionaries (e.g., Wallace 1869, Holle c. 1900 via Stokhof 1980, Fortgens 1921), and the number of Sula dialects is not known.<sup>4</sup> James T. Collins (1981, 1982) concluded that there are likely only two main dialects of Sula, and my own fieldwork confirms this; consultants indicated only the one main dialect division, and that was borne out by lexical comparison. There are, however, most likely a number of definable but more subtle dialect divisions scattered throughout the region. The present paper describes the relationship between the dialects, henceforth referred to as Mangon, and Sanana.

The choice to use the names, *Mangon* and *Sanana*, is intended to limit ambiguity in this paper. Mangon is a Sula endonym for the Mangon tribe and one of the oldest villages on the island; the island is also known as Mangole/Mangoli. It is unclear if Sanana is an exonym or endonym. It refers to the primary administrative town in the region as well as the island it is situated on. That island is also known by the name Sula/Sua. However “Sula/Sua” is too ambiguous a term to use as a dialect designation, since it also refers to the entire island group as well as to the parent ethnic group of both dialects. So for this paper, and without any social or political bias:

---

<sup>1</sup> I would like to thank Robert Blust, Andrea Berez, Victoria Anderson, Gary Holton, Ken Rehg, and Albert Schütz for their help and comments—none of whom bear responsibility for any errors in this paper.

<sup>2</sup> Collins writes “Fagudu and Falahu” for what I refer to as Sanana dialect and “Mangoli” or “Sula” (generally) for what I call Mangon. The reason for my choice in vernacular is explained in section 1 below.

<sup>3</sup> These three islands are referred to by numerous alternate names and spellings. They are situated near the following Google Maps searchable coordinates: -2, 125.40.

<sup>4</sup> This is evident by Ethnologue’s multiple entries for Sula (ISO szn) and “Mangole” (ISO mqc) (Lewis 2015).

- Sula/Sua: the general name for the language, ethnic group, and island archipelago
- Mangon: the northeasternmost island; the dialect and tribe that settled this island
- Sanana: the southernmost island; the general dialect of the Falahu, Fagudu, Facei tribes that settled the southernmost island of the Sula archipelago

Sanana is the main town on the southern island. While it is not urban by objective measures, there are a few banks, and several local shops where staples can be purchased. There are no chains or franchises, nor recreational locations such as movie theaters or malls. The area does provide an urban infrastructure, including paved roads, and in many neighborhoods, intermittent electric power, and cellular coverage with limited Internet data availability. This is the most populous center within the region and also the area with the most conspicuous language attrition. Most public interactions are observed to be in Malay, and it is rare to hear youth in particular interacting in Sula even in more private domains. In rural communities, on the other hand, the use of Sula language is more vigorous; however the populations are smaller, and children still frequently tend to speak to one another in Malay; elderly language consultants often remark that children no longer speak the language well.

Transmission of Sula is almost entirely oral: Bahasa Indonesia (and informally, Moluccan Malay) is the medium of academic instruction, and the Sula language has neither a body of literature nor formal orthography. Political slogans and homemade signs are the main places written Sula is encountered. There is a Facebook group dedicated to the language, but conversational discourse within even that group is primarily in Malay, and the use of Sula tends to be limited to postings expressing support for the online group, expressions of nostalgia among members of the Sula diaspora, and answers to questions about how to say particular words.

The Sula ethnic group consists of four tribes—Falahu, Fagudu, Facei, and Mangon. According to local knowledge, the first three originally settled on the island of Sanana, while the fourth, Mangon, settled on Mangon island.<sup>5</sup> The physical separation of the islands is the likely reason for the Sanana–Mangon dialect division. Since original settlement, the Mangon tribe has formed two additional settlements on Sanana Island: a neighborhood in the region that has become the greater urban area of Sanana, and a village named Malbufa, on the northern part of the island’s west shore. The Malbufa dialect of Sula is significantly different from other Sanana dialect(s); however the Mangon village seems to have adopted a dialect that is of a typical Sanana variety.<sup>6</sup> Sanana tribes have also settled a number of villages around the islands, the oldest of these may be the Facei tribe villages of Capalulu, Wai U, and Orifola on the island of Mangon, where a contact dialect appears to have developed (Bloyd, in prep.). Additional Falahu and Fagudu villages were established on Taliabo Island and particularly along the southern coast of Mangon.<sup>7</sup>

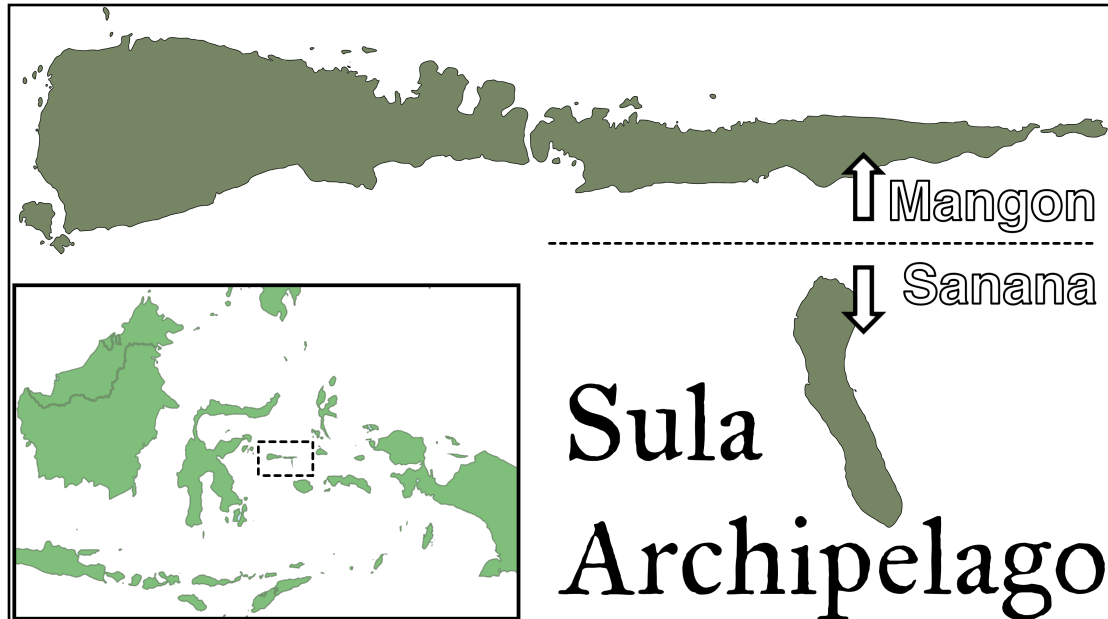
---

<sup>5</sup> I do not know of any published histories of the region; however the broad demographic and historical information I present here seems to be agreed upon by all Sula groups.

<sup>6</sup> My time spent in the village has given me the impression that the dialect is a form of Mangon that has undergone significant contact-induced dialect leveling.

<sup>7</sup> This was reported to me by a representative at the Sanana Badan Pemberdayaan Masyarakat dan Pemerintahan Desa (BPMPD) office.

MAP 1. Sula archipelago: The Mangon tribe settled the northeastern island, and the Falahu, Fagudu, and Facei tribes settled the southern island



Few publications analyze the Sula language directly. Holton 1996 lists eight entries for Sula, of which four contain primary lexical data: Collins 1976, 1981,<sup>8</sup> Holle c. 1900 via Stokhof 1980, and Wallace 1869. As a result of this lack of descriptive data, comparative linguistic work has been limited (e.g., Blust 1981, Collins 1983). However Fortgens (1921) described Soboyo on the neighboring island of Taliabo within the Sula archipelago, and Devin (1989) and Grimes (1992) described the main indigenous language of Buru island immediately to the south, which has a sizable Sula community in the village of Namlea.<sup>9</sup> Based on the material available at the time, Blust (1981) and Collins (1981) argued for an Austronesian subgroup that includes Buru, Sula, and Taliabo.

Blust (1981) used the Soboyo data in Fortgens 1921 to show that PAN \*S, which has disappeared in all other languages of eastern Indonesia for which comparative data is available, is reflected consistently as *h*, and to argue for the inclusion of Soboyo in a subgroup that contains Sula and Buru. Collins (1989) picked up the analysis of Taliabo and explained that field research showed the island to contain a single native language spreading over a long dialect continuum. The languages of the B-S-T subgroup were next taken up by Blust (1993), who disputed a claim in Esser 1938 proposing a subgroup containing Taliabo–Sula–Bacan, pointing out that Esser did not support the proposal with linguistically based reasoning.

Mark Donohue and Charles Grimes (2008) challenged the Central Eastern subgrouping favored by Blust, and the Blust (2009) follow-up defended the grouping. No new data source was cited either as basis for the challenge or the follow-up. The primary data and analyses contained within this paper will help to develop the academic literature in an area of Austronesian linguistics with significant debate and little data available.

**2. METHODOLOGY.** Data for this paper were gathered in North Maluku during two three-month trips. Language consultants were chosen as described below to help ensure that data would be reliable for comparative work. Elicitations were conducted at seventeen sites on Mangon and Sanana. Before selecting the sites, I consulted the regional development office, the bureau of statistics, and community elders; I inquired about the settlement history for each village to learn which tribe it was settled by and which dialect

<sup>8</sup> Collins 1976 is unpublished.

<sup>9</sup> It is unclear to what degree Sula language may still be spoken in this village.

it was reported to speak. A minimum of ten sessions were conducted for each targeted dialect. Sessions included a series of three videos that were watched by two speakers at a time within three age ranges as well as elicitation of a 230-word basic vocabulary list adapted from Greenhill et al.’s (2008) Austronesian Basic Vocabulary Database (see table 1). Typically, various other data collection and documentation sessions were also conducted. These ranged from ethnographic interviews to performance recordings; demonstrations of traditional medicine and farming/hunting/fishing methods; and topical vocabulary elicitation.

TABLE 1. Elicitation sessions by tribe

*Sessions conducted to determine whether there is tribe-based variation; this is an ongoing line of research. To control for geographic variation, research did not cross village boundaries. In addition to the four Sula tribes, the same elicitation materials were used with Sanana’s Bajo community, and also the Facei tribe community settled on Mangon’s southern coast (CMF).*

<b>Dialect</b>	<b>Video 1</b>			<b>Video 2</b>			<b>Video 3</b>			<b>Swadesh</b>
	18-29	30-49	50+	18-29	30-49	50+	18-29	30-49	50+	
<b>Falahu</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Fagudu</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Facei</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Mangon</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>CMF<sup>10</sup></b>		✓			✓			✓		✓
<b>Bajo</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

At least two speakers participated in vocabulary elicitation sessions. This helped to counteract effects of code switching, because all participants are fluent Malay speakers, and Malay was the language of elicitation. Sessions had three or more participants when group dynamics facilitated thoughtful responses. All sessions—vocabulary and video alike—were conducted in the target geographical area. For example, an elicitation session for Facei could not be conducted in a Fagudu village, even if the consultants were native Facei speakers, and the two main participants for each targeted elicitation session were required to have been born and raised in the particular village where the session was being conducted (i.e. they could not come from another village that reported to be of the same tribe or speak the same dialect). During vocabulary elicitation sessions, dialect-external participants were invited to join as tertiary observer-consultants when available. Their task was to listen and identify when unexpected or interesting dialect differences were encountered. These participants did not tend to contribute a large quantity of material; however their observations tended to be fascinating and relevant.

Three videos were created and used to elicit conversation controlled for topic. Videos were produced locally to (a) keep discussion from centering on any foreign elements rather than scenes and actions being depicted and (b) allow contextually prime participants to speak Sula rather than Malay which they typically hear accompanying media content depicting the outside world. The videos include a movie sans audio depicting a hungry young man earning money to buy a bread roll (Video 1) and two compilation videos consisting of various scenes filmed around the island (Video 2 and Video 3). Each video was presented in the same manner to two participants at a time. Videos were filmed and presented in HD 1080 on an iPad Retina.

Nine sessions were conducted for each target dialect. During a session, a video was watched and responded to two times by two participants who were: (a) born and raised in the community where the session took place, and (b) within a given age range (either 18–29, 30–49, 50+). It became clear that broad

<sup>10</sup> Central Mangon Facei is a dialect found on Mangon island’s southern coast that has undergone significant contact-induced leveling. See Bloyd (in prep) for more information.

age ranges were needed, because many participants were unsure of their chronological age.<sup>11</sup> To correct for age uncertainty, a reported age was recorded alongside an age estimate made by the researcher. A small number of elicitation sessions were excluded from the study and repeated with different participants. This happened if either a reported or observed age was outside the target range. The study included men and women; however it was impractical to balance for gender due to communal circumstances.

A video was watched by only one participant during the first viewing. The viewer described what was seen to the second participant, who would ask questions to clarify and gather additional information. Instructions were provided primarily in the Sula language. After the first viewing of video 1, the non-viewer recounted what was remembered from the story. The video was then played a second time when participants watched it together and discussed it freely. Videos were shown in this way to help collect conversational language data from which vocabulary was later extracted and added to Swadesh list terms for analysis with the Comparative Method.

### 3. FINDINGS.

**3.1 SANANA – MANGON CORRESPONDENCES.** Phoneme correspondences between Sanana and Mangon are given in table 2.

TABLE 2. Sanana–Mangon phoneme correspondences

Consonants:			Vowels:				
Sanana		Mangon	PSM	Sanana	Mangon	PSM	
p	–	p	*p	i	–	i	*i
b	–	b	*b	Ø	–	i	*i
t	–	t	*t	(word final)			
d	–	d	*d	u	–	u	*u
r	–	d	*d	a	–	u	*u
(intervocalic)				(if preceding syllable contains o)			
k	–	k	*k	Ø	–	u	*u
g	–	g	*g	(word final if preceding syllable			
ʔ	–	Ø	*ʔ	does not contain /o/)			
m	–	m	*m	e	–	e	*e
n	–	n	*n	o	–	o	*o
n	–	ŋ	*ŋ	a	–	a	*a
r	–	r	*r	ei	–	e	*ei
f	–	f	*f	oa	–	o	*ou
s	–	s	*s	ao	–	o	*ao
h	–	Ø	*h				
y	–	y	*y				
h	–	l	*l				
(subset 1)							
l	–	l	*l				
(subset 2)							
c	–	c	*c				
j	–	j	*j				
w	–	w	*w				

<sup>11</sup> The custom of keeping track of age is widespread only among the young.

**3.2 PSM PHONEME INVENTORY.** The phoneme inventory of Proto Sanana–Mangon is given in table 3.

TABLE 3. The phoneme inventory of Proto Sanana–Mangon

Consonants				Vowels		
p b	t d	k g	ʔ	i		
	c ([tʃ])	j ([dʒ])			u	
m	n	ŋ		e	o	
	r					
f	s		h	a		
	l					
w		y ([j])				

**3.2.1 CLUSTERS.** Prenasalized consonant clusters appear to be reflecting a fossil morpheme *N-*. It is uncertain whether they arose in Mangon after the dialect split (likely from the possessive marker *-in*), or if they reflect an earlier affix and were reduced in Sanana. If the morpheme was present in PSM, the following clusters would have been present word-initially: Ny, Nb, Nl, Nc, Np

VOWEL CLUSTERS (PHONEMIC STATUS UNCERTAIN):

ia, ui, ua  
ei, eu, oi, ou  
ai, au, ae, ao

**3.3 CHANGES IN PROTO SULA.** One important change must have occurred at a level higher than Proto Sanana–Mangon. I refer to this stage as Proto Sula.

**3.3.1 PAN \*uCi,u > oCi,u.** During fieldwork, an unexpected vowel correspondence was noticed between Sula and my contact language. Malay cognates containing uCu such as *kutu* ‘louse’, *bulu* ‘fur’, and *sepuluh* ‘ten’, corresponded to oCu in Proto Sula (oCu in Mangon; oCa in Sanana). This correspondence appeared to reflect a change from PAN \*u to Proto Sula \*o in syllables preceding \*u. Comparison to other Austronesian language data from the Austronesian Comparative Dictionary (Blust and Trussel, on-going), the Austronesian Basic Vocabulary Database (Greenhill et al. 2008), and pers. comm. with Robert Blust, revealed this correspondence to be a change to PAN \*u in the penult when the ultimate vowel is [+high].<sup>12</sup> That is, \*u became o in the penultimate syllable *only* where the final vowel was \*i or \*u. A uCV[+high] sequence does not violate the phonotactics of either dialect, and the presence of PSM form \*kuli ‘right’ (*kuli* (M), *kul* (S)) in modern Sula further implies that the change must have occurred at a higher level than PSM. The \*u > o change is demonstrated in the following examples (list in Appendix 1):

*uCi				
PMP	PSM	Mangon	Sanana	English
*bukij <sup>13</sup>	*fa-ʔoki	faoki	faʔok	‘forested mtn. areas’
*buni	*daʔufoŋi	daufoŋi	daʔufon	‘to hide’
*duRi	*loi	loi	hoi	‘thorn’/‘bone’

<sup>12</sup> This change appears to be connected to a subsequent and equally bizarre \*u > a change that occurred to the same lexical items in Sanana (discussed in section 3.5.2).

<sup>13</sup> This form is somewhat problematic, but it seems likely when we consider that the intermediate Proto Sula (PS) stage changed PAN \*b > f and added Sula’s environment prefix, *fa-*. The problem is that intervocalic *f* must then be reduced to a glottal. Thus: (PAN) \*bukij > (PS) \*fa-foki > (PSM) \*fa-ʔoki > (M) *faoki*, (S) *faʔok*. I cannot find other instances of \*f reducing to a glottal in Proto Sula, so change by analogy may be a better explanation for the form. That is, the \*f in *foki* became *faʔ* to match the series of trisyllabic environmental words beginning with *fa-* (*fasina* ‘moon’, *faoki* ‘forest’, and (*p*)*ajara* ‘cloud’).



<b>*uCu</b>				
*bubu	*fofu	fofu	fofa	‘bamboo fish/eel trap’
*buku	*foku	foku	foka	‘joint, e.g. finger’
*bulu	*fou	foʔ	foa	‘hair, feathers’

When we compare words like *ma-tua* ‘old’ (people) from \*tuqa, *uha* ‘shrimp, lobster’ from \*qudan, *uma* ‘house’ from \*Rumaq, *uya* ‘rain’ from \*quzan, and *fua* ‘fruit’ from \*buaq, we find that that \*u changed only before high vowels. Furthermore, we can see that the change did not affect other vowels in the same environment by comparing forms like *ga-pitu* ‘seven’ from \*pitu, *nih* ‘tooth’ from \*ŋisi, *nui* ‘coconut’ from \*niuR (< met.), *timu* ‘cucumber’ from \*qatimun, and *winu* ‘to drink’ from \*inum.

This particular change poses a problem to the fundamental Neogrammarian hypothesis: sound change is regular and without exception (Osthoff and Brugmann 1878). The Neogrammarian position is that sound change can be conditioned only by phonetic factors (Hock 1991); change is postulated to be an unconscious result of speakers slightly missing their targets when attempting to repeat a mental representation of an articulatory gesture. These repeated mistakes over time are said to cause modifications to the underlying mental representation of the sound (Blust n.d.); however, there is no identifiable biological reason why repeatedly missing a \*u target in the environment of /\_Ci,u would move a speaker’s prototypical mental representation of the segment closer to *o*. As it stands, the data at hand are most in keeping with the Neolinguists’ assertion that change is possible in the absence of a generalizable underlying phonetic motivation (Bartoli 1925). Blust (under review) further discusses Sula’s PAN \*u > o change in an expanded discussion on the theoretical implications of sound changes that are not phonetically conditioned and why they are *odd*.

**3.4 SOUND CHANGES FROM PSM TO SANANA.** This section identifies the changes that occurred between Proto Sanana–Mangon and the present-day Sanana dialect.

### 3.4.1 CONSONANTS.

**3.4.1.1 \*d>r /V\_V.** Intervocalic \*d became a trill in Sanana (usually produced as a flap surface allophone). This change must have occurred before Sanana’s final-vowel reduction, otherwise the Sanana form for ‘horn’ would be *tad*. Examples:

PSM	Sanana	English
*badagana	baragana	‘to dream’
*gad(i,e)ha	gareha	‘four’
*padomu	paroma	‘knee’
*tadu <sup>14</sup>	tar	‘horn’

**3.4.1.2 \*l>h/[V\_V], [#\_].** There are many instances of *l* in Sanana basic vocabulary; however PSM \*l became *h* in intervocalic and onset positions in the words in the following list. This change is apparently sporadic rather than regular; a regular sound change applies across the board. A sporadic sound change affects only a subset of candidate words in a language, and there is no identifiable regularity determining where a change of this sort will occur (Campbell 2004). Examples:

---

<sup>14</sup> Blust (pers. com. 2015) questions whether this is a Malay loan based on the fact that there are not any/many horned animals native to Maluku. To my knowledge, the anoa and Javan rusa (and possibly rhinoceros beetles) are the only horned animals native to Wallacea. Additionally, babirusa are native to the Sula islands, and although they are not true horned animals, all four incisors grow upward into long hornlike tusks. Deer, buffalo, and other horned animals have been introduced to Sula, but it is unclear at what date. However, since *tadu* (Mangon) – *tar* (Sanana) show expected sound correspondences, it is likely that a borrowing would have happened prior to the separation of Mangon and Sanana dialects, and thus \*tadu belongs in the PSM lexicon.

PSM	Sanana	English
*geli	gehi	‘to stand’
*lai	hai	‘earth/soil’
*lama	hama	‘eye’
*lani	-han	‘near’
*loi	hoi	‘bone’
*(N-)losa	hosa	‘leaf’

Collins (1981, 1982) speculates that PAN \*d/D became *h* in Sanana and *l* on Mangon. He provides the following evidence:

- PAN \*(dD)uRi ‘thorn’ > *loi*
- PAN \*dakep ‘embrace’ > *hakʔkotʔ*
- PAN \*DuSa ‘two’ > *guu*
- PAN \*ke(dD)ej ‘stand’ > *keli*
- PCM \*dama ‘eye’ > *lama*

I cannot account for all of Collins’ segment reconstructions; however the forms themselves are mostly consistent with my data for the Mangon dialect. One difference concerns the form, ‘embrace’, which he identifies as having a Fagudu origin; his use of ʔ represents “an unspecified devoiced vowel” (1981:42). I heard closures rather than devoiced vowels and recorded *hak* and *hak kot* ‘to hug’ for the Fagudu dialect.<sup>15</sup>

The form, *guu* is also problematic: Collins (1981:42) notes that loss of *l* from ɖ is regular in Sula but that there is an unexpected loss of final *-a* and that there is an unexpected appearance of *u* for *a* in the numeral prefix *ga-*. I reconstruct PSM \*gahu ‘two’. In the Mangon dialect, PSM \*h is deleted, resulting in \*gau. However, the form is still problematic, as \*\*au is then reduced to *u* (and lengthened for syllabic weight), yet *au* > *u* is not attested elsewhere.

**3.4.1.3 \*l – l.**<sup>16</sup> The /l/ phoneme was apparently retained in a number of forms. It is not clear that all additional forms I collected with an l–h correspondence reflect PAN \*d/D; however the forms with an l–l correspondence do appear to reflect PSM \*l.

One explanation for why some instances of \*l did not become *h*, could be that *l* was retained in intervocalic and onset position in early loans and forms that are very similar to Malay cognates/false cognates. These forms might have been ‘protected’ from the spread of \*l > *h*. That is, nearly all Sula speakers have been bilingual in Malay for many generations, and repeated production of forms with *l* while conversing in Malay could have halted the spread of \*l > *h* on these forms. This subset of vocabulary is demonstrated in the following examples:

PSM	Sanana	English
*bahali	bahal (cf. Malay <i>malu</i> )	‘shy, ashamed’
*bet pila	bet pila (cf. Malay <i>bila</i> )	‘when?’
*galima	galima (cf. Malay <i>lima</i> )	‘five’
*kalo	kalo (cf. Malay <i>kalau</i> )	‘if’
*koli	kol (cf. Malay <i>kulit</i> )	‘skin’
*la	la (cf. Malay <i>melayang</i> )	‘to float, fly’

<sup>15</sup> Fagudu dialect as spoken in Waibau village.

<sup>16</sup> This is not a sound change, but rather, a peculiar example of what appears to be the halting of a sound change across a particular stratum of words in a bilingual population due to the words’ similarity to cognates in the community’s other language.

*laka	laka (cf. regional Malay <i>laka</i> )	‘to walk’
*laŋi	lan (cf. Malay <i>langit</i> )	‘sky’
*lawa	lawa (cf. Malay <i>labah-labah</i> )	‘spider’
*lika	lika (cf. Malay <i>memilih</i> )	‘to choose’
*lima	lima (cf. Malay <i>lima</i> )	‘hand’
*lua	lua (cf. Malay <i>meluahkan</i> )	‘to vomit’
*manutelu	mantel (cf. Malay <i>telur</i> )	‘egg’
*nonu boli	nona bol (cf. Malay <i>goler</i> )	‘to lie down’
*sagila(ʔat)	sagilaʔat (cf. Malay <i>halilintar</i> )	‘lightning’
*tilu	til (cf. Malay <i>telinga</i> )	‘ear’

This similar-word hypothesis is not fully satisfactory though, because the following forms are not transparently similar to Malay. Examples:

PSM	Sanana	English
*balela	balela	‘to laugh’
*dalena	dalena	‘wide’
*gatelu	gatel	‘three’
*kila	kila	‘liver’
*lepa	lepa	‘above’
*lifi	lif	‘to turn’

**3.4.1.5 \*ŋ>n.** PSM \*ŋ became *n* in Sanana in all instances. Examples:

PSM	Sanana	English
*baifoŋi	baifon	‘to hide’
*daufoŋi	daʔufon	‘to hide’
*laŋi	lan	‘sky’
*maŋa	mana	‘sharp’
*maŋa pau	mana pau	‘beat, pound’
*meŋa	mena	‘to cry’
*(N-)caŋa	sana	‘branch’
*naŋu	nan	‘to swim’
*ŋa	na	‘name’
*ŋapu	nap	‘head’
*ŋihi	nihi	‘tooth’
*saŋa-petu	sanapet	‘thatch/roof’
*yaŋa	yana	‘to see’

### 3.4.2 VOWELS.

**3.4.2.1 \*u>a/o<sub>σ</sub> #.** Lowering of \*u to a where the preceding syllable contains a mid back vowel is a particularly interesting sound change which has no apparent phonetic motivation—similar to the lowering of PAN \*u before high vowels described in the previous section. In an o<sub>σ</sub>u environment, the tongue raises between the first and second vowel. A change from PSM \*o<sub>σ</sub>u > o<sub>σ</sub>a requires the tongue to lower between the first and second vowel, so this change not only has no obvious phonetic motivation, it appears to behave in an opposite manner to phonetic expectation and even more soundly corroborates the Neolinguists’ assertion that change is possible in the absence of a generalizable underlying phonetic motivation. As mentioned above, see Blust (under review) for an expanded discussion on the theoretical implications of sound changes that are not phonetically conditioned. This sound change had to occur before final high

vowel deletion, otherwise forms like ‘to suck’, ‘to burn’, and ‘to hit’ would be *bos*, *don*, and *dot* instead of *bosa*, *dona*, and *dota*. Examples:

PSM	Sanana	English
*bagou	bagoa	‘cold’
*bosu	bosa	‘to suck’
*donu	dona	‘to burn’
*dotu	dota	‘to hit’

### 3.4.2.2 \*i,u>Ø/ [-cont] \_# [-glottal]

Sanana deleted word final high vowels following non-glottal consonants.<sup>17</sup> They remain following vowels and glottal consonants (e.g., *tui* ‘snake’, *yau* ‘far’, *behi* ‘to throw’, *gahu* ‘two’). They may also remain following affricate consonants; however there are too few examples to state this conclusively (e.g., *gaji* ‘grease’, which could be a Malay loan). Examples:

PSM	Sanana	English
*(t,d)ufi	duf	‘to stab’
*aku	ak	‘1SG’
*api	ap	‘fire’
*asu	as	‘dog’
*bafei	bafei	‘green’
*bagu	bag	‘thick’
*bahali	bahal	‘shy’
*baifonji	baifon	‘to hide’
*bamapu	bamap	‘to cook’

### 3.4.3 NON-GENERALIZABLE.

1. Proto \*c corresponds to /s/ in ‘branch’, but \*c is so infrequent that this form is almost certainly a loan. Interestingly, the Malay form *cabang* ‘branch’ also contains the affricate; however these forms are not transparently cognate or even similar.

PSM \*(N-)caŋa>sana ‘branch’

2. -y- epenthesis occurred in ‘mouth’

PSM \*baØoni>baØon>bayon ‘mouth’

3. -l- epenthesis occurred in ‘to steal’. The *l* could not have been in PSM, otherwise it would still be present in Mangon.

PSM \*biØnaka>bilnaka ‘to steal’

4. The numeral prefix *ga-* was dropped in the morpheme for ‘one’ and all ordinal derivatives. It is retained however in the lexical item, *gahia* ‘alone’ (Mangon deleted the *h* and reduced *gaia* to *gØia*). No analogous forms are present.

PSM \*fatu-**ga**-hia>fat-**ga**-hia>fat-Ø-hia ‘one unit’

PSM \***ga**-hia>Ø-hia ‘one’

PSM \*ca-**ga**-hia>ca-Ø-hia ‘one thousand’

5. Proto \*g was deleted in the following two forms. It does exist in onset position elsewhere in the language, however.

PSM \*gami>gam>Øam ‘to squeeze’

<sup>17</sup> Forms like *tapa* ‘left’ and *laka* ‘walk’ demonstrate that this change occurred only to high vowels (*i* and *u*).

PSM \*gifu>gif>Øif ‘to uncover’

6. The consonant cluster \*mf was reduced to /f/ in ‘to swell’. No analogous forms are present.

PSM \*mamVfai>mafai ‘to swell’

7. Final ŋ was deleted in the Dutch loan meaning ‘to count’. This occurred prior to final vowel deletion.

PSM \*rekiŋ (L)>rekiØ>rekØØ ‘to count’

8. PSM \*s either became *h* in the Sanana form for ‘flesh’ or \*h became *s* in Mangon. The first option is more likely.

Option 1: PSM \*(N-)isi>(Sanana)ihi ‘meat/flesh’

Option 2: PSM \*(N-)ihi>(Mangon)nisi ‘meat/flesh’

9. Proto \*y became a glottal in the form meaning ‘heavy’ (the reverse is not possible, because the Mangon form would have deleted the segment if the ancestral form had contained a glottal stop).

PSM \*fayata>faʔata ‘heavy’

**3.5 SOUND CHANGES FROM PSM TO MANGON.** This section identifies the changes that occurred between Proto Sanana–Mangon and the present-day Mangon dialect.

### 3.5.1 CONSONANTS.

**3.5.1.1 \*h>Ø.** With few exceptions, PSM \*h was lost in Mangon in all positions. The following forms demonstrate \*h deletion. Examples:<sup>18</sup>

PSM	Sanana	English
*baha	baɪ	‘to buy’
*bahali	bali	‘shy’
*behi	beɪ	‘to throw’
*gad(i,e)ha	gadia	‘four’
*gahu	guɪ	‘two’
*gatahua	gatua	‘eight’
*han	an	‘who?’
*kahiku	kiku	‘grass’
*mahi	mai	‘sea’
*nahu	nau	‘long’
*ŋihi	ŋiɪ	‘tooth’
*pougahu	pogu(ɪ)	‘twenty’
*sahafa	fafa	‘rat’
*samohu	samo	‘needle’
*tahaga <sup>19</sup>	taga	‘lake’
*tahun	taun	‘year’

**Exceptions:** A few counter examples have been recorded in which forms are produced optionally with *h*. All are high frequency words, and all have variants without *h*. The forms could either be back borrowings from Sanana or examples of \*h retention. Examples:

<sup>18</sup> See numbers 2 and 4 in section 3.5.3 Non-generalizable for discussions about the Sanana forms *kiku* and *fafa*.

<sup>19</sup> This is a Malay loan word, but *taga* – *tahaga* show expected sound correspondences, so it is likely that the borrowing occurred prior to the separation of Mangon and Sanana dialects.

PSM	Sanana	English
*dahi	dahi	‘true’
*deha	deha	‘be located’
*hapa	hapa	‘what’

**3.5.1.2 \*ʔ > Ø.** Surface glottal stops are produced in Mangon (e.g., [faʔ.ko] ‘dog’, [saʔ.ka.fi] ‘to hold, as a baby’); however these do not appear to be phonemic: no minimal pairs have been found, and hypothetical pronunciations with and without surface glottal stops have been judged equally correct. The following forms demonstrate \*ʔ deletion. Examples:

PSM	Sanana	English
*daʔufoŋi	daufonji	‘to hide’
*fa-ʔoki	faoki	‘forest’
*kiʔi	ki:	‘3SG’
*maʔana	mana	‘man’

**3.5.1.3 /N-/ PREFIX.** Many Mangon forms have prenasalized word-initial onsets. It is common in Sula to frame nouns in isolation using the morpheme *in*.<sup>20</sup> It is unclear whether the nasal segment in Mangon’s prenasal clusters are reflecting a fossilized morpheme that was lost in Sanana, or whether they arose in Mangon subsequent to the dialect split (likely derived from the *in* morpheme). The following forms demonstrate word initial prenasalized clusters present in Mangon but not in Sanana. Examples:

PSM	Sanana	English
*(N-)caŋa	ncanja	‘branch’
*(N-)losa	nlosa	‘leaf’
*(N-)yai	nyai	‘leg/foot’
*(N-)boyu	mboyu	‘tail’

There are several exceptional forms as well: *manpani* ‘wing’ contains a prenasalized cluster internally. The form ‘wing’ is also produced as, *npani*, and *manpani* is likely a contraction of *manu-npani* ‘bird wing’. The form *fantui* ‘star’, which also contains an internal prenasalized cluster, belongs to a class of environment nouns that all include a non-analyzable morpheme, *fa-/pa-*. Other nouns in this set are *fasina* ‘moon’, *faoki* ‘forest’, and *(p,f)anara* ‘cloud’. Evidence that the nasal segment in these clusters derives from a morpheme is seen in the fact that it also occurs on vocalic syllables such as *nisi* ‘meat’.

**3.5.2 VOWELS.** Many vowel clusters are reduced in Mangon; however it is difficult to find a single phonetic condition to account for all of the changes. In instances of \*V<sub>1</sub>V<sub>1</sub>>V<sub>1</sub>, it is unclear whether the first or second segment was deleted. With \*ei>e and \*ou>o, the second vowel is deleted, and it is a higher vowel than the first, but in \*ao>o, the first vowel is deleted, and it is a lower, less back vowel.

**3.5.2.1 \*V<sub>1</sub>V<sub>1</sub>>V<sub>1</sub>.** Sequences of like vowels are reduced to a single vowel with length determined by a minimal word requirement. Examples:

PSM	Sanana	English
*baha>baa>	ba:	‘to buy’
*bahali>baali>	bali	‘shy’
*ki’i>kii>	ki:	‘3SG’
*maʔana>maana>	mana	‘man’
*ŋihi>ŋii>	ŋi:	‘tooth’

<sup>20</sup> *in* is used to indicate possession, and it also frequently serves an expletive-like function.

*sahafa>saafa>	fafa	‘rat’
*tahaga>tahaga>	taga	‘lake’

**3.5.2.2 \*ei>e.** Sequences of \*ei were reduced to e. That is the vowel that was retained is lower, and it is the first vowel in the sequence. Examples:

PSM	Sanana	English
*apfei	apfe(ɪ)	‘smoke’ (fire)
*bafei	bafe(ɪ)	‘smoke, fog’ (general)
*behi>bei>	beɪ	‘to throw’
*ganei	gane(ɪ)	‘six’

**3.5.2.3 \*ao, \*ou>o.** In sequences of \*ao, a was deleted: that is, the vowel that was retained is higher and it is the second in the sequence. In \*ou sequences, u was deleted: that is, the retained vowel is lower and it is the first in the sequence. Examples:

PSM	Sanana	English
*bagou	bagoɪ	‘cold’
*baoni	boni	‘mouth’
*fou	foɪ	‘hair, feather’
*pou	poɪ	‘blood’
*pou	poɪ	‘ten’
*pougahu	poguɪ	‘twenty’
*pougalima	pogalima	‘fifty’
*samohu>samou>	samo	‘needle’
*saotu	sotu	‘dry’

### 3.5.3 NON-GENERALIZABLE.

- \*au reduced to /u/ in forms containing the number ‘two’  
 PSM \*gahu>gaŋu>guɪ ‘two’  
 PSM \*gatahua>gataŋua>gatua ‘eight’ (from ‘minus two’)  
 PSM \*pougahu>pougaŋu>poguɪ ‘twenty’
- \*ai reduced to /i/ in ‘grass’. The /ai/ sequence is only found in final position in other forms.  
 PSM \*kahiku>kaŋiku>kiku ‘grass’
- l- epenthesis; and \*ai to /a/ reduction (note: word final ai is permitted)  
 PSM \*baifoŋi>baŋofoŋi>balfoji ‘to hide’
- \*s became /f/ in ‘rat’<sup>21</sup>  
 PSM \*sahafa>saŋafa>saŋofofa>fafa ‘rat’

**3.6 CONCLUSION.** Sula is a language with a clear primary division between the Sanana and Mangon dialects, and this paper describes the correspondences and sound changes responsible for Sula’s synchronic variation. The data in this paper will aid in refining Sula’s position within the Austronesian family. In Proto Sula (before the Sanana–Mangon split), PAN \*u was lowered in words where the following syllable contained a high vowel. This is a sound change that challenges the neogrammarian assertion that all sound change must be phonetically motivated.

<sup>21</sup> It is not likely that the proto form was \*safafa, because *sahafa*, *saafa*, and *safa* all occur in Sanana, where intervocalic *h* deletion appears to have begun but is still optional and limited mostly to rapid, non-deliberate speech.

In the Sanana dialect, (a) \*d became an alveolar trill, (b) \*l became *h* in intervocalic and onset position in native vocabulary, but it appears that a subset of the lexicon with surface similarity to Malay was shielded from this change, (c) \*ŋ became *n* in Sanana in all environments, (d) peculiarly, *u* was lowered to *a* where the preceding syllable contains a mid back vowel—a change that not only challenges the neogrammarian assertion that all sound change must be phonetically motivated but seems to display the opposite of phonetic motivation, (e) word-final high vowels were deleted following non glottal consonants but not following vowels and glottals.

In the Mangon dialect, (a) \*h was lost in Mangon in all positions (with some exceptions), (b) \*ʔ was deleted, (c) initial nasals, which might reflect an early expletive (possibly genitive) marker, were fused to many nouns. These often resulted in prenasalized consonant clusters (in a dialect that avoids clusters otherwise). (d) Sequences of like vowels were reduced to long vowels, (e) \*ei was reduced to *e*, and (f) \*ou and \*ao were reduced to *o*.

The study of the Sula language is still very much in its infancy; however this paper builds on the groundbreaking work begun by Robert Blust and James T. Collins three and a half decades ago and narrows the gap in the academic literature. It provides a more complete picture of the inner structure of Sula—especially with regard to the confusing vowel correspondences, and it provides data from which hypotheses about higher branches of the Austronesian language can be evaluated.

## APPENDIX 1. Section 3.3.1. examples

<b>*uCi</b>				
<b>PMP</b>	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
*bukij	*fa-ʔoki	faoki	faʔok	‘forested mountain areas’
*buni	*daʔufonji	daufonji	daʔufon	‘to hide’
*duRi	*loi	loi	hoi	‘thorn’/‘bone’
*kulit	*koli	koli	kol	‘skin’
*ma-putiq	*boti	boti	bot	‘white’
*puki	*poki	poki	pok	‘vulva, vagina’
*qutin	*oti	oti	ot	‘penis’
<b>*uCu</b>				
<b>PMP</b>	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
*bubu	*fofu	fofu	fofa	‘bamboo fish/eel trap’
*buku	*foku	foku	foka	‘joint, finger or bamboo’
*bulu	*fou	fo:	foa	‘hair, feathers’
*kutu	*kotu	kotu	kota	‘hair louse’
*puluq	*pou	po:	poa	‘ten’
*pusuq	*pou	po:	poa	‘banana inflorescence’
*susu	*sosu	sosu	sosa	‘female breast’
*tunu	*donu	donu	dona	‘to burn’
*tuktuk	*dotu	dotu	dota	‘to strike’



APPENDIX 2. Section 3.4.2.1 examples

**\*u>a/o<sub>σ</sub>\_#**. Lowering of \*u to a where the preceding syllable contains a mid back vowel

<b>PSM</b>	<b>Sanana</b>	<b>English</b>
*bagou	bagoa	‘cold’
*bosu	bosa	‘to suck’
*donu	dona	‘to burn’
*dotu	dota	‘to hit’
*fou	foa	‘hair, feather’
*kotu	kota	‘louse’
*momu	moma	‘to hit’
*moru	mora	‘wind’
*moyu	moya	‘no, not’
*(N-)boyu	boya	‘tail’
*nonu	nona	‘to sleep’
*padomu	paroma	‘knee’
*pou	poa	‘blood’
*pou	poa	‘ten’
*pougahu	poagahu	‘twenty’
*pougalima	poagalima	‘fifty’
*samohu	samoha	‘needle’
*saotu	saota	‘to dry’
*sosu	sosa	‘breast’
*yotu	yota	‘to hunt’

APPENDIX 3. Section 3.4.2.2 examples

**\*i,u > Ø / [-cont] \_#**  
**[+glottal]**

<b>PSM</b>	<b>Sanana</b>	<b>English</b>
*(t,d)ufi	duf	‘to stab’
*aku	ak	‘1SG’
*api	ap	‘fire’
*asu	as	‘dog’
*bafei	bafei	‘green’
*bagu	bag	‘thick’
*bahali	bahal	‘shy’
*baifonji	baifon	‘to hide’
*bamapu	bamap	‘to cook’
*banapi	banap	‘to shoot’
*baoni	bayon	‘mouth’
*betu	bet	‘day’
*boti	bot	‘white’

*daeti	daet	‘branch’
*dagati	dagat	‘narrow’
*daufoŋi	daʔufon	‘to hide’
*deti	det	‘to cut, hack’
*dogi	dog	‘to grow’
*duki	duk	‘to come’
*eki	ek	‘neck’
*gami	am	‘to squeeze’
*gapitu	gapit	‘seven’
*gasi	gas	‘salt’
*gatelu	gatel	‘three’
*jubi	jub	‘to shoot’
*kabaresi	kabares	‘bad, evil’
*kafini	kafin	‘mosquito’
*kagi	kag	‘to fear’
*kahiku	kahik	‘grass’
*kiti	kit	‘1PL.INCL’
*koli	kol	‘skin’
*kuli	kul	‘right’
*lani	-han	‘near’
*laŋi	lan	‘sky’
*lifi	lif	‘to turn’
*maki	mak	‘tongue’
*manipi	manip	‘thin’
*miti	mit	‘black’
*naŋu	nan	‘to swim’
*nibu	nib	‘to sit’
*nonu boli	nona bol	‘to lie down’
*ŋapu	nap	‘head’
*rekiŋ (L)	rek	‘to count’
*saku	sak	‘to pierce’
*samamu	samam	‘to chew’
*saŋa-petu	sanapet	‘thatch/roof’
*waki dabu	wak dab	‘to think’
*winu	win	‘to drink’
*yopu	yop	‘to suck’

Appendix 4. Words: PSM – MANGON – SANANA

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
1	*afu-	afumai	aftuka	‘ash’
2	*aku	aku	ak	‘I’
3	*api	api	ap	‘fire’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
4	*asu	asu	as	‘dog’
5	*baba	baba	baba	‘father’
6	*badagana	badagana	baragana	‘to dream’
7	*bagou	bago(ɔ)	bagoa	‘cold’
8	*bagu	bagu	bag	‘thick’
9	*baha	baɔ	baha	‘to buy’
10	*bahali	bali	bahal	‘shy’
11	*bama	bama	bama	‘to split’
12	*bamapu	bamapu	bamap	‘to cook’
13	*banapi	banapi	banap	‘to shoot’
14	*baoni	boni	bayon (see 3.4.3)	‘mouth’
15	*basa	basa	basa	‘other’
16	*baumata	baumata	baumata	‘to kill’
17	*behi	beɔ	behi	‘to throw’
18	*bena	bena	bena	‘to climb’
19	*betu pila	betu pila	bet pila	‘when?’
20	*betu	betu	bet	‘day’
21	*beu	beu	beu	‘to tie up’
22	*binaka	binaka	bilnaka (see 3.4.3)	‘to steal’
23	*bisa	bisa	bisa	‘good’
24	*bo	bo	bo	‘at’
25	*bosu	bosu	bosa	‘to suck’
26	*boti	boti	bot	‘white’
27	*bua	bua	bua	‘to fall’
28	*ca-gia	ca-gia	ca-hia	‘one thousand’
29	*daeti	badaeti	daet	‘branch’
30	*dagati	dagati	dagat	‘narrow’
31	*dalena	dalena	dalena	‘wide’
32	*daʔufoŋi	daufonji	daʔufon	‘to hide’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
33	*deti	deti	det	‘to cut, hack’
34	*dogi	dogi	dog	‘to grow’
35	*donu	donu	dona	‘to burn’
36	*dotu	dotu	dota	‘to pound, beat’
37	*duki	duki	duk	‘to come’
38	*eki	eki	ek	‘neck’
39	*(fa-N-)tui	fantui	fatui	‘star’
40	*fa-ŋara	paŋara	faŋara	‘cloud’
41	*fa-sina	fasina	fasina	‘moon’
42	*fa-ʔoki	faoki	faʔok	‘woods, forest’
43	*fata	fata	fata	‘wife’
44	*fatu	fatu	fat	‘stone’
45	*fatugia	fatugia	fathia	‘piece of~’
46	*feu	feu	feu	‘new’
47	*fina	fina	fina	‘woman’
48	*fofu	fofu	fofa	‘bamboo trap for fish and eels’
49	*foku	foku	foka	‘node in bamboo, sugarcane’
50	*fou	foʔ	foa	‘hair, feather’
51	*gad(i,e)ha	gadia	gareha	‘four’
52	*gahu	guʔ (see 3.5.3)	gahu	‘two’
53	*galima	galima	galima	‘five’
54	*ganei	gane(ʔ)	ganei	‘six’
55	*gapitu	gapitu	gapit	‘seven’
56	*gasi	gasi	gas	‘salt’
57	*gatahua	gatua	gatahua	‘eight’
58	*gatasia	gatasia	gatasia	‘nine’
59	*gatelu	gatelu	gatel	‘three’
60	*geka	geka	geka	‘painful, sick’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
61	*geli	geli	gehi	‘to stand’
62	*gia (see 3.4.3)	gia	hia	‘one’
63	*han	an	han	‘who?’
64	*hapa	apa, hapa	hapa	‘what?’
65	*ik(i,a)	ika	ik(i)	‘this’
66	*jubi	jubi	jub	‘bow’
67	*kabaresi (L?)	kabaresi	kabares	‘bad, evil’
68	*kafini	kafini	kafin	‘mosquito’
69	*kagi	kagi	kag	‘to fear’
70	*kalo	kalo	kalo	‘if’
71	*kam	kam	kam	‘we (excl)’
72	*kau	kau	kau	‘to cut, hack’
73	*kena	kena	kena	‘fish’
74	*kim	kim	kim	‘you’
75	*kiti	kiti	kit	‘we (incl)’
76	*kiʔi	kiː	kiʔi	‘he/she’
77	*koli	koli	kol	‘skin’
78	*kotu	kotu	kota	‘louse’
79	*kuli	kuli	kul	‘right’
80	*la	laː	la	‘to fly’
81	*lai	lai	hai	‘earth, soil’
82	*lai mai	lai mai	hai mai	‘dust’
83	*laka	laka	laka	‘to walk’
84	*lama	lama	hama	‘eye’
85	*lani	lani	han	‘near’
86	*laŋi	laŋi	lan	‘sky’
87	*lawa	lawa	lawa	‘spider’
88	*lepa	lepa	lepa	‘above’
89	*lifi	lifi	lif	‘to turn’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
90	*lika	lika	lika	‘to choose’
91	*lima	lima	lima	‘hand’
92	*loi	loi	hoi	‘thorn, bone’
93	*mahi	mai	mahi	‘sea’
94	*maki	maki	mak	‘tongue’
95	*manipi	manipi	manip	‘thin’
96	*manu	manu	man	‘bird’
97	*manutelu	manutelu	mantel	‘egg’
98	*maŋa	maŋa	mana	‘sharp’
99	*mata	mata	mata	‘dead’
100	*mata-pia	matapia	matapia	‘human’
101	*ma-tua	matua	matua	‘old’
102	*maʔana	mana	maʔana	‘man’
103	*meŋa	meŋa	mena	‘to cry’
104	*meta	meta	meta	‘wet’
105	*meu	meu	meu	‘rope’
106	*mia	mia	mia	‘red’
107	*miti	miti	mit	‘black’
108	*momu	momu	moma	‘to pound, beat’
109	*mon	mon	mon	‘you’
110	*moru	moru	mora	‘wind’
111	*moyu	moyu	moya	‘no, not’
112	*muamua	muamua	muamua	‘all’
113	*(N-)boyu	mboyu	boya	‘tail’
114	*(N-)losa	nlosa	hosa	‘leaf’
115	*(N-)yai	nyai	yai	‘leg’
116	*nahu	nau	nahu	‘long’
117	*naŋu	naŋu	nan	‘to swim’
118	*nau	nau	nau	‘to know’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
119	*nibu	nibu	nib	‘to sit’
120	*nonu	nonu	nona	‘to sleep’
121	*ŋa	ŋa	na	‘name’
122	*ŋapu	ŋapu	nap	‘head’
123	*ŋau	ŋau	nau	‘cat’
124	*ŋihi	ŋiː	nihi	‘tooth’
125	*padomu	padomu	paroma	‘knee’
126	*pia	pia	pia	‘alive, good, safe’
127	*pou	poː	poa	‘blood’
128	*pou	poː	poa	‘ten’
129	*pougahu	pogu(:)	poagahu	‘twenty’
130	*rekiŋ (L)	rekiŋ	rek (see 3.4.3)	‘to count’
131	*saku	saku	sak	‘to stab’
132	*samamu	samamu	samam	‘to chew’
133	*samohu	samo	samoha	‘needle’
134	*sanjapetu	sanjapetu	sanapet	‘thatch/roof’
135	*saotu	sotu	saota	‘dry’
136	*soba	soba	soba	‘wing’
137	*sosu	sosu	sosa	‘breast’
138	*tadu	tadu	tar	‘horn’
139	*tahaga (L)	taga	tahaga	‘lake’
140	*tahun	taun	tahun, taun	‘year’
141	*tapa	tapa	tapa	‘left’
142	*tilu	tilu	til	‘ear’
143	*timu	timu	tim	‘cucumber’
144	*tua	tua	tua	‘husband’
145	*tui	tui	tui	‘snake’
146	*tuka	tuka	tuka	‘intestines’
147	*uha	ua	uha	‘shrimp, lobster’

	<b>PSM</b>	<b>Mangon</b>	<b>Sanana</b>	<b>English</b>
148	*uma	uma	uma	‘house’
149	*uya	uya	uya	‘rain’
150	*wai	wai	wai	‘water’
151	*(wai)ewa	(wai)ewa	(wai)ewa	‘to flow’
152	*waka	waka	waka	‘root’
153	*waki dabu	waki dabu	wak dab	‘to think’
154	*wama	wama	wama	‘to breathe’
155	*winu	winu	win	‘to drink’
156	*yaŋa	yaŋa	yana	‘to see’
157	*yau	yau	yau	‘far’
158	*yotu	yotu	yota	‘to hunt’

## REFERENCES

- BARTOLI, MATTEO. 1925. *Introduzione Alla Neolinguistica (principi--scopi--metodi)*. Genève: L. S. Olschki.
- BLOYD, TOBIAS. In prep. Facei settlements on Mangon: Dialect leveling and koineization in Sula.
- BLUST, ROBERT. n.d. “Pysiological” theories of sound change. LING-645. Honolulu, University of Hawai‘i at Mānoa.
- BLUST, ROBERT. 1981. The Soboyo reflexes of Proto-Austronesian \*S. In *Historical Linguistics in Indonesia*, part 1, ed. by Robert A. Blust, 21–30. NUSA 10. Jakarta: Lembaga Bahasa, Universitas Katolik Indonesia Atma Jaya.
- BLUST, ROBERT. 1993. Central and Central-Eastern Malayo-Polynesian, *Oceanic Linguistics*. 32(2):241–93. Honolulu: University of Hawai‘i Press.
- BLUST, ROBERT. 2009. The position of the languages of Eastern Indonesia: A reply to Donohue and Grimes. *Oceanic Linguistics*. 48(1):36–77. Honolulu: University of Hawai‘i Press.
- BLUST, ROBERT. Under review. Odd conditions: Context-sensitive sound changes in unexpected contexts. *Journal of Historical Linguistics*.
- BLUST, ROBERT, AND STEPHEN TRUSSEL. Ongoing. Austronesian comparative dictionary. ([www.trussel2/acd.com](http://www.trussel2/acd.com))
- CAMPBELL, LYLE. 2004. *Historical linguistics: An introduction*. 1st ed. Cambridge, Mass.: MIT Press.
- COLLINS, JAMES T. 1976. Field notes in languages of Indonesia. Unpublished MS.
- COLLINS, JAMES T. 1981. Preliminary notes on Proto-West Central Maluku: Buru, Sula, Taliabo and Ambelau. In *Historical Linguistics in Indonesia*, part 1. ed. Robert A. Blust. 31–45. Jakarta: Badan Penyelenggara Seri Nusa, Universitas Atma Jaya.
- COLLINS, JAMES T. 1982. Linguistic research in Maluku: A report of recent field work. *Oceanic Linguistics* 21(1/2): 73–146. Web.



- COLLINS, JAMES T. 1983. *The historical relationships of the languages of Central Maluku, Indonesia*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, the Australian National University. Pacific Linguistics. Ser. D ; No. 47.
- COLLINS, JAMES T. 1989. Notes on the language of Taliabo. *Oceanic Linguistics* 28(1): 75–95. Web.
- DEVIN, CHAUMONT. 1989. *A Buru-English dictionary*. Honolulu.
- DONOHUE, MARK, AND CHARLES E. GRIMES. 2008. Yet more on the position of the languages of Eastern Indonesia and East Timor. *Oceanic Linguistics* 47(1):114–58.
- ESSER, S. J. 1938. Talen. *Atlas van tropisch Nederland*. (sheet 9b). Amsterdam: Koninklijk Nederlands Aardrijkskundig Genootschap.
- FORTGENS, J. 1921. *Bijdrage tot de Kennis van het Sobojo (Eiland, Taliabo, Soela-Groep)*. M. Nijhoff. Web.
- GREENHILL, SIMON J.; ROBERT BLUST; AND RUSSELL GRAY. 2008. The Austronesian basic vocabulary database: From bioinformatics to lexomics. *Evolutionary Bioinformatics* 4:271–83. Web.
- GRIMES, CHARLES E. 1992. The Buru Language of Eastern Indonesia. Australian National University PhD dissertation.
- HOCK, HANS HENRICH. 1991. *Principles of historical linguistics*. 2nd rev. and updated ed. Berlin; New York: Mouton De Gruyter.
- HOLTON, GARY. 1996. Bibliography of language and language use in North and Central Maluku. Honolulu: Center for Southeast Asian Studies, School of Hawaiian, Asian, and Pacific Studies, University of Hawai'i at Mānoa. Southeast Asia Paper 40. Web.
- LEWIS, M. PAUL; GARY F. SIMONS; AND CHARLES D. FENNIG (eds.). 2015. *Ethnologue: Languages of the world*, 18<sup>th</sup> edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- OSTHOFF, HERMANN, AND KARL BRUGMANN. 1878. Vorwort. In *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, vol. 1, Leipzig: Hirzel, p. iii–xx.
- STOKHOF, W.A.L.; LIA SALEH-BRONKHORST; AND K. F. HOLLE. 1980. *Holle lists: Vocabularies in languages of Indonesia*. Vol. 2. *Sula and Bacan Islands, North Halmahera, South and East Halmahera*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
- WALLACE, ALFRED RUSSEL. 1869. *The Malay archipelago: The land of the orang-utan, and the bird of paradise. A narrative of travel, with studies of man and nature*. New York: Harper & Brothers.

bloyd@hawaii.edu