

Journal of the Linguistic Society of Papua New Guinea

ISSN: 0023-1959

Vol. 35, 2017

Skou Languages Near Sissano Lagoon, Papua New Guinea

S. A. Miller (SIL)

Abstract

To date, the speech varieties spoken in the villages of Barapu, Pou, Ramo, and Sumo have been classified together as the language of Warapu (ISO 639-3 code [wra]). Varied evidence presented in this paper indicates that at least three distinct languages are spoken among these villages. These languages are Bauni, Uni, and Bouni. The evidence presented includes reported intelligibility, cognate percentages, and corroborating research. Some information will also be given about Bobe, the language once spoken in the village of Nouri, as well as a note about a previously undocumented language, Bounino.

"Warapu" is currently known to the world as the name of a language on the north coast of Papua New Guinea, on the shores of Sissano Lagoon, in Sandaun Province. It is listed in *Ethnologue* as Warapu, with the ISO 639-3 code [wra] (Lewis 2014).

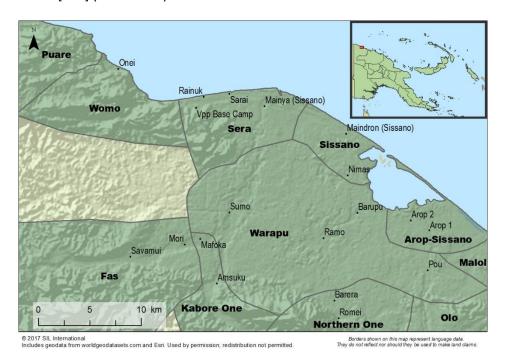


Figure 1: Map of the "Warapu" area.

The problem with this map is that Warapu is a village name, not a language name. The location of the village is not marked on the map above, but it is close to Sissano Lagoon on the coast. For more than a year I thought "Warapu" was an alternate pronunciation of "Barapu", but I have since learned from an expatriate linguist working in the area¹ that Warapu is a village on the coast, whereas Barapu is a village located more inland. So if Warapu is not the language in the area, what language is actually spoken there?

During 2015 and 2016, my wife and I periodically lived in the village of Arop 1, for a total of about five months. We worked with educated national translators from ten different languages and dialects, who

-

¹ Matthew Woods, 2016, p.c.

regularly meet together to work in the Aitape West Translation Project (AWTP).² I worked primarily with the translators from the villages of Ramo and Sumo. I also worked somewhat with their languages while I was away from the area over the course of about seven months. Moreover, I have traveled by foot and by boat throughout the area labeled as Warapu on the map, as well as through parts of the Arop-Sissano and Serra language areas. I learned that at least three languages exist in the area currently marked as the Warapu area. This paper presents evidence that they are indeed separate languages. I will focus on three of those languages: Bauni, Bouni, and Uni. Some additional information will also be given about the Bobe variety.

1. Languages

I have revised the map above to show the approximate locations of Bauni, Uni, and Bouni (moving from east to west) in Figure 2 below:³



Figure 2: Revised map showing approximate locations of the languages Bauni, Bouni, and Uni.

The fourth language in the area is Bobe. The speakers of this language left their village of Nouri and now live in the villages of Onei, Sumo, and Sarai. ⁴ More information can be found about this situation in Appendix A below.

Donohue and San Roque (2004) placed Bauni, Bouni, Uni, and Bobe in the Skou language family tree. Corris cited this tree in the 2005 version of her dissertation (Corris 2005).

² For the history of this very unique project, see Nystrom and Nystrom (2012).

³ The language boundaries drawn here should by no means be considered official or authoritative. The boundaries were simply placed between villages where the languages are known to be spoken.

⁴ The people left Nouri village after an attack.

Donohue, San Roque, Corris, and other scholars used village names as a proxy for language names in their research. Some may have done so without realizing it. This is not surprising to those with experience with languages from the area. Speakers in the area very often use village names as a proxy for their language names. Someone from the village of Ramo might say "Ramo understands Barupu," meaning that people from the village of Ramo understand the people from the village of Barupu. This is similar to saying, "Melbourne understands Detroit." However, neither "Ramo" nor "Barupu" are language names; they are instead the names of villages. The actual names speakers use for the languages spoken in these two villages are instead "Uni" and "Bauni," respectively.

I discovered this firsthand when I first visited the village of Ramo. I said that I wanted to learn Ramo. After saying this a couple of times, I was quietly corrected. I was told that I was in Ramo, but I wanted to learn to speak Uni.

In June 2001, a number of speakers of these and other languages in the AWTP worked together in an orthography workshop. During the workshop, they produced "trial spelling guides," and these are the first known documents to carry their respective language names instead of village names (Aikon et. al. 2001; Marum et. al. 2001; Nugia et. al. 2001).

1.1. Bauni

Bauni is spoken in the villages of Barapu and Warapu. In addition, the people of Pou village has shifted to using Bauni, probably as a result of intermarriage between the villages. Pou speaks a different dialect than Barapu and Warapu. It has been reported that the people of Pou village previously spoke what is known an the Bounino language variety (see section 1.5).

1.2. Uni

Uni is spoken in the village of Ramo. The people I talked from this village are proud of their language. An educated speaker estimated more than a thousand people live there.

Speakers say Uni vitality is strong. My own informal observations in the village of Ramo over a period of nine days corroborates that assessment. On one occasion, men ranging in age from teens to elders were discussing their history and forgot I was present. They spoke only Uni for roughly forty-five minutes before they remembered I was sitting in the room.

Speakers recognize three divisions within the Uni language area: Oro Kuku (meaning "Big Place"); Oro Raumo (meaning "Little Place"); and Oro Toma (meaning "Middle Place"). I am told the people in the first two divisions, Oro Kuku and Oro Raumo, speak the same; there is no difference in speech between them, and only geography determines the distinction between the two. In contrast, Oro Toma is said to use tone while the other two do not. I started a study to understand how tone is used by the Uni speakers of Oro Toma, but was not able to complete it. It is certainly an intriguing puzzle worthy of further research.

1.3. Bouni

Bouni is spoken in Sumo. Two Sumo villages actually exist, one on either side of the river that runs through the middle of their territory.

It is reported that some children understand Bouni, others cannot. One educated speaker told me that some children are learning to speak Bouni, but in my first visit to Sumo, a "big man" (or leader) in the

village told me they are not. I heard children speaking only Tok Pisin. When I asked a few of them about speaking their language, they appeared to be ashamed. In a visit to the other village, I listened to the men speaking while they assumed I was sleeping inside the house, and they freely mixed Tok Pisin with Bouni. Speakers know their language is threatened by Tok Pisin, and it is a considerable concern for some of them.

I was told by an AWTP national translator that the language area has two divisions: Ore and Biria. I have not learned whether these divisions reflect distinct dialects.

1.4. Bobe

The village of Nouri, where the language of Bobe was originally spoken, no longer exists. According to Sebby Nugia, a Bouni speaker from Sumo, Bobe speakers from Nouri continue to live in the Sumo, Onei, and Sera.⁷

I have been to Onei and I did find Bobe speakers there, contrary to what might be expected from prior research.⁸ My main purpose was to determine the language vitality of Bobe, and I learned that none of the children are learning the language; they speak only Tok Pisin (Miller 2016). The language vitality of Bobe in Sumo and Sera, the other two locations speakers migrated to, is unknown.

1.5. Bounino

An AWTP translator from the village of Pou told me that the people of Pou originally spoke Bounino, before they adopted Bauni as their language. He also told me that the elders still speak Bounino in addition to Bauni, but that Bounino is nearly extinct. I have heard nothing more about Bounino, and to my knowledge nothing has else ever been documented about it. Any further discussion of the language is thwarted by the absence of research or data.

2. Reported Mutual Intelligibility

Most of the information on mutual intelligibility comes from national translators involved in the Aitape West Translation Project, all of whom are educated and are native speakers of their languages for which they report. I drew up a chart of what I learned from them, and all of them verified its accuracy. I also asked a few speakers outside the translation project about intelligibility, and their responses confirmed what the AWTP translators had reported. That chart has proven to be confusing to most people unfamiliar with the area, so a short synopsis is as follows:

4

⁵ Tok Pisin, or Melanesian Pidgin, is the language of wider communication in the area, as it is throughout much of Papua New Guinea.

⁶ I later told them I was listening.

⁷ See either of the maps in the introduction above for locations. Onei is a village in what is currently marked as the "Womo" language area in the northwest of the map. Sarai village is in the Sera language area.

⁸ See section A.2 Bobe below.

Speakers of Bauni understand Uni, and Uni speakers understand Bauni; however, it is a learned intelligibility. One Uni speaker told me, "If you have French relatives, you'll speak some French." Whether that statement is actually true for me as an American is beside the point; it is true for him. Only the "big men" (or leaders) from the village Pou can understand Bouni, but the people from the village of Barapu do not understand anything of Bauni.

2.2. Uni

2.1. Bauni

As stated above in 2.1, Uni speakers (who live in Ramo) have a learned intelligibility with Bauni speakers. Some Uni speakers can understand a little of Bouni, but those from Ora Toma cannot, because of the Ora Toma use of tone.

2.3. Bouni

Only the "big men" (or leaders) can understand Bauni speakers from the village of Pou. They can speak a little of Bauni from the village of Barapu, even though the reverse is not true.

Likewise, Bouni speakers can understand a little of Uni speakers. They cannot, however, understand Uni speakers from Oro Toma, again because of the latter's use of tone.

Bouni speakers are the only ones who can understand Bobe. Again, it is unknown whether Bobe is a separate language, or if Bobe and Bouni are dialects of each other.

2.4 Bobe

Sebby Nugia said that the Bobe is mutually intelligible with Bouni⁹. Questions remain as to whether Bobe is a dialect of Bouni or a separate language.

3. Cognate Levels

Both Corris (2003) and Robertson (2009a-2009d) made recordings of word lists from languages in the area. Neither researcher documented which word list was used, nor could Robertson tell me when I asked her years later. However, the word lists bear a strong resemblance to the SIL Survey List, and were probably derived from it, given that both researchers had interaction with the AWTP and SIL. Recordings were made of three Bauni speakers, two from Pou and one from Barapu. Other recordings were done of two Uni speakers and two Bouni speakers.

⁹ Sebby Nugia, 2016, p.c.

¹⁰ Helen Robertson, 2016, p.c.

¹¹ SIL Survey Word List, English/Tok Pisin version. August 1999 (1999 Revision, 120 Adaptation).

I transcribed the seven recordings¹² and analyzed them with the Blair Method (Blair 1990:31-33), using Cog,¹³ a software program. Both¹⁴ lexical¹⁵ and phonetic similarities¹⁶ were compared, and their respective charts presented below were generated by the software. While the phonetic similarities scored higher than the lexical similarities, the patterns of the two forms of similarity nearly matched each other.

3.1. Lexical Similarities

Lexical similarities range between 38% and 85%:

	Bauni (in Pou) 1	Bauni (in Pou) 2	Bauni (in Barapu)	Uni 2	Uni 1	Bouni 1	Bouni 2
Bauni (in Pou) 1		<u>72</u>	<u>67</u>	<u>59</u>	<u>53</u>	<u>39</u>	<u>45</u>
Bauni (in Pou) 2	<u>72</u>		<u>47</u>	<u>50</u>	<u>59</u>	<u>52</u>	<u>45</u>
Bauni (in Barapu)	<u>67</u>	<u>47</u>		<u>52</u>	<u>53</u>	<u>38</u>	<u>38</u>
Uni 2	<u>59</u>	<u>50</u>	<u>52</u>		<u>82</u>	<u>56</u>	<u>56</u>
Uni 1	<u>53</u>	<u>59</u>	<u>53</u>	<u>82</u>		<u>50</u>	<u>52</u>
Bouni 1	<u>39</u>	<u>52</u>	<u>38</u>	<u>56</u>	<u>50</u>		<u>85</u>
Bouni 2	<u>45</u>	<u>45</u>	<u>38</u>	<u>56</u>	<u>52</u>	<u>85</u>	

Figure 3: Lexical Similarity Matrix

¹² See Appendix B for data.

¹³ For more information about Cog, see its website: https://github.com/sillsdev/cog/wiki/Cog-Tutorial.

¹⁴ For a brief, overall discussion of phonetic similarities and lexical similarities, the Cog help documentation refers to a SurveyWiki.info article; see "Blair Method" (2016).

¹⁵ Regarding lexical similarities, Damien Daspit, the developer of Cog, wrote (1 February 2017 p.c.) that Blair (1990:32 table 8) "presents the criteria for lexical similarity for a word pair. Blair refers to it as 'linguistic similarity.' The table can be generalized as 'if 50% or more of the segment pairs are category 1 AND 75% or more of the segment pairs are category 1 or 2, then the word pair are lexically similar.' Once Cog determines which word pairs are lexically similar (likely cognates), it calculates an overall lexical similarity for the variety pair. This is simply calculated as '# of lexically similar word pairs / total # of word pairs', so you get a percentage of lexical similarity."

¹⁶ Regarding phonetic similarities, Damien Daspit, the developer of Cog, wrote (1 February 2017 p.c.) that "Cog uses a method called ALINE to calculate the phonetic similarity for a word pair. It is basically a phonetically-aware version of the Levenshtein distance. A phonetic distance is calculated for each segment pair based on the phonological features of the segments. For example, if you have two consonants that have the same manner and place of articulation but one is voiced and the other is not, the phonetic distance would be very small. It takes the distances for all of the segment pairs and uses it to compute a phonetic similarity as a percentage for the entire word pair. Once Cog has the phonetic similarity for all word pairs, it calculates the overall phonetic similarity for a variety pair by averaging the phonetic similarity percentages for all of the word pairs. ALINE was originally proposed in the paper [by Kondrak (2000)]. The paper [(Downey, et. al. 2008)] uses ALINE to compute an overall phonetic similarity for a variety pair in a similar fashion to Cog."

The greatest similarities, unsurprisingly, were between speakers of the same language. The two Bouni speakers are rated 85%, the two Uni speakers are rated at 82%, and the two Bauni speakers in Pou 72%. The results for Bauni may be skewed because so few words from the second Bauni speaker were elicited. The similarities between the speech of the Bauni speaker in Barapu and the speech of those in Pou was 67% and 47%. Again, the latter may be skewed because so few words from the second Bauni speaker were elicited. The similarities between all the others were 60% or less.

The greatest divergence was between the speech of the Bouni speakers and those of the other languages, scoring as low as 38% with speech of the Bauni speaker in Barapu.

3.2. Phonetic Similarities

Phonetic similarities range between 54% and 88%, which is higher than lexical similarities:

	Bauni (in Pou) 1	Bauni (in Pou) 2	Bauni (in Barapu)	Uni 2	Uni 1	Bouni 1	Bouni 2
Bauni (in Pou) 1		<u>84</u>	<u>72</u>	<u>69</u>	<u>66</u>	<u>53</u>	<u>57</u>
Bauni (in Pou) 2	<u>84</u>		<u>68</u>	<u>64</u>	<u>69</u>	<u>63</u>	<u>61</u>
Bauni (in Barapu)	<u>72</u>	<u>68</u>		<u>62</u>	<u>64</u>	<u>54</u>	<u>55</u>
Uni 2	<u>69</u>	<u>64</u>	<u>62</u>		<u>86</u>	<u>68</u>	<u>66</u>
Uni 1	<u>66</u>	<u>69</u>	<u>64</u>	<u>86</u>		<u>65</u>	<u>68</u>
Bouni 1	<u>53</u>	<u>63</u>	<u>54</u>	<u>68</u>	<u>65</u>		<u>88</u>
Bouni 2	<u>57</u>	<u>61</u>	<u>55</u>	<u>66</u>	<u>68</u>	<u>88</u>	

Figure 4: Phonetic Similarity Matrix

Again, the greatest similarities were between speakers of the same language. The two Bouni speakers were rated at 88% similarity, the two Uni speakers were rated at 86% similarity, and the two Bauni speakers in Pou were rated at 84% similarity.

The similarities between the speech of the Bauni speaker in Barapu and one of the Bauni speakers in Pou was 72%. The similarities between all the others was 70% or less.

Again, the greatest divergence was between the speech of Bouni and the other languages, and the greatest divergence was between them and the speech of the Bauni speaker in Pou, scoring as low as 53%.

3.3. Donohue's ASJP Comparison and Language Trees

The similarity matrices above are in agreement with findings Donohue (2010) presented as part of the Automated Similarity Judgment Program (ASJP),¹⁷ providing more evidence of the differences between the languages. (See Figure 5 below.) The speech of Bauni in Barupu is most closely linked with Bauni in Pou (Poo). Uni, the speech of Ramo, is close to the two of them, but less so. Bouni, the speech of Sumo, is related to the others, but not so much as to be mutually intelligible. Bobe was not charted.

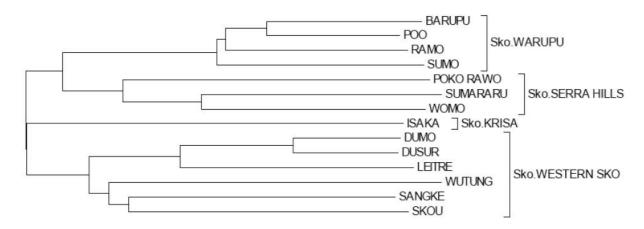


Figure 5: Donohue's ASJP Tree for Skou Languages (Donohue 2010). Barupu = Bauni spoken in Barapu; Poo = Bauni spoken in Pou; Ramo = Uni spoken in Ramo; Sumo = Bouni spoken in Sumo.

Mark Donohue started work on the western side of the Skou languages and eventually made the language family an object of study. The tree he produced for the "Macro-Skou linkage" with San Roque (2004:6) is below in Figure 6. Of interest here are: Bobe (No for Nouri), Bouni (So for Sumo), Uni (Rm for Ramo), and Bauni (Ba for Barapu). All appear to be recognized as languages under "Warapu":

(http://en.wikipedia.org/wiki/Automated Similarity Judgment Program). This says, in part: "The Automated Similarity Judgment Program (ASJP) is a collaborative project applying computational approaches to comparative linguistics using a database of word lists. The database is open access and consists of 40-item basic-vocabulary lists for well over half of the world's languages. It is continuously being expanded."

 $^{^{17}}$ For more information about the ASJP database, see Wichman and Brown (2016). For background information, the ASJP website refers to the Wikipedia page

Macro-Skou linkage Skou-Serra-Piore linkage I'saka Skou Serra Hills Piore River (see below) Rawo-Main Serra Pu Warapu Rw Rm Ba Main Serra chain So Su Wm Mo

Figure 6: Donohue and San Roque's (2004) Skou language tree. No = Nouri, where Bobe was once spoken. So = Sumo, where Bouni is spoken. Rm = Ramo, where Uni is spoken. Ba = Barapu, one of the places where Bauni is spoken.

In 2010, Donohue updated this tree. He combined Barapu and Pou together under one node, which makes sense, since they both speak Bauni. He tentatively shifted Nouri over to the Main Serra tree, and replaced "Warapu" with "Piore River" (Figure 7):

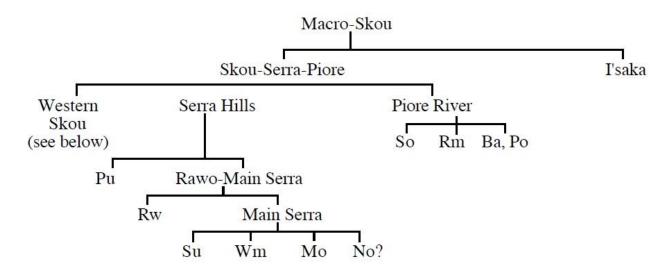


Figure 7: Donohue's "Comparative Method phylogeny of the Skou languages" (Donohue 2010:2). So = Sumo where Bouni is spoken. Rm = Ramo, where Uni is spoken. Ba = Barapu and Po = Pou, the two places where Bauni is spoken. No = Nouri, where Bobe was once spoken.

3.4. Cognate Levels Conclusion

The results in the lexical and phonetic similarity matrices substantiate the mutual intelligibility reported by speakers in section 2 above. Both the similarity matrices and mutual intelligibility reports agree with Donohue's ASJP tree and language tree (2010).

One expatriate working with these languages in the AWTP remarked to me more than once, "How can anyone think these are only one language?" ¹⁸

4. An Updated Language Tree

Given the findings and research presented above, I propose some small alterations to Donohue's 2010 language tree (Figure 7 above): language names replace village names, and "Lagoon" replaces "Piore River."

For some years, Donohue used "Piore River" in his language trees. When I asked people who live and work in the area for the location of this river, no one could give me an answer. After consulting several maps, I discovered that it is a tributary of the river that runs between the two locations of Sumo village, to the west of all the languages discussed here. Bouni speakers living in the area have a name for each section of the river, but "Piore" is not one of them. While I have found agreement that the name "Piore" should be replaced, finding a replacement name has been more difficult. Sissano Lagoon is by far the most widely known geographic feature in the region, and all know it by that name. However, "Sissano" does refer to another group of languages. When I discussed the issue with Matthew Woods of AWTP, we thought the best compromise is to simply call the node "Lagoon". The node above Lagoon in turn becomes "Skou-Serra-Lagoon".

I moved Bobe away from the Serra Hills branch and back onto the Lagoon branch, because it is reportedly mutually intelligible with Bouni. More analysis needs to be done on Bobe to adequately categorize it.

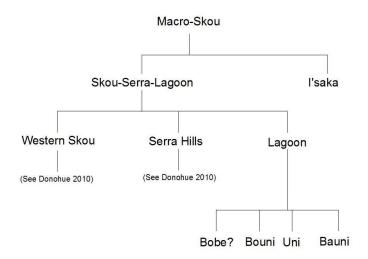


Figure 8: Updated Skou language tree.

_

¹⁸ Matthew Woods, 2016, p.c.

Conclusion

Given Donohue's work and findings presented here, three languages are presented: Bauni, Uni, and Bouni. The classification of Bobe is unknown. Historically these languages have been grouped together as Warapu [wra], but are now grouped under the "Lagoon" node in the Macro-Skou tree as separate languages. Donohue's language tree has been modified somewhat to reflect this.

Appendix A. Previous Research

The practice of using village names as proxies for language names makes finding previous research about the languages more challenging. This section is included to help other researchers trace previous research on these languages through time.

A.1. "Warapu" (Bauni, Uni, and Bouni)

The best sketch of the earliest academic work done can be found in Laycock (1975:849-851). This research extends as far back as 1912. Donohue and Crowther (2005) offer empirical evidence for linguistic interaction and change.

The languages Bauni, Uni, Bouni, and Bobe have long been identified together by researchers as the language of "Warapu." For instance, Laycock wrote (1973b:250-251): "Warapu is also spoken in three more villages, Po, Ramo, and Sumo, all lying south and inland of Sissano, and of the original village of Warapu."

Two years later Laycock (1975:851) recognized eight languages of the Skou (Sko) family. The chart is reproduced below in Figure 9 below, with population figures. He listed the language "Warapu" under the "Krisa Family" under the "Skou Phylum-Level Stock". He also listed the population of Warapu as 2,991.

SKO PHYLUM-LEVEL STOCK	6,570+
Vanimo Family	2,355+
Sko	350+
Sangke	200+
Wutung	410
Vanimo	1,395
Krisa Family	4,215
Krisa	437
Rawo	506
Puari	371
Warapu	2,991

Figure 9: Laycock's chart of languages, with population figures.

As has been noted above, Donohue, Corris, and San Roque have produced the greatest amount of research of these languages. (See section 3.3. above.)

A.2 Bobe

In published literature, Bobe has consistently been identified by its place name "Nori", spelled here and in Donohue as "Nouri."

Capell (1962:38) wrote that Nori was a language used by a single village, with no apparent relation to the languages on the coast, such as Vanimo. This was indeed the case at one time, but the people have since migrated to three other villages: Onei, Sumo, and Sarai.

Laycock (1975:851) did not add Nori to his Skou language tree. He did not explicitly state his reason for this.

Baron (1986:3) said the people of Nori joined the Puindu village of Sera, but he, Sugu Afoke, and Steve Whitacre could find only three elderly speakers there in 1979.

Donohue and San Roque (2004:6-7) wrote that Bobe was "the western-most member of the Piore River family, which under strong pressure from the adjacent Serra Hills languages has moved so far in their direction that is it now appropriate to refer to it as a mixed language." However, six years later, Donohue (2010) amended this: "Note that there is no clear subgrouping for Nouri, an extinct language of the southeastern Serra Hills. The limited material available shows Nouri sharing some traits with the Serra Hills languages, and some with the Piore River languages." He further added, "Nouri appears from a phonological and morphological perspective to be liminal between these two categories [of Serra Hills and Piore River (=Warapu)] [sic], and may provide a link between these two major subgroups, though questions of its 'best-fit' affiliation in one or the other of these two groups remain uncertain." (Dononhue 2010).

According to Sebby Nugia, a Bouni speaker from Sumo, Bobe speakers from Nouri continue to live in Sumo, Onei, and Sarai, and the language is mutually intelligible with Bouni. I have spoken with Bobe speakers in Onei (Miller 2016). Questions do remain as to whether it is a dialect of Bouni or a separate language.

Appendix B: Word Lists

Corris (2003) recorded word lists of speakers from the villages of Ramo, Sumo, and Pou. Robertson (2009a-2009d) later recorded word lists of speakers from those villages and Barupu, but did not publish them. Both word lists show significant similarities to the SIL Survey Word List (1999) and were probably based on it. I transcribed the word lists using a variety of software applications, including SayMore and ELAN¹⁹, then analyzed them using Cog.

Semantically similar words with the same stem were eliminated from the list, so as not to skew the results.²⁰ For instance, nearly all the words in the languages for "leaf" and "bark" had the same stem as

19

¹⁹ SayMore is produced by SIL for organizing recordings, transcribing, and other language documentation tasks. For more information, see its website: https://saymore.palaso.org. SayMore is designed to work with ELAN, which creates annotations on video and audio recordings. ELAN is produced by the Max Planck Institute for Psycholinguistics. For more information, see its website: https://tla.mpi.nl/tools/tla-tools/elan/.

²⁰ This recommendation was from the SIL Survey Department. The term originally used was "doublets," but it has been pointed out that this use of "doublet" has an established meaning in historical linguistics, one in which two words are derived from the same historical source.

the word for "tree." Similarly, the words for "green coconut juice" had the same stem as the word for "coconut."

B.1 Issues with Elicited Verbs

At least Uni and Bouni mark verbs for tense, and I suspect the other Skou languages do, too, but the word list was elicited using Tok Pisin, which is not marked with tense. I think at least one speaker used past tense in his responses, but the tense used in most of the others is not known. When I used the SIL Survey Word List with Bobe speakers in Onei, they discussed at length whether to use the imperative or some other verbal form. I expect the recordings in this study suffer from similar issues with verbs. Furthermore, Uni verbs reflect the gender of the speaker. It is possible the verbs of the other languages do as well. Both word list compilers were female, and all speakers were male. I believe one male speaker responded as though he were female for the benefit of the compiler, but it is unknown whether the other speakers responded for themselves as male, or as female on behalf of the compilers.

Given these problems with the verbs, I removed most of them from list to produce the lexical similarity matrix shown above (Figure 4, repeated below as Figure 10). I then produced another lexical similarity matrix for just these verbs (Figure 11 below). Only two verbs in total were recorded by Uni 2 and Bouni 2, so they are not shown. Based on what I heard in the recordings, my guess is that researchers recognized these problems with verbs as their research progressed and abandoned further attempts to elicit them.

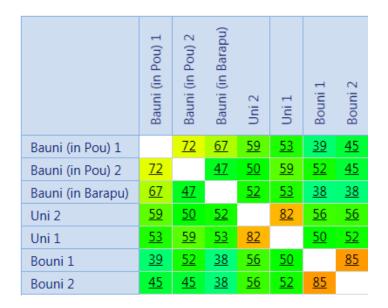


Figure 10: Lexical Similarity Matrix (without verbs)

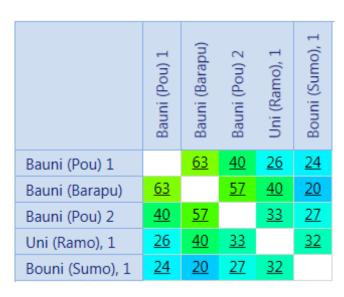


Figure 11: Lexical Similarity Matrix of just verbs

Based on these findings, I believe lexical comparisons of verbs are suspect at best.

B.2. Word List Tables

Words were elicited using Tok Pisin, and the Tok Pisin words are included in the table below. I provided an English translation from Tok Pisin for reference. In addition, I correlated the seven recordings into a single table.

Phonetically, Uni has at least two tones, high and low, and the tones contrast phonemically. Bauni has more. However, Cog ignores the use of diacritics in IPA transcriptions, and so tone does not have a bearing in these results.²¹

A full language survey would have recordings and transcriptions of more speakers.

Two tables are presented below. The first is a table of elicited words, minus verbs, that were used in the lexical similarity matrix in the body of the paper. The second is a table of elicited verbs used in the second chart. The POS column is for part of speech, and has the following key: adj = adjective, adp = adposition, det = determiner, n = noun, num = number, q = question. Uni 1 is the Corris (2003) recording of Vincent Yove; Bouni 1 is her (2003) recording of William Sipai Raikos; and Bauni 1 is her (2003) recording of Clement Tonini. The four other recordings are from Robertson (2009a-2009d) of unidentified speakers.

²¹ Cog evaluates IPA tone marks differently than diacritics, and Damien Daspit, Cog's developer, is aware of that behavior. For that reason, I changed the phonetic transcription from using tone marks to diacritics.

²² Cog uses the part of speech in its calculations.

B.2.1 Word List Table (without Verbs)

Englis h	Tok Pisin	P O S	Bauni 1 (in Pou)	Bauni 2 (in Pou)	Bauni (in Barapu)	Uni 1	Uni 2	Bouni 1	Bouni 2
head	het	n	t͡ʃa		t∫a	ˈti.a	ˈti.a	t͡ʃa.ˈpó.ɾ e	t͡ʃa.ˈpó.r i
eye	ai	n	'í.ne		ˈʔí.ne	'í.ne	ˈʔí.ne	ˈʔi.ne	ˈʔi.ne ^j
ear	yau	n	ˈte.βi		ˈti.βi	ˈti.βi		'te.be	ˈté.βi
nose	nus	n	'ʔu.θ	ˈʔu.we	ˈʔu.wə	'ʔu.θ	'ʔu.θ	wu.e	ˈʔu.wθ
mout h	maus	n	ro	ro	го	ro	LO	ιο	Ro
teeth	tit	n	?e	?e		'?e.θ	?e.u	?e.u	?e.u
hand	han.tasol	n	ka ^j .ni.ja. tu			ˈkaʷ.ta	'ʔe.nθ	ˈʔe.nə	ˈʔé.nu
joint	skru	n	po.ro.ˈtá .pu			ki.ka.ˈp âʷ	gi.ga.ˈb áʷ	ki.ka.ˈpâ w	ti.ka.ˈpâ w
finger	pinga	n	ka ^j . 'ni.tu			'ʔe.nø.t e.me		ˈʔe.no. me	
breast	susu	n	tuː			tu:		tu?	
belly	bel	n	ˈrí.mu		gri	riː	te	to?	to
heart 23	lewa	n	'to.ton	to.ˈtó m	to.'tóm	to.ton		to.'ton	to'tom
knee	skru.bilong.l ek	n	po.ro.ˈtá .pu		?a.nin.t á.po	ө.ni.ˈtá. po		?u.ˈra.ta	
skin	skin	n	ˈtâʷ	re.ˈká		ˈní.ti	ní.di	té.pa	ˈté.pa
blood	blut	n	ne	ne	no	ne		nu	nê
bone	bun	n	ˈrí.ka	ri.ga	ˈri.ka	ˈrí.ka	ri.da	ˈre.βa	ˈre.βa
muscl e	mit	n	bê ^j			bê ^j	be ^j	be ^j	
urine	pispis	n	'tiru			ˈti.ɾu		ˈti.ru	
feces	pekpek	n	?aː			?aː		?a:	
mama	mama	n	muː		k ^w an	mûː	mûː	na.ˈní	
рара	рара	n	ˈʔa.ka	ˈta.ta	?á.ka	ˈta.ta		na.ηo	ˈná.ni
child	pikinini	n	ma			ma		maː	
Englis h	Tok Pisin	P O S	Bauni 1 (in Pou)	Bauni 2 (in Pou)	Bauni (in Barapu)	Uni 1	Uni 2	Bouni 1	Bouni 2
male in-law	tumbuna.ma n	n	ˈá.pu			ˈá.pu	?á.pu	ˈʔaː.pu. wɔ	

_

²³ At least three different native speakers of Tok Pisin I consulted were most comfortable with this gloss.

wind	win	n	pû		pû:	pû:	bû		pû:
sun	san	n	ˈʔu.me			ˈʔu.me	ˈʔu.me	ˈʔú.mɔ	ˈʔú.mɔ
moon	mun	n	ˈʔu.ra		ˈʔu.ra	ˈʔu.ra	'ʔu.ra	'ʔú.ra	ˈʔú.ra
night	nait	n	pru.ru		ˈrú.ru	ru.ru		ru.ru	ˈru.ru
morni	moning	n	ru.ná ^j		ru.ru.ˈn	ru.ɾu.ˈn	ru.ŗu.ná	?in.ti.ki.	?in.ti.ki.
ng					á ^j	á ^j	j	na ^j	ˈná ^j .i
aftern oon	apinun	n			ba.ˈrí.ri		ˈba.ɾi.ɾi	ba.ˈrí.ri	ba.ˈrí.ri
star	sta	n	ʔa.βa.ro ^j		ka. ˈmú	ˈbi.r ^j u		ka ^w .mɔ	ʔa.ma.r ô ^j
smok e	smok	n	'ʔú.θ		ˈʔu.mə	'ʔu.θ		'ʔú.βɔ	ˈte.u
wilder ness	bus	n	ˈʔú.ga			ˈʔu.a		ˈʔu.ka	
moun tain	mauntan	n	to		tow	ˈtú.ru	ˈtu.ru	'tu.ru	ˈtu.ru
garde n	gaden	n	?e.ˈrú.ra			'ʔé		ˈʔéː	
house	haus	n	ˈʔú.ru		?ó.ru	'ʔu.ru		'?e.re	7ú.ru
thatc h roof	rup.mota	n	mo:			mo.'pe:		ˈʔu.pe	
village	ples	n				'ʔu.ru		ku.ˈrú.m	
fence	banis	n	ri			ˈʔa ^j .a.te		ˈpoʲ.me	
line ²⁴	lain	n	tî:			tî:		ti	
canoe	kanu	n	ˈpó.ro			ˈpo.ro		ˈpó.ɾe	
? ²⁵	pun	n	bi.je.tu			bi.ˈjo.te		bi.ˈjó.te	
road	rot	n	ˈrá.ra	га.га	ˈrá.ra	ˈta.ɾa		ˈtá.ɾa	ˈtá.ra
Englis h	Tok Pisin	P O S	Bauni 1 (in Pou)	Bauni 2 (in Pou)	Bauni (in Barapu)	Uni 1	Uni 2	Bouni 1	Bouni 2
axe	tamiak	n	?a.me.ˈg			?a.me'.	?a.ma.ˈ	?am.ˈké.	?a.m ^j a.ˈ
			á.me			ga.me	gá.me	mo	ké.mo
bow	boanaro	n	ˈrú.a			ˈtú.a		ˈtú.a	
arrow	spia	n	ˈko.ka ^j n		ˈʔáʷ.ra	ˈkə.kə	ˈke.ge	?a.ˈkru.t u	?a.ˈkrú.t u
string bag	bilum	n	é ^j		?e.kó.k	?e ^j .ˈko. ko	?е	ˈʔe.ko.k	?e.kө.kө
this	dispela.hia	de t	bi:		bé	in.if'	'ʔi.ɾi	ˈʔu.ri.a	ˈʔo.wa. wo
that	dispela.long wei	de t	bi.ri			in.in.iS'		?o.βa ^j	

 $^{^{24}}$ Tok Pisin "lain" is polysemous. It is uncertain which meaning is meant here. 25 It has been suggested that this is "paddle".

cold

kol

maˈrí.rí

mo.ri.ri

mo.ri.ri

mo.ˈrí.ri

bo.mo.'

rí.ri

ma.rí.ri

ˈma.rí.rí

ad

Barapu)

Pou)

S

	1			1		T T
full	pulap	ad	ro.ro.k	ro.ˈrí		ˈró.bú.ti
		j	ó.re.u			
fat /	gris	n		're.re		ˈre.re
greas						
е						
sugar	suga.ken	n		kó	kô:	kôw
cane	(sic)					
yam	yam	n		ˈmi.je	me	?i.⊖
bean	bin	n		ba.ˈnú.n	ˈba ^j .nu.	?a.ka.p
				u	ţи	on
yester	asde	n		ba.ri.ri.t	ˈba.ɾi.ɾi	ba.ˈrí.ri
day				a'táre		
tomor	tumora	n		'ú.ri	?ú.ri	?u.ˈtí.gi
row						
all	olgeta	ad		ma ^w .'m	mu.ˈmu	gu.na ^j . 'j
		j		á ^w		á.wa
where	we	q		ˈro.pi	'ʔi.ɾi	?ɔ.ni
yes	yes			?á.wө	ˈʔa.wa ^j	ʔaʷ.ˈmá
						.ti
bat /	pilai.pokis	n	ma ^j .tu.		ˈma ^j .tu.	
flying			ˈpá		ра	
fox						

B.2.2 Word List Table: Verbs

English	Tok Pisin	Bauni 1 (in Pou)	Bauni 2 (in Pou)	Bauni 1 (in Barapu)	Uni 1	Bouni 1
look	lukim	ˈjá.rá	-	ka.ˈjá.ra	ˈri.bu	ti.ˈáː.ɾaʷ
I look	mi lukim	nu.mi.ˈja1.ra1	ka. ˈprí.ki	,	?a.na.ˈri1.bu	ba.na.ti.ˈa1.raʷ
hear	harim	k ^w a.βé	ka.rí.βo		ˈʔe.kə	
talk	toktok	ka ^j .re	·		ˈaʷ.i.a ^j	ro
I talk	mi tok		?a.wa.ka ^j .re	ˈka ^j ra ^j	ʔa.na ^j .a ^j	rɔ.ˈpi1.ka
sleep	aisleep	ka.na.na.ˈnaʲ.ɲa			ˈʔi.na.to.ju	?i.no.bo.ˈto.ba
I fear	mi pret	ˈɾi.ɾi			ˈɾi.ɾi	ba.na.ˈra.bu
I (MASC) sit	mi (man) sindaun	ka.ˈké	ko.ˈké	ka'ké	?akeni	ba.ke.ni
I (MASC) walk	mi (man) wokabaut	ˈkaʷ.ti	ká ^w .tu	ˈkáʷti	tu.a	ra.na.ˈtu1.ba
swim	swim	ˈta.ɾu.ta.ɾu		ka ^w 'úra	'ro.to	ˈru1.a
fly	flai	ˈpu.pu			ˈpu.pu	ˈpu.pu
wash	waswas	ˈkɔ.ni1			ˈni.ti	pi.la.nin
dig	dig	me.ri.ku ^j			'ʔu.ti	ˈʔu.mi
cut/stab	sutim-long-naip	tu.ˈta1.ri			k ^j u.ˈte1r.i	?e.mo.ˈki1.a
fall	pundaun	ku.ré	ka.ré		re ^j	re ^j .mi
eat	kaikai		?a.wa.ka.ra	ˈkárá	?a [₩]	ma:
I (MASC) eat	mi (man) kaikai	a.βu.ka.ra	?a.wa.ka.ra		?a.na	ˈra.ma
he/she eats	em kaikai				ˈʔa.wa.ta	ˈba.ta
2 of us eat	mitupela kaikai				ˈʔe.pi.pa	re.pa
drink	dring	ko.tún	ka. 'tún	kaˈtún	tun	pun
die	dai	ˈko.ɾai	ka.ˈrá ^j		ra ⁱ V	ba.ra ^j
give	givim	lum.wa			?a.t ^w a	kum.wa
come	kam	na.ˈkam1	ka.ˈká ^j	kaˈká ^j	ˈka.mi	ˈka1.mi
go up	go-antap	'no.kum.na.ki			?o.ˈko1.mi	ˈko.mi
laugh	lap	ˈʔa.wo.to	?aʷˈto.ka.ˈráʷ	a.wo.to	ˈʔa.wo.to	ˈʔo.ta
cook by fire	kukim long paia	ra ^j .ni			ki1	ki:
blow on fire	winim paia	nom. pu1.tu			ˈpu1.tu	ˈpu.tu
hit	paitim					mi.ja
find / hunt	painim	nimoˈɾi1naʷ			?өге	ˈʔɔri
bite (by mosquito)	kaikai (natnat)	?e.jo.ˈji1.na				'?eː.bo.ti.na
no	nogat	bɔn		ˈbaʷ.ni	?u.ni	bon
cry	krai	ˈku1.ru			ˈpə.ru	
go down	go daunbilo	mara pi1ta			?o.ˈma1.ru	

Bibliography

- Aikon, Chris, Clement Tonini, Michael Kawai, Leonard Pundi, Gebbie Avera, Jack Roipe, Deborah Larkins, and Hannah Marie Hoel. June 2001. Trial spelling guide for the Bauni language. Manuscript.
- Baron, Wietze. 1986?. *Malay Influence on West Sepik Kinship Terminology*. Linguistic Society of Papua New Guinea. http://www.kwomtari.net/malayinf.pdf. Accessed 2 February 2017.
- Blair, Frank. 1990. Survey on a Shoestring. Dallas: Summer Institute of Linguistics.
- "Blair Method". 13 July 2011. http://surveywiki.info/index.php/Blair_Method. Last accessed 17 January 2017.
- Capell, Arthur. 1962. *A linguistic survey of the south-western Pacific*. South Pacific Commission Technical Paper No. 136, Noumea.
- Corris, Miriam (compiler), Miriam Corris (recorder), William Sipai Raikos (speaker), Clement Tonini (speaker), Vincent Yove (speaker). 2003. Ramo, Sumo and Pou Wordlists. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). http://www.language-archives.org/item/oai:paradisec.org.au:MC1-200303. Accessed 30 August, 2015.
- Corris, Miriam. 2005. *A grammar of Barapu*: *a language of Papua New Guinea*. Ph.D. thesis. Sydney: University of Sydney.
- Donohue, Mark and Lila San Roque. 2004. *I'saka: a sketch grammar of a language of north-central New Guinea*. Pacific Linguistics 554. Canberra: Research School of Pacific and Asian Studies, Australian National University. Pages 6-8.
- Donohue, Mark and Melissa Crowther. 2005. Meeting in the middle: interaction in North-Central New Guinea. In Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide, eds., *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*: 167-184. Canberra: Pacific Linguistics.
- Donohue, Mark. 2010. *Skou*. Canberra: Australian National University.

 <u>papuan.linguistics.anu.edu.au/Skoufiles/Donohue</u> 2010 <u>Skou</u> <u>ASJP.pdf</u>. Accessed 25 September 2015.
- Downey, Sean S. and Brian Hallmark, Murray P. Cox, Peter Norquest, J. Stephen Lansing. 2008. Computational feature-sensitive reconstruction of language relationships: Developing the ALINE distance for comparative historical linguistic reconstruction. Journal of Quantitative Linguistics, vol. 15, num. 4, 340-369.
- Kondrak, Grzegorz. 2000. A new algorithm for the alignment of phonetic sequences. NAACL 2000 Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference. Pages 288-295. Online: http://dl.acm.org/citation.cfm?id=974343. Accessed 2 February 2017.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. Ethnologue: Languages of the World, Seventeenth Edition. Dallas, Texas: SIL International. Online: http://www.ethnologue.com.

 Accessed 29 September 2015.
- Laycock, Donald C. 1973a. *Sepik languages: checklist and preliminary classification*. Canberra: Pacific Linguistics B-25.
- Laycock, Don 1973b. Sissano, Warapu and Melanesian Pidginization. Oceanic Linguistics 12(1/2), 245-277
- Laycock, Donald C. 1975. Sko, Kwomtar, and Left May (Arai) phyla. In Wurm et al., eds *New Guinea area* languages and languages study, vol. 1: Papuan languages and the New Guinea linguistic scene, 849-858. Canberra: Pacific Linguistics C-38.
- Marum, Maris, Joe Maikon, Justine Roy, Patrick Nono, Vincent Yove, Deborah Larkins, and Hannah Marie Hoel. June 2001. Trial spelling guide for the Uni language. Manuscript.

- Miller, S. A. June, 2016. The Languages of the Village of Onei in Sandaun Province: Bobe (Nori/Nouri) and Pin (Womo[wmx]). Online: www.PngLanguages.org.
- Nugia, Sebby, Florian Krope, Michael Wamu, Deborah Larkins, and Hannah Marie Hoel. June 2001. Trial spelling guide for the Bouni language. Manuscript.
- Nystrom, John and Bonnie Nystrom. 2012. Sleeping Coconuts. Orlando: Wycliffe Bible Translators.
- Robertson, Helen (compiler), Helen Robertson (recorder), unknown speaker. 2009a. Recording of word list of a speaker from Barapu. Unpublished.
- Robertson, Helen (compiler), Helen Robertson (recorder), unknown speaker. 2009b. Recording of word list of a speaker from Pou. Unpublished.
- Robertson, Helen (compiler), Helen Robertson (recorder), unknown speaker. 2009c. Recording of word list of a speaker from Ramo. Unpublished.
- Robertson, Helen (compiler), Helen Robertson (recorder), unknown speaker. 2009d. Recording of word list of a speaker from Sumo. Unpublished.
- SIL Survey Word List, English/Tok Pisin version. August 1999 (1999 Revision, 120 Adaptation).
- Wichmann, Søren, Eric W. Holman, and Cecil H. Brown (eds.). 2016. The ASJP Database (version 17). http://asjp.clld.org/. Accessed 23 January 2017.