



Historical Relationship of the Sahu, Gamkonora, and Waioli Languages in West Halmahera, North Maluku

Ety Duwila^{1*}, Farida Maricar², Rahma Do Subuh³, Agus Supriyadi⁴

^{1,2,3}Faculty of Cultural Sciences, Khairun University, Indonesia

⁴Faculty of Teachers Training and Education, Khairun University, Indonesia

Corresponding Author: Rahma Do Subuh rahmadosubuh@unkhair.ac.id

ARTICLE INFO

Keywords: Historical Relation, Sahu, Gamkonora and Waioli Language, West Halmahera

Received : 28, April

Revised : 25, May,

Accepted: 23, June

©2026 Duwila, Maricar, Subuh, Supriyadi: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

The indigenous languages of North Maluku Province generally exhibit distinct characteristics. While some belong to the Austronesian language family, others are classified as non-Austronesian. Despite shared origins, several of these languages display significant differences in grammatical structure. The focus of this research is on three isolects in West Halmahera: Sahu, Gamkonora, and Waioli, whose speech areas are geographically separated. There exists a classification ambiguity where language authorities categorize the three as dialects of the Sahu language, whereas native speakers claim independent linguistic identities. By employing comparative historical linguistic methods and lexicostatistical techniques applied to the 200 Swadesh basic word list, this study aims to reconstruct the genetic relationships and phonological change patterns among them. The analysis reveals regular phonemic correspondences, such as $[k] \approx [g]$ and $[r] \approx [l]$, as well as a unique innovation in the Waioli language involving the process of final vowel paragoge. Lexicostatistical results show cognate percentages ranging from 64% to 71%, which places these three isolects in the "Language Family" category rather than merely being dialects. These findings provide an empirical basis for the necessity of reviewing language mapping and revitalization efforts in North Maluku.

INTRODUCTION

Many comparative linguistic studies have been conducted on languages in western Indonesia, such as those by Nothofer (1975), Mills (1975), Sneddon (1978), and Adelaar (1985) in Fernandez (1996). Furthermore, comparative studies of the history of languages in eastern Indonesia have also been conducted by Collins (1981), Wurm (1978), and Anceaux (1970), who studied languages in Papua and Maluku. This relates to previous statements that several languages show very close kinship with one another, such as the Ibo, Sahu, Waioli, and Gamkonora languages found in West Halmahera Regency (Andrews et al., 1996). Similar data is shown in the language map from the Ministry of Education and Culture's Language Agency (2025), which states that Sahu and Wayoli (Waioli) are two dialects of Sahu with a 69% difference. Sahu itself is a distinct language from Ibo, Galela, and Ternate. This statement still creates a gap regarding the status of languages in North Maluku, especially in West Halmahera. Research on languages in North Maluku using comparative historical linguistics is also still very limited. Based on this gap, this study focuses on the linguistic (phonological) characteristics of Sahu, Waioli, and Gamkonora, which show the kinship of the three languages, as well as the kinship of Sahu, Waioli, and Gamkonora, using quantitative (lexicostatistical) calculations.

Traditional linguistic genetics research, such as that conducted by Campbell (2013) and Heggarty (2010), confirms that language kinship can be analyzed through phonological, morphological, and lexical data. Campbell emphasizes the importance of regular sound correspondences and lexical evidence in proto-language reconstruction (*Historical Linguistics: An Introduction*, 3rd ed.). Meanwhile, in the Austronesian and Papalian regions, early comparative studies such as those by Nothofer (1975), Sneddon (1978), Collins (1981), and Wurm (1978) descriptively classified the languages of Indonesia, but these were generally macroscopic and did not focus on the internal relationships of three geographically close varieties. Several recent studies emphasize the importance of dialectal data and quantitative approaches in clarifying the disputed status of varieties as languages vs. dialects (e.g., Ross, Pawley & Osmond (1998) in Austronesian studies).

In general, the languages found in North Maluku are divided into two large families, namely the languages in southern Halmahera which are languages in the Austronesian family, such as Buli, Gane, Taba, Sula and the languages in the northern part of Halmahera which are included in the non-Austronesian family (SIL groups them into the West Papuan family), such as Ternate, Tidore, Tobelo, Sahu, Ibo, Waioli, and Gamkonora. The last four languages are very closely related in terms of sound and vocabulary. In fact, Waioli and Gamkonora have lexical similarities of 81% (SIL: 2005). This percentage identifies that Waioli and Gamkonora are one language, meaning that both languages are variations or dialects of one language. Ibo, Waioli, Gamkonora, and Sahu languages according to SIL (2005) have similarities in terms of form and meaning. However, the status of the four languages as distinct languages or dialects has not been determined. The North Maluku Provincial Language Office (2019) identified the four languages as dialects of the Sahu language, while the Ibo language has been

declared nearly extinct. However, according to speakers of each language, they use different languages. From a dialectological perspective, it is said that if two neighboring dialects are different, but the speakers of both dialects can still understand each other, then it can be said that the two dialects are one language. Conversely, if there is no mutual understanding, then the two isolects are not different dialects of the same language but are different languages (Mahsun: 1995). Another explanation from Kleden (Kleden & Imelda; Gamkonora & Waioli Language, n.d.) through research on the origins of languages in West Halmahera using several approaches, it is said that Gamkonora and Waioli are two different languages. From the several articles above, it is clear that there are still contradictions regarding the status of the languages in West Halmahera. Specifically, the status of the three languages (Waioli, Gamkonora, and Sahu) still requires further research. This research itself was conducted to examine the level of phonological similarity between the three languages. It is also clear that there is not much research that specifically focuses on the internal historical relationships between the three languages. Most studies are still descriptive, classificatory, or ethnolinguistic, so they have not yet produced a detailed picture of sound correspondence patterns, lexical retention, and shared innovation.

LITERATURE REVIEW

This study analyzes the kinship seen from the linguistic (phonological) characteristics of the three languages, Sahu, Waioli, and Gamkonora, and finds a precise percentage of kinship between the three languages using quantitative calculations (lexicostatistics). There is no empirical agreement regarding the status of the languages or dialects of Sahu, Waioli, and Gamkonora based on quantitatively measured phonological and lexicostatistical evidence. In other words, this study places the three languages within a more focused historical-comparative framework. The lexicostatistical approach is often used by language researchers, such as that conducted by Mulae and Sarif (2021), but only uses the similarity of 200 Swadesh vocabulary words, without considering the level of phonological similarity. This means that a lexicostatistical analysis has never been carried out consistently combining phoneme correspondence and patterns of innovation or phonological similarity, specifically for the three languages in a single study. Language research and documentation, including various elements such as dictionary development, language structure books, language identification, and language mapping, are important efforts to revitalize regional languages. One way to ensure language survival is by identifying languages and determining their status through comparative historical linguistic studies.

Comparative historical linguistics studies data from one or more languages and compares them to obtain the rules of change that occur in these languages which show their kinship. A pair of words being compared is considered a kin word if they are identical, differ by one phoneme, have phonetic similarities, and have phonemic correspondence, either regularly or irregularly or in varying ways (Keraf, Fitrah & Afria, 2017).

Comparative historical linguists assume that related languages have a historical relationship. Related words derived from a proto-language demonstrate: first, a large number of words from a particular language group

show relatively high similarities compared to other groups; second, phonetic changes in the history of certain languages also show regular characteristics; third, the deeper the history of related languages is traced, the more similarities there are between the items being compared (Keraf, 1996).

Thomason and Kaufman (Putri & Qalyubi, 2025) explain that language relatedness or language contact is one of the main drivers of linguistic change, especially in multilingual environments such as Indonesia. This phenomenon highlights the importance of studying language kinship to understand historical relationships. and the development of related languages. Historical relationships between languages can be traced through both quantitative and qualitative approaches, according to Fernandes (Mulae & Sarif, 2021).

The quantitative approach was conducted by utilizing data collection instruments in the form of the Morris Swadesh basic vocabulary list and lexicostatistical techniques to determine the level of kinship between the compared isoles. The lexicostatistical formula is as follows:

Number of Glosses

Description: Number of related words (identical, one phoneme different, corresponding, phonetically similar)

Number of Glosses (Vocabulary after subtracting loanwords and empty glosses from a total of 200 words)

N = percentage of related words

If the similarity level is above 81%, the two languages can be said to be the same, but if the percentage is below that figure, the two languages can be said to be different (Keraf, 1996). The following is a classification of the percentage of related words.

Table 1. language level and percentages of relatives

Language Level	Percentage of Relatives
Language	81-100
Family	37-80
Stock	12-36
Microphylum	4-11
Mesophylum	1-3
Macropylum	≤1%

METHODOLOGY

This research employed a qualitative descriptive method with a historical-comparative approach. According to Sugiyono (2007: 87), qualitative research generally involves both qualitative and quantitative data. Qualitative data itself is data in the form of words that have a descriptive basis for expressing or understanding phenomena in greater depth. Data were obtained using observation techniques, structured and unstructured interviews, as explained by Sudaryanto (1993: 3). In this technique, researchers are flexible, being both active (engaging in dialogue) and receptive (simply listening), as well as literature studies to support the data diachronically.

The research data consisted of basic vocabulary, pronouns, numerals, and several other lexical elements collected from native speakers of Waioli, Sahu, and Gamkonora in West Halmahera Regency. The data were analyzed using comparative techniques to identify cognates, sound correspondences, retentions, and shared innovations. The results of the analysis were then used to explain the historical relationship between the three languages. The data in this study consists of 200 basic vocabulary words from Blust's revised Morris Swadesh vocabulary, as well as additional phrases and sentences. All data were sourced from fieldwork at three locations: Sahu, Gamkonora, and Waioli sub-districts. The novelty of this research lies in its comprehensive comparative-historical investigation of the Sahu Language, Gamkonora Language, and Waioli Language languages within a single integrated framework. Previous studies have generally discussed these languages separately, focused only on sociolinguistic identity, or employed limited lexicostatistical approaches.

RESEARCH RESULT

The Sahu language (hereinafter abbreviated as BS), the Gamkonora language (hereinafter abbreviated as BG), and the Waioli language (hereinafter abbreviated as BW) are three regional languages found in West Halmahera Regency. These three languages are spoken by their speakers in adjacent areas. Sahu is spoken by the Sahu community in Sahu and East Sahu Districts, specifically in the villages of Worat-Worat, Tacici, Golo, Todahe, Awer, Hoku-Hoku, Idamgamlamo, and Taraudu. Gamkonora is spoken by its people in several villages around Mount Gamkonora in Ibu District, and Waioli is spoken by its speakers in the south of Ibu District, specifically in Tosoa Village and also in parts of Sahu District. The use of these three languages in such close areas is supported by phonological and lexical similarities, suggesting that they originally descended from a single language. This can be seen based on the following qualitative and quantitative studies.

DISCUSSION

Comparison of Linguistic Characteristics of Sahu, Gamkonora, and Waioli

The linguistic descriptions of Sahu, Gamkonora, and Waioli from a phonological perspective include vowels, diphthongs, consonants, vowel sequences, consonant sequences, and phonotactic systems. The comparison shows that all three languages have five vowels [a, e, i, o, u], which are distributed in the initial, medial, and final positions of words, except in BW, where [e] is not found in the initial position, as shown in the following example.

Table 2. linguistic characteristics

		Start	Middle	End
BS:	[a]	ayi 'tongue'	camu 'suck'	so'a 'leaves'
	[e]	eno 'skin'	pesa 'rain'	sae 'head'
	[i]	ijit 'tooth'	opid 'spit'	taki 'walk'
	[o]	orom 'eat'	ropos 'swell'	waro 'know'
	[u]	u'u 'fire'	dume 'kiss'	capu 'chew'
BG:	[a]	alo 'cold'	gasi 'salt'	ŋala 'moon'

	[e]	eno 'skin'	anel 'dream'	okay 'drink'
	[i]	isenj 'listen'	aid 'tongue'	godu 'eat'
	[o]	obonj 'bone'	la'o 'eye'	poloto 'bite'
	[u]	u'u 'fire'	tutu 'grow'	teu 'sit'
BW :	[a]		waro 'know'	faja 'dirty'
	[e]	adi 'crying'	yeno'o 'skin'	sae' head
	[i]	i'sene 'listen'	idi'i 'choose'	tag 'walking'
	[o]	toli 'steal'	lobos 'swell'	'cut' bibs
	[u]	udu 'mouth'	dudunj 'back'	wuru 'intestine'

Apart from vowels, in BS and BG there are also long vowel sounds which are located in both the final syllable (ultima) and prepenultima, while in BW there are no such long vowels, such as:

	BS	BG
[a:]	lala: 'good'	-
[e:]	-	nane: 'this'
[i:]	di:n 'sew'	-
[o:]	so:loro 'fly'	po:l 'stomach'

The diphthongs [au] and [oi] are found in all three languages, the diphthong [ai] only in BW and BG as in the following example.

Table 3. diphthongs [au] and [oi]

	BS	BG	BW
[au]	ηau 'ear'	sau 'hot'	ηau'u 'ear'
[ai]	-	bairere 'Spider	sa'ai 'cook'
[oi]	ηoi 'i'	ηamoi 'one'	ηoi 'i'

Apart from diphthongs, there are also vowel series as follows.

Table 4. vowel

	BS	BG	BW
[au]	ηaun 'blood'	woet 'laugh'	manaudu 'husband'
[oa]	soa 'leaf'	loane 'three'	ηoa'a 'people'
[oi]	koin 'throw'	loin 'hitung'	maroimi 'count'

The total number of consonants in BS, BG, and BW is 20, namely: [b, c, d, f, g, h, j, k, l, m, n, η, ŋ, p, r, s, t, w, y, and ?]. The distribution of consonants above occurs in all positions (initial, medial, and final) in BS and BG except [b, c, f, g, h, j, y, and ŋ], which are not found in final position. Meanwhile, in BW, consonants are found only in initial and medial positions. The phonotactic systems of BS and BG are the same: V, CV, VC, and CVC, while in BW, they are V, CV, and CVC. It can be said that words in BW are open syllables.

The phonological comparison between BS, BG, and BW can also be seen in the presence of corresponding and variant forms.

1) Correspondence and Variation Between BS and BG

The corresponding forms between BS and BG are only consonants, either perfect or imperfect. Variations can be either consonants or vowels. Examples of corresponding forms and variations are as follows.

a) Consonant Correspondence

Table 5. consonant correspondence

Korespondensi	BS	BG	glos
[k] ≈ [g] :#K(V)-	kosi	gosi	'egg'
	kumi	gumi	'rope'
	kane	gane	'lice'
[k] ≈ [g] :-K(V)#	taki	tagi	'walk'
	koko	gogo	'hair'
	kokon	gogon	'hide'
	leku	legu	'other'

The correspondence [k] ≈ [g], although less than perfect, occurs in two positions, namely the ultima and penultima, as illustrated in the example above. In addition, there is a correspondence [r] ≈ [l] which occurs in two positions, the ultima and penultima, as well as the penultima in three-syllable forms.

Table 6. penultima in three-syllable forms

	BS	BG	glos
[r] ≈ [l] :#K(V)-	rou	lou	'foot'
	ropos	lobos	'swollen'
	roaje	loaje	'three'
	rata	lata	'four'
	romdidi	lomodidi	'two'
[r] ≈ [l] :#-K(V)-#	koro:u	koloud	'dust'
	were'a	wele'a	'wife'
	peletor	pereto	'narrow'
	so:loro	solol	'fly'
[r] ≈ [l] :-K(V)#	tori	toli	'steal'
	lowor	powor	'smoke'
	wajer	wajel	'day'

The following correspondence [p] ≈ [b] also occurs in the ultimate position as in the example below.

Table 7. correspondence [p] ≈ [b]

	BS	BG	glos
[p] ≈ [b] :-K(V)#	o'pon	obon	'bone'
	opid	obid	'saliva'
	ropos	lobos	'swollen'

The correspondence [∅] ≈ [ŋ] occurs in the ultima as in the following example.

Table 8. correspondence [Ø] ≈ [ŋ]

	BS	BG	glos
[Ø] ≈ [ŋ] : -K(V)#	una	unaŋ	'he'
	ana	anaŋ	'they'
	maddi	madiŋ	'stone'

b) *Consonant Variation*

Consonant variation occurs in the variation of [n] ~ [ŋ] in the ultimate position as follows.

	BS	BG	glos
[n] ~ [ŋ] :-K(V)#	dudun	duduŋ	'back'
	ŋaun	laun	'blood'

Apart from the correspondence [p] ≈ [b] which occurs in the ultima, the consonants [p] and [b] are also variations in the penultimate position in three- and two-syllable words as follows.

	BS	BG	glos
[p] ~ [b] :#-K(V)-#	dupuso	tubuso	'heavy'
[p] ~ [b] :#K(V)-	peleas	beleas	'shoulder'
	pesa	besa	'rain'

c) *Vowel Variations*

Vowels do not have corresponding forms, but only variations, as follows.

Table 9. vowel variations

	BS	BG	glos
[i] ~ [Ø] : -(K)V#	paiti	pait	'dig'
	suŋi	suŋ	'new'
[e] ~ [Ø] : -(K)V#	isene	isen	'hear'
[u] ~ [e] : -(K)V#	malulu	malule	'lying down'
[a] ~ [o] :#(K)V-	sawo	sowo	'fruit'
	ŋawu	ŋowen	'when'

The vowel variations [i] ~ [Ø], [e] ~ [Ø], and [u] ~ [e] occur in the ultima while the variations [a] ~ [o] occurs in the pentate.

2) *Correspondence and Variation Between BS and BW*

Unlike BS and BG, which correspond only to consonants, in Sahu and Waioli there are forms with perfect and imperfect correspondences for both consonants and vowels. Likewise, the forms vary as follows

a) *Consonant Correspondence*

Table 10. consonant correspondence

Korespondensi	BS	BW	glos
[k] ≈ [g] :#K(V)-	kiyam	giyama	'hand'
	kubali	gubadi	'rope'
	kuwida	guwida	'lice'

	kodi	godi	'bite'
	ku'u	gu'u	'hold'
	kosi	gosi	'egg'
	koko	gogo	'book'
	kane	gane	'lice'
[p] ≈ [b] :-K(V)#	o'pon	obonjo	'bone'
	opid	obidi	'saliva'
	ropos	roboso	'swollen'
	lopi	lobidi	'night'
[p] ≈ [b] :#K(V)-	peleas	beleasa	'shoulder'
	pesa	besa	'rain'
	poko	boko'o	'short'
[∅] ≈ [w] :#K(V)-	o'e	woe	'drink'
	ate	wate	'wood'
	utu	wutu'u	'rootr'
	u'u	wu'u	'fire'

The correspondence [k]≈[g] is a perfect correspondence for the penultimate and some prepenultimate positions, as in the examples above. Some disyllabic words in the SL become trisyllabic words in the SL with the addition of a final vowel, such as *giyam* with *giyama*, meaning 'hand', *lopi* with *lobidi* 'night', *poko* with *boko'o* 'short', and *ropos* with *roboso* 'swollen'. The correspondence [p] ≈ [b] occurs in both the ultima and penultimate positions. In the ultima, the correspondence [p] ≈ [b] in the SL is for disyllabic words, while in the SL it is for trisyllabic words. The correspondence [∅] ≈ [w] occurs in the penultimate position, as seen in the examples above.

b) Consonant Variation

In addition to the corresponding forms between the SL and SL, there are also variations as follows.

Table 11. consonant variation

	BS	BW	glos
[d] ~ [t] :#K(V)-	dupuso	tobuso	'heavy'
	dume	tumese	'kiss'
[∅] ~ [y] :#K(V)-	eno	yeno'o	'skin'
	eta	yeta'a	'fall'
[c] ~ [j] :-K(V)#	mocon	mojonjo	'Afraid'
[w] ~ [p] :#K(V)-	wari'u	pari'u	'tie'
[h] ~ [w] :#K(V)-	hutu	wutu	'hair'

As previously stated, there are corresponding forms and variations in the BS and BW that occur between two-syllable words in the BS and three-syllable words in the BW. In the BW, the addition of a final vowel (paragog) often occurs. Similarly, there is variation in [c] ~ [j] in ultima in the BS and penultima in the BW in three-syllable words.

c) *Vowel Correspondence*

The vowel correspondence found in the comparison of BS and BW is [Ø] in the BS with the vowel in the BW in ultima, as shown in the following data.

Table 12. vowel correspondence

	BS	BW	glos
[Ø] ≈ [o] : -(K)V#	nelo	nelo'o	'laugh'
	eno	yeno'o	'skin'
	orom	oro'mo	'Eat'
	teos	teoso	'stand'
	karaŋos	karaŋoso	'scratch'
	ropos	roboso	'swollen'
	nao	nao'o	'fish'
	sawo	sowo'o	'buah'
	lowor	loworo	'smoke'
	lamo	lamo'o	'big'
	poko	boko'o	'short'
[Ø] ≈ [u] : -(K)V#	dudun	duduŋu	'nose'
	ŋunuŋ	ŋunuŋu	'nose'
	ŋau	ŋau'u	'ear'
	tutu	tutu'u	'pound'
	muduŋ	muduŋu	'star'
[Ø] ≈ [a] : -(K)V#	camal	camala	'neck'
	Peleas	beleasa	'hair'
	ŋunaŋ	ŋunaŋa	'vomit'
	bab	baba	'father'
	lomaŋ	lomaŋa	'name'
	eta	yeta'a	'fall'
	baŋan	baŋana	'forest'
	diwaŋ	diwaŋa	'sky'
	muras	murasa	'thunder'
	ŋinar	ŋinara	'thin'
	roat	roata	'wide'
[Ø] ≈ [e] : -(K)V#	sae	sae'e	'head'
	Dume	tumese	'kiss'
	nanel	nanele	'dream'
	senen	senene	'dead'
	waneŋ	waneŋe	'day'
[Ø] ≈ [i] : -(K)V#	opid	obidi	'saliva'
	di:n	dini	'sew'
	di:m	di'imi	'tail'
	kapiriŋ	kapiriŋi	'thick'
	tori	tori'i	'steal'

The vowel correspondence in the ultima between BS and BW occurs in all vowels [a, e, i, o, and u] as in the example above.

d) *Vowel Variation*

Table 13. Vowel Variation

	BS	BW	glos
[a] ~ [i] : #K(V)-	ɲana	ɲini	'engkau'
[a] ~ [e] : #K(V)-	pari'i	peri'i	'old'

The vowel variations [a] ~ [i] and [a] ~ [e] between BS and BW occur in the pentima, as seen in the two examples above.

3) *Correspondence and Variation Between BG and BW*

Comparison between BG and BW shows a correspondence between consonants and vowels. Variation occurs only in consonants, as shown below.

a) *Consonant Correspondence*

Table 14. consonant correspondence

	BG	BW	glos
[l] ≈ [r] : #K(V)-	lou	rou	'foot'
	lobos	roboso	'swollen'
	lata	rata	'four'
[l] ≈ [r] : -K(V)#	toli	tori'i	'steal'
	ɲala	ɲara	'month'
	pele	pere	'go on'
	solol	soloro	'fly'
	gudal	gudara	'grass'
[∅] ≈ [w] : #K(V)-	utu	wutu	'hair'
	oke	wo'e	'drink'
	u'u	wu'u	'fire'

The correspondence of [l] and [r] between BG and BW occurs in both penultimate and ultimate tenses, as in the examples above. However, in some ultimas, the corresponding form is at the end of the ultima, and in BW, it occurs in trisyllabic words. This occurs because in BW, the addition of a vowel at the end of a word (paragog) is often found compared to BG. This can be seen in the word (BG) solol with (BW) soloro 'fly' and the word (BG) gudal with gudara 'grass'.

b) *Consonant Variation*

Consonant variations are only found in the consonants [g] and [k] in the penultimate, as well as the nazal consonants as in the following example.

Table 15. consonant variation

	BG	BW	glos
[g] ~ [k] : #K(V)-	gaud	kaudu	'long'
	goloud	koroudu	'ash'
[ŋ] ~ [w] : -K(V)-	gɲida	guwida	'kanan'
[m] ~ [w] : #K(V)-	mate	wate	'wood'
[n] ~ [m] : -K(V)#	gane	game	'lice'
[n] ~ [ŋ] : #K(V)-	nage	ɲage	'that'
	nane	ɲane	'this'

Nasal consonants vary with both nasal and approximant sounds, as seen in the examples above.

c) *Vowel Correspondences*

The corresponding vowel forms in BG and BW can be seen as follows.

Table 16. vowel correspondences

	BG	BW	glos
[Ø] ≈ [u] : -(K)V#	dudun̩	dudun̩u	'panjang'
	ɲunun̩	ɲunun̩u	'nose'
	ɲau	ɲau ^ʔ u	'ear'
	tutu	tutu ^ʔ u	'pound'
[Ø] ≈ [e] : -(K)V#	sae	sae ^ʔ e	'head'
	nanel	na ^ʔ nele	'dream'
	seɲen	seɲene	'dead'
[Ø] ≈ [i] : -(K)V#	aid	a ^ʔ idi	'tongue'
	wiwis	wowisi	'suck'
	idi	idi ^ʔ i	'suck'
[Ø] ≈ [a] : -(K)V#	camal	camala	'leher'
	beleas	beleasa	'shoulder'
	ɲoa	ɲoa ^ʔ a	'child'
[Ø] ≈ [o] : -(K)V#	obon̩	obon̩o	'bone'
	te ^ʔ os	te ^ʔ oso	'stand'
	solor	soloro	'fly'

The zero form in BG in the two-syllable ultima corresponds to all the vowels [u, e, i, a, and o] in the ultima in three-syllable words, as seen in the example above.

A comparison of BS, BG, and BW, with their innovative forms in terms of phoneme distribution, long vowel forms, and corresponding and variant forms, shows that BS and BG share similarities in phoneme distribution and both have long vowels. BW differs in that it lacks long vowels and the vowel [e] is not present at the beginning of words. Unlike BS and BG, Waioli has an open syllable in the ultima. Therefore, BW has a phonotactic system of V, CV, and CVC in the penultima, while BS and BG have a phonotactic system of V, CV, VK, and CVC in both the ultima and the penultima.

The corresponding form can be seen in BG and BW showing an innovative form, namely the presence of the consonant [k] in BS corresponding to [g] in BG and BW both in the ultima and penultima. Likewise, the vowels in BS and BG in words ending in consonants in these two languages will correspond to the vowels in BW. This means that the consonants at the end of words in BS and BG will always get an additional vowel at the end of the word in BW. [Ø] in BS and BG corresponds to all vowels in BW in the ultima.

Table 17. BS and BG corresponds

	BG	BW	glos
[Ø] ≈ [u] : -(K)V#	duduŋ	duduŋu	'long'
[Ø] ≈ [e] : -(K)V#	sae	sae'e	'hesd'
[Ø] ≈ [i] : -(K)V#	aid	a'idi	'tongue'
[Ø] ≈ [a] : -(K)V#	camal	camala	'neck'
[Ø] ≈ [o] : -(K)V#	obonŋ	obonŋo	'bone'
[Ø] ≈ [o] : -(K)V#	nelo	nelo'o	'laugh'
[Ø] ≈ [u] : -(K)V#	tutu	tutu'u	'pound'
[Ø] ≈ [e] : -(K)V#	wanŋer	wanŋere	'day'
[Ø] ≈ [i] : -(K)V#	opid	obidi	'saliva'
[Ø] ≈ [a] : -(K)V#	diwanŋ	diwanŋa	'sky'

The zero form in BG in the two-syllable ultima corresponds to all the vowels [u, e, i, a, and o] in the three-syllable word, as shown in the example above.

A comparison of BS, BG, and BW, with their innovative forms in terms of phoneme distribution, long vowel forms, and corresponding and variant forms, shows that BS and BG share similarities in phoneme distribution and both have long vowels. However, BW differs in that it lacks long vowels and the vowel [e] is not present at the beginning of words. Although BW has undergone a different development compared to BS and BG, the kinship between BS and BW shows a closer kinship. This is consistent with the following lexicostatistical calculations.

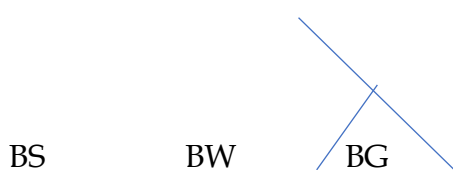
Relationship between BS, BG, and BW using Lexicostatistical Calculations

Based on lexicostatistical calculations using a list of two hundred basic Swadesh vocabulary words, the percentage of Sahu cognates with Gamkonora is 64%, Sahu cognates with Waioli cognates with Waioli cognates with Gamkonora.

Table 18. BS, BG, and BW using Lexicostatistical Calculations

estimation (%)	Bahasa Waioli (BW)	Bahasa Gamkonora (BG)
Bahasa Sahu (BS)	71	64
Bahasa Waioli (BW)		65

The results of the percentage of cognate words indicate that the kinship relationship between Sahu and Waioli is closer than the kinship relationship between Sahu and Gamkonora or Waioli and Gamkonora. The close kinship between BS and BW is also shown geographically that BW is also spoken in several villages which are Sahu-speaking areas. The kinship status of Sahu with Waioli, Sahu with Gamkonora, and Waioli with Gamkonora is a language family relationship. The kinship tree between the three languages is as follows.
Prabahasa SGW



The prelingual SGW first split into BS and BG. This is indicated by the presence of *k (PSGW) in BS, where [k] corresponds to [g] in BG and BW, as in the word *taki* (BS) with *tagi* in (BG and BW). BW subsequently developed, with the most common characteristic being the addition of a vowel at the end of the word (paragog), as in the word *te'os* (BS and BG) with *te'oso* (BW), meaning 'to stand'.

Comparing BS with BG, BS with BW, and BG with BW, it can be seen that BS and BG still retain the same forms as BW. In this case, BW has undergone innovation in several ways, such as the addition of phonemes in the middle of words (epenthesis), at the end of words (paragog), and at the beginning of words (prothesis).

a) Prothesis

Prothesis is the addition of phonemes to the beginning of words. The PSGW * \emptyset found in BS and BG will appear as the vowel [w] in BW, as in **u'u* (BS, BG) with *wu'u* (BW), **o'e* /*oke* 'drink' (BS, BG) with *wo'e* (BW). Likewise, the PSGW * \emptyset in BS **eta* 'fall' with *yeta'a* in BW.

b) Epenthesis

Epenthesis is the addition of a phoneme in the middle of a word. The PSGW * \emptyset in BS **romdidi* 'two' will appear in BG as [o] in *lomodidi*.

c) Paragoge

Paragoge is the addition of a phoneme to the end of a word. This form of word-ending addition is common in BW, with PSGW * \emptyset in BS and BG appearing as the vowels [a, e, i, o, and u] in BW. For example, the word **camal* (BS, BG) appears as *camala* 'neck' (BW). The word **opid/obid* (BS, BG) appears as *opidi* (BW).

CONCLUSIONS AND RECOMMENDATIONS

The results of this analysis were conducted by combining sound correspondences (phonology) and lexicostatistics within a single analytical framework. The finding that BW undergoes systematic prothesis, epenthesis, and paragoge is novel in this local study. Previous research has not demonstrated systematic phonological differences between these varieties. The results of this study are not merely descriptive research or observations, but also contribute typologically to the study of cognate languages, which are often claimed as distinct languages by their speakers in the West Halmahera region. In more detail, the results of this study show that between BS and BW languages have the highest percentage of cognates (71%), while lexically located between BS and BW around (65%). The existence of this correspondence evidence shows support for the interpretation of the kinship of the three languages. In other words, this study can be said to provide a typological contribution by proving quantitatively that the three are one language family. From the findings in the field, it is advisable to conduct linguistic data collection and revitalization of languages in West Halmahera. Considering that one language (Ibo) has been declared extinct.

ADVANCED RESEARCH

The historical relationship among the Sahu, Gamkonora, and Waioli languages demonstrates a strong genetic connection rooted in the North Halmahera Papuan linguistic tradition. Comparative evidence from phonology, lexicon, morphology, and syntax supports the hypothesis that the three languages descended from a common ancestral language while evolving through distinct sociocultural and geographical processes.

The research further establishes that Sahu and Waioli maintain a closer linguistic affinity, whereas Gamkonora reflects more extensive innovation and external influence. Continued documentation and reconstruction efforts are essential for preserving the linguistic heritage of West Halmahera and deepening scholarly understanding of Eastern Indonesian language history.

ACKNOWLEDGMENT

First and foremost, the researcher would like to express sincere gratitude to all those who contributed to the completion of this research entitled "*Historical Relationship of the Sahu, Gamkonora, and Waioli Languages in West Halmahera, North Maluku.*" Special appreciation is extended to the academic supervisors, lecturers, and researchers whose guidance, insights, and constructive suggestions greatly enriched this study.

The researcher is deeply grateful to the native speakers and community leaders of the Sahu, Gamkonora, and Waioli speech communities who generously shared linguistic data, cultural knowledge, and valuable time during the fieldwork process. Without their cooperation and hospitality, this research would not have been possible.

The researcher also acknowledges the support provided by the institution, colleagues, friends, and family members who offered encouragement, assistance, and motivation throughout the research process. Their moral and practical support has been invaluable in completing this study successfully.

Finally, the researcher hopes that this work will contribute to the development of comparative historical linguistics, particularly in the study of regional languages in West Halmahera and the preservation of local linguistic heritage in North Maluku.

REFERENCES

- Andrews, Beth A. dkk 1996. *Atlas Bahasa Tanah Maluku*. Pusat Pengkajian dan Pengembangan Maluku Universitas Pattimura dan Summer Institute of Linguistics.
- Badan Bahasa Pengembangan dan Pembinaan bahasa. 2025 Peta Persebaran Bahasa
<https://petabahasa.kemendikdasmen.go.id/bahasasastra/infobahasa.php?id=491&id2=0>
- Fernandez, Inyo Yos. 1986. *Relasi Historis Kekerabatan Bahasa Flores*. Flores Nusa Tenggara Timur: Nusa Indah.
- Fitrah, Yundi dan Rengki Afria. 2017. *Kekerabatan Bahasa-Bahasa Etnis Melayu, Batak, Sunda, Bugis, dan Jawa di provinsi Jambi*: Sebuah Kajian Linguistik

- Historis Komparatif. Jurnal Titian. FIB, Universitas Jambi.
<https://repository.unja.ac.id/2511/>
- Kantor Bahasa Provinsi Maluku Utara. 2022. *Kantor bahasa Malut Memetakan 19 Bahasa Daerah, Ada yang terancam Punah*.
<https://rri.co.id/daerah/87291/kantor-bahasa-malut-petakan-19-bahasa-daerah-ada-yang-terancam-punah>
- Keraf, Gorys. 1996. *Linguistik bandingan Historis*. Jakarta: Gramedia Pustaka Utama.
- Kleden Ninuk, Imelda (2011). Gamkonora dan Waioli; Bahasa dalam Konstruksi Identitas Etnik. <https://media.neliti.com/media/publications/150364-ID-gamkonora-dan-waioli-bahasa-dalam-konstr.pdf>
- Lestari, D. A., Akbarjono, A., & Heriadi, M. (2024). *Quantifying dialect relatedness in Serawai (Bengkulu, Indonesia): A 200-item lexicostatistical study*. *Jurnal Pustaka Indonesia*, 5(3), 392 (DOI online).
<https://doi.org/10.62159/jpi.v5i3.392> (Siducat)
- Mahsun. 1995. *Dialektologi Diakronis: Sebuah Pengantar*. Yogyakarta: Gadjah Mada University Press.
- Putri, J. A., & Qalyubi, I. (2025). *Lexical relationship between Dayak Ngaju and Dayak Sampit languages: A lexicostatistical study*. *JiIP - Jurnal Ilmiah Ilmu Pendidikan*, 8(7), 8347–8353. <https://doi.org/10.54371/jiip.v8i7.8715> (Iiip)
- Rizki, Fahmi Anggia dan Dwi Widyawati. 2021. *Perubahan Bunyi Bahasa Proto Austronesia Ke Dalam Bahasa Jawa Dialek Sumatra: Kajian Linguistik Historis Komparatif*.
<https://ejournal.warmadewa.ac.id/index.php/kulturistik/article/view/3380>
- Saputri, T. D., Misrita, & Qalyubi, I. (2026). *Tracing historical linguistic kinship: A lexicostatistical study on Dayak Ngaju and Banjar Kuala languages in Kalimantan*. *English Journal of Indragiri: Studies in Education, Literature, and Linguistics*, 10(1), 160–177. <https://doi.org/10.61672/eji.v10i1.3064> (ejournal-fkip.unisi.ac.id)
- Sintia, D., Qalyubi, I., & Misrita. (2025). *A lexicostatistical study: Language kinship of Dayak Ngaju and Dayak Maanyan language*. *English Journal of Indragiri: Studies in Education, Literature, and Linguistics*, 9(2), 486–497. <https://doi.org/10.61672/eji.v9i2.3001> (ejournal-fkip.unisi.ac.id)
- Suciati, T. A., Qalyubi, I., & Misrita. (2026). *Comparing basic words and using lexicostatistics to identify language connections among Indonesian, Dayak Ngaju, and Javanese*. *English Journal of Indragiri: Studies in Education, Literature, and Linguistics*, 10(1), 1–9. <https://doi.org/10.61672/eji.v10i1.3010> (ejournal-fkip.unisi.ac.id)
- Sudaryanto. 1993. *Metode dan Teknik Analisis bahasa* (<http://teknikcakapdantekniksimak.blogspot>)
- Sugiono. 2007. *Metode Penelitian Kuantitatif, Kualitatif*. Bandung: alfabeta
- Summer Institute of Linguistics. (2006). *Bahasa-Bahasa di Indonesia*. SIL International Indonesia Branch.