

# Lengilo and Lun Bawang Languages of Sarawak are Sister Languages: Lexicostatistic Study

Hamidah Abdul Wahab<sup>1\*</sup>, Nur Ardini Jian Abdullah<sup>1</sup>, Su-Hie Ting<sup>1</sup>

<sup>1</sup> Faculty of Education, Language and Communication, Universiti Malaysia Sarawak, Kota Samarahan, Malaysia

\*Corresponding Author: [awhamidah@unimas.my](mailto:awhamidah@unimas.my)

Received: 22 February 2025 | Accepted: 6 June 2025 | Published: 30 June 2025

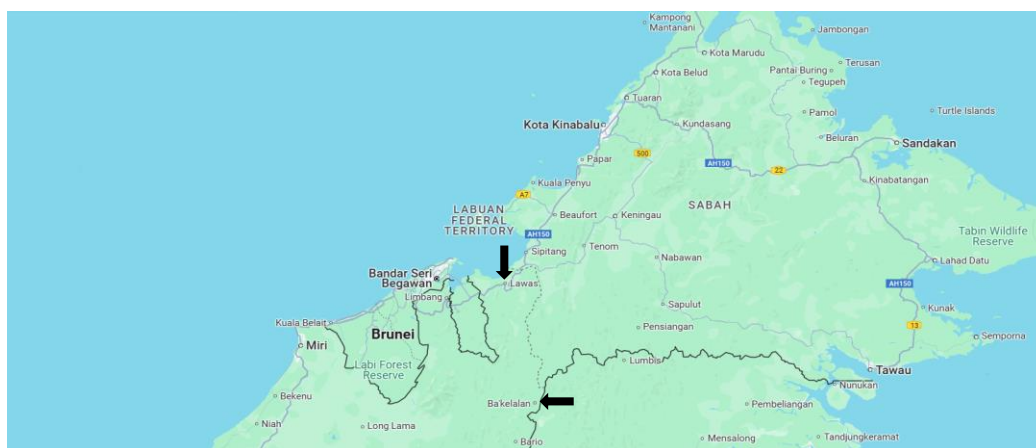
DOI: <https://doi.org/10.55057/ajress.2025.7.5.35>

**Abstract:** *The Lengilo is one of the minority indigenous groups living in Ba’Kelalan and Lawas, Sarawak. As most people are not aware of the existence of the group, the Lengilo people identify themselves as Lun Bawang. Intermarriages bring the two distinct groups together but it is not known how similar their languages are. The study compares the basic vocabulary of the Lengilo and Lun Bawang languages to determine the level of kinship. The data were collected through interviews with 10 native speakers of Lengilo and Lun Bawang in Ba’Kelalan and Lawas using a list of 100 Swadesh vocabulary items. The 10 informants, consisting of men and women, were aged 56 to 68. The lexicostatistical method was used to identify the percentage of cognate words involves comparing the similarities and differences in the vocabulary of these languages based on Gudschinsky’s (1956) proposed language kinship level table. The results show 94.6% similarity in the form and meaning of Lengilo and Lun Bawang words. The findings suggest that since the divergence between Lengilo and Lun Bawang is minimal, they probably come from the same ancestor language, likely even forming part of a dialect continuum.*

**Keywords:** Lengilo, Lun Bawang, basic vocabulary, lexicostatistics, Sarawak

## 1. Introduction

Lengilo is spoken by the Lengilo people, who reside mostly in the Lawas and Ba’Kelalan areas of Sarawak (Figure 1).



**Figure 1: Geographical location of Ba’Kelalan and Lawas**  
(Source: Google maps)

Lawas is a small town and the capital of the Lawas District located in Limbang Division of northern Sarawak. The geographical coordinates of Lawas are 4° 51' 37.44" N and 115° 24' 32.4 E. Lawas is located along the Lawas River and is only 26 metres above sea level. Lawas can be reached by air, water or land.

On the other hand, Ba'Kelalan is an interior town, located close to the border of Sarawak and Kalimantan. The geographical coordinates of Ba'Kelalan are 3° 59' 0" North, 115° 37' 0" East. Ba'Kelalan is 910 metres above sea level, and can be reached by air or land (about 6 hours of journey via four-wheel drive from Lawas in the year 2024).

The Lengilo migrated from the Highlands of Krayan in Kalimantan, Indonesia before the Japanese Occupation of Sarawak in 1941-1945. In Indonesia, they live near the Krayan River, mostly in the Central Krayan and Upper Krayan areas (Tirusel & Da'a, 2021). As most people are not aware of the existence of the group in Sarawak, the Lengilo people identify themselves as Lun Bawang in Sarawak and Lun Dayeh in Sabah.

The Lengilo population is estimated at 120 based on the number of members in a Lengilo WhatsApp group but this also includes Lengilo from mixed marriages. It is difficult to obtain official population statistics for Lun Bawang and Lun Dayeh, let alone Lengilo, because they are grouped as "Other Indigenous" in Department of Statistics Malaysia record. However, some estimates are obtained from other sources. The Lun Bawang population is 17,676 in Sarawak and 6,941 in Sabah based on the 2020 census (Wikipedia, 2025)<sup>1</sup>. The Lun Bawang people in Sarawak also self-label themselves as Murut (Langub, 1987). Interestingly, in Sabah the Lengilo people identify themselves as Lun Dayeh (Eghenter & Jok, 2012). The Lun Dayeh population in Sabah is estimated at 10,000 and mostly living in Sipitang district (Sidek, 2022). To place these population figures in context, the Sarawak population in 2020 is 2.8165 million (see Appendix 1).

Based on Jian Abdullah et al.'s (2024) preliminary study, the Lengilo language has six types of consonants, three types of vowels and diphthongs. There are six consonant types in Lengilo: plosive, nasal, fricative, trill, lateral, and partial vowel. These consonants do not consistently appear at the beginning, middle, and end of words. For instance, the velar plosive [g], alveolar fricative [s], glottal fricative [h], and lateral-alveolar [l] illustrate this selective distribution. Lengilo has eight vowel sounds: the narrow front vowel [i], semi-narrow front vowels [e], semi-wide front vowels [ɛ], wide front vowels [a], a semi-wide central vowel [ə], the narrow back vowel [u], semi-narrow back vowels [o], and semi-wide back vowels [ɔ]. In terms of diphthongs, the language includes [ai], [ui], and [oi]. Among them, [ai] appears to occur more frequently, while [ui] and [oi] are relatively rare.

The Lengilo is a small community living within a larger Lun Bawang (Sarawak) or Lun Dayeh community (Sabah), which is why they have to assimilate. The assimilation to the local communities may take the form of culture such as way of life, food, and customs, as well as language. Intermarriages bring the two distinct groups together but it is not known how similar their languages are. In this regard, there is a need to study the kinship or cognate between Lengilo and Lun Bawang languages.

This study compares the basic vocabulary of the Lengilo language with the Lun Bawang language to determine the level of kinship.

## 2. Literature Review

### 2.1 Swadesh list

The Swadesh list is a set of basic vocabulary terms found across all languages, designed for use in lexicostatistics and glottochronology to assess genealogical relationships and estimate language divergence. Named after the linguist Morris Swadesh (1950, 1952, 1955), the Swadesh list includes universally shared concepts such as water, star, and hand. Researchers compare Swadesh list translations across languages to quantify linguistic similarity and infer degrees of relatedness between the languages.

The Swadesh list was originally compiled by Morris Swadesh based on intuitive selection of culturally neutral, universally available vocabulary, without prioritising word stability. However, stability has since been recognised as relevant for glottochronology, and later analyses (e.g., Lohr, 2000) have assessed the resistance of these terms to change. More recent lists, such as the Dolgopolsky (1964) and Leipzig-Jakarta (Sakel & Everett, 2012) lists, are based on cross-linguistic data but are not widely used. Swadesh's (1971) final 100-word list is the most commonly used reference.

Having established the purpose and structure of the Swadesh list, the next step in comparative analysis involves identifying cognate word pairs—words in different languages that share a common historical origin. By examining which Swadesh list terms are cognates across languages, researchers can assess the depth of linguistic relatedness, quantify lexical similarity, and trace patterns of divergence. This process forms the core of lexicostatistical comparison and provides the empirical basis for evaluating language relationships.

Cognate word pairs may not be exactly the same in two languages in terms of form and meaning. Based on Mohamed (1999), there are at least three types of cognate word pairs:

1. Identical word pairs are pairs that show 100% similarity in form (phonemes and the arrangement of phonemes), and 100% similarity in meaning;
2. Similar word pairs are pairs that are 100% the same in form but similar in meaning, similar in form but 100% the same in meaning, and similar in form and meaning; and
3. Corresponding word pairs are pairs with phonemes that 100% the same in the same position, and must correspond in meaning.

Phoneme correspondence, particularly the position of sounds within a word, plays an important role in identifying cognate word pairs. While some word pairs may appear similar due to having identical or nearly identical phonemes, such similarities can often be attributed to regular patterns of sound change influenced by the linguistic environment. These patterns are not random; they tend to occur systematically and repeatedly across related words, providing evidence of a historical relationship between languages.

In most Swadesh-based or lexicostatistical studies, the focus is on true (genetic) cognates. However, when identifying cognate pairs, cognates that may distort the measurement of linguistic relatedness include loanwords, partial cognates, and false cognates.

1. Loanwords are words that are similar due to borrowing, not shared genetic ancestry. For example, *sekolah* is a word in Malay and Iban but this word is not native to either language. It originated from the Portuguese word, *escola*, and entered the Malay language during the colonial period, and subsequently entered the Iban language through language contact;

2. Partial cognates are words pairs of words in the same language that come from the same original root (usually from another language), but entered the language in different ways — one might be borrowed, and the other inherited or altered internally. For example, both “cow” and “beef” refer to the same animal and share the ancient Indo-European root but “cow” is an old English word while “beef” is borrowed from the French word, *boeuf*; and
3. False cognates are words that look or sound similar across languages but are not related in origin. For example, the English word “air” means the air we breathe while the Malay word “*air*” means water we drink. They sound identical but have completely different meanings.

For the comparison of the basic vocabulary of the two languages, cognate words are counted. Next, to obtain the percentage of cognate words, the number of cognate words is divided by the number of words compared. The number of words being compared is the number of words from which loanwords, empty words, and compound words have been excluded, that is, words that are not needed (Mohamed & Abdul Wahab, 2004; Mohamad & Mohd Yusof, 2004).

## 2.2 Lexicostatistical comparison of languages

In this section, several lexicostatistical studies are reported to show the lack of studies on the similarity of cognate words among these Greater North Borneo languages.

In West Malaysia, lexicostatistical studies have been conducted to compare the similarity of Malay and Banjar language spoken in Kalimantan (Abdul Wahab, 2022; Abdul Wahab & Che Halin, 2021). The researchers found 91.6% similarity in cognates. Kasim (2021) found that the similarity between old Malay and Jakun language is 80.77%. Muda and Hamzah (2021) found 76% cognate similarity between Malay and Che Wong language (part of Orang Asli Senoi), and concluded that the two languages probably diverged five to 25 centuries earlier. The studies were conducted using either the 100 or 200 Morris Swadesh vocabulary lists, and involved locations in Peninsular Malaysia. The studies also concentrated on making comparisons with Malay.

Some lexicostatistical studies have been conducted in Sarawak to compare similarity of Malay and indigenous languages. Jimbai et al. (2024) compared Malay and Iban vocabulary and found that all derived words that have the same meaning are combined with different morphemes to form prefixes, suffixes, and infixes. Malay and Iban have different word formation processes. Jimbai et al. (2024) built upon the work of Omar (2008) who found 63% similarity between Malay and Iban. According to Blust (2010), the languages in Greater North Borneo are a subgroup of “Western Indonesia” which includes all Austronesian languages in mainland Southeast Asia, Madagascar, and the Greater Sunda Islands, but not the languages of Sulawesi. Mohamed and Mohd Yusof (2004) places the similarity of Malay and Iban at 68%, slightly higher than Omar (2008). It is not certain whether the word list compared differs or whether the increased similarity is due to changes in the vocabulary of the two languages over time.

Mohamed and Abdul Wahab (2004) examined the kinship level of Malay and Bidayuh dialects, and concluded that the four Bidayuh dialects originated from the same language. The Bidayuh dialect most similar to Malay is Selako-Lara spoken in Lundu (78.4%) while the other three Bidayuh dialects have slightly less than 50% similarity with Malay: Bukar-Sadong, 43.8%; Biatah, 45.3%, Bau-Jagoi, 47.9%. Selako-Lara belongs to the Malayic Dayak subgroup, the same group that includes Iban, Kendayan, and Malay itself. However, the other Bidayuh dialects belong to the Land Dayak subgroup of Austronesian languages.

The similarity of several other indigenous languages to Malay have also been studied. Mohamed and Mohd Yusof (2004) analysed the basic vocabulary of 10 Austronesian languages in Sarawak, namely, Malay, Iban, Bidayuh, Selakau, Melanau, Miriek, Bintulu (Segan), Bukit/Buket, Kiput, and Lun Bawang. Table 1 shows the cognate similarities among 10 languages spoken in Sarawak. Based on Mohamed and Mohd Yusof (2004), Melanau is only 34%-46% similar to Lun Bawang, Kiput, Bukit, Bintulu, and Miriek. However, Miriek and Kiput are 69.4% similar. Miriek and Kiput languages are similar because they both belong to the North Sarawakan subgroup of the Austronesian language family, specifically within the Berawan–Lower Baram group. This means they share a relatively recent common ancestor and have likely been in close geographic and cultural contact for centuries. Miriek and Kiput can be considered sister languages.

**Table 1: Cognate similarities among 10 languages spoken in Sarawak**

Bahasa	B Melayu	B Iban	B Bidayuh	B Selakau	B Melanau	B Miriek	B Bintulu	B Bukit	B Kiput
<b>B Lunbawang</b>	41.23	37.2	32.9	38.1	40.2	41.8	40.2	38.5	39.1
<b>B Kiput</b>	40.2	35.1	29.8	30.9	34.0	69.4	40.2	34.0	
<b>B Bukit</b>	33.3	35.1	34.0	34.0	41.7	30.2	37.5		
<b>B Bintulu</b>	51.0	40.4	35.4	43.3	46.4	46.4			
<b>B Miriek</b>	45.4	38.4	32.9	43.3	42.3				
<b>B Melanau</b>	50.0	44.7	40.6	46.9					
<b>B Selakau</b>	77.3	67.0	40.2						
<b>B Bidayuh</b>	42.7	36.8							
<b>B Iban</b>	68.0								

The literature review shows that the kinship relationship between Lengilo language and Lun Bawang languages are not known, and needs to be studied to understand the similarities in core vocabulary (e.g., body parts, natural elements), phonological systems, and grammar and syntax.

### 3. Research Method

The data were obtained through fieldwork in the Ba’Kelalan and Lawas districts of Sarawak. Ten informants consisting of men and women aged 56 to 68 years were involved. Interviews were conducted by the research team based on a list of 100 Swadesh vocabulary items, which involved basic vocabulary related to ordinary life. Table 2 shows the Morris Swadesh basic vocabulary list in English and Malay but only the Malay list was used during the fieldwork. The present study focusses on the 100 Swadesh word list, and the words are in the singular such as “tooth” instead of “teeth”.

**Table 2: Swadesh Basic Vocabulary List (in English and Malay)**

No.	English	Malay
1.	all	semua
2.	ashes	abu
3.	bark (of tree)	kulit pokok
4.	belly	perut
5.	big	besar
6.	bird	burung
7.	bite	gigit
8.	black	hitam
9.	blood	darah
10.	bone	tulang

11.	breast	susu
12.	burn	terbakar
13.	nail	kuku
14.	cloud	awan
15.	cold	sejuk/dingin
16.	come	datang
17.	die	mati
18.	dog	anjing
19.	drink	minum
20.	dry	kering
21.	ear	telinga
22.	earth	tanah
23.	eat	makan
24.	egg	telur
25.	eye	mata
26.	fat	lemak
27.	feather	bulu
28.	fire	api
29.	fish	ikan
30.	fly	terbang
31.	full	penuh
32.	foot	kaki
33.	give	beri
34.	good	baik
35.	green	hijau
36.	hair	rambut
37.	hand	tangan
38.	head	kepala
39.	hear	dengar
40.	heart	jantung
41.	horn	tanduk
42.	I	aku
43.	kill	bunuh
44.	knee	lutut
45.	know	tahu
46.	leaf	daun
47.	lie down	berbaring
48.	liver	hati
49.	long	panjang
50.	louse	kutu
51.	man-male	lelaki
52.	many	banyak
53.	meat-flesh	daging
54.	moon	bulan
55.	mountain	gunung
56.	mouth	mulut
57.	name	nama
58.	neck	leher
59.	new	baru
60.	night	malam
61.	nose	hidung
62.	not	tidak
63.	one	satu
64.	person	orang
65.	rain	hujan
66.	red	merah
67.	road	jalan
68.	root	akar
69.	round	bulat
70.	sand	pasir



71.	say	sebut/ungkap
72.	see	lihat
73.	seed	benih
74.	sit	duduk
75.	skin	kulit
76.	sleep	tidur
77.	small	kecil
78.	smoke	asap
79.	stand	berdiri
80.	star	bintang
81.	stone	batu
82.	sun	matahari
83.	swim	berenang
84.	tail	ekor
85.	that	itu
86.	this	ini
87.	you	kau/kamu/engkau/ awak
88.	tongue	lidah
89.	tooth	gigi
90.	tree	pokok/pohon
91.	two	dua
92.	walk	berjalan
93.	hot	hangat/panas
94.	water	air
95.	we	kami
96.	what	apa
97.	white	putih
98.	who	siapa
99.	woman	perempuan
100.	yellow	kuning

The basic vocabulary comparison method or lexicostatistics begins by comparing the basic vocabulary of the Lengilo language with the Lun Bawang language in pairs, to determine which vocabulary can be used or rejected. Vocabulary that needs to be rejected are empty words (present in only one language), loanwords, and compound words. In addition, if there are derived words in the word list, the affixes are removed and only the root word is used. If the word has appeared in another word before, then that root word also needs to be rejected. After all these unnecessary words are rejected, the cognate word pairs between the Lengilo language and the Lun Bawang language is determined.

Table 3 shows Gudshinsky's (1956) proposed language kinship level which is used to determine the kinship level between the Lengilo language and the Lun Bawang language based on the percentage of cognate words obtained.

**Table 3: Language Kinship Levels**

Kinship Level	Separation duration	Percentage of cognate words
Language	0-5	100-81
Family	5-25	81-36
Branch	25-50	36-12
Microphylum	50-75	12-4
Mesophylum	75-100	4-1
Macrophylum	Over 100	Less than 1

(Source: Gudschinsky, 1956)

## 4. Results

Table 4 shows the comparison of the Lengilo language data with the Lun Bawang language and the cognate status of both languages based on the list of 100 Swadesh basic vocabulary items. The calculation of percentage to determine the language family status is also included after the vocabulary comparison analysis. Cognate words are marked with +, while non-cognate words are marked with -.

**Table 4: Comparison of Basic Vocabulary of Lengilo and Lun Bawang Languages**

No.	English	Lengilo Language	Lun Bawang Language	Cognate Level
1.	all	mong	Amung	+
2.	ashes	abu	abuh	+
3.	bark (of tree)	kulit kayu	kubil kayuh	reject (compound word)
4.	belly	batek	batek	+
5.	big	rayeh	rayeh	+
6.	bird	suit	suit	+
7.	bite	ngetep	ngetep	+
8.	black	mitem	mitem	+
9.	blood	darak	darak	+
10.	bone	tulang	tulang	+
11.	breast	itik	itik	+
12.	burn	ngeseb	ngeseb	+
13.	nail	lisun	lisun	+
14.	cloud	laput	laput	+
15.	cold	meteneb	meteneb	+
16.	come	mesing	mecing	+
17.	die	matai	mate	+
18.	dog	ukok	ukok	+
19.	drink	ngirup	ngirup	+
20.	dry	tekering	tekering	+
21.	ear	lalid	lalid	+
22.	earth	tanak	tanak	+
23.	eat	kuman	kuman	+
24.	egg	terur	terur	+
25.	eye	mateh	mateh	+
26.	fat	lemek	lemek	+
27.	feather	bulu	buluh	+
28.	fire	apui	apui	+
29.	fish	lawid	lawid	+
30.	fly	temulud	temulud	+
31.	full	penok	penuk	+
32.	foot	kukud	kukud	+
33.	give	merai	mere	+
34.	good	duak	do	+
35.	green	mebatak	mebatak uduh	reject (compound word)
36.	hair	fuk	apuk	+
37.	hand	tisuk	tichuk	+
38.	head	uluh	uluh	+
39.	hear	ninger	ninger	+



40.	heart	pusuk	pusuk	+
41.	horn	unga	unga	+
42.	I	iak	uih	-
43.	kill	ngatai	ngate	+
44.	knee	aleb	aleb	+
45.	know	kelik	kelik	+
46.	leaf	da'un	daun	+
47.	lie down	telubid	telubit	+
48.	liver	niat	niat	+
49.	long	makadang	mekadang	+
50.	louse	kutu	kutuh	+
51.	man-male	dela'ih	delai	+
52.	many	mulak	mulak	+
53.	meat-flesh	wang	wang	+
54.	moon	bulan	bulan	+
55.	mountain	surud	surud	+
56.	mouth	ta'ang	tang	+
57.	name	ngadan	ngadan	+
58.	neck	di'er	dier	+
59.	new	meberuh	meberuh	+
60.	night	malem	malem	+
61.	nose	isung	icong	+
62.	not	na'am	nam	+
63.	one	satok	aceh	-
64.	person	lemulun	lemulun	+
65.	rain	mudan	mudan	+
66.	red	mesiak	mesiak	+
67.	road	nalan	nalan	+
68.	root	war	wat	+
69.	round	melibuh	melibuh	+
70.	sand	bada	bada	+
71.	say	neten	neten	+
72.	see	ni'er	nier	+
73.	seed	ilung	ilung	+
74.	sit	tudau	tudo	+
75.	skin	kubil	kubil	+
76.	sleep	rudap	rudap	+
77.	small	di'it	sut	-
78.	smoke	refun	repun	+
79.	stand	tuped	tuped	+
80.	star	gitu'en	gituen	+
81.	stone	batu	batuh	+
82.	sun	mate sau	mate co	reject (compound word)
83.	swim	lemangui	lemangui	+
84.	tail	yur	yur	+
85.	that	ineh	ineh	+
86.	this	neh	inih	+
87.	you	ikeh	iko	+
88.	tongue	dilak	dilak	+

89.	tooth	lipen	lipen	+
90.	tree	lawa kayuh	pu'un kayu	reject (compound word)
91.	two	dueh	dueh	+
92.	walk	nalan	nalan	reject (67)
93.	hot	melauk	melauk	+
94.	water	fak	abpak	+
95.	we	kitam	kai	-
96.	what	nun	anun	+
97.	white	mebudak	mebudak	+
98.	who	ai	ide	-
99.	woman	desor	decur	+
100.	yellow	mebirar	birar	+
Total number of rejected words				= 5
Total number of compared words (100-5)				= 95
Total number of cognate words (95-5)				= 90
Percentage of cognate words (90/95)				= 94.7

Five basic vocabulary items were rejected because they consist of compound words (3, 35, 82, 90) and repeated words (92). The cognate comparison shows that the cognate relationship between Lengilo and Lun Bawang languages is close, which is 94.7%. Based on Gudschinsky's (1956) language kinship levels, both languages can be grouped at the language family kinship level. Thus, Lengilo and Lun Bawang languages can be considered sister languages.

## 5. Conclusion

This study provides valuable insight into the linguistic relationship between the Lengilo and Lun Bawang languages, two communities often assumed to be culturally integrated but linguistically distinct. Through a lexicostatistical comparison based on the Swadesh list, the findings reveal a high degree of cognate similarity (94.7%), indicating a close genetic relationship between the two languages. This supports the classification of Lengilo and Lun Bawang within the same language family subgroup, suggesting that their differences are minimal and likely shaped more by sociocultural factors than by linguistic divergence.

By addressing the previously undocumented linguistic kinship between these groups, this study fills a critical gap in Bornean language research. It not only contributes to the understanding of Austronesian language relationships in the region but also emphasises the importance of examining small, assimilated communities like the Lengilo to better understand language continuity, change, and identity in multilingual settings. Future research can build on the proposed genetic classification of Lengilo and Lun Bawang by exploring phonological, syntactic, and sociolinguistic dimensions to understand the connection between the two languages and cultures.

## Acknowledgements

The study is funded by the Dayak Chair Grant, Universiti Malaysia Sarawak. Grant No.: UNI/F09/DRC/85543/2023.

## References

- Abdul Wahab, M. K. (2022). Pengelompokan bahasa Melayu dengan bahasa Banjar dalam salasilah hipotetikal bahasa: Kajian perbandingan leksikostatistik. *Jurnal Melayu*, 21(1), 1-21.
- Abdul Wahab, M. K., & Che Halin, A. K. (2021). Leksikostatistik dan glotokronologi antara bahasa Banjar dengan bahasa Melayu: Kajian linguistik sejarah dan perbandingan. *Jurnal Kesidang*, 6, 44-61.
- Blust, R. (2010). The greater north Borneo hypothesis. *Oceanic Linguistics*, 49(1), 44-118.
- Dolgopolsky, A. B. (1964). Gipoteza drevnejšego rodstva jazykovych semej Severnoj Evrazii s verojatnostej točki zrenija [A probabilistic hypothesis concerning the oldest relationships among the language families of Northern Eurasia]. *Voprosy Jazykoznanija*, 2, 53-63.
- Eghenter, C., & Jok, D. (2012). *Formadat: Forum masyarakat adat dataran tinggi Borneo*. Formadat.
- Gudschinsky, S. C. (1956). The ABC'S of lexicostatistics (Glottochronology). *WORD*, 12(2), 175-210, DOI: 10.1080/00437956.1956.11659599
- Jian Abdullah, N., Abdul Wahab, H., & Ting, S. H. (2024). Sounds of Lengilo based on Swadesh list. *International Journal of Business and Technology*, 6(2), 216-223.
- Jimbai, J., Banseng, S., & Handrianto, C. (2024). Formation process of derived words in Malay and Iban. *International Journal of Education, Information Technology, and Others*, 7(1), 325-340. <https://doi.org/10.5281/zenodo.10871263>
- Kasim, M. R. (2021). Hubungan kekerabatan antara bahasa Jakun dengan bahasa Melayu purba: Satu analisis perbandingan kosa kata dasar. *Jurnal Bahasa*, 21(1), 41-66.
- Langub, J. (1987). Ethnic self-labelling of the Murut or "Lun Bawang" of Sarawak. *Sojourn: Journal of Social Issues in Southeast Asia*, 2(2), 289-299.
- Lohr, M. (2000). New approaches to lexicostatistics and glottochronology. In C. Renfrew, A. McMahon & L. Trask (Ed.), *Time depth in historical linguistics*, 1 (pp. 209-223).
- Lun Bawang*. (2025, June 1). Wikipedia. [https://www.detailedpedia.com/wiki-Lun\\_Bawang](https://www.detailedpedia.com/wiki-Lun_Bawang)
- Mohamed, N. (1999). Pengelompokan bahasa. *Jurnal Dewan Bahasa*, 43, 1077-1107.
- Mohamed, N., & Abdul Wahab, H. (2004). Tingkat kekerabatan bahasa Bidayuh di Sarawak. *Jurnal Bahasa*, 4(3), 452-483.
- Mohamed, N., & Mohd Yusof, R. (2011). *Pengelompokan leksikostatistik bahasa-bahasa peribumi di Sarawak*. Penerbit Universiti Sains Malaysia.
- Mohamed, N., & Mohd. Yusof, R. (2004). Persamaan dan kemiripan bahasa: Perbandingan kosa kata dasar 10 bahasa Austronesia di Sarawak. *Jurnal Bahasa*, 4(2), 278-359.
- Muda, N., & Hamzah, Z. A. A. (2021). Aspek fonologi dan leksikal bahasa Orang Asli Che Wong. *Jurnal Antarabangsa Dunia Melayu*, 14(1), 69-100.
- Muda, N., & Hamzah, Z. A. A. (2021). Leksikostatistik bahasa Che Wong dan bahasa Melayu melalui aspek perbandingan kosa kata. *PENDETA*, 12(2), 11-28.
- Omar, A. (2008). *Ensiklopedia Bahasa Melayu*. Dewan Bahasa dan Pustaka.
- Omar, A. (2015). *Kaedah penyelidikan bahasa di lapangan* (2nd ed.). Dewan Bahasa dan Pustaka.
- Sarawak Government. (2018). Sarawak population by District, Administrative District, Ethnic and Age Group from 2010 - 2020p. Population Projections (Revised) based on the Population and Housing Census of Malaysia 2010. <https://data.sarawak.gov.my/home/data/dataset/bdc23af1-4525-42eb-8580-8f9233522ce4>
- Sakel, J., & Everett, D. L. (2012). *Linguistic fieldwork: A student guide*. Cambridge University Press.

- Sidek, A. (2022, August 22). *The Lunbawangs and Lundayehs of Borneo*.  
<https://cj.my/9900/the-lunbawangs-and-lundayehs-of-borneo/>
- Swadesh, M. (1950). Salish Internal Relationships. *International Journal of American Linguistics*, 16, 157-167.
- Swadesh, M. (1952). Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society*, 96, 452-463.
- Swadesh, M. (1955). *Concept list Swadesh 1955 100*. <https://concepticon.clld.org/contributions/Swadesh-1955-100>
- Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics*, 21, 121-137.
- Swadesh, M. (1971). *The origin and diversification of language*. Aldine.
- Tirusel, S. T. P., & Da'a, U. (2021). *Jejak peradaban: Manusia Dayak Krayan (Lengilo') di Sungai Krayan*. Lembaga Literasi Dayak.

**Appendix 1. Sarawak population in 2020**

Group	Population	Percentage
<b>Malaysians</b>	<b>2,657,000</b>	<b>94.3</b>
Malay	655,500	23.3
Iban	812,600	28.9
Melanau	142,600	5.1
Bidayuh	223,200	7.9
Other Indigenous	179,200	6.4
Chinese	627,800	22.3
Indians	7,700	0.3
Others	8,300	0.3
<b>Non-Malaysians</b>	<b>159,500</b>	<b>5.7</b>
<b>TOTAL</b>	<b>2,816,500</b>	<b>100.0</b>

Note: Population Projections (Revised) based on the Population and Housing Census of Malaysia 2010.  
(Source: Sarawak Government. (2018). Sarawak data. <https://data.sarawak.gov.my/home/data/dataset/bdc23af1-4525-42eb-8580-8f9233522ce4>)