IJEAL (*International Journal of English and Applied Linguistics*)
Volume : 5 | Number 2 | August 2025 | E-ISSN: 2787-9482 | DOI: doi.org/ijeal.v5i2.6546

THE KINSHIP RELATIONSHIP BETWEEN THE MALAY LANGUAGE OF RIAU AND THE SUNDANESE LANGUAGE A COMPARATIVE HISTORICAL LINGUISTIC STUDY

Syahrul Efendi [1], Evan Afri[2], Dardanila[3], Dwiwidayati[4], Aron Meko MBete [5]
Politeknik Gihon[1], Politeknik Ganesha Medan[2], Universitas Sumatera Utara[3,4,5]

syahrulfendi19@gmail.com[1], evanafri@gmail.com[2], dardanila@usu.ac.id[3], dwiwidayati@usu.ac.id[4], aronmbete@yahoo.com[5]

Received: 17/07/2025
Accepted: 28/07/2025
Publication: 01/08/2025

Abstract
This study aims to reveal the kinship relationship between Riau Malay and Sundanese languages seen from some vocabularies similarities. The method used is descriptive qualitative and quantitative with note-taking technique. The research data is formed of vocabulary in BMR and BS obtained from various written sources and informant interviews. Based on 200 Swadesh words for BMR and BS, there are 12,5%, namely 25 identical pairs words. This data proves that BMR and BS are not closely related because the vocabulary similarities are less than 36% or part of the sub-grouping. Vowel relatives couples different by 16 data or about 18%; 3 relatives couples have different pairs of consonants or about 1.5%; 5 relatives couples or about 2.5% who have one syllable difference; and 9 relatives couples or about 4.5% who have one consonant phoneme difference. The results of the study found that the percentage of kinship between the two languages was 29% which indicated that the BMR and BS were the languages of the family from one family (stock).BMR and BS are thought to be the same prelanguage around 2,856 years ago and are estimated to have started to separate from their parent languages in 833 BC (calculated in 2023)

Keywords: Kinship, Riau Malay Language, Sundanese Language

Introduction

Because of its cultural and linguistic diversity, Indonesia is a multicultural and multilingual country. There are around 726 regional languages spoken in the country (Montolalu, 2005) . The Austronesian or Polynesian Malay language family includes the languages spoken in Indonesia (Nababan, 1991) . Furthermore, there are 5,445 languages spoken worldwide, according to (Parera, 1993). Several languages are spoken in Indonesia. There are as many as 741 regional languages in Indonesia, according to data (Schane, 1992). All languages currently spoken in Indonesia are thought to be related to each other because they have kinship between languages.

The pattern of language development results from the process of cooperation and communication between various social groups, and over time causes the language to develop and change (Nababan, 1991). Language modification is done to ensure effective communication in sociocultural life, language change results in similarity in form and meaning. If two or more groups of language speakers communicate rarely or not at all, a language will experience very diverse changes and developments (Sudarno, 1994 ; Hafizah, 2018) .

Riau Province is one of the regions in Indonesia that has a distinctive language. In this province, Malay is the language that has the greatest influence on the development of the

Volume : 5 | Number 2 | August 2025 | E-ISSN: 2787-9482 | DOI: doi.org/ijeal.v5i2.6546

Indonesian language. Riau residents, as well as mainland Malays and Coastal Malays, speak Riau Malay. Coastal Malays live on the coast or on the islands of Riau province, while mainland Malays live on the mainland. Riau Malay is a simple language that can be understood by everyone. Considering that Riau Province is a destination area for the migration of a number of ethnic groups, including Sundanese, Banjar, Bugis, and others, it is necessary to conduct research on the relationship between the languages used by its people. This is because the Riau region has a friendly and cultured population, a strategic geographical location, and of course economic management that can be taken into account (Hamidi, 2017) . Banjar, Batak, Malay, Minangkabau, Sundanese, and Bugis are among the six languages spoken in Riau Province, according to the National Language and Book Development Agency (2020). The diversity of other regional languages found in Riau Province is undeniable; there may be other regional languages that exist in Riau Province, it is just that linguistic research is still limited.

In addition, compared to regional languages in eastern Indonesia, Riau Malay and Sundanese have a very high number of speakers. Both of these regional languages are still very actively used by their speakers. As a result, these two languages are dynamic and can change both in their phonetic structure and grammar. Geographically, the Sundanese and Riau Malay are far apart. Riau Malay and Sundanese are found in the Austronesian geographical region, according to Salzner's 1960 classification of language families (see (Keraf, 1991) . Judging from Salzner's scale, the two languages are very closely related to each other. However, to the best of the author's knowledge, it has not been determined to what extent the languages are related to each other.

Based on the above, the author is interested in knowing how closely related Riau Malay and Sundanese are, as well as the similarity of word sounds between the two languages. Looking for similarities in word forms by comparing all phonemes to determine the level of kinship. In addition, by comparing the word forms of the two languages phonetically, namely by comparing consonant and vowel variants (Jahdiah,2011) conducted similar research by comparing Banjar and Sundanese languages. In addition, there is research from (Sofiatunnida, 2021) titled "Lexicostatistics of Mandailing Language and Malay Language" which compares Wowonii, Morene, and Kulisusu languages. This research basically succeeded in determining the degree of kinship and finding related word forms from each of the languages being compared. According to his findings, there is 58% relatedness between Mandailing and Malay. Moreover, between 1,419 and 1,101 years ago, Mandailing Batak and Malay were one language. In the period between 601 and 919 AD (measured from 2020), Mandailing Batak and Malay began to deviate from protolanguage.

Literature Review (Times New Roman 14)

This study is based on the same theory as historical comparative linguistics, which argues that phonetic, morphological and grammatical similarities between two languages can be demonstrated. The morphological and grammatical structures of the two languages are not discussed in this study; rather it only compares the phonetic similarities and similarities of each word (lexicon). Historical Comparative Linguistics is the theory used in this study. According to (Zulaeha, 2010) and (Dalimunthe, 2018) , historical comparative linguistics is a theory that looks for similarities (historically) between the linguistic components of the dialects or languages being compared. In linguistic theory, Historical Comparative Linguistics is known as lexicostatistics, which is a technique to determine the level of kinship of two or more languages to be compared.

Lexicostatistics, according to (Keraf, 1991) , is a method of language grouping that prioritizes statistical observation of words (lexicon) to establish groupings based on the percentage of similarities and differences between one language and another. Several techniques, including those listed below, are used to establish relationships between two or more languages/dialects. 1) Identical word pairs, i.e. word pairs in which all phonemes are the same; 2) Phonemically correspondent pairs, i.e. if the balanced forms between the two languages occur reciprocally, frequently, and regularly, the two languages are considered related; 3) Phonetically similar, meaning that phonetic features alone are sufficient for a pair of words to be considered

Volume : 5 | Number 2 | August 2025 | E-ISSN: 2787-9482 | DOI: doi.org/ijeal.v5i2.6546

related if they are phonetically similar at the same position of articulation; 4) One phoneme is different, meaning that if a pair of phonemes is different but the change can be attributed to environmental influences, it can be said that the two are related (see Jahdiah ;, 2011 Fauzi, 2019) . In contrast, the strategy used in this study to establish kinship between Riau Malay and Sundanese is to look for word similarities (both identical and based on phonetic similarity). Regardless of how distant the two languages are from each other, the languages that are compared in lexicostatistics are those that come from the same protolanguage. According to (Crowley, 2010) , lexicostatistics gives different names to several levels of clustering, among others:

Table 1. Language Grouping Categories Based on Kinship Percentage

Kekerabatan

No.	Tingkat Pengelompokan	Persentase Kata Kerabat
1.	Dialek dari satu bahasa	81-100
2.	Bahasa dari satu keluarga	36-81
3.	Keluarga dari satu rumpun (stock)	12-36
4.	Turunan dari sebuah mikropilum	4-12
5.	Mikropila dari sebuah mesopilum	1-4
6.	Mesopila dari sebuah makropila	0-1

According to (Widayati, 2015) , statistical similarity is estimated based on the percentage of similarity and similarity of 200 base words between related languages. Based on the data above, the following formula can be used to determine how close the kinship relationship is between languages.

$$c = \frac{k \times 100\%}{n}$$

k= number of related words

n= number of words being compared

Furthermore, (in Dardanela, 2010) calculated the separation time of the compared languages, with the formula:

$$t = \frac{\log c}{2 \log r}$$

t= time depth (separation time)

r= retention or constant percentage in 1000 years, or also called an *index*, in this case the retention is 80.5%, (Swadesh).

When compared to other techniques, lexicostatistical techniques have a number of advantages, according to Nothofer (in Ino, 2015) . These benefits include:

1. it serves as a basic vocabulary list that can be used to identify kin or kinship languages;
2. as a grouping tool for related languages or dialects whose protobahasa is not too old or ancient.
3. as a grouping tool or procedure that can be applied early on to determine language classification.

Glotochronology, according to (Widayati, 2015) , is a technique of estimating the amount of time that separates two language groups. Glotochronology is a method for classifying languages by emphasizing the age of the languages concerned or the time of separation between two languages. It is used in comparative historical linguistics.

Table 2. Categories of Language Grouping Based on Kinship Percentage

No.	Tingkat Pengelompokan	Waktu Pisah
1.	Dialek dari satu bahasa	Kurang dari 500 tahun
2.	Bahasa dari satu keluarga	500-2500 tahun
3.	Keluarga dari satu rumpun (stock)	2500-5000 tahun
4.	Turunan dari sebuah mikropilum	5000-7500 tahun
5.	Mikropila dari sebuah mesopilum	7500-10000 tahun
6.	Mesopila dari sebuah makropila	Lebih dari 10000 tahun

Research Method

Determining kinship relationships and reconstructing ancient languages that gave rise to related languages are two fundamental tasks carried out in the study of comparative linguistics. Quantitative methods and qualitative methods are two techniques used to identify linguistic relationships or groupings. While qualitative methods describe the characteristics of similar languages, quantitative methods describe lexicostatistical methods (Mahsun, 2017). The lexicostatistical and glottochronological approaches used in this study are quantitative research with a descriptive qualitative methodology. The descriptive approach is a technique that uses natural language to explain facts and generate linguistic rules. Although it is called qualitative because the data collected are words rather than numbers, it strives to make a systematic and accurate description of the data analyzed based on actual phenomena and empirical facts. According to Bogdan and Taylor (in Alijah, 2016), this strategy is used to generate descriptive data in the form of spoken or written words from people and observable behavior.
The data collection techniques used in this study are:

a. Record: In this study, the recording method was carried out using a list of 200 basic Indonesian words modified from the basic words found by Morris Swadesh. In addition, the informants translated the 200 words into the appropriate regional language, which was then explicitly recorded by the researchers.

b. Recording: All vocabulary presented by informants was captured using a recording approach.

Therefore, the relationship between the two languages, particularly BMR and BS, will be discussed in this study. The objects were taken from the entire extensive discourse or discourse that had already been conducted, whether selected as samples or not. Data was taken from all statements made by informants to be used as research data sources. Furthermore, BMR and BS in their respective language user areas are the objects of study in this research. 200 Swadesh words are used as data in the research. The research is conducted using the theory of Comparative Historical Linguistics to answer research questions that include determining the kinship relationship between BMR and BS.

The data were analyzed using historical comparative techniques, and the two languages were compared descriptively and comparatively. Descriptive comparison was used to identify changes in the related languages that are still used by speakers today. Since unrelated forms are considered to have different etymologies, both in form and meaning, related word sets (cognate sets) were compared by selecting related forms. This means that words that are constructed and borrowed, rather than being root words, are not considered in the glossary.

To calculate the degree of kinship and the time of separation between BMR and BS, the number of forms with phonetic similarity and correspondence was counted from the compared cognate sets. Without considering how different the two languages are from each other, the languages compared in this lexicostatistics must be languages that developed from the same protolanguage. The criteria for identifying similar words are as follows: (1) The words being compared are the same; (2) The words are phonetically related; and (3) One phoneme difference.

Volume : 5 | Number 2 | August 2025 | E-ISSN: 2787-9482 | DOI: doi.org/ijeal.v5i2.6546

The lexicostatistical calculation was done using the following formula: C= k/n x 100%. The number of words compared is n, and K denotes related words. The separation time between languages was then calculated using the following formula: t = Log c/ 2 log r. T stands for separation time, and r for vocabulary durability.

Results and Discussion

The relationship between BMR and BS was analyzed based on the identification of the two languages using Swadesh's basic vocabulary of 200 basic vocabulary. The Swadesh list is considered good in this study as it consists of non-cultural words. Moreover, the retention of the base words has been tested in a language that has a written script. The data from the 200 Swadesh base vocabularies, identified related word pairs which were compared through identical pairs. The comparison results are presented in the following table:

Table 3. Identical pairs

No	Gloss	Riau Malay	Sundanese
1	dog	anjiG	anjiG
2	stone	stone	stone
3	street	road	road
4	fruit	fruit	fruit
5	feather	bulU	bulU
6	worm	caciG	caciG
7	meat	dagiG	dagiG
8	come	dataŋ	dataŋ
9	leaf	leaf	leaf
10	at	at	at
11	two	dUa	dUa
12	the	the	it
13	evil	evil	evil
14	needle	needle	needle
15	dirty	dirty	dirty
16	skin	skin	skin
17	lice	kutU	kutU
18	spit	spit	spit
19	knock	Keto?	Keto?
20	dig	dig	dig
21	inflatable	inflatable	inflatable
22	hot	hot	hot
23	long	panjaG	panjaG
24	seven	seven	seven
25	white	white	white

From the list above, there are 25 identical pairs between BMR and BS or about 12.5%. Furthermore, the vocabularies of the two languages were also identified based on phonemic correspondence. The pairs that have phonemic correspondence are divided into vowel- different, consonant-different, and one-syllable- different pairs. The results of the analysis are presented in the following table.

Table 4. Vowel-different pairs

No	Gloss	Riau Malay	Sundanese
1	land	land	tanəƟh
2	wind	angən	wind
3	star	star	béntang

4	where	in manə	where
5	far	jaC:h	far
6	lightning	kilət	lightning
7	yellow	yellow	konəng
8	Male	ləlaki	lalaki
9	sky	langət	sky
10	sea	laot	sea
11	five	limə	five
12	eye	matə	eye
13	short	pendék	cottage
14	wet	wet	basəʊh
15	heavy	heavy	bəʊrat
16	split	split	bəʊlah

From the table above, there are only 16 data with different vowels or about 8%. The vowel differences identified include the correspondence a ~ əʊ; ə ~ i; i ~ é; ə~ a; C: ~ u; i ~ ə; o ~ u; e ~ o; é ~ o; e ~ əʊ.

Table 5. Consonant-Different Pairs

No	Gloss	Riau Malay	Sundanese
1	dream	dream	ŋ dream
2	rotten	rotten	bad
3	promise	promise	jaŋji

Different consonant pairs amount to 3 words or about 1.5%. The consonant difference identified is the variation of consonants that are close in articulation, such as [m] with [ŋ] where [m] is bilabial and [ŋ] is velar, [s] with [r] which are both alveolar and [n] with [ŋ] where [n] is alveolar and [ŋ] is velar (Schane, 1992).

While those that differ by one syllable are found in the following table 6, which amounts to 9 words or 4.5%.

Table 6. One Syllable Different Pair

No	Gloss	Riau Malay	Sundanese
1	Correct	bəna	bənər
2	four	əmpat	opat
3	heart	jantuG	jajantuG
4	wood	wood	kai
5	dead	mali	maot
6	drink	drink	nginum
7	vomit	vomit	Utah
8	three	three	tilu
9	buy	buy	məʊli

Furthermore, there are 5 words or around 2.5% that have different pairs of one consonant phoneme as shown in the table below:

Table 7. Different Pair of One Consonant Phonemes

No	Gloss	Riau Malay	Sundanese
1	ayam	ayam	hayam
2	Mr.	father	father

3	rain	rain	rain
4	open	bukak	open
5	thin	thin	ipis

One consonant differences are generally dominated by BS, although they also exist in BMR, namely bapak ~ bapa, bukak ~ buka, and tipis ~ ipis. This one consonant phoneme difference is found in the initial position, and the final position. This means that Sundanese has longer phonemes than Riau Malay. The overall similarity of vocabulary from BMR and BS that can be concluded from the five tables above is 58 vocabularies or about 29%. This means that the percentage of kinship between the two is 29%. This percentage shows that BMR and BS are families of the same stock.

Based on the kinship classification data between BMR and BS that has been done previously, it can be seen that the percentage of kinship between the two languages is:

C = $\frac{VT}{VD}$ x 100% = $\frac{58}{200}$ x 100% = 29%

From the data above, we can know the calculation of the separation time of BMR and BS, namely

C = 29% Log r (retention/retention time of vocabulary) = 80.5% Split Time = ?

Split Time = $\frac{log.C}{2.\log r}$ = $\frac{\log 0,29}{2.log.\ o,805}$ = $\frac{-0,537}{2.-0,094}$ = $\frac{-0,537}{-0,188}$ = 2,856

Separation time = 2.188 x 1000 = 2.856

From the calculation, it can be seen that the initial separation time between BMR and BS is 2,856 years ago.

From the calculation of the initial separation time between BMR and BBT above, it can be concluded as follows:

1. BMR and BS are estimated to be the same pre- language about 2,856 years ago
2. BMR and BS are estimated to have begun to separate from their parent language in 833 BC (calculated in 2023).

Conclusion

Based on data analysis and discussion, the results of this study can be concluded as follows.

1. Based on the calculation of lexicostatistical techniques, out of 200 Riau Malay vocabularies with Sundanese consist of 58 related vocabularies and 142 unrelated vocabularies. So, the percentage of kinship between the two languages is 29%. The relationship between Riau Malay and Sundanese can be established as languages from the family of one family (stock).
2. Based on the calculation of the glotochronology technique, the separation time between Riau Malay and Sundanese is 2,856 years ago, starting from 2023 (833 BC).
3. Evidence of sound correspondence between Riau Malay and Sundanese based on Swadesh vocabulary can be seen in several criteria, 25 identical pair words, 16 different vowel pair words, 3 different consonant pair words, 9 different syllable pair words, and 5 different pair words with one consonant phoneme. The separation of the language is caused by migration, culture, and identity factors of the speakers.

From the entire discussion of cases in the form of Mandailing language above, Mandailing language is typologically not a polysynthesis or incorporation language, but Mandailing language has an incorporation form. The form of Mandailing language incorporation turns out to be quite extensive and varied which is almost the same as Indonesian, and can also be applied clearly without changing the meaning of each incorporation. Incorporation with verb obliteration is an incorporation that occurs because one of the cases is incorporated into the verb after experiencing verbalization, as a result the original verb is obliterated. Based on the results of the analysis above, it can be concluded that there are four types of incorporation of verb deletion found in Mandailing language, namely objective incorporation such as marsalaor panjang (wearing long pants), instrumental incorporation such as sabuni (clean with soap), locative incorporation such as celengkon (put in a piggy bank), and state incorporation such as mangotori (dirtying).

References

Alijah, S. (2016). *Kinship between Bugis and Muna languages*. *Journal of Humanities*,*16(1) , Mar*. Crowley, T. and C. B. (2010). *An Introduction to Historical Linguistics.* Oxford: Oxford University Press.

Dalimunthe, S. R. (2018). *Kinship Relationship Mandailing Batak Language and Tanah Ulu (A Historical Linguistic Study).Comparative Linguistic Study)*. *Journal of Medan Makna, Vol XVI* No. 1 pp. 84 - 91.

Dardanela, et al. (2010). *Cognates among the Karo, Alas and Gayo Languages. International. Journal of Humanities and Social Science,* (55-58).

Fauzi, M. (2019). *Kinship Relationship between Riau Malay and Akit Dialect (A Historical Approach. Journal of Cultural Sciences, Vol. 16*, *No. 1 August 2019*.

Hafizah. (2018). *Statistical Lexicon of Indonesian with Minang Language of Bukit Tinggi Dialect (A Comparative Historical Linguistic Study). Journal Deixis, Vol. 10 No.03, September- December 2018.*

Hamidi, U. (2017). *Malay Assimilation with Bugis and Banjar in Riau.* http://bilikkreatif.com/asimilasi-melayu- dengan-bugis-dan-banjar-di-riau-oleh-uu-%0A hamidy.%0A

Ino, L. (2015). *Utilization of Comparative Historical Linguistics in Mapping the Languages of the Archipelago. Retorika: Journal of Language Science*, *1(2),* (365-378).

Jahdiah. (2011). *Kinship relations between Banjar and Sundanese. Bandung: Journal of Metalingua West Java Language Center.*

Keraf, G. (1991). *Historical Comparative Linguistics.* Jakarta: PT Gramedia Pustaka Umum.

Mahsun. (2017). *Language Research Methods, Stages, Strategies, Methods and Techniques.* Rajagrafindo Persada.

Montolalu, D. (2005). *Sociolinguistics.* Jakarta. PT. Gramedia Pustaka Utama.

Nababan, P. W. J. (1991). *Sociolinguistics An Introduction.* Jakarta: PT Gramedia Main Library.

Parera, J. D. (1993). *General Linguistic Studies, Comparative Historical, and Structural Typology.* Jakarta: Erlangga Publisher.

Schane, (1992). *Phonology Workbook Generative.* Jakarta: Summer Institute of Linguistics.

Sofiatunnida and Hendrokumoro. (2021). *"Lexicostatistics of Mandailing and Malay". Nusa Journal, Vol. 16 No.. 2*

Sudarno, M. E. (1994). *Language Comparison Nusantara.* Jakarta: Arikha Medika Cipta.

Widayati, D. (2015). *Historical Linguistics Comparative, Textbook, Teaching Materials.* Medan: CV. Mitra Medan.

Zulaeha, I. (2010). *Dialectology: Dialects Geography and Social Dialects.* Yogyakarta: Graha Ilmu.