

# Lexicostatistics Calculation on Manggarai and Bima Languages: A Comparative Historical Linguistics Study

Salahuddin

Gadjah Mada University, Yogyakarta, Indonesia

## ABSTRACT

This study aims to provide quantitative and qualitative evidence through lexicostatistical calculations to determine the separation distance and possible relationship between the Manggarai and Bima languages. Manggarai and Bima languages are two Austronesian languages belonging to the Central-Eastern-Malayo-Polynesian located on two adjacent islands. Manggarai is the language spoken by the Manggarai tribe on West Flores Island, while the Bima is the language spoken on the island of Sumbawa. This is based on evidence of the discovery of several vocabularies that have similar forms and meanings in both languages. The researcher uses secondary data provided in Lexirumah and Austronesian Basic Vocabulary databases that contain Austronesian lexical data based on 200 Swadesh word lists promoted by Swadesh (1952). Based on lexicostatistical calculations, the results show that the two languages show a kinship level of 27%, which means that the Manggarai and Bima languages are two different languages from different family group (stock level), or, in other words, the Bima and Manggarai languages were a single language 3.016±242 years ago. Meanwhile, the qualitative results were obtained by considering the related vocabulary between the two languages and showing the existence of sound correspondences, such as [w]~[v], [f]~[p], [d]~[t], [r]~[l], and [a]~[e].

## ARTICLE INFO

*Paper type:*  
Research Article

*Article history:*  
Received: 16/10/2022  
Revised: 05/05/2023  
Accepted: 22/06/2023

*Keywords:*

- Bima
- Glottochronology
- Historical Linguistics
- Lexicostatistics
- Manggarai

## 1. Introduction

Language reconstruction in the archipelago has long been carried out by grouping hypothetical genetic relations among languages. This relatedness can be found in the forms of meaning that correspond to each other by referring to the PAN (Proto-Austronesian) form compiled by Denpwofff. Research on the historical relationship of Western Austronesian languages has been carried out for a long time. It has rapidly developed until it has succeeded in reconstructing several proto-languages at a lower level (Fernandez, 1996: 14). Language studies in Eastern Indonesia are still uncommon. Those languages are less popular and spoken by a small number of people in the archipelago.

According to Fernandez (1996), Comparative Historical Linguistics is a branch of linguistics whose objective is to determine the facts and levels of relatedness between different languages speakers closely related to grouping these languages. Languages are considered to have a joint development, originating from members of a language group. This study examines the relationship between two languages: the Bima and the Manggarai languages.

The Bima language, or what the Bima people are more familiar with, is called *Nggahi Mbojo* (Arafiq, 2020: 13; Karim, 2022: 21). It is a language spoken by the Bima community group located on the island of Sumbawa, West Nusa Tenggara province. According to language mapping data released by the Ministry of Education and Culture (Sunendar, 2019), the Bima language is spoken by the Bima (also known as *Mbojo*) ethnic group who inhabit the Bima Regency area, including Bima City and Dompu Regency on the Eastern Sumbawa Island, NTB (Budasi, 2018). It was also found that a small number of community groups in East Nusa Tenggara used the Bima language as their first language, such as Reo in Manggarai Regency, Pota in East Manggarai Regency, Labuan

Bajo, and Nanga Lili in West Manggarai Regency (Sanjaya et al., 2018: 62). The Bima-speaking community groups are also found in Sabu Island and Sumba Island, Flores areas. The existence of the Bima community group in the East Nusa Tenggara region is due to political factors that occurred hundreds of years ago. Under the auspices of the Goa kingdom, the Bima kingdom had the authority to collect tribute from the Manggarai (Manggarai in general before the expansion) in 1667 BC.

The Bima ethnic group in the Sumbawa archipelago consists of immigrant Malays and indigenous tribes (Sulistyo, 2014: 156; Astuti, 2019: 111). The immigrant Malays come from different backgrounds; some are from Minangkabau, Bugis, Javanese, and many others. Meanwhile, the indigenous tribe that inhabits the island of Sumbawa is the *Donggo* tribe, a tribe that inhabits mountainous areas. The book of *Bo Sangaji* explains that before entering the Hindu kingdom in the 10th century, people in Bima lived in tribes that became known as the *mbojo*. This community group has settled, and cultivated crops, and the members of the community group are only around 45 to 250 people in each clan. However, no historical records explain the origin of the tribe in detail. Bima has a good relationship with the Goa kingdom in Makassar, but both still show different characteristics, for instance, seen from the pattern of traditional houses and the language used.

Furthermore, the Manggarai language is spoken by people in Manggarai, West Manggarai, and East Manggarai Regencies. Before being divided into these three regions, the name Manggarai was better known in general. Long before that, the area was better known as Nuca Lale, Nuca, which means island, and Lale is the name for a type of tree called the Kerbang Tree (Erb, 1997: 49; Bustan et al., 2020: 16). Manggarai was only mentioned after the Bima kingdom invaded the Nuca Lale region (Djakariah et al., 2019). Mentioned a young man named Mangga Macing; this figure was the eldest son whom Bima sent to conquer Manggarai with his three brothers. The term Manggarai is given after one call for the word *Manggar rai*. The word *manggar* means an anchor raised as a call to run (Erb, 1997: 49). By that time, the name Manggarai is used to this day.

There is no definite record of the origins of the Manggarai tribe in West Flores. However, some popular opinions say that the Manggarai people are said to have come from Minangkabau, specifically those who are settled in Todo Village, Manggarai Regency (Daeng, 1997: 17). However, there is no empirical evidence that can prove this, including in terms of language. However, it is explained that there are similarities in folklore about two brothers who disagreed and decided the difference by fighting buffalo. Thus, the word Minangkabau according to the Manggarai people comes from *Menang Kerbau* which means 'the winning buffalo'. The youngest brother who lost in the buffalo fight finally decided to go and travel east, to Todo, West Flores Island.

Another opinion says that the Manggarai people have strong relations with the Sumbanese who are on the opposite side of the island (Moses, 2018: 4). On the other hand, The Manggarai people are believed to have descended from the Nggae Sawu, who migrated from Mando Sawo (Suwondo, 1978: 8-9). This is in line with the multivariate stochastic correlation result carried out by Glinka (1073), who found that the Manggarai, Sumba, Belu and Kemak Marai tribes had the same physical characteristics, namely long heads, tall looks, long, short noses and tall bodies (Suwondo, 1978: 11). This last opinion is indeed more widely accepted because apart from their close proximity to each other, the Manggarai and Sumba tribes have the quality of their kinship system and cultural practices. Blanc (2024) in the website <https://peoplegroups.org/> also includes the Manggarai people in the Flores-Sumba-Alor group. This is also in line with the language grouping carried out by Esser (1938) who included the Manggarai language in the Bima-Sumba group.

Taking into account the geographical location factor, it is possible that the Manggarai people and their language come from adjacent islands. The researcher hypothesizes that it is possible that the Manggarai language has a strong relationship with the Bima language that existed on the island of Sumbawa long before Manggarai was brought under the rule of the Bima kingdom. This is also supported by Grimes' statement (1997: 5), based on archaeological evidence, it is estimated that the people from Taiwan migrated south through the Philippines and moved to Eastern Indonesia about four thousand years ago. Fernandez (2007) also argues that if two or more languages are separated by several adjacent islands, then these languages can be traced to their kinship. Dyen (1962: 39) explains that when two languages share a common ancestor, the more recently they diverged from each other, the more genetically similar they are likely to be. As time passes, different forms of a language can develop and become increasingly distinct from each other. Departing from this assumption, thus, the researcher tries to trace the kinship relationship between the two opposing tribes by providing quantitative and qualitative language evidence.

There are several previous studies that specifically examined the relationship and grouping of the Bima language and other languages in East Nusa Tenggara. Brandes (1884: 84) divided the languages of the archipelago into groups of the West Archipelago and the Eastern Archipelago. Meanwhile, the languages in East Nusa Tenggara are grouped into two parts, namely the languages of West Flores and East Flores. The West Flores language is grouped into the West Archipelago language, while the East Flores language is included in the East Nusantara language. Esser (1938) grouped Manggarai into the Bima-Sumba (BS) group in addition to several other Flores languages, which were grouped into the Ambon-Timor (AT) group.

Mbete (1990) tried to test the results of language grouping conducted by Dyen (1982) regarding the kinship of Balinese, Sasak, and Sumbawa languages which were grouped into (the Bali Subgroup). In his research, he compared the languages in the east, namely Bima and Manggarai, and the languages in the west, Javanese and Madurese. The research results are that the languages compared are from the same stock and show a tendency for the Balinese, Sasak, and Sumbawa languages to have a higher degree of similarity with the comparison language group in the west than the comparison language group in the east. Thus, the research confirms Brandes' hypothesis regarding grouping Polynesian Malay languages into subgroups of the West Archipelago and the Eastern Archipelago.

Research conducted by Syamsuddin (1996) attempted to test the hypothesis of language grouping conducted by Esser regarding the Bima-Sumba language grouping. The results show that the Bima language is closer to the languages in the east than the languages in the west, namely the Sasak language and the Sumbawa language. Thus, Syamsuddin validated Esser's hypothesis.

Fernandez (1996) conducted a study on the historical relationships between languages belonging to the Bima-Suma group, which was proposed by Blust (2008) in his previous research. The Bima-Sumba subgroup includes the Bima language, seven Sumba languages, and four other languages - Manggarai, Rembong, Ngadha, Lio, and Palue. The findings revealed that the languages in Flores fall into two subgroups - East Flores and West Flores. Fernandez (1996) emphasized that within the West Flores language subgroup, the Manggarai, Rembong, and Komodo languages together form a separate group called the MRK group, while the Ngadha, Palu'e, and Lio languages form a separate group called the NPL group. Furthermore, the study highlighted that within the MRK group, the Manggarai language bears a closer resemblance to the Rembong language than to the Komodo language.

Table 1. Percentage of kinship relations between Flores languages in East Nusa Tenggara (Fernandez, 1996)

Lh	62														
Sk	60	65													
Pl	48	50	49												
Li	35	42	48	62											
Ng	41	45	47	56	61										
Rb	40	42	49	50	52	48									
Mg	41	47	50	49	55	50	70								
Km	39	41	43	52	50	48	58	59							
<b>Bm</b>	21	26	24	26	28	27	25	<b>27</b>	29						
Kb	23	25	25	25	22	26	28	29	27	20					
Rt	22	20	21	21	21	19	-	21	-	25	27				
Hl	21	24	22	-	20	21	-	22	39	17	24	39			
Dw	20	20	19	-	18	17	-	19	-	20	23	34	30		
Tt	21	24	24	-	20	20	-	20	-	25	28	31	31	33	
Kd	23	23	22	-	21	19	-	22	-	20	22	27	29	24	28
	Kd	Lh	Sk	Pl	Li	Ng	Rb	<b>Mg</b>	Km	Bm	Kb	Rt	Hl	Dw	Tt

Besides the languages within the Flores islands, Fernandez (1996) also examines the connection between the Manggarai language, spoken on the westernmost island of Flores, and the Bima language, spoken on the eastern tip

of Sumbawa Island. According to Table 1, the two languages share 27% of kinship, meaning they belong to the same 'stock' language level, which means they are a parent language and all its derived dialects and languages (Harcourt, 2010). However, Fernandez did not provide an in-depth explanation of the relationship between the two languages in terms of their linguistic characteristics (such as phonological similarities in the form of sound correspondence) and the distance between them. Hence, this research aims to examine these aspects in detail, which Fernandez did not reveal in his previous research.

In addition, this research was also motivated by the discovery of the similarity of the vocabulary used in the two languages, especially regarding family relations and the vocabulary used to interact in everyday life. In vocabulary related to kinship, for example, there are similarities in the form and meaning of the words.

Table 2. Examples of similarities in terms of forms and meanings

<b>Kinship Terminology</b>		
<b>Manggarai</b>	<b>Bima</b>	
<i>Ema/Ame</i>	<i>Ama</i>	Father
<i>Ine/Ende</i>	<i>Ina</i>	Mother
<i>Empo</i>	<i>Ompu</i>	Grandfather
<i>Kae</i>	<i>Sae</i>	Old brother/sister
<b>Polite/Courteous Address</b>		
<b>Manggarai</b>	<b>Bima</b>	
<i>Ite</i>	<i>Ita</i>	You
<i>Hia</i>	<i>Sia</i>	She/he

This similarity could be coincidental, considering that Manggarai was once under the rule of the Bima kingdom, so it is possible that the Manggarai language also absorbs a lot of vocabulary from the Bima language. As Blust (2008: 83) points out, like other areas west of Lesser Sunda, Manggarai was subject to the dominance of the Makassar-speaking kingdom of Goa for several centuries, as well as the sultanate of Bima after about 1727.

The basic paradigm that must be possessed in comparing languages is that all languages worldwide have specific universal characteristics. However, in studying historical-comparative linguistics, the research only focuses on the characteristics that indicate certain similarities, which are seen from the similarity of forms and meanings of the languages being compared. Related languages originating from the same proto will show similarities, such as the similarity of the sound system (phonetics) and sound arrangement (phonological), morphological similarity (the same form of words and grammatical), and syntactic similarity (similarity of relations between words in sentences).

The similarity of form is found in the corresponding phonemic distribution between the compared languages. This similarity can also be strengthened by the similarity of meaning and grammatical similarity of these languages. The similarity of form and meaning is caused by three factors, the direct inheritance factor (from the same proto), the chance factor, and the loan factor. The direct inheritance factor is the basis for determining the word relatives (cognates). Keraf (1984) explains that correspondence is a regular sound change between the languages of relatives. Correspondence can be seen from successive forms of the same word, between corresponding words in relatives' languages, and from areal contacts (borrowing from a donor language).

Basic vocabulary is essential in the grouping of languages in the archipelago. Comparative linguists agree that grouping should be based on qualitative evidence in the form of phonetic retention and innovation. Retention is an element of inheritance from the original language that has not changed in the current language (Fernandez, 1996: 22). However, if there are doubts, it can apply innovative methods, both innovations in basic vocabulary and innovations in grammatical and phonological elements. Innovation is closely related to renewal, namely the existence of phonemic changes derived from the old proto.

After referring to the basic vocabulary, it is uncomplicated to determine whether the languages being compared are experiencing innovation or maintaining their relic form (retention). Furthermore, it needs to determine the related word forms (cognates) between the languages compared by paying attention to the correspondence of sounds or phonemes they have. Each phoneme in the same position is documented in a correspondence set (Keraf, 1984).

One quantitative approach of grouping languages is lexicostatistics (Blust, 2013: 277). Lexicostatistics can be distinguished from glottochronology (Campbell, 2013: 448) even though they have often been used interchangeably (Hymes, 1960: 4). Lexicostatistics is a technique in grouping languages using statistical numbers to determine the percentage of similarities and differences between the languages being compared (Sarah, 1956: 1; Crowley and Bown, 2010; Zhang and Gong, 2016). Meanwhile, Glottochronology focuses on determining the evolutionary relationships between different languages (Campbell, 2013). Glottochronology is carried out by prioritizing the calculation of time (time depth) or the age of relatives' languages using general calculations such as thousands of years (millennium) (Lehmann, 1992). Glottochronology is always juxtaposed with lexicostatistical calculations. The percentage of similarity between these languages determines the distance between the languages being compared.

Four basic assumptions are used as benchmarks to find answers regarding language age, such as first, part of the vocabulary of a language is tough to change when compared to other parts (Renfrew, 1994: 116), including terms for pronouns, numerals, body parts, and geographical features among others (Sarah, 1956: 177); second, retention (endurance) of basic vocabulary is fixed (unchanged) over time (Swadesh, 1952: 452; Crowley and Bown, 2010: 138); third, changes in basic vocabulary in all languages are the same; and fourth, if the percentage of relatives' vocabulary is known, then the separation time between the two languages can be calculated.

The lexicostatistics calculation goes through several necessary steps, such as collecting the basic vocabulary of relatives' languages, determining which pairs of the two languages are cognates, calculating the age or separation time of the two languages, and calculating the error range to determine the appropriate possible split time (Sarah, 1956: 178). The researcher utilizes the 200 Swadesh word list, which can be found in Campbell's work (2013: 449-451) or also can be accessed in ComparaLex website (<https://comparalex.org/>). The basic vocabulary compiled by Swadesh (in Hoijer, 1956: 50) consists of consisting of body parts, numerals, certain objects of nature, simple universal activities. This list offers multiple advantages (Sarah, 1956: 179). Firstly, it includes non-cultural items that are part of the core vocabulary and have been specifically selected for this purpose. Secondly, the list has been used in various studies, allowing for easy comparison of results. Finally, obtaining the complete list in all the studied languages may be difficult, so comparisons may have to be made with fewer items than in the original list.

After collecting vocabulary from the languages being compared, the process continues by determining vocabulary with a cognate form. The determination of the vocabulary of relatives is carried out with the following conditions (Keraf, 1984):

- 1) The pair is identical, that is, a pair of words that all have the same phoneme;
- 2) Pairs that have phonemic correspondence occur regularly and reciprocally between the languages being compared;
- 3) Phonetic resemblance, phonetic resemblance in the same articulatory position; and
- 4) One different phoneme.

To find the percentage of related words, divide the number of cognate words by the number of base words being compared and multiply by 100 (Darman, 2022: 335; Humaidi and Kasmilawati, 2023: 68). The kinship percentage (i.e., the level of relatedness to each other) is calculated using the formula:

$$C = \frac{\text{Number of cognate words}}{\text{Number of vocabularies being compared}} \times 100\%$$

After determining the relative vocabulary based on the provisions above and calculating the percentage of related words, the next step is calculating the separation time (Swadesh, 1952: 456). Calculating the split time is done using the formula:

$$W = \frac{\log C}{2 \log r}$$

$W$  is the time of separation in thousands (millennium) years ago,  $r$  is the retention or percentage constant in 1000 years (index),  $C$  is the percentage of relative words, and  $\log$  means 'logarithm of', so that  $\log C$  means the

logarithm of the percent of probable cognates registered, and  $2 \log r$  means twice the logarithm of the constant (Sarah, 1956: 201).

Considering the assumption that languages cannot separate at a particular time and, of course, over a relatively long period (gradually), it is necessary to calculate the limit error. Calculating the limit error avoids errors in statistical figures by estimating a particular time apart between the languages being compared. During that period, the process of separation between languages slowly occurs. Thus, a statistical method is used to calculate the error range by using the formula:

$$S = \frac{\sqrt{C(1-C)}}{n}$$

$S$  is the standard error in the percentage of relative words,  $C$  is the percentage of related words, and  $n$  is the number of words compared (kin and non-relatives).

After going through a series of steps, it is possible to determine the grouping of the languages being compared. Greenberg (1957) mentions three criteria for classifying languages. The first criterion is non-arbitrary, i.e., the criteria applied to distinguish, relate, and group languages will result in the same grouping. Second, specific criteria, all languages without exception must be included in one of the groups. The third criterion is uniqueness, which means no language can be grouped into more than one group or classification.

There are four methods of classifying language, genetic classification, typological classification, areal classification, and sociolinguistic classification (Parera, 1991). The classification carried out by Swadesh is a form of application of the genetic language classification method, which is carried out based on the correspondence criteria of sound and meaning. Furthermore, Swadesh (1952) proposed levels of subgrouping, as shown in Table 2 below.

Table 3. Different levels of subgrouping

Language Level	Type Split in Ages	Percentage of Relatives
Dialects of a language	0-5	100-81
Languages of a family	5-25	81-36
Families of a stock	25-50	36-12
Stocks of a microphylum	50-75	12-4
Microphyla of a mesophylum	75-100	4-1
Mesophyla of a macrophylum	Up to 100	1-less

Swadesh proposed a classification of languages that are grouped into languages; that is, when two languages are compared, they show 81-100% similarity. Two or more languages are called families if the two languages show similarities between 36-81%. It is called a clump (stock), the languages being compared show 12-36% similarity, and so on according to the description of the table. A family is a type of subgrouping where the members share more than 36 percent of their core vocabularies. Languages that are not closely related but are still believed to have descended from a common ancestor are not considered to be in the same family but in the same stock or phylum (Crowley and Bower, 2010: 139).

## 2. Methods

This study seeks to reconstruct the kinship relationship between Bima and Manggarai. Appropriate methods and techniques are applied to observe the kinship between the two languages. This research is included in diachronic language research. Mahsun (2019) explains that diachronic linguistics is a branch of linguistics that investigates the development of language from one time to another and investigates the comparison of one language with another. This study uses a comparative method approach that is quantitative and qualitative. The quantitative method is applied by using lexicostatistic and geochronological calculations. The results obtained from these quantitative calculations will be used as a reference for hypotheses which are then proven by the results obtained from testing

qualitative methods (Fernandez, 1996: 29). This result is related to the phonemic correspondence of the two languages being compared.

This study uses secondary data provided in Lexirumah and (<https://lexirumah.model-ling.eu/>) and Austronesian Basic Vocabulary Database (ABVD) (<https://abvd.eva.mpg.de/austronesian/>) as a comparison. Lexirumah is a website that provides information about the corpus of Austronesian languages, and also from the Language and Literature Diversity Laboratory based on the 200 Swadesh word list proposed by Morris Swadesh (Kaiping and Klamer, 2018; Kaiping et al., 2019), whereas ABV database contains 327,549 lexical items from 1,701 languages throughout the Pacific region. Most of these languages belong to the Austronesian language family, the most prominent family in the world, containing around 1,000 and 1,200 languages (Greenhill et al., 2008). Each database provides data collected by previous researchers.

In analyzing the data, the researcher applied the descriptive analytical method, which was then followed by the comparison-equalizing technique. At first, the data were analyzed using lexicostatistics and glottochronology techniques to determine the percentage of kinship and the distance between the two languages. Furthermore, the researcher describes the phonemic correspondences between the two languages based on the data obtained from the online database. The researcher also describes the types of phoneme inheritance in the two languages being compared.

### 3. Results and Discussion

#### 3.1. Quantitative Evidence (Lexicostatistics and Glottochronology)

##### 3.1.1. Collecting Basic Vocabulary

Fifty-five words are considered to be cognates of the 200 words compared. Of the 54 cognates, three words from Bima and Manggarai are identical. In contrast, the rest are 51 cognates that have phonemic correspondence.

##### 3.1.2. Calculating the Percentage of Cognate Words

When two languages share words with the same meaning, those words are called homo-semantic cognates (Dyen, 1962: 39). Suppose the lists of two languages contain a significantly higher number of homo-semantic cognates than either language shows with any third language. In that case, it can be concluded that the two languages were probably the same language in the recent past. By knowing the number of cognates, the lexicostatistics can be calculated as follows:

$$C = \frac{\text{Number of cognate words}}{\text{Number of vocabularies being compared}} \times 100\% = \frac{54}{200} \times 100\% = 27\%$$

The result of 27% indicates that Bima and Manggarai are two different languages in the same family. However, it is necessary to test the reliability by finding the separation time and the calculation error period. Quantitative evidence by Fernandez (1996: 43) also shows the percentage of the same level of kinship between Bima and Manggarai, which is 27%. In his research, besides comparing the languages found on the island of Flores, Fernandez also took several languages in the vicinity, one of which was the Bima language.

##### 3.1.3. Calculating The Time Split

After finding the kinship percentage, it can be then calculated the separation period between Bima and Manggarai by applying the formula described above.

$$W = \frac{\log.C}{2 \log.r} = \frac{\log. 0,27}{2 \times \log. 0,805} = \frac{-1,309}{2 \times -0,217} = \frac{1,309}{0,434} = 3,016$$

The value of 0.805 for  $r$  is derived from the second assumption, which states that the rate of lexical replacement in the core vocabulary is more or less stable (Crowley and Bower, 2010: 148). However, the rate of lexical replacement in peripheral vocabulary is not constant at all and may vary depending on the nature of cultural contact between speakers of different languages. This second assumption has been tested in 13 languages for which written records exist over long periods. The results show that, on average, there is 80.5 percent vocabulary retention every

1,000 years. In other words, after 1,000 years, a language would have lost approximately one-fifth of its original basic vocabulary and replaced it with new forms (Lees, 1953: 124; Campbell, 2013: 452).

With that calculation, the distance between the two languages is 3,016 thousand years. Alternatively, in other words, the Bima and Manggarai languages were the same language about 3000 years ago. Then, the two languages separated from the proto-language around 994 BC or the X century BC.

### 3.1.4. Calculating Limit Error

The calculation of the limit error is carried out to minimize the possibility of errors in the calculation of the time split between the two languages (Lees, 1953: 124). In order to avoid an error in the calculation of the separation period with a definite result, it is necessary to recalculate to find out the estimated period of the separation. On the pretext that if the research is conducted by narrowing the year, it means that it has also reduced the possibility of its accuracy.

$$S = \frac{\sqrt{C(1-C)}}{n}$$

$$S = \frac{\sqrt{0,27(1-0,27)}}{200} = \frac{\sqrt{0,27 \times 0,73}}{200} = \frac{\sqrt{0,1971}}{200} = \sqrt{0,00098} = 0,03130 = 0,03$$

The results of the standard errors are then summed with the percentage of relative words to obtain the new value of  $C$ , so  $0.27+0.03 = 0.3$ . By knowing the new value of  $C$ , it is possible to recalculate the split time using the split period formula.

$$W = \frac{\log.C}{2 \log.r} = \frac{\log. 0,3}{2 \times \log. 0,805} = \frac{-1,204}{2 \times -0,217} = \frac{1,204}{0,434} = 2,774 \text{ years ago}$$

After knowing the new value of  $W$ , the old split time is then subtracted by the new value of  $W$ , so that:

$$W_1 - W_2 = 3,016 - 2,774 = 242$$

This number must be added and subtracted by the old value of  $W$  to obtain the age or time of departure of the two languages. Thus, the following conclusions can be drawn using the formula for calculating the standard error range (70%).

- 1) Bima and Manggarai were a single language 3,016±242 years ago,
- 2) Bima and Manggarai were a single language from 3,258-2,774 years ago, and
- 3) The Bima and Manggarai languages began to separate from a proto-language between 1235-751 BC (calculated from 2023).

## 3.2. Qualitative Evidence (Phonemic Correspondence and Phonetic Change)

### 3.2.1. Sound Correspondence between Bima and Manggarai Language

The most effective way to find similarities is to look for consistent and meaningful correspondences in lexical items, morphology, and grammar, which are also used for reconstruction (Crowley and Bower, 2010: 109). Correspondences in lexical items should be regular and not limited to a single area of the language. Shared suppletive forms are more indicative of a relationship than random shared items. When it comes to comparing remotely related languages, the sounds used in their respective proposals are usually very similar, if not identical (Campbell, 2013: 350).

Based on the cognates of the two languages, several sound correspondences were found between Bima and Manggarai. The sound correspondences include /w~v/, /f~p/, /d~t/, /r~l/, and /a~e/.

#### a) Sound Correspondence [w]~[v]

The sound correspondence /w~v/ can occur in the ultimate position as in [awu] and [avu], as well as in the penultimate position as in [wari] and [valek], [wadu] and [vatu], etc.



Table 4. Sound Correspondence [w]~[v]

Gloss	Bima	Manggarai
Ash	[awu]	[avu]
Return	[wari]	[valek]
Stone	[wadu]	[vatu]
Fruit	[wua]	[vua]
Moon	[wura]	[vulaŋ]
Right	[wana]	[vanaŋ]

b) *Sound Correspondence [f]~[p]*

It can be found in the ultimate position as in [afi] and [api], and the penultimate position as in [fuʔu] and [puʔu].

Table 5. Sound Correspondence [f]~[p]

Gloss	Bima	Manggarai
Fire	[afi]	[api]
Tree	[fuʔu]	[puʔu haju]

c) *Sound Correspondence [d]~[t]*

This type of sound correspondence occurs at the ultima and penultimate positions, such as [wadu] and [vatu] in the ultimate position, and [dana] and [tana] in the penultimate position.

Table 6. Sound Correspondence [d]~[t]

Gloss	Bima	Manggarai
Stone	[wadu]	[vatu]
Lice	[hudu]	[hutu]
Eye	[mada]	[mata]
Die	[made]	[mata]
Liver	[ade]	[ati]
Soil	[dana]	[tana]
Dull	[dumpa]	[dempul]

d) *Sound Correspondence [r]~[l]*

Sound correspondence [r]~[l] occur in the ultima position as in the examples of [wari] and [valek], as well as in the penultimate position as in the examples of [rima] and [lime].

Table 7. Sound Correspondence [r]~[l]

Gloss	Bima	Manggarai
Float	[karente]	[lenteŋ]
Return	[wari]	[valek]
Star	[ntara]	[ntala]
Moon	[wura]	[vulaŋ]
Hand	[rima]	[lime]

e) *Phoneme Correspondence [a]~[e]*

Unlike the previous form of sound correspondence, the sound correspondence /a~e/ in Bima and Manggarai is only found in the ultima position as shown in the examples [siʔa] and [ciʔe], [rima], and [lime].

Table 8. Sound Correspondence [a]~[e]

Gloss	Bima	Manggarai
Salt	[sia]	[ciʔe]
You (polite)	[ita]	[ite]
Mother	[ina]	[ine]
Die	[made]	[mata]
Hand	[rima]	[lime]

### 3.2.2. Types of Sound Change

The inheritance of proto phonemes into Bima and Manggarai occurs in several inheritance patterns. The proto-Austronesian in this research is retrieved from Greenhill et al. (2008) in the Austronesian Basic Vocabulary Database and Lexirumah and also supported with Proto Central Malayo Polynesian reconstructed by Blust (1993: 280-284).

#### a) Linear Inheritance

Linear inheritance means that today's inherited languages still retain their prototype form.

Table 9. Linear Inheritance

Index	Gloss	PAN	Bima	Manggarai
4	1sg	*aku	-	[aku]
6	Child	*anak	-	[anak]
19	Father	*ama	[ama]	-
126	Lelaki	*laki	-	[laki]
132	Five	*lima	[lima]	[lima]
138	Eye	*mata	-	[mata]
191	Three	*təlu	-	[təlu]

From the data shown above, Manggarai retains much of its prototype form. Of the ten data showing direct (linear) inheritance, seven words are found in Manggarai, which retains their complete form or, in other words, without any change in form. Meanwhile, the other three data are included in the Bima vocabulary, which maintains the prototype form.

#### b) Inheritance with Phoneme Change

Changes can occur when the phonemes in the proto-language change the derived language.

Table 10. Inheritance with Phoneme Change

Index	Gloss	PAN	Bima	Manggarai
1	Ash	*abu	[awu]	[avu]
8	Dog	*asu	-	[acu]
17	Return	*balik	-	[valek]
19	Father	*ama	[ama]	[əma]
21	New	*bayu	-	[vəru]
23	Stone	*batu	[wadu]	[vatu]
39	Feather	*bulu	-	[vulu]

51	Lake	*dano	-	[sano]
72	Dig	*gali	[ŋari]	-
74	Scratch	*garuk	-	[kərok]
75	Fat	*gəmək	-	[gemok]
83	Nose	*ʔiduŋ	-	[isuŋ]
89	Rain	*udan	-	[usaŋ]
98	Sew	*jait	-	[jaik]
108	Right	*wanan	-	[vanaŋ]
117	Nail	*kuku	-	[vuku]
120	Lice	*kutu	[hudu]	[hutu]
143	Drink	*inum	-	[inuŋ]
166	One	*sa	-	[ca]
170	Who?	*sai	-	[cai]
173	Husband	*laki	[rahi]	-
180	Soil	*tana	[dana]	-
193	Thin	*nipis	-	[mipis]
195	Stick	*toŋkat	-	[doŋkar]

By looking at the data in the table, the following proto phoneme inheritance can be formulated as follows:

- 1) The proto phoneme \*b is reflected as the phoneme /w/ in Manggarai and phoneme /v/ in Manggarai;
- 2) The proto phoneme \*s is reflected as the phoneme /c/ in Manggarai, as in [acu], [ca], and [cai];
- 3) The proto phoneme \*n is reflected as the phoneme /ŋ/ in Manggarai;
- 4) The proto phoneme \*t is reflected as the phoneme /d/ in Bima and /t/ in Manggarai; and
- 5) The proto phoneme \*d is reflected as the phoneme /s/ in Manggarai;

c) *Inheritance with Phoneme Omission*

Phoneme omission can occur in the ultima and penultimate positions. In the ultima position, only one gloss occurs, namely the proto form \*əpat, and the phoneme deletion in Manggarai becomes [pat]. The unique feature of proto-phoneme inheritance to Manggarai is the omission of phonemes in the final syllable. This omission does not have a fixed rule, but from the data shown in the table, the omission occurs in closed consonant phonemes such as nasal /ŋ/ and /n/, and stop consonants /t/, /k/, and /s/.

Table 11. Inheritance with Phoneme Omission

Index	Gloss	PAN	Bima	Manggarai
6	Child	*anak	[ana]	-
7	Wind	*aŋin	[aŋi]	-
35	Animal	*binataŋ	[binata]	-
67	Two	*duwa	[dua]	-
70	Four	*əpat	-	[pat]
73	Salt	*sira	[siʔa]	-
108	Right	*wanan	[wana]	-
122	Sky	*laŋit	[laŋi]	-
140	Die	*matay	-	[mata]
144	Mouth	*mulut	-	[muʔu]
150	Hot	*panas	[pana]	-
179	Rope	*tali	[ai]	-
193	Thin	*nipis	[nipi]	-
196	Old	*tuha'	[tuʔa]	[tuʔa]

d) *Inheritance with Phonemes Addition*

The example below adds a voiced alveolar nasal /n/ at the beginning of the word, both in Bima and Manggarai. In both languages, a form of phonological innovation was found in the form of an obstructive nasal consonant cluster at the initial position of the word.

Table 12. Inheritance with Phonemes Addition

Index	Gloss	PAN	Bima	Manggarai
51	Lake	*dano	[ndano]	-
177	Year	*tahun	-	[ntauŋ]

e) *Inheritance with Phoneme Change and Deletion*

The examples shown below show a change in phonemes accompanied by the omission of phonemes in Bima and Manggarai. As in inheritance in the form of phoneme omission, the omission accompanied by a phoneme change occurs in the last syllable. However, there were also forms of omission of phonemes at the beginning of the syllable (ultima) and the end of the syllable (penultima) simultaneously, such as the proto form \*hatay/ being /ade/ in Bima and /ati/ in Manggarai. The first phoneme is omitted in the penultimate position as in the proto form to /huu/ in Bima.

Table 13. Inheritance with Phoneme Change and Deletion

Index	Gloss	PAN	Bima	Manggarai
2	Water	*wayer	-	[vae]
10	Fire	*apuy	[afi]	[api]
27	Seed	*bənih	-	[vini]
29	Swim	*laŋuj	-	[laŋe]
37	Fruit	*buah	[wua]	[vua]
38	Moon	*bulan	[wura]	-
52	Blood	*dayah	-	[dara]
69	Tail	*əkɔr	-	[iko]
70	Four	*əpat	[upa]	-
82	Liver	*hatay	[ade]	[ati]
83	Nose	*ʔiduŋ	[ilu]	-
89	Rain	*udan	[ura]	-
95	Wife	*binay	-	[vina]
101	Fall	*dabuh	[mabu]	-
117	Nail	*kuku	[huʔu]	-
118	Skin	*kulit	[huri]	-
140	Die	*matay	[made]	-
147	Breathe	*napas	[nawa]	-
155	Squeeze	*pəras	-	[kəra]
158	Think	*pikir	[fiki]	-
179	Rope	*tali	[ai]	-
185	Egg	*tolur	[dolu]	-

f) *Inheritance with Phoneme Change and Addition*

The addition of phonemes only occurs in Bima, and the adder appears in the initial position, namely the nasal /m/ sound in [mbei] and [mbeça], and the addition of the phoneme /i/ in the word [ica]. Data numbers 32 and 22 show an addition of obstructive nasal consonants in the initial position.

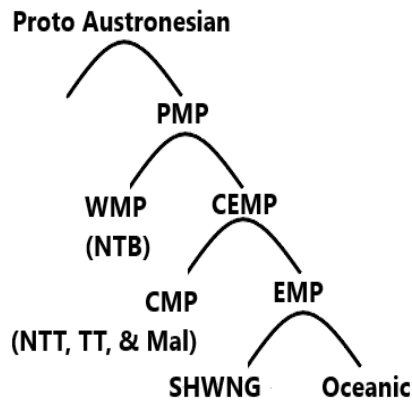
Table 14. Inheritance with Phoneme Change with Phoneme Addition

Index	Gloss	PAN	Bima	Manggarai
32	Give	*beri	[mbei]	-
22	Wet	*batʔah	[mbeça]	-
166	One	*sa	[ica]	-

3.2.3. Language Classification

After knowing the percentage value of cognate words and calculating the separation time between the two languages, a conclusion can be drawn regarding the relationship between the two languages. If the percentage of related words is high, it may indicate that the languages being compared belong to the same group. If two languages have very similar core vocabularies, it can be assumed that they have diverged recently and belong to a lower-level subgroup. However, if their core vocabularies are pretty different, it can be assumed that they diverged much earlier and belong to a much higher level of subgrouping (Bower, 2010). This principle calculates the percentage of cognate words by comparing the number of highly similar words with 200 Swadesh word lists. The results are then applied in the language subgroup depicted in Table 2 above.

Figure 1. Austronesian Language Genealogy (Blust, 1978)



Using lexicostatistical principles, Swadesh proposed a classification to determine the kinship of languages. It can be concluded that Bima and Manggarai are two languages from the same larger family (stock). Language family refers to all languages descended from a common ancestor language (Crowley and Bower, 2010: 139). It includes languages that are related to each other, whether closely or distantly. However, according to a lexicostatistical classification, a family is a specific level of subgrouping where members share more than 36% of their core vocabularies. Languages that share lesser degrees of relationship but are still presumed to have descended from a common ancestor are not considered to be in the same family but in the same stock or phylum. According to Salzner (in Keraf, 1984: 205), the Austronesian languages are divided into two prominent families, such as the Western Austronesian languages (Indonesian languages), divided into Hesperonesian or Western Indonesian and Eastern Indonesian languages such as Moluccas (Tyron, 1995: 79); and the Eastern Austronesian languages (Oceanian

languages); are divided into Heonesian languages (Polynesian and Micronesian) and Melanesian languages (Melanesia and the east coast of West Papua).

Languages spoken in the western region of the island of Sumbawa (located in Nusa Tenggara Barat), as well as in Sulawesi, western Indonesia, Malaysia, and the Philippines, fall under the category of Western Malayo-Polynesian languages (WMP). On the other hand, languages spoken in the eastern region of the island of Sumbawa, including the Austronesian languages found in Nusa Tenggara Timur, Timor Timur, Maluku, and Irian Jaya, are classified as Central-Eastern-Malayo-Polynesian (CEMP) languages (Grimes, 1997: 6).

#### 4. Conclusion

The Bima language and the Manggarai language are grouped into clumps (stock) with a kinship percentage of 27%. These two languages were once a single language 3,258-2,774 years ago, then began to split from a proto-language between 1236-752 BC (calculated from 2022). The kinship of these two languages can be proven by their sound correspondence. Some of the sound correspondences include [f]~[p], [d]~[t], [r]~[l], and [a]~[e]. Many of the vocabularies of Manggarai still retain their prototype form when compared to Bima. Thus, it is possible that a group of people once spoke a particular proto-language, then later spread and separated to Sumbawa Island and the western part of Flores Island. Due to certain factors, such as environmental influences, the two languages change word form but retain their meaning. Lexicostatistical data should be supplemented with other forms of evidence, such as archaeology, comparative ethnography, and linguistic paleontology. These various lines of the study corroborate and refine one another, providing a more complete picture of the subject at hand (Swadesh, 1952: 453). Therefore, it is hoped that future researchers will combine data from multiple disciplines to strengthen the result of this present study.

#### Data availability

The datasets analysed during the current study are available in the zenodo repository: <https://zenodo.org/records/10677379>.

#### References

- Arafiq. (2020). The Syntax of Personal Pronouns in the Bima Language. *International Linguistics Research*, 3(2), 13-21. <https://doi.org/10.30560/ilr.v3n2p13>
- Astuti. (2019). Eksistensi Budaya Rimpu Masyarakat Suku Mbojo di Kabupaten Bima. *JUPE: Jurnal Pendidikan Mandala*, 4 (25), 111-114. <http://dx.doi.org/10.58258/jupe.v4i5.843>
- Badan dan Peta Bahasa di Indonesia. *Bima (Mbojo): Provinsi Nusa Tenggara Barat (Nusa Tenggara Barat)*. Accessed in April 20th, 2022, from <https://petabahasa.kemdikbud.go.id/infobahasa.php?idb=210>
- Blanc, F. L. (2024). *People Cluster: Flores-Sumba-Alor*. Accessed in February 15th, 2024, from <https://peoplegroups.org/>
- Blust, R. (1993). Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics*, 32(2), 241–293. <https://doi.org/10.2307/3623195>
- Blust, R. (2008). Is there a Bima-Sumba Subgroup? *Oceanic Linguistics*, 47(1), 45-113. <https://doi.org/10.1353/ol.0.0006>
- Blust, R. (2013). *The Austronesian Languages*. The Australian National University: SEAsian Mainland Languages E-Series (SEAMLES). <http://hdl.handle.net/1885/10191>
- Budasi, I. G. (2018). Kekerabatan Secara Ekolinguistik Bima di NTB dan Bahasa Sabu di NTT. *Deskripsi Bahasa*, 1(2), 130-137. <https://doi.org/10.22146/db.v1i2.330>
- Brandes, J. L. A. (1884). *Bujdrage tot de Verglijkende Klankeer der Westere Afdeeling van de Maleische Polynesische Taalfamilie*. Utrecht.
- Bustan, F., Semiun, A., & Bire, J. (2020). The Cultural Conceptualisation of Manggarai People on The Origin of Life: Cultural Linguistic Analysis. *Academic Journal of Educational Sciences*, 4(1), 13-19. <https://doi.org/10.35508/ajes.v4i1.2351>
- Campbell, L. (2013). *Historical Linguistics: An Introduction (NED-New edition, 3)*. Edinburgh University Press. <http://www.jstor.org/stable/10.3366/j.ctt1g0b5gq>

- Crowley, T., & Bowern, C. (2010). *An Introduction to Historical Linguistics (4th Edition)*. Oxford University Press.
- Daeng, H. (2013). Manggarai Daerah Sengketa antara Bima dan Goa. *Humaniora*, 0(2), 16-22. <https://doi.org/10.22146/jh.1970>
- Darman, F. (2022). Leksikostatistik Bahasa Alune dan Wemale di Pulau Seram, Maluku: Lexicostatistics of Alune and Wemale Languages on Seram Island, Maluku. *TOTOBUANG*, 10(2), 331-343. <https://doi.org/10.26499/totobuang.v10i2.420>
- Djakariah, D., Utomo, S. S., & Dambut, K. G. (2019). Sejarah Masuknya Kekuasaan Bima di Nuca Lale Pada Tahun 1727-1930. *Jurnal Sejarah*, 16(2), 36–51. <http://publikasi.undana.ac.id/index.php/js/article/view>
- Dyen, I. (1962). The Lexicostatistical Classification of the Malayopolynesian Languages. *Language*, 38(1), 38–46. <https://doi.org/10.2307/411187>
- Dyen, I. (1982). The Present Status of Some Austronesian Subgrouping Hypothesis. *Papers from the TICAL*. <https://malaycivilization.com.my/omeka/items/show/151077>
- Erb, M. (1997). Contested Time and Place: Constructions of History in Todo, Manggarai (Western Flores, Indonesia). *Journal of Southeast Asian Studies*, 28(1), 47-77. <https://doi.org/10.1017/S0022463400015174>
- Esser, S. J. (1938). *Atlas van Tropisch Nederland*. Batavia Centrum.
- Fernandez, I. Y. (1996). *Relasi Historis Kekerabatan Bahasa Flores: Kajian Linguistik Historis Komparatif Terhadap Sembilan Bahasa di Flores*. Jakarta: PT Nusa Indah.
- Fernandez, I. Y. (2007). Inventarisasi Bahasa-Bahasa Daerah Di Propinsi Nusa Tenggara Timur. *Humaniora*, 19(3), 241-247. <https://doi.org/10.22146/jh.907>
- Greenberg, J. H. (1957). *Essay in Linguistics*. Chicago: University of Chicago Press.
- Greenhill, S.J., Blust, R., & Gray, R.D. (2008). The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics. *Evolutionary Bioinformatics*, 4:271-283.
- Grimes, C. E. (1997). *A Guide to the People and Languages of Nusa Tenggara*. Artha Wacana Press.
- Harcourt, H. M. (2010). *Definition of Linguistic Stock in American English [Webster's New World College Dictionary, 4th Edition]*. Accessed in 16th February, 2024, from <https://www.collinsdictionary.com/us/dictionary/>
- Hendrokumoro. (2019). Hubungan Kekekerabatan Bahasa Ma'anyan dan Bahasa Malagasi. *Jurnal Metalingua*, 17(2). <http://dx.doi.org/10.26499/metalingua.v17i2.312>
- Hojjer, H. (1956). Lexicostatistics: A Critique. *Language*, 32(1), 49–60. <https://doi.org/10.2307/410652>
- Humaidi, A., & Kasmilawati, I. (2023). Deah, Maanyan, and Banjarnese Languages Kinship in Tabalong Regency South of Kalimantan. *Tunas: Jurnal Pendidikan Guru Sekolah Dasar*, 9(1), 67–74. <https://doi.org/10.33084/tunas.v9i1.6009>
- Hymes, D. H. (1960). Lexicostatistics So Far. *Current Anthropology*, 1(1), 3–44. <https://doi.org/10.1086/200074>
- International Mission Board. (2022). *People Cluster: Flores-Sumba-Alor*. Accessed in April 27th, 2022, from <https://peoplegroups.org/explore/ClusterDetails.aspx?rop2=C0075#topmenu>
- Kaiping, G. A., & Klamer, M. (2018). LexiRumah: An online lexical database of the Lesser Sunda Islands. *PLoS ONE*, 13(10): e0205250. <https://doi.org/10.1371/journal.pone.0205250>
- Kaiping, G. A., Edwards, O., and Klamer, M. (2019). *LexiRumah 3.0.0*. Leiden: Leiden University Centre for Linguistics. Accessed on 18th February, 2024, from <https://lexirumah.model-ling.eu/>
- Karim, S. (2022). Personal Naming System of Bima Tribe in West Nusa Tenggara: Usage and Naming Behaviour. *Register Journal*, 15(1), 20-41. <https://doi.org/10.18326/rgt.v15i1.20-41>
- Keraf, G. (1984). *Linguistik Bandingan Historis*. Jakarta: Penerbit PT Gramedia Pustaka Utama.
- Laboratorium Kebinekaan Bahasa dan Sastra. (2022). *Kosakata Swadesh*. Accessed in April 27th, 2022, from <https://laBimaineka.kemdikbud.go.id/bahasa/sebaran>
- Lees, R. B. (1953). The Basis of Glottochronology. *Language*, 29(2), 113–127. <https://doi.org/10.2307/410164>
- Lehmann, W. P. (1992). *Historical Linguistics an Introduction (3rd Edition)*. Routledge. <https://doi.org/10.4324/9780203416433>
- Mahsun. (2019). *Metode Penelitian Bahasa (Tahapan, Strategis, Metode, dan Tekniknya)*. Depok: Rajawali Pers.
- Mbete, A. M. (1990). *Rekonstruksi Protobahasa Bali-Sasak-Sumbawa*. [Doctoral Dissertation in Universitas

- Indonesia]. Accessed in February 15th, 2024, from <https://lib.ui.ac.id/detail?id=74944>
- Moses, F. (2018). *Mengenal Manggarai di Nusa Tenggara Timur*. Jakarta: Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan.
- Parera, J. D. (1991). *Kajian Linguistik Umum Historis Komparatif dan Tipologi Struktural*. Jakarta: Penerbit Erlangga.
- Renfrew, C. (1994). World Linguistic Diversity. *Scientific American*, 270(1), 116–123.  
<https://doi.org/10.1038/scientificamerican0194-116>
- Sanjaya, I. P. K., Rupa, I. W., & Tangkas, M. V. (2018). *Bugis dan Bajo di Labuhan Bajo Manggarai Barat Perspektif Sejarah dan Budaya*. Penerbit Kepel Press.
- Sarah C. G. (1956). The ABC'S of Lexicostatistics (Glottochronology). *WORD*, 12(2), 175-210.  
<https://doi.org/10.1080/00437956.1956.11659599>
- Sulistyo, B. (2014). Multikulturalisme di Bima pada Abad X – XVII. *Paramita: Historical Studies Journal*, 24(2).  
<https://doi.org/10.15294/paramita.v24i2.3120>
- Sunendar, D. (2019). *Sekapur Sirih Bahasa dan Peta Bahasa di Indonesia (6th edition)*. Bahasa dan Peta Bahasa di Indonesia. Accessed in February 14th, 2024, from  
<https://petabahasa.kemdikbud.go.id/infobahasa.php?idb=210>
- Suwondo, B. (1978). *Sejarah Daerah Nusa Tenggara Timur*. Departemen Pendidikan dan Kebudayaan, Pusat Penelitian Sejarah dan Budaya, Proyek Penelitian dan Pencatatan Kebudayaan Daerah.
- Swadesh, M. (1952). Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society*, 96(4), 452–463.  
<http://www.jstor.org/stable/3143802>
- Syamsuddin A. R. (1996). *Kelompok Bima-Sumba: Kajian Linguistik Historis Komparatif* [Doctoral Dissertation]. Retrived from Universitas Padjadjaran Bandung.
- Tim Penyusun Kamus Mbojo Indonesia. (2015). *Kamus Mbojo Indonesia*. Kantor Bahasa Provinsi NTB.
- Tryon, D. (1995). *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies*. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110884012>
- Zhang, M., & Gong, T. (2016). How Many Is Enough? —Statistical Principles for Lexicostatistics. *Front. Psychol*, 7(1916). <https://doi.org/10.3389/fpsyg.2016.01916>