IJLHE: International Journal of Language, Humanities, and Education
ISSN: 2986-0369 (e) I 2963-4520 (p)
2026, Vol. 9, No. 2, page 297-308
https://jurnal.stkippgribl.ac.id/index.php/ijlhe/index

Lexical and Phonological Reflexes of Proto-Austronesian in Sanggar, Bima Regency: A Historical-Comparative Linguistic Study

Ni Putu Dania Meiliana Taurusia[1*], I Gede Budasi[1], Nyoman Karina Wedhanti[1]

[1]Universitas Pendidikan Ganesha, Bali, Indonesia
dania@student.undiksha.ac.id *

ABSTRACT

This study aimed to examine the lexical relationship and phonological inheritance from Proto-Austronesian (PAn) to the Sanggar language spoken in Bima Regency, West Nusa Tenggara. The study employed a sequential explanatory mixed-method design within a historical-comparative linguistic framework. The primary data consisted of 200 basic vocabulary items collected from three native speakers of the Sanggar language through interviews and audio recordings. The secondary data comprised PAn etyma documented in English Finderlist of Reconstructions in Austronesian Languages by Wurm and Wilson (1975). The quantitative analysis used a lexicostatistical method to calculate the percentage of lexical relationship, while the qualitative analysis examined phonological inheritance, regular sound correspondences, and phonological changes. The findings revealed 50 cognate words out of 200 vocabulary items, indicating a lexical relationship of 25.0%. Of these cognates, four showed full retention, while 46 showed retention with phonological changes. The identified phonological changes included sound deletion, sound insertion, vowel changes, consonant changes, lenition, and metathesis. These findings show that the Sanggar language shares cognate vocabulary with the selected PAn reconstructions, although most of the identified cognate words have undergone phonological changes.

ARTICLE INFO

Keywords:
lexical relationship;
proto-austronesian;
phonological inheritance;
phonological change;
sanggar language

How to Cite in APA Style:
Taurusia, N. P. D. M., Budasi, I. G., & Wedhanti, N. K. (2026). Lexical and Phonological Reflexes of Proto-Austronesian in Sanggar, Bima Regency: A Historical-Comparative Linguistic Study. *IJLHE: International Journal of Language, Humanities, and Education, 9*(2), 297-308. https://doi.org/10.52217/yj8td552

This is an open access article under CC-BY-NC 4.0 license.

INTRODUCTION

Languages change over time as part of their historical development. Historical-comparative linguistics is used to study relationships between languages by comparing words, identifying cognates, and examining regular sound

correspondences (Campbell, 2013; Crowley & Bowern, 2010). Through this approach, researchers can investigate how languages develop from their ancestral languages and identify which linguistic features are retained or changed. Proto-Austronesian (PAn) is a reconstructed ancestral language that helps researchers trace the historical development of Austronesian languages, including those spoken in Indonesia (Blust & Trussel, 2013).

The Sanggar language is spoken in Sanggar District, Bima Regency, West Nusa Tenggara, and contributes to the linguistic diversity of eastern Indonesia. Previous studies have examined historical relationships and phonological inheritance in eastern Indonesian languages. Budasi (2018) examined the ecolinguistic relationship between Bima and Sabu, including cognate vocabulary and reflexes of Proto-Austronesian etyma. Budasi (2012) investigated phonological evidence separating and uniting Mamboro and Proto Wanokaka–Anakalang in Sumba. These studies provide regional comparative context, although neither directly examines Proto-Austronesian reflexes in Sanggar. Anggayana et al. (2020) investigated lexical relationships among Lipang, Langkuru, and Waisika in Alor, while Salahuddin (2023) examined lexical relationships and sound correspondences between Manggarai and Bima.

Although previous studies have examined lexical relationships and phonological inheritance in eastern Indonesia, the studies reviewed here do not provide a detailed account of individual Proto-Austronesian reflexes in Sanggar. Budasi (2018), for example, focused on Bima and Sabu. A closer examination of Sanggar cognates and their phonological differences from selected reconstructions can extend this regional documentation. Identifying cognates and regular sound correspondences is central to historical-comparative analysis because these features help establish historical relationships and trace changes in inherited words (Campbell, 2013; Crowley & Bowern, 2010). The present study examines these features to describe lexical correspondences and phonological retention and change in Sanggar.

Based on this research gap, this study focuses on three main aspects: the percentage of lexical relationship between Proto-Austronesian and the Sanggar language, the phonological reflections of PAn in Sanggar, and the types of phonological changes identified in the data. Addressing these issues, the present study contributes to earlier regional research by providing a preliminary inventory of Sanggar lexical forms compared with selected PAn reconstructions and describing the phonological differences identified in the comparison. Unlike studies primarily concerned with broader genetic relationships or regional proto-language reconstruction, this research focuses on the individual lexical reflexes found in a 200-item Swadesh list, collected through interviews and audio recordings with native speakers of the Sanggar language. The data are then compared with PAn etyma recorded in English Finderlist of Reconstructions in Austronesian Languages by Wurm and Wilson (1975). Using a historical-comparative linguistic approach, this study is expected to

contribute to the documentation of the Sanggar language by explaining its lexical relationship with PAn and describing the phonological features that have been retained or changed over time.

METHOD

This study employed a sequential explanatory mixed-method design combined with a historical-comparative linguistic approach (Creswell & Plano Clark, 2007). The study used quantitative analysis to determine the percentage of lexical relationship between Proto-Austronesian (PAn) and the Sanggar language, followed by qualitative analysis to examine phonological inheritance, regular sound correspondences, and phonological changes (Campbell, 2013; Crowley & Bowern, 2010).

The primary data consisted of 200 basic vocabulary items from a 200-item Swadesh basic vocabulary list (Swadesh, 1952), collected from three native speakers of the Sanggar language in Sanggar District, Bima Regency, West Nusa Tenggara. The participants were selected through purposive sampling, with one primary informant and two supporting informants who helped verify the accuracy and consistency of the lexical and phonological data. The secondary data consisted of reconstructed PAn etyma obtained from English Finderlist of Reconstructions in Austronesian Languages compiled by Wurm and Wilson (1975).

The research instruments included the researcher as the primary instrument, a vocabulary elicitation guide based on the 200-word Swadesh list, audio recording devices, a basic vocabulary list, a data inventory table, and a list of reconstructed PAn etyma. Data were collected through direct interviews with the informants. Their responses were audio-recorded, transcribed, and verified with the supporting informants to ensure the accuracy of the lexical forms and pronunciations. The use of basic vocabulary for lexical comparison follows the lexicostatistical tradition discussed by Swadesh (1952, 1955).

Data analysis was conducted in two stages. First, quantitative analysis employed the lexicostatistical method through a bottom-up approach. The Sanggar vocabulary items were compared with their corresponding PAn forms to identify cognates based on their meanings and phonological forms. The percentage of lexical relationship was calculated using the calculation described by Mbangi and Marafad (2018):

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

In this formula, **C** represents the percentage of lexical relationship, **k** represents the number of cognate words identified, and **n** represents the total number of

vocabulary items compared (Mbangi & Marafad, 2018). The calculation is performed by dividing the number of identified cognate words by the total number of vocabulary items and then multiplying the result by 100 percent (Mbangi & Marafad, 2018).

Second, qualitative analysis employed a historical-comparative method through a top-down approach. The reconstructed PAn etyma were compared with their corresponding Sanggar forms to identify candidate sound correspondences in comparable phonological positions. Historical-comparative principles informed the analysis of retention and sound change (Anttila, 1972; Campbell, 2013; Crowley & Bowern, 2010; Millar & Trask, 2015), while phonological terminology was informed by Katamba (1989). The analytical categories considered were assimilation, dissimilation, lenition, fortition, deletion, metathesis, haplology, monophthongization, diphthongization, insertion, palatalization, nasalization, cluster reduction, rhotacism, and spirantization. These categories served as descriptive labels for examining the dataset; they were not treated as a single taxonomy proposed jointly by the cited authors. Finally, the cognate items were classified into two patterns of phonological inheritance: full retention and retention with phonological changes.

To ensure trustworthiness, the recorded lexical forms and pronunciations were cross-checked and verified with the native speakers, particularly the supporting informants. The verified Sanggar vocabulary was then compared with the reconstructed PAn forms documented by Wurm and Wilson (1975). This process helped ensure the consistency of the data and supported the accuracy of the lexical and phonological analysis.

RESULTS

Results

Based on the analysis of the 200 basic vocabulary items in the Swadesh list, the lexical relationship between the Sanggar language and Proto-Austronesian (PAn) was found to be 25.0% of the 200 vocabulary items analyzed. The complete data are presented in Table 1 below.

Table 1. Percentage of the Lexical Relationship of the Sanggar Language

Language	Number of Words Analyzed	Full Retention and Retention with Phonological Change	Lexical Innovation (Non-Cognates)	Percentage of Relationship
Sanggar	200	50	150	25.0%

The data in Table 1 show that out of the 200 basic vocabulary items in the Swadesh list analyzed in the Sanggar language, 50 words (25.0%) can be identified as cognates with Proto-Austronesian (PAn). Meanwhile, 150 words (75.0%) were not identified as cognates with the reconstructed PAn forms used in this study. This

indicates that although the Sanggar language has retained some basic vocabulary that can be traced to Proto-Austronesian, it has also undergone significant lexical development through processes of innovation.

Types of Phonological Inheritance

The analysis identified two patterns of phonological inheritance in the Sanggar language: full retention and retention with phonological changes. Of the 50 cognate words, four showed full retention, while 46 showed retention with phonological changes. These results indicate that most of the identified cognate forms have undergone phonological changes.

Full Retention

Full retention refers to cognate forms that show no identifiable phonological changes between the reconstructed PAn form and the corresponding Sanggar form. Four words were classified as full retention: *ama* 'father', *ina* 'mother', *asu* 'dog', and *tua* 'old'.

Table 2. Example of Full Retention in the Sanggar Language

No.	English	Sanggar form	Reconstruction and source attribution
1	father	ama	*ama (PANDYPMPL, Dyen 1953a (sc))
2	mother	ina	*ina (PANDYTV, Dyen 1949 (sc))
3	dog	asu	*asu (PANS, Stresemann 1927)
4	old	tua	*tua (PANS, Stresemann 1927)

Note. Historical source attributions are reported indirectly through Wurm and Wilson (1975), the compilation identified in the study. The original historical works are not presented as directly consulted sources. Attributions dated after 1975 require separate source verification, as indicated in the Word comments.

As shown in Table 2, the four forms retain the same phonological structure in the comparison used in this study. These examples demonstrate that some basic vocabulary in Sanggar remains unchanged in relation to the selected PAn etyma.

Retention with Phonological Changes

Retention with phonological changes occurs when Sanggar words still have a relationship with their PAn forms, although some of their sounds have changed. In this study, 46 cognate words were found in this category. These words were then grouped based on the number of phonological changes they underwent.

A. Retention with a Single Phonological Change

There are 17 words that involve a single phonological process in their inheritance, such as vowel changes or the deletion of a phoneme.

Table 3. Example of Retention with a Single Phonological Change

No.	English	Sanggar form	Reconstruction and source attribution
1	wife	sawa	*asawa (PANC, CAPELL 1943)
2	name	ngarkan	*ngaran (PANC, CAPELL 1943)
3	hand	nima	*lima (PANC, CAPELL 1943)
4	liver	ate	*atay (PANC, CAPELL 1943)
5	feather	wulu	*bulu (PANC, CAPELL 1943)
6	wood	kaju	*kaju' (PAND, Dempwolff 1983)
7	fire	api	*apuy (PANDYPMPL, Dyen 1953a (sc))
8	stone	watu	*batu (PANDYTV, DYEN 1949 (sc))
9	sky	langi	*langit (PAND, Dempwolff 1983)
10	hit	pala	*palu (PANB, BLUST 1992)
11	open	buka	*buka' (PAND, Dempwolff 1983)
12	die	mate	*matey (PANDTV, DYEN 1949 (SC))
13	hot	pana	*panas (PANC, CAPELL 1943)
14	thin	tipi	*tipis (PANS, Stresemann 1927)
15	three	ntelu	*telu (PANDYMT, DYEN 1946)
16	right	wana	*wanaN (PANL, LEE 1972)
17	what	ape	*'apa' (PAND, DEMPWOLFF 1938)

Note. Historical source attributions are reported indirectly through Wurm and Wilson (1975), the compilation identified in the study. The original historical works are not presented as directly consulted sources. Attributions dated after 1975 require separate source verification, as indicated in the Word comments.

Table 3 presents examples of cognate words that have undergone only one phonological change from PAn to the Sanggar language. The changes include simple processes such as vowel change, consonant change, or sound deletion. Even though each word has experienced one phonological change, the relationship between the Sanggar forms and their PAn etymons can still be clearly identified. This suggests that the inherited words have been preserved while undergoing only minor phonological modifications.

B. Retention with Two Phonological Changes

There are 18 words that involve two phonological processes in inheritance, such as the combination of sound deletion and sound addition.

Table 4. Example of Retention with Two Phonological Changes

No.	English	Sanggar form	Reconstruction and source attribution
1	person	ata	*taw (PANC, Capell 1943)
2	male	mone	*manay (PANC, Capell 1943)
3	who	sei	*stai (PANC, Capell 1943)
4	you	komi	*kamiw (PANC, Capell 1943)
5	tooth	ngi	*ngipen (PANC, Capell 1943)
6	house	huma	*yumah (PAND, Dempwolff 1983)
7	oil	ola	*lana (PANDLRD, DEMPWOLFF 1925)

No.	English	Sanggar form	Reconstruction and source attribution
8	moon	wula	*bulan (PAND, Dempwolff 1983)
9	come	maet	*mai (PANBLAA, BLUST 1973)
10	laugh	male	*pali (PANS, Stresemann 1927)
11	stand	ntendi	*tendey (PANDLRD, DEMPWOLFF 1925)
12	buy	pembeli	*belih (PANDLO, LOPEZ n.d.)
13	small	deke	*dikih (PAND, Dempwolff 1983)
14	long	naru	*labu (PANDLRD, DEMPWOLFF 1925)
15	thick	teba	*tebel (PANDLO, LOPEZ n.d.)
16	new	wau	*bayu (PAND, DEMPWOLFF 1938)
17	two	ndua	*duva (PANDN, DEMPWOLFF 1922)
18	white	bara	*burak (PANBLWO, BLUST 1971)

Note. Historical source attributions are reported indirectly through Wurm and Wilson (1975), the compilation identified in the study. The original historical works are not presented as directly consulted sources. Attributions dated after 1975 require separate source verification, as indicated in the Word comments.

C. Retention with 3 or More Phonological Changes

There are 11 words that have undergone 3 or more phonological changes in their inheritance. These words typically show several changes at once, such as sound deletion, vowel insertion, sound insertion, and consonant changes.

Table 5. Examples of Retention with Three or More Phonological Changes

No.	English	Sanggar form	Reconstruction and source attribution
1	skin	huri	*kulit (PAND, Dempwolff 1983)
2	fruit	wua	*buwah (PANC, Capell 1943)
3	eat	ngaka/ ngeke	*ka’ (PAND, Dempwolff 1983)
4	drink	ngino	*inum (PANDYPMPL, Dyen 1953a (sc))
5	dust	kalubu	*buk (PANDN, DEMPWOLFF 1922)
6	ash	rawu	*habuh (PANDLO, LOPEZ n.d.)
7	mist	apu	*kabut (PAND, Dempwolff 1983)
8	turn	peko	*pikul (PANBLAA, BLUST 1973)
9	one	ica	*hisah (PANDLO, LOPEZ n.d.)
10	four	mpapa	*pat (PANDF, DEMPWOLFF, 1929)

No.	English	Sanggar form	Reconstruction and source attribution
11	day	ai	*waRih (PANL, LEE 1972)

Note. Historical source attributions are reported indirectly through Wurm and Wilson (1975), the compilation identified in the study. The original historical works are not presented as directly consulted sources. Attributions dated after 1975 require separate source verification, as indicated in the Word comments.

The Most Dominant Pattern of Phonological Retention and Change

Based on the analysis of 200 basic vocabulary items from the Swadesh list, the Sanggar language was found to have 50 cognate words from Proto-Austronesian (PAn). Among these cognate words, the most dominant pattern was Retention with Phonological Changes, which occurred in 46 words. Meanwhile, 4 words showed Full Retention without Phonological Changes. These findings indicate that the Sanggar language has retained some forms from Proto-Austronesian (PAn), but most of the inherited vocabulary has undergone various phonological changes, such as sound deletion, sound addition, vowel changes, and consonant weakening. The distribution of the complexity of phonological changes in the category of retention with phonological changes is presented in Table 6 below.

Table 6. Distribution of the Complexity of Phonological Changes in the Retention with Change Category

Type of Change	Word Count	Percentage (%)
1 Change	17 words	37.0%
2 Changes	18 words	39.1%
3 or More Changes	11 words	23.9%
Total	46 words	100%

Discussion

The findings show that the lexical relationship between the Sanggar language and the reconstructed PAn forms examined in this study was 25.0%, based on 50 cognate words out of 200 Swadesh items. This result provides a quantitative description of the lexical relationship within the selected vocabulary list and reconstruction reference. However, the percentage should not be interpreted as a complete measurement of the historical relationship between Sanggar and all forms of Proto-Austronesian because the analysis was limited to the selected 200 vocabulary items and the PAn etyma documented by Wurm and Wilson (1975)

The phonological analysis revealed two patterns of inheritance: full retention and retention with phonological changes. Full retention was found in four words, namely ama, ina, asu, and tua. These forms showed no identifiable sound changes in the comparison conducted. In contrast, the other 46 cognate words showed phonological differences while retaining a cognate relationship with the selected PAn forms. This distinction illustrates that inherited vocabulary may preserve its original phonological

form or undergo changes during language development.

The changes identified in the data included sound deletion, sound insertion, vowel changes, consonant changes, lenition, and metathesis. For example, langit > langi illustrates the deletion of a final consonant, while bulu > wulu illustrates lenition from /b/ to /w/. The comparison atay > ate illustrates the simplification of the final diphthong into a single vowel. These examples show that phonological inheritance can involve different types of changes rather than only the complete preservation of an ancestral form.

These findings are consistent with the historical-comparative perspective that inherited vocabulary and regular sound correspondences can be examined through systematic comparisons of linguistic forms (Campbell, 2013; Crowley & Bowern, 2010). The analysis of the Sanggar data therefore considers both lexical similarity and phonological differences when identifying cognate forms. This is important because surface differences alone do not necessarily rule out a historical relationship between two forms.

The distribution of phonological changes further shows that the identified cognate forms do not all undergo changes of the same degree. Eighteen words showed two changes, 17 showed one change, and 11 showed three or more changes. This variation provides a more detailed description of how the selected Sanggar forms correspond to the reconstructed PAn etyma. It also supports the use of qualitative historical-comparative analysis after the quantitative identification of cognates.

Overall, the findings indicate that the Sanggar language retains some vocabulary comparable with the selected PAn etyma while also showing phonological changes. The results contribute to the documentation of Sanggar vocabulary and its phonological patterns. Further research using additional vocabulary and comparative references may provide a broader understanding of the historical development of the Sanggar language.

CONCLUSIONS

This study found a 25.0% lexical relationship between Proto-Austronesian (PAn) and the Sanggar language, based on 50 cognate words out of 200 basic vocabulary items. The findings revealed two patterns of phonological inheritance: full retention and retention with phonological changes. Four words showed full retention, while 46 cognate words showed phonological changes. The identified changes included sound deletion, sound insertion, vowel changes, consonant changes, lenition, and metathesis. These findings indicate that the Sanggar language shares cognate vocabulary with the selected PAn reconstructions, although most of the identified cognate words have undergone phonological changes. Further research using a larger vocabulary list and additional reconstruction sources is recommended to provide a broader understanding of the historical relationship between PAn and the Sanggar language.

Another limitation concerns the uncertain etymological status of some lexical comparisons and the fact that the identified sound correspondences have not yet been established as systematic patterns across a sufficiently broad comparative dataset. Therefore, future research should extend the comparative analysis by examining Sanggar alongside neighboring languages, particularly Bima, Sambori, and Kolo, to establish more reliable regional sound correspondences and evaluate the historical relationships among these languages. Researchers should also consult additional Proto-Austronesian reconstruction sources to verify questionable cognates, consider alternative etymologies, and refine the identification of inherited forms. Investigating language contact and contact-related vocabulary would help clarify the distinction between inherited forms, innovations, and possible borrowings. These directions are expected to strengthen the historical-comparative evidence and contribute to a more comprehensive and reliable documentation of the Sanggar language.

REFERENCES

Anggayana, I. W. A., Suparwa, I. N., Dhanawaty, N. M., & Budasi, I. G. (2020). Lipang, Langkuru, Waisika language kinship: Lexicostatistics study in Alor Island. *International Journal of Psychosocial Rehabilitation, 24*(4), 301–319. https://doi.org/10.37200/IJPR/V24I4/PR201010

Anttila, R. (1972). *An introduction to historical and comparative linguistics*. Macmillan.

Blust, R., & Trussel, S. (2013). The Austronesian comparative dictionary: A work in progress. *Oceanic Linguistics, 52*(2), 493–523. https://doi.org/10.1353/ol.2013.0016

Budasi, I. G. (2012). Phonological evidences which separate and unite Mamboro language from Proto Wanokaka-Anakalang in Sumba group of languages. *k@ta, 14*(2), 67–74. https://doi.org/10.9744/kata.14.2.67-74

Budasi, I. G. (2018). Kekerabatan secara ekolinguistik bahasa Bima di NTB dan bahasa Sabu di NTT. *Deskripsi Bahasa, 1*(2), 130–137. https://doi.org/10.22146/db.v1i2.330

Campbell, L. (2013). *Historical linguistics: An introduction* (3rd ed.). Edinburgh University Press.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. SAGE Publications.

Crowley, T., & Bowern, C. (2010). *An introduction to historical linguistics* (4th ed.). Oxford University Press.

Katamba, F. (1989). *An introduction to phonology*. Longman.

Mbangi, L. I., & Marafad, L. O. S. (2018). The lexicostatistic study of Culambacu language with Tolaki language. *Cakrawala Linguista, 1*(1), 28–33. https://doi.org/10.26737/cling.v1i1.496

Millar, R. M., & Trask, R. L. (2015). *Trask's historical linguistics* (3rd ed.). Routledge.

Salahuddin, S. (2023). Lexicostatistics calculation on Manggarai and Bima languages: A comparative historical linguistics study. *PAROLE: Journal of*

Linguistics and Education, 13(2), 33–48. https://doi.org/10.14710/parole.v13i2.33-48

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.

Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics, 21*(2), 121–137. https://doi.org/10.1086/464321

Wurm, S. A., & Wilson, B. (1975). *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)* (Pacific Linguistics, Series C, No. 33). Department of Linguistics, Research School of Pacific Studies, Australian National University.