

The Genetic Relationship Between Balinese and Madurese

Ni Putu Parmini ¹

¹ IKIP Saraswati, Indonesia

I Wayan Mawa ²

² IKIP Saraswati, Indonesia

I Wayan Soper ³

³ IKIP Saraswati, Indonesia

I Made Suparta ⁴

⁴ IKIP Saraswati, Indonesia

Ni Made Sueni ⁵

⁵ IKIP Saraswati, Indonesia

I Gede Bagus Wisnu Bayu Temaja ⁶

⁶ IKIP Saraswati, Indonesia

Corresponding Author: parminiputu2019@gmail.com

Abstract

This study aimed to identify the genetic relationship between Balinese and Madurese through historical linguistics study. The results show that the percentage of genetic relationships in both languages is 47% and was presumed to diverge around 278 - 443 AD (as in 2018). Qualitatively, there are fifteen sets of sound correspondences, i.e., / a - e /, / i - e /, / ə - a /, / a - ə /, / ə - e /, / u - o /, / Ø - ʔ /, / k - ʔ /, / h - r /, / t - ʔ /, / h - Ø /, / b - bh /, / g - gh /, / d - dh /, and / j - jh /. Balinese maintained 23 proto lexicons in Proto-Austronesian lexicons (12%), a retention more than Madurese's five lexicons (2.5%).

Keywords: Genetic Relationship; Lexicostatistics; Glottochronology; Sound Correspondences.

INTRODUCTION

Balinese and Madurese are two geographically adjacent languages. Balinese is spoken in Bali, Indonesia (Temaja, 2018a; Temaja, 2018b; Temaja, 2018c), while Madurese is spoken on the

139



island of Madura and the eastern part of East Java Province, Indonesia. The proximity of the two languages geographically certainly creates relatedness that leads to linguistic similarities (Georgi, Xia, & Lewis, 2010), especially in terms of phonological and morphological aspects. In addition, historically, both languages belong to the same proto, Proto-Austronesian (PAN) family. As members of the same language family, the genetic relatedness of both languages can be confirmed (Parera, 1991: 107).

The proximity of the two languages can be seen from the field of historical linguistics. Some lexicons that cause both languages 'related' can be seen from the same word pair of 'ash' (Balinese *abu*; Madurese *abu*) and a similar word like 'smooth' (Balinese *alus*; Madurese *alos*). The same and identical word pairs derived from a common ancestor (Aridawati et al., 2023; Blust, 2013: 687) or proto are known as cognate or word of genetic relationship (the term 'relationship' will be used throughout the article to refer to 'genetic relationship'). The findings were a prerequisite in comparing the language in this case to find its relationship (Keraf, 1996: 34). However, the level of relationship between the two languages cannot be ascertained in detail yet. Therefore, the relationship's status needs to be examined further. The discussions of the relationships between two or more languages can be studied through the historical linguistics field using qualitative and quantitative approaches (Hendrokumoro & Temaja, 2018). The relationship that was summed up at the beginning, the relationship on the level of stock, resulted from quantitative calculation through lexicostatistics. A qualitative approach also needs to be implemented to strengthen the status of the language relationship between Balinese and Madurese. Such relationship qualitatively can be examined by identifying sound correspondence sets from cognates (Blust, 2013: 690). In addition, since both languages came from the same proto, the relationship of both languages can be traced to it (Turchin, Peiros, & Gell-Mann, 2010), as in identifying PAN lexicons that remain in both languages as well as to find which languages retain the most proto-forms. Therefore, the search for answers to this study can be traced by identifying 1) lexicostatistics and glottochronology, 2) sound correspondence sets, and 3) PAN's lexicon retention.

Related research has already been conducted by Mbete (1990) and Adelaar (2005). Mbete (1990) stated that Balinese and Madurese have a cognate percentage of 41%. Adelaar (2005), in his article on the subgrouping of Malayo-Sumbawan languages, grouped Balinese and Madurese as two members. His study underlies the results of Mbete's (1990) study on Bali-Sasak-Sumbawa proto-languages. Reflecting on both studies, they only discovered the relationship between both languages using lexicostatistics. Several aspects have not received attention, e.g.,



glottochronology (quantitative), sound correspondence, and retention of the PAN lexicon (qualitative). Furthermore, reexamining the lexicostatistics results of both studies is paramount.

Lexicostatistics was applied by finding out the cognates of compared languages to obtain the cognate percentage statistically (Keraf, 1996; Crowley & Bowern, 2010). Percentage results can be used further to identify compared languages' time-depth by applying glottochronology (Keraf, 1996; Parera, 1991). Both techniques were used to seek the relationship quantitatively. Then, a qualitative search for a relationship was implemented by finding sound correspondence sets based on cognates (Keraf, 1996). This correspondence can be established if phonemic recurrence exists or there is more than one pair of cognate's appearance per correspondence set, as well as if the rules or environment of correspondences were formulated. Later, the retention of a lexicon of a language can be derived from its lexicon maintenance using innovative methods or renewal of the compared lexicon of its proto lexicon. Innovation is applied to find out the form of relic and innovation. Relic represents a way to find the retention of all protoforms and their semantic meaning that still survives in the present (Keraf, 1996: 115-118). In contrast, innovation identifies the essential vocabulary development from its proto form phonologically, morphemically, and semantically (Keraf 1996: 115-118). Loan words do not influence innovation because the language per se developed naturally.

METHOD

This study was conducted in three stages: data collection, analysis, and display (Temaja, 2017; Temaja, 2018d; Temaja, 2021). At the data collection stage, the researchers applied the interview method through techniques of note-taking and recording (Sudaryanto, 2015) and literature study. The instrument for data collection was 200 Swadesh's basic vocabulary obtained from informants of both Balinese and Madurese. To support the data collection, researchers used additional references from the two languages, e.g., the Balinese dictionary from Denes (1997; 1998) and the multilingual dictionary by Sugiarto et al. (2001), which Madurese is one of them.

This study was conducted in three stages: data collection, analysis, and display. At the data collection stage, researchers applied the interview method through note-taking and recording techniques (Sudaryanto, 2015) and literature study. The instrument for data collection was 200 Swadesh's basic vocabulary obtained from informants of both Balinese and Madurese. To support the data collection, researchers used additional references from the two languages, e.g., the Balinese dictionary from Denes (1997; 1998) and the multilingual dictionary by Sugiarto et al.



(2001), which Madurese is one of them.

Quantitative data analysis was carried out by applying lexicostatistics and glottochronology techniques from Keraf (1996: 126-136) through the following steps: 1) removing unaccounted vocabulary; 2) isolating bound morphemes; 3) determining the cognates; 4) calculating cognates' percentage using lexicostatistics formula; 5) calculating time-depth using glottochronology formula; 6) calculating time-depth error; 7) calculating a new time-depth; 8) subgrouping languages relationship. The cognate percentage and diverging language time can determine the subgrouping, as shown in Table 1 below.

Table 1. level of language subgrouping

Language Level	Time-depth (in Century)	Cognate Percentage (%)
Language	0 - 5	100 - 81
Family	5 - 25	81 - 36
Stock	25 - 50	36 - 12
Microphylum	50 - 75	12 - 4
Mesophylum	75 - 100	4 - 1
Macrophylum	100 above	less than 1

Qualitative data analysis was conducted by applying sound correspondence and innovation methods. Sound correspondence sets were obtained by searching the correspondence of sounds that appear in more-or-less similar cognate pairs. For example, the word 'smooth' in both languages is a couple of cognates (Balinese *alus*; Madurese *alos*). From two the only difference between them is the ultimate vowel sounds of /u/ and /o/. Both cognates can be stated that high back vowel /u/ *alus* corresponds to mid back vowel /o/ *alos*; therefore, it can be formulated rules as follows: / u - o / C- (read: in gloss 'smooth', Balinese vowel /u/ is equivalent to Madurese vowel /o/ which appears in the nucleus of the second syllable preceded by dental approximant consonant /l/).

In the innovation method, Dempwolff's dictionary of PAN vocabulary (1938) was used to identify which languages still retain more proto lexicon based on data from the 200 vocabularies of both languages. In order to examine language retention, the innovation method was applied in identifying forms of relic and innovation. Sometimes, relic also occurs in innovation forms as an idiom, but the semantic proto still survives entirely, followed by additional elements. For example,



in Old English, proto *shep is still preserved as a 'shepherd' while 'sheep' is an innovation. The lexicon innovations later formed from proto can be seen, e.g., PAN *hulu 'head', in other languages developed its semantic innovation into *hulu* 'upstream' (Indonesian), *duur* 'head' (Balinese), and others.

The analysis results were presented in formal and informal manners and a combination of both (Sudaryanto, 2015). In formal display, researchers presented the analysis with tables and formulas of rules, especially on formulating the rules of sound correspondence sets. Meanwhile, the informal display was applied by formulating the results through the word-by-word description.

RESULTS AND DISCUSSION

Percentage of Genetic Relationship and Time-Depth

By removing unaccounted vocabulary and isolating bound morphemes, 94 cognates of 199 (one gloss not found, 'some' Balinese) were found. Those cognates were calculated then to obtain the percentage of cognate (C) by using the following lexicostatistics formula.

$$C = \frac{94}{199} \times 100\% = 47\%$$

The calculation shows that the relationship between both languages equals 47% or 0.47. The two languages show genetic relationship levels at the language 'family' based on the percentage range (36 - 81%) as in Table 1.

Time-depth

The cognate percentage was then used to calculate the time-depth (W) of the two languages as in the following glottochronology formula.

$$W = \frac{\log. 0.47}{2 \times \log. 0.805} = 1.740 \times 1,000 = 1,740$$

The calculation results show that both languages split 1,740 years ago. Since time-depth is a hypothetical result, it is impossible to determine the time of languages separation only at one exact time; therefore, it is necessary to analyze the range of time-depth to estimate a more reliable separation of the proto (Hanna, Mursidin, & Jannah, 2018). Subsequently, a further calculation is needed, which in glottochronology can be identified by determining the error range.

Range of error of the time-depth

By taking the calculation of the cognate percentage, 0.47, then it was implemented to



determine the range of error (S).

$$S = \frac{\sqrt{0.47(1-0.47)}}{200} = 0.035$$

Thus, the range of error was 0.035. This score was then used to determine the new cognate percentage (C1).

$$C1 = 0.47 + 0.035 = 0.505$$

The calculation obtained a new cognate percentage between the two languages equal 0.505 or 50.5%.

New time-depth

The new cognate percentage was used to examine a new time-depth of both languages (W1) by implementing the lexicostatistics formula again.

$$W1 = \frac{\log 0.505}{2 \times \log 0.805} = 1.575 \times 1,000 = 1,575$$

The calculation denoted a new time-depth of both languages was 1,575 years ago. From it, a possible range of time-depth of both languages was 1,740 - 1,575 years ago or a year of 278 - 443 AD as in the current year (2018). This time-depth period indicates that Balinese and Madurese belong to the language 'family' subgrouping according to the time-depth range of 5 - 25 (in century) as in Table 1.

Sound Correspondence

Based on the analysis, fifteen (15) sound correspondence sets had recurrences or the appearance of more than a pair of cognates. The sets are: / a - e /, / i - e /, / ə - a /, / a - ə /, / ə - e /, / u - o /, / ø - ʔ /, / k - ʔ /, / h - r /, / t - ʔ /, / h - ø /, / b - b^h /, / g - g^h /, / d - d^h /, and / j - j^h /. The description of distribution and rules of each correspondence set is described as follows (abbreviations used are consonant (C) and vowel (V)).

For the beginning, sound correspondence / a - e / in Table 2 occurs in the initial syllable position of the cognate pairs. Such occurrence can be seen in cognates of *barak* : *mera* 'red' and *sa* : *settoŋ* 'one'. The vowel is preceded by consonants /b/ and /s/, except the data of Balinese *ake* 'I', which is not preceded by any sound, which considered as an empty sound or zero (ø). However, if it is related to PAN *aku > *senkoʔ* in Madurese, it resulted from a prothesis or initial sound addition, dental fricative consonant /s/.



Table 2. sound correspondence / a - e /

Gloss	Balinese	Madurese	Rule
I	ake	seŋkoʔ	/ a - e / C-
Red	barak	Mera	/ a - e / C-
One	sa	Settoŋ	/ a - e / C-

Table 3 shows the sound correspondence set of / i - e / occurs in the syllable nucleus preceded by consonant as the onset. There is an exception to the rule, i.e., 1) cognate of *ai* : *are* 'sun', this correspondence occurs in the ultimate position, and 2) cognate *inəm* : *enom* 'to drink', their correspondence occurs in the initial position. However, on the one hand, the first exception might be explainable since syncope, or middle sound loss of glide, occurred in Balinese, **vayi* > *ai*. On the other hand, gloss for 'to drink' is truly an initial open syllable, even for PAN **inum*.

Table 3. sound correspondence / i - e /

Gloss	Balinese	Madurese	Rule
Wind	aŋin	aŋen	/ i - e / C-
To split	sibak	ŋebəʔ	/ i - e / C-
To kill	mati	mateʔe	/ i - e / C-
Worm	caciŋ	caceŋ	/ i - e / C-
Liver	ati	ate	/ i - e / C-
Small	cənik	keneʔ	/ i - e / C-
Skin	kulit	koleʔ	/ i - e / C-
Yellow	kuniŋ	koneŋ	/ i - e / C-
Five	limə	lemaʔ	/ i - e / C-
To die	mati	mate	/ i - e / C-
White	putih	pote	/ i - e / C-
Rope	tali	tale	/ i - e / C-
Ear	kupiŋ	kopeŋ	/ i - e / C-
Thin	tipis	tepes	/ i - e / C-
Sun	ai	are	/ i - e / -#
To drink	inəm	enom	/ i - e / #-



The sound correspondence set of / ə - a / occurs at the position of the syllable nucleus preceded by a consonant. It is an exciting finding *vis-à-vis* correspondence set of / a - e /, which somehow kind of its reversal. Mid central vowel or *schwa* /ə/ at the ultimate position in Balinese is being reflected by low central vowel /a/ at the penultimate position in Madurese, such as *apə* : *apa* 'what', *bapə* : *appa?* 'father', or *dijə* : *ed^himma* 'where'. However, take into account that quite a different phenomenon appears at the cognate of *kəbər* : *ḡabb^hər* 'to fly', which is positioned at the penultimate. Such exception may cause by cognates ended by labial trill consonant /r/. Consider the statement is hypothetical; hence counter data are needed to support such a claim. The following data in Table 4 may ease up the explanation.

Table 4. sound correspondence / ə - a /

Gloss	Balinese	Madurese	Rule
What	apə	apa	/ ə - a / C-
Father	bapə	appa?	/ ə - a / C-
Where	dijə	ed ^h imma	/ ə - a / C-
Hot	panəs	panas	/ ə - a / C-
Five	limə	lema?	/ ə - a / C-
Eye	matə	mata?	/ ə - a / C-
To fly	kəbər	ḡabb ^h ər	/ ə - a / C-
Old	tuwə	towa	/ ə - a / C-

Similar to the first correspondence set of / a - e /, it occurs in the syllable nucleus preceded by a consonant, but, in contrast, the vowel /a/ in Balinese corresponds to schwa /ə/. However, general rules regarding occurrence position may not be drawn quickly since the data show sporadic distribution as in *kambəḡ* : *kəmb^həḡ* 'to float', *sibək* : *ḡebə?* 'to split', or *buwəḡ* : *buwə* 'fruit' occurs at the ultimate position. Contrast environment appears at the penultimate position such as *batu* : *bəto* 'stone', *jait* : *j^həi?* 'to sew', *taluh* : *təllor* 'egg', and others. Furthermore, both position, ultimate and penultimate, also occur in such correspondence, as in *bahət* : *bərrə?* 'heavy' and *jalan* : *j^hələn* 'to walk'.

Apart from the previous generalization, there are exceptions to *appa?* 'father' in Madurese, where the vowel /ə/ is not preceded by a consonant. It makes vowel /ə/ occur at the penultimate position followed by the labial stop consonant /p/. It is a result of prothesis in



Balinese *ama > *bapə*. Table 5 provides the whole data that represents the correspondence.

Table 5. sound correspondence / a - ə /

Gloss	Balinese	Madurese	Rule
To float	kamban	kəmb ^h ən	/ a - ə / C-
Stone	batu	bəto	/ a - ə / C-
To split	sibak	ɲebəʔ	/ a - ə / C-
Heavy	bahat	bərrəʔ	/ a - ə / C-
To give	bahan	bərriʔ	/ a - ə / C-
Fruit	buwah	buwə	/ a - ə / C-
Moon	bulan	bulən	/ a - ə / C-
And	lan	bən	/ a - ə / C-
Blood	rah	d ^h ərə	/ a - ə / C-
Salt	uyah	bujə	/ a - ə / C-
Rain	ujan	oɟ ^h ən	/ a - ə / C-
To sew	Jait	ɟ ^h əiʔ	/ a - ə / C-
To walk	jalan	ɟ ^h ələn	/ a - ə / C-
Egg	taluh	təllor	/ a - ə / C-
Father	bapə	əppaʔ	/ a - ə / -C

In Table 6, the correspondence set of / ə - e / occurs in the syllable nucleus preceded by a consonant. Balinese vowel /ə/ equivalent to Madurese /e/ at the ultimate position *sələm* : *cellən* 'black' and *cupək* : *copeʔ* 'narrow', as well as the penultimate position, represents by *cənɪk* : *keneʔ* 'small'. Since penultimate recurrence is only represented by one data, it is presumed as the result of metathesis, *cənɪk* metathesized into *keneʔ*, where the first syllable *ce-* reversed its position to be *ke-* as well as at the same time palatal stop /c/ corresponds to velar stop /k/.

Table 6. sound correspondence / ə - e /

Gloss	Balinese	Madurese	Rule
Black	sələm	cellən	/ ə - e / C-
Small	cənɪk	keneʔ	/ ə - e / C-
Narrow	cupək	copeʔ	/ ə - e / C-



Correspondence set of / u - o / positioned at the syllable nucleus preceded by a consonant. Firstly, it appears at the ultimate position for both open and closed syllables, e.g., *batu* : *bəto* 'stone' or *gunuŋ* : *g^hunoŋ* 'mountain'. The penultimate position only covers open syllables, as in *kulit* : *koleʔ* 'skin', *kuniŋ* : *koneŋ* 'yellow', *putih* : *pote* 'white', and others. Double-correspondence occurs at both ultimate and penultimate positions, as in *kuku* : *koko* 'fingernail', *kutu* : *koto* 'louse', *susu* : *soso* 'breast', and others.

However, there is an exception toward cognates of *ujan* : *oʝ^hən* 'rain' and *utah* : *ŋota* 'to vomit' where their correspondences appear at the penultimate position followed by a consonant. Such exception is acceptable due to reflecting its PAN lexicons, *udan and *untah, that show quite similar forms. Table 7 below provides the entire data for the correspondence set.

Table 7. sound correspondence / u - o /

Gloss	Balinese	Madurese	Rule
Stone	batu	bəto	/ u - o / C-
Mountain	gunuŋ	g ^h unoŋ	/ u - o / C-
Smooth	alus	alos	/ u - o / C-
To count	ituŋ	bituŋ	/ u - o / C-
Fingernail	kuku	koko	/ u - o / C-
Skin	kulit	koleʔ	/ u - o / C-
Yellow	kuniŋ	koneŋ	/ u - o / C-
Louse	kutu	koto	/ u - o / C-
Breast	susu	soso	/ u - o / C-
White	putih	pote	/ u - o / C-
Narrow	cupək	copeʔ	/ u - o / C-
Ear	kupiŋ	kopeŋ	/ u - o / C-
Egg	taluh	təllor	/ u - o / C-
Three	təlu	təllor	/ u - o / C-
Old	tuwə	towa	/ u - o / C-
Bone	tulaŋ	tolaŋ	/ u - o / C-
Dull	puntul	tompol	/ u - o / C-
Rain	ujan	oʝ ^h ən	/ u - o / -C



To vomit	utah	ŋota	/ u - o / -C
----------	------	------	--------------

This set / $\emptyset$ - η / appears in the ultimate position, which is preceded by a vowel or can be declared its appearance in the final position. The correspondence still questions its status since Madurese glottal stop consonant / η / has no equivalent to Balinese. Balinese has no such consonant; therefore, the researchers argue it possesses zero phonemes represented by / $\emptyset$ /. As seen in Table 8, whole data that shows such claim, e.g., *ake* : *seŋko* 'I', *bapə* : *əppa* 'father', and others. Despite those, there are exceptions for rules based on the data for *mate* η e in Madurese, where consonant / η / is followed by mid-front vowel /e/.

Table 8. sound correspondence / $\emptyset$ - η /

Gloss	Balinese	Madurese	Rule
I	ake	seŋko η	/ $\emptyset$ - η / -#
Father	bapə	əppa η	/ $\emptyset$ - η / -#
Here	dini	ed ^h inna η	/ $\emptyset$ - η / -#
Two	duwə	d ^h uwə η	/ $\emptyset$ - η / -#
Five	limə	lema η	/ $\emptyset$ - η / -#
Eye	matə	mata η	/ $\emptyset$ - η / -#
Three	təlu	təllə η	/ $\emptyset$ - η / -#
To kill	mati	mate η e	/ $\emptyset$ - η / V-V

In Table 9, the correspondence set of / k - η / occurs in the ultimate position of the cognate pairs. A vowel sound precedes the consonants. This set can counter data for correspondence / $\emptyset$ - η / , which is the zero sound substituted by the velar stop consonant /k/ in Balinese. However, researchers consider that both were distinct sets supported by prominent data representing each.

Table 9. sound correspondence / k - η /

Gloss	Balinese	Madurese	Rule
Child	panak	ana η	/ k - η / -#
Many	bək	bəŋŋa η	/ k - η / -#
To split	sibak	ŋebə η	/ k - η / -#
To sit	təgak	toju η	/ k - η / -#
Small	cənik	kene η	/ k - η / -#



Hair	bok	obu?	/ k - ʔ / -#
Few	bədik	diddi?	/ k - ʔ / -#
Narrow	cupək	cope?	/ k - ʔ / -#
Horn	tanduk	tandu?	/ k - ʔ / -#
To laugh	kədek	gʰəllə?	/ k - ʔ / -#

The correspondence set of / h - r / in Balinese and Madurese supported van der Tuuk's RGH law of Austronesian languages. Table 10 shows that correspondence occurs in the ultimate and penultimate positions initiated by vowel sounds. Entire data are in the form of closed syllables for correspondence occurrence.

Table 10. sound correspondence / h - r /

Gloss	Balinese	Madurese	Rule
Correct	bənəh	bʰəndʰər	/ h - r / V-
Heavy	bahat	bərrə?	/ h - r / V-
To give	bahanj	bərri?	/ h - r / V-
Egg	taluh	təllor	/ h - r / V-

Apart from the two previous correspondence that involves Madurese glottal consonant /ʔ/, it also shows equivalents sound by Balinese consonant /t/ in forming a correspondence set of / t - ʔ /. Its environment is in the ultimate position of cognates preceded by a vowel. Thus, Madurese /ʔ/ is equivalent to two Balinese consonants /k/ and /t/, as well as one zero sound. Despite Madurese having 26 consonants and 9 vowels *vis-à-vis* Balinese's 18 consonants and 6 vowels, however, such surprising equivalence on certain sounds may not be influenced by several phonemes that a language has. The distribution of this correspondence is displayed in Table 11.

Table 11. sound correspondence / t - ʔ /

Gloss	Balinese	Madurese	Rule
Heavy	bahat	bərrə?	/ t - ʔ / -#
Four	pat	əmpa?	/ t - ʔ / -#
To sew	jait	jʰəi?	/ t - ʔ / -#
Skin	kulit	kole?	/ t - ʔ / -#



Sky	lanjit	lanje?	/ t - ʔ / -#
-----	--------	--------	--------------

Similar to the previous correspondence, the corresponding set of / h - Ø / occurs in the ultimate position of the word initiated by vowel sound. The final loss sound phenomenon occurs in the Madurese lexicon if it is tracked from the origin in PAN, as in **banih* > *big^{hi}* 'seed', **tanah* > *tana* 'earth', etc.

Table 12. sound correspondence / h - Ø /

Gloss	Balinese	Madurese	Rule
Seed	binih	big ^{hi}	/ h - Ø / -#
Earth	tanah	tana	/ h - Ø / -#
Blood	rah	d ^h arə	/ h - Ø / -#
Salt	uyah	bujə	/ h - Ø / -#
Fall	labuh	Labu	/ h - Ø / -#
Far	joh	j ^h əu	/ h - Ø / -#
To vomit	utah	ŋota	/ h - Ø / -#
White	putih	Pote	/ h - Ø / -#

Table 13 below shows the sound correspondence set / b - b^h /. This correspondence appears in the syllable onset part and is followed by the vowel. Since Balinese has no aspirated consonants, therefore there are plenty of correspondence sets involving aspirated consonants.

Table 13. sound correspondence / b - b^h /

Gloss	Balinese	Madurese	Rule
To float	kambaŋ	kəmb ^h əŋ	/ b - b ^h / -V
Correct	bənəh	b ^h ənd ^h ər	/ b - b ^h / -V
To hunt	boros	b ^h uru	/ b - b ^h / -V
To fly	kəbər	ŋabb ^h ər	/ b - b ^h / -V

The sound correspondence set of / g - g^h / occurs on the syllable onset followed by a vowel. Only two pieces of data represent it, which can be seen in Table 14.



Table 14. sound correspondence / g - g^h /

Gloss	Balinese	Madurese	Rule
Tooth	gigi	g ^{hi} g ^{hi}	/ g - g ^h / -V
Mountain	gunuŋ	g ^h unonŋ	/ g - g ^h / -V

The correspondence set / d - d^h /, in Table 15, appears in the open syllable onset position followed by the vowel sound. Most of the data show that the correspondence appears at the initial consonant position.

Table 15. sound correspondence / d - d^h /

Gloss	Balinese	Madurese	Rule
Leaf	don	d ^h əun	/ d - d ^h / -V
Where	dijə	ed ^h imma	/ d - d ^h / -V
Here	dini	ed ^h innaʔ	/ d - d ^h / -V
There	ditu	ed ^h issa	/ d - d ^h / -V
Two	duwə	d ^h uwəʔ	/ d - d ^h / -V

Table 16 presents the corresponding sound set of / j - j^h /, which appears at the onset of a syllable initiated by a vowel sound as its nucleus. Similar to previous correspondence, it also occurs at the initial consonant position.

Table 16. sound correspondence / j - j^h /

Gloss	Balinese	Madurese	Rule
Rain	ujan	o ^j hən	/ j - j ^h / -V
To sew	Jait	j ^h əiʔ	/ j - j ^h / -V
To walk	Jalan	j ^h ələn	/ j - j ^h / -V
Far	Joh	j ^h əu	/ j - j ^h / -V

PAN's Lexicon Retention

PAN's lexicon retention was obtained by applying the innovation method by determining relic and innovation forms within 200 vocabularies of both languages. However, three PAN vocabularies, i.e., 'because', 'to throw', and 'to breath' were hard to find; hence only 197 data



were analyzed. The results show that Balinese has PAN's lexicon retention of 23 vocabularies (12%) of 196 (gloss for 'some' is unavailable). Meanwhile, Madurese maintains 5 vocabularies (2.5%) out of 197. It can be drawn that Balinese maintains its proto lexicon more than Madurese. Following 23 Balinese lexicons that reflect PAN's, i.e., *abu > *abu* 'ash', *aŋin > *aŋin* 'wind', *kamban > *kamban* 'to float', *batu > *batu* 'stone', *bintaŋ > *bintaŋ* 'star', *bulan > *bulan* 'moon', *bulu > *bulu* 'feather', *tanah > *tanah* 'earth', *di > *di* 'at', *diŋin > *diŋin* 'cold', *gigi > *gigi* 'tooth', *gunuŋ > *gunuŋ* 'mountain', *daki > *daki* 'dirty', *kuku > *kuku* 'fingernail', *kulit > *kulit* 'skin', *kuniŋ > *kuniŋ* 'yellow', *kutu > *kutu* 'louse', *laŋit > *laŋit* 'sky', *putih > *putih* 'white', *tali > *tali* 'rope', *tanduk > *tanduk* 'horn', *təbəl > *təbəl* 'thick', and *təlu > *təlu* 'three'. Meanwhile, Madurese retention of PAN's lexicon, i.e., *abu > *abu* 'ash', *apa > *apa* 'what', *bintaŋ > *bintaŋ* 'star', *bulu > *bulu* 'feather', and *maləm > *maləm* 'night'. Thus, it suggests that Madurese considerably evolved lexically since their separation from its proto rather than Balinese. Such involvement is presumed due to the 'richness' of phonemes that exist in Madurese. However, such a claim may still open a question of whether the phenomenon might begin with the evolution of phonological or lexical aspects.

CONCLUSION

This study attempted to provide quantitative and qualitative evidence to establish Balinese and Madurese relationships and support the previous research, especially in lexicostatistics results (Mbete, 1991; Adelaar, 2005). Despite the lexicostatistics results having been calculated, both languages' time-depth dramatically needs to be examined. Historically, they diverged from the common proto, Austronesia. Thus, these results may provide the answer. Several cognates show regular sound correspondence that might be recognizable by both language speakers while making contact. By recognizing it, intelligibility may increase among the speakers. Furthermore, the inheritance of some proto lexicons toward both languages greatly supports their membership status in Austronesian languages.

REFERENCES

- Adelaar, A. (2005). Malayo-Sumbawan. *Oceanic Linguistics*, 44(2), 357-388.
- Aridawati, I. A. P., Temaja, I. G. B. W. B., Putra, I. B. R., Tama, I. W., Sudiana, I. M., Sudiarta, I. W., & Armini, I. G. A. (2023). The Genetic Relationship Between Balinese and Balinese Malay: A Lexicostatistics and Phonemic Correspondence Study. *International Journal of Education*,



- Vocational and Social Science*, 2(1), 283-295. <https://doi.org/10.99075/ijevss.v2i01.169>
- Blust, R. (2013). *The Austronesian Languages (Revised Edition)*. Canberra: Asia-Pacific Linguistics.
- Crowley, T., & Bower, C. (2010). *An Introduction to Historical Linguistics (Fourth Edition)*. New York (NY): Oxford University Press.
- Dempwolff, O. (1938). *Vergleichhen de Lautlehre des Austronesischen Wortschatzes*. Hamburg: Friedderichsen, de Gruyter.
- Denes, I. M. (1997). *Kamus bahasa Indonesia-Bali A-K*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Denes, I. M. (1998). *Kamus Bahasa Indonesia-Bali L-Z*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Georgi, R., Xia, F., & Lewis, W. (2010). Comparing Language Similarity Across Genetic and Typologically-Based Groupings. In Coling (Eds.), *23rd International Conference on Computational Linguistics 2010, Beijing, China: Proceedings* (pp. 385-393).
- Hanna, M., & Jannah, S. N. R. (2018). Language Kinship Between Tulambatu and Tolaki Language in North Konawe. *Mediterranean Journal of Social Sciences*, 9(1), 91-102.
- Hendrokumoro, & Temaja, I. G. B. W. B. (2019). The Genetic Relationship Between Ma'anyan and Malagasy (Hubungan Kekerabatan Bahasa Ma'anyan dan Bahasa Malagasi). *Metalingua*, 18(2), 113-124. <https://doi.org/10.26499/metalingua.v17i2.312>
- Keraf, G. (1996). *Linguistik Bandingan Historis*. Jakarta: Gramedia Pustaka Utama.
- Mbete, A. M. (1990). *Rekonstruksi Protobahasa Bali-Sasak-Sumbawa*. Dissertation for Fakultas Pascasarjana Universitas Indonesia, Jakarta.
- Parera, J. D. (1991). *Kajian Linguistik Umum Historis Komparatif dan Tipologi Struktural*. Jakarta: Erlangga.
- Sudaryanto. (2015). *Metode dan Aneka Teknik Analisis Bahasa: Pengantar Penelitian Wahana Kebudayaan Secara Linguistik*. Yogyakarta: Sanata Dharma University Press.
- Sugiarto, et al. (2001). *Kamus Indonesia-Daerah: Indonesia, Jawa, Bali, Sunda, Madura*. Jakarta: Gramedia Pustaka Utama.
- Temaja, I. G. B. W. B. (2017). Sistem Penamaan Orang Bali. *Humanika*, 24(2), 60-72. <https://doi.org/10.14710/humanika.v24i2.17284>
- Temaja, I. G. B. W. B. (2018a). Cecimpedan: Teka-Teki dalam Bahasa Bali. Kongres Bahasa Indonesia XI. <http://repositori.kemdikbud.go.id/id/eprint/10138>
- Temaja, I. G. B. W. B. (2018b). Intransitive Verb Markers in Balinese. *Jurnal Arbitrer*, 5(2), 60-66.



<https://doi.org/10.25077/ar.5.2.60-66.2018>

Temaja, I. G. B. W. B. (2018c). Pola Reduplikasi Bahasa Bali: Perbandingannya dengan Pola Reduplikasi Bahasa-Bahasa Austronesia. *Prasasti: Journal of Linguistics*, 3(2), 190-199.

<https://doi.org/10.20961/prasasti.v3i2.17520>

Temaja, I. G. B. W. B. (2018d). Sapaan Kekerabatan dalam Bahasa Bali (Kinship Addresses in Balinese Language). *Metalingua*, 16(2), 211-220.

<https://doi.org/10.26499/metalingua.v16i2.238>

Temaja, I. G. B. W. B. (2021). Pentingnya Makna dan Nilai Budaya Konsep Kaja Kangin (Utara Timur) di Bali pada Anak. *Widya Kumara: Jurnal Pendidikan Anak Usia Dini*, 2(2), 189-198.

<https://doi.org/10.55115/widyakumara.v2i2.1566>

Turchin, P., Peiros, I., & Gell-Mann, M. (2010). Analyzing Genetic Connections Between Languages by Matching Consonant Classes. *Journal of Language Relationship*, 3, 117-126.

