www.bircu-journal.com/index.php/birci
Budapest International Research and Critics Institute-Journal (BIRCI-Journal)
Humanities and Social Sciences
ISSN 2615-3076 Online)
ISSN 2615-1715 (Print)
Budapest Institute

Lexicostatistics of Toba Language, Sibolga Language, and Minangkabau Language

Rezky Khoirina Tarihoran[1], Dwi Widayati[2]
[1]*Universitas Muslim Nusantara Al Washliyah / University of Sumatera*
[2]*University of Sumatera*
rezkykhoirina@umnaw.ac.id, *dwiwidayati@usu.ac.id*

Abstract

This study examines the Toba Batak language, the Sibolga Coastal language, and the Minangkabau language belonging to the Austronesian or Polynesian Malay family. The assumption regarding the kinship of the three languages is in the fact that there are similarities and similarities in form and meaning which are a reflection of the same historical heritage which is thought to have come from the mixing of natives with immigrants. This research uses the lexicostatistical method. The first stage, collects 200 basic vocabulary compiled by Morris Swades. The method used in providing this data is the referential method, while the technique used is the note-taking technique. Second, determine the pairs which of the three languages is a cognate language. Third, calculate the age and separation time of the three languages. Fourth, calculate the error term to determine a more precise separation time. From this research, it was found that the language status of the Sibolga language and the Minangkabau language is a family category, while the language status of the Minangkabau - Toba Batak and Sibolga - Toba Batak languages is a stock category.

Keywords
lexicostatistics; kinship; Toba batak language; Sibolga language; Minangkabau

I. Introduction

Language is one of the most important things in the life of every human being (Purba, N. et al. (2020). Language is a tool used by humans to communicate and interact. Over the course of human history, language has undergone many changes based on interactions made through exploration and trade. This ultimately affects the language in an area that has a relationship or resemblance to other languages. In its journey, the language of the Sibolga Coastal tribe, especially the Sumando Indigenous Language, is a language created from the assimilation of the Toba Batak language brought by the Toba Batak people from Silindung named Dorong and the Hutagalung surname who came to the coast in the 1700s to farm and settle there, and also the Minangkabau language brought by Minang immigrants to the coastal areas (Naim, 1979). This can be proven from the close kinship of the Toba Batak language and the Minangkabau language in the Sumando Indigenous Language.

In linguistic terms, language kinship is a relationship formed from the same source which includes two or more languages (KBBI, 2018). While related language is defined as a language that is genealogically related to other languages. Based on this, it can be concluded that language kinship is the relationship between languages between one another. This relationship may be formed from the same source so that each variety of language has similarities or shows the same characteristics. In terms of language, similarities can be shown in terms of morphology, syntax, and even phonology. In the linguistics dictionary, Kridalaksana (2008:116) explains that kinship (genetic relationship)

is a relationship between two or more languages that are derived from the same parent language source which is called an proto language.

II. Review of Literature

There are three articles that are used as references or literature reviews in this journal, the first is from Mifani Septriani (2013) with the title "Batak Toba Muslim: Study of Cultural Change in Coastal Communities in Sibolga (1970-2000)" which analyzes the history, formation of culture, religion, and language on the coast of Sibolga, especially the Simando Custom which is the result of assimilation from the Toba Batak and Minangkabau.

The paper "The Kinship of Javanese and Gayo Language: A Comparative Historical Linguistic Study" by Muhammad Surip and Dwi Widayati (2019), which was conducted using the lexicostatistical approach, is another work linked to linguistic kinship and the lexicostatistical method. According to Muhammad Surip and Dwi Widayati, just four of the 200 Swadesh terms examined lacked a partner. Comparatively, only 13 percent of the 196 terms that have a spouse are linked. The fact that the two languages share less than 36 percent of their core vocabulary demonstrates that BJ and BG are not closely related. However, it is still possible to demonstrate that the two languages share a common ancestor. There were 16 pairs of relatives with identical phonemes, 12 pairs with phonemically corresponding phonemes, 9 pairs with phonetically comparable phonemes, and 13 pairs with only one distinct phoneme out of a total of 25 relatives.

Another article that discusses linguistic kinship is "The Kinship of Karo Language Vocabulary, Nias Language, and Simalungun Language in Medan City: A Study of Comparative Historical Linguistics" by Sherly Novita and Dwi Widayati (2019), which was conducted using the comparative method and according to Sherly Novita and Dwi Widayanti, a comparison of the kinship level between Karo, Nias, and Simalung Language or genetic lineage. The level of kinship between CB and CB is 37%, indicating that the two languages belong to the same family or language family, while the level of kinship between BN and CB is 24%, indicating that the two languages belong to the same family or stock family.

Theoritical Review

The ideas utilized in this study have a connection to the theory of Comparative Historical Linguistics (CHL), wherein CHL studies on language are presented, one of which is to query related languages by undertaking a comparative examination of the elements that can demonstrate their kinship (Crowley, 1987; Keraf, 1991). Historical comparative theory, according to Keraf (1984:22), is the discipline of linguistics that examines language concerns in a certain time field as well as changes in language aspects that occur within that time period.

To compare two or more languages, lexicostatistical approaches are utilized, which allow us to assess the size of the relationship between these languages by analyzing the similarities in each word (Crowley: 1992: 168). Lexicostatistics is a language comparison method based on two fundamental assumptions. The first is that some sections of the basic vocabulary are more resistant to change than others, and the second is that the changes in each language's fundamental vocabulary are same (Crowley: 1987: 191-192).

Lexicostatistics is a method for classifying languages based on the proportion of cognates and relations (Mahsun, 1995:115). In lexicostatistical computations, cognates are words that are classified as related based on their phonetic or morphological similarity to

other languages being analyzed. According to Keraf (1984:121), lexicostatistics is the classification of languages that tend to prioritize the lexicon or statistical observation of words, and then apply the observation according to the proportion of similarities and differences between one language and another.

In lexicostatistics the naming of language subgroups is carried out as follows:

1. If the percentage of kinship in the core vocabulary is 81-100%, then the language is included in the sub-group “Language”;
2. If the percentage of kinship in the core vocabulary is 36-81%, then the language is included in the “family” sub-group;
3. If the percentage of kinship in the core vocabulary is 12-36%, then the language is included in the “Stock” sub-group;
4. If the percentage of kinship in the core vocabulary is at 2-13%, then the language is included in the "Microphylum" sub-group;
5. If the percentage of kinship in the core vocabulary is at 1-4%, then the language is included in the “Mesophyllum” sub-group;
6. If the percentage of kinship in the core vocabulary is 0-1%, then the language is included in the "Macrophylum" sub-group.

According to Keraf (1984:172), the formula used to calculate the degree of kinship between languages is as follows:

$$C = \frac{K}{G}\ x\ 100\%$$

Where
C= Cognates or related words
K = total vocabulary of relatives
G= amount of gloss

Following the reference from Crowley and Keraf, the following formula is used to calculate the separation time between the three languages:

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

Where
W = split time
Log = logarithm of log
r = percentage constant over 1000 years (retention)

III. Research Method

This study will utilize numbers to demonstrate the degree of resemblance and relationship between the investigated languages. Consequently, the method of comparison is accompanied by lexicostatistical calculations, which are then utilized to construct the matching feature components in Toba Batak language (BT), Coastal Sibolga language (BS), and Minangkabau language (BM) by employing a vocabulary list from each studied language that has been conformed to the reference from Morris Swadesh. The chosen vocabulary list already contains non-cultural words and has been evaluated for retention of basic words in languages that already have a written script. Considering the value of the chosen vocabulary, it is obvious that it will be beneficial to this research.

Several approaches were used to determine the relationship between each vocabulary in the BT, BS, and BM languages, namely, first, for vocabulary that was not considered, such as gloss with no words or empty words, followed by vocabulary that was borrowed and vocabulary complicated. The second step is to separate bound morphemes from the source word. In this instance, the bound morphemes in the chosen vocabulary will be split from the base word in order to make it easier to determine whether or not the words to be studied are same. The third criteria is that the relationship between the studied vocabulary pairs must satisfy one of the following: 1) all phonemes are identical or identically similar; 2) phonetically corresponds; 3) have the same articulation even in their respective positions, or are phonetically comparable; 4) there are diverse phonemes that are affected by their surroundings. After the preceding steps were completed, the kinship calculations for BT, BS, and BM were performed. This computation is based on a previously generated list of words.

IV. Results and Discussion

This investigation was conducted utilizing linguistic grouping and lexicostatistical approaches. The first level consists of 300 words prepared by Morris Swades. This information is provided using a referential manner, whereas the technique employed is a note-taking technique (Kesuma, 2007:48; Sudaryanto, 1993:13-16; Sudaryanto, 1988:5). Second, categorise the word kin (cognate) based on: (a) identical pairings, (b) phonemic correspondence pairs, (c) phonetically similar pairs, and (d) distinct pairs to obtain the proportion of kinship in the three languages. one phoneme. Calculate the age and separation time of the three languages, as well as the error range, to determine the separation time with greater precision. Fourth, the classification of kinship systems as a single language (language), language family (subfamily), language family (stock), microphylum, mesophyllum, or macrophylum (Keraf, 1984:126-128).

4.1 Lexicostatistics on Sibolga and Minangkabau Languages

After determining the word kin, the next step is to find the percentage of relatives with the formula:

$$C = \frac{Vt}{Vd} x100\%$$

Information:
C= relative word;
Vt= number of relatives' vocabulary;
Vd= amount of gloss calculated

$$C = \frac{Vt}{Vd} x100\% = \frac{146}{200} x100\% = 0.73x100\% = 73\%$$

After knowing the percentage of relatives, we can calculate the separation time from Sibolga and Minangkabau languages

Given: C = 73%
r = 80.5%
asked: W1 = …?

Answer:

$W_1 = \frac{Log\ C}{2 log 805} = \frac{Log\ 73}{2 log 805}$

$= \frac{315}{2x217} = \frac{315}{434} = 0.725$

The separation time is multiplied by 1000 so that the result becomes 725

So, the calculation of the initial separation time of the Sibolga language and the Minangkabau language is 725 ago.

In other words, the calculation of the initial separation time for the Sibolga language and the Minangkabau language can be stated as follows:

1) Sibolga and Minangkabau are thought to have formed a single language about 725 years ago.
2) Sibolga and Minangkabau languages are estimated to have started to separate from their mother tongues around 1297 AD (calculated in 2022).

After the separation time from the Sibolga language and the Minangkabau language is known, the next step is to calculate the error term. This is done to avoid miscalculations and to set a more precise separation time. It should be remembered that to anticipate errors in statistics is to give an estimate, ie not in a certain time, but in a certain time period. To calculate the error term can use the formula:

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

Information:

S = Standard error in the percentage of relative

C = Percentage of relatives kata

n = Number of words compared, both relatives and non-relatives

known: C = 0.73

n = 200

asked: S = …?

$S = \frac{\sqrt{C(1-C)}}{n}$ $S = \frac{\sqrt{0.73(1-0.73)}}{200}$

$S = \frac{\sqrt{0.73(1-0.73)}}{200}$ $S = \frac{\sqrt{0.73(0.27)}}{200}$

$S = \frac{\sqrt{0.1971}}{200}$ $S = \sqrt{0.0009855} = 0.0314\ (Simplified\ to\ 0.03)$

The result of this standard error (0.03) is summed with the percentage of initial relatives (C1) to get C2 (C2 = C1 + S). So C2 the result is 0.73+0.03=0.76. With C2, the separation time can be calculated again, using the same formula:

Given: C2 = 76% log r = 0.805

Asked: W2= .....?

Answer: $W_2 = \frac{Log\ C2}{2 log 805} = \frac{Log\ 76}{2 log 805}$

$= \frac{274}{2x217} = \frac{274}{434} = 0.631$

The separation time is multiplied by 1000 so that the result is 631
Thus, the error term = W1-W2= 725-631= 94

So, the age of the Sibolga language and the Minangkabau language can be expressed as follows:

1) Sibolga and Minangkabau languages are thought to have formed a single language about 725 ± 94 years ago.
2) Sibolga and Minangkabau were a single language from 631-819 years ago.
3) Sibolga and Minangkabau languages are estimated to have started to separate from their mother tongues around 1203-1391 AD (calculated in 2022).

4.2 Lexicostatistics of the Minangkabau and Toba Batak languages

After determining the word kin, the next step is to find the percentage of relatives with the formula:

$$C = \frac{Vt}{Vd} x100\%$$

Information:
C = relative word; Vt= number of relatives' vocabulary; Vd= amount of gloss calculated

$$C = \frac{Vt}{Vd} x100\% = \frac{47}{200} x100\% = 0.235x100\% = 23.5\% = 24\%$$

After the percentage of relatives is known, we can calculate the separation time from the Minangkabau language and the Toba Batak language

Given: C = 24%
r = 80.5%
asked: W1 = …?

Answer: $W_1 = \frac{Log\ C}{2\log 805} = \frac{Log\ 24}{2\log 805}$

$$= \frac{1427}{2x217} = \frac{1427}{434} = 3.288$$

The split time is multiplied by 1000 so that the result becomes 3288

So, the calculation of the initial separation time of the Minangkabau language and the Toba Batak language is 3288 ago.

In other words, the calculation of the initial separation time of the Minangkabau language and the Toba Batak language can be stated as follows:

1) Minangkabau language and Toba Batak language is estimated to be a single language around 3288 years ago.
2) The Minangkabau language and the Toba Batak language are estimated to have started to separate from their mother tongue around 1266 BC (calculated in 2022).

After the results of the separation of the Minangkabau language and the Toba Batak language are known, the next step is to calculate the error term. This is done to avoid miscalculations and to set a more precise separation time. It should be remembered that to anticipate errors in statistics is to give an estimate, ie not in a certain time, but in a certain time period. To calculate the error term can use the formula:

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

Information:
S = Standard error in the percentage of relative
C = Percentage of relatives kata
n = Number of words compared, both relatives and non-relatives
known: C = 0.24
n = 200
asked: S = …?

Answer: $S = \frac{\sqrt{C(1-C)}}{n}$ $S = \frac{\sqrt{0.24(1-0.24)}}{200}$

$S = \frac{\sqrt{0.24(1-0.24)}}{200}$ $S = \frac{\sqrt{0.24(0.76)}}{200}$

$S = \frac{\sqrt{0.1824}}{200}$ $S = \sqrt{0.000912} = 0.0301$ *Simplified to* 0.03)

The result of this standard error (0.03) is summed with the percentage of initial relatives (C1) to get C2 (C2 = C1 + S). So C2 the result is 0.73+0.03=0.27. With C2, the separation time can be calculated again, using the same formula:

Given: C2 = 27% log r = 0.805
Asked : W2= .....?

Answer: $W_2 = \frac{Log\ C2}{2\log 805} = \frac{Log\ 27}{2\log 805}$

$= \frac{1309}{2x217} = \frac{1309}{434}$ = 3.016

The separation time is multiplied by 1000 so that the result becomes 3016
Thus, the error term = W1-W2= 3288-3016= 272
So, the age of the Minangkabau language and the Toba Batak language can be stated as follows:

1) Minangkabau language and Toba Batak language is estimated to be a single language around 3288 ± 272 years ago.
2) The Minangkabau language and the Toba Batak language were a single language from 3016-3560 years ago.
3) The Minangkabau language and the Toba Batak language are estimated to have started to separate from their mother tongue around 1538-994 BC (calculated in 2022).

4.3. Lexicostatistics of Sibolga and Toba Batak languages

After determining the word relatives, the next step is to find the percentage of relatives with the formula:

$$C = \frac{Vt}{Vd} x100\%$$

Information:
C= relative word; Vt= number of relatives' vocabulary; Vd= amount of gloss calculated

$$C = \frac{Vt}{Vd} x100\% = \frac{53}{200} x100\% = 0.265x100\% = 26.5\% = 27\%$$

After the percentage of relatives is known, we can calculate the separation time from the Sibolga language and the Toba Batak language.

Given: C = 27%
r = 80.5%
asked: W1 = …?

Answer: $W_1 = \frac{Log\ C}{2\log 805} = \frac{Log\ 27}{2\log 805}$

$= \frac{1309}{2x217} = \frac{1309}{434} = 3.016$

The separation time is multiplied by 1000 so that the result becomes 3016

So, the calculation of the initial separation time of the Sibolga language and the Toba Batak language is 3016 years ago.

In other words, the calculation of the initial separation time of the Sibolga language and the Toba Batak language can be stated as follows:

1) The Sibolga language and the Toba Batak language are estimated to have formed a single language around 3016 years ago.
2) The Sibolga language and the Toba Batak language are estimated to have started to separate from their mother tongue around 994 BC (calculated in 2022).

After the separation time from the Sibolga language and the Toba Batak language is known, the next step is to calculate the error term. This is done to avoid miscalculations and to set a more precise separation time. It should be remembered that to anticipate errors in statistics is to give an estimate, ie not in a certain time, but in a certain time period. To calculate the error term can use the formula:

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

Information:
S = Standard error in the percentage of relative words
C = Percentage of relatives kata
n = Number of words compared, both relatives and non-relatives
known: C = 0.27
n = 200
asked: S = …?

Answer: $S = \frac{\sqrt{C(1-C)}}{n}$ $S = \frac{\sqrt{0.27(1-0.27)}}{200}$

$S = \frac{\sqrt{0.27(1-0.27)}}{200}$ $S = \frac{\sqrt{0.27(0.73)}}{200}$

$S = \frac{\sqrt{0.1971}}{200}$ $S = \sqrt{0.0009855} = 0.0314\ (simplified\ tp\ 0.03)$

The result of this standard error (0.03) is summed with the percentage of initial relatives (C1) to get C2 (C2 = C1 + S). So C2 the result is 0.27+0.03=0.30. With C2, the separation time can be calculated again, using the same formula:

Given: C2 = 30% log r = 0.805
Asked : W2= .....?

Answer: $W_2 = \frac{Log\ C2}{2\log 805} = \frac{Log\ 30}{2\log 805}$

$= \frac{1204}{2x217} = \frac{1204}{434} = 2.774$

The separation time is multiplied by 1000 so that the result becomes 2774

Thus, the error term = W1-W2= 3016-2774= 242

So, the age of the Sibolga language and the Toba Batak language can be stated as follows:

1) Sibolga language and Toba Batak language is estimated to be a single language around 3016 ± 242 years ago.
2) The Sibolga language and the Toba Batak language were a single language from 2774-3258 years ago.
3) The Sibolga language and the Toba Batak language are estimated to have started to separate from their mother tongue around 1236-752 BC (calculated in 2022).

The lexicostatistic technique not only serves to determine the percentage of kin words and calculate language age, but can also be used for grouping kin languages. Languages that show a high percentage of kinship are groups that are closer in membership, while those with a low percentage of kinship are groups whose membership level or kinship is more distant.

Swadesh proposed a classification of language kinship systems, namely:

Table 1. Classification of language kinship systems

Level Subgrouping	Years of Separation	Percentage of Cognate
Dialects of Language	0-5	100-81
Language of Family	5-25	81-36
Family of Stock	25-50	36-12
Stocks of a microphylum	50-75	12-4
Microphyla of a mesophylum	75-100	4-1
Mesophyla of macrophylum	More than 100	1- less than 1 %

(Source: Keraf, 1984)

This classification is used to find out how the position or relationship between one relative's language and another relative's language is. Before grouping the languages, it would be nice if we first looked at the percentage level of kinship of each language.

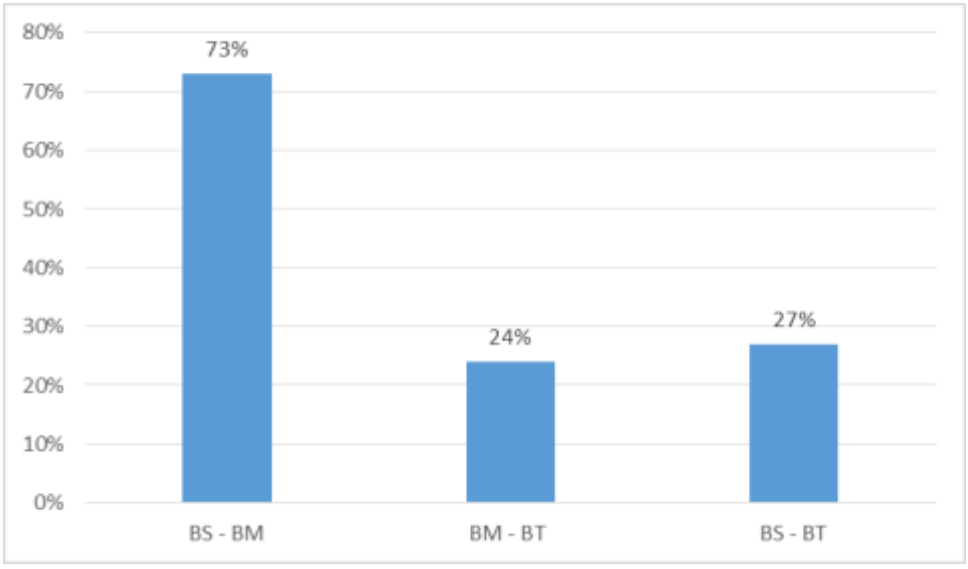

Figure 1. *Percentage Chart of Kinship of Toba Batak Language, Sibolga Coastal Language, and Minangkabau Language*

The graph above shows that the percentage of relative words from the Sibolga language, Minangkabau language, and Toba Batak language is quite varied. The largest percentage of kinship words is between Sibolga and Minangkabau languages by 73%, Sibolga and Batak Toba languages by 27%, and Minangkabau and Batak Toba languages by 24%. Based on the percentage comparison, we can conclude that the Sibolga and Minangkabau languages have a closer kinship level than the Sibolga - Batak Toba and Minangkabau - Toba Batak languages.

After knowing the percentage of kinship words from each language, the next step is to compile a branching graph or kinship lineage (stammbaum) from the Sibolga language, Minangkabau language, and Toba Batak language.

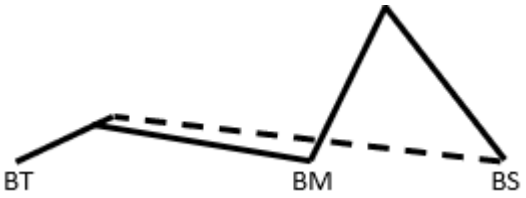

Figure 2. *Graph of Sibolga, Minangkabau, and Toba Batak kinship*

Through the description of the graph above, it can be seen that (i) the kinship tree of the Sibolga language and the Minangkabau language is closer than that of the Sibolga - Batak Toba and Minangkabau - Batak Toba languages, (ii) lexicostatistically the language status of the Sibolga and Minangkabau languages is the family category, while the language status of the Minangkabau – Batak Toba and Sibolga – Toba Batak languages are in the stock category.

V. Conclusion

Based on the description above, the authors conclude as follows:

1. Of the 200 vocabularies for Sibolga and Minangkabau, all are complete pairs. Of the 200 complete pairs, there are 146 pairs of related words or only 73% of relatives. The Sibolga language and the Minangkabau language are estimated to have formed a single language around 725 ± 94 years ago or 631-819 years ago. The Sibolga language and the Minangkabau language are estimated to have started to separate from their mother tongue around 1203-1391 AD (calculated in 2022).
2. Of the 200 vocabularies for the Minangkabau language and the Toba Batak language, all of them are complete pairs. Of the 200 complete pairs, there are 47 pairs of kin or only 24% of kin. Minangkabau language and Toba Batak language is estimated to be a single language around 3288 ± 272 years ago or at 3016-3560 years ago. The Minangkabau language and the Toba Batak language are estimated to have started to separate from their mother tongue around 1538-994 BC (calculated in 2022).
3. Of the 200 vocabularies for Sibolga and Batak Toba, all of them are complete pairs. Of the 200 complete pairs, there are 53 pairs of related words or only 27% of relatives. The Sibolga language and the Toba Batak language are estimated to have formed a single language around 3016 ± 242 years ago or 2774-3258 years ago. The Sibolga language and the Toba Batak language are estimated to have started to separate from their mother tongue around 1236-752 BC (calculated in 2022).
4. The largest percentage of kinship words is between Sibolga and Minangkabau languages by 73%, Sibolga language and Toba Batak language by 27%, and Minangkabau language and Toba Batak language by 24%. Based on the percentage comparison, we

can conclude that the Sibolga and Minangkabau languages have a closer kinship level than the Sibolga - Batak Toba and Minangkabau - Toba Batak languages.

5. The language status of the Sibolga language and the Minangkabau language is a family category, while the language status of the Minangkabau - Batak Toba and Sibolga - Toba Batak languages is a stock category.

References

Crowley, T. 1992. *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.

Keraf, Gorys. 1984. *Linguistik Bandingan Historis*. Jakarta: Gramedia Pustaka Utama.

Keraf, Gorys. 1991. *Linguistik Bandingan Tipologis*. Jakarta: Gramedia Pustaka Utama.

Kesuma, Tri Mastoyo Jati. 2007. *Pengantar (Metode) Penelitian Bahasa*. Yogyakarta: Carasvatibooks.

Mahsun. 1995. *Dialektologi Diakronis: Sebuah Pengantar*. Yogyakarta: Gadjah Mada University Press.

Mahsun. 2005. *Metode Penelitian Bahasa*. Jakarta: Raja Grafindo Persada.

Manalu, M.S. 2013, *Batak Toba Muslim: Studi Perubahan Budaya Pada Masyarakat Pesisir Di Sibolga (1970-2000)*, Departemen Sejarah Fakultas Ilmu Budaya Universitas Sumatera Utara Medan

Naim Mochtar, *Merantau, Pola Migrasi Suku Minangkabau*, Yogyakarta, Gadjah Mada University Press, 1979

Novita, Sherly., Widayati, Dwi. 2019. *Kekerabatan Kosa Kata Bahasa Karo, Bahasa Nias dan Bahasa Simalungun di Kota Medan: Kajian Linguistik Historis Komparatif.* Linguistika Buletin Ilmiah Program Magister Linguistik Universitas Udayana · September 2019

O.H.S Purba. 1997. *Migrasi Spontan Batak Toba (Marserak) Sebab, Motip dan Akibat Perpindahan Penduduk dari Dataran Tinggi Toba*, Medan, Penerbit Monora, 1997, hlm.124

Panitia Hari Jadi Kota Sibolga ke-307, *Sibolga Dalam Lintasan Sejarah*, 2007, dalam makalah tanggal 2 April 2007, hlm. 2

Purba, N. et al. (2020). Language Acquisition of Children Age 4-5 Years Old in TK Dhinukum Zholtan Deli Serdang. *Linglit Journal: Scientific Journal of Linguistics and Literature*. P.19-24

Sudaryanto. 1988. *Metode Linguistik: Bagian Kedua Metode dan Aneka Teknik Pengumpulan Data*. Yogyakarta: Gadjah Mada University Press.

----------. 1993. *Metode dan Aneka Teknik Bahasa*. Seri ILDEP. Yogyakarta: Duta Wacana Universitas Press.

Surip, Muhammad., Widayati, Dwi. *Kekerabatan Bahasa Jawa dan Bahasa Gayo: Kajian Linguistik Historis Komparatif.* Jurnal Bahasa Lingua Scientia, Vol. 11, No. 1, Juni 2019