PRELIMINARY STUDY OF THE DIALECTS OF KAMBERA

Raynesta Mikaela Indri Malo
Universitas Kristen Wira Wacana Sumba
raynesta@unkriswina.ac.id

ABSTRACT

Kambera language is a regional language used by speakers who live in the eastern part of the island of Sumba, in this case East Sumba Regency. In previous studies, the claim stated by the researchers is that there is only 1 language spoken in East Sumba. However, the people of East Sumba, as native speakers, claim that there are several languages used depending on the villages. This study aims to map these languages/dialects. This descriptive preliminary study is limited to make an inventory of basic vocabulary between 11 different locations determined based on the number of paraingu namely kabihu unions or ethnic clans in the past. The basic vocabularies are taken from Sawdesh 200-list and Leipzig-Jakarta 100-list. The combination of these lists results in 223 basic vocabularies which become the main data for this study. The data were collected through recording and note-taking.

Keywords: dialect; Kambera; Sumba

INTRODUCTION

Kambera is a native language spoken by roughly 150,000 speakers in the eastern part of Sumba Island of the province Nusa Tenggara Timur in eastern Indonesia. Blust classified it as belonging to the Central-Malayo-Polynesian subgroup. Kambera is included in the Austronesian language family (Blust, 2013). In Sumba, there are several other indigenous languages spoken including Weyewa (75,000), Kodi (40,000), Lamboya (15,000), Wanukaka (10,000), Anakalang (14,000) and Mamboru (16,000) (Wurm, 1994). Of those languages, Weyewa and Kodi, in particular, appear to be incomprehensible by speakers of Kambera (Klamer, 2011). It means that speakers of Kambera cannot understand Weyewa and Kodi. However, they can understand the other languages (or dialects) spoken in other Sumba regions. It is not yet decided whether the others are languages or dialects.

Kambera itself is spoken in all area of Kabupaten Sumba Timur, the region of the eastern part of Sumba Island. Since 2007, this region is divided into 22 kecamatan or subdistricts for governmental importance ("http://www.sumbatimurkab.go.id/sejarah.html," n.d.). However, as generally happens in Indonesia and other countries, areas are not only differentiated geographically but also culturally. People are not only classified because they live in different geographic areas but also because they are culturally different. It means that they hold certain kinds of cultures, customs, and languages that are different from other areas. These phenomena also happen in Sumba Timur where geographic borders lead to more differences such as for languages. People of Sumba Timur are culturally and historically divided into two legal institutions or groups. The first is Kabihu which refers to the clans or gens. The second is Paraingu which refers to the territorial groups consisted of several kabihu. In 1970, as the Dutch came to Sumba, the region was divided into 7 swapraja which consisted of one or several paraingu, (i) Lewa – Kambera, (ii) Kanatang, (iii) Tabundungu, (iv) Masu – Karera, (v) Melolo, (vi) Rindi – Mangili, and (vii) Waijilu (Kapita, 1976).

Klamer mentioned that "Kambera is spoken in the whole eastern region of Sumba with different degrees of dialectical variations" (Klamer, 1998). It means that, although this region is now divided into 22 subdistricts geographically, 9 at the time Klamer did her project in 1998, and 7 swapraja in 1910, the limit is not always about the geographic area but also in language problems. It is clear that

Klamer mentioned it as dialects, not languages. The language spoken in all eastern area of Sumba Island is just one language with different dialects. A rule that is applied to this notion is that "when dialects become mutually unintelligible – when the speakers of one dialect group can no longer understand the speakers of another dialect group – these 'dialects' become different languages" (Fromkin, 2000). This rule cannot be applied in Sumba Timur since the speakers of Kambera from different areas of kecamatan or swapraja can understand each other although not fully. "We shall …… refer to dialects of one language as mutually intelligible versions of the same basic grammar, with systematic differences between them" (Fromkin, 2000). Wakelin (1989) defined dialect as all the linguistic elements in one form of a language including phonological, grammatical, and lexical elements.

This study is a comparative linguistic study that aims at comparing the dialects of Kambera language. The comparison discussed here is limited to the lexical elements or the basic vocabularies, not including the grammatical and phonological elements. However, as this is a preliminary study, it is hoped that the data will lead to a more comprehensive discussion, which might include the grammatical and phonological aspects of Kambera language. This study is aimed at making inventory of the basic vocabularies of Kambera based on the 11 areas mentioned previously. These comparisons, differences, and similarities are hoped to be the basis for further analysis of this language, for instance to map out the dialects of Kambera spoken in Sumba Timur and finding the percentage of lexical cognates between the dialects which then can help to determine whether they are dialects or separate languages.

Dialect refers to a variety of language, which is related to a certain regional or social group (Fought, 2006). A language may have several dialects when it is spoken in different areas by different social or ethnic groups. Each group has its own way of speaking the language. It is what a dialect is. A dialect is the characteristic of a particular group of the language's speakers. The dialects of a particular language are closely related and, although they are different in particular aspects, they are most often largely mutually intelligible to speakers of the same language, especially if close to one another on the dialect continuum. A dialect does not only emerge due to regional speech patterns but also due to other aspects such as social class or ethnicity.

Wakelin (1989) defined dialect as all the linguistic elements in one form of a language including phonological, grammatical, and lexical elements. It is in line with Finegan who defined a dialect as a language variety in its totality including vocabulary, grammar, pronunciation, and even pragmatics aspects (Finegan, 2004). It can be concluded that a dialect is a form of a language that is spoken in a particular area and that uses some of its own words (lexicon), grammar, and pronunciations.

As a characteristic of a particular social group, dialects can be the characteristic of regional, ethnic, socioeconomics, or gender groups. A dialect that is associated with a particular social class can be termed a sociolect, a dialect that is associated with a particular ethnic group can be termed as ethnolect, and a regional dialect may be termed a regiolect (Wolfram and Schilling, 2016).

It is not always easy to decide whether a dialect should be regarded as a dialect of a particular language or as an independent language. Mutually intelligible is the key to deciding it. It means that "when the dialects become mutually unintelligible – when the speakers of one dialect group can no longer understand the speakers of another dialect group – these 'dialects' become different languages" (Fromkin, 2000). The contrary means that the dialects are from the same language.

The cause of dialectical differences might be found in the fact that people or groups of people are separated from each other geographically and socially. The separation leads to the lock of certain language features and changes of particular language use in that group only. Some features or aspects might be used only and are spread and inherited in that group and cannot be accessed by other groups due to the separation. "When some communication barriers separate some groups of speakers – be it physical barriers such as an ocean or a mountain range, or social barriers of political, racial, class, or religious kind –

linguistics changes are not easily spread and dialectical differences are reinforced" (Fromkin, 2000).

There are previous researches that studied the languages in Sumba and Kambera in particular. Klamer mentioned some of them in her work

"The oldest publication in which a language of Sumba is mentioned is Heymering (1846), which contains a brief word list of the dialect of Mangili. Other word lists are given by Roos (1872), De Roo van Alderwerelt (1891), Vermast (1895), Pos (1901), and Wielenga (1917). Van der Velden (1900) is a brief description of the language of the Laura district in western Sumba. Wielenga (1909), Onvlee (1925) and Kapita (1983) are short grammars of Kambera. Klamer (1998a) is a recent grammar. The two Kambera dictionaries are Kambera–Indonesian (Kapita, 1982) and Kambera–Dutch (Onvlee 1984). Kambera texts with Dutch translations are Wielenga (1913) and Onvlee (1925). Kambera traditional ritual speech, songs, stories, and sayings are documented by Kapita (1977, 1979, 1986, 1987). These four publications, together with the Kambera New Testament ('Na Paràndingu Bidi', 1961) and Hymnbook ('Ludu Pamalangu', 1979), make up the Kambera written literature. Kambera ritual speech is discussed in Forth (1981) and Fox (1988), while Kuipers (1998) deals with aspects of ritual speech in Weyewa. Forth (1985) is an antropological study of the Kambera-speaking community of Rindi" (Klamer, 2011).

Klamer herself did a study on Kambera language that focused on the phonology, morphology, and morpho-syntax aspects of the language. However, since the study was conducted in only one village, it did not discuss the dialect issues. It is mentioned in the study that Kambera "is spoken in the whole eastern region of Sumba with different degrees of dialectical variation". Though, there is no explanation of the variation.

Another study on Kambera and other languages in Sumba was conducted by Budasi (Budasi, 2009) who studied the genetic relationship of isolects in Sumba. There were 7 isolects studied in this research Kodi (Kd), Wewewa (Ww), Laboya (Lb), Kambera (Km), Mamboro (Mb), Wanokaka (Wn), and Anakalang (An). There were also 2 other languages from outside Sumba that were used for comparative matter, Sawu language from Sabu Island, and Bima language from Nusa Tenggara Barat Province. The study was aimed at describing whether the 7 isolects can be regarded as dialects or as different languages. It also described the genealogy relationship of the 7 isolects. The data of this study consisted of linguistic data in the form of lexicons from the 7 isolects and the other two languages which were captured using the Swadesh 200 list vocabulary, images and recording devices in the form of tape recorders. The type of data identified consists of secondary data and primary data. Secondary data was taken from the dictionaries of the 7 isolects and of the other two languages, while the primary data were the lexicon taken from 3 informants from each isolect and the two languages studied. The data obtained were analyzed by the lexicostatistic technique, namely the same or similar lexicon (cognate) divided by 200 after being reduced by the amount of empty gloss. After the percentage of kinship level between isolects and the two languages can be determined, the status of the seven isolects is determined whether as dialect or language through the formula for determining the level of language grouping determined by Swadesh (1952). This study concludes that the 7 isolects are different languages and are in one language group, the Sumba language group.

Leif Asplund (2010) in his research on languages in Sumba did a lexicostatistics study on languages and dialects spoken in Sumba and made a list of categorizations the languages into groups and sub-groups. However, the basis for the groupings was not mentioned in the paper. It might be based on the lexicostatistics count for a cognate percentage between the languages and dialects. The percentage amount determines a particular language or dialect that refer to which group or subgroup. Asplund grouped the languages into 3 big groups namely (1) Central-East Sumbanese (CESu) (with subgroups of Central Sumbanese (CSu), Mamboru, and East Sumbanese (ESu)), (2) Wewewa – Laboya (with subgroups of Wewawa Group (WeG) and Laboya (Lab), and (3) Kodi Group (KoG) (with subgroups of Kodi (Ko) and Gaura (Gau)) (Asplund, 2010).

Kambera is included in the subgroup of East Sumbanese which comprises all of the following dialects (1) South-eastern Sumbanese, (2) Mangili, (3) Kambera-Kanatang-Prai Paha, (4) Lewa-Tidas, (5) North-eastern Sumbanese, (6) Umbu Ratu Nggai, (7) Praiwatana, (8) Maderi (Asplund, 2010).

The groupings above might seem confusing because Asplund did not do it based on paraingu or kabihu, nor the administrative governmental groups. From the grouping, we can see that in south part of east Sumba, there are many dialects spoken. It also happens in other parts. We can also see those places such as Umbu Ratu Nggay, Praiwatana, and Maderi are included in this group of "East Sumbanese" although administratively based on cultural grouping or the governmental grouping they belong to the central Sumba. This study can be good basis for dialects analysis because it uses lexicostatistics technique to count the cognate words between the dialects. This study also helps researcher to anticipate more data needed or more data shown during the research that is based on administrative groupings. This can happen because it is not clear yet how historically the areas or parts of Sumba Island were divided into groups. Areas that are included in the same group today, might originally belong to other groups. As Asplund did the lexicostatistics counting for all the languages and dialects, then his result can also be used in this research as data for comparison.

Aritonang et al. (2002) did a study on Swadesh basic vocabularies in various locations in Indonesia, including in East Sumba by taking 7 observation locations. This research was limited to making a list of 200 Swadesh words and not to look for diachronic relationships from the data obtained.

METHOD

This study is a descriptive study to show the language use in 11 different areas in east Sumba, represented by chosen vocabularies. This study will show the similarities and differences of the languages/dialects.

The data for this research are the Kambera version of the Swadesh 200 words-list. Swadesh list, created by Morris Swadesh, is a classic compilation of basic concepts for the purposes of historical-comparative linguistics. It includes 200 words that are regarded to be universal in concepts and exist in all cultures around the world. Translations of the Swadesh list into a set of languages allow researchers to quantify the interrelatedness of those languages. It is used in lexicostatistics (the quantitative assessment of the genealogical relatedness of languages) and glottochronology (the dating of language divergence). There will be 11 informants taken from 11 paraingu mentioned in Kapita (1) Lewa, (2) Kambera, (3) Kanatang, (4) Napu, (5) Tabundungu, (6) Masu, (7) Karera, (8) Melolo, (9) Rindi, (10) Mangili, and (11) Waijilu.

This study is aimed at making an inventory of the vocabularies of Kambera found in 11 locations. The data will be collected through recording and note taking. The informants will be asked to mention each word in Kambera using their own pronunciation related to their respective paraingu. The words are from the Swadesh 200-word list. The criteria of the informants are (1) adult male or female East Sumbanese between 40 – 65 years old, (2) speak Kambera as their first language and understand Bahasa Indonesia well, (3) able to read and write in both languages, and (4) have not lived outside their respective Paraingu.

DISCUSSION

The result of this study is the inventory of vocabularies of Kambera based on the Swadesh list used as the basis of this study. As the analysis is done, there are some modifications of data used in this analysis to adjust the need of this research. The basis data used is not only the 200 words of Swadesh because Swadesh himself made some modifications for the list. For example, there is another Swadesh list consists of 100 words. There are 7 words from the 100-word list that were not in the 200-word list. The final number words used from Swadesh list in this study are 207 words.

According to Keraf (1996: 126) Swadesh List "consists of non-cultural words, and the retention of basic words has been tested in languages that have written texts". However, in reality not all words in the Swadesh List are universal and apply to all cultures in the world. Therefore, in this study, 2 words from the list were discarded ('ice', 'snow') because

geographically and culturally it did not exist in Sumba. In addition, in this study also added 20 words from The Leipzig - Jakarta List compiled in 2009 which is a development of the Swadesh List by taking into account various data from other languages. This list consists of 100 words and was created by Max Planck Digital Library. Within these 100 words, there are words that also appear in the Swadesh List.

The total words from the two lists are 227 words. However, because the words ice and snow are not words that are culturally derived from Sumba, they are discarded. There are also two other words from the list that should be omitted. First is 'foot' because as in Bahasa Indonesia, in Kambera, foot is just a part of leg and does not have its own term. Second is the third singular pronoun. This item is omitted because it should be combined with another item referring to the same meaning. In Kambera, as in Bahasa Indonesia, there is no difference between 'he', 'she', and 'it' and they are only represented by the word 'dia'. The total of all words used is 223 words. Each informant was interviewed and asked to state the meaning of each word in the language they use in the place they live, East Sumba. The informants are from 11 paraingu or villages as it was mentioned previously.

The 223 words include different lexical categories. There are 99 nouns, which mostly consist of nouns related to body parts of human and animal and also related to nature. Other nouns are related to people, names of animals, plants, time, numbers, and direction. Another lexical category is adjectives which include 40 words that related to different thing such as distance, taste, condition, weather, color, amount, and size. There are 15 pronouns which include personal pronoun, interrogative pronoun, and demonstrative pronoun. There are also 63 verbs and 6 partciples.

CONCLUSION

Besides the inventory, there are some important notes:
1. The word foot and leg in English culture might have different concept culturally in Sumba. While the two words refer to partially different things in English, foot refers to a particular part of a leg and the leg refer to the whole part of the limb used to walk, they do not act the same way in Sumbanese culture. The limb is only expressed using one word. It is the reason to omit the word 'foot'.
2. There are some words that have several meanings based on the contexts. For example, action verbs based on tools used in the action. The word stab refers to the act of stabbing using a tool. However, the use of different tools also causing the use of different verb.
3. Another example is the word 'fall'. The point where someone or something fall from, decides which word to be used in certain context. For example, the water that falls in the waterfall is explained using different word from the fall of a fruit from the tree.
4. From all the vocabularies above, there are words that are attached to affixes. These facts are also important to be addressed since lexicostatistics do not allow the use of bound morphemes such as affixes in counting the cognate.

The result of this study can be the basis of other research for example to find the cognates and to decide the separation of the dialects.

REFERENCES

Aritonang, Buha, Non Martis, Hidayatul Astar, and Wati Kurniawati. 2002. *Kosakata Dasar Swadesh Di Kabupaten Belu, Ngada, Sumba Barat, Sumba Timur, Dan Timor Tengah Utara*. Jakarta: Pusat Bahasa Departemen Pendidikan Nasional.

Asplund, Leif. 2010. *The Languages of Sumba*.

Blust, Robert. 2013. *The Austronesian Languages*. Australian National University.

Budasi, I. Gede. 2009. "RELASI KEKERABATAN GENETIS KUANTITATIF ISOLEK-ISOLEK SUMBA DI NTT: Sebuah Kajian Linguistik Historis Komparatif." Pp. 216–24 in *Masyarakat Linguistik Indonesia*. Malang: Kongres Internasional Masyarakat Linguistik Indonesia (KIMLI).

Finegan, E. 2004. *Language: Its Structure and Use*. Massachusetts: Thomson Wadsworth.

Fought, Carmen. 2006. *Language and Ethnicity*. Cambridge: Cambridge University Press.

Fromkin, Victoria. 2000. *An Introduction to Language / Victoria Fromkin, David Blair, Peter Collins*. edited by D. Blair 1942- and P. C. Collins. Marrickville, N.S.W: Harcourt Australia.

Kapita, U. H. 1976. *Masyarakat Sumba Dan Adat Istiadatnya*. Panitia Penerbit Naskah-Naskah Kebudayaan Daerah Sumba, Dewan Penata Layanan Gereja Kristen Sumba.

Klamer, Marian. 1998. *A Grammar of Kambera*. edited by G. Bossong and B. Comrie. Berlin: Mouton Grammar Library.

Klamer, Marian. 2011. "Kambera." in *The Austronesian Languages of Asia and Madagascar*, edited by A. Adelaar and N. P. Himmelmann. London: Routledge.

Wakelin, M. F. 1989. *The Archaeology of English*. Barnes & Noble Books.

Wolfram, Walt, and Natalie Schilling. 2016. *American English: Dialects and Variation*. West Sussex: Wiley Blackwell.

Wurm, S. 1994. *Australasia and the Pacific*. In Moseley, Christopher and R. E. Asher (eds.), Atlas of the World's Languages, 93-130. Cambridge: Cambridge University Press.