CHAPTER II

LITERATURE REVIEW

In order to have better understanding about the lexical differences in Madurese variety, these theories are presented as guidance in describing the lexical differences and determining the status of Madurese variety spoken by people in Kangean Island, Sumenep Regency. Furthermore, these theories provide comprehensive conceptual understanding to discover the details about dialects in guiding the study.

2.1 Related Theories

2.1.1 Variety, Language, Dialect, and Accent

According to Chambers and Trudgill (2004, p. 5), variety is the neutral term to apply to any particular kind of language which is considered as a single entity. Talking about variety will lead people to think about language, dialect and accent because variety can be in the level of language, dialect, and accent. People will also question the difference between language, dialect and accent. Chambers and Trudgill (2004, p.5) have differentiated language based on its linguistic contents which are distinct, codified, standardized form with its orthographies, grammar, and literatures. For example, British English and American English are considered as different languages because of the standardized form with its orthographies. In British English,

the term *colour* is the standard form, while in American English, the standard form is *color*. However, according to Max Weinreich cited in Campbell (1988, p. 193), language is not only about linguistic matter but also political matters. He added example of Swedish and Norwegian that are mutually intelligible but they are still considered as different languages because of political reasons. A language is also claimed as a collection of mutual intelligible dialect (Chambers & Trudgill, 2004, p.3). It means that language is divided into some dialects. The definition is in accordance with the definition of language based on Holmes (2008, p. 135). It is stated that language is a collection of dialects that is linguistically similar used by different social groups who choose to say that the speakers of a language that functions to unite and represent them compare with other groups. Based on the definitions, dialect is then considered as the subdivision of language. According to Trudgill (2004, p. 2), dialect is the particular combination of words, pronunciation, and grammatical form spoken by people who are from same area and social background and that combination differs from other people from different area and social background. The dialect definition is related to the concept of accent. Accent refers to the way in which speakers pronounce their characteristic of speech (Chambers & Trudgill, 2004, p. 5). The accent, then, can help people to convey information about the speakers" dialect, such as what region or social class the speakers are from. For instance, people can identify that particular speaker uses regional dialect of Britain if he or she pronounces „luxury" as [lʌgʒəri] and use regional dialect of America if they pronounce it as [lʌkʃəri]. Another example is that

in Britain, people will assume particular speaker is from lower class society if he or she pronounces –*in* in „swimming‟ as [in]. Meanwhile, they will assume the particular speaker is from higher class society if he or she pronounces the –*in* as [iŋ].

2.1.2 Types of Dialect

There are two types of dialect as explained by Trudgill (2004, p. 2). The first type of dialect is social dialect. It is a dialect that emerges as a result of particular kind of social background (Trudgill, 2004, p. 2). For example, people who work in official government tend to use the higher variety of their dialect, while people who work as farmer, for example, tend to use the lower variety of their dialect. Another example can be seen in America where African descent speaks different dialect than those of European, Asian or Hispanic descent (Fromkin, Rodman & Hyams, 2011, p. 439). The differences can be seen from the different social background between African descent and those other descents. The dialect used by the Africans is likely the lower varieties because mostly Africans in the past were slaves.

The second type of dialect is regional dialect, which refers to linguistic differences that accumulate in a particular geographic region (Trudgill, 2004, p. 2). For example, English which is spoken in Britain is different from the one which is spoken in America. The difference can refer to lexical, phonological, semantic and other linguistic varieties. For example in terms of lexical aspect, in Britain, people will use term *sweet*; while, in America, people will use term *candy*. However, in this

study, I focused on regional dialect because the phenomenon of lexical differences in this study is found and differentiated by particular region.

2.1.3 Regional Dialect

Regional dialect refers to linguistic differences that accumulate in a particular geographic region (Trudgill, 2004, p. 2). It means that the dialect which is used in particular region is different from other region. The linguistic differences accumulated in particular region can also be the result of language and dialect contact. The language and dialect contact can emerge not only because of outsider but also geographical reason where such area is located near other area whose people speak different language or dialect. It is in line with the „wave theory". According to Campbell (1998, p. 189), changes due to contact among languages and dialects are spread from central area of the spreading like a wave. It is like a stone which is thrown into it and result in wave. It will not stop in dialect border but continue to language border. The area, which is located nearer to the central area of the spreading, shows higher number of similarity. Furthermore, if there is more than one central area of the spreading, it will result in transition area. Then, transition area shows the borrowing features which also show the existence of contact. In addition, the number of borrowing features depends on the speaker, situation, and the native speakers" openness towards the contact (Romaine cited in Laksono, 2004, p. 12).

Furthermore, people from different regions can still understand each other because there is a chain of mutual intelligibility that connects them. The mutual

intelligibility may also be not equal in both directions because of some factors, such as listener˙s degree of exposure to the language, their degree of education and willingness to understand (Chambers & Trudgill, 2004, p. 4). In regional dialect, the degree of mutual intelligibility is also related to the distance among regions. Chambers & Trudgill (2004, p. 3) added that the greater the geographical separation is, the greater the difficulty of comprehension of linguistic differences is. For example, in Madurese varieties, people in Kangean might have greater difficulty in understand Bangkalan varieties than Sumenep varieties. It is caused by the condition where Kangean Island has greater geographical distance to Bangkalan than Sumenep.

The linguistic differences that accumulate in particular region and differentiate it from other regions are studied by researcher by using synchronic research. Synchronic research is conducted to describe language phenomenon in particular time (Mahsun, 2005, p. 84). It is different from another kind of research in dialectology, called diachronic research. Diachronic research aims to analyze the historical development of language (Mahsun, 2005, p. 84). The linguistic features in dialectology are based on phonology, lexical, syntax, morphology and semantic (Mahsun, 2005, p. 137). Phonological difference refers to how the speakers in particular area pronounce word in different way from other speaker in other areas. For example, in the southern area of America, the term *creek* is pronounced with tense [i] as [krik]. On the contrary, in the north Midlands, it is pronounced with lax [ɪ] as [krɪk]. Meanwhile, a lexical difference is the difference of word that people use to name the same object (Chambers & Trudgill, 2004, p. 54). Furthermore, lexical

difference refers to contrast in the words used by different speakers to characterize the same object ore action (Chambers & Trudgill, 2004, p. 54). For example, British people use the term *flat* to refer to rented room, while Americans use the term *apartment*. In addition, syntax difference describes different sentence formation (Chambers & Trudgill, 2004, p. 54). For example, the grammar of Scottish Standard English is different from the grammar in Standard English of England. The speaker of Scottish Standard English will say *"Had you a good time last night?"*; while the speaker of Standard English of England will say *"Did you have a god time last night?"* (Trudgill, 2004, p. 7). Meanwhile, morphological difference discusses the different word forms (Chambers & Trudgill, 2004, p. 54). The example of morphological differences in dialect is that past tense form of *learn* in America is commonly attached by suffix *–ed*, such as *learned*, while it is commonly attached by suffix *–t* in Britain, such as *learnt*. Semantic involves the difference in meaning of same words (Chambers & Trudgill, 2004, p. 56). For example, in Britain, *biscuit* means a sweet and dry flat cake, while in America; *biscuit* is a small airy roll (Roca-Varela, 2006, p. 133). In this study, the linguistic difference that was being studied is lexical differences because the varieties in Kangean Island differ in term of lexical items existing conspicuously among villages in Arjasa and Kangayan Sub-districts.

2.1.4 Lexical Differences

A lexical item or lexeme is a term that refers to the minimal distinctive unit conventionally listed in the dictionaries (Crystall, 2008, p. 276). On other words, all

words listed in dictionary are called lexical items. There are word categorizations that people know, such as nouns, verbs, adjectives, adverbs, pronouns, propositions, articles, and conjunctions. They are divided into two kinds of word. Fromkin, Rodman & Hyams (2011, p.38) divided words into content words and function words.

Content words involve nouns, verbs, adjectives, and adverbs and have portion to do with meaning. Objects, actions, attributes, and ideas are denoted by content words (Fromkin, Rodman & Hyams, 2011, p. 39). The examples of content words are *table, chair, doctor* for nouns; *kick, walk, sing* for verbs; *beautiful, warm, happy* for adjectives; and *quickly, yesterday, here* for adverbs.

Function words involve conjunctions, prepositions, articles and pronouns. Function words function to connect words to the larger grammatical context (Fromkin, Rodman & Hyams, 2011, p. 40). The examples of function words are *and, or, but* for conjunctions; *in, at, of* as prepositions; *the, a, an* as articles; and *it, her, his* as pronouns.

Lexical difference refers to contrast in the words used by different speakers to characterize the same object or action (Chambers & Trudgill, 2004, p. 54). To decide whether the words belong to different word or not, it should be analyzed from the root (Mahsun, 2005, p. 122). A root is the base form of a word which cannot be further analyzed without total loss of identity, such as the word *meaningfulness*. Removing *-ing, -ful* and *–ness* results in the root *mean* (Crystal, 2008, p. 419). For example, in Sumenep varieties, some people in western part of Sumenep call

„moming‟ by *ghulagghu* or [gʰʊlakgʰʊ], and others in eastern part of Sumenep call it by *lagghu* or [lakgʰʊ]. Then, they are considered as the same word because the first word is added by affix -gʰʊ. In this study, I only focused on noun, verb, adjective, adverbs, conjunction, prepositions and pronoun because those words are considered to be the representative of lexical difference presented in dialect areas.

The lexical difference is written in both Roman alphabet and International Phonetic Alphabet (IPA). The IPA functions to explain reader how to pronounce a word that is used in a particular area. Here are the tables of list of the IPA symbols of American English with the example of the word that use such symbol in its sound. The part of the words whose such sound is in bold font (Fromkin, Rodman & Hyams, 2011, p. 216):

Table 1: Table of Consonant Sound Symbols of American English IPA

Consonants							
Symbol	Example	Symbol	Example	Symbol	Example	Symbol	Example
p	ti**p**	k	s**k**in	θ	**th**rough	r	**r**eef
pʰ	a**pp**ear	kʰ	**c**riti**c**	ð	wea**th**er	j	**y**ou
b	**b**it	g	**g**irl	ʃ	mu**sh**	w	**w**itch
m	s**m**ack	ŋ	si**ng**	ʒ	vi**s**ion	ʍ	**wh**ere
t	kiss**ed**	f	**f**at	ʧ	mat**ch**	h	**h**at
tʰ	a**tt**ack	v	**d**ove	ʧʰ	**ch**oke	ʔ	bu**tt**on
d	ri**d**e	s	pa**ss**	ʤ	mi**dg**et	ɾ	wri**t**er

n	snow	z	jazz	l	call

Table 2: Table of Vowel Sound Symbols of American English IPA

Vowels							
Symbol	Example	Symbol	Example	Symbol	Example	Symbol	Example
i	beet	æ	laugh	o	grow	aɪ	bit
ɪ	bit	u	boot	ɔ	core	aʊ	about
e	gate	ʊ	foot	a	father	ɔɪ	boy
ɛ	bet	ʌ	cut	ə	sofa		

In addition, the consonant [ñ] is also added as one of the consonant sound that found in Bahasa Indonesia and Madurese (Hasibuan, 205, p. 20). The consonant [ñ] represents sound /ny/ as in word *minyak* in Bahasa Indonesia. According to Sofyan (2010, p. 207), there are additional consontant and vowel sounds in Madurese. For consonant sounds, they are [bʰ], [dʰ], [Dʰ], [gʰ], [jʰ], and [T]. Meanwhile, for vowel sounds, they are [aj], [oj], and [uj].

Furthermore, the lexical difference is drawn on a map by using isogloss. Isogloss is a line drawn on the map to separate an area which has one word from an area which has another (Trudgill, 2004, p. 23). It means that one line depicts one lexical. The line is drawn in one display map. The isogloss can show obviously the lexical difference among regions which was being observed. Chambers & Trudgill

(2004, p. 54) described five categories of isogloss based on linguistic structure. The first category is pronunciation isogloss, which refers to describing the contrasting pronunciation. The second is phonetic isogloss, which is used to contrast the phonetic differences of different regions. Phonemic isogloss is the third category of isogloss which describes dialect resulted from the phonemic variety. The fourth category is lexical isogloss that describes the difference word used by speaker naming same object or action. The last isogloss category is syntactic isogloss which involves some aspects of sentence formation. In this study, I focused on lexical variety that led me to focus on lexical isogloss as well. Furthermore, group of isogloss that forms a bundle is then called bundle of isogloss (Mahsun, 2005, p. 164). In line with that, Chambers & Trudgill (2004, p. 60) added that the more bundling it has, the more contrastive it is on either side.

The lexical differences in a particular are further calculated in order to determine the status of the lexical differences. The percentage of lexical differences becomes the determiner of status of variety. Guiter cited in Mahsun (2005, p. 154) categorized the status into five, as the following:

0% - 20%	= no difference
21% - 30%	= different speech
31% - 50%	= different sub-dialect
51% - 80%	= different dialect
80% - 100%	= different language

Furthermore, the formula used to determine the lexical status is called Dialectometry formula (Mahsun, 2005, p. 154). The percentage of comparison among villages determined the status of lexical differences in Kangean Island. Here is the dialectometry formula used to determine the lexical status:

$$\text{Index } d\% = \frac{s x\ 100\%}{n}$$

Details:

d = vocabulary distance in percentage (%)

s = number of lexical differences compared to other OP

n = number of total lexical difference

In addition, there are some principles which are important to be considered in determining the lexical status (Mahsun, 2005, p. 156). They are:

a) If one observation point has one or more lexical items for one meaning and one of them is used in other observation points also, then it can not be considered as lexical differences.

b) If one observation has no lexical form for a particular meaning, then it is considered as lexical differences

c) If all observation points do not have lexical form of a particular meaning, then it is considered as equal or no lexical difference

d) In the calculation of lexical dialectometry, the difference in phonology and morphology that appears is not taken into lexical difference.

2.1.5 The Monograph of Kangean Island

According to Fattah in *Asal Usul Pulau Kangean Sumenep* (2013), the name "Kangean" is obtained from Madurese word *ka-aengan* which means "submerged". Formerly, this island was an as exile area for the political and local authority"s enemies and also the prisoners (Farjon cited in Bustami, 2003, p. 72). In addition, Bustami (2003, p.74) stated that people who had come to Kangean Island are from Madura Island, Poday Island, China, Arab, Banjar, Malay, Bawean, Java, Bali, Bugis-Makassar and Mandar.

Kangean Island itself is located in the eastern part of Sumenep Regency across the sea. Its width is about 446, 67 hectare and the number of population has reached 81.427 inhabitants in 2012 (Sumenep Regency Central Bureau of Statisticss, 2012). There are two sub-districts in Kangean Island, which are Arjasa and Kangayan Sub-districts. Arjasa Sub-distrcit is located in the western part of Kangean Island, while Kangayan Sub-district is in the eastern part of Kangean Island.

Arjasa Sub-district is comprised of 19 villages. They are Kalisangka, Laok Jang-jang, Sumber Nangka, Bilis-bilis, Duko, Kali Anyar, Angkatan, Kolo-kolo, Buddi, Gellaman, Sawah Sumur, Paseraman, Pandeman, Pabiyan, Sambakati, Angon-angon, Kalikatak, Arjasa and Pajanangger. Its width is 241, 99 hectare and the number of population of this sub-district in 2012 is 60.592 inhabitants (Sumenep

Regency Central Bureau of Statistics, 2012). The majority of people in Arjasa Sub-district work in agriculture. It is shown by Sumenep Regency Central Bureau of Statistics (2012) where 16.034 lands are used for agriculture and 8.165 are used for non-agriculture practice.

Meanwhile, there are seven villages in Kangayan Sub-districts, which are Dandung, Timur Jang-jang, Jukong-jukong, Batu Putih, Torjek, Tembayangan, Torjek, Kangayan and Saobi. According to Sumenep Regency Central Bureau of Statistics (2012), the width of Kangayan Sub-district is 204, 68 hectare with 20.835 inhabitants live there. The use of land in Kangayan Sub-district is similar to Arjasa Sub-district that is dominated by agriculture practices. The number of agriculture land in this sub-district is 11.318 and non-agriculture is 9.150 (Sumenep Regency Central Bureau of Statistics, 2012). Kangayan sub-district was just legitimized as the expansion of Arjasa sub-district on 31 January 2005 (Yashik, 2005).

2.1.6 Madurese Varieties in Kangean Island.

Madurese is spoken by 13, 7 million speakers and becomes the language with the fourth largest speaker number in Indonesia after Bahasa Indonesia, Javanese and Sundanese (Lauder cited in Sofyan, 2010, p.208). One of the reasons is that people who live in some regions outside Madura Island, such as Gresik, Surabaya, Pasuruan, Lumajang, Probolinggo, Situbondo, Bondowoso, Banyuwangi and Jember, also speak Madurese (Sofyan cited in Izzak (2012, p. 1).

There are three speech levels in Madurese. Those three speech levels are *Enja' Iya* to address younger people*, Engghi Enten* to address people of same age and

Engghi Bunten to address older people. For example, when a speaker wants to say "Where are you going?" in *Enja' Iya*, he will say"*Mangkadha* [are going to] *ka dhimma* [where] *be'na* [you]*?*". The speaker also will say "*Mangkadha* [are going to] *ka imma* [where] *dika* [you]*?*" in *Engghi Enten* level and "*Meyosa* [are going to] *ka'emma* [where] *panjennengan* [you]*?"* in *Engghi Bunten* level.

Based on the linguistic aspect, Madurese is divided into six dialects. Sofyan, Wibisono, Hendrik cited in Izzak (2012, p. 2) divided Madura into four dialects which are Sumenep dialect, Pamekasan dialect, Bangkalan dialect, and Kangean dialect. In addition, Zainuddin et al cited in Izzak (2012, p. 2) added Pinggir Papas dialect as one of the dialects of Madurese. Bawean dialect is also claimed as the sixth dialect of Madurese (Adiana et al cited in Izzak, 2012, p. 2). Besides, there are two dialects based on sociolingustic aspect. Sociolinguistics aspect emerges because of social background in relation to Sumenep Kingdom in Sumenep Regency. The nearer to kingdom the people live, the more usage of higher variety of Madurese is. Those two dialects are *Madhura Bara'* (West Madura), which is used by people who live in Bangkalan, Sampang, Pamekasan and some border areas between Pamekasan and Sumenep, and *Madhura Temor* (East Madura), which is used in Sumenep and some islands of Sumenep. The sociolingustic aspect here refers to social dialect. It is a dialect that emerges because of particular social background (Trudgill, 2004, p. 2). In this case, Sumenep dialect is distinguished from other dialects based on social class reasons. The speakers of Sumenep dialect tend to use the higher variety of Madurese.

The use of higher variety by the speaker of Sumenep dialect might be the result of the history where Sumenep was formerly a kingdom.

One of the Madurese varieties is Kangean varieties. Kangean varieties are used by people who live in Kangean Island. Kangean Island belongs to Sumenep Regency and is located in the eastern part of Sumenep across the sea. Although Kangean Island is part of Sumenep Regency, the varieties used in Kangean Island is different from the varieties used in Sumenep. The example of the difference of word between Sumenep varieties and Kangean varieties is that people in Sumenep call „later" by *dagghi'* or [dʰəggʰɪʔ], while people in Kangean call „later" by *jhuan* or [dʒʰʊɑn].

There are also speech levels in Kangean varieties, which are *Ako-Kao*, *Nira-Nae* and *Kaula-Panjennengan* (Bustami, 2003, p. 74). The first level, *Ako-Kao* is used to address people in same age. It is also called by *Eson-Sede* and *Eson-Kake*. Furthermore, *Nira-Nae* has another term, called *Die-Dika,* and is used by son or daughter-in-law to father or mother-in-law. The last speech level, *Kaula-Panjennengan*, is used to address older people. This level is dominated by Madurese For example, when a speaker of Kangean varieties wants to say "Where are you going?" in *Ako-Kao*, he will say "*Alakoa* [are going to] *kamma* [where] *kao* [you]*?*". The speaker also will say "*Alakoa* [are going to] *kamma* [where] *situ* [you]*?*" in *Die-Dika* level and "*Meyosa* [are going to] *ka' kamma* [where] *hampeyan* [you]*?"* in *Kaula-Panjennengan* level. The variety in Kangean Island differs in term of lexical items existing conspicuously among villages in Arjasa and Kangayan sub-district. For

example, in Arjasa Sub-district, people in Duko call „watermelon‟ as [somɑŋka], while people in Pajanangger call it as [gʊləŋgʊlən]. On the contrary, in Kangayan Sub-district, people in Dandung use the term [tabʰəntor] to refer to „hit‟,while people in Torjek use the term [tadʒʰəroŋgoŋ].

2.2 Related Studies

There have been previous studies that discussed the dialect variety in Madurese. The first previous study entitled *Stratigrafi Isolek-Isolek Bahasa Madura Dialek Pamekasan* was conducted in Pamekasan Regency by Asyatun (2005). She tried to describe the isolect stratigraphy of Madurese spoken by people in Pamekasan Regency by using synchronic research. The study was conducted by comparing 233 glosses in seven OPs (Observation Point) to find the lexical differences among those sub-districts. The OPs are Tlanakan, Pademawu, Pamekasan, Proppo, Larangan, Peganten and Pakong Sub-district. In each OP, she had three informants. In total, she had 21 informants in her study. In this study, she conducted the research by doing interview, taking notes and recording. The study found that out of 233 glosses, there were 92 lexical differences and they were assumed as speech different. The status was determined as different speech based on the percentage result. The gap of her study and my study is that her study focused on the varieties in Pamekasan, while my study focused on the varieties in Kangean Island. The formula used in determining the status of lexical variety of her study is based on Lauder‟s calculation, while my

study used Guiter‟s formula which is cited by Mahsun. In addition, the number of glosses, which were being questioned in my study, was more developed.

The second previous study that studied the dialect variety of Madurese was conducted by Soegianto et al (1982) and (1985). The first study was conducted in 1982 entitled *Pemetaan Bahasa Madura di Pulau Madura.* The study aimed to describe the mapping of Madurese based on phonological, morphological, syntactical, lexical, and intonation aspect to documentary and development of theory of local languages in Indonesia. This study was conducted in four regencies in Madura Island, which are Bangkalan, Sampang, Pamekasan and Sumenep. In Bangkalan, the writers took 18 sub-districts, and in Sampang, they had 12 sub-districts as the location of the research. They also had 11 sub-districts in Pamekasan and 17 sub-districts in Sumenep. The informants were 120 and the number of isogloss that the writers used in this study were 600 words. It consisted of 550 words of the lowest speech level and 50 words of the highest speech level. This study‟s result was that there is only idiolect variation based on the phonological aspect. For morphological aspect, there is variation of suffix /-na/ in the word with glottal stop ending. The writers found no variation of syntactical aspect in Madurese spoken in Madura Island, while for the variation of intonation, it was found that Sumenep dialect have slow tempo and utter the words completely. Then, for lexical aspect, the writers found 26 different variations of words based on the synonym and word form change. The difference between the previous study and my study is that the study above described and compared Madurese varieties which were spoken in Bangkalan, Sampang,

Pamekasan and Sumenep. Moreover, the study did not take the islands of Sumenep as observation point, for example, Kangean Island. Whereas, Kangean Island is part of Sumenep Regency and the varieties used by people there is categorized as the variety of Madurese.

For the next study, Sugianto et al (1985) conducted study entitled *Struktur Bahasa Madura Dialek Kangean*. The study described and compared between the linguistic structure of Madurese varieties used in Kangean and other regions in Madura Island. The linguistic structure studied is based on the phonological, morphological and syntactical aspects. The writers had 15 informants and different number of gloss in each linguistics aspect. From phonological aspect, the writers used 450 words as the instrument. The result shows that the phoneme and its distribution is the same as the phoneme of Madurese; there is variation of lexical sound, which is different from lexical of Madurese in same meaning; and there is variation of utterance of phoneme /i/, phonem /ö/ and phonem /u/. The morphological aspect has 180 derived words and the result shows that the affix, repetition and compound morpheme is the same as Madurese. The last result also provides result of syntactical aspect. The writers used 60 phrased and the study shows the basic sentence pattern in Kangean varieties is the same as Madurese. There is a probability of sentence variation emergence based on its structure variation, and the syntactic of sentence structure is expandable by using phrase and clause.

The study above is contrast with this study because I did not analyze Kangean varieties based on those aspects, but I focused more on lexical differences. In

addition, I also studied Kangean varieties in Kangean Island based on dialectology perspective which means that I also determined the status of the lexical differences of Kangean varieties. Therefore, the status of the lexical differences of Kangean varieties was not revealed by the two studies of Sugianto et al (1982) and (1985).

Moreover, the study about Kangean varieties was conducted when there was only one sub-district, which was Arjasa Sub-district. However, in 2005, Kangayan Sub-district was legitimized as the expansion of Arjasa Sub-district (Yashik, 2005). Then, the influence of other local language, such as Bajo and Mandar language spoken by people in Sapeken and Pagerungan Island, into varieties in Kangean Island, particulary in Kangayan sub-district, is necessary to be analyzed. Furthermore, the study that discusses lexical differences in Arjasa and Kangayan sub-district of Sumenep Regency and its status have not been conducted yet, especially in English Department of Universitas Airlangga Surabaya. Thus, it is necessary to be conducted because it relates to the impact of mutually intelligibility among villages in Kangean Island.