LANGUAGE KINSHIP BETWEEN MANDARIN, HOKKIEN CHINESE AND JAPANESE (LEXICOSTATISTICS REVIEW)

KEKERABATAN ANTARA BAHASA MANDARIN, HOKKIEN DAN JEPANG (TINJAUAN LEXICOSTATISTICS)

Abdul Gapur[1], Dina Shabrina Putri Siregar[2], Mhd. Pujiono[3]

[1,2,3]Faculty of Cultural Sciences, University of Sumatera Utara
Jalan Universitas, No. 19, Medan, Sumatera Utara, Indonesia
Telephone (061) 8215956, Facsimile (061) 8215956
E-mail: mhd.pujiono@usu.ac.id

Abstract

Mandarin and Hokkien Chinese are well known having a tight kinship in a language family. Beside, Japanese also has historical relation with China in the field of language and cultural development. Japanese uses Chinese characters named kanji with certain phonemic vocabulary adjustment, which is adapted into Japanese. This phonemic adjustment of kanji is called Kango. This research discusses about the kinship of Mandarin, Hokkien Chinese in Indonesia and Japanese Kango with lexicostatistics review. The method used is quantitative with lexicostatistics technique. Quantitative method finds similar percentage of 100-200 Swadesh vocabularies. Quantitative method with lexicostatistics results in a tree diagram of the language genetics. From the lexicostatistics calculation to the lexicon level, it is found that Mandarin Chinese (MC) and Japanese Kango (JK) are two different languages, because they are in a language group (stock) (29%); (2) JK and Indonesian Hokkien Chinese (IHC) are also two different languages, because they are in a language group (stock) (24%); and (3) MC and IHC belong to the same language family (42%).

Keywords: *language kinship, Mandarin, Hokkien, Japanese*

Abstrak

Bahasa Mandarin dan Hokkien diketahui memiliki hubungan kekerabatan dalam rumpun yang sama. Disamping itu, bahasa Jepang juga memiliki keterkaitan sejarah dengan Tiongkok mengenai perkembangan bahasa dan kebudayaannya. Di dalam bahasa Jepang digunakan aksara Tiongkok yang disebut 'kanji' dengan penyesuaian bunyi kosakata yang diserap ke dalam bahasa Jepang. Kosakata dengan penyesuaian bunyi dari kanji ini disebut dengan kango. Oleh karena itu penelitian ini membahas kekerabatan bahasa Mandarin, Hokkien di Indonesia dan kango bahasa Jepang dengan tinjauan lexicostatistics. Metode yang digunakan adalah metode kuantitatif dengan teknik lexicostatistics. Dalam metode kuantitatif ini dicari persentase kesamaan dari sejumlah 200 kosakata dasar Swadesh. Metode kuantitaif dengan lexicostatistics akan menghasilkan pohon diagram kekerabatan bahasa. Dari hasil perhitungan lexicostatistics pada tataran leksikon diketahui bahwa, (1) MC dengan JK merupakan bahasa yang berbeda karena berada dalam kelompok rumpun (stok) (29 %), (2) JK dengan IHC juga merupakan bahasa yang berbeda karena berada dalam kelompok rumpun (stok) (24%), dan (3) MC dengan IHC merupakan bahasa yang se-Keluarga (family) (42%).

Kata kunci: kekerabatan, bahasa Mandarin, bahasa Hokkien, bahasa Jepang

Language Kinship Between Mandarin, Hokkian Chinese and Japanese: Lexicostatistics Review
(Abdul Gapur, Dina Shabrina Putri Siregar, Mhd. Pujiono)

Halaman 301 – 318

INTRODUCTION

For Indonesian people, Mandarin and Japanese are quite popular foreign languages to be studied. It deals with Mandarin which is being the second international language and according to UNESCO, it has been used by more than one billion people in the world (Rohmatillah, 2012).

In Indonesia, it was marked as new beginning of history called reformation or reform in May 1998, Mandarin was massively tought and learned after 30 years for being banned. As Gus Dur was served as president of Indonesia in 2000, a presidential decree was launched which associated with the right for Chinese to run all kinds of expression of Chinese culture, thus including of Mandarin language learning and its literature. This presidential decree is being a mark that Mandarin language activities, includes teaching and learning is accomodated as in formal and informal. This teaching activity is performed in some forms, such as course, private course, training, extraculicullar, an optional study or compulsory study in any Chinese department in many Indonesian's University. (Sutami, 2016) Numbers of Mandarin learner increase and it is proved by numbers of Indonesian students who study in China has reached 13.700 people since 2010, with an annual rise by 10 percent and the the majority of them studied in the department of Chinese language (Jegho, 2015).

Meanwhile, according to Japanese Embassy in Indonesia Yasuaki Tanizaki, Japanese has more than 870 thousand learners in Indonesia (Rini, 2016). It shows that Mandarin and Japanese are the most interested language in Indonesia.

Beside Mandarin, the Chinese people in Indonesia, especially in Medan, North Sumatra are the native speakers of Hokkien Chinese. It is related to the history of the Chinese migration into Indonesia. The ancestors of Chinese people in Indonesia are from China, precisely from Guangdong, Hokkien and Hainan. In Indonesia, they live and marry to local people (Christian, 2017, p. 12).

Therefore, Hokkien Chinese as a cultural practice of Chinese who have settled in Medan is a form of maintained activities even for extended years. Japanese also has a related history with Chinese. First, the former Japanese did not have their own characters, thus they adopted Chinese characters into Japanese then it was known as *kanji*. The accession of *Kanji* into Japan is predicted in the 5th century, which was brought by Chinese people who are called *Toraijin* (渡来人) through the Korean peninsula Chōsenhantō (朝鮮半島). (Kasahara, 1992, p.27). In progress of time, Japanese gradually started to modify and simplify several forms of *kanji* and then they were created as their own characters which are called *kana* (*Hiragana* and *Katakana*).

Even though the modifications to *kanji* have been created, the phonetics of Chinese are still used in Japanese with some adjustment to the Japanese's pronunciation ability. It makes some Japanese vocabularies are the result of phonetic in *kanji* which derived from China. This authentic phoneme of *kanji* is called *onyomi* (読) and the form of vocabularies uses the phonetics of *onyomi* is called *kango*.

Based on the description above, it is known that there is a hypothesis of kinship between Mandarin, Hokkien Chinese in Indonesia and Japanese *Kango*. Mandarin and Hokkien are surely from one stock languages, that is proto Sino-Tibet. (Matthew, 2002) Meanwhile, the Japanese language is reviewed by the

history is influenced mostly by the Chinese language. This research discusses about the kinship between Mandarin, Hokkien Chinese in Indonesia and Japanese *Kango* by using lexicostatistics approach.

Linguistics is traditionally distinguished for three branches namely, descriptive linguistics, historical linguistics and comparative linguistics (Antila, 1989, p. 20). The descriptive linguistics is concerned with finding and formulating the structure of a language in particular time.The synchronic term in general has similar meaning with descriptive, that includes dialectology and sociolinguistics. Synchronic is the opposite of Diachronic, while the diachronic is the synonym of historical linguistics. Comparative linguistics has two main concerns. They are (1) to determine the facts and degrees of two languages kinship or more and (2) to reconstruct the proto of language itself. in way of linguistics studies, language kinship has relationship between an effort in searching for the history of a language through the reconstructionof its proto language and the determining the fact of and the rank of its kinship. Fernandez (1996, p. 33) states that the study about kinship between languages is available through the study of historical comparative.

Comparative-historical linguistics is a branch of linguistics that studies about the language in the field of time with the change of language elements which occurs by the time (Keraf, 1996, p. 22). According to Fernandez (1996, p. 14) comparative historical linguistics is a branch of linguistics which has the main tasks, among others, to establish facts and levels of interrelatedness of kinship that is closely related to the grouping of language kinship.

Furthermore & Mbete (2009, p. 1) says that comparative historical linguistics is a branch of linguistics that studies and examines language in time dimensions, especially the past. Within this time dimension, the study is diachronic, which the language has developed and evolved through time, and it is contrast to synchronic descriptive linguistics.

This theory is developed by some experts, namely Jacob Grimm (1787--1863), Lehman (1972), Hock (1988), Bynon (1979). This theory is also called as diachronic theory, which discusses about form analysis and continuous language changing that equipped with sound changing. For examples to reconstruct a former language, such old-time language (proto) existed for thousand years ago. This old language had changed and been splited into variety of languages in case an effect of place and time changing. These derivative languages inherited the norms of its origin and will change later because of the development (innovation) occurred after knowing that those languages are different (Ino, 2015).

The kinship between groups of language in comparative study is essentially proved based on the legacy elements of the origin or proto language. The language concept of origin or proto language is actually not a grammatical form for what it looks like, but it is redesigned or reassembled as a portrait of any languae past. In other words, this concept is built to be a very simple idea for connecting the systems of language which has language kinship one to another by using particular norms. The facts of linguistics in form of regularity, comparability which are found in the languages kinship shows the evidence of shared authenticity inherited from the same ancestor.

Within the similar inherent of characteristics, the close relationship between language kinship are found and the language system is easily to be marked (Ino, 2015).

Comparative historical linguistics aims to explain the relationship between kinship and the history of languages in a particular region. The relation of the basic is abstracted in the genealogical form. Behind the relationship, implied facts of linguistics become the basic of determination and proof of kinship relationship.

These facts illustrate the historical process of the languages of the relatives in the course of time. The historical process is related to the change and forwarding of elements (static) and the structure of language.

Lexical landscape is one of the most important in the study of comparative. This looked especially on the observations of an early level in the grouping of efforts between the related language. By using evidence quantitative centered oriented on observation glance at some basic vocabulary can be determined a group of related based on prosentasenya, while the phonological landscape can be used in a higher level to determine grouping proto language Based on the change of phones which is regularly occurred in each language kinship can be formed the phonemic correspondence (Ino, 2015).

Basically an efforts of the investigation to the kinship of language for the purpose of language grouping (sub-grouping) or research origin language, is as a discovery to the conclusive proof, that was found in every language which is being compared. The evidence can be quantitative and qualitative. Quantitative evidence is in the form of a number of relative vocabulary (cognate set) pertaining to the retention altogether (shared retention). Qualitative evidence in the form of innovation together (shared innovation and the phonological correspondence (Crowley, 2010).

The historical comparative of several languages can be reached by quantitative approach apart from using a qualitative approach. The quantitative approach has been done through the procedure of grouping the language in accordance with the percentage of many calculation lexicostatistics applied by the scholars in setting the group of language kinship beside a qualitative approach (Nothofer in Fernandes, 1996, p. 23).

Lexicostatistics technique in a quantitative approach can be considered equal and sometimes distinguished by some linguists with glotochoronology. But when viewed of the stage with the use of lexicostatistics glotochronology, both has connectivity directly. Glotokronologi technique used follow lexicostatistics technique.

Lexicostatistics technique is used to determine the percentage of the kinship several languages compared. After the percentage of the kinship several languages determined new technique glotochronology followed by.The research is used just one term enclosing procedure is lexicostatistics analysis

The technique of lexicostatistics is a method of language grouping in a study of historical linguistics comparative which is already used by lot of researchers of language in the world , the most famous used by Dyen in his work in grouping the austronesian languages based on 250 famous language. Technique lexicostatistics uses statistics numbers as basic for the choice with the kinship relation of some languages in the research from the aspect of vocabulary. Fernandes (1996, p. 23). states that quantitative approach with lexicostatistics technique uses the main data such the list of swadesh (two hundreds basic vocabulary) to find out the cognitive intrument in language which is in the research. Afte the preparing list is filled by, the percentage of cognate is determined by using the understanable about the rules of systemical phonetics changing inter-languages

Nothofer in Langgole (1992, p. 53). that Lexicostatistics has several advantages compared to other methods. They are: (1) as a list of basic vocabularies which is quickly to determine the language kinship; (2) as a means of grouping languages or dialects whose language is not old enough; (3) as a tool or method that can be used at the initial stage to determine language classification. So the lexicostatistics technique with several

advantages is used to determine the level of kinship in Japanese, Chinese, and Mandarin.

Crowley (2010, pp. 191--192) confirms that the lexicostatistics technique is used to depart two basic assumptions. First, some parts of the vocabulary of a language are subjects which are more or less changed or replaced by any other vocabulary.

The main changing of vocabulary is relatively similar and applied for all languages. Second, the gap during the time of vocabulary changes is little bit similar and occured for all languages.

For this assumption, it has found that the main vocabulary exists at least for 80.5 percent in a thousand years or it means that the main vocabulary of a particular language will be replaced by 19.5 percent of new vocabulary in a thousand years.

Keraf (1996) explains the way of lexicostatistics technique works based on several principles. First, collect a number of words from basic vocabulary. Secondly, determine the pairs of basic vocabulary which are close. In determining the basic cognate of vocabulary. It is viewed based on the rules of phonemic change.

In addition, in Historical Comparative Linguistic, phonemics change is also reviewed based on the regularity of changes or not (Fernandes, 1996, p. 68). Phonetics changes which is regularly occured also be sorted in substitutions, splits, mergers and zero change.

The substitution of phonetic changes is a phonemic change in form of replacing a phoneme with another phoneme. Phonemic changes in the form of splits are phonemic changes with several phonemes or separation. Changes in the form of mergers are changes in some phonemes onto a phoneme which is also called a phonemic composite. Zero change is a none change of phonem or phonemic dismiss.

Change in phonemes is also distinguished from conditional phonemic changes and unconditional phonemic changes. Conditional phonemic changes occur when the phonemic changes do not occur in all positions (initial word position, word center position, word end position), but only occurs in certain position. Unconditional phonemic changes occurs when the phonemic changes occurs in the overall position, which occurs at the initial position of the word, the middle position of the word, from the final position of the word.

Considering the level of the comparative historic linguistic scope, Keraf (1996, p. 23) points out that the purpose and importance of comparative historical linguistics are namely (1) questioning the cognate languages by comparing the elements that show their kinship. The area used to make the comparison are phonology and morphology. The temptation in making the comparisons in the field of syntax gets unsatisfactory results. Two (2) establish the reconstruction of existing languages to ancient languages (proto languages) or languages that derive contemporary languages. In other words, comparative historical linguistics tries to find the proto language in modern languages nowadays.

METHOD

The method used in the research is the quantitative method. Quantitative method in comparative historical linguistics uses lexicostatistics technique. Quantitative research is research that emphasizes numerical data (numbers), which is processed by statistical methods (Azwar, 2007, p. 5). According to Subana and Sudrajat (2005, p. 25), quantitative research is viewed in terms of objectives, this study is used to reviewed a theory, presents a fact or describe statistics, and to show the relationships between variables and those that developing concepts, understanding or describe things.

Meanwhile the quantitative method with lexicostatistics technique is used to determine

the kinship of language in purpose to gain a brief description about the level of historical relationship and the connection based on the kinship and its instrument in form of 100--200 Swadesh basic vocabularies. The quantitative method looks for the percentage from 200 Swadesh basic vocabularies. Quantitative method with lexicostatistics will result in a tree diagram of language kinship.

Data was collected by using the refer method, recording technique and note-taking technique as the optional advanced techniques. Thus, the informant was asked to say their own regional language vocabulary for each basic vocabulary contained in the Swadesh list.

The recording technique was used at the time the informant speaks the vocabulary. After recording is finished, then followed by the note-taking technique, so that the research data in the form of two hundred basic vocabulary from the three languages studied is able to be collected.

Note-taking is used by phonemic transcription according to the purpose of the study, which is to found the level of kinship of three languages. in accordance with Sudaryanto's view (2015, p. 133-136). For the validity of the data from the informant, the data is consulted from the dictionary.

Lexicostatistics is a technique in grouping the languages. It tends to give priority to look up the words statistically. Then it tries to determine the grouping based on the percentage of the similarity and the difference of one language to another.

Lexicostatistics were aligning the language or dialect which prioritizes calculation to know the number of words in statistics in common words relatives rates compared (Grimes, 1987 in Ino, 2015). Lexicostatistics contains the list of vocabulary which is compared in languages. Morris swadesh proposed 200 a vocabulary of the base that universal, which includes the words pronoun, said the number of words the

limbs (the nature and their activities), nature and surrounding areas and daily cultural activity.

Lexicostatistics as one technique grouping language has been widely used by the scientists / linguist in this world. This technique uses statistic of numbers as the basic of its analysis. The technique is able to find all the two languages or more by taking into account the equation in vocabulary. According to Nothofer (1992), lexicostatistics techniques have some excellences when it is compared with other methods. Certain excellences that referred to include (1) as vocabulary list which is quick to find and determine in certain single language (language kinship); (2) as a tool for grouping the kinship of language /dialect whch its proto-language is nott ancient or quite old times; and (3) as a method of classification which can be used on a preliminary step to determine classifications of language.

Furthermore, Nothofer (1992), sets out three basic assumptions of the lexicostatistics method, namely (1) basic vocabulary is replaced with the same gap in all languages at the same time. According to this assumption that every 1000 years is around 18-20% of the basic vocabulary that changes and applies to all languages simultaneously, (2) all the basic vocabulary contained in the basic vocabulary list is likely to change the words simultaneously, and (3) there is a particular basic vocabulary which is considered to be valid in every language in the world. The way lexicostatistics works as follows the pattern proposed by Keraf (1996), namely, (1) collects number of words from the basic vocabulary, and (2) determines the basic vocabulary which match are equal. Furthermore, efforts to determine the basic basic vocabulary that follows the steps proposed by Keraf (1996), namely, (1) looking for vocabulary that is not from a language/ loan word, (2) experiencing a single/ free morpheme by isolating all bound morphemes, and (3) comparing all pairs of words to determine a

pair of related words based on recurrence, co-occurence, and analogy.

An effort to determine the same pair of words is done by identifying all pairs of words that are similar, pairs of words with similar phonetics, pairs of words corresponding to the phone, and pairs of words which have only one phoneme difference in one phoneme. Step of implementation in lexicostatistics technique.

1) To collect the basic vocabularies of cognate language which are being researched. This technique is the first step in comparing the cognate languages. List of vocabularies used in the research consists of 200 Swadesh vocabularies.

2) To determine the cognate words.

3) To calculate the percentage of kinship by defining and counting the same and similar pairs of cognate words based on the following formulation.

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

k = number of cognate words
n = number of compared words

Further, the relation of kinship status among the languages is done based on the percentage in the table below.

Tabel 1 Classification of Kinship Percentage

Classification of kinship percentage	
Dialect of language	81-100%
Language of familly	36-81%
Families of stock	12-36%
Stock of microphilum	04-12%
Microphyla of esophyulum	01-4%
Mesophyla of acrophylum	00-1%

(Source: Crowley, 2010)

After obtained the percentage of the cognate of the three languages, the next step

is the way to prove the different languages and dialect. Using this method , quantitative evidence in the form of their resemblances and in common vocabulary of the same origin in every language compared can be counted .The calculation of the number of the percentage based on the number of pairs of vocabulary of same origin a sum divided gloss are filled times a hundred percent.

RESULT AND DISCUSSION
Mandarin Language

Word 'Mandarin' in English is taken from Portuguse (mandarim), which is from Malay. Another source states that Mandarin is literally from a designation from foreign to the elites of Dinasty Qing. Dinasty Qing is established by Manchu. This dynasty usually called the elites of the empire as *Mandaren* (*man da ren* /滿 大 人), which means the elites of Manchu. Hence, the language use for the Manchu's authorities at that time was also called 'Mandaren language'. Then, its transliteration has changed further into "Mandarin'.

There are other names of Mandarin, (1) *Guoyu*, is a term for the north dialect of Han language, which is known as Mandarin. *Guoyu* literally means 'national language'. In fact that Mandarin has been set as a governmental legal and the national language in some states, such as People's Republic of China and Republic of China in Taiwan. Two (2) *Huayu* is a term for the north dialect of Han, which is known as Mandarin nowadays. Huayu literally means "Hua language", which is the common language used by the Chinese, which shows Mandarin is spoken broadly.

Mandarin uses specific characters called *Hanzi*. Once, all Chinese literatures use traditional characters. In 1949, China gained freedom, thus the authorities consider the traditional characters are quite difficult for those people who interest to learn Mandarin. From that moment, the Chinese authority assembled

all best teachers in the entire state to change the complicated characters into the easier ones, thus the simplified writing appeared If the traditional writing had 15 lines, then the simplified writing only has 5 lines and it seems efficient.

In 1956, the simplified writing and its spelling were inaugurated in a congress, thus it is used and applied nowadays. Singapore is the first country outside China which uses the writing up to now. Taiwan, as the main holder of traditional characters has the authority since China switched the writing in 1956. Nowadays, Taiwan uses traditional characters as their daily note. The majority people of Taiwan have less understanding about traditional characters. They are feeling awkward for abbreviating the simplified characters because of the hazzy history of the characters. Traditional characters are created by people in thousand years ago. Traditional and simplified characters have similar tone and meaning. One difference is only on the writing system, even not much. Mostly they have the same form.

A Chinese word has many similar pronunciations with different meanings. The difference of these homophones is caused by the tonal quality for rising and falling tone in particular morphemes.

Mandarin has four tones, namely high, rising, falling-rising and falling, and the fifth additional tone is the neutral tone. For illustration, a word *ma* has four different meanings based on its tones.

High	ma	'mother'
Rising	má	'hemp' or 'numb'
Falling-rising	ma	'horse'
Falling	mà	'to scold or 'swear'

(Chinese Basic Lessons for Invitees, 2007).

Hokkien Language

Hokkien or Hokkien is a branch of South Min language, which is a part of Han language. This language is broadly used in the Fujian Province (Hokkien), Taiwan, North side (Kengtang) and Southeast Asia where all the groups of Chinese immigrants are majority from the Fujian Province. Hokkien is also known as Holo language in China and Taiwan.

Their ancestors had migrated to Indonesia after living for years in qing dynasty (between 1600-an until 1900). The remaining people which ha migratied early already been assimilated with the locals, this is mainly because of an ancestor of the Chinese in the old times did not usually bring their family (women). Therefore, if they would not back to their homes, they usually who had been settled in new living area would marry, especially with the local women.

In Medan, they who speak hokkien are mostly are Ciangchiu, this is also similar with the hokkien in Penang, Malaysia. The hokkien dialect in Medan is quite similar with those which is pronounced in Penang. But he hokkien on the east coast sumatra, for example Bagan siapi-api, mostly come from coanchiu. The dialect they speak are similar dialect with hokkien in taiwan

Hokkien ethnic with its language is the majority of Chinese population in Indonesia. Numbers of Hokkien Chinese speakers are estimated for about 50 million all around the world. Hokkien people are concentrated in several areas such North Sumatra, Riau (Pekan Baru), West Sumatra (Padang), Jambi, South Sumatra (Palembang), Bengkulu, Jawa, Bali, Kalimantan (Banjarmasin, Kutai), Sulawesi (Makassar, Kendari, Manado) and Ambon (www.tionghoa. info.com).

Japanese Language

Japanese or *Nihongo* is the national language of Japan and it is used by people in daily. Japanese has a tight kinship of history with its surrounding states. One of those states is China, which has given the biggest influence to the Japanese language and culture in past. Therefore, Japanese *kanji* is likely similar to

Chinese *hanzi*, but both have distinctive tones. Japanese *kanji* is obviously the Han characters which were brought to Japan archipelago. This characters are then borrowed as one of written varieties (Guānghuá, 2002, p. 87).

This is related to the development of Japanese which gained much influence by the Buddhist religion. Conceded before the Buddhist religion, Japanese used to communicate with them by using sign language or spoken.

According to a representative from kingdom of China to Japan, it was gained report that at the end of a century 6th, they had no written letters and learned only by the use of notched sticks and knotted ropes'. Japanese develop into the systematic writing only in the 8th through bringing together all writen texts and books which writen by popular Buddhist monks.

This situation helped the development of the language use and Chinese literature in systemic way. For those Sinto, which is the origin religion of Japanese has no writen text or literature, there were no experts , holy books and doctrins in doing the pray to the Gods (*kami*), so the descenting religion followers are not from books, or such reccords but only through the memory of the parent's that been given to the new generation. An intimate service ahead of exultation and the worship is run by Shinto's holy man by read the spells who was memorized. It is quite different with Buddha, Sinto is a kind of worship for the nature and it is not a systemic religion. It puts only to worship for nature and the formers and it has many *kami* (God).

Kami or God of Sun (*Amateresu Oomikami*) is concsidered as the descendant from Japanese emperor. before the Buddhist's coming, the interest about literature is not extremely needed for Japanese. Almost in 1100 years in the beginning of Jimmu emperor, Japanese had put the emperor so hight in position, thus the emperor with support from his people

are able to conquer half of Korean area. This conquering has an impact for the entrance of Buddhist into Japan .At that time, language system and sytematic literature and also the influence of Chinese culture came into Japan through the Korean. The influence of religion and the cultural also has been rooted for long time in Korea. In 552, *Raa* in area of *Pakche* (one of Korean kingdom) was surrender and gave a holy book which was known as Buddhist doctrine, the arca of Buddha, some waves and umbrellas as the gifts for Japanese emperor. Two years later, the learners and Buddhist monks who had the knowledge about classical literature of China, science and music arrived in Japan. Even there was a contra for this Buddhist entrance. the silk chinese (book), bronze statue of Buddhist , colorful clothing ceremony , the way of speaking the pray,tinkling and cremation has attracted the interest of the Japanese to Buddhism".

Massive among Japanes have intention to learn about this newest religion. Regarding to the monks which are from China only speak Chinese. So thet they who are being the new followers should learn Chinese and its sinogram (kanji) at the first (Bin Jambi, 2003).

For this reason, the former Japanese people took some parts of *Hanzi* and reformed them into new forms of Japanese characters. Then, it became the last three of the Japanese characters, namely *hiragana, katakana* and *kanji*. Beside these three forms, Japan created an easiness for beginners who know less about *kana* and *kanji*. Japan writes its language into phonetic transcription which is known as *romaji*. Thus, the Japanese writing itself is formed by four types of letter.

Japanese phonetic vocabulary retains the original phonetic of *kanji* named *Kango*. Tanimitsu (1995, p. 62--63) states that *Kango* is initially conveyed from China. The Japanese people used them as their own language although it is unclear when it happened. It is known that

kango had been used in Nara era, and in the era of Heian. *kango* have been used in literary works such as *monogatari* 'story'. Altogether, within the developing era, *kango* is broadly used and now almost half words collected in dictionary of *Kokugo Jiten* are *Kango*. Thus, *kango* is the words which roots deeply into the life of Japanese through extended periods.

Based on the history, *kango* does not seem different from *garaigo* (loan-word) because they are from foreign languages. However, *kango* has a particular characteristic which is quite different from *garaigo*. Therefore, *kango* is special types of vocabulary in Japanese. Ishida (in Dahidi & Sudjianto, 2012, p. 87) explains the characteristics of *kango* as follows:

a) *Kango* are words which is read by way of *on'yomi* (Chinese pronunciation) which consists of one *kanji* or a combination of two or more *kanji*, e.g : 文 (*bun*)、本 (*hon*)、金 (*kin*)、茶 (*cha*), etc.

b) Because in way of reading *onyomi* consists *go'on* (dinasty of Wu's pronounciation), kan'on (dinasty of Han's pronounciation), too'on (dinasty of Tang's prononciation), there are variety ways of reading. For example: *Gakki* ➜ 学期、楽器、月忌. Saigo ➜最期、最後.

c) The first letter is mostly used syllable *dakuon* (voiced sound), but there are none used syllable *handakuon* (semivoiced) such as ぱ(pa)、ぴ(pi)、ぷ(pu). For example学部 (*gakubu*)、言語 (*gengo*), 言語 (*gengo*) etc.

d) There are lots of phonemes *yoo'on* (diphthong) and *choo'on* (long vowel). Example of *yoo'on* : きゅ(*kyu*)、きょ(*kyo*)、しゃ(*sya*)、しゅ(*syu*)、しょ(*syo*) etc. Example of *choo'on*: ゆうびん (*yuubin*)、ねえさん (*neesan*)、おとおさん(*otoosan*).

e) Long words can be constructed by combining several *kango*. For example: 対共産圏輸出統制委員会則違反事件 (*Taikyōsan-ken wa-de tōsei iinkai-soku ihan jiken)*. Violation case of Community Sphere Risk Control Committee Committee.

f) There are numbers of nouns, especially for human activity concrete & abstract nouns. For example: 勉強 (*benkyou/*study), 運動 (*undou/*sport料理(*ryouri/*cuisine), 生活(*seikatsu/*Life),愛(*ai/*love),正義 (*seigi/* justice).

Cognate word calculations between Mandarin, Hokkien Chinese and Japanese
The definition of cognate words is narrowly, and most ussually, one of two or more words or morphemes which are directly descended from a single ancestor for, in the single common ancestor of language in which the words or morphemes are found, with no borrowing. The cognate words has function to identify the grouping or sub-group of comparative language. According to Bellwood (2000, p. 147), a defimition of cognate is one or more words which have similarity in meaning and phone has been derived from the origin language to one or more derivative language, and it is not kind of a borrowing words from any particular language outside. The procedure in determining the cognate or noncognate words by comparing phoneme with phoneme and phoneme with cluster of phoneme (Gudschinsky in Panggabean, 2004).

After knowing the comparable pairs of phoneme-phoneme or phoneme-cluster phoneme pairs, the criteria for the determination of related pairs are as follows:

1) The couples of phonem are identically similar.
2) The couples of phonem have similarity in comparing the phonetics.
3) The couples of phonems are totally differen in case of environmental effect.
4) The couples of phonmes found continuously in another matching couples which is in comparative position.

Two words compared are only named cognate if there are at least three couples of phonem–phonem, phonem– cluster phonem, or cluster phonem– cluster phonem found as cognate. If the compared words have less than three phonems, then all the couples which is compared should be cognate.

The calculation of kinship percentage between the Mandarin, Indonesian Hokkien and Japanese *Kango* with lexicostatistics technique is conducted by applying procedures as follows:

(a) To decide the uncounted gloss (unrelated words or loanwords).
(b) To isolate bound morphemes.
(c) To determine the cognate words.

Finding the cognate words in these three languages is started from characters with the similar meaning. That step is crucial because there are several characters of *kanji* which have similar character with different meaning in Mandarin or Hokkien characters, for example 「聞」. In Japanese 聞 (*shin*) means 'listen' but in mandarin聞 (*wén*) means 'smell'. Thus, the determination of the cognate vocabularies follows the conditions as follows:

(1) Identical pair of words with all similar pronunciation.

Gloss	Manadarin	Hokkien	Jepang	
Hair	*Máo*	*Mo*	*mou*	毛

(2) Phonetic correspondence
If the phonemic change between the two languages is occured reciprocally and orderly, and its frequency is high. Therefore, the equal form between the two languages is considered as a cognate.

Gloss	Mandarin	Hokkien	Jepang	
Water	*Shuǐ*	*Cui*	*sui*	水

(3) Phonetic similarity
If there are similarities in articulatory positttion, herefore the pair of words may be considered as the cognate words.

Gloss	Mandarin	Hokkien	Jepang	
Fire	*Huǒ*	*hoi*	*Ho*	火

(4) One different phoneme
If a pair of words have a one different pronunciation but the diffrence can be explained by the environmental influences, while in another language the influences do not change its phoneme. Therefore, that pair is determined as cognate words.

Gloss	Mandarin	Hokkien	Jepang	
Bunga	*Huā*	*hua*	*Ka*	花

In order to avoid errors in statistic is to give an assumption that an event of language contact happens not in certain period, but during the particular per iods which is quite long. During this language contacts, it occurs an accumulation of differences between two languages, by the day it is more develops , and slowly but sure it becomes a mark of separation between the two languages. To calculate the errors is used a standard errors that is 70% from the correctness which is estimated.

Data soure of the research is spoken language which is collected by using list of 200 Swadesh vocabularies. The data consist 200 vocabularies and thus it is rechecked by looking up the dictionary (Mandarin and Japanese *Kango*).

The result of lexicostatistics percentage is calculated from the amount of word matching in 200 pairs of Swadesh vocabularies from

Language Kinship Between Mandarin, Hokkian Chinese and Japanese: Lexicostatistics Review
(Abdul Gapur, Dina Shabrina Putri Siregar, Mhd. Pujiono)

Halaman 301 − 318

Mandarin, Indonesian Hokkien Chinese and Japanese *Kango* shows the percentage as follow:

Hokkien		
42%	Mandarin	
24%	29%	Japanese *Kango*

The obtained comparisons are:

Hokkian ⟶ Mandarin ⟶

$$c = \frac{85}{200} \, x \, 100\% \ = 42\%$$

Mandarin ⟶ Japanese ⟶

$$c = \frac{59}{200} \, x \, 100\% \ = 29\%$$

Hokkian ⟶ Japanese ⟶

$$c = \frac{48}{200} \, x \, 100\% \ = 24\%$$

Based on the result, it can be stated that the language kinship between Hokkien and Mandarin is closer each other than the Japanese. It can be illustrated by the tree diagram as follow:
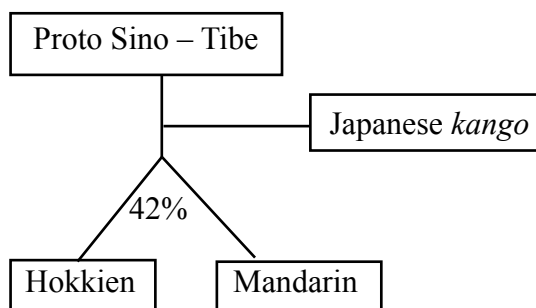
CONCLUSION

Mandarin, Hokkien in Indonesia and Japanese *Kango* which have been reviewed by lexicostatistics method have a relation of language kinship. This kinship is possible with the percentage of cognate words between Mandarin and Hokkien are 42%, between Mandarin and Japanese *Kango* are 29%, and between Hokkien and Japanese *Kango* are 24%. From this percentage, it can be concluded that Mandarin and Hokkien in Indonesia belong to the same Language Family. Meanwhile, the kinhip relation between Mandarin and Hokkien

is that both belong to the same Families of Stock with Japanese *Kango*.

REFERENCES

Antila, R. (1989). *Historical and Comparative Linguistics*. Amsterdam: John Benjamins Publisshing Company.

Azwar, S. 2007. *Metode Penelitian*. Pustaka Pelajar: Yogyakarta.

Bellwood, P. (2000). *Prasejarah Kepulauan IndoMalaysia*. Jakarta: PT. Gramedia Pustaka Utama.

Christian, S. A. (2017). Identitas budaya orang Tionghoa Indonesia. *Jurnal Cakrawala Mandarin*, *1*(1), 11–22.

Crowley, T. & Claire B. (2010). *An Introducing to Historical Linguistic: Fourth Edition*. New York: Oxford University Press.

Dahidi, A. & Sudjianto. (2012). *Pengantar Linguistik Bahasa Jepang* (p. 103). Jakarta: Kesaint Blanc.

Fernandez, I.Y. (1996). *Relasi Historis Kekerabatan Bahasa Flores Kajian Linguistik Historis Komparatif terhadap sembilan bahasa di Flores* (p. 14). Ende: Penerbit Nusa Indah.

Guānghuá, X (许光华). (2002).中国文化概要 (p. 87). 上海：汉语大词典出版社.

Ino, L. (2015). "Pemanfaatan Linguistik Historis Komparataif Dalam Pemetaan Bahasa-Bahasa Nusantara". *RETORIKA: Jurnal Ilmu Bahasa*, 1(2), 365. doi:10.22225/jr.1.2.41.365-378

Jegho, L. (2015, May). *Indonesians Studying in China Rising in Number. Global Indonesian Voice*. From http://www.

globalindonesianvoices.com/20935/
indonesians-studying-in-china-rising-
in-number/ (Accessed on October 15,
2018)

Kasahara (笠原). (1992). 許説日本史 (p. 27).
東京：講談社.

Keraf, G. (1996). *Linguistik Bandingan Historis*
(p. 22-33). Jakarta: Gramedia Utama.

Matthew S.D. (2002). *Word Order in Sino-
Tibetan Languages from a Typological
And Geographical Perspective
from:* The Sino-Tibetan Languages
Routledge. Accessed on 05 Jun 2018
https://www.routledgehandbooks.com/
doi/10.4324/9780203221051.ch3

Mbete, A.M. (2009). *Sekilas tentang Linguistik
Historis Komparatif* (p. 1). Denpasar:
Universitas Udayana.

Nothofer, B. (1992). Tinjauan Sinkronis dan
Diakronis Dialek-Dialek Bahasa Jawa di
Jawa Barat dan di Jawa Tengah (Bagian
Barat). *Kumpulan Bahan Penataran dan
Seminar Lokakarya yang diselenggarakan
Pusta Studi Bahasa Bahasa Asia Tenggara
Pasifik Fakultas Sastra Universitas Gadjah
Mada*, 27 Februari-11 Maret 1992.

Panggabean, H. (1994). *Telaah Bahasa-
Bahasa Batak Dari Segi Lexicostatistics*.
Bandung: Program Pascasarjana Universitas
Padjadjaran. Bandung.

Panggabean, H. (1994). *Telaah Bahasa-
Bahasa Batak Dari Segi Leksikostatistik*.
Bandung: Program Pascasarjana Universitas
Padjadjaran. Bandung.

.L. (2016). *Jumlah Orang Indonesia
Belajar Bahasa Jepang Terbanyak Kedua
Dunia*. From https://www.republika.co.id/
berita/pendidikan/eduaction/16/01/27/
o1m85x299-jumlah-orang-indonesia-
belajar-bahasa-jepang-terbanyak-kedua-
dunia (Accessed on October 15, 2018).

Subana, M & Sudrajat. (2005). *Dasar-Dasar
Penelitian Ilmiah*. Bandung: Pustaka.

Sudaryanto. (2015). *Metode dan Aneka Teknik
Analisis Bahasa*. Yogyakarta: Duta wacana
University Press.

Suku Tionghoa Indonesia. (n.d). tionghoa.info
http://www.tionghoa.info/suku-tionghoa-
di-indonesia/

Sutami, H. (2016). Fungsi dan Kedudukan
Bahasa Mandarin di Indonesia. *Paradigma,
Jurnal Kajian Budaya*, 2(2), 212.
doi:10.17510/paradigma.v2i2.28

Tanimitsu, T. (1995). *Nihongo gaku* (p. 62-
63). Tokyo: Sakai Shoten.

Tanimitsu, T. (2007). *Chinese Basic Lessons
for Invitees*. Peace Corps: China.

Language Kinship Between Mandarin, Hokkian Chinese and Japanese: Lexicostatistics Review
(Abdul Gapur, Dina Shabrina Putri Siregar, Mhd. Pujiono)

Halaman 301 − 318

ATTACHMENT
Swadesh List Data with Cabinet-ordered Romanization System
(*Kunrei-shiki Rōmaji*)

NO	Gloss	Hokien Chinese	Mandarin	Letter	Japanese *Kango*	Letter	HCI	MC	JK
	ash	huei	huī	灰	hai	灰	A	A	A
	water	cui	shuǐ	水	sui	水	A	A	A
	root	kin	gēn	根	gen	根	A	B	B
	flow	lau	liú	流	ryuu	流	A	A	B
	child	kia	háizi	孩子	ko	子	A	B	C
	wind	hong	fēng	风	fuu	風	A	B	C
	dog	kau	gǒu	狗	kou	狗	A	B	A
	what	hamik	shénme	什么	nani	何	A	B	C
	fire	hoi	huǒ	火	ho	火	A	A	A
	floating	pui	piào	漂	hyou	漂	A	A	A
	smoke	ien	yān	烟	en	烟	A	A	A
	cloud	hun	yún	云	un	雲	A	A	A
	father	papa	bàba	爸爸	fu	父	A	A	B
	how	ancua	zěnme yàng	怎么样	dou/ikaga		A	B	B
	good	ho	hǎo	好	ryou	良	A	A	B
	burn	sio	shāo	烧	shou	焼	A	A	A
	back	peng	hòu	后	ko/gou	後	A	B	B
	many	ce	duō	多	ta	多	A	B	C
	lie down	to	fǔwò	俯卧	yokoni naru		A	B	C
	new	sin	xīn	新	shin	新	A	A	A
	wet	tam	shī	湿	shitsu	湿	A	B	C
	stone	ciok	shítou	石头	seki	石	A	B	C
	some	kuice	yīxiē	一些	ikutsu ka		A	B	C
	split (me)	chiet	qiè	切	setsu	切	A	A	B
	correct	tiok	duì	对	sei	正	A	A	B
	swollen	ceng	zhǒng	肿	shou	腫	A	A	B
	seed	ceng	zhǒng	种	shushi	種子	A	A	B
	weight	tang	zhòng	重	chou	重	A	A	A
	swim	iu cui	yóuyǒng	游泳	yuuei	游泳	A	B	C
	give	ho	gěi	给	ageru	挙げる	A	B	C
	walk	kia lo	zǒulù	走路	yuuho	遊歩	A	B	C
	big	tua	dà	大	tai	大	A	A	A
	when	na si	shí	时	ji	時	A	A	A
	animals	khim siu	dòngwù	动物	doubutsu	動物	A	B	C
	star	ciok	xīng	星	sei	星	A	B	C
	fruit	cui ko	shuǐguǒ	水果	ka	果	A	A	B
	month	guek nio	yuè	月	gatsu/getsu	月	A	B	C
	fur	mo	máo	毛	mou	毛	A	A	A
	flower	hua	huā	花	ka/ke	花	A	A	A

kill it	phak si	shā sǐ	杀死	satsu	殺	A	A	B
hunt	...	shòuliè	狩猎	shuryō	狩猟	A	B	C
bad	phai kua	huài	坏	aku	悪	A	B	C
bird	ciao	niǎo	鸟	chou	鳥	A	A	A
rotten	tho thang	zāo	糟	fu	腐	A	B	C
worm	cim	chóngzi	虫子	mushigera		A	B	C
kiss	se	qīn	亲	kuchidzuke		A	B	C
wash	bak	xǐ	洗	sen	洗	A	B	C
meat	kak	ròu	肉	niku	肉	A	B	C
and	...	hé	和	soshite		A	B	C
lake	hoek	hú	湖	ko	湖	A	A	A
blood	lai	xuè	血	ketsu	血	A	B	C
come	hiok	lái	来	rai	来	A	B	C
leaf	thau hun	yèzi	叶子	ha	葉	A	B	C
dust	kin	huīchén	灰尘	ai	埃	A	B	C
close	kak	jìn	近	kin	近	A	B	B
with	kak	gēn	跟			A	B	C
hear	thia	tīng	听	bun	聞	A	A	B
in the	khi lai bin	lǐ	里	naka	中	A	B	C
where	khi ta lok	nǎlǐ	哪里	doko	何处	A	B	C
here	khi koi lai	zhèlǐ	这里	koko	此处	A	B	C
there	khi ho peng	nàlǐ	那里	soko	其处	A	B	C
on, at	khi	zài	在			A	B	C
cold	leng	lěng	冷	rei	冷	A	A	A
self	khia	zìjǐ	自己	jiko	自己	A	B	C
push	tua	tuī	推	sui	推	A	A	A
two	ji	èr	二	ni	二	A	B	C
sit	ce	zuò	座	za	座	A	A	B
tail	boi	wěibā	尾巴	bi	尾	A	B	C
four	si	sì	四	shi	四	A	A	A
you	lu	nǐ	你	anata	彼方	A	B	C
dig	o	wā	挖	kutsu	掘	A	A	B
salt	iam	yán	盐	en	塩	A	A	A
scratch	pe	zhuā	抓	sou	抓	A	B	C
fat	pui	pàng	胖	shibo	脂肪	A	A	B
tooth	ge	yáchǐ	牙齿	shi	歯	A	B	C
bite	ka	hé	龁	kou	噛	A	B	A
scrub	bua	shuā	刷	satsu	刷	A	A	B
mountain	sua	shān	山	san	山	A	A	A
hell	phak	dǎ	打	da	打	A	B	B
delete	chat	shānchú	删除	sakujo	削除	A	B	C
heart	sim	xīn	心	shin	心	A	A	A
nose	pi khang	bízi	鼻子	bi	鼻	A	B	B

Language Kinship Between Mandarin, Hokkian Chinese and Japanese: Lexicostatistics Review
(Abdul Gapur, Dina Shabrina Putri Siregar, Mhd. Pujiono)

Halaman 301 − 318

life	uak	shēnghuó	生活	seikatsu	生活	A	B	C
green	che	lǜ	绿	ryoku	绿	A	B	C
suction	suk	xīlì	吸力	kyuuin	吸引	A	B	C
black	o	hēi	黑	koku	黑	A	B	C
count	sui	shù	数	suu	数	A	A	A
rain	ho	yǔ	雨	u	雨	A	B	B
forest	pa	sēnlín	森林	shinrin	森林	A	B	C
he	i	tā/tā	它／他	hi	彼	A	B	C
mother	mama	māmā	妈妈	bo	母	A	A	B
fish	hu	yú	鱼	gyo	鱼	A	A	B
tie	pak	bǎzi	把子	baku	縛	A	B	C
this	ce ie	zhè	这	kore	之	A	A	B
wife	bo	qīzi/ lǎopó	妻子／老婆	sai	妻	A	B	C
that	he le	nà	那	kore	其れ	A	B	C
sewing	chia	féngrèn	缝纫	housei	縫製	A	B	C
street	kia lo	mǎlù	马路	douro	道路	A	A	B
heart	sim	xīn	心	shizo	心臓	A	A	B
fall down	puak to	pūdǎo	扑倒	rakka	落下	A	A	B
far	heng	yuǎn	远	en	遠	A	B	C
fog	ien	wù	雾	mu	霧	A	B	C
feet	kha	jiǎo	脚	zoku	足	A	B	C
if	na si	rúguǒ	如果	baai		A	B	C
we us	wa lang	wǒmen	我们	watashitachi	私達	A	A	A
you	lu	nǐ	你	anata	貴方	A	B	C
right	cia peng	yòu	右	yuu	右	A	B	B
because	ing hui	yīnwèi	因为	楢		A	A	B
said	ua	huà	话	wa	話	A	A	A
small	se	xiǎo	小	shou	小	A	A	A
fight	sio phak	dòuzhēng	斗争	kenka	喧嘩	A	B	C
head	thao	tóu	头	tou	頭	A	A	A
dry	ta	gàn	干	kan	干	A	A	A
left	to peng	zuǒ	左	sa	左	A	B	C
dirty	la sam	zàng	脏	o	汚	A	B	C
nail	ceng kak	zhǐjiǎ	指甲	sou	爪	A	B	C
skin	phoi	pí	皮	hi	皮	A	B	B
yellow	eng	huáng	黄	kou	黄	A	A	B
fleas	sat	shī	虱	shitsu	虱	A	A	A
other	pakkai	bié de	别的	betsu	別	A	A	B
sky	thi	tiān	天	ten	天	A	A	A
sea	hai	hǎi	海	kai	海	A	A	A
wide	khuak	kuān	宽	fuku	幅	A	A	B
neck	am kun	bózi	脖子	shuu	首	A	B	C
man	ta po	nán	男	nan	男	A	B	B

throw	tim	diū	丢	tou	投	A	A	B
slippery	kut	huá	滑	kotsu	滑	A	B	C
tongue	cik	shétou	舌头	setsu	舌	A	B	B
see	khua	kàn	看	kan	看	A	A	A
five	go	wǔ	五	go	五	A	B	C
saliva	nua	kǒushuǐ	口水	ta	唾	A	B	C
straight	tit	zhí de	直的	choku	直	A	B	C
knee	khau	xīgài	膝盖	shitsu	膝	A	B	C
play	thik tho	wán	玩	gan	玩	A	B	B
eat	ciak	chī	吃	shoku	食	A	A	B
night	ame	wǎnshàng	晚上	ban	晚	A	B	B
eye	bak ciu	yǎnjīng	眼睛	gan	眼	A	B	C
sun	jit thau	tàiyáng	太阳	taiyou	太陽	A	B	B
die	si	sǐ	死	shi	死	A	A	A
red	ang	hóng	红	kou	红	A	A	B
they	e lang	tāmen	他们	anatatachi	貴方達	A	B	C
drink	lim	hē	喝	in	飲	A	B	C
mouth	chui	zuǐ	嘴	kou	口	A	A	B
gag	au	ǒu	呕	ou	嘔	A	A	B
name	mia	míngzì	名字	myou	名	A	B	C
breath	khui	hūxī	呼吸	kokyuu	呼吸	A	B	C
sing	chio kua	chànggē	唱歌	shouka	唱歌	A	B	C
person	lang	rén	人	nin/jin	人	A	B	C
hot	juak	rè	热	netsu	熱	A	B	C
long	teng	zhǎng	长	chou	長	A	B	C
sand	sua	shā	沙	sa	沙	A	A	A
hold it	gim	názhe	拿着	kaku	捆	A	B	C
short	ke	duǎn	短	tan	短	A	B	B
squeeze	te	jǐ	挤	kou	絞る	A	B	C
women	cabo	nǚ	女	nyo	女	A	B	B
stomach	pakto	dùzi	肚子	fuku	腹	A	B	C
think	sio	sī	思	shi	思	A	A	A
tree	chiu cang	shù	树	moku	木	A	B	C
cut	chiet	qiè	切	setsu	切	A	A	B
back	pa ciak	bèi	背	hai	背	A	B	C
navel	to cai	dùqí	肚脐	sei	臍	A	A	B
white	pek	bái	白	haku	白	A	A	B
hair	thamo	tóufǎ	头发	hatsu	髮	A	B	C
grass	chao	cǎo	草	sou	草	A	A	A
one	it	yī	一	ichi	一	A	A	A
i	wa lang	wǒ	我	watashi	私	A	A	A
wing	sit	chìbǎng	翅膀	yoku	翼	A	B	C
a little	tampok	shǎo	少	shou	少	A	B	B

Language Kinship Between Mandarin, Hokkian Chinese and Japanese: Lexicostatistics Review
(Abdul Gapur, Dina Shabrina Putri Siregar, Mhd. Pujiono)

Halaman 301 − 318

	narrow	ek	xiázhǎi	狭窄	kyouai	狭隘	A	B	C
	all	kaliao	dōu	都	zen	全	A	B	C
	noon	e po	zhōngwǔ	中午	chuu	昼	A	B	C
	who	hamang	shéi	谁	sui	誰	A	B	B
	husband	ang	zhàngfū	丈夫	fuu	夫	A	A	A
	river	kang	hé	河	ka	河	A	B	C
	know	cai	zhīdào	知道	chi	知	A	B	B
	year	ni	nián	年	nen	年	A	A	A
	sharp	ciam	jiān	尖	sen	尖	A	A	B
	afraid	kia	hàipà	害怕	kyoufu	恐怖	A	B	C
	rope	sok	shéngsuǒ	绳索	jou	縄	A	B	C
	soil	tho	de tǔ	地土	tochi	土地	A	A	A
	hand	chiu	shǒu	手	shu	手	A	B	B
	pull	thua	lā	拉	in	引	A	B	C
	thick	kao	hòu	厚	kou	厚	A	A	A
	ear	hia	ěrduǒ	耳朵	ji	耳	A	B	C
	egg	nui	dàn	蛋	ran	卵	A	B	B
	fly	poei	fēi	飞	hi	飛	A	A	C
	laugh	chio	xiào	笑	shou	笑	A	A	A
	tetek	leng	nǎi	奶	nyuu	乳	A	B	C
	no	bo	bù	不	fu	不	A	A	B
	sleep	khun	shuìjiào	睡觉	suimin	睡眠	A	B	B
	three	sa	sān	三	san	三	A	A	A
	stabbed	chiam	cì	刺	shi	刺	A	A	A
	thin	pok	báo	薄	haku	薄	A	A	B
	inflatable	choi	chuī	吹	sui	吹	A	A	A
	stick	cha	zhàng	杖	jou	杖	A	A	B
	old	lao	lǎo	老	ro	老	A	A	A
	bone	kut	gǔ	骨	kotsu	骨	A	A	A
	blunt	tun	dùn	钝	don	鈍	A	A	A
	snake	cua	shé	蛇	ja	蛇	A	B	C
	intestines	teng	chángzi	肠子	chou	腸	A	B	B