

Classification of Specific Words of Oversea Chinese and its Implication: A Case Study of Malaysian Chinese

Li Jiwei¹ and Yao Xiaodong²

¹College of Chinese Language and Culture, Jinan University
Tianhe District Guangzhou, PRC.

²School of Foreign Languages, Beijing Forestry University
Haidian District, Beijing, PRC.

Received December 2014; revised December 2014

ABSTRACT. The objects of study for oversea Chinese vocabulary are their specific words, which can be divided into specific words of word items, sememe, collocation and frequency. The choices and treatment of the oversea Chinese specific words vary with the goals of study.

Key words: Oversea Chinese; specific words; classification, Malaysia

1. **Introduction.** Oversea Chinese (hereafter OC) has received growing attention in Chinese linguistic study for the past years and becomes a hot topic for the moment. Huáyǔ(Chinese), roughly taken as “the common language shared by all oversea Chinese” in the past, is in essence “Chinese in its narrow sense”, or the so called “the inner circle Chinese”. The current understanding is that “Huáyǔ(Chinese) is the language shared by Chinese people based on Mandarin Chinese”^[1], or “common language based on mandarin shared by all Chinese people all over the world”^[2,3], which is “Chinese in its broad sense” or “the Outer circle Chinese”. Compared with the concepts of “Chinese in its narrow sense” and Mandarin Chinese, this notion goes beyond the country boundary. Now with the understanding of Chinese has been updated and it is realized that Chinese in and out of China are integral components of Huáyǔ(Chinese), and they are the regional variations of Huáyǔ(Chinese). The characteristics, differences and similarities, and the interaction between these variations all call for in-depth study.

Descriptive adequacy is the base for further research. Two approaches are to be employed for the description of OC. The first is to take a certain Chinese variation as an independent, isolated system, and describe it comprehensively the way as we do to other language or dialect, presenting a holistic picture of it. A non-Chinese language user shall

take this method when he or she wants to have a good knowledge of the Chinese variation in Malaysia. The second approach, reveals the differences between OC and modern Mandarin by way of describing the specialties of the variation with a contrastive analysis of the mandarin and Chinese, which is the routine adopted by Chinese scholars to describe OC.

Great process has been witnessed in the field of OC lexicon description. A good understanding of OC lexicon has been achieved with the publishing of “Global Chinese Dictionary” edited by Li^[2]. Description and research of the overseas Chinese lexicon vary with their respective goals, such as for the purpose of compiling and editing overseas Chinese dictionary, aiming at making the testing standards for OC^[4], as well as presenting the lexical system underlying the overseas Chinese variation and its differences with that of modern Mandarin Chinese, and so on. The focus of study varies with different goals.

What is the object of OC lexicon study? One easy and quick answer is “Overseas Chinese specific words” (OCSWs), which however is hard to define. In the foreword to “Global Chinese Dictionary”, it is suggested that “as many local specific words and their senses as possible should be recorded”, while in “guide to using the dictionary”, “mainly to record the frequently seen words peculiar to various Chinese communities since 1980s”. Obviously, “specific words and their senses” does not necessarily equal to “the frequently seen words peculiar to the various Chinese communities”, with the latter a subset of the former.

Our conception of OC and the objectives of our study factor into the definition of “OCSWs”. This paper, grounded on the study of Chinese variation in Malaysia, classifies the OCSWs into four categories: (a) specific lexical-item words, (b) specific sememe words, (c) specific collocation words, and (d) specific frequency words. It is aimed at deepening our understanding of the differences in the vocabulary systems between OC and modern mandarin, shedding light on the modification of “Global Chinese Dictionary”, on the compiling of “Great Dictionary of Global Chinese”, and in this way paves the way for the making of OC testing standards. The data come from Overseas Chinese Literature Series (Malaysia section, 10 volume), and Southeast Asia Overseas Chinese Media Corpora in Overseas Chinese Research Center, Jinan University (Malaysia section, the abbreviation for this corpus is SAOC), with the former sourced, and the latter not.

2. Specific lexical-item words in Overseas Chinese. A lexical item means an item from a lexicon of a certain language or a language variant. Therefore, what we called “specific lexical-item words in overseas Chinese”, which are based on the contrast to modern mandarin Chinese, are unique in overseas Chinese only. A single word item includes the phoneme, the form, and the lexical meaning of the word; at the present time, the definition of specific lexical-item words mainly lies on the word form and the lexical meaning. In other words, if the form and the lexical meaning (it means all the sememes of a word rather than a single one) of a word item is distinguished from that of modern mandarin Chinese, it should be defined as a “specific lexical-item word”.

Recent researches about overseas Chinese mainly focus on “specific lexical-item words”,

which can be categorized from different aspects. For instance, from the perspective of mandarin-oversea Chinese contrast and the coordination of intra and external Chinese, it can be divided into “unique words”, “the words that have different forms but have same referent”, and “the words that have same form but have different referents”^[5]; from the perspective of source, “specific lexical-item words” can be classified into ‘loan words’, “dialect words”, “classical Chinese words”, and “community words”. The following part of this section is going to discuss “specific lexical-item words” in Malaysian Chinese from the perspective of source.

First, let’s talk about the loan words in “specific lexical-item words”. In Malaysian Chinese, there are a lot of words borrowed from Malaysian, English, and other languages, such as “甘榜(gānbǎng, village, translated form Malaysian)”, “巴冷(bālěng, a kind of sword, translated form Malaysian)”, “巴列(bāliè, creek, translated form Malaysian)”, “巴刹(bāshā, market, translated form Malaysian)”, “罗里(luólǐ, lorry, translated form the English word ‘lorry’)", “杯葛(bēigě, boycott, translated form the English word ‘boycott’)", “巴仙(bāxiān, percent, translated form the English word ‘percent’)", “依格(yīgé, acre, translated form the Malaysian word ‘ekar’, which is translated form the English word ‘acre’)". Here are several sample sentences:

(1) 由于禽流感在甘榜居民后院饲养的鸡群中散播，而非商业养鸡场，无法一举灭鸡，为灭鸡行动增添了一定的难度。

Yóuyú qínliúgǎn zài gānbǎng jūmín hòuyuàn sìyǎng de jīqún zhōng sànbò, ér fēi shāngyè yǎng-jīchǎng, wúfǎ yījǔ miè jī, wèi mièjī xíngdòng zēngtiān le yīdìng de nándù.

Since avian flu spread among the chickens kept in the back yard of the residents in the village rather than commercial chicken farms, it is hard to kill all the chickens at one stroke, increasing the difficulty to sweep the virus.

(2) 他透露，吉兰丹每天共有5万只肉鸡被在巴刹售卖。

Tā tòulù, Jilándān měitiān gòngyǒu 5 wàn zhī ròujī bèi zài bāshā shòumài.

He revealed that everyday in Kelantan, there are 50,000 chickens being sold in the markets.

(3) 我们的生意做得大，三架罗里也不够用，加上载黄梨常要塞车。(马崙《黄梨成熟时》)

Wǒmen de shēngyì zuòdé dà, sān jià luólǐ yě bù gòuyòng, jiāshàng zài huánglí cháng yào sāichē. (Mǎ lún, “huáng lí chéngshú Shí”)

We do a land office business, so even three lorries won't be enough. What's worse, we always come across traffic jam. (Malun, When the Yellow Pears Mature)

(4) 只是，马币脱钩至今，币值一直在狭窄的一两巴仙之间波动，近期印尼盾狂贬，甚至引发了马币贬值的可能。

Zhǐshì, mǎbì tuōgōu zhìjīn, bìzhí yīzhí zài xiázhǎi de yīliǎng bā xiān zhījiān bōdòng, jìnqī

yìnnídùn kuángbiǎn, shènzhì yǐnfā le mǎbì biǎnzhí de kěnéng.

The only thing is, since MYR decoupled, its currency value have been fluctuating within a narrow percent by one to two. Recently, the sharp devaluation of IDR may also bring the possibility of the devaluation of MYR.

The second category is the dialect words in “specific lexical-item words”. Li Rulong once pointed out that overseas Chinese settled abroad with the mother tongue of southeast China dialects. In Southeast Asia, Fujian dialect and Cantonese were mighty ones in the early period. The vocabularies of these two dialects easily became a part of overseas Chinese, in the mean time, lacking strict regulations, which resulted in the fact that there are many vocabularies from Chinese dialects in overseas Chinese.^[6] Take Malaysian Chinese as an example, there are numerous dialect words from Fujian dialect and Cantonese.

(5) 郑云城也是评论人，他非议说，上述文章暴露了典型的马华公会“头家政治”心态。

Zhèngyúnychéng yěshì pínglùn rén, tā fēiyì shuō, shàngshù wénzhāng bàolù le diǎnxíng de mǎhuá gōnghuì “tóujiā zhèngzhì” xīntài.

Zhen Yuncheng is also a commentator. He censured that the article above reveals the typical “boss politics” mindset of members in MCA.

(6) 只讲述马共的立场和斗争目标，俾社会同情及珍惜他们的贡献。

Zhǐ jiǎngshù mǎgòng de lìchǎng hé dòuzhēng mùbiāo, bǐ shèhuì tóngqíng jí zhēnxī tāmen de gòngxiàn.

Just talk about the position and target of CPM to make the society feel sympathy and cherish their contribution.

(7) 大批记者今日在沙里尔和居民代表的带领下，巡查数间受影响住家。

Dàpī jìzhě jīnrì zài Shāliěr hé jūmín dài biǎo de dàilǐng xià, xúncá shùjiān shòu yǐngxiǎng zhùjiā.

Today, a large group of journalists led by Shahrir and the representatives of the resident patrolled the residence influenced.

(8) 风帆并没有停下脚步。秋宁听到的是他沉重脚步的回音。闰中仿佛是隔了千山万水。(李忆著《困境》)

Fēngfān bìng méiyǒu tíngxià jiǎobù. Qiūníng tīngdào de shì tā chénzhòng jiǎobù de huíyīn. Jiǎnzhōng fǎngfú shì géle qiānshānwànshuǐ. (Lǐyìjūn, “kùnjìng”)

Fengfan did not stop. Qiuning could only hear the echo of his heavy foot step. It seem that there were a rather far distance between them. (Li Yijun, Dilemma)

(9) 这则专访出街后，国内各语文报纸纷纷转载。

Zhè zé zhuānfǎng chūjiē hòu, guónèi gè yǔwén bàozhǐ fēnfēn zhuǎnzài.

After this interview was published, it have been reprinted on many Chinese newspapers.

(10) 原棕油是马来西亚表现标青的领域。

Yuán zōngyóu shì mǎláixīyà biǎoxiàn biāoqīng de lǐngyù.

Palm oil industry is Malaysia's outstanding field.

“头家 (tóu jiā)” in (5) means “boss”; “俾 (bǐ)” in (6) appears both in Cantonese and Fujian dialect, which means “let”; “住家 (zhùjiā, residence)”, “间中 (jiānzhōng, between, middle)”, “出街 (chūjiē, go into the street)”, “标青 (biāoqīng, outstanding)” in (7) - (10) are from Cantonese.

Now we turn to classical Chinese words in “specific lexical-item words”. Due to the fact that Malaysian Chinese is deeply influenced by south Chinese dialects and Cantonese and Fujian dialect keep more ancient Chinese vocabularies than modern mandarin Chinese, there is a part of overlap between classical Chinese words and dialect words in Malaysian Chinese. “俾 (bǐ, let)” in (6) also belongs to the category of classical Chinese.

(11) 她说，新山宽柔五小发生两百名家长漏夜排队为孩子抢学额的事件，反映了华小不足的严重性。

Tā shuō, Xīnshān Kuānróu wǔxiǎo fāshēng liǎngbǎi míng jiāzhǎng lòuyè páiduì wèi háizi qiǎng xué'é de shìjiàn, fǎnyìng le huáxiǎo bùzú de yánchóngxìng.

She said that it showed the strict problem on the shortage of Chinese primary schools that 200 parents lined a whole night in order to get the limited planned places in the five primary schools in Xinshan and Kuanrou.

(12) 他们于是献议我出任政治部人员一职。

Tāmen yúshì xiànyì wǒ chūrèn zhèngzhìbù rényuán yīzhí.

They suggested to me that I can take the job in the politics department.

(13) 此外，也希望他对我崎岖的命运有所提点，为我指引一道明路。(柏一《红疯子》)

Cǐwài, yě xīwàng tā duìwǒ qíqū de mìngyùn yǒusuǒ tídiǎn, wèiwǒ zhǐyǐn yīdào mínglù. (Bǎi yī “hóng fēngzi”)

In addition, I hope that he can remind me something in my rough destiny and lead me to a bright way. (Baiyi, The Red Lunatic)

(14) 泰半马来人族群仍然贫富悬殊，处在赤贫水平。

Tàibàn mǎláirén zúqún réngrán pínfù xuánshū, chù zài chìpín shuǐpíng.

Over half of the Malaysian ethnic group still face the deep wealth gap, suffering in abject poverty.

“学额 (xué'é, planned place in school enrollment)” in (11) derives from the number of the students admitted by government through the imperial competitive examinations. In Malaysian Chinese, it has the similar meaning as what we called “学位(xuéwèi, the limited study position)” in mainland China. “献议 (xiànyì)” in (12) equals to “建议 (jiànyì,

advise)”. “提点 (tídiǎn)” in (13) means “remind and direct”. It was also used in nearly ancient Chinese novels, such as in the episode 6 of *Èrshí nián mùdǔ zhī guài xiànzhuàng* (*Odd Things Witnessed Over Twenty Years*): “教导呢，我是不敢说；不过处处都想提点你，好等你知道些世情。 *Jiàodǎo ne, wǒ shì bùgǎn shuō; bùguò chùchù dōu xiǎng tídiǎn nǐ, hǎo děng nǐ zhīdào xiē shìqíng.* (I dare not say that I can teach you something, but I will remind you sometimes, for you to get to know about the society.) ” We can find it in modern Cantonese as well. “泰半 (tàibàn)” means “over a half or more than a half”. This word has been existing in ancient Chinese for quite a long time. The four words above are not included in *Modern Chinese Dictionary (sixth edition)*.

The last category is community words in “specific lexical-item words”. Community words is, Tian Xiaolin commented, the abbreviation of social community words. Along with dialect words, loan words, and classical Chinese words, community words are an important origin and component of the normal words in modern Chinese vocabulary. Community words break a new path for us to find that in different Chinese communities, there is a group of special words reflecting the certain formation of a certain community, including the words on politics, economy, and culture.^[7] There are huge amounts of “specific lexical-item words” from community words, such as “通译员 (tōngyìyuán, translator)”, “图测 (tú cè, layout)”, “组屋 (zǔwū, residence built by the government)”, “娘惹 (niáng rě, local female Chinese)”, “终站 (zhōngzhàn, terminal)”, “准证 (zhǔnzhèng, permission note)”, “东协 (dōngxié, ASEN)”, “积犯 (jī fàn, recidivist)”, “波道 (bō dào, channel)”, “滑鼠 (huáshǔ, mouse)”, “电眼 (diànyǎn, road camera)”, “交通圈 (jiāotōngquān, roundabout)”, “廉宜 (liányí, cheap)”, and so on.

(15) 他也表示，工程局不会因大城堡居民反对，而修改相关高路的建造结构图测。

Tā yě biǎoshì, gōngchéngjú bù huì yīn dàchéngbǎo jūmín fǎnduì, ér xiūgǎi xiāngguān gāolù de jiànào jiégòu túcè.

The also said that the engineering bureau will not modify the layout of the highway simply because the residents are against to it.

(16) 幸亏昨天，在他第四次的考试大关上，他居然“过关”了，领到了一张准证。
(马汉《平安夜》)

Xìngkuī zuótiān, zài tā dì sì cì de kǎoshì dàguān shàng, tā jūrán “guòguān” le, lǐngdào le yī zhāng zhǔnzhèng. (Mǎ Hàn “píng'ān yè”)

Luckily, he passed his fourth test last night and got a permission note. (Mahan, Silent Night)

(17) 当马来西亚人再度来到未来的十字路口，华裔朝野政治，似乎又绕回 18 年前的交通圈。

Dāng Mǎláixīyǎ rén zàidù láidào wèilái de shízì lùkǒu, huáyì cháo yě zhèngzhì, sìhū yòu ràohuí 18 nián qián de jiāotōngquān.

When the Malaysian meet the crossroad of the future once again, the Chinese

government politics seems to return to the roundabout 18 years ago.

(18) 不过采用这种最廉宜的方法清芭，仍然是油棕公司的惯行。

Bùguò cǎiyòng zhè zhǒng zuì liányí de fāngfǎ qīng bā, réngrán shì yóuzōng gōngsī de guàn xíng.

But taking the cheapest method to clean the plantain is still the convention of the oil palm companies.

3. Specific sememe words in Oversea Chinese. Sememe is the basic unit of word senses. One word can only have one sense, but in most cases, one word has a variety of senses. Generally speaking, one sememe is one sense item in dictionary. There is such a kind of word, existing in both modern mandarin and some oversea Chinese variant. Nevertheless, it has some specific sememes in Oversea Chinese, while in modern mandarin it doesn't. We define this kind of word as Specific sememe word. Let's take “素质(sùzhì)”, “报效(bàoxiào)”, “投入(tóurù)”, “离开(líkāi)”, “夜(yè)”, “慢(màn)”, “浓(nóng)”, “把(bǎ)” as examples.

In modern Chinese, “素质(sùzhì)” has three senses as a noun. The first is “the nature of something”, the second is “(human) accomplishment”, the third is “the innate characteristics of one person's nervous system and sense”, during which the second sense is more common. Although these senses of “素质(sùzhì)” are also used in Malaysian Chinese, the more common is its another meaning, “质量(zhìliàng, quality)”.

(19) 对普腾而言实在是难堪的决定，也是进一步反映出，普腾的产品素质实在有待改善。

Duì Pǔténg éryán shízaishì nánkān de juédìng, yěshì jìnyìbù fǎnyìng chū, Pǔténg de chǎnpǐn sùzhì shízài yǒudài gǎishàn。

It is really an embarrassing decision for Proton, and further reflects the fact that the quality of Proton's products is yet to be improved.

(20) 全马各地区的空气素质正常，不过，印尼及马来西亚的林火热点剧增。

Quánmǎ gèdìqū de kōngqì sùzhì zhèngcháng, búguò, Yìnnì jí Mǎláixīyà de línhuǒ rèdiǎn jùzēng。

The air quality of all regions in Malaysia is normal while the number of forest fire hotspot increases dramatically in Indonesia and Malaysia.

The verb “报效(bàoxiào)” only has one sense in modern Mandarin Chinese, that is “to do one's best in order to repay one's kindness.” This meaning is used occasionally in Malaysian Chinese, the more common sememe is “提供(tígòng, to provide)” and “捐赠(juānzèng, to donate)”.

(21) 我们放一个大架子，放菜呀、米呀、油呀、盐呀、米粉呀，都是村民们报效。

Wǒmen fàng yíge dà jiàzi, fàng cài ya, mǐ ya, yóu ya, yán ya, mǐfěn ya, dōushi

cūnmínmen bàoxiào。

We set a big shelf where we can put vegetables, rice, oil, salt, and rice noodles, all of which are provided by villagers.

(22) 这名善心人士答应捐献一块位于士毛月的土地，及报效一间价值超过一亿元的中医学院。

Zhè míng shànxīn rénrén de dāyìng juānxiàn yíkuài wèiyú Shìmàoyuè de tǔdì, jí bàoxiào yìjiān jiázhí chāoguò yíyìyuán de zhōngyīxuéyuàn。

This generous person promised to donate a piece of land located in Semenyih and an institute of traditional Chinese medicine which valued more than 100 million yuan.

According to “Modern Chinese Dictionary (sixth edition)”, “投入(tóurù, to devote)” can be used as noun, verb and adjective, and has four senses. But beyond that, the word has a more-often-used verb meaning in Malaysian Chinese that is “加入(jiārù, join a organization/team /industry)”.

(23) 现年30岁的法蒂玛自1996年投入警队，当时为辅警，2005年被调到特别行动部队。

Xiànnián 30suì de Fǎdìmǎ zì 1996nián tóurù jǐngduì, dāngshí wéi fǔjǐng, 2005nián bèi tiáodào tèbiéxíngdòngbùduì。

The 30-year-old Fatima joined the police force since 1996 as an auxiliary officer, and was transferred to Special Action Force in 2005.

(24) 参与及投入华人社团的，真的是剩下“有心人”了。

Cānyù jí tóurù huárén shètuán de, zhēnde shì shèngxià “yǒuxīnrén” le。

Only the remaining “caring people” participate and join associations of oversea Chinese.

The verb “离开(líkāi)” only has one sense in modern Chinese, that is “separate from someone/ something/some place”. But in Malaysian Chinese, there’s another verb meaning, which is equal to “距离(jùlí, be away from)” and “离(lí, be away from)”.

(25) 我没有了母亲，只有一位父亲，在乡下靠割胶维生，离开这里有一百多公里路。

(孟沙《遗爱》)

Wǒ méiyǒule mǔqīn, zhǐyǒu yíwèi fùqīn, zài xiāngxià kào gējiāo wéishēng, líkāi zhèlǐ yǒu yìbáiduō gōnglǐ lù. (Mèngshā 《yí’ài》)

I lost my mom, and only have one father, who lives by tapping rubber in countryside, which is more than one hundred kilometers away from here. (Mèngshā “Eternal Love”)

(26) 目前，离开2006年新学年开学日期只剩下区区两个多月的时间。

Mùqián, líkāi 2006nián xīn xuénián kāixué rìqī zhǐ shèngxià qūqū liǎnggè duōyuè de shíjiān。

There are merely 2 months from the 2006 year semester.

In modern Chinese, “夜(yè, night)” is a noun, while in Malaysian Chinese, it can be used as an adjective, means “深(shēn, late)” or “晚(wǎn, late)”.

(27) 真对不起，这么夜还把你吵醒。(李忆著《痴男》)

Zhēn duìbùqǐ, zhème yè hái bǎ nǐ chǎoxǐng. (Lǐyìjūn 《chīnān》)

I'm terribly sorry. It's so late but I still wake you up. (Yijun Li "crazy man")

(28) 不要太夜打烊，不熟的顾客可以不要做。

Búyào tài yè dǎyáng, bùshú de gùkè kěyǐ búyào zuò.

Do not close her shop too late or do business with unfamiliar customers.

The adjective “慢(màn, slow)” is opposite to “快(kuai, fast)” in modern Chinese, means “the speed is slow” or “cost too much time to walk or do something”. Besides, “慢” also means “晚(wǎn, late)” or “迟(chí, late)” in Malaysian Chinese, which is more commonly used.

(29) 易言之，虽然少年安华出身学生运动，但论年龄却比纳吉慢了12年从政。

Yìyánzhī, suīrán shàonián Ānhuá chūshēn xuéshēng yùndòng, dàn lùn niánlíng què bǐ Nàjǐ màn le 12nián cóngzhèng.

In other words, although young Anwar engaged in students movement at very early time, as to the years of politics, he is 12 years later than Naji.

(30) 就是说黄梨厂吃不完黄梨，载去也无法下车，慢几天再砍。(马喙《黄梨成熟时》)

Jiùshìshuō huánglichǎng chībùwán huánglí, zàiqù yě wúfǎ xiàchē, màn jǐtiān zài kǎn.

That is to say, the pineapple plant can't finish all pineapples, even if you transport them, they can't be unloaded, so you'd better cut them a few days later.

“浓(nóng)” is an often-used adjective in modern Mandarin Chinese, which has three senses, they are: opposite to “淡(dàn, weak)”, the color is “深(shēn, dark)”, and “深(shēn, deep)”. To match the three senses there are following collocations: “浓茶(nóngchá, strong tea)”, “浓绿(nónglǜ, strong green)” and “兴趣很浓(xìngqù hěnnóng, great interests)” etc.. In Malaysian Chinese, there is another new meaning, which express “(the possibility of opportunities is)大(dà, great), 高(gāo, high)”.

(31) 是的，听说我升级的机会很浓，但是事情还没有尘埃落定，谁也不敢肯定。(朵拉《胜利者》)

Shìde, tīngshuō wǒ shēngjí de jīhuì hěn nóng, dànshì shìqìng hái méiyǒu chén'āilùodìng, shuí yě bùgǎn kěndìng. (DuōLā 《ShèngLìZhě》)

Yes, I heard that I have got great opportunities for promotion, but things are still unfolding, no one can make sure. (Dora "The winner")

(32) 消息说，三名副主席人选中，前三名候选人中选的机会最浓。

Xiāoxishuō, sānmíng fūzhǔxí rénxuǎn zhōng, qiánsānmíng hòuxuǎnrén zhòngxuǎn de jīhuì zuì nóng。

News said that, the top three vice-presidential candidates have got the best chance for the three seats.

According to *Modern Chinese Dictionary(sixth edition)*, the classifier “把(bǎ)” could be used in the following five situations: ① to collocate with something with a handle or appliances that can be grabbed by hand, “一把刀(yībǎdāo, a knife)”; ② to count the quantity that one hand can grab, “一把米(yībǎmǐ, a handful of rice)”; ③ to collocate with something abstract, “一把年纪(yībǎniánjì, long in the tooth)”; ④ to describe hand’s action, “拉他一把(lātāyībǎ, help him a hand)”; ⑤ to count the number of doing something, “过把瘾(guòbǎyǐn, getting high)”. These collocations are also used in Malaysian Chinese, besides, the classifier “把” could count “嘴(zuǐ, mouth)” and “声音(shēngyīn, voice)”.

(33) 人民公正党鹅麦区国会议员阿兹敏阿里也插一把嘴：“国会成为巴刹，因为我们
有来自林茂的巴刹主席！”

Rénmínggōngzhèngdǎng émàiqū guóhuì yìyuán Āzīmǐn'ālǐ yě chā yībǎzuǐ : “guóhuì chéngwéi bāshà, yīnwèi wǒmen yǒu láizì línmao de bāshà zhǔxí。 ”

Mohamed Azmin Ali, the PKR congressman from Gombak district, also interrupted, “Congress has become Pasar, because we have Pasar chairman from Lin Mao”

(34) 在巫统代表大会第二天，代表口中难得地出现了一把自省的声音。

Zài wūtǒng dàibiǎo dàhuì dì'èrtiān, dàibiǎo kǒuzhōng nándé de chūxiàn le yì bǎ zìxǐng de shēngyīn。

On the next day of UMNO congress, the delegates rarely show their introspection.

4. Special collocation words in Oversea Chinese. Special collocation words of Oversea Chinese are the words which exist both in modern Chinese and some variant of oversea Chinese, and these words have almost the same meaning in the two language system, but some collocations can only be found in variant of oversea Chinese. Different collocations make the features of language different. If we want to have an in-depth understanding of the difference between oversea Chinese and modern Chinese, “Special collocation words” are one thing that we must know. Let me show you by giving the following examples: “课题”(kètí), “帮忙”(bāngmáng), “灵光”(língguāng), “优越”(yōuyuè), “圆满”(yuánmǎn).

In modern Chinese, “课题”(kè tí) means “a problem someone should investigate or discuss or an important issues which needs to be done immediately”. When it comes to specific collocation, it is related to scientific research or a serious problem. However, in Malaysian Chinese, it barely is associated with scientific research. For example:

(35) 因此，我们可以看到马来西亚执政当局在处理汽油涨价课题时，也特别谨慎。

Yīncǐ, wǒmen kěyǐ kàndào Mǎláixīyà zhízhèng dāngjú zài chǔlǐ qìyóu zhǎngjià kètí shí,

yě tèbié jīnshèn.

Therefore, we could see the Malaysian authority was very cautious while it was handling the problem of the rising price of petroleum.

(36) 他认为，目前情况很混乱，对马航未来首脑人选的课题，他也不愿置评。

Tā rènwéi, mùqián qíngkuàng hěn luàn, duì Mǎháng wèilái shǒunǎo rénxuǎn de kètí, tā yě bùyuàn zhìpíng.

He thought the current situation was a chaos, and he would make no comment on the issue of “who would be the best for the future leader of Malaysia Airlines”.

“帮忙”(bāngmáng) is an intransitive verb in modern Chinese. If we want to say “help somebody” in Chinese, we should say “bāng somebody a máng”. But in Malaysian Chinese, “帮忙” is a transitive verb. It has the same use as “帮”(bāng, help) and “帮助”(bāng máng, help).

(37) 身为家中长子，我小时候还得不时帮忙父亲载送鸡蛋到市区的商店。

Shēnwéi jiāzhōng zhǎngzǐ, wǒ xiǎoshíhòu hái dēi bùshí bāngmáng fùqīn zài sòng jīdàn dào shìqū de shāngdiàn.

When I was little, I have to help my father deliver eggs to the downtown shop as the oldest son.

(38) 如果你不要帮忙我们，就直接告诉我们，那就算了。

Rúguǒ nǐ bù yào bāngmáng wǒmen, jiù zhíjīe gàosù wǒmen, nà jiù suàn le.

If you don't help us, just tell us, and that's it.

The explanation of the adjective “灵光(língguāng)” from *Modern Chinese Dictionary(The Sixth Version)* is: “good”; “good effect”. In modern mandarin, “灵光” is always collocated with “头脑(tóunǎo, brain)”. However, in Malaysian Chinese, it is also can be collocated with “语言(yǔyán, language)”.

(39) 一名来自吉打州居林区的年老代表陈楚江，由于马来语不灵光，只能以生硬的马来语发言。

Yī míng láizì Jídǎ zhōu jū línqū de niánlǎo dàibiǎo Chén Chǔ jiāng, yóuyú mǎláiyǔ bù língguāng, zhǐnéng yǐ shēngyìng de mǎ láiyǔ fāyán.

Chen Chujiang, an old representative from Kulim, Kedah, spoke Malay in a very unskilled way as his Malay was not “good” enough.

The explanation of the adjective “优越(yōuyuè)” from *Modern Chinese Dictionary (The Sixth Version)* is: “superior”, “excellent”. In modern Chinese, “优越” is often used with “条件(tiáojiàn, condition)”, “地位(dìwèi, status)”, “地理位置(dìlǐ wèizhì, geographical location)”, “地理环境(dìlǐ huánjìng, geographical environment)” and affix like “感(gǎn, feeling)”, “性(xìng, nature)”. These collocations also can be seen in Malaysian Chinese, but there are many more which cannot be found in modern Chinese. For example:

(40) 汉达和利道安从中一到中三都由于成绩优越而被编在同一班上课。(碧澄《牵》)

Hàndá hé Lidào ān cóng zhōngyī dào zhōngsān dōu yóuyú chéngjì yōuyuè ér bèi biān zài tóng yī bān shàng kè. (bìchéng, qiān)

Handa and Lidaoan were arranged in the same class from grade one to grade three in junior high school because of their excellent academic records.

(41) 该国总统于今日会晤首相拿督斯里阿都拉后，宣布提供 6 名大马优越生至其国进修的奖学金作为答谢。

Gāiguó zǒngtǒng yú jīnrì huìwù shǒuxiàng nādū sīlǐ Ādūlā hòu, xuānbù tíngōng 6 míng dà mǎ yōuyuèshēng zhì qíguó jìnxiū de jiǎngxuéjīn zuòwéi dáxiè。

After the president met the prime minister Datuk sri Abdullah, the president announced that six excellent Malaysian students taking a further study in his country would receive scholarship in return.

(42) 沙菲说，为了把大马打造成为教育优越的国家，政府将会在马来西亚第 9 计划下，在吉兰丹，丁加奴及雪兰莪各设立 1 间工艺学院。

Shāfēi shuō, wèile bǎ dà mǎ dǎzào chéngwéi jiàoyù yōuyuè de guójiā, zhèngfǔ jì āng huì zài Mǎláixīyà dì jiǔ jìhuà xià, zài Jilándān, Dīngjiānú jí Xuělánē jié gè shèlì yī jiān gōngyì xuéyuàn。

Shafie said, the government will execute the “Malaysian 9th plan” to set a polytechnic in Kelantan, Terengganu and Selangor separately for making Malaysia a country of superior education.

In modern Chinese, “圆满(yuánmǎn, satisfactory)” is often used in “圆满的答复(yuánmǎn de dáfù, a satisfactory reply)”, “圆满地解决(yuánmǎn de jiějué, be settled well)”, “圆满结束(yuánmǎn jiéshù, a successful close)”. These collocations are also widely used in Malaysian Chinese. Besides, “圆满” can modify time phrases in Malaysian Chinese as well. For example:

(43) 前副首相安华被革职圆满六年的那一天，终于赢回了个人的人身自由与尊严。

Qián fùshǒuxiàng Ānhuá bèi gézhí yuánmǎn liùnián de nà yī tiān, zhōngyú yíng huí le gèrén de rénsēn zìyóu yǔ zūnyán。

Ex vice-premier Anwar finally won back his freedom and dignity today, whole six years after he was dismissed.

(44) 目前距离第 2000 天只剩下 175 天，“慈祥、中庸、温和”的阿都拉上台圆满两年两个月。

Mùqián jùlǐ dì 2000 tiān zhǐ shèngxià 175 tiān, “cíxiáng, zhōngyōng, wēnhé” de Ādūlā shàngtái yuánmǎn liǎngnián liǎnggè yuè。

Now it's only left 175 days to the 2000th day, the “amiable, moderate and gentle”

Abdullah has taken the office for a complete 2 years plus 2 months.

5. **Special frequency words in oversea Chinese.** To better distinguish oversea Chinese and modern mandarin Chinese, attention should be paid not only to the topic of “presence or absence”, but also the topic of “frequency”. The topic of “presence or absence” refers to some linguistic phenomena, which exist in oversea Chinese but not exist in modern mandarin Chinese. The topic of “frequency” refers to the different frequency of use in some certain lexical items and grammar items between oversea Chinese and modern mandarin Chinese. For example, some numeral abbreviations like “五讲四美(wǔjiǎngsìměi, five attentions and four beauties)” and “双规(shuāngguī, double ruling)” are frequently used in modern Chinese but rarely used in oversea Chinese. In the same way, some lexical items are frequently used in oversea Chinese but rarely used in modern Chinese; these differences shouldn’t be ignored in the process of studying oversea Chinese. This paper names the words frequently used in oversea Chinese but rarely or less used in modern Chinese “频率特色词 (pínlǜ tèshè cí, special frequency words)”.

Firstly, let’s take a look at the temporal noun “当儿(dāng’er)”. It is usually used in spoken Chinese, referring to “时候(shíhou, moment). The expression forms like “在…的当儿(zài…de dāng’er, at…moment)”, “当…的当儿(dāng … de dāng’er, when)”, “…的当儿(…de dāng’er, the moment…)” in modern Chinese are also used in Malaysian Chinese, but the frequency of use is much higher. We can find these expressions in SAOC (about 100,000,000 words) 409 times. However, in the MCCPU (abbreviation for “Modern Chinese corpus of Peking University”, about 300,000,000 words), we can only find these expressions 410 times. That is to say, the usage of these expressions in modern Chinese could be commensurate with Malaysian Chinese with 3 times larger corpus scale. Take a look at the use cases in Malaysian Chinese:

(45) 半岛的天空被烟霾笼罩的当儿, 马来西亚的经济看来也要明年方能放晴。

Bàndǎo de tiānkōng bèi yānmái lǒngzhào de dāng’er; Mǎláixīyà de jīngjì kànlái yěyào míngnián cáinéng fàngqíng.

When the peninsula is being capped by the smoke. It seems that the economy of Malaysia won’t clear up until next year.

(46) 我常在他输血的当儿, 坐在幽黯的角落自言自语。(李忆著《红颜》)

Wǒ cháng zài tā shūxiě de dāng’er; zuòzài yōu’àn de jiǎoluò zìyánzìyǔ.

As his transfusion progressed, I often sit at the dark corner and took to myself. (Li Yijun, the Beauty)

The verb “恋栈(liànzhàn)” can be explained as “留恋(liúliàn, be reluctant to leave)”. In MCCPU, it is used 26 times, such as “恋栈北京的士兵, 则不可能成为未来的元帅! (Liànzhàn běijīng de shìbīng, zé bùkěnéng chéngwéi wèilái de yuánshuài, the soldier reluctant to leave Běijīng is unlikely to become the marshal in the future!)”, which shows that the word is used in modern Chinese. However, in SAOC, “恋栈” is used 109 times, the

use frequency of which in Malaysian Chinese is much higher than modern Chinese. Let's see the cases:

(47) 那些恋栈政府高职的马华公会高层，更不会轻言放弃难得的升官机会。

Nàxiē liànzhàn zhèngfǔ gāozhí de mǎhuá gōnghuì gāocéng, gèng búhuì qīngyán fàngqì nándé de shēngguān jīhuì.

Certainly, those senior members of the Malaysian Chinese society union, who are unwilling to leave the high position in the government, won't give up the rare opportunities of promotion.

The verb “撙节 (zǔnjié) often used in written Chinese, means “节约 (jiéyue, to save)” and “节省 (jiéshěng, to economize). In SAOC, “撙节” is used 195 times, while in MCCPU, it is only used 12 times, which shows that the word is used much more frequently in Malaysian Chinese than contemporary Mandarin Chinese.

(48) 槟州首席部长林冠英在槟州行政议会后表示，为了撙节开支槟州政府决定暂缓购买供议员们，还有议长使用的手提电脑和打印机。

Bīnzhōu shǒuxí bùzhǎng Lín Guānyīng zài bīnzhōu xíngzhèng yìhuì hòu biǎoshì, wèile zǔnjié kāizhī bīnzhōu zhèngfǔ juédìng zànhuǎn gòumǎi gōng yìyuánmen, háiyǒu yìzhǎng shǐyòng de shǒutídiànnǎo hé dǎyìnjī.

Lin Guanying, the chief minister of Penang, expressed after the Executive Council that the government of Penang decided to suspend the purchase of laptops and printers for senators and speakers to cut the budget.

In modern Chinese, “予(yǔ, give)”, a classical Chinese verb, always used with verbs like “给(gěi, give)”, “赋(fù, grant)” or “交(jiāo, hand)” to form double-syllable words “给予(jǐyǔ, give)”, “赋予(fùyǔ, grant)” or “交予(jiāoyǔ, hand)”. Occasionally, there are forms like “V+NP1+予+NP2” (such as “送画予人[sòng huà yǔ rén, give a picture to someone]”), “予+NP1+NP2” (such as “予人无限的想象空间[yǔ rén wúxiàn de xiǎngxiàng kōngjiān, give someone infinite imaginary space]”). But these forms are rarely used. However, in Malaysian Chinese, these two forms of “予 (yǔ, give)” are used frequently. In the aspect of the use numbers, we find that “予+人+NP” (yǔ +rén+NP, give sb. sth.)” is only used once in MCCPU's first 50 cases. However, in SAOC, these forms appears 17 times in the first 50 cases, which shows that “予 (yǔ, give)” is uses much more frequently in Malaysian Chinese than modern Chinese. Let's look at the cases of the use “V+NP1+予+NP2” and “予+NP1+NP2”.

(49) 南洋报业控股曾于2003年10月份，发出一批红股予股东。

Nányáng bàoyè kònggǔ céngyú 2003nián10yuèfèn, fāchū yīpī hónggǔ yǔ gǔdōng.

Nanyang Press Holdings gave out some bonus stocks to stockholders in October 2003.

(50) 警察部队自然比其他公共服务部门更易予人纪律败坏的印象。

Jǐngchá bùduì zìrán bǐ qítā gōnggòng fúwù bùmén gèngyì yǔrén jìlǜ bàihuài de yìnxiàng

Compared to other public service sectors, it is easier for the police force to give an undisciplined impression on people.

The adjective “熟络(shúlù, familiar)” is used in modern Chinese. It is used 21 times in MCCPU while 48 times in SAOC, which is rather much. “熟络” is the synonym with “熟悉 (shúxī)”, such as:

(51) 这时刘先生就献计表示自己与移民厅内部人员非常熟络, 并且愿意帮助小毛脱离牢狱之苦。

Zhèshí Liúxiānshēng jiù xiànjì biǎoshì zìjǐ yǔ yímíntīng nèibù rényuán fēicháng shúlùò, bìnghuān yuànyì bāngzhù Xiǎomáo tuōlí láoyùzhīkǔ.

At this time, Mr. Liu made a suggestion, showing that he knew the insiders of the immigration very well and was willing to help Xiaomao out of jail.

The adjective “淡静 (dànjìng)” is the near synonym with “平静 (píngjìng, tranquility). It was used 24 times in MCCPU while 320 times in SAOC, which shows that it is much more frequently used in Malaysian Chinese than modern Chinese. Let’s look into the specific cases:

(52) 大马股市今早开市在淡静的交投下呈上扬趋势, 市场焦点落在特定建筑股项。

Dàmǎ gǔshì jīnzǎo kāishì zài dànjìng de jiāotóuxià chéng shàngyáng qūshì, shìchǎng jiāodiǎn luòzài tèdìng jiànzhù gǔxiàng

Malaysian stock market, opened this morning, shows the rising tendency in quiet trade. The market focus is on specific building stocks.

6. Conclusion: Implications of the Classification of OCSWs. On the basis of the description of the OC vocabulary, this study has classified OCSWs into four categories, specific lexical-item words, specific sememe words, specific collocation words and specific frequency words, and the selection and treatment of these four types of OCSWs vary with the goals of study.

For the purpose of dictionary compiling, such as “Global Chinese Dictionary” or “Great Dictionary of Global Chinese”, “specific lexical-item words” is the top priority to be taken into consideration and must be collected, while whether to collect “specific sememe words” depends on the target readers and the compiling technology. Meanwhile consideration shall also go to the sememe shared by modern Mandarin and OC. On the one hand, it would be redundant for the native Chinese readers and the size of the dictionary will be enlarged. On the other hand, it will easily impression the readers that the word only bears specific meaning in OC. Specific sememe words are occasionally collected in “Global Chinese Dictionary” in 2010, with only the specific sememe collected. Take “美(měi, beautiful)” for instance, the senses offered include “with good quality”, “fresh” and “delicious”, leaving out other senses such as “beautiful” and happy” present in modern Mandarin. As mentioned before, “报效(bàoxiào)” is absent in “Global Chinese Dictionary”, in whose modification, if only specific sense “to provide” and “to donate” are offered, readers will

most possibly curious about whether the sense “to try one’s best to reward the other party’s kindness” is absent from the Chinese variation in Malaysia. “Co-reference method” can be a solution to such a dilemma. For instance, in the entry of “报效(bàoxiào)”, the specific sememe will be offered first, and then other sememes shared with the modern Mandarin Chinese. If possible, offer the respective frequencies of specific sense and common senses. Such as in Chinese variation in Malaysia, the sense of “to provide” and “to donate” of “报效” is more frequent than the commonly used sense “to try one’s best to reward the other party’s kindness” in modern Mandarin.

Specific Collocation words need not be taken into consideration when modifying and editing “Global Chinese Dictionary”, for they do not bring about new sememe, but they shall be highlighted by the Chinese language teaching and testing. Specific collocation words have implication for the collocation error analysis in the oversea Chinese students’ language learning process, and they serve as an important reference for the making of Chinese proficiency testing. Wang et. al. argues for an integration of OC study and Chinese testing, making a “Oversea Chinese Proficiency Standard” by way of adjusting and modifying the “standards of native Chinese speaker” based on the typical oversea Chinese convention.^[4] A case in point is the adjective phrase “优越(yōuyuè, superior)”, which is often employed to modify the noun “成绩(chéngjì, academic records)” and “表现(biǎoxiàn, performance)”. This kind of phenomenon shall be taken into consideration when the “oversea Chinese Standard” is issued.

The issue of “specific frequency words” reminds us that the differences between modern Mandarin and OC variants lies not only in the presence or absence of linguistic phenomena, but also in their frequency, which confirms to the “usage-based perspective” in functional linguistics, where roles played by frequency in linguistic changes and linguistic landscape is highly stressed, which at the same time is a perspective for the differences between various OC variants. A case in point is that a general assumption is that Chinese variants in Malaysia and Singapore have a close tie and differences between them are meager. When comes to the term “撙节(Zǔnjié, retrench)”, frequencies of its usage differ greatly even in the same corpus, i.e. Southeast Asia Chinese Media Corpora in Oversea Chinese Research Center, Jinan University, with 195 items in Malaysia Chinese corpus, and only 2 in the Singapore Chinese corpus. In the near future, with the growing influence from modern Mandarin on global Chinese variants, specific frequency words, hopefully will be a point of view penetrating into OC variants and changes.

Compared with the relatively mature description and study of English variants in the world, study on the OC is still in its infancy. This research, taking Chinese in Malaysia as example, proposes a new method for the classification of specific words, so as to update and deepen our view of the landscape of oversea Chinese lexicon, hence greatly broaden the scope for OC lexicon study.

Difficulties ahead are evident. For an adequate description, especially one including qualitative study, large-scale accurate corpora is essential. What we have is far from being sufficient. The data in this study, from “Southeast Asia Chinese Media Corpora” at Jinan University, is confined to Media text and short of stylistic variations and diversities. The

building of large-scale accurate multilingual corpora with country-specific is the urgent task ahead of us for the moment.

REFERENCES

- [1] Guo, Xi. A few words on “oversea Chinese”[J]. *Journal of Beihua University*, 1: 37-42, 2010.
- [2] Li. Yuming (Ed.). *Global Chinese Dictionary* [M]. Beijing: The Commercial Press, 2010.
- [3] Li, Yuming. Hierarchical variation of Chinese [R]. A Lecture at linguistics Institute of China. given at Beijing Language and Culture University, Aug. 3, 2013.
- [4] Wang, Hanwei. et. al. The overall design of HSC[J]. *TCSOL studies*, 4: 84-89, 2013.
- [5] Guo, Xi. On vocabulary coordination and standard of modern mandarin Chinese and Singapore-Malaysia Chinese [J]. *Social Sciences in Nanjing*, 12: 78-83, 2002.
- [6] Li, Rulong. A new horizon on oversea Chinese dialects[J]. *Lexicographical Studies*, 1: 42-47, 2013.
- [7] Tian, Xiaolin. 2004. A study on Hongkong Community words[J]. *Language Science*, 3: 29-37.