

การศึกษาเปรียบเทียบเชิงปริมาณของคำพ้องความหมายระหว่าง ภาษาจีนในเอเชียตะวันออกเฉียงใต้กับภาษาจีนกลาง

ซ่งเฟย

คณะภาษาจีน มหาวิทยาลัยภาษาต่างประเทศที่ 2 แห่งปักกิ่ง กรุงปักกิ่ง 100024 สาธารณรัฐประชาชนจีน
อีเมล: bestsongfei@126.com

เกียรติกศักดิ์ ฟงปรีชากุล

วิทยาลัยนานาชาติปรีดี พนมยงค์ มหาวิทยาลัยธรรมศาสตร์ กรุงเทพฯ 10200
อีเมล: fengjialao3@hotmail.com

บทคัดย่อ: บทความนี้ได้รวบรวมคำพ้องความหมายระหว่างภาษาจีนในเอเชียตะวันออกเฉียงใต้กับภาษาจีนกลางจำนวน 900 คำจากฐานข้อมูล “พจนานุกรมภาษาจีนรอบโลก”(全球华语词典) จากนั้นได้ทำการศึกษาเปรียบเทียบเชิงปริมาณจากหำมิติ ได้แก่ประเทศและพื้นที่ที่ใช้คำพ้องความหมาย จำนวนของคำพ้องความหมาย ที่มาของคำพ้องความหมาย ประเภทและจำนวนพยางค์ของคำพ้องความหมาย ขอบเขตความหมายของคำพ้องความหมาย จากการศึกษาพบว่า ประเทศที่ใช้คำพ้องความหมายมากที่สุดคือสิงคโปร์ มาเลเซีย และไทย คำพ้องความหมายส่วนใหญ่สัมพันธ์กันแบบ 1ต่อ1 แต่คำพ้องความหมายที่ต้องระมัดระวังเป็นพิเศษคือคำที่มีสัมพันธ์แบบ 1 ต่อ M(Many) คำพ้องความหมายที่มาจากภาษาต่างประเทศส่วนใหญ่มาจากภาษาอังกฤษ ภาษาไทยและภาษามาลายู ส่วนคำพ้องความหมายที่มาจากภาษาถิ่นของจีนส่วนใหญ่มาจากภาษาแคว้นต้ง ภาษาฮกเกี้ยน ภาษาเจียงซู(อุ) ในด้านประเภท คำพบว่าคำพ้องความหมายส่วนใหญ่เป็นคำนาม ในด้านจำนวนพยางค์พบว่าส่วนใหญ่มีสองพยางค์ ในด้านขอบเขตความหมายพบว่าคำพ้องความหมายส่วนใหญ่เป็นคำที่หมายถึงสิ่งของที่เป็นรูปธรรม สิ่งที่เป็นนามธรรม สิ่งมีชีวิต และกิจกรรมทางสังคม

คำสำคัญ: ภาษาจีนในเอเชียตะวันออกเฉียงใต้; ภาษาจีนกลาง; คำพ้องความหมาย

Quantitative and Comparative study on the Synonym between Southeast Asian Chinese and Mandarin

Song Fei

Faculty of Arts, Beijing International Studies University, Beijing, 100024, P.R. China

Email: bestsongfei@126.com

Kiattisak Fongprechakun

Pridi Banomyong International College, Thammasat University, Bangkok, 10200, Thailand

Email: fengjialao3@hotmail.com

Abstract: Based on the "Global Chinese Dictionary", we establish a database and extract 900 synonyms between Southeast Asian Chinese and Mandarin through the study on the definition of Exclusive Words, Homonyms, Synonyms and the relationship between them. Then we conduct a Quantitative and Comparative study on those words from the dimensions of countries and regions, the Quantity of Synonyms, the source from foreign language and Chinese dialect, the lexical category and word length, and the semantic categories. The study found that most synonyms are used in Singapore, Malaysia and Thailand; most synonyms are in one-to-one relationship; the main foreign language sources of Synonyms are English, Thai, Malay, and the main Chinese dialect sources are Guangdong dialect, Fujian dialect, and Wu dialect; most Synonym are nouns and disyllabic words; most Synonym belong to the concrete matter, the abstraction, the living beings and the social activities.

Keywords: Southeast Asian Chinese; Mandarin; Synonyms

东南亚华语与汉语普通话异名同实词汇的量化比较研究*

宋飞

北京第二外国语学院汉语学院，北京，100024，中国

电子邮箱：bestsongfei@126.com

冯忠芳

泰国法政大学比里·帕依荣国际学院，曼谷，10200，泰国

电子邮箱：fengjialao3@hotmail.com

摘要：本文基于《全球华语词典》建立数据库，通过界定异名同实词的概念及其外部关系，提取出了 900 个东南亚华语与汉语普通话异名同实词，并从国别与区域分布、数量情况、外语及方言词源、词类与词长、义类范畴五大维度，对这些异名同实词进行了量化比较研究。研究发现，新加坡、马来西亚、泰国的异名同实词数量最多；大部分异名同实词都是一对一的关系，但需要特别注意一对多的关系；异名同实词的外语来源主要为英语、泰语、马来语，汉语方言来源为粤方言、闽方言、吴方言；从词类来看，异名同实词中名词占主体，从词长来看，异名同实词中双音节词占主体；异名同实词的义类范畴主要集中于具体物、抽象物、生物、社会活动。

关键词：东南亚华语；普通话；异名同实词汇

* [收稿日期] 2017-03-05

本研究受国家社科基金重点课题“东南亚汉语传播的国别比较研究”（14AYY011）项目资助。

一、研究背景

词汇差异是华语变体间的主要差异之一，表现为异名异实、同名异实、异名同实三种情况，这三种情况分别对应于特有华语词、异名同实词、同名异实词三类。这三类词一方面是当地独特的自然、社会因素在语言中的反映，另一方面也构成了不同华语社区间交流沟通的显著障碍。因此十分有必要对各地特有华语词、同名异实词、异名同实词进行量化比较研究。

此前针对东南亚华语词汇的研究主要集中在东南亚华语的汉语方言词源研究（刘新中，韩慧玲，2008）和东南亚某一特定国家华语词汇的研究（刘全惟，2013）。此外，王世凯（2012）对全球华语中异名同实词语的语源、趋同性进行了统计，但并没有针对东南亚华语的统计。本人前期对东南亚特有华语词、东南亚华语与汉语普通话同名异实词进行了量化的对比研究，界定并提取出了东南亚特有华语词 1183 个，同名异实词 174 个（宋飞，2016）。目前还没有针对东南亚华语与汉语普通话异名同实词汇的研究。因此，迫切需要对这部分词汇进行提取和分析，以打造东南亚华语与汉语普通话间词汇差异研究的完整链条。

东南亚是全球华人、华侨最集中的地区之一，包括泰国、缅甸、越南、柬埔寨、老挝、印度尼西亚、文莱、新加坡、马来西亚、菲律宾、东帝汶（下文将分别简称为“泰、缅、越、柬、老、印、文、新、马、菲、东”）十一个国家。中国与东盟互为重要贸易伙伴，双方政治、经济、文化交流日益密切。因此，建成东南亚华语与汉语普通话异名同实词库，对东南亚各国与中国之间的词汇差异进行量化比较研究，对扫清双方交流障碍具有重要意义。

二、异名同实词汇的界定及其外部关系

本文认为，特有华语词、同名异实词、异名同实词是不同华语社区间的三种差异词汇类型。本文将通过界定特有华语词、同名异实词、异名同实词的概念，并理清三者间的关系，来研究东南亚华语与汉语普通话异名同实词汇的提取方法，进而提取出这部分词汇，建成东南亚华语与汉语普通话异名同实词库。

东南亚特有华语词指的是以特定语音、书写形式映射东南亚特有的概念形成的词。各华语社区特有的自然、人文环境会催生出其他华语社区没有的概念，即所谓的“异实”，以特定的语音和书写形式（“异名”）与之发生映射，就形成了特有华语词。所以如果要跟同名异实词、异名同实词相对应，特有华语词在本质上也可为“异名异实”词，只是人们习惯上并不这么说。例如，“巴冷刀”是新、马、印、文对当地一种刀的称呼，“峇迪”是新、马、泰、印、文当地一种蜡染的印花布，“白马”是新加坡特有的对父母有权势、有地位、有影响力的现役军人的称呼，“卜基”是新、马对非法接受赛马等赌注的人的称呼，等等。以上概念在汉语普通话均没有与之完全对应的词汇，是东南亚相对于中国大陆特有的概念。

同名异实词指的是不同华语社区间语音、书写形式相同，所指概念不同的词。东南亚华语与汉语普通话的同名异实词指的就是东南亚和中国大陆之间词形、发音相同，但词义不同的词。例如：“爱人”在新、马、泰指恋人、情人（如“这部剧里的欢喜冤家原来是一对爱人。”），在中国大陆则指丈夫或妻子（如“上周我陪爱人回娘家看望她的父母了。”）；“白金”在新、马、泰、印指吊丧时给丧家的礼金（如“法师逝世后，近日来善男信女所送的白金已有 34 万元。”），在中国大陆本指贵金属，也用来指很高的价值、

等级等（如“林先生长期以来都是我们俱乐部的白金会员。”）；“单位”在新、马、泰指一套单元楼房（如“这座公寓每一层有四个单位。”），在中国大陆则通常指机关、企业、团体或属于一个机关、企业、团体的各个部门（如“企业单位应当为职工办理住房公积金。”），等等。

本文中的异名同实词指的是不同华语社区用来表述同一概念，但语音或书写形式不同的词，在相关研究中又被称为异名同物词、同物异名词、异名同指词、异名词等。东南亚华语与汉语普通话异名同实词则指东南亚华语中，针对同一概念与汉语普通话表述方式不同的词。异名同实词又分为异写型词（如新、马、泰的“雷射”和中国大陆的“镭射”，新、马、泰、印的“榴槿”和中国大陆的“榴莲”等）和异说型词（如新、马的“报生纸”和中国大陆的“出生证”，新、马、泰的“奇异果”和中国大陆的“猕猴桃”等）。其中，异写型词因为彼此发音相同，使用的字形也往往较为接近，造成的沟通障碍通常较小。而异说型词因为发音不同，甚至词的结构形式也各异，往往造成较为显著的沟通障碍。

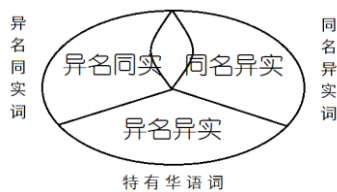


图 1 华语社区间差异词汇构成示意图

如图 1 所示：特有华语词、同名异实词、异名同实词分别对应了不同华语区词汇差异的三种情况：异名异实、同名异实、异名同实。三者互相补充，共同构成了不同华语社区间的差异词汇。同名异实词和异名同实词间有时存在交叉。如“代差”一词，在新加坡华语和汉语普通话中都存在，但在新加坡指年轻一代与老一代在思想方法、价值观念、生活态度、兴趣爱好方面存在的心理距离或心理隔阂（如“由于社会、文化背景的代差所衍生出的对管教子女时的莫衷一是，常会为时下的父母带来极大的困扰。”），而在中国大陆则指年代差异（如“该国军方内部取得默契，对外出售武器将始终维持 10 年以上代差。”），形式相同，意义不同，因此属于两个华语区的同名异实词。但新加坡华语的“代差”和汉语普通话的“代沟”二词虽然形式不同，表达的含义却是相同的，因此二者又互为异名同实词。把握好这种关系对于提取东南亚华语与汉语普通话异名同实词具有重要意义。

三、东南亚华语与汉语普通话异名同实词汇的提取

（一）数据准备

本文基于《全球华语词典》（李宇明，2010）建立了数据库，以此作为提取东南亚华语与汉语普通话间的异名同实词汇的基础。《全球华语词典》（下称“《词典》”）收词近万条，记录了各华语社区的词汇使用情况。每个词条包含“词目、释义、例句、使用地区、异名词语、知识窗”等几个方面内容。基于《词典》中词条的文本特点，本文设定的数据库结构包括“ID、词、词长、译法（外来词）、来源（外来词）、异名词语、使用地区、释义、词类、词类数量、义项数量、知识窗”12个字段，并将《词典》中的文本数据化后存入库中。库中的“异名词语”字段的数据来源于《词典》本身，包

含部分异名同实词汇信息。但这部分信息并非专门只针对东南亚华语与汉语普通话间的异名同实词，因此很多信息与本文研究无关。此外，除了《词典》中显示有异名同实词的词条，其余词条中也大量存在异名同实词，但《词典》并未进行说明。因此在提取过程中并不仅以《词典》中的“异名词语”信息为依据，而是在此基础上进一步人工标注，最终得到完整的异名同实词集合。

在东南亚 11 个国家中，东帝汶独立较晚（2002 年），语料较缺乏。因此，《全球华语词典》吸收了除东帝汶之外的其他十国（也是“东盟”十国）的华语词汇，可以较为全面地反映东南亚华语词汇使用情况。此外，由于本文涉及的词汇均来自于《词典》，因此词汇的界定均直接采用《词典》的意见。《词典》主要收录了 20 世纪 80 年代以来各华人社区常见的特色词汇，即各华人社区间的差异词汇，也收录了少量共有词汇。

（二）东南亚华语与汉语普通话异名同实词汇的提取步骤

由东南亚华语与汉语普通话特有华语词汇、同名异实词汇、异名同实词汇关系可知，异名同实词汇是这两个华语社区间差异词汇的一部分。要提取异名同实词汇，必须先提取东南亚华语与汉语普通话间的差异词汇。提取步骤如下：

第一，在基于《词典》建立的数据库中，先按“使用地区”字段筛选出包含“新、马、泰、老、缅、越、柬、文、印、菲”中任一国名缩写的词，这些词中既包含了使用范围仅限东南亚的特有华语词，也包含了东南亚华语与汉语普通话间的同名异实词和异名同实词，以及东南亚华语与汉语普通话共同使用，但并非全球通用的词汇。筛选结果包含 2734 个词。

第二，按“使用地区”字段排除东南亚华语与汉语普通话共同使用的词，保证剩余词汇在两个华语社区间的“异名”性。这时要注意异名同实词和同名异实词的交叉关系，如果一个词有多个义项，若有任一义项使用范围不包括中国大陆，则认为该词是东南亚华语与汉语普通话间的异名同实词，如上文所述，这种特殊的异名同实词也是同名异实词。如“贴士”，虽然在“提示”的义项上同时使用于东南亚和中国大陆，但“小费”义项只在东南亚使用。因此在第二个义项上，东南亚的“贴士”和中国大陆的“小费”属于异名同实词。与此同时，“贴士”在东南亚可指小费，但在中国大陆则指提示，又是同名异实词。此类异名同实、同名异实的交叉词共有 9 个，包括“代差、饭盒、公车、黑油、看衰、老爷、龙头、上手、贴士”。此次筛选结果包含 2121 个词。

第三，按“释义”字段人工标注部分异名同实词。人工标注的标准为：第一，若释义为单个词，则认为该释义词是被释词的异名同实词，如用于新、马的“表表者”的释义为“佼佼者”，则认为“佼佼者”是“表表者”的异名同实词。第二，若释义形式为“义同‘……’”，也认为“义同”后的词是被释词的异名同实词，如“巴士道”的释义为“义同‘公交车道’”，则认为“公交车道”是“巴士道的异名同实词。第三，若释义形式为“‘……’的简缩”，也认为“的简缩”前的词是被释词的异名同实词。如“按”的释义为“‘按揭’的简缩”，则认为“按揭”是“按”的异名同实词。由此对 2121 个词添加了 1151 个异名同实词标注。

第四，将 1151 个新加异名同实词标注数据合并到“异名词语”字段。原“异名词语”字段中没有数据，新加数据来自于人工标注的词共有 518 个。

合并规则为：若新加数据在“异名词语”字段已经存在，则不再进行合并；

若新加数据在“异名词语”字段不存在，则合并至该字段数据中。如“阿兵哥”的“异名词语”字段数据为“兵哥（台湾、新马）”，新加的异名同实词“军人”在原数据中不存在，则合并到“异名词语”字段数据中，成为“兵哥（台湾、新马）、军人（中国大陆）”。

第五，在数据合并后的“异名词语”字段，筛选出使用地区包括“大陆”的数据。筛选后共得到 900 个词，这 900 个词可认为是最终得到的东南亚华语与汉语普通话间的华语异名同实词。由此建成包含 900 条数据的“东南亚华语与汉语普通话异名同实词库”。

四、东南亚华语与汉语普通话异名同实词汇的比较研究

（一）国别与区域分布

基于“东南亚华语与汉语普通话异名同实词库”，本文将对东南亚华语与汉语普通话异名同实词汇的国别与区域分布情况进行量化比较研究。

从国别分布情况来看，泰国华语与汉语普通话异名同实词最多，达到 612 个（如“昂场”和“会场”、“波立”和“警察”等）；其次是新加坡的 583 个（如“插车”和“叉车”、“公坟”和“公墓”等）和马来西亚的 515 个（如“路墩”和“减速带”、“停表”和“秒表”等），三个国家的异名同实词汇数量相加占全部数量的 91.69%。因为一个词可能同时在几个国家使用，因此各国使用的异名同实词总数相加大于 900。

表 1 东南亚华语与汉语普通话异名同实词汇国家分布情况

使用地区	泰	新	马	印	文	柬	越	菲	緬	老
数量	612	583	515	90	32	15	10	4	3	1
比例	32.82%	31.26%	27.61%	4.83%	1.72%	0.80%	0.54%	0.21%	0.16%	0.05%

以上数据表明，泰、新、马三国华语与汉语普通话在表达同一概念时词汇差异最大，这三个国家的历史、自然、人文环境造就的当地华语词汇变体数量最多。

从区域分布情况来看，基于数据的区域划分类别有 26 个。这 26 个区域指共同使用某一异名同实词的若干国家组成的整体。数据显示，新、马、泰地区作为一个区域整体，其华语与汉语普通话间的异名同实词数量最多，达到 228 个；其次是新、马地区作为一个区域整体，异名同实词数量达到 166 个，二者相加占全部异名同实词数量的 79.12%。

表 2 东南亚华语与汉语普通话异名同实词汇前 10 个区域分布情况

使用地区	新马泰	新马	新马泰印	新马印	新泰	新马泰印文	马泰	新马印文	新马泰印文柬	新马泰印文菲越
数量	228	166	32	17	14	13	13	8	4	3
比例	45.78%	33.33%	6.43%	3.41%	2.81%	2.61%	2.61%	1.61%	0.80%	0.60%

这说明近一半东南亚华语与汉语普通话异名同实词使用于新、马、泰地区，该地区作为一个区域整体，与中国大陆间的华语异名同实词最多。该结果与之前国别分布情况的统计结果相印证，说明新、马、泰三国无论作为单独国家，还是作为区域整体，其异名同实词汇最多，与中国大陆在表达同一概念时差异最大。

（二）数量情况

不同华语社区的异名同实词之间会形成一对一、或一对多的关系。针对汉语普通话中的单个词，考察其在东南亚华语中异名同实词的数量情况，从而定位到异名同实词数量较多的那部分词。这部分普通话词汇在东南亚华语中变体很多，给基于其概念的交流增加了很多障碍，值得特别注意。

经统计，汉语普通话词汇在东南亚华语中的异名同实词数量从 1 至 6 个不等。其中 68.67% 的词异名同实词数量为 1 个，占大多数（如大陆的“开发商”和新、马、泰、印的“发展商”、大陆的“甩站”和新、马的“飞站”等）；16.33% 的词异名词数量为 2 个（如大陆的“下班”在新、马的异名词为“放工”和“收工”，大陆的“敞篷车”在新、马、泰为“开篷车”，在新、马又叫“敞车”）；7.67% 的词异名词数量为 3 个（如大陆的“到达”在新、马、泰有“抵埗、抵步、抵埠”3 个异名词，又如大陆的“东盟”在马、印被称为“东合”，在马、泰被称为“东协”，在新加坡又被称为“亚细安”）；4.22% 的词异名词数量为 4 个（如大陆的“方便面”在泰国称为“即食面、速食面”，在新、马、泰称为“快熟面”，在新加坡称“泡面”，）；1.78% 的词异名词数量为 5 个（如大陆的“摩托车”在泰、越被称为“机车、摩托机车”，在新、马、泰、印被称为“电单车”，在新、马又叫“摩哆西卡”，在泰国被称为“摩托西”）；1.33% 的词异名词数量为 6 个（如泰国的“冬养贡”，又称“冬荫功、冬央贡、辣供汤”，在新、马称作“冬炎、东炎”，在大陆则被称作“酸辣汤”）。

表 3 东南亚华语与汉语普通话异名同实词汇数量分布

异名同实词数量	1	2	3	4	5	6
词汇数量	618	147	69	38	16	12
比例	68.67%	16.33%	7.67%	4.22%	1.78%	1.33%

由此可见，以汉语普通话向东南亚华人传达特定概念，以及接收、理解东南亚华人传达的概念时，要特别注意这些异名同实词，特别是数量达到 4 个、5 个甚至 6 个的异名同实词，这些词虽只占整体的不到 10%，却是极易造成误解的词。

（三）外语及汉语方言词源分布

东南亚华语与汉语普通话异名同实词的来源大致可分为三类，一类来自于当地的语言生活，是当地特有的历史、自然、人文环境产生的，这部分词汇是异名同实词汇的主体；另外两类则分别来源于外语和汉语方言，分别形成了东南亚华语词汇中的外来词和汉语方言词。

1. 外来词源分布

经统计，900 个异名同实词中共有 128 个词属于外来词，可以找到明确的外来词源，占全部异名同实词数量的 14.22%，略高于东南亚差异词汇整体的外来词比例 11.96%（宋飞，2016）。外来词的来源包括英语（如“雅痞士、恤衫”等）、泰语（“突家、娘少”等）、马来语（如“甘榜、隆帮”等）、日语（如“卡哇伊、一级棒”）、印地语（如“布都”）、葡萄牙语（如“嘟嘟”）、法语（如“卡叻”），共 6 种。

表 4 东南亚华语与汉语普通话异名同实词汇外来词来源分布

来源	英语	泰语	马来语	日语	印地语	葡萄牙语	法语
词汇数量	91	21	11	2	1	1	1
比例	71.09%	16.41%	8.59%	1.56%	0.78%	0.78%	0.78%

表 4 数据显示，外来词最主要的来源是英语、泰语、马来语，来自三种语言的外来词数量相加，占全部异名同实词的 96.09%。外来词中又包括音译词 125 个，如“部落格”（网络日记，英语 blog 的音译）、“叻叻”（市场，一般指菜市场，泰语音译）、“基罗”（公斤；千克，英语 kilo 的音译）、“銮抱”（称呼佛寺师父；方丈，泰语音译）、“慕丝”（一种用明胶凝结乳酪和鲜奶油制成的奶油冻，英语 mousse 的音译）、“惹兰”（街道；道路；小巷；小路，马来语音译）、“亚曾”（老师。泰语音译）等等；音译兼意译词 3 个，包括“倾斗罗厘”（自卸卡车，英语 tip lorry 的音译兼意译）、“吴哥寺”（柬埔寨著名寺院，英语 Angkor Wat 的音译兼意译）、“一级棒”（最好；最优秀；最高级，日语音译兼意译）。

2. 汉语方言词来源分布

除 128 个外来词，异名同实词中还有 94 个词来源于汉语方言，占全部异名同实词数量的 10.45%，比例略低于外来词。其中包括来自粤方言的 55 个，如“落货”（卸货。落，粤方言指“下（到处所）”）、“衰仔”（兔崽子，浑小子。衰，粤方言指坏、缺德）；来自闽方言的 37 个，如“补梃”（补胎。梃，闽南方言指橡胶）、“厝边”（邻居。厝，闽南方言指房子、屋子）等；来自吴方言的 2 个，如“放鸽子”（集体活动中故意丢下旁人不管，使受到冷落）、“写字楼”（办公楼），共 3 个来源。

表 5 东南亚华语与汉语普通话异名同实词汇中的汉语方言词来源分布

来源	粤	闽	吴
词汇数量	55	37	2
比例	58.51%	39.36%	2.13%

在 3 个汉语方言词来源中，来自粤方言和闽方言的异名同实词所占比例最高，二者相加达到 97.87%，占绝大多数。其中来自闽方言的词中，绝大多数又来自闽南方言。

因此从东南亚华语与汉语普通话异名同实词汇的外语及汉语方言词源来看，外来词主要来源于英语、泰语、马来语，汉语方言词主要来源于粤方言和闽方言。

（四）词类与词长

1. 词类分布

从词类来看，东南亚华语与汉语普通话异名同实词汇包括名词（如“大假、单形浦”等）、动词（如“倒账、电发”等）、形容词（如“耳轻、激心”等）、副词（如“间中、好得”等）、量词（如“丘、仙”等）、代词（如“他者、乜”）、助词（如“到”，用在动词后面，相当于“得”）、介词（如“透过”，和中国大陆的“透过”词义有所区别），共 8 种。

其中，名词数量最多，占全部词汇数量的 74.00%，其次是动词和形容词，分别占全部词汇数量的 19.78%和 7.00%。从东南亚华语与汉语普通话的差异词汇整体来看，三者的比例分别为 76.12%、19.90%、3.95%（宋飞，2016），因此，异名同实词汇的词类分布与差异词汇整体分布情况类似。

此外，部分异名同实词属于兼类词（如“守阁”是动词和名词兼类，分别指看门和看门的人；“甘”是形容词和量词兼类，分别指“甜”和重量单位“克”。），因此各词类的词汇数量之和多于 900 个。

表 6 东南亚华语与汉语普通话异名同实词汇的词类分布

合并词性	名,	动,	形,	副,	量,	代,	助,	介,
词汇数量	666	178	63	13	10	2	1	1
比例	74.00%	19.78%	7.00%	1.44%	1.11%	0.22%	0.11%	0.11%

2. 词长分布

从词长来看，东南亚华语与汉语普通话异名同实词从单音节到八音节不等。如单音节词“番、承”等，双音节词“甘榜、福物”等，三音节词“风化区、咖喱啡”等，四音节词“骨痛热症、好马人物”等，五音节词“已开发国家、银角电话机”等，六音节词“天盖建筑系统、工商保安警察”等，七音节词“电动行人输送带、健力士世界纪录”，八音节词“生物认证科技护照”。本文将三音节至八音节统一归并为多音节。其中，双音节词汇数量最多，达 55.67%，其次是多音节词，达 35.33%，最少的是单音节词，占 9.00%。从东南亚华语与汉语普通话的差异词汇整体来看，三者的比例分别为 52.96%、42.50%、4.54%（宋飞，2016），异名同实词与之相比，单音节和双音节词所占比例略高，多音节词所占比例略低，平均词长略短。

表 7 东南亚华语与汉语普通话异名同实词汇的词长分布

词长	单音节	双音节	多音节
词汇数量	81	501	318
比例	9.00%	55.67%	35.33%

(五) 义类范畴

为考察东南亚华语与汉语普通话异名同实词产生在哪些领域，涉及到哪些社会存在和社会意识，本文对其义类分布进行了考察。考察方法为，根据《现代汉语分类词典》（苏新春，2013）中的 9 个一级义类划分，对 900 个东南亚华语与汉语普通话异名同实词进行义类标注，进而对这些词的义类分布进行统计和考察。

经统计，东南亚华语与汉语普通话异名同实词汇在 9 个一级义类范畴都有分布。如表具体物的“山芭鸡、霜榭”等，表抽象物的“位阶、饷当”等，表生物的“相士、心贤”等，表社会活动的“桌诈、义标”等，表性质与状态的“中庄、标青”等，表生物活动的“呷醋、食风”等，表辅助词义类的“络续、好得”等，表运动与变化的“找欠、直透”等，表时空的“部域、港乾”等。

表 8 东南亚华语与汉语普通话异名同实词汇的义类分布

义类	具体物	抽象物	生物	社会活动	性质与状态	生物活动	辅助词	运动与变化	时空
词汇数量	275	236	147	121	56	29	13	12	11
比例	30.56%	26.22%	16.33%	13.44%	6.22%	3.22%	1.44%	1.33%	1.22%

按照所占比例从高到低，9 个义类的顺序为：具体物—抽象物—生物—社会活动—性质与状态—生物活动—辅助词—运动与变化—时空。按照该顺序，《现代汉语分类词典》中 9 个义类的比例分别为：16.38%、20.07%、11.90%、13.19%、11.57%、13.75%、2.95%、6.34%、3.86%。可见，与《现代汉语分类词典》中收录的汉语词汇整体相比，东南亚华语与汉语普通话异名同实词中表具体物的词所占比例明显很高，几乎是前者的两倍；表抽象物、

生物、社会活动的词所占比例也都较高；表性质与状态、生物活动、辅助词、运动与变化、时空的词所占比例普遍较低。

数据对比可知，东南亚华语与汉语普通话异名同实词主要分布于具体物、抽象物、生物、社会活动这些义类，特别是具体物义类。这反映出东南亚华语与汉语普通话的词汇差异主要集中于描述事物的词，特别是描述具体事物的词。

五、结语

本文基于《全球华语词典》建立数据库，通过界定特有华语词、同名异实词、异名同实词的概念及其关系，提取出 900 个东南亚华语与汉语普通话异名同实词汇，并从国别与区域分布、数量情况、外语及方言词源、方言词源、词类与词长、义类范畴五个维度，对这些异名同实词汇进行了量化的比较研究。由此发现：新、马、泰无论单独作为国家还是作为一个整体区域，其与中国大陆间的异名同实词数量都是最多的；大部分异名同实词都是一对一的关系，但需要特别注意一对多的关系；异名同实词的外语来源主要为英语、泰语、马来语，汉语方言来源为粤方言、闽方言、吴方言；从词类来看，异名同实词中名词占主体，从词长来看，异名同实词中双音节词占主体；异名同实词的义类范畴主要集中于具体物、抽象物、生物、社会活动。

以上研究结果有助于摸清东南亚华语与汉语普通话异名同实词汇的整体情况，打破东南亚华语与汉语普通话语言交际中的词汇沟通障碍。研究过程中建设的“东南亚华语与汉语普通话异名同实词库”也为东南亚华语的信息化处理提供了重要条件，在机器翻译和舆情分析领域都可加以利用。

参考文献

- 郭熙. 2002. 普通话词汇和新马华语词汇的协调与规范问题——兼论域内外汉语词汇协调的原则与方法. 南京. 南京社会科学。
- 郭熙. 2010. 新加坡中学生华语词语使用情况调查. 华文教学与研究。
- 李慧. 2005. 从《两岸现代汉语常用词典》看两岸的同实异名词语. 修辞学习。
- 李如龙. 2013. 海外汉语方言研究的新视野——读《全球华语词典》. 辞书研究。
- 李宇明. 2010 《全球华语词典》. 北京：商务印书馆。
- 刘全惟. 2013 《马来西亚华语词语使用情况调查》. 导师：李红印. 北京大学。
- 刘文辉、宗世海. 2006 《印度尼西亚华语区域词语初探》. 暨南大学华文学院学报。
- 刘新中、韩慧玲. 2009 《闽粤客三大方言对东南亚华语语音词汇的影响例说》载《首届海外汉语方言国际研讨会论文集》. 暨南大学汉语方言研究中心。
- 邱克威. 2014 《叻报》的词语特点及其词汇学价值管窥. 语言研究。
- 施春宏. 2015 《从泰式华文的用词特征看华文社区词问题》载《语文研究》。
- 宋飞. 2016 《东南亚特色华语词汇的区域和国别比较研究》载《语言文字应用》。
- 王世凯、方磊. 2012 《全球华语词典》中异名词语的调查分析. 语言文字应用。
- 周荐. 1997 《异名同实词语研究》载《中国语文》。