

基于《日本语源大辞典》的汉源日词历时分析

李 斌 陈 静

(南京师范大学 文学院,江苏 南京 210097)

摘 要:汉源日词是语言学的重要研究对象,能够揭示出汉字词对日语词汇发展的影响。本文基于日本出版的《日本语源大辞典》中所收录的6000多个词语,人工录入和标注词条的年代、词性、汉字比例和与汉词词义相似度等信息建立数据库。统计得出,在词典的所有词项中,词语出现数量最多的朝代分别是奈良时代、平安时代以及江户时代。同时,词典中能写成全汉字的词占一半以上,而包含汉字的日词则占词典的九成以上。最后,从社会历史方面探讨新词产生及汉字词比例涨落的原因。

关键词:借词 汉源日词 历时分析

中图分类号:H36 **文献标识码:**A **文章编号:**1008-9853(2017)03-177-07

汉源日词是指在日语中来自汉语的词语,属于日语外来语的一种。汉源日词在中国和日本都是重要的研究对象,有利于厘清日语词汇的发展史和中日两国语言文化交流的历史。近年来,大陆学者在研究汉源日词时,多偏重于使用定性方法的研究(汪维辉,2009),如研究某个具体时代引入汉语的原因(蔡凤林,2016)、汉语词汇在引入日语后发生的字形、语义变化(史有为,2005)等。但是这些研究缺乏整体性和大数据量的支撑。特别是日语词到底在哪些时代从中国借用了多少词语,一直模糊不清。

为了提供资料作论据,我们利用了日本出版的《日本语源大辞典》一书,对当中所收录的6000多条词语进行标注及分类,针对以下三个问题进行了统计分析:(1)观察词典内词汇的时代分布状况;(2)统计词典内词条的汉字构成比例,以得出汉字在日语词汇中的地位;(3)通过日语词与汉语源词的意义相似度分析,进一步确定每个时代引进汉字词和自造日语词的情况。据此,我们对汉源日词的整体状况和历时演进情况进行描写和讨论。

一、汉源日词的研究现状

对汉源日词的研究属于外来语研究的一部分,有助于中日两国的历史语言文化交流,一直受到国内外学者的关注。新村出(1995)肯定了日语中汉源词存在的积极意义并指出:“即使在当代,汉源词也是形成新词汇和构词的重要数据和原料。”国内学者从不同的角度对汉语中日语借词进行研究。从历史角度来说,蔡凤林(2016)指出,中国历史对日本的记载最早是在《汉书·地理志》中,汉语带来的影响主要是以中国文化的渗透为基底或先导。历史上汉语对日语的影响也是以中国儒释文化传入日本为重要前提,自538年佛教传入日本后迅速普及,为汉字在古代日本社会大显身手创造了优越条件。而潘钧(2000)尝试追溯日本辞书的发展历史,发现新代以前的日本辞书的编纂也是在中国辞书的影响下发展起来的。以完成于平安中期的《新撰字镜》为例,它是现存最早及初步具备了汉和辞典特征的重要辞典,所收录的词汇有大量的“国字”出现,反映了当时汉字在日语词汇中的地位。

在文字方面,史有为(2005)指出,日本所用汉字不但基本上来自汉语和中国汉字,而且从属于中国

收稿日期:2017-05-03

作者简介:李斌,男,1981年生,江苏徐州人,南京师范大学文学院副教授。陈静,女,1995年生,香港人,南京师范大学文学院学生。
基金项目:教育部社科青年项目(16YJC740034),山东省语言资源开发与应用重点实验室开放课题。

的传统汉字系统。可大略分为四部分：中国传统正体汉字、中国传统汉字的异体字、日本根据中国传统汉字或中国异体自行发展出的异体字、日本根据本国情况按照汉字六书原则自行创制的汉字。其中，前三者占日本汉字的97%以上。陈红(2002)指出，日语使用汉字的历史悠久且数量众多，日本学习研究社编写的《学研汉和大辞典》就收录了1.1万个汉字。飞田良文、菅原义三编著的《国字辞典》中收录了1553个日本人根据社会生活需要模仿汉字的造字规律而创制出来的“国字”。日语国字绝大多数是会意字，约占9成左右。汪维辉(2009)以日语中的汉语字词作例，探讨了日语中汉字借词的情况。分别是：借字借词保存古义；传入的多数是文化词，固有词一般不易更改；日语借词可反映汉语的基本词历史更替；以及保留古时字形。

在语音方面，李莉(1998)指出在日语中的汉语词(日本多称之为汉语)的概念是指过去由中国传入成了日本语的词语，以及被广泛使用后模仿汉语读音的词语。汉语词的意义一是由汉字组成，二是用汉语音读。对此，李无未(2005)也有提及日本的汉音是来自唐代的长安音的，与《切韵》音系有一些细微的差别，基本上可以认为《切韵》音系与日本的汉音音系是相等的。

日本也有不少汉语借词的研究。桃井惠一(2006)曾以《汉语外来词词典》为例，汉语的10272个外来词中只有861个(8.38%)是来自日本的，而在这些源自日本的词语中，由日本产生的有671个，剩下的190个皆是来自中国古籍，说明历史上汉字词对日语词汇的影响较大。铃木俊二(2005)研究了进入琉球方言的汉词主要是在14世纪左右因交易传入的，这些词包含了动植物、饮食、生活、艺术、文化等多方面。

综上所述，可见目前国内外对汉源日词已经有了较多研究，但在历时量化分析上仍有不足。为了进一步探究汉语对日词带来的影响，本文从各时代的日词产生情况来推断汉语对日语的影响过程。

二、语料来源和说明

为了了解历史上各时代的日词产生的数据，最好的办法就是利用大而全的日语历时语料库。然而这样的语料库的缺位，使得学者们暂时无法进行历时统计分析。本文退而求其次，依据日本出版的高质量历时语源词典——日本小学馆株式会社出版的《日语源大词典》进行统计分析。该书一共收录6013个日语词及7067个义项，包含动植物、食物、服饰、道具、动词、形容词等多类型的词语，每一个词条都标有词性、词义、造词理据、最早出现的正式文献名称及出版年份，征引的文献多达850部，从古时的风土记至近代的作品，包括现代的部分辞典。这部词典的数据可信性及准确性较高，能够帮助我们对词汇最早出现的时间进行界定并进行历时分析。

为了便于量化分析，我们对该词典建立了电子数据库。首先，我们对词典内所有的词条进行了录入和信息标注。除了词典已有的词义、最早出现年代和词性外，人工补充标注了每个词条的最早出现年代对应的日本朝代，便于更直观地从历史朝代上进行分析。然后对每个词条进行汉字比例的标注，分成全汉字、半汉字和全假名三类。对与汉语的词义相似度标注分成全假名、完全相同、部分相同、完全不同和日语独有词五类。这样就形成了信息较为丰富的日语词汇语源数据库，每个词条的信息参见表1。在这个数据库的基础上，可以统计出词典内的所有词条的时代分布、汉字所占比重及与汉语词典的相似度等分布信息。然后进行汉源日词的资料分析，可以更客观地得到汉源日词的分布数据，同时也方便利用图表查看每个词条的时代、词性、汉字比例及与汉语词义相似度等。

表1：词条录入和标注示例

词典原有信息	词	呜呼		
	词义	用于表示对某种事物感到惊讶、悲伤、欢喜、疑问等情绪。		
	最早出现年份	720	词性	叹词
人工标注	对应朝代	奈良时代(710~794)		
	汉字比例	全汉字词		
	与汉语词义相似度	部分相同		

需要说明的是，在下文的统计中词条、义项(词项)的数量会有所不同。词典共收录6013个日语词

及 7067 个义项,词汇的时代分布统计时去掉了缺少时代信息的词条,剩余 6626 个词条。在对比日汉语相似度时,只抽取当全汉字及全假名的词条做比对,一共只有 4447 个词条。

三、汉源日词的历时分析

通过对这本《日本语源大词典》进行录入及标注,我们可从中获得日语词的各种数据分布。首先,为了统计源自汉语的日语词汇,我们需要用这个数据库对日词进行整体性的历时分析,用以推断汉源日词的历时借用情况。为了更好地观察日语词的历史朝代分布情况,我们按照日本及中国双方的朝代年份进行了对照(表 3)。进而从日词产生的时代分布信息、汉字的构成比例和与汉语的意义相似度三方面来推断出日词与汉语的关联性。

(一)日语词汇的产生时代分布。时代分布是当中最为直接的角度。时代信息是词典中原有的信息,准确地指出了每个词条最早出现的年份,因此只要按照中日双方的朝代进行对照,便能统计出所有词条的时代分布。根据数量较突出的时代进行探究,就能发掘其背后与汉语的关联。图 1 给出了词典内 6000 多个词条的时代分布状况,新词数量最多的顺序分别是奈良时代、平安时代以及江户时代,分别占到词条总数的 35% (2320 个)、28% (1846 个)和 21% (1386 个)。

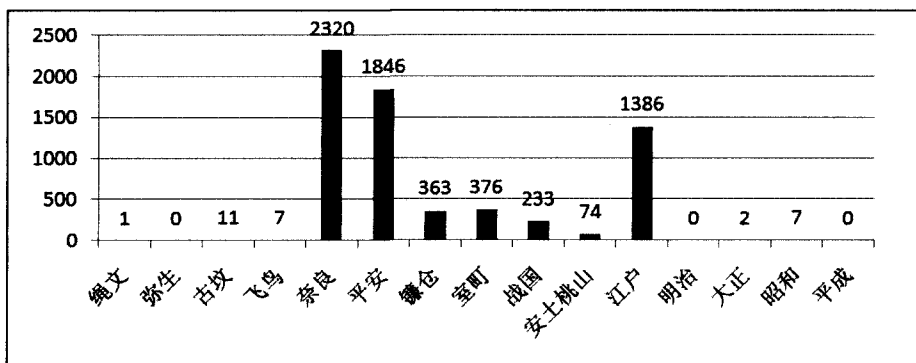


图 1:词汇的产生时代分布图

这三个朝代的词语数量特别突出,加起来已经超过总词数的八成以上,这与三个朝代的历史情况基本相符:(1)奈良时代的新词数量最多,可能是遣唐使的制度所致。据《日本文化史》记载,遣唐使的制度一共持续了 200 多年。当中覆盖的年份就包括了奈良时代(710~794)。第一次是 607 年圣德太子所派的遣隋使。隋代前后共派出了 5 次。在 618 年唐取代隋后,日本在 630 年开始再次派出遣唐使,每次约有五至六百人,带回大量儒佛经典,广为流传。至 894 年,日本共派出 19 次遣唐使,成功抵达长安 15 次。直到下一个时代被当时的遣唐大使菅原道真废除为止。“汉字和汉籍儒典的传人为汉字在日本的传播起了很大的推进作用(叶渭渠,2003)。”从中国传入的典籍把大量汉词带入了日本。日本初期儒学以中国传入的典籍的原文为主,产生了许多全汉字词。(2)平安时代时沿袭奈良时代,派遣了两次遣唐使。《日本文化史》一书指出,这个时代是汉风文化在日本最为灿烂的时期。据《日本古代汉文学与中国文学》一书,在平安时代因为遣唐使的制度而有许多来自日本的僧人进入中国,他们从中国带回了大量的汉籍,这些汉籍除了经典的著作外,还有碑、赞以及传、记等,为日本文学带来巨大的影响。(3)江户时代的新词产生现象则主要因为宋学在日本的本土化和文学创作的大众化。早在 13 世纪的镰仓时代,宋学就已经传入了日本,但在当时只被视作学问的一种。直至江户时代,日本刚结束了战国时代的长期战乱,社会生活也趋向安定。人民面对太平时代,充满了讴歌现实生活的热情。此外,自室町时代的庶民地位提升后,民间文学和民间技艺得到发展。旧有的以儒教文化作背景的理想主义与新出现的重视现实生活的享乐主义两者的冲突和互相融合,加以文学大众化的现象,两者相辅相成,促成了许多新词的出现。

(二)词条的汉字构成比例分析。只知道日语词产生最多的时代是不够的,还要观察日语词中汉字的构成比例,以确定哪些是源自汉语,哪些是日本用假名或汉字自造的词语。表 2 给出了词典内所有日词的汉字构成比例,含假名日词的比例要低于纯汉字的日词。这说明大部分日词来源于汉字词。

表 2:假名日语与汉字日语的比例关系

类型	义项数	例 词
全假名	554	あえか(柔弱)
半假名	2262	赤垂れ(指男女之间有很深的交情)
全汉字	4251	兄、匕首
合计	7067	

表 3 和图 2 进一步给出了全假名、半假名和全汉字日语在各朝代的分布情况。表 3 是详细的朝代对应关系及词语类型数据,图 2 则更清晰地勾画出日语词产生的历史趋势。

表 3:日语在各朝代中出现的汉字构成比例

日本朝代 (时代)	年份(公元)	对应中国朝代	全假名	半假名	全汉字
绳文	前 10000 年 ~前 3 世纪	夏、商、周	0	0	1
弥生	前 3 世纪 ~ 3 世纪	秦、西楚、西汉、新、玄汉、东汉、三国、晋	0	0	0
古坟	3 世纪后 ~ 7 世纪	晋、十六国、南北、隋	1	2	8
飞鸟	6 世纪末 ~ 710	隋、前唐、周、中唐	1	1	5
奈良	710 ~ 794	中唐	59	845	1416
平安	794 ~ 1192	中唐、五代十国、辽、南北宋	103	700	1043
镰仓	1192 ~ 1333	辽、南北宋、西夏、金、元	24	123	216
室町	1333 ~ 1467	元、明	22	101	253
战国	1467 ~ 1573	明	26	68	139
安土桃山	1573 ~ 1603	明	7	11	56
江户	1603 ~ 1867	明、清	236	328	822
明治	1868 ~ 1912	清	0	0	0
大正	1912 ~ 1926	中华民国	1	0	1
昭和	1926 ~ 1989	中华民国、中华人民共和国	3	0	4
平成	1989 ~	中华人民共和国	0	0	0
合计	-	-	483	2179	3964

可以发现词典中的日语在各朝代中出现的汉字构成比例在新词数量出现最多的三个时代如下:1. 奈良时代的日语汉字构成比例是:汉字 > 半假 > 全假。2. 平安时代的日语汉字构成比例是:汉字 > 半假 > 全假。3. 江户时代的日语汉字构成比例是:汉字 > 半假 > 全假。

从数据中可见奈良时代的全汉字词大大地超过了其他两个类别,其全汉字词与其他分类之间的差距在所有时代中也是最大的。在平安时代(794 ~ 1192),遣唐使的制度被当时的遣唐大使菅原道真所废除,从中国输入的汉文化也开始渐渐减少,但这正是日本开始消化汉文化,成功完成了汉风文化向和风文化的过渡,并形成具有特色的平安文化的时候。日本文化亦开始走向成熟,九世纪左右假名文字系统的兴起也推动了日本文学与词汇的发展。因此,平安时代的日语类型顺序里,半假名词和全汉字词都比奈良时代的少,但全假名词却有所增长。

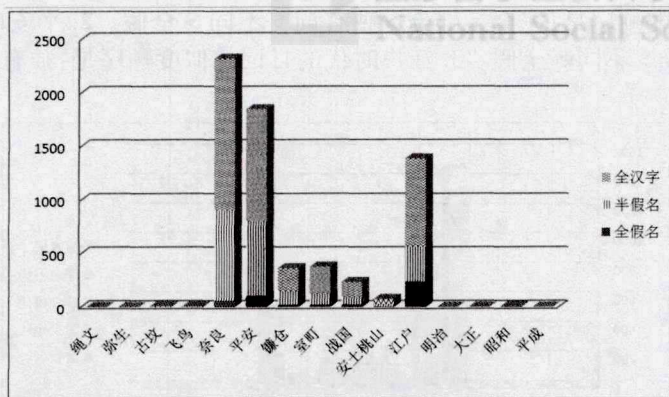


图2: 日词的汉字构成比例

江户时代虽然已没有了遣唐使的制度,但在当时古典主义盛行的时代背景影响下,日本的儒学者纷纷掀起了对中国古典的研究热潮。叶渭渠(2003)指出,在江户时代前期,幕府为了防止天主教及反对新政权势力结合组织“宗教暴动”,对之实行了禁止乃至镇压。相较之下,选择传播及普及对稳定社会能起一定作用的宋学,使宋学渐渐取代了佛教的主导地位,以新儒学的面貌成为江户占统治地位的意识形态。不少日本的学者开始探讨宋学,并促使儒学和神道融合,完成了学问的日本化,也因此产生了许多无法以汉字意译的全假名的日词,使江户时代的全假名日词比例有较大的增长。

此外,全汉字词在江户时代的比例还是很大,这是因为就算到了江户时期,中国文学对于日本来说仍然是非常值得研究探讨的对象。蔡毅(2002)指出,“在这一时代接受中国文学的情况,从诗、文到白话小说,是很全面的。”可见在17世纪到19世纪前半,日本依旧持续研究着中国的文学作品,因而产生大量汉字词不足为奇。

以词典中的日语词为例,“白无垢”,是指“一种日本特有的白色衣装,一般指婚礼中新娘穿着的衣物”,“世话女房”,意指“能把家里弄得井然有序,又能很好地照顾丈夫的女性”。两者都是在汉语中没有的全汉字日词。而就像平安时代的“御手洗”一样,这种全汉字的日词代表了日语吸收汉字词后自创日词的情况。而比起平安时代,日本在江户时代的文字系统已经较为成熟,已经开始懂得按需要造词,因此这种现象要较平安时代多。

(三)日文词语与汉语源词的意义相似度分析。我们已经分析了日语词大量产生的时代信息及汉字在日词中的占有比例,但光靠这两者仍然不够充分,因此我们还对日语词和对应的汉语词的意义进行了相似度的比对。只有在时代较早又意义相同的情况下,才更能显示出哪些日词是源自汉语的。

与汉语词典进行比对,分辨词典内所有全汉字类的日语与相应汉词的词义和词形是否相同来判定两者之间的相似度。根据资料进行统计,得出词典内日词与汉词在整体上的一致性如下。

表4: 日词与汉词的意义相似度

相似度	词项数	例词
全假名	554	あえか(柔弱)
词形词义完全相同	1545	雨、兄
词义部分相同	585	母、泡(义项不如汉语丰富)
词义完全不同	597	出雲(日本神话中的国名)
日语独有的汉字词(独有词)	1524	白無垢(日本服饰)
合计	4805	

上表显示了词典中所有全汉字日词经过比对后的词义相似度结果,数量最高的两个类别是完全相同及日语独有词,两者的数值相差无几。图3进一步用柱状图展示了词语相似度的历时分析,其结果如下:

1. 奈良时代的日词相似度顺序是:完全 > 部分 > 独有词 > 不同 > 全假。2. 平安时代的日词相似度顺序是:完全 > 独有词 > 部分 > 不同 > 全假。3. 江户时代的日词相似度顺序是:独有词 > 全假 > 完全 > 不同 > 部分。

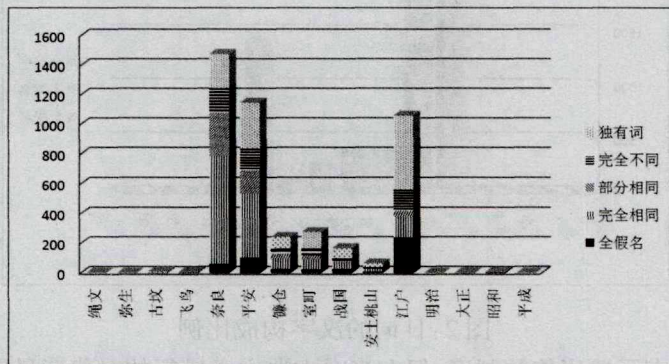


图3:各朝代日词与汉词的词义相似度

整体上,在三个新词数量最大的时代中有较明显变化的分别是独有词和完全相同两个分类。能用汉字意译却在现代汉语词典中没有成词的日词从奈良时代开始渐渐增加。另一方面,与现代汉语词典同形同义的日词却在同一时段不断减少,恰好展示了日本文字从引进到自造的发展过程。

在汉字传入日本以前,日本是没有文字的(何志明,2010)。从本文的统计数据来看,日本的文字系统是从奈良时代时借用汉字开始,到平安时代消化并创造出日本特有的假名文字系统,再至江户时代时已能较为熟练运用自己的文字系统造词。与汉词同义的日词数量不断减少的现象正说明了日本内部一直在消化当时现有的汉字词,并尝试本土化。而没有在汉语中成词的日词数量不断增加则证明了后期日本已开始给予部分汉字的除原义以外的字义,并把其运用于造词。

奈良时代的日词相似度中,完全相同的词居多。这个现象可能跟儒学在日本的盛行有关。中国的儒学在奈良时代之前就已在日本发展了起来,随着儒学在日本的盛行,必然会有大量的汉字词也跟着进入日本,并被原样引用。只因为日本最早的文学作品产生于奈良时代,才使得许多与汉词形义相同的全汉字词在奈良时代初次出现。王新生(2005)写道,“以假名写成的《万叶集》收集了759年以前约4500首和歌。作者范围广泛,包括从天皇到平民、关东的农民及服兵役的防人等。”

平安时代的日词相似度数据也反映出该时代日词与汉词的关联性,当时日本正在文化过渡期,以及日本文学的变迁有关。比起奈良时代,平安时代的独有词类有所增加,这很有可能是因为当时日本减少了对汉文化的吸收,开始把汉文化转为本土特有的和风文化时,产生了符合本土特色的汉字新词。江户时代产生的汉字新词在词义上与汉词完全相同的比例减少了,未成词的汉字新词数量则有较大增加。从日本的历史上看,江户时代正是日本开始重视本土特有文化的重要时期,此时的日本不再一味致力于学习来自中国的各种文学风格和思想,而是把这些都融合到当地的文学中,转而创造属于自己的风格。许多新的著作、论说都如雨后春笋般露出头来。他们使用汉字去表达日本特有的概念,才会形成许多未在汉语词典中成词的汉字新词。

借助日本出版的《日语源大辞典》和人工标注的数据,我们得出的结论如下。(1)整体上看,汉字词占的比例最高,日语假名词语比例较低。(2)时代上看,奈良时代、平安时代及江户时代等三个时代引进的汉字词数量最为庞大。(3)词条的汉字构成比例上,过半数以上的日词属于全汉字日词,约占词典中的六成,而其他含汉字的日词占词典中的三成多,含汉字的日词在词典中共占九成以上。(4)各朝代的词条的汉字构成比例上,新词数最高的三个时代均是全汉字词的数量最多,比例也最大。(5)词义相似度方面,全汉字词中比例最高的两个类型是日汉完全相同的词和日语独有词,两者各占的比例相近,都是三成。(6)各朝代的词义相似度上,全汉字新词数量最多的两个朝代均是与汉语拥有相同词形词义的词汇数量最多,而江户时代最多的却是在汉语中未成词的独有词词汇。

总的来说,首先,汉字词在日本的高峰期是奈良时代、平安时代及江户时代。因为这三个时代出现的新词数量最多。其次,汉字对日词的影响很深,根据词典,九成以上的日词是含汉字的。同时,有不少

汉语在传入日本后仍然保有原本的词义。同时,汉字传入日本后改变了当地的造词方式,创造了不少汉语中没有的新词。最后,以各朝代为分界来看,奈良时代、平安时代及江户时代的新词都是全汉字词居多。前两者的全汉字新词与汉语拥有相同词形词义的词汇数量最多。而江户时代最多的却是在汉语中未成词的独有词词汇。猜测可能是因为在奈良时代及平安时代主要以吸收汉字汉语为主,直到江户时代才是利用汉字造词的主要时期。所以在那个时候新造出来的汉字词很多是汉语中没有的。

我们将研究日本更大规模词表的词源分析,以确定日本历史上从中国借入的词语。另一方面,量化地研究日语对汉语的影响,特别是日本用汉字造词以后反传中国的情况。进而从整体上建立日汉互动的词源数据库,为中日语言文学文化方面的研究提供数据支撑。

参考文献:

- [1] 蔡凤林. 试述汉字对奈良时代以前的日本民族文化发展的历史贡献[J]. 中央民族大学学报(哲学社会科学版), 2016(1).
- [2] 蔡毅. 中国传统文化在日本[M]. 北京: 中华书局, 2002.
- [3] 陈红. 汉字造字模式与日语的“国字”[J]. 解放军外国语学院学报, 2002(7).
- [4] 冯玮. 大国通史·日本简史[M]. 上海: 上海社会科学院出版社, 2008.
- [5] 高兵兵. 日本古代汉文学与中国文学[M]. 北京: 中华书局, 2006.
- [6] 何志明. 从日语看日本文化[M]. 香港: 中文大学出版社, 2010.
- [7] 金田一京助. 新明解国语辞典[M]. 日本: 三省堂, 1999.
- [8] 金田一京助. 新选国语辞典[M]. 日本: 日本小学馆, 2002.
- [9] 李莉. 关于日语词汇中的汉语词[J]. 日语学习与研究, 1998(1).
- [10] 李无未. 日本汉字音的时间层次及其确认的证据和方法[J]. 当代语言学, 2005(3).
- [11] 鈴木俊二. 借用語の理論(3): 日本列島の歴史言語学[Z]. 国際短期大学紀要, 2005.
- [12] 潘钧. “汉和辞典”的性质、种类及发展方向[J]. 解放军外国语学院学报, 2000(1).
- [13] 前田富祺. 日本语源大辞典[M]. 日本: 日本小学馆, 2009.
- [14] 邱惠悠. 你一定想知道的日本文学简史[M]. 台北: 大是文化有限公司, 2014.
- [15] 史有为. 日本所用汉字的汉语“转型”初探[J]. 世界汉语教学, 2005(4).
- [16] 唐月梅. 日本文学[M]. 上海: 三联书店, 2006.
- [17] 桃井惠一. 中国語に入った日本語表現[Z]. 京都創成大学紀要, 2006(6).
- [18] 汪维辉. 域外借词与汉语词汇史研究[J]. 江苏大学学报(社会科学版), 2009(1).
- [19] 王保田. 日本简史[M]. 上海: 人民出版社, 2006.
- [20] 王新生. 日本简史[M]. 北京: 北京大学出版社, 2005.
- [21] 王仲涛, 汤重南. 日本史[M]. 北京: 人民出版社, 2008.
- [22] 新村出. 外来語の話[M]. 日本: 教育出版, 1982.
- [23] 叶渭渠. 日本文化史[M]. 广西: 广西师范大学出版社, 2003.
- [24] 赵基天. 详解日本外来语辞典[M]. 吉林: 吉林教育出版社, 2002.

责任编辑: 曰 若

A Diachronic Analysis of Loanwords from Chinese in Japanese based on *Japan Etymology Dictionary*

Li Bin Chen Jing

Abstract: The loanwords from Chinese in Japanese are always an important object of linguistics studies. Based on the data in the *Japan Etymology Dictionary*, this paper analyzes the time, the part of speech, the proportion of the Chinese characters and their similarity with Chinese words. The result shows that, most of the loanwords emerge in the Nara, the Heian and the Edo period. In addition, more than half of the words in the dictionary are written in Chinese characters, while more than 90% of Japanese word contains Chinese characters. It also explores the causes of the loanwords, the phenomenon of the high proportion of Chinese characters, and the phenomenon of a sharp increase or decrease from time to time.

Key words: Loanwords; Japanese Loanwords from Chinese; Diachronic Analysis; Etymology