

## 语言文字学研究

## 基于“HSK 常用词汇”的现代汉语外来词统计研究

张 彤

(中国传媒大学 外国语学院,北京 100024)

**摘要:** 本研究以《汉语水平词汇与汉字等级大纲》(1992)的词表为总体样本,以《汉语外来词词典》《词库建设通讯》“汉语外来概念词库”词条选刊等8个“鉴别依据”中的外来词研究结论为范本,按照统一的分类和鉴别原则、以广义外来词的视角,对词表中各种词源和借用类型的外来词进行鉴别,共提取989条外来词,对它们在各常用级别、词源、词类、借用类型、音节/语素数目的分布状况和特点进行描述与分析。

**关键词:** 外来词;现代汉语常用词;借用类型;鉴别依据;分布

**中图分类号:** H136.5      **文献标识码:** A      **文章编号:** 1005-6378(2017)04-0014-11

**DOI:** 10.3969/j.issn.1005-6378.2017.04.003

## 一、研究的缘起与词表的确定

关于外语对汉语影响的研究存在着两种截然相反的观点,一种认为外来词进入汉语太多,对汉语构成了威胁;另一种认为汉语比较保守,不容易吸收外来词。然而对于汉语中目前有多少外来词还没有一个明确数据。英语总词汇量的约80%是外来词<sup>[1]108</sup>,俄语中外来词占10%<sup>[2]</sup>,现代德语非派生的基本词汇中原德语词仅占四分之一<sup>[3]</sup>。日语词汇中,汉语(从汉语借入及利用汉字自创的词汇)占41.3%,外来语(主要指从西方语言引进的词汇,基本用片假名书写)占2.9%,混种语(由前述二种或三种词汇混合构成的词汇)占1.9%<sup>[4]</sup>。

要对汉语在语言竞争中的处境作出准确的判断、制定合理的语言政策,首先应该对汉语中已有的外来词量有一个客观的认识,进而对汉语自身吸收外来词的特点有一个客观的认识。因此本文计划以一个具有代表性的现代汉语词表为基础,进行外来词量和词语特征的统计分析,以定量研究的结果来说明汉语在语言竞争中的处境。

与一些常用词词典(词汇量4000—6000条)相比,《汉语水平词汇与汉字等级大纲》(北京语言学院出版社,1992)的“HSK常用词汇”8822词表收词量较大,兼顾了词汇的分布和频率,且配有甲、乙、丙、丁四个常用级别的标注,便于进行统计,因而被确定为本文的统计词表;该表第6471和6472词条“吐,乙(级),动(词)”重复,实际应为8821词,统计中按8821词计算。

## 二、统计中外来词的范围与分类

历史比较语言学在探寻汉语源头的过程中不断通过同源和借用的考证来分析早期华夏周边语言的

收稿日期:2016-12-20

基金项目:教育部人文社科青年基金项目“现代汉语常用词汇中的广义外来词统计研究”(13YJC740140)

作者简介:张彤(1975—),女,山东滕州人,博士,中国传媒大学外国语学院副教授,主要研究方向:社会语言学。

接触情况和汉语的形成情况,汉语借词的历史上限已经推到了三千年前的民族融合时期。据考证,“谷”和“禾”可能来自壮侗语族<sup>[1]39</sup>，“老虎”来自南亚语<sup>[5]33</sup>，“江”和“河”分别来自南亚语和阿尔泰语<sup>[5]9</sup>。本文选定的词表中就有“河”“禾苗”“江”，为了统计上便于操作，有必要确定外来词的历时范围和参照。

根据张德鑫对于外来词进入汉语的历史分期,汉语有过三次大规模吸收外来词的高潮。第一次在汉唐时期,汉族与周边少数民族或国家交流密切,并且大规模翻译佛教经;第二次在清末至民国时期,中国大量翻译西方著作,且有大批留学生去国外学习,回国后传播新知识;第三次在改革开放后,特别是 90 年代以来<sup>[6]</sup>。对这三三次汉语外来词高潮学界有较为充分的研究材料,本文选用的词表是 1992 年编制的,因此把统计的外来词上限定在上古汉语,主要考察汉代以来外来词经历过前两次高潮后在 90 年代初现代汉语常用词汇中的沉积状况,对汉语吸收外来词的能力和特点进行评价,第三次外来词高潮不是本文统计的主要方面。

国内外研究领域没有相对统一的汉语外来词分类体系,为了避免定义和分类上的混乱,本研究采用广义的划分,将借入的音、形(含字母)、意、语素、词、词组等统称作外来词(或借词),从借用方式上作进一步划分。这一划分的依据主要有两个方面:(1)《汉语外来词》第一章“外来词的界定”中指出,在外来词的三个基本层次——借音、借形<sup>①</sup>、仿译——之外还有外来语素的借用<sup>[1]8-10</sup>; (2)《语言接触与英汉借词研究》中还加入了“词组借用”这一层次<sup>[7]203-204</sup>。音、形、义的借用形式在语素、词、词组不同的单位层面上存在交叉,很难将音译和意译截然分开,所以本文以连续统的方式将西语外来词和日语外来词统一进行分类(见表 1)。

表 1 外来词形、音、义借用形式连续统

类型序号	借用方式	语素	词	词组
类型 1	字母词 借形+借义	Tel <sup>②</sup> , No., \$		
类型 2	字母词借形+借音 +借义	cc, kg, km, e/E (如“E 时代”), -G	AA <sup>③</sup> , ABC, CEO, DNA, E-mail, Fax, GRE, HIV, IQ, MTV, SARS, VC, wifi, WTO, X(X)L	B to B / B2B, P to P/ P2P
类型 3	借形+音译 / 音译 +借形		T 恤(T-shirt), 维他命 C(Vitamin C)	
类型 4	借形+借音+意 译/意译+借形+ 借音		ATM 机(automatic teller machine), B 超(type-B ultrasonic), 三 K 党(Ku Klux Klan 简称 K.K. K.), 4S 店(4S store/shop), IC 卡(integrated circuit), L 码(Large), PM2.5 (细颗粒物)(Particulate Matter 2.5), X 光(X-ray), α 波(awave), 维生素 C(Vitamin C)	AA 制, F1 赛车, IP 卡/电话, IP 地址, PDF 文档/格式, SOS 儿童村, 3D 打印, 4G 网络, VIP 卡/会员
类型 5	(无意译成分的)纯 音译 / 无意谐音	的(taxi): 打的、面的、的哥; -巴(bus): 大/中/小 巴; 吧(bar): 酒吧、网吧、泡吧、吧台; 秀(show): 脱口秀、时装秀、宠物秀、做秀	英格兰(England), 玛丽(Mary), 打(dozen), 米(meter), 安培(ampere), 尼古丁(nicotine), 苏打(soda), 尼龙(nylon), 福尔马林(formalin), 克隆(clone), 沙发(sofa), 拷贝(copy), 坦克(tank), 探戈(tango), 迪斯科(disco: discotheque); 卡拉 OK(カラオケ/カオケ/カタオタ)	

① 史有为先生所指的“借形”指的是日语汉字词进入汉语的情况, 本文将这一概念层次也应用于进入汉语的西文字母。

② “Tel”“No.”“\$”通常用于书面语中, 在读到这些内容是人们通常换以汉语。“Tel”读作“电话(号码)XX”; “No.”通常读作“编号 XX”或“(第)XX 号”; “\$”读作“美元”。

③ 根据《新华新词语词典》的解释, AA 为拉丁语的缩写, 是“各”“各各”的意思。汉语借用的是英语字母 A 的读音。

续表 1

类型序号	借用方式	语素	词	词组
类型 6	以偏旁部首释义的音译		钙(calcium)、硅(silicon)等化学名称,咖啡(coffee)	
类型 7	通过表意语素释义的音译		仙客来(cyclamen),芒果(mango),香波(shampoo),绷带(bandage),的确良(dacron),奔腾(Pentium),可口可乐(Coca-Cola),舒服佳(Safeguard),维他命(vitamin)	
类型 8	音译+意译/意译+音译		魔术(magic),轮胎(tire),啤酒(beer),霓虹灯(neon),冰激凌(ice cream),摩托车(motorcycle),雪茄烟(cigar),芭蕾舞(ballet),拓扑学(topology),因特网(Internet)	特洛伊木马(Trojan horse),潘多拉魔盒(Pandora's box),德谟克利斯之剑(Damocles' sword)
类型 9	借译/仿造/仿译(逐字意译)		黑板(blackboard),足球(football),热狗(hot dog),甜心(sweetheart),机关枪(machine gun),超市(supermarket),键盘(keyboard),黑马(dark horse),蜜月(honeymoon),鸡尾酒(cocktail)	斑马线(zebra crossing),象牙塔(ivy tower),死亡之吻(a kiss of death),武装到牙齿(armed to the teeth),冰山一角(the tip of the iceberg)
类型 10	(无音译成分的)纯意译	多-(multi-/poly-):多媒体;兆-(mega-):兆赫,兆米	纤维(fiber),计算机/电脑(computer),火车(train),先生(gentleman),煤油(kerosene),水泥(cement),出租(汽)车(taxi)	
类型 11	借形+借义	—主义(主義/しゅぎ),腺(腺/せん)	吨(噸/トン),瓦斯(瓦斯/ガス),淋巴(淋巴/リンパ),写真(写真/しゃしん) <sup>①</sup> ,科学(科学/かがく),电话(電話/でんわ),手续(手続/てつづき),场合(場合/ばあい),议会(議會/ぎかい),俱乐部(俱樂部/クラブ)	
类型 12	借义 <sup>②</sup>	电—(原指闪电,现指 electric 或 tele-);电扇、电视、电报	知识(knowledge),龙头(tap),点(point),肥皂(soap);寿司(すし),民主(みんしゅ),经济(経済/けいざい),革命(かくめい),柔道(じゅうどう)	

### 三、统计中的外来词鉴别依据和原则

统计时,对 8821 词表中外来词的确定以下列 8 种外来词研究比较权威的成果为依据。

鉴别依据:① 刘正埭、高名凯、麦永乾、史有为主编《汉语外来词词典》(上海辞书出版社,1984 年)。② 香港中国语文学会主办《词库建设通讯》(1993 年 7 月—2000 年 6 月,共 22 期)

① 写真(しゃしん)“在日语中指“照片”,在汉语中指“黄色照片”,意义有所变化。

② “借义”指汉语(包括古汉语)中原本没有该词的现代意义,由于受到外来事物影响,汉语用原有词汇表示“外来概念”;对于日语外来词,日语借用了古汉语的词汇并赋予了新的意义,现代汉语使用的是日制汉语的语义。而对于日语用古汉语词汇对译西语且与古汉语语义保持不变的词语(如“政治”“消费”等),本文不将其作为汉语向日语借词处理。

“汉语外来概念词词库”词条选刊。③ [意]马西尼著,黄河清译《现代汉语词研究汇的形成——十九世纪汉语外来词研究》(香港汉语大词典出版社,1997年)185—274页“附录1”、“附录2”。④ 商务印书馆辞书研究中心主编《新华新词语词典》(商务印书馆,2003)中的外来词。⑤ 李晋霞、杨书俊(2004)从清华大学孙茂松 87448 词条的大词表中人工辨认的 1017 个印欧音译词词表。⑥ 香港中国语文学会主编《近现代汉语新词词源词典》(香港汉语大词典出版社,2000年)。⑦ 史有为著《外来词——异文化的使者》(上海辞书出版社,2004年)317—355页外来词索引。⑧ 高名凯、刘正琰著《现代汉语外来词研究》(文字改革出版社,1958年)33—111页第三章“现代汉语的外来词”。

需要特别说明的是:“鉴别依据⑥”《近现代汉语新词词源词典》“序”第2页第二段指出“我们收词的对象主要是外来词。……为了避免这些理论上的纠缠,我们选用了‘新词’这个名称,实际包括了音译词和意译词”;第3页末段又说明词典“给出了外语对应词,主要是英语对应词,但是这仅仅是个读者提供一种方便,并不表示该词一定译自英语。”因此统计中凡仅仅依据该词典确定的外来词,词源后加“?”表示“存疑”;对于没有给出对应词的条目,统计中不作为外来词,例如“演奏家”。

由于不同的外来词研究成果对某些外来词的结论存在分歧,在统计开始前和进行中本文先后设定了下列“鉴别原则”以解决这些分歧。

鉴别原则:

1. 广义外来词原则。统计中将音译词和意译词均视为外来词。如果“鉴别依据①—⑧”中经研究有充分依据判定为外来词的,在本统计中即认定为外来词。例如:“冰棍儿”仅在“鉴别依据②”中有正式的收录和考证,统计中则依此确定其为外来词。

2. 新结论优先原则。对于同一个词的考证,当不同“鉴别依据”存在分歧时,如果理据充分且逻辑线索清晰则按时间顺序取舍,首选最新研究结论。例如:现代意义的“知识”在“鉴别依据①”中列为日源外来词,而“鉴别依据②”的考证证实现代意义的“知识”先产生于中国,用于意译“knowledge”,然后传入日本并广泛使用,又回流到中国,因此统计时采取“鉴别依据②”的结论。

3. 证据不足并存原则。对于同一个词的考证,如果考证结论不同且本文没有充分的依据做取舍,则按理据的充分性做降序排列。例如:“常识”一词“鉴别依据①”将其作为日源外来词,认为是日语意译英语的日创汉语词;“鉴别依据②”中的条目显示其词源为英语。统计中将两个不同的鉴别结果按顺序列出。

4. 古汉语的借义原则。对于受到外来事物影响而使古汉语词语产生的新义,统计中将其视为“借义”,作为外来词“类型12”处理,例如“知识”、“地质”等。

5. 术语借词的确定原则。对于“鉴别依据③”中“附录1汉语化学元素名称对照表”的词,凡属新造词按外来词对待,例如Oxygen曾先后以“养气”、“养”来翻译,最后出现了“氧”。

6. 非外来词的确定原则。对于日语用古汉语词汇对译西语,且汉语古今意义基本保持不变的词语,统计中不将其作为日源外来词处理,例如“政治”、“消费”等。对于曾被认为是外来词,而有充分的新证据表明是本族新词的,统计中不将其作为外来词,例如“资料”,“鉴别依据①”判定为日源外来词,而“鉴别依据②”证明该词先出自我国,判定为本族新词,统计中以新的结论为准,不将其作为外来词。

#### 四、基于 8821 常用词表的外来词统计结果

##### (一)外来词在各常用级别和词源中的分布

统计结果显示(表2),在8821条常用词汇中外来词为989条(不计异体),约占11.2%。从总体上看,汉语外来词量的比例与英语、日语、德语相比明显较低,与俄语的情况比较接近。

表2 外来词在各常用级别词汇中的数量和比例(不计异体)

常用级别	总表词条	外来词词条	外来词比例
甲	1 035	122	11.8 %
乙	2 018	273	13.5 %
丙	2 198	253	11.5 %
丁	3 570	341	9.6 %
合计	8 821	989	11.2 %

在甲、乙、丙、丁四个常用级别中,乙级词汇中的外来词量略高,然后随常用级别的下降而下降,甲、丙两级中的比例几乎持平,丁级词汇中的外来词量略低。本文认为,外来词在四个级别的常用词汇中形成这种分布的原因在于,甲级词汇中有一部分是“基本词汇”,基本词汇以稳定的本族原生词或根词(primitive)为主,借用的含量较低,因而甲级词汇中的外来词不是四个级别中比例最高的;然而日常生活中较为常用的事物或概念中外来的又比较多,所以乙级词汇是外来词比例最高的,随着常用级别的下降,外来词的比例也相应下降。

8821常用词表中外来词的词源分布表3显示历史上的少数民族语言对汉语常用词汇的影响很微弱。按我国目前的版图来定义少数民族语言,统计中的外来词涉及到蒙古语、匈奴语、女真语、伊兰语、苗语共8个词条,一半是动植物的名称。少数民族语言对汉语常用词汇极为微弱的影响反映了汉语在历史上的强势地位并没有因少数民族入主中原而有所改变,这种语言强势很可能是由于汉语所承载的科技和文化在我国历史上一直是处于领先地位而造就的。

表3 外来词的词源分布

常用级别	外来词	亚洲语言							欧洲语言							
		日语	日语存疑	梵语	梵语存疑 <sup>1</sup>	阿拉伯语	蒙古语	其他语言	英语	英语存疑	法语	德语	德语存疑 <sup>2</sup>	俄语	西班牙语	其他语言
甲	122	61	1	1	0	0	1	0	25	33	0	0	0	0	0	0
乙	273	154	0	1	1	1	0	3	22	83	1	2	1	0	1	3
丙	253	118	0	0	0	1	2	3	28	92	4	2	1	2	0	0
丁	341	170	1	1	1	1	0	0	40	124	1	0	0	0	1	1
合计	989	503	2	3	2	3	3	6	115	332	6	4	2	2	2	4

说明:

1. 梵语或波斯语,梵语或于阐语。
2. 德语或英语。

从总体上看,常用词表中外来词的词源分布特点是以日源和英源词为两大主流(图1),日源词是日创汉字词,占了外来词的51%,日本人对汉语创造性地使用为现代汉语词汇做了很大贡献,日语对汉语的影响最大。此外,本文在“鉴别依据”中特别说明凡仅仅根据“鉴别依据⑥”《近现代汉语新词词源词典》确定的外来词在统计中视为“词源存疑”,这种情况在332条英源存疑的词汇中占328条,在2条日源存疑的词汇中占1条,在其他词源中没有。如果排除所有词源存疑的日源和英源词,则8821常用词表中外来词的数量和比例如图2所示。去掉日源和英源存疑的外来词后,外来词在常用词汇中的比例下降约4个百分点,为7.3%;日源和英源外来词依旧是常用词汇中外来词的主流,只是日源词的比例更大,超过四分之三。存疑的英源外来词主要是意译词。

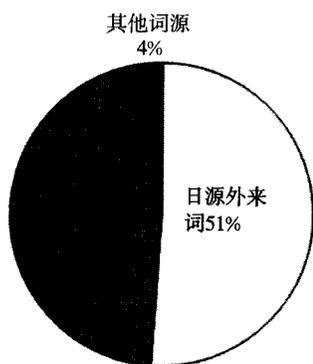


图1 外来词的词源分布(比例)

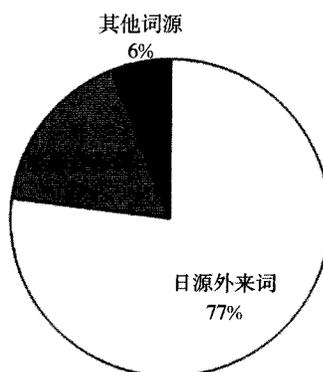


图2 不含日源、英源存疑的外来词词源分布(比例)

### (二)外来词的词类分布特点

8821 常用词表中很多词条没有标注词类,对于统计中鉴别出的所有外来词,凡缺少词类标注的本文均依据《现代汉语词典》(2005)确定了词性。表4 所显示的是外来词在各词类中的分布情况,表5 是将兼类词折合成词条数后的统计结果。从表5 中可以看出,名词是外来词中的主力军(8821 词表不含人名、地名等专有名词),其总量接近于外来词的四分之三,动词略大于总量的五分之一,形容词和其他词类不足10%。外来动词比例在各常用级别中比较接近。外来词名词和形容词在甲、丙、丁三个常用级别中的比例差别不大,却在乙级词汇中最富戏剧性。

表4 各常用级别中外来词的词类分布

常用级别	总表词条	外来词	名词	动词	形容词	量词	代词	连词	介词	兼类词					词头 <sup>①</sup>
										动名	名形	动形	名量	形副	
甲	1035	122	84	13	2	1	1	0	0	17	2	0	1	1	0
乙	2018	273	173	32	18	3	1	0	1	29	8	7	0	0	1
丙	2198	253	187	35	7	3	0	0	0	15	2	4	0	0	0
丁	3570	341	241	70+2 <sup>②</sup>	15	0	0	2	0	9	1	1	0	0	0
合计	8821	989	685	152	42	7	2	2	1	70	13	12	1	1	1

表5 兼类词折合为词条数后各词类在外来词中所占比例(不含非自由语素)

常用级别	非兼类词条数	兼类词折合词条数 (兼类词数×2)	累计	名词比例	动词比例	形容词比例	其他词类比例
甲	101	42	143	72.7%	21.0%	3.5%	2.8%
乙	228	88	316	66.5%	21.5%	10.4%	1.6%
丙	232	42	274	74.5%	19.7%	4.7%	1.1%
丁	330	22	352	71.3%	23.3%	4.8%	0.6%
合计	891	194	1 085	70.9%	21.6%	6.3%	1.3%

乙级词汇是各常用级别中外来词比例最高的,但却是外来名词比例最低的,且低于总体水平,然而外来形容词的比例明显超过其他常用级别两倍多。外来形容词80%来自日语,乙级在日源形容词方面比总体略高4个百分点。对于乙级词汇中外来名词与形容词的一低一高这种异常现象,本文又进一步分别

① “词头”为非自由语素“阿”。

② 动词词组2条:“合资”,“通货膨胀”。

统计了 8821 词表中的名词、形容词(含兼类,包括本文统计时为外来词添加的词类标注)比例,并分别与外来词占名词、形容词的比例相比较。名词和外来名词的对比情况见表 7 和图 3,形容词和外来形容词的对比情况见表 7 和图 4。

表 6 名词和外来名词在各常用级别中的分布情况

常用级别	有词类标注 词条数	有词类标注 词条比例	名词词条数 (含兼类)	名词比例	外来名词 词条数	外来词占 名词比例
甲	968	93.5%	441	45.6%	104	23.6%
乙	1 907	94.5%	893	46.8%	210	23.5%
丙	2 059	93.7%	1 030	50.0%	204	19.8%
丁	3 203	89.7%	1 521	47.5%	251	16.5%
合计	8 137	92.2%	3 885	47.7%	769	19.8%

从图 3 上可以清楚地看到,各常用级别中,甲、乙两级词汇分别是名词比例最低和次低的,但外来名词占名词比例却分别是最高和次高的(比例仅相差 0.1%)。这说明乙级词汇中的外来名词实际上并不少,只是由于外来形容词的比例相对高于其他级别一倍或更多而使外来名词表现出相对的低位。那么乙级词汇中外来形容词的异常高比例才是更值得关注的。

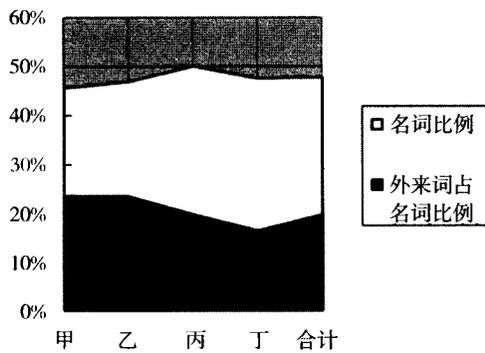


图 3 名词比例与外来名词比例比

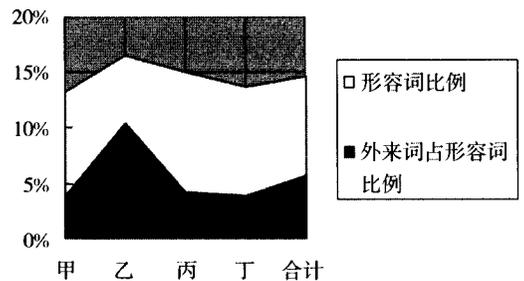


图 4 形容词比例与外来形容词比例比

在图 4 上,两组比例所形成的线条并不平行,虽然乙级词汇中的形容词比例本来就略高一些,但在外来形容词占形容词的比例明显超过了乙级词汇中形容词占所有词类比例的相应范围,而甲、丙两级中外来形容词占形容词的比例均小于相应的比例,所以才有了乙级词汇中外来形容词相对集中的异常情况。虽然就数值而言,10.5%本身并不算很大,但与其他级别同类词汇相比,乙级形容词中的这一外来比例还是比较突出的。乙级外来容词的超常现象表明,随着科技和文化的发展,贴近人们生活的新事物和新概念不断增加,日常交际中人们需要更多的分工更细的形容词来描述它们,这些描述性词汇使用比较频繁,而原有词汇系统不能满足人们描述新生事和概念物的需要,外来词便填补了这些空白。

表 7 形容词和外来形容词在各常用级别中的分布情况

常用级别	有词类标注 词条数	有词类标注 词条比例	形容词词条数 (含兼类)	形容词比例	外来形容词 词条数	外来词占 形容词比例
甲	968	93.5%	128	13.2%	5	3.9%
乙	1 907	94.5%	315	16.5%	33	10.5%
丙	2 059	93.7%	309	15.0%	13	4.2%
丁	3 203	89.7%	438	13.7%	17	3.9%
合计	8 137	92.2%	1 190	14.6%	68	5.7%

(三)外来词的借用类型特点

表 8 不同借用类型的外来词数量与比例

常用 级别	外来词	类型 5		类型 6		类型 7		类型 8		类型 9		类型 10		类型 11		类型 12	
		数量	比例	数量	比例	数量	比例	数量	比例								
甲	122	1	1%	2	2%	1	1%	4	3%	9	7%	39	32%	38	31%	28	23%
乙	273	5	2%	2	1%	1	0%	5	2%	19	7%	83	30%	114	42%	44	16%
丙	253	6	2%	4	2%	0	0%	9	4%	24	9%	86	34%	96	38%	28	11%
丁	341	11	3%	7	2%	2	1%	8	2%	36	11%	101	30%	147	43%	29	9%
合计	989	23	2%	15	2%	4	0%	26	3%	88	9%	309	31%	395	40%	129	13%

8821 常用词表中不含字母词,因而统计中实际涉及到的外来词是表 1 中的“类型 5—12”。表 8 显示了不同借用类型的外来词数量与比例情况。在所有的外来词类型中,几种音译形式“类型 5—8”普遍较低,都在 5% 以下,4 种类型合计也不过只占外来词的 7%(占 8821 词表的 0.77%),其中“类型 8”(音译+意译/意译+音译)比例最高,“类型 5”(无意译成分的纯音译/无意谐音)次之,“类型 6”(以偏旁部首释义的音译)其次,“类型 7”(通过表意语素释义的音译)比例最低,接近于 0。

意译是汉语吸收外来词的主要途径。“类型 9”(借译/仿造/仿译/逐字意译)除“幼儿园”一条出自德语外,其他均源自英语,这一类在意译形式中比例最低为 9%,说明中西语言间词素和语义能直接一一对应的情况比较少;“类型 10”(纯意译)占外来词的 31%,除“眼镜”一条出自阿拉伯语外,其他均是对英语等西语的意译;“类型 11”(借形+借义)占外来词的 40%,全部源自日语,是日创汉字词,可以说是现代汉语直接接纳了日语对西语的意译;“类型 12”(借义)占外来词的 13%,六分之五源自日语意译西语对古汉语的语义创新,六分之一是我国用古汉语意译西语(表 9)。统计中的 989 条外来词中有 13.4%(133 条)出自古汉语,此外还有 38 条本文不计入外来词的古汉语,但有的研究把它们作为“回归/回流借词”或意译词,可见古汉语是现代汉语丰富和发展的宝贵的财富。但假如我们站在古人的角度来看待这些古为今用的语义变化,恐怕先人们要骂这些子孙后代滥用他们的语言了。

如果去掉日源外来词(即“类型 11”和“类型 12”中的六分之五),不含音译的意译约占外来词的 42%,是各类音译(“类型 5—8”)总和(7%)的六倍。从常用词汇中音译与意译比例上的悬殊可以看出,在接纳其他语言的成分方面,汉语极大程度上受到单音节意音文字和语义型语言特点的制约<sup>[8]</sup>。“类型 6”(以偏旁部首释义的音译)涉及到造字,对于汉字和词汇量相对较大的汉语,这种增加负荷的方式可能不受欢迎,而“类型 7”(通过表意语素释义的音译)匹配的难度比较大,在一些商标中较为常见,但常用级别很低。“类型 5”(无意译成分的纯音译/无意谐音)与汉语语义型语言的特点不吻合,所以量很少。“类型 8”(音译+意译/意译+音译)这种以意译调和音译的方式比较适合汉语的特点,所以在音译方式中比例最高。

有些学者认为意译、仿译、用古汉语意译西语看不出词汇的外来特征,不属于借词。按照这种狭义外来词观点来统计 8821 词表,则外来词的数量为 575 条,占常用词汇的比例是 6.5%,略高于广义外来词统计结果 11% 的一半。对于日源汉字词在汉语中的外来身份,很少有学者明确否认<sup>①</sup>。同样,还有通过英语进入各种语言的法语、拉丁语成分,通常也被人们认为是英源外来词。可见,假如语言 A 的成分 Xa 被借入语言 B,加工改造后成为 Xb,该成分在 B 语言中的身份成为既定的事实,即便再输出到其他语言,也不会改变其 Xb 的词源身份 b。

(四)外来词的音节数目特点

现代汉语词汇以双音节为主流的特点也决定了外来词要进入常用词汇中也必须适应这一语言特点。

① 参见史有为(2013:238—241)7 位学者的外来词分类体系对比。

表9显示双音节借词占外来词总数的88.7%，单音节外来词占2.4%，三音节及三音节以上借词共占8.9%，且常用级别越低三音节借词越多，四音节借词也有类似的趋势，唯一进入甲级词汇的四音节借词是与人们生活联系紧密的“出租汽车”。

表9 外来词的音节特点

常用 级别	外来词	单音节		双音节		三音节		四音节		五音节	
		数量	比例	数量	比例	数量	比例	数量	比例	数量	比例
甲	122	8	6.6%	108	88.5%	5	4.1%	1	0.8%	0	0.0%
乙	273	4	1.5%	251	91.9%	17	6.2%	1	0.4%	0	0.0%
丙	253	4	1.6%	229	90.5%	16	6.3%	3	1.2%	1	0.4%
丁	341	8	2.3%	289	84.8%	36	10.6%	7+1 <sup>1</sup>	2.3%	0	0.0%
合计	989	24	2.4%	877	88.7%	74	7.5%	13	1.3%	1	0.1%

注:1.词组1条:“通货膨胀”

单音节借词在甲级词汇中为“点、电、分、米、她、系、站、周”，在乙级词汇中为“阿(词头)、吨、秒、塔”，在丙级词汇中为“癌、棒、磅、铝”，在丁级词汇中为“钙、硅、磷、镁、氢、权、氧、州”；以名词和量词为主，名词中化学元素占约三分之一，但常用级别不高，其他单音节外来名词则有较强的派生或组合构词能力；音译自法语的形容词“棒”使人联想到目前流行的英语音译词“酷”，或许随着时间的推移“酷”也会进入常用词中。单音节借词中的63%出现在“类型6”(以偏旁部首释义的音译)中，17%出现在“类型5”(无意译成分的纯音译/无意谐音)中，可见有80%的单音节借词是以音译的方式进入汉语常用词汇的。

对三音节借词借用类型的分析显示(表10)近三分之二三音节词外来词源自西方语言，意译是音译的三倍多。只有两种音译类型出现在三音节外来词中，“类型5”(无意译成分的纯音译/无意谐音)是5种借用类型中数量和比例最低的，仅有三个丁级词；“类型8”(音译+意译/意译+音译)是音译的主要方式。在四音节和五音节借词(共14个)中也只有3个词含音译，均属于“类型8”。可见即便是在多音节词汇中汉语也是主要通过意译吸收外来词的，音译也多是配以意译进行的，纯音译只是极个别的情况，而且使用频率也不高。

表10 三音节外来词的借用类型分布

常用 级别	三音节外来词 词条数	音译		意译		日源借形+借义
		类型5	类型8	类型9	类型10	类型11
甲	5	0	0	0	3	2
乙	17	0	1	5	8	3
丙	16	0	3	2	7	4
丁	36	3	4	6	6	17
合计	74	3	8	13	23	27
占三音节外来词比例		15%		49%		36%

为了更清楚地分析多音节(三、四、五音节)外来词在借用方面的特点，本文将三、四、五音节借词在借用类型分布方面做了合并统计表11。通过表10与11的对比可以看出，三音节借词的特点代表了多音节借词在借用类型和常用级别方面的特点。多音节借词近三分之二来自西方语言，意译是音译的三倍。从表13也可以看出，意译程度越高多音节借词越多，音译程度越高多音节借词越少，常用级别越低多音节借词越多。

表 11 多(3个≤)音节外来词的借用类型分布

常用 级别	多音节外来词 词条数	音译		意译		日源借形+借义
		类型 5	类型 8	类型 9	类型 10	类型 11
甲	6	0	0	0	4	2
乙	18	0	2	5	8	5
丙	20	0	5	2	7	4
丁	44	3	4	7	10	20
合计	88	3	11	14	28	32
占多音节外来词比例		16%		48%		36%

意译程度越高多音节借词越多,音译程度越高多音节借词越少,这一特点有点出乎作者的预料。究其原因,音译词音节越多,记忆的难度越大,交际成本高,也就是说交际双方的语言文化知识必须对等,否则要花些功夫解释。如果是意译,即使音节多一点,也容易理解和记忆,所以一些音译词刚刚被借入汉语时或许风靡一时,一旦经过一段时间的使用,有了可以匹配的意译词,音译词使用的频率就会大大下降,因此,意译在多音节外来词的使用当中有绝对的优势。例如,俄语音译外来词“布尔什维克”,本意为“多数派”,但我国主要使用的是它所负载的文化意义“共产主义者”或“共产党”“共产主义”是源自日语的借词,同为外来词但日源词更符合中国人的认知习惯,所以“共产主义”成为丁级常用词汇,“布尔什维克”则不常被提到。

外来词借用类型和音节数目方面的统计结果表明,语言的特点决定人们对词汇的选择;音译单音节和多音节借词的情况表明,其他语言词汇的音节特征将决定它们能否被汉语接纳和以何种方式接纳。意译不但能满足语义型语言对“义”的要求,而且能缓解其他语言词汇与汉语词汇音节数量方面的矛盾。日源外来词能占据汉语常用词汇中外来词的半壁江山,主要原因就在于它们满足了现代汉语对“形”“音(节)”“义”的综合要求。

### 五、统计结果小结

为了探究前两次外来词高潮对现代汉语的影响程度,本文以 8821 现代汉语常用词词表为总体样本,以 8 个鉴别依据中的外来词研究结论为范本,按照统一的分类和鉴别原则、以广义外来词的视角,对词表中涉及的各种词源和音、形、义各种类型借用方式的外来词,进行了穷尽式的鉴别,共提取甲、乙、丙、丁四个常用级别中 989 条外来词,对这些外来词在各常用级别、词源、词类、借用类型、音节数目、鉴别依据中的分布状况进行分析,显示以下主要特点:

(1)8821 词表中包含音译和意译的广义外来词共 989 个,占词表总量 7.3%—11.2%(排除与不排除日源、英源存疑词条);去掉意译、仿译、用古汉语意译西语这些有争议的借词后,狭义外来词占词表总量 5.4%—6.5%。

(2)在四个常用级别中,乙级词汇中的外来词量最大,从乙级开始外来词量随常用级别的下降而下降。

(3)不排除日源、英源存疑词条的话,日源、英源、其他词源外来词的比例分别为 51%、45%、4%,排除日源、英源存疑词条三者比例为 73%、20%、7%。

(4)外来词在不同词类中的分布分别为:名词 70.9%、动词 21.6%、形容词 6.3%、其他词类 1.3%。乙级词汇中,名词比例略低于平均水平,形容词比例显著高于平均水平。

(5)借用形式方面,两种意译类型占外来词总量 40%,日源借形+借义占 40%,利用古汉语进行借义的占 13%(六分之五为日源),四种音译类型的外来词共占 7%(占 8821 词表的 0.77%)。

(6)音节数目方面,双音节词占外来词总量 88.7%,单音节占 2.4%,三音节及三音节以上借词共占 8.9%。常用级别越低多音节借词越多,意译程度越高多音节借词越多,音译程度越高多音节借词越少。

从统计的结果来看,虽然经历过两次外来词进入汉语的高潮,但是大浪淘沙能最终成为汉语稳定成

份的外来影响比较有限,外语并没有对汉语构成很大威胁,汉语自身的特点如单音节意音文字、语义型语言、词汇的双音节化等结构因素对进入汉语的外来词有一定的过滤作用<sup>[8]</sup>。

[参 考 文 献]

- [1] 史有为.汉语外来词(增订本)[M].北京:商务印书馆,2013.
- [2] 孙汉军.俄语外来词研究[J].外语与外语教学,2002(11):5-8.
- [3] 何涛.德语中外来词的演变与规范[J].首都师范大学学报(社会科学版),1998(2):33-38.
- [4] 吴侃.日语词汇研究[M].上海:上海外语教育出版社,2005.
- [5] 王士元.汉语的祖先[C].李褒嘉,主译.北京:中华书局,2005.
- [6] 张德鑫.第三次浪潮——外来词引进和规范刍议[J].语言文字应用,1993(3):70-76.
- [7] 胡兆云.语言接触与英汉借词研究[M].济南:山东大学出版社,2001.
- [8] 张彤.语音、文字特点对词汇发展的影响[J].澳门语言学刊,2009(2):54-55+76.

【责任编辑 卢春艳】

## A Statistic Research of Loan Words in Modern Chinese Based on the Vocabulary of HSK Words in Common Use

ZHANG Tong

(School of International Studies, Communication University of China, Beijing 100024, China)

**Abstract:** This paper demonstrated findings about loan words in Modern Chinese based on the statistic analysis of “Commonly Used Words List for HSK Test” (1992) as the sample. Taking 8 authoritative loan words studies like “Dictionary of Loan Words in Chinese” etc. as “identification proof”, adopting unified loan words classification and identification principles, and approaching from a broad loan word perspective, the research made an exhaustive identification of all possible etymology and borrowing style in the sample and sorted out 989 loan words. The paper described loan words distribution in terms of frequency level, etymology, part of speech, features of borrowing style, syllable / character number, etc.

**Key words:** loan words; commonly used Modern Chinese words; borrowing style; identification proof; distribution