



《中国语学文库》 总主编：邢福义

现代汉语三音节词语研究

A Study of the 3-syllable Words and
Expressions in the Contemporary Chinese

杨书俊◎著



中国出版集团



世界图书出版公司

《中国语学文库》

总主编：邢福义

副总主编：汪国胜 朱 斌

本书是教育部人文社会科学重点项目“二十世纪书面汉语使用状况统计与分析(2004ZDI105-B01)”和“中国人民大学文学院科学研究基金项目(RMULA-14002)”的项目成果之一。

现代汉语三音节词语研究

A Study of the 3-syllable Words and
Expressions in the Contemporary Chinese

杨书俊◎著



中国出版集团

世界图书出版公司

广州·上海·西安·北京

图书在版编目(CIP)数据

现代汉语三音节词语研究 / 杨书俊著. —广州:
世界图书出版广东有限公司, 2015.4
ISBN 978-7-5100-9580-1

I. ①现… II. ①杨… III. ①现代汉语—音节—研究
IV. ①H116.3

中国版本图书馆 CIP 数据核字 (2015) 第 072580 号

现代汉语三音节词语研究

策划编辑 刘婕妤

责任编辑 黄 琼

出版发行 世界图书出版广东有限公司

地 址 广州市新港西路大江冲 25 号

[http:// www.gdst.com.cn](http://www.gdst.com.cn)

印 刷 北京天正元印务有限公司

规 格 710mm × 1000mm 1/16

印 张 12

字 数 167 千

版 次 2015 年 4 月第 1 版 2015 年 4 月第 1 次印刷

ISBN 978-7-5100-9580-1/H · 0939

定 价 36.00 元

版权所有, 翻版必究

汉语词汇复音化的新趋势

——序杨书俊《现代汉语三音节词语研究》

农历乙未年大年初四，收到杨书俊电邮，说她的博士论文做了较大修改后，入选“中国语学文库”，请作序。假期，对于大学教师来说，是以特殊方式工作的时期。书俊的春节，显然是一个专心科研的春节。

书俊的家乡在长江西陵峡畔的秭归，那是三闾大夫屈原的故乡，是三峡大坝的矗立地。书俊早年就读于湖北民族学院，之后负笈彩云之南，获取云南师范大学文学硕士学位。2002年考入华中师范大学攻读博士学位，与我便有了师生之缘。2005年，书俊获取博士学位，谋得中国人民大学的一席教职。她报到后跑来见我，把工作证给我看，十分激动。书俊从学校到学校，这是她的第一份工作，也是她愿意努力做好的工作。

我常告诫学生，女士为学，要比男人艰辛数倍。是的，转眼十年，书俊成家生子，还受命到哥本哈根商学院执教两年。这期间，教书理家，可以想见有多少困难。但出人意料，她竟没有忘记科研，没有忘记她的博士论文，积累资料，反复思考，终于改定，交付出版。

书俊的论文《现代汉语三音节词语研究》，最大的特点就是采用了定量研究方法。她选取《现代汉语词典》（2002，增补本）、《二十世纪书面汉语使用状况统计与分析》和北京大学现代汉语语料库中的三音节词语作为具体分析

对象,利用语料库语言学的方法,得到了大量有价值的数 据,通过这些数据得到了不少规律。此次修订最明显的,是增加了汉语方言和民族语言的内容。方言选取了长沙、苏州、广州和厦门地区的,民族语言选取了壮侗、藏缅、苗瑶等三个语族的。这增加了巨大的工作量,也扩大了学术比较的视野,能够更为全面地了解三音节词语。

汉语词汇从古到今发生的最大变化就是“复音化”。上古汉语以单音节词为主,之后逐渐增加双音节词,或是将单音节词变为双音节。大约到东汉魏晋南北朝的中古汉语时期,双音节化现象就较为繁盛,新生的双音节词数量已远超新生单音节词数量。之后,历经唐宋元明清,由近代汉语发展到现代汉语,双音节已经居于汉语词汇的主体地位。

双音节词语中,有“蝴蝶”“凤凰”“咕咚”“哗啦”等单纯词,但更多的是合成词。有了合成词,就有了构词法;词汇双音节化从另一角度看,也是词法的产生和发展。单音节的造词能力、单纯词的产生数量,都是有限的,因为每一个事物都须用一个词表示。由于语言的音节数量有限,人的记忆能力有限,故而只有单纯词,特别是单音节词的语言,必然集聚大量的同音词,使语言的学习和应用较为困难。而合成词则不然,它将两个或多个语素用词法组合起来,不仅造词能力呈几何级数般的增长,而且词语内部因有语法结构而使词义较为透明,词语之间因有语法关系而易成系统。比如乙未年是中国的羊年,英语中没有“羊”这个上位词语,“ram(公羊)”“goat(山羊)”“sheep(绵羊)”又都是单纯词,词形之间没有语法联系。而汉语不仅有上位词“羊”,“公羊”“母羊”“山羊”“绵羊”还都是偏正结构,用“区别性语素+羊”构成词汇场。中国人过羊年,没人计较是“公羊年”“山羊年”,还是“绵羊年”,而英语要表述中国的羊年,便因弄不准是何种羊而产生点小郁闷或是小情趣。美国总统奥巴马 2015 年 2 月 18 日(中国农历除夕)向全球华人祝贺新年,幽默地说:“管它是公羊、山羊还是绵羊,我祝你们新年快乐。”

双音节合成词,分复合词与派生词两大类。复合词的词法结构与汉语的句法结构基本一致,都有联合、主谓、偏正、动宾等结构关系。词法与句法的这

种一致性,是汉语语法的一大特点,也是一大优点。当然,这并不是说汉语的词法与句法是等同的,比如“纸张”“车辆”“灯盏”“马匹”这种“名词性+量词性”的结构,句法中就没有。

双音节词的大量产生并逐渐成为词汇的主体,与汉语语音系统的发展演变也具有相关性。有研究认为,汉语语音系统的简化导致了词汇的复音化;也有人认为,或许是词汇复音化才使语音系统有可能简化;或许还可以认为,语音、词汇的发展各有其内在规律,但也相互影响,即语音系统的简化与词汇的复音化相互促进。词汇与语音的共变关系如何,本书难以结论,也不必结论。此时更关心的是,21世纪汉语词汇的复音化又出现的新趋势:

为了给杨书俊的著作写序,2月23日(大年初五),我发电邮请教对新词语持续进行研究的侯敏教授。侯教授也是以科研度假的语言学家,正在撰写《语言监测十年》和咨询报告《网络语言规范热议》。她慷慨提供了2006—2014年的有关数据。近九年来,从中国语言资源监测中心的海量语料库中,共提取、甄别了5264个新词语。这些新词语以三音节居多,占46.26%;双音节、四音节分别占21.56%和22.89%。三音节占优势与这些年来多利用词语模造词有关,以后缀“族”造成的“~族”词语就有333个,“~门”162个,“~哥”69个,“~客”65个,“微~”212个,“被~”66个,“云~”57个。据统计,利用词语模造成的新词有1300多个,占新词语的25%左右。

这些数据很有意思。在现代汉语词汇的存量中,双音节当然还是优势音节。但从新造词的角度看,单音节几乎丧失了造词能力,三音节新词超过双音节、四音节新词的总和,是双音节新词的2倍以上,也是四音节新词的2倍以上。

三音节词语当然古已有之,但成为新词语的优势音节模式却是汉语复音化的新趋势。汉语由单音节优势转化为双音节优势,引发了学者的一系列探讨。在此三音节的新趋势下,也有许多问题更值得探问,比如:新造词由双音节向三音节转变的动因是什么?假以时日,汉语词汇总量是否也会以三音节为主体?这种变化会对语音系统、语法系统产生什么影响?会不会影响长久形成的韵律感知及语言审美?将来会不会出现“四音节优势”?

面对汉语复音化的新趋势,更觉得书俊当年关于三音节词语的选题有意义,为此所花费的十余年的精力有价值。今日恰逢元宵节,自然想起了当年读过的唐代崔液《上元夜》:“玉漏铜壶且莫催,铁关金锁彻明开;谁家见月能闲坐,何处闻灯不看来。”崔诗以委婉笔法,妙写元宵夜之盛:谁能见月闲坐?谁能闻灯不看来?对于三音节词语的上述话题,我想就如同十五观灯一样,书俊也许不忍“闲坐”,不忍“不看”,还会继续进行观察与思考,有些新发现,做出新贡献。

李宇明

2015年3月5日

农历乙未年元宵节

序于北京惧闲聊斋

目 录

第1章 绪 论	/001
1.1 问题的提出	/001
1.2 三音节词语研究现状	/002
1.3 语料来源和研究内容	/008
第2章 三音节词语溯源	/013
2.1 三音节词语发展轨迹	/013
2.2 现代汉语三音节词语发展原因探索	/017
第3章 现代汉语三音节词语统计与分析	/027
3.1 现代汉语三音节词语分布情况统计	/027
3.2 汉语三音节不平衡性及其原因分析	/046
3.3 《二十世纪》与《现汉》三音节比较	/054
第4章 汉语三音节造词法及其发展	/067
4.1 汉语三音节造词法分析	/067
4.2 汉语三音节造词方式的发展	/080
第5章 三音节词语的压缩与扩充	/088
5.1 三音节词语的压缩	/088
5.2 三音节词语的扩充	/095
5.3 结 语	/101

第6章 三音节词语的语法功能及其演变	/102
6.1 三音节词语的词法和句法功能	/102
6.2 三音节词语结构的重新分析	/105
6.3 双音节+“着”“到”“于”的分析	/114
6.4 补充结构 $V_1 + \text{不} + V_2 (\text{adj})$ 的虚化过程及机制	/119
6.5 三音节连词的语法化过程及机制	/124
6.6 结 语	/131
第7章 基于比较视角的三音节词语考察	/133
7.1 现代汉语共同语和方言里的三音节词语比较	/133
7.2 跨语言的三音节词比较研究	/156
参考文献	/172
后 记	/180

第 1 章 绪 论

1.1 问题的提出

现代汉语从音节角度来说，双音节词语在一般情况下分布频率都在 60% ~ 70%。单音节数量虽不大，但也仍然十分活跃，尤其是单音节动词和形容词活跃度更高。三音节词发展较快，其在现代汉语中的比例逐步提高，已经成为汉语新词新语产生的重要音节形式。四音节除成语较为凝固外，其他大多是两个双音节的组合（少数为 3+1 或 1+3 式，如“自来水笔”“副研究员”），由于现代汉语双音节自主性较强，两个双音节组合时通常都被看作短语。四音节以上除了歇后语、谚语以及专业术语外，一般都是短语。也就是说，汉语词的音节数主要是一到四个音节。

单音节由于是汉语构词的基座和要件，所以成果很多。双音节是现代汉语的主体，研究成果更是“汗牛充栋”。三音节惯用语、四音节成语以及四音节及以上的歇后语、谚语等，由于其特性也受到较为广泛的关注，成为教材词汇教学内容的一部分。相对而言，三音节词语在汉语词汇中研究较少、较薄弱，尤其是普通的三音节词语研究成果在 20 世纪 80 年代以前更是不多。随着社会节奏的加快，现代汉语产生了大量的新词新语，其中有很多都是三音节词语，所以 20 世纪 80 年代以后，三音节词语开始受到学界关注。三音节词语研究成果逐渐增多。但从目前来看，三音节词语研究成果较少，且多倾向于定性、历

史性以及验证性研究,研究涉及面较宽,而集中度不足。鉴于此,本书专门以三音节词语为研究对象,从定量角度出发,静态地统计了《现代汉语词典》(第4版)中三音节词语的分布情况,然后从动态角度出发,统计了《二十世纪》(“二十世纪现代汉语语料库”)中的现代汉语三音节词语分布情况,并对两者进行比较研究,探讨三音节分布不平衡和三音节成词难易度的原因。然后在定量研究基础上,从定性角度出发,探讨三音节成词原因、三音结构词方式以及某些三音节词语的句法问题。

1.2 三音节词语研究现状

20世纪80年代以前,有关三音节词语的研究成果很少,主要散见于论文和个别专著的章节中,但是这些研究成果见解深刻,提出了一些令人深思的问题。较早论述三音节的是陆志韦先生,他在《北京话单音词词汇》序中认为北京话中:“多音词多半是两个音的。四音的也不在少数……三音的并不多。五个音以上的简直是少极了”。“我也明知汉语的多音词会一天一天地加多,特别是多音的名词,其次是形容词……”^[1]在随后出版的《汉语构词法》里,陆志韦先生(1957)详细列举了汉语的各种构词类型,其中也较大篇幅地讨论了汉语三音节的各种构词类型。吕叔湘先生在《现代汉语单双音节问题初探》一文中,从单音节的活动受限制、双音化的倾向以及双音节四音节搭配等角度讨论了现代汉语单双音节的构成以及互相转化问题,三音节动宾结构和偏正结构的数量及成词问题,三音节在句子中的扩充和缩略问题,此外还特别讨论了一种特殊的“截搭式”造词方式。尤其值得一提的是,吕先生在文中还提出了一些很有见地的思想,如在讨论到2+1式比1+2式多得多时,他提到“其中2+1式(如‘动物学’,‘示意图’,‘辩证法’,‘可见度’)比1+2式(如‘副作用’,‘手风琴’)要多得多,跟在前或在后的单字的性质和可以这样用的单字数量有关系,还可以进一步研究”;在论述到动宾结构1+2式多于2+1式时,他又说到“这跟常用动词中单音的较多有关系,可是是否完全由于这个因素,也还需要进一步分析”。此外,吕先生还谈到了按音节分词的想法,“如果把数目作为一个重要条件,似乎也可以把双音节(和三音节的)和四音节的(和四音

[1] 陆志韦编著:《北京话单音词词汇》序,科学出版社1956年版,第9—10页。

节以上的)分别对待,不按同样的标准处理”^[1]。这些思想都是富有启发意义的,也成为我们进一步研究的起点和探索的基点。70年代末,另一位语言学大师赵元任在《汉语口语语法》里也论述了三音节问题,他从词性的角度列举三音节结构类型,如A-NN型“红十字、咸鸭蛋”;AN-N型“白皮书、孤儿院”;VO-N型“起重机、显微镜”;NN-N型“帆布鞋、海岸线”;VV-N型“驱逐舰、降落伞”;AA-N型“长短句、酸辣汤”;VR-N型“放大镜、漂白粉”;其他复合词SP-N型“地震仪、人来疯”,V-NN型“炒牛肉”V-AN型“拌黄瓜、失乐园”等。还特别从节奏角度论述了汉语三音节的1+2式和2+1式的重音情况^[2],如三音节复合词,“重轻重”的节奏类型总是2+1,如“豆腐浆”。“重重轻”的类型一般是1+2,如“炸锅巴”。没有轻声字的时候,1+2比2+1略多。此外赵先生还特别从音节和节奏角度论述了汉语三音节词语的变调情况。^[3]他认为在三音节组中第二声变成第一声。如在三音节词或词组ABC中,A是第一声或第二声,B是第二声,C是轻声之外的任何声调,那么在一般会话速度的语流中,B就会变成第一声,但在慢速说话中不会改变。陆先生、吕先生和赵先生,从构词类型、单双音节配合、节奏和变调角度,对三音节词语进行了基础性但同时又是开拓性的研究,研究角度涉及构词、音节搭配、轻重音和变调,也基本涵盖了后学的研究重点。

80年代以后,尤其是近年来,随着三音节词语的日益增加,一些学者开始关注三音节研究,并从各个角度深化研究三音节词语。大致可以归纳为如下几个方面。

1.2.1 从历时角度考察三音节

胡敕瑞(2002)在对《论衡》和东汉佛典两份语言资料进行了专书研究和相互比较以后,认为东汉“表示一般事物名的三音词语较之上古有明显增多,特别是表示一般事物名的三音词语,东汉以前很少见到”,并指出“大量三音节的外来译词似乎对汉语三音词的发展起到了推波助澜的作用”,“夫馀国”“摩尼珠”等音意兼译的三音词语的出现“不但增添了三音词语的数量,而且巩固

[1] 吕叔湘:《现代汉语单双音节问题初探》,载《中国语文》1963年第1期。

[2] 赵元任《汉语口语语法》,商务印书馆1979年版,第223页。

[3] 赵元任《汉语口语语法》,商务印书馆1979年版,第22页。

了一种三音节构词的范式”。

陈宝勤(2002)在其《汉语造词研究》一书中更是对远古、上古、中古、近古等不同时期的偏正、支配、陈述、补充和附加方式构造三音词进行了系统历时研究,分别考察了各种语法造词法的发展线索、内部构成、内部词性等问题。如她在对偏正构造三音节词语考察后,认为“偏正造三音词语萌发于远古,生长于上古,发展于中古,成熟于近古;新的偏正式三音词系统在近古晚期形成;旧的偏正式三音词系统在近古晚期衰亡,只存留于古文献中”,“偏正所造三音词一般为名词,且一般为离合名词;偏正式三音离合名词的意义一般是其构成成分意义的融合”。^[1]“汉语三音以上的多音词自古及今也在逐渐生长,日趋增多。汉语三音以上多音词的日益增多主要是汉语自身发展之因,语词的发展一般是由单音词到双音词而再到多音词;另是受外族词的影响,上古匈奴语的传入,中古梵语的传入,近古蒙语、满语的借入,现代英语、法语、德语、俄语等的传入都促进了汉语三音以上多音词的发展。”^[2]

黄凤显(2003)从韵律的角度考察了“屈辞”中的“三字结构”后,认为“屈原大量使用‘三字结构’所创造的奇偶相谐的诗歌句式,不仅构成了屈辞新诗体的要素,而且对于中国古代诗歌句式及其语言形式的发展产生了重要而深远的影响”。

1.2.2 从内部结构角度研究三音节

李赓钧(1992)针对汉语三音节词语中的“三语素合成词”日益增多的趋势,对“凡在结构上和意义上都已定型的三语素结构”进行了考察。考察后发现“它们绝大多数包括两个层次,其构造格式为:A式(xy)z, B式x(yz),以A式为多。构词方式包括复合式和附加式两大类,其中偏正关系复合式占大多数,以名词为最多。因此,三语素合成词的主体是A式偏正关系复合式名词”。

刘丹青(1996)在其《词类和词长的相关性》一文中谈道:“在名词的各项功能中,双音节、三音节是自由的”,而“动词和动词性短语的基本界限在双音节内部,三音节和三音节以上单位将继续保持短语性、难以进入词的行列”。

郑庆君(2003)通过考察三音节合成词尤其是三音节偏正式合成词的结构

[1] 陈宝勤:《汉语造词研究》,巴蜀书社2002年版,第180页。

[2] 陈宝勤:《汉语造词研究》,巴蜀书社2002年版,第380页。

类型及层次,也得出了“汉语三音节合成词强烈的名词化倾向”的结论。

差不多与此同时,汤志祥(1994)对广州方言中的“三字格词语”的语法功能、构成和语源、语义和语用等进行了考察,并把它们与普通话比较,认为“广州话三字格词语不但动宾式多,而且偏正式也多,相比之下普通话主要是偏正式,很少是动宾式,这也不能不说是广州话的一大特色”。

上述有关三音节词语的共时论述大多集中在现代汉语平面,杨爱姣(2002)独辟蹊径,采用以“传统的结构形式派、传统的语义派”为主的分析方法,从“三音词的发展原因”“三音词的结构方式”“三音词的语义构成”“三音词的语法功能”“三音词的修辞特点”等五个方面对近代汉语中的三音节词语进行了共时平面的考察。

1.2.3 三音节是否成词的问题

三音节是词还是短语学术界一直众说纷纭,大部分学者认为有一部分三音节是词,大多数是短语,有的认定三音节是否为词应视三音节内部结构的差异而定,偏正结构倾向于成词,而动宾结构主要为短语,但是这些看法都是一种个人的感知,并不具有理论基础和认知理据。近年来,一些语言学家开始重新考虑这些问题,并尝试从韵律和社会语言学的角度来解决这一问题。冯胜利从韵律角度提出,由三个音节组成的音步叫“超音步”。“超音步”一般遵循“左向造语,右向构词”规律。他以形式句法为基础,强调了韵律对句法的主动作用,并参照其中成分的句法关系可分为1+2、2+1的形式,其中2+1式可以成词,而1+2式由于不能构成韵律词,因此要么不能说,要么只能是短语。吴为善(1986、1989、2003)则分别从三音节的组合规律、三音节中单音节的黏附性以及韵律角度讨论了三音节是否成词的问题,他认为三音节内部结构差异以及能否成词与三音节节律层面的松紧和句法结构层面的松紧相适应,而最终决定因素则是认知语义。端木三、陆丙甫(1991、2002)从“辅重原则”和“辅长原则”出发,比较和探讨了现代汉语中的2+2、1+1、2+1、1+2式成立与否以及转换的限制因素。王洪君(1994、2000、2001)从音节搭配的角度对三音节进行了较为详尽的分析,她认为三音节内部结构差异是因为汉语史上双音节化历程中名词与动词、形容词、副词等的发展差异和单双音节的语法功能分化造成的。王灿龙(2002)从认知语言学“相邻原则”和“相似原则”出发,认为

动宾结构中 1+1 和 2+2 式是常规的无标记组合形式,而 1+2、2+1 式则是有标记的组合形式。柯航(2012 年)则从单双音节搭配的关联标记模式角度出发,综合了前人的研究成果。

但是他们的判定也遭到了部分语言学家的质疑和反对,如周荐(1998)在《词汇学问题》一书中指出,判定一个三字组合是词还是短语,不能根据其结构是 1+2 式还是 2+1 式,而要看该组合的语法属性和结构关系的性质,2003 年他又在《三字组合与词汇单位的确定》一文中明确提出了“在三字组合词语的判定上起根本作用的还是意义的整体性”。尽管目前三音节是词是短语的问题没有得到很好的解决,但是他们对汉语三音节是否成词问题的研究,比前人研究还是前进了一大步。

另一些语言学者则从社会语言学角度对人们的“语感”或“词感”中的三音节是否成词问题进行了调查研究。王立(1998、2002、2003)在调查问卷的基础上,对公众观念中什么是“词”的问题进行了探索,发现公众眼中的“词”和专家眼中的“词”存在着很大的差距;同时也发现虽然三音节词语词感倾向远远不及双音节强烈,而且内部各种结构词感倾向也不尽一致,但是大量三音节的产生,使许多三音节词语词感倾向增强。“三音节定中结构呈现强烈词化倾向,这与早期双音化的发展极为相似;三音节数量结构(如“一片片”“数千年”等)呈显著词化倾向,三音节动补(动趋)结构所表现的准词化倾向,说明有向词化异动的可能,但发展前景并不十分明朗。”^[1]陈松岑(2004)在北京大学校园里也进行了一项词语社会调查,调查主体是大学生,结果同样发现,专家眼中的“词”与公众眼中的“词”相去甚远。^[2]胡明扬先生(1999)则从理论的角度阐释了专家眼中和公众眼中的“词”和“短语”差别的种种依据。并且主张在对待双音节和三音节偏正结构中一些既可以判定为词也可以判定为语的词语应该从宽。

上面两个角度,一个重理论,一个重社会调查,但他们的研究成果却有“异曲同工”之妙。从韵律、句法和语义角度研究表明,偏正结构的三音节词语是最容易成词的,动宾结构以及其他结构成词难度较大;从社会语言学调查的情

[1] 王立:《汉语词的社会语言学研究》,商务印书馆 2003 年版,第 255 页。

[2] 陈松岑:《规范与变异——关于〈现代汉语规范词典〉的收词》,载《语文研究》2004 年第 3 期。

况来看,三音节偏正结构在公众心目中也是最容易成词的,其他结构成词的节拍明显慢于偏正结构。而且韵律词所能包含的词也远比一般专家眼中的“词”要大,而与公众眼中的“词”有相当的契合。

1.2.4 三音节发展趋势的预测问题

三音节将如何发展,也是一些研究者关注的焦点,姚汉铭(1998)在统计了近3000个新词语后认为,如果说“双音化是现代汉语的主要节奏倾向”只是对“上古汉语向中古汉语、近代汉语乃至五四后的相当时期现代汉语发展的词汇外形特点的概括,那还是正确的。但是,新时期新词语的发展事实却修正了这一观点:当今的汉语词汇正在向以四音节、三音节为主的多音词发展”。

在稍后的一段时间里,杨爱姣、王军、王立等人发表了与之相反的看法:

杨爱姣(2002)在考察了近代汉语中的三音节词语以后,预测了三音词的发展趋势:“当双音词扩展为三音词、四音词压缩为三音词的同时,三音词压缩为双音词的现象也就同时开始了。双音化是汉语发展的主要趋势……这种矛盾斗争的结果就是汉语词汇的双音化依然占优势,三音词有节制地发展。”

葛本仪主编的《汉语词汇学》(2003)一书中谈到“双音化趋势”时提到“当前汉语中,不但双音词大量增加,而且也出现了不少的三音节词和四音节词等”,“在词的音节数量上,一般要经历一个‘单音——双音——多音——简缩为双音或三音’的过程”,并指出“就现在的情况看,双音词的数量仍然占着优势”^[1]。

王立(2003)在《汉语词的社会语言学研究》一书中提到,目前三音结构只是在局部范围内成词,即只在名词性三音结构上实现。“三音词是在双音词的基础上发展起来的,初步的发展迹象显示,三音词将有节制地发展”,“汉语构词的双音格局将不会被轻易打破,汉语三音构词只是双音构词的一种补充形式”。

他们虽然在多音节是否将成为汉语词汇增长的主要趋势问题上存在争议,但是他们也都同意多音节(主要是三音节和四音节)将会有节制地发展。

以上关于三音节词语的研究虽然角度新颖和多样,而且提出了许多新问题,但是这些研究大多是建立在研究者的内省式或举例式阐述的定性分析基础之上的,虽然也有个别研究者把三音节词语的研究建立在规模化穷尽性的定量考察

[1] 葛本仪主编:《汉语词汇学》,山东大学出版社2003年版,第698页。

基础之上,如:刘云(2001)“为研究文本信息处理中无可避免的未登录词(即机器词典中没有的词)的识别问题,为研究现代汉语合成词的构造规律问题”开发了一个三音节词的构词库,对词库中的三音节词的词性、构词方式和构造层次进行了统计,并把三音节词的各种构词方式所占的比例与双音节词进行了比较,发现“三音节词与双音节词的构造是有较大差别的”。但是总的来说,这些研究大多是以双音节作为主要基点,论述三音节的某些特殊性质,而没有将三音节作为一个研究基点,系统地统计三音节词语内部结构各种分布比例和形式特征,全面分析三音节词语的基本属性,三音节词语内部发展不平衡的原因,以及三音节词语的发展趋势等。

1.3 语料来源和研究成果

1.3.1 语料来源

本书以定量研究为主,辅以定性研究,语料主要来源于以下三个语料库。

1.《现代汉语词典》(2002,增补本)中的所有三音节词条

语料库性质:熟语料、中型现代汉语词典。

语料库容量:60 000 多条,其中三音节 5 100 余条。

语料时段:现代。

收词地域:中国大陆。

语料内容:一般语汇;常见的方言词语;不久以前还使用的旧词语;书面上常见的文言词语;某些习见的专门术语;地名、人名等。

编列方式:按拼音字母次序排列。

出版单位:商务印书馆。

语料库用途:鉴于该语料库是经过专家及群众认可了的熟语料库,主要用来统计和描述现代汉语三音节词语的基本属性,并作为其他章节的用例。

2.《二十世纪书面汉语使用状况统计与分析》中离析出来的三音节词语

语料库性质:分词语料库。

语料库容量:7 000 万字,词条总数 87 448 条,其中三音节词语约为 11 894 条。