

外国学生汉语字词学习的影响因素^{*}

——兼论《汉语水平大纲》字词的选择与分级

江 新 赵 果 黄慧英 柳燕梅 王又民

提要 本文针对《汉语水平大纲》修订的需要,对外国学生汉语字词学习的影响因素进行了实验研究和语料统计研究,在此基础上对《汉语水平大纲》字词选择和分级提出了建议。论文分为四个部分,首先对前人关于外国学生汉语字词学习影响因素的研究进行概述,然后对外国学生字词学习影响因素进行实验研究,再对外国学生汉语字词使用情况进行统计分析,最后对《大纲》字词选择和分级的原则进行讨论。

关键词 字词学习;字频;词频;构词数

《汉语水平大纲》(以下简称《大纲》)中字词的选择与分级是一项非常重要的研究课题。大纲汉字、词汇的选择和分级标准的制定,要考虑的因素很多,不但要依据基于汉语母语者语料的常用字词的统计学研究(字或词的出现频率、字的构词能力等),也要考虑外国人汉语字词学习的规律。影响外国人字词学习的因素有哪些?哪些字词先学会、容易学会,哪些字词后学会、不容易学会?对这些问题进行实证的定量研究和定性研究,可以从外国人字词学习的角度为《大纲》字词选择、分级、排序提供科学的理论依据。

一 相关研究概述

已有的关于本族人学习汉语的研究表明,从汉字属性特征的角度看,影响汉字学习的主要因素有:(1)出现频率。出现频率高的汉字比出现频率低的汉字易学。(2)字形复杂性。字形复杂性包括笔画数、部件数、笔画类型、字形对称性等特征。笔画数少的汉字比笔画数多的汉字易学,部件少的汉字比部件多的汉字易学,笔画横直的汉字比笔画斜曲的汉字易学,字形对称的汉字比字形不对称的汉字易学。(3)结构类型。左右结构的字比上下结构的字易学,左右、上下结构字比包孕结构字易学。外国人汉字学习是否具有类似的规律,还需进一步研究。

近十年来,针对外国人的汉语字词学习的研究逐渐增多。该领域进行的实证研究主要包括以下几个方面:(1)汉字的频率、字形复杂性、构词数等属性特征对汉字学习效果的影响(Sergent & Everson 1992,周清海、梁荣基 1995,柳燕梅 2002,江新 2004,赵果 2003);(2)汉字声旁意识和形旁意识(Shen 2000,江新 2001,陈慧、王魁京 2001);(3)汉字学习策略(McGinnis

^{*} 本研究为国家汉办专题研究项目(项目号 HB K01 - 05/072),语料收集和整理工作得到北京语言大学研究生王蕾、吕婵、王锦玉、李传燕、高珊、鄢胜涵、朱昱、李俊岩等的协助,在此深表谢意!

1999, Ke 1998, 江新、赵果 2001, 赵果、江新 2002); (4) 汉字识别与汉字书写的关系 (Ke 1996); (5) 汉语字词知音与知义的关系 (Everson 1998, Yang 2000, 江新 2003a); (6) 汉字教学方法 (柳燕梅、江新 2003); (7) 接受型词汇与表达型词汇之间的距离 (Lin 2000, 鹿士义 2001); (8) 汉字正字法意识 (鹿士义 2001, 江新 2003b); (9) 汉字书写错误 (江新、柳燕梅 2004); (10) 汉语虚词的学习 (Wen 1995, 1997, Teng 1999, Polio 1994, Kuo 2000, Bourgerie 1996)。汉语虚词的学习常常被看成是语法习得的范围。我们研究的重点为实词的学习。

外国人汉语字词学习研究取得了一些成果,但迄今为止,还很少有人从《大纲》字词选择和分级标准的制定需要出发,来系统总结和归纳已有的字词学习研究成果,并就外国人字词学习的影响因素进行有针对性的实证研究。本文拟从两个方面进行研究:

(1) 实验研究:对外国人字词学习的影响因素进行实验研究。包括汉字使用频率、构词能力、字形复杂性(笔画数、部件数)等因素对汉字学习的影响;词语的使用频率、字的使用频率等因素对词语学习的影响。

(2) 语料分析:对外国人汉字和词语使用的情况进行统计和分析。包括不同阶段留学生实际拥有的汉字量和词汇量(教学字词量不一定等于学生实际拥有的字词量),并将留学生作文常用词与《大纲》词汇进行对比。

二 外国人汉语字词学习的影响因素的实验研究

2.1 实验 1: 汉字出现频率、构词数、笔画数、部件数与汉字学习的关系

2.1.1 方法

被试为在北京语言大学学习汉语的一年级第一学期的非汉字圈国家留学生 28 人,他们是以英语、法语或俄语等语言为母语、汉语作为第二语言的汉语初学者,学习的教材为杨寄洲主编的《汉语教程》(北京语言文化大学出版社 1997),每周 24 学时,测验时学习时间为 4 个月,近 400 个学时,学完了 1 至 33 课。

从《汉语教程》的前 33 课课文和生词表中出现的汉字中随机选取 102 个字进行测验。要求被试根据字形写出汉字的拼音并组词或短语。然后计分,拼音计时声母和韵母正确就计为正确,忽略声调;意义计时词或短语的意义正确就计为正确,忽略字形。

同时,统计每个汉字的四个属性:(1)出现频率。统计汉字在前 33 课的课文和生词表中出现的次数,作为汉字在教材中的出现频率(也称为教材字频);(2)构词能力。统计每个汉字在前 33 课的生词中参与构词的次数,作为汉字的构词能力;(3)笔画数;(4)部件数,统计以上海交通大学汉字编码组主编的《汉字信息字典》中的数据为准。

2.1.2 结果

采用相关分析的统计方法对数据进行分析,探讨汉字的出现频率、构词能力、笔画数、部件数与汉字学习效果的关系,结果见表 1。

表 1 汉字的频率、构词数、笔画数、部件数与汉字拼音、意义成绩之间的简单相关

	频率	构词数	笔画数	部件数
拼音正确率	0.416	0.562	-0.333	-0.328
意义正确率	0.414	0.590	-0.326	-0.319

(所有 $p < 0.01$)

表 1 简单相关分析的结果表明,汉字的频率、构词数、笔画数、部件数与汉字测验成绩都有

显著相关。从拼音成绩看,汉字的频率、构词数与汉字拼音成绩有显著的正相关关系,即汉字频率越高,构词数越高,汉字测验成绩越好。汉字的笔画数、部件数与汉字拼音成绩有显著负相关,即笔画数越多、部件数越多,汉字拼音成绩越差。意义成绩也有相同的统计结果。

但是,由于频率、构词数、笔画数和部件数这几个变量之间可能存在相互关系,因此,为了研究在其中两个变量的影响保持恒定的情况下另一个变量与汉字成绩之间的相关,我们对汉字的频率、构词数等因素与汉字测验成绩之间的关系进行了偏相关分析,见表 2。

表 2 汉字的频率、构词数、笔画数、部件数与汉字拼音成绩、意义成绩之间的偏相关

	频率	构词数	笔画数	部件数
拼音正确率	0.140	0.369 ^{**}	-0.091	-0.055
意义正确率	0.114	0.415 ^{**}	-0.089	-0.039

(^{**} $p < 0.01$)

由表 2 可见,对于频率、构词数、笔画数和部件数这四个因素,当控制了其中三个因素对某个因素以及汉字拼音成绩的影响,与拼音成绩的相关系数最大的就是构词数,其次是频率,最小的是笔画数、部件数。具体地说,当控制了频率、笔画数、部件数对构词数与汉字拼音成绩的影响,构词数与汉字拼音成绩的相关系数仍然非常显著;但是,当控制了构词数、笔画数、部件数对频率与汉字拼音成绩的影响,频率与汉字拼音成绩之间的相关变得很弱,相关系数不显著;同样,当控制了其他三个因素的影响,笔画数、部件数与汉字拼音成绩之间的偏相关都不显著。对四个因素与意义成绩的偏相关分析也得到了相同的结果。

从简单相关看,汉字的出现频率、构词数、笔画数、部件数与汉字学习效果都是相关的;从偏相关看,汉字的构词数与汉字学习效果的关系最为密切。本实验只是一个相关性研究,要知道这些因素与汉字学习效果之间是否存在因果关系,还必须进行控制比较严格的实验研究。

2.2 实验 2: 整词出现频率、单字出现频率与整词学习的关系

以非汉字圈国家的留学生为被试,从留学生已学过的教材的词表中随机选择一定数量双字词和多字词进行测试,要求被试进行识别(写出词语的拼音并组成短语),采用相关分析的统计方法对数据进行分析,探讨整词的出现频率、单字的出现频率与整词学习效果的关系。

2.2.1 方法

被试与实验 1 的被试相同。测验材料是从《汉语教程》的前 33 课的生词表(包括人名、专名和补充生词)中随机选取的 76 个双音节和三音节词。要求被试根据词形写出词的拼音并写出意义(用被试熟悉的语言来写)。计分原则同实验 1。

同时,统计每个词的两个属性:(1)出现频率。统计整词在前 33 课的课文中出现的次数,作为整词在教材中的出现频率;(2)单字出现的频率。统计词中的每个汉字在前 33 课的课文中出现的次数,作为单字的出现频率。

2.2.2 结果

表 3 整词频率、单字频率与整词学习成绩之间的简单相关

	整词频率	单字频率
拼音正确率	0.276	0.464
意义正确率	0.322	0.481

(所有 $p < 0.01$)

表 3 表明,整词频率、单字频率与整词测验成绩都有显著相关,频率越高,测验成绩越好。

但是,由于整词频率、单字频率这两个变量之间可能存在相互关系,因此,为了研究在其中一个变量的影响保持恒定的情况下另一个变量与测验成绩之间的相关,我们对整词频率、单字频率与整词测验成绩之间的关系进行了偏相关分析,见表 4。

表 4 整词频率、单字频率与整词学习成绩之间的偏相关

	整词频率	单字频率
拼音正确率	- 0.052	0.391 ^{**}
意义正确率	0.000	0.378 ^{**}

(** $p < 0.01$)

由表 4 可见,当控制了单字频率对整词频率与整词拼音成绩的影响,整词频率与汉字拼音成绩的相关系数不显著;但是,当控制了整词频率对单字频率与整词拼音成绩的影响,单字频率与整词拼音成绩之间的相关仍然非常显著。对两个因素与意义成绩的偏相关分析也得到了相同的结果。

从简单相关看,整词的出现频率、单字的出现频率与整词学习效果都是相关的;从偏相关看,单字出现频率与整词学习效果的关系最为密切。实验 2 也只是一个相关研究,要知道整词的出现频率、单字的出现频率与整词学习效果之间是否存在因果关系,还必须进行控制比较严格的实验研究。

2.3 结论

综合以上两个实验的结果,我们得到以下初步结论。

(1) 外国人的汉字学习效果受汉字出现频率、构词数、字形复杂性(笔画数、部件数)的影响。出现频率越高、构词数越多、笔画数和部件数越少,学习效果越好。这个结论提示我们,要选择高频的常用的汉字教给学生,而且要按照频率高低和常用程度来决定汉字在大纲中的等级,高频的常用的字先学习,等级较低;低频的不常用的字后学习,等级较高。还要选择构词能力强的汉字教给学生,而且要通过提高生字在学生已有词汇中的构词数,来提高生字的学习效果。此外,字形简单的字先学习,字形复杂的字后学习。

(2) 外国人的词语学习受整词出现频率和单字出现频率的影响。整词出现频率越高,单字出现频率越高,词的学习效果越好。因此,我们要选择出现频率高的合成词教给学生,同时也要重视选择出现频率高的汉字教给学生。选择基础的常用字教给学生,既有助于单字的学习,也有助于词汇的学习。

三 外国学生汉字和词汇使用的统计分析

3.1 留学生语料的选择

我们收集了不同国家、不同年级的外国学生汉语作文语料。所有语料都是学生在考试时独立完成的作文,是比较自然、真实的语料。本研究以其中一部分语料为基础,来统计分析留学生作文用字和用词情况。

从所收集的作文语料中分别随机选择不同国家(韩国、日本、东南亚、非汉字圈)、不同水平(初级、中级、中高级) 12 - 20 人的主题相似的语料(共 107333 字)进行统计分析。选取语料时尽可能使语料来源在学生的国籍、汉语水平这两个因素上是平衡的,语料的主题也基本一致(理想的做法是,使语料包括多个主题、而且语料的主题在不同国家和学生的汉语水平上平

衡,遗憾的是,目前我们还做不到这一点)。见表 5。

表 5 各个阶段留学生的语料字数、主题、国家、人数

	年级	语料字数	语料主题	国家及人数
初级	B班	18601	写信	韩国 20/日本 20/东南亚 20/非汉字圈 20
中级	二年级	48139	写信	韩国 20/日本 20/东南亚 19/非汉字圈 20
高级	三年级	40593	论学习汉语的意义;论生与死	韩国 20/日本 13/东南亚 12/非汉字圈 15

3.2 留学生汉字和词语使用数量统计结果

选取适当的语料样本后,我们对留学生作文中使用的汉字和词语的数量、使用频率进行了统计。汉字的统计比较容易,词语的统计牵涉词切分的问题,所以比较困难。本文采用机器自动切分然后进行人工校对的方法。自动切分后,请从事对外汉语教学和研究的语言学专业的研究生、教师进行人工校对,对机器切分的明显错误进行改正。

我们分别统计了初级、中级、中高级三个阶段学生的汉字数量和词语数量。表 6 是不同阶段留学生汉字和词语数量的统计结果。

表 6 不同水平留学生汉字和词语使用数量

水平	汉字数量	词汇数量
初级	907	1161
中级	1491	2572
中高级	1296	2403
总计	1847	3985

下面将留学生汉字和词汇使用量与《大纲》的要求进行比较。由于我们只统计了从初级到中高级(三年级)学生的语料,所以只与《大纲》丙级的要求比较。通过比较可以看到,在统计的语料中,留学生使用的汉字总量(1847个)还达不到《大纲》丙级要求的汉字量(2205个),留学生使用的词语总量(3985个)也远未达到《大纲》丙级要求的词汇量(5253个)。

统计得到的汉字和词语使用总量与《大纲》要求的数量存在差距,有三个可能的因素:一是学生的实际用字量和用词量不止这么多,但由于所统计语料的主题比较单一,样本量也不够大,所以统计结果并不能完全反映留学生实际的用字量和用词量。二是该方法统计出来的只是学生的产生型汉字量和词汇量,而学生的识别型汉字量、词汇量比产生型的要大。三是学生的汉字量和词汇量确实没有达到《大纲》的要求,这是因为学生学得不够多、输入汉字量和词汇量不够大,还是因为《大纲》要求过高?这一问题值得研究。

3.3 留学生汉字和词语使用频率统计结果

我们统计了初级到中高级学生的汉字和词语使用频率,得到不同阶段留学生汉字使用频率表、词语使用频率表。比较发现,留学生经常使用的一些词语《大纲》并未收入。例如:

双字组:

表姐 地狱 点菜 敦厚 法语 刮风 汉族 好汉 好玩 回归 回国 回家
 吉他 记住 解脱 京剧 烤鸭 厉害 聊天 临死 魅力 面条 母语 难忘
 期末 期中 钱包 清淡 缺课 人才 日语 入世 软卧 思路 寺庙 听力
 外语 吸烟 下次 下雪 阳台 英文 英语 油腻 早上 总是 走路 足迹

三字组:

比如说 火车票 拉肚子 旅行社 上个月 实际上 手提包 售货员 售票员
俗话说 太极拳 下星期 音乐会 怎么办 只不过 主人公

四字组:

大吃一惊 各种各样 哈哈大笑 好不容易 胡思乱想 另一方面 没完没了
名胜古迹 宁死不屈 万事如意 一般来说 一无所有 总而言之

此外还有不少常用的专名《大纲》也没有收入(但教材中一般都有),例如“中国、英国、德国、法国、欧洲、亚洲、非洲、瑞典、东京、汉城、广州、苏州、杭州、黄河、天坛、故宫”等。

上述词语在我们统计的语料中出现的次数都不少,从直觉上看也是常用的,是学生应该掌握的,但是《大纲》没有收入。《大纲》选词时一方面依靠母语者语料的频率统计,另一方面依靠专家的直觉判断。但仅这两个方面还不够,还应该对学习者的词汇使用情况进行分析。本研究对留学生汉字和词语使用频率所做的统计,有助于了解留学生的常用字和常用词,为汉字和词汇大纲的编制提供参考。

当然,由于语料、词切分等诸多限制,本研究提供的留学生字频和词频资料(特别是词频资料)还存在许多不足,因此本研究在方法上的意义可能大于实际资料的意义。

四 《大纲》字词选择和分级

4.1 《大纲》汉字和词语选择及分级的优缺点

《大纲》所收的词应是常用词,这一点大家已达成共识。确定现代汉语的常用词汇或基础词汇,是编制词汇大纲的基础,但如何确定常用词却不是一个简单的问题。

刘英林、宋绍周(1992)指出了已有词汇大纲选择词语遵循的8条原则:常用性原则、均匀性原则、科学性原则、规范性原则、实用性原则、联想性原则、包容性(节省性)原则、序列性(等级性)原则。从实际操作的角度看,以上原则可以概括为以频率和使用度为主、以专家主观判断为辅。这种客观与主观相结合的方法,是合理的,也是可行的。但《大纲》采用时存在以下问题。

(1)语料的问题。频率统计所依据的语料存在的主要问题是:书面语体语料较多、口语语体语料较少(孙德金 1995),语料过时,语料的数量不够大(仅一百多万)。赵金铭等(2003)也认为用于词频统计的语料在数量、范围和语体等方面存在不足。因此,根据这些语料统计所得的词频选择常用词,必定导致有些常用词没有收入,有些非常用词却收了进去。例如《大纲》有“根据地”,没有“骗子”(孙德金 1995),有“法西斯”而没有“小偷”,有“帝国主义”而没有“电脑”。因此,新的词汇大纲依据的词频统计应该基于及时更新的语料。

(2)专家干预方法的问题。主观判断可以弥补频率统计方法的不足,但仅局限于几十位专家联想干预是不够的,应该借鉴国外基础词汇研究的方法,采用调查法(孙德金 1995),进行大样本的词汇联想测验或主题联想测验。

(3)对留学生日常生活中常用词的研究不足。有的词语在词频统计中可能属于低频词,但在留学生的日常生活中却是常用词,例如“米饭”。哪些词属于这样的词,这个问题靠专家联想干预的方法可以在一定程度上解决,但不能完全解决。

专家联想干预是一种主观方法,应结合客观方法进行。研究留学生常用词的客观方法有二:一是以广泛使用的对外汉语教材的课文语料和词表为基础统计词频和常用词。二是以留

学生的语料为基础统计词频,根据留学生词语使用频率来确定常用词。赵金铭(1989)曾使用过第一种方法。确定常用字时,他除了依据字的频率(基于母语者语料的统计)、字的构词能力,还依据对四部基础汉语教材用字的统计。后一种方法,已发表的论文中至今未见采用。本研究采用该方法对留学生常用词语进行了研究,在统计留学生词语使用频率方面提供了一个样例,以期词汇大纲修订提供方法上和资料上的参考。

(4)处理字词关系时采用以词定字的方法存在不足。只根据词的出现等级来决定汉字的等级,可能对汉字学习不利,最终导致对词汇学习不利。在处理字词关系时,是否可以考虑针对汉语词多字少的特点、采取以字定词的方法?这个问题下面还要详细讨论。

4.2 新大纲字词选择和分级应遵循的原则

综上,我们为等级大纲的字词选择和排序、字词关系的处理等问题的解决提出两种方法。

4.2.1 以字定词,字词协调

第一步:选择汉字。根据汉字的出现频率(依据母语者语料的统计)和留学生使用频率(依据留学生语料的统计)来选择汉字。例如,先确定各个等级应该学习的汉字量(例如前1000、1500、2000、2500个高频字)。

第二步:以字找词。在不同频率等级的词中找出由这些汉字组成的所有的词。例如,确定各个等级应该掌握的词汇量(例如前1000、3000、5000、10000个高频词),并分别在不同词频等级的词中找出由这些字组成的词。表7是以北京语言大学信息科学系网站提供的以小说为文本统计的字频和词频资料为例,对各个频率等级汉字在各个频率等级词汇中的构词数进行统计的结果。这里的构词数指某个等级的汉字在某个等级的词汇中构词的个数,非单音节词的汉字必须全部属于指定等级才计算在内。

表7 各个等级汉字在不同等级词汇中的构词数

		1级	2级	3级	4级
		前 1000个高频词	前 3000个高频词	前 5000个高频词	前 10000个高频词
	前 500个高频字	769	1449	2048	3366
1级	前 1000个高频字	979	2330	3445	4977
2级	前 1500个高频字	1000	2436	3551	7438
3级	前 2000个高频字	1000	2732	3847	8460
4级	前 2500个高频字	1000	3000	4435	9199

对于前10000个高频词来说,如果能认识500个汉字,就可以认识其中3366个词的汉字(33.66%);如果认识了1000个汉字,就能认识其中4977个词的汉字(49.77%)。依此类推。

第三步:适当调整。根据这些词的出现频率、留学生使用频率适当调整各等级的词汇量。

由此看来,“以字定词”的方法既是可行的,也是有效的,它既可以有效地减轻学生汉字学习的负担,又能迅速扩大学生的词汇量。例如赵金铭(1989)所说,找出外国人基础汉语用字,让学生先学习这些字,比无计划地任意学要好得多。学了基础字,不但可以认识那些“熟字熟义”的所谓语义透明词,也有助于认识那些“熟字生义”的所谓不透明词,而词义的学习,反过来又可加深字义的理解,“字义和词义是辗转相生的”(赵金铭1989)。

4.2.2 各自为政,互相协调

具体做法是,汉字和词汇按各自标准进行选择 and 排序,最后分级时再进行协调,达到统一。

选择汉字时可以依据的标准是:出现频率高、构词数高、留学生使用频率高。选择词汇时

可以依据的标准是:整词出现频率高、单字出现频率高、留学生整词使用频率高、留学生单字使用频率高。这一标准考虑了外国人汉字和词汇学习的主要影响因素,汉字选择的标准也与赵金铭(1989)制定“外国人基础汉语用字表”时采用的三个原则(频率、构词能力和常识)一致。这种方法比较传统,但与更传统的“以词定字”相比有所改进,既重视词频,又兼顾字频。

字词选择和分级无论采用哪个方法,都有一个重要的前提,就是要从对外汉语教学的实际需要出发,分别以汉语母语者的语料(有代表性和时代性)、汉语教材语料、留学生的语料等三类语料为基础进行字频、词频和汉字构词数的统计。

参考文献

- 陈慧、王魁京 2001 外国学生识别形声字的实验研究,《世界汉语教学》第2期。
- 江新 2001 外国留学生形声字表音线索的意识及其发展的实验研究,《世界汉语教学》第2期。
- 江新 2003a 不同母语背景的外国学生汉字知音和知义之间关系的研究,《语言教学与研究》第6期。
- 江新 2003b 初学汉语的美国学生汉字正字法意识的实验研究,载赵金铭主编《对外汉语研究的跨学科探索》,北京语言大学出版社。
- 江新 2004 汉字复现率对拼音文字背景的外国学生汉字学习的影响,载《北京地区第三届对外汉语教学学术研讨会论文集》,北京大学出版社。
- 江新、柳燕梅 2004 拼音文字背景的外国学生汉字书写错误的研究,《世界汉语教学》第1期。
- 江新、赵果 2001 初级阶段留学生汉字学习策略的调查研究,《语言教学与研究》第3期。
- 李清华 1999 《汉语水平考试和汉字等级大纲》的词汇量问题,《语言教学与研究》第1期。
- 刘英林、宋绍周 1992 论汉语教学字词的统计与分级(代序),载《汉语水平词汇与汉字等级大纲》,北京语言学院出版社。
- 柳燕梅 2002 生词重现率对欧美学生词汇学习影响的实验研究,《语言教学与研究》第5期。
- 柳燕梅、江新 2003 欧美学生汉字学习方法的实验研究——重复抄写与回忆默写法的比较,《世界汉语教学》第1期。
- 鹿士义 2001 词汇习得与第二语言能力研究,《世界汉语教学》第3期。
- 上海交通大学汉字编码组 1988 《汉字信息字典》,科学出版社。
- 孙德金 1995 《HSK词汇等级大纲》问题浅见,载《第四届国际汉语教学讨论会论文集》,北京语言学院出版社。
- 张凯 1997 汉语构词基本字的统计分析,《语言教学与研究》第1期。
- 赵果 2003 初级阶段欧美留学生识字量与字的构词数,《语言文字应用》第3期。
- 赵果、江新 2002 什么样的汉字学习策略最有效?——对基础阶段留学生的调查研究,《语言文字应用》第2期。
- 赵金铭 1989 《外国人基础汉语用字表》草创,载南开大学对外汉语教学中心主编《汉语研究》第2集,南开大学出版社。
- 赵金铭、张博、程娟 2003 关于修订《汉语水平》词汇等级大纲的若干意见,《世界汉语教学》第3期。
- 中国社会科学院语用所 1988 《汉字问题学术讨论会论文集》,语文出版社。
- 周清海、梁荣基 1995 字词频率与语文学习成效的相关研究,载《第四届国际汉语教学讨论会论文集》,北京大学出版社。
- Bourgerie, D. S 1996 Acquisition of Model Particles in Chinese Second Language Learners In S. McGinnis (ed.), *Chinese Pedagogy: An Emerging Field*, 107 - 134, Columbus, OH: Ohio State University Foreign Language Publications
- Everson, M. E 1998 Word Recognition among Learners of Chinese as a Foreign Language: Investigating the Rela-

- tionship between Naming and Knowing *Modern Language Journal*, 82: 194 - 204.
- Ke, C. 1996 An Empirical Study on the Relationship between Chinese Character Recognition and Production *The Modern Language Journal*, 80: 340 - 349.
- Ke, C. 1998 Effects of Strategies on the Learning of Chinese Characters among Foreign Language Studies *Journal of the Chinese Language Teachers Association*, 33 (2), 93 - 112.
- Kuo, J. Y. 2000 Strategies for Learning Classifiers In B. Swierzbinska et al (eds), *Social and Cognitive Factors in Second Language Acquisition*, 424 - 442. Somerville, MA: Cascadia Press
- Lin, Y. 2000 Vocabulary Acquisition and Learning Chinese as a Foreign Language *Journal of the Chinese Language Teachers Association*, 35 (1): 85 - 108.
- McGinnis, S. 1999 Students Goal and Approaches In M. Chu (ed), *Mapping the Course of the Chinese Field*, 151 - 175, Kalamazoo, MI: Chinese Language Teachers Association
- Polio, C. 1994 Non-native Speakers' Use of Nominal Classifiers in Mandarin Chinese *Journal of Chinese Language Teacher Association*, 29: 51 - 66.
- Sergent, W. K., & Everson, M. E. 1992 The Effects of Frequency and Density on Character Recognition Speed and Accuracy by Elementary and Advanced L2 Readers of Chinese *Journal of the Chinese Language Teachers Association*, 27 (1/2): 29 - 44.
- Shen, H. H. 2000 Radical Knowledge and Character Learning among Learners of Chinese as a Foreign Language *Proceedings of the International Conference on Chinese Pedagogy*, 85 - 93.
- Teng, S. H. 1999 The Acquisition of “了 le” in L2 Chinese, 《世界汉语教学》第 1 期。
- Wen, X. 1995 Second Language Acquisition of the Chinese Particle le *International Journal of Applied Linguistics*, 5 (1): 45 - 62.
- Wen, X. 1997 Acquisition of Chinese Aspect: An Analysis of the Interlanguage of the Learners of Chinese as a Foreign Language *Review of Applied Linguistics*, 1 - 26.
- Yang, J. 2000 Orthographic Effect on Word Recognition by Learners of Chinese as a Foreign Language *Journal of Chinese Language Teacher Association*, 35 (2): 1 - 17.

The Effects of Frequency, Productivity and Complexity on Learning Chinese Characters and Words by Foreign Students

Abstract This paper includes four parts. Firstly, the previous studies on learning Chinese characters and words are shortly reviewed. Secondly, the present research on the effects of frequency, productivity and complexity on learning Chinese characters and words by foreign students is reported. Thirdly, the analysis of the characters and words based on the foreign students' composition corpus is conducted. Lastly, the principles of selecting characters and words for designing *Hanyu Shuiping Dagang* (《汉语水平大纲》) are discussed.

Key words learning Chinese characters and words; frequency of character; frequency of word; productivity

(江 新 100083 北京语言大学对外汉语研究中心)
 (赵 果 王又民 100083 北京语言大学汉语进修学院)
 (黄慧英 柳燕梅 100083 北京语言大学汉语速成学院)
 (责任编辑 张世方)