



国家汉办/孔子学院总部
Hanban/Confucius Institute Headquarters

汉语国际教育用 音节汉字词汇等级划分 (国家标准·应用解读本)

The Graded Chinese Syllables,
Characters and Words for the Application of
Teaching Chinese to the Speakers of Other Languages

(National Standard: Application and Interpretation)

国家汉办 教育部社科司

《汉语国际教育用音节汉字词汇等级划分》课题组



北京语言大学出版社
BEIJING LANGUAGE AND CULTURE
UNIVERSITY PRESS



国家汉办/孔子学院总部
Hanban/Confucius Institute Headquarters

汉语国际教育用 音节汉字词汇等级划分 (国家标准·应用解读本)

The Graded Chinese Syllables,
Characters and Words for the Application of
Teaching Chinese to the Speakers of Other Languages
(National Standard: Application and Interpretation)

国家汉办 教育部社科司

《汉语国际教育用音节汉字词汇等级划分》课题组



北京语言大学出版社
BEIJING LANGUAGE AND CULTURE
UNIVERSITY PRESS

图书在版编目(CIP)数据

汉语国际教育用音节汉字词汇等级划分：国家标准
·应用解读本 / 中国国家对外汉语教学领导小组办公室，
教育部社科司《汉语国际教育用音节汉字词汇等级划分》
课题组编. — 北京：北京语言大学出版社，2010. 10
ISBN 978-7-5619-2902-5

I. ①汉… II. ①中… ②教… III. ①汉语—音节—
对外汉语教学—教学参考资料②汉语—词汇—对外汉语教
学—教学参考资料 IV. ①H195.4

中国版本图书馆 CIP 数据核字 (2010) 第 212048 号

书 名：汉语国际教育用音节汉字词汇等级划分（国家标准·应用解读本）
责任印制：汪学发

出版发行：北京语言大学出版社
社 址：北京市海淀区学院路 15 号 邮政编码：100083
网 址：www.blcup.com
电 话：发行部 82303650/3591/3651
编辑部 82303647
读者服务部 82303653/3908
网上订购电话 82303668
客户服务信箱 service@blcup.net
印 刷：北京画中画印刷有限公司
经 销：全国新华书店

版 次：2010 年 10 月第 1 版 2010 年 10 月第 1 次印刷
开 本：880 毫米×1230 毫米 1/16 印张：24
字 数：769 千字
书 号：ISBN 978-7-5619-2902-5/H·10285
定 价：75.00 元

凡有印装质量问题，本社负责调换，电话：82303590

研制《汉语国际教育用音节汉字词汇等级划分》

探寻汉语国际教育新思维*（代序）

刘英林 北京语言大学

马箭飞 国家汉办

○ 引言

这次研制《汉语国际教育用音节汉字词汇等级划分（国家标准）》（以下简称《等级划分》），不是对 1992 年《汉语水平词汇与汉字等级大纲》（以下简称《等级大纲》）简单的、一般性的修订^①，而是在相隔 17 年之后，在社会语言生活发生了巨大变化、对外汉语教学出现了重大转变的时代背景下，在“面向全球汉语教育”“正处在抉择的紧要关头”（白乐桑、张丽，2008）进行的“一项意义十分重大的工程”（赵金铭、张博、程娟，2003），是一项具有代表性、标志性的重大研究课题。

完成这一重要课题，必须贯通这个知识领域的过去、现在和未来，必须摆脱既有的思维定势和框架，改变传统、单一的对外汉语教学视角和尺度，由海内外多学科专家“跨界”参与研讨，探求汉语国际教育的新思维。在准备和研制的整个过程中，要作认真的探索和不断的积累，尽量少一点对现有知识的重复和复述，尽量多一点符合实际特别是符合汉语自身特点的独立思考和知识创新，追求较高的知识附加值。完成这个重大课题，最紧要的是根据时代发展的需求，找准几个最重要的关节点，从汉语国际教育的战略目标出发，立足于破解所谓“汉语难学”的瓶颈问题，立足于汉语国际教育的通俗化、大众化、普及化，立足于构建新标准、新模式、新体系，立足于自主创新、现代科技手段的运用并与未来长远发展接轨。

汉语国际教育要“面向现代化，面向世界，面向未来”^②，要创立世界知名品牌。

一 主要依据

《等级划分》研制的主要依据包括 30 多亿字次的当代大型动态语料库和多种具有代表性、针对性的词典、词表、字表。

*《汉语国际教育用音节汉字词汇等级划分（国家标准）》是国家汉办/孔子学院总部、教育部社科司的重要科研项目，得到了国家汉办许琳主任的大力支持，得到了教育部杨光司长、李宇明司长和北京语言大学王路江书记、崔希亮校长等领导的大力支持，也得到了侯敏、杨尔弘、靳光瑾、别良峰等先生的大力协助。课题组主要成员有：陈瑞端、傅永和、林茂灿、佟乐泉、张家驷、张振兴、侯晓辉、李方延。另外，曹文、施家炜、张劲松、张维佳、袁礼、张彤辉、周高宇、慕亚东等参加了第二阶段的研究工作。在此一并向他们致谢！

① 2001 年对《等级大纲》进行了简单的修订，主编刘英林（那时正在香港）既没有参与也不知情，谨此说明。

② 1983 年 9 月，邓小平同志给北京景山学校题词：“教育要面向现代化，面向世界，面向未来。”我们认为，“三个面向”也应成为汉语国际教育这一学科的指导方针。

1.1 当代大型动态语料

(1) 大型对话节目动态语料。这是最重要、最核心的选材依据,是课题组特别请求相关科研机构制作的。《广播电视对话节目语料》基本数据(2005~2008)包括中央电视台、北京电视台等广播电视对话性质节目转写文本 9153 个文件,包含汉字 49,899,862 字次,5735 字种;词语 30,689,302 条次,词种 201,161 个。

(2)《有声媒体语料》基本数据(国家语言资源监测与研究有声媒体语言分中心,2005~2008)。包括中央电视台等 8 家电视台、中央人民广播电台等 7 家广播电台的转写语料,4 年共计 2.6 亿余字次。^③

(3)《报纸、广播电视、网络新闻综合语料》基本数据(国家语言资源监测与研究,2005~2008)。语料源自报纸、广播电视、网络新闻三种媒体,4 年共计 37.1 亿余字次。①广播电视。参见上文《有声媒体语料》基本数据。②报纸基本数据(国家语言资源监测与研究平面媒体语言分中心,2005~2008)。选择了《北京青年报》等 15 种报纸的语料,4 年共计 19.1 亿余字次。③网络新闻基本数据(国家语言资源监测与研究网络媒体语言分中心,2005~2008)。选择了新华网、人民网、新浪网、腾讯网等网站的新闻文本作为语料,4 年共计 15.39 亿余字次。

(4)《九年义务教育(新课标)语文科语料》基本数据(郑州航空工业管理学院文字信息化研究所,2009)。统计现行“九年义务教育(新课标)”中小学 1~9 年级语文教材共 90 本(包括人教版、北京版、苏教版、北师大版、语文版等 5 个版本),统计语料共计 2,684,345 字次。

(5)《国家语委现代汉语语料库(平衡语料库)》基本数据(教育部语言文字应用研究所,2009)。包括 1919~1992 年时间段的大部分语料以及 1992~2002 年时间段的部分语料,共计 1 亿字次。其中以近 20 年的语料为主。已经完成词语切分和标注加工的约 5000 万字语料。

1.2 现代汉语词典、等级大纲、常用字表和词表

(1)《现代汉语词典》(第 5 版,中国社会科学院语言研究所词典编辑室,2005)音节表及注音、词类标注;

(2)《汉语语言文字启蒙》1、2(白乐桑、张朋朋,1997);

(3)《汉语水平词汇与汉字等级大纲》(国家对外汉语教学领导小组办公室汉语水平考试部、北京语言大学汉语水平考试中心,刘英林主编,1992);

(4)《汉语水平等级标准与语法等级大纲》(国家对外汉语教学领导小组办公室汉语水平考试部、北京语言大学汉语水平考试中心,刘英林主编,1996);

(5)《高等学校外国留学生汉语言专业教学大纲》(国家对外汉语教学领导小组办公室,2002);

(6)《香港地区普通话教学与测试词表》(陈瑞端主编,首席顾问刘英林,2008);

(7)《汉语教材常用词表及常用字表统计分析报告》(国家汉办/孔子学院总部教材处,2008);

(8)《现代汉语常用字表》(国家语言文字工作委员会,1988);

③ 本课题是现代汉语规范口语主导型的新型《等级划分》,为了突出口语特点与权重,特别把包含在《报纸、广播电视、网络新闻综合语料》之内的《有声媒体语料》单列出来,作为一项重要的语料数据。

(9)《普通话水平测试实施纲要》(国家语言文字工作委员会普通话培训测试中心, 2008);

(10)《现代汉语常用词表(草案)》(《现代汉语常用词表》课题组, 2008)。

二 研制概要

2.1 研制准备阶段

2007年10月, 课题组在京自发组建, 自行开始研发工作。大家根据汉语国际教育的新形势、新情况, 决定先组成“民间”专家组, 确定研究课题, 进行总体设计; 充分论证、探索汉语音节吸纳、引入“国家标准”的可行性及途径和方法; 对一级汉字表先期进行预选试验, 待条件成熟、确有把握时, 再向有关部门、有关单位正式提出申请立项。本阶段的工作主要有:

(1) 初步完成《汉语国际教育用基本音节表》的筛选工作。通过《基本音节表》的初步筛选与粗略分级, 我们意识到, 把音节的概念引进汉语国际教育, 很可能成为自主研发“国家标准”的突破口和创新点, 成为创立音节、汉字、词汇“三合一”的新型《等级划分》的关键、核心要素之一, 成为逐步破解所谓“汉语难学”这个问题的最佳途径之一。大家觉得这项研究工作很有价值, 也很有创意。

(2) 一级汉字表的预选试验。我们选取4个字表作为预选试验的材料: ①国家语委绿皮书1200字(2006年); ②清华大学王良辰1000字(网上下载, 2007年); ③白乐桑、张朋朋900字(1997年); ④国家汉办、北京语言大学汉语水平考试中心《等级大纲》甲级字800个(1992年)。用计算机将这4个不同数量的字表进行交集、比对, 得出如下基本数据: 4个字表收录的字557个, 3个字表收录的字255个, 这812(557+255)个字都进入了预选字表。另外, 两个字表收录的字有301个, 从中选取88个字进入预选字表, 选取依据主要有两个: 一是日常生活中使用频度高, 二是与以上812个字音节不同, 其目的是增加汉字表最常用音节的数量和比例。最终得出一级900字的预选字表和相对应的一级预选音节表(591个)。

通过一级汉字表的预选试验, 我们进一步增强了信心, 积累了实践经验, 找到了研制分级汉字表和音节表可靠、有效的途径。我们认为: ①汉字表和音节表的筛选必须以当代大型动态语料库和多种具有代表性的字表、词表为依据。②一级汉字表的基本量, 从“汉字覆盖率”、“音节的统计与覆盖率”和“与未来发展接轨”三个角度综合考虑, 定为900字为好^④。这个预选字表也成为后来正式研制一级900汉字表的重要依据之一。

实践证明, 组建民间课题组进行先期准备与论证是一条切实可行的合作发展之路。

2.2 一级(初级, 也称为普及化等级, 下同)汉字表、音节表、词汇表的筛选

(1) 一级汉字表的筛选。①筛选的主要依据。一级汉字表的筛选是建立在9种汉字表的基础上

^④ 大型对话体动态语料900字的覆盖率为92.10%; 一级900字所含608个音节对《等级划分》1206个基本音节的覆盖率为50.41%, 也就是说, 在普及化等级(初级), 学习608个音节就掌握了《基本音节表》50%以上的音节; “与未来发展接轨”重点考虑与即将研制的自动化汉语口语水平考试(HKK)普及化等级A、B、C三等衔接, C等的汉字为300个, 是900字的1/3。

的，它们涵盖了各种不同的类别，既有大型动态语料，又有针对性很强的字表。前6种字表是：广播电视对话节目字频统计前900字，对话节目语料词频统计前1538个词中包含的900字，有声媒体字频统计前900字，中小学新课标语文科字频表前900字，多媒体（报纸、广播电视、网络新闻）字频统计前900字，国家语委平衡语料库字频统计前900字。后3种字表是：本课题组预选试验字表900字，白乐桑、张朋朋900字，《等级大纲》中甲级800字。②筛选的基本方法。首先，用计算机对前8个900字表进行比对，共得到1305个字，它们成为一级汉字表选字的基本范围。其次，交集、比对8个字表、7个字表和6个字表共有的字，绝大多数直接进入900字表；5个字表和4个字表共有的字，大多数选入900字表；而3个字表、2个字表和1个字表有的字，少量经过专家干预，选入900字表。专家干预的原则有三条：A. 选入最基本、最常用的交际用字；B. 选入最基础、适于交际的衣、食、住、行用字；C. 以下面两类字表作为筛选的重要参照。第一类是上述9个字表中针对性很强的3个字表，它们是：预选900字表，白乐桑、张朋朋900字表，《等级大纲》甲级800字表；第二类是新近出现的字表和“统计分析报告”，它们是：许琳、姚喜双800字表，国家汉办《高等学校外国留学生汉语言专业教学大纲》中的一级字795个，国家汉办/孔子学院总部教材处《汉语教材常用词表及常用字表统计分析报告》中的常用字表等。上述两类6种字表或材料作为“查漏补缺”的依据，用来弥补大型频度统计的不足。实践证明，这样“有据可依”的“查漏补缺”是非常必要的，也是可行的。③汉字表与音节表的筛选相结合。

（2）一级音节表的筛选。一级音节表的筛选原则和方法有三个：①常用汉字和常用音节的统计频率、使用频率基本上是一致的。我们以高频度的音节覆盖作为取舍的重要标准。②一级音节表和汉字表的筛选基本上是结合在一起进行的。也就是说，在进行汉字分级时，兼顾音节的频率统计。一级音节表要尽量覆盖更多的常用音节。③对于多音字，要根据相对应汉字的常用度和汉字构成词的常用度进行衡量与分级。

（3）一级词汇表的筛选。①词汇筛选依据的大型频率统计词表有5个：广播电视对话节目词频表，有声媒体词频统计表，多媒体（报纸、广播电视、网络新闻）词频统计表，中小学新课标语文科词频表，国家语委平衡语料库词频表。所有的词表均去除人名、地名和英文字母词（下同）。②一级词汇筛选的两个档次。将一级常用词根据常用性的差异分为两个档次：第一档次为最常用词，第二档次为常用词。因此，词汇筛选也相应分为两步走。第一步：用一级900字与5个词表的前1600词进行交集（据我们统计，广播电视对话节目词频统计前1600词里出现了900字种，因此我们以5个词表的前1600词作为第一个共同的档次进行比对），保留那些完全由这900字组成的词。随后，将这5个词表进行比对，分别得出5表、4表、3表共有的词，这些词大多进入第一档次的最常用词，部分进入第二档次的常用词。第二步：用一级900字与5个词表中频度位置在1601~3000的词作为第二个档次进行交集，同样保留那些完全由这900字组成的词。随后，将这5个词表进行交集，分别得出5表、4表共有的词，这些词少量进入第一档次的最常用词，部分进入第二档次的常用词。另外，将第一步中2表、1表有的词，与第二步中3表、2表共有的词再进行交集，根据比对结果，少数进入第一档次的最常用词，部分进入第二档次的常用词。初步的筛选结果是，第一档次的最常

IV

用词 1342 个, 第二档次的常用词 903 个。

2.3 二级(中级, 下同)汉字表、音节表、词汇表的筛选

(1) 二级汉字表的筛选。①二级汉字表的筛选是在 8 种字表的基础上进行的。其中 5 个表是汉字频度统计, 包括: 广播电视对话节目字频统计前 1800 字, 有声媒体字频统计前 1800 字, 中小学新课标语文学科字频表前 1800 字, 多媒体(报纸、广播电视、网络新闻)字频统计前 1800 字, 国家语委平衡语料库字频统计前 1800 字。为了减少人名、地名用字对比对结果的干扰, 其他 3 个字表使用的是去除人名、地名的词汇频度统计结果中包含的字种。举例来说, 广播电视对话节目词频统计结果中, 按频度排列的前 5238 个词是由 1800 个字种组成的, 这 1800 字就作为二级汉字表的筛选依据。以此方法生成的后 3 种字表包括: 广播电视对话节目词频统计前 1800 字种, 中小学新课标语文学科词频统计前 1800 字种, 国家语委平衡语料库词频统计前 1800 字种。②筛选的基本方法。与一级汉字表的方法基本相同, 先用计算机对这 8 个 1800 字表进行比对, 共得到 2538 个汉字, 从这 2538 个字中除去一级 900 字, 剩下的 1638 字就是二级字的选字范围。汉字的取舍主要看对比结果, 如 8 个表、7 个表都有的字, 原则上要进入二级 900 字。如前文所示, 一级 900 字是从 1305 个字中选取的, 因此剩下的 405 个字会优先考虑进入二级 900 字。此外, 《等级大纲》中的甲级字, 许琳、姚喜双 800 字, 白乐桑、张朋朋 900 字中未进入一级 900 字的, 原则上也优先考虑进入二级 900 字。二级字的选取依然考虑了音节频度、构词能力等相关因素。

(2) 二级音节表的筛选, 情况与一级基本相同。

(3) 二级词汇表的筛选。①词汇筛选依据的大型频率统计词表有 5 个, 与一级相同。②筛选的基本方法。一级 900 字和二级 900 字共 1800 字, 分别与 5 个词表的前 5500 词进行交集(广播电视对话节目词频统计前 5238 词包含了 1800 字种), 保留那些完全由这 1800 字组成的词。将这 5 个词表进行比对, 共得出 8911 个词。从这 8911 个词中去掉一级 2245 个词, 剩下的 6666 个词就是二级词的主要选词范围。在筛选过程中, 专家进行了适当的人工干预。由于这部分词语大部分是日常生活、日常交际中常用的词语, 专家干预的比例明显少于对一级词的干预。

2.4 三级(高级, 下同)汉字表、音节表、词汇表的筛选

(1) 三级汉字表的筛选。①筛选的主要依据与二级相同, 只是字频字数、词频字种数都定为 2700 字。②筛选的基本方法。经过 8 个 2700 字表比对, 共得到 3525 个汉字, 从这 3525 字中除去一级、二级汉字共 1800 字, 剩下的 1725 字就是三级 900 字的选字范围。

(2) 三级音节表的筛选, 情况与前两级基本相同。

(3) 三级词汇表的筛选。①词汇筛选依据的大型频率统计词表有 5 个, 与前两级相同。②筛选的基本方法。一、二、三级字相加共 2700 字, 分别与 5 个词表的前 12000 词进行交集(广播电视对话节目词频统计前 12112 词包含了 2700 字种), 保留那些完全由这 2700 字组成的词。随后, 将这 5 个词表进行比对, 共得出 20007 个词。从比对结果中去掉一级、二级共 5456 个词, 剩下的 14551 个词就是三级词的主要选词范围。此外, 将一级、二级词筛选后剩下的词与三级词进行比对, 其中大部分词优先考虑进入三级词汇范围。

与一、二级词相比，三级词略倾向于普通话书面语。

2.5 高级“附录”汉字表、音节表、词汇表的筛选

高级“附录”是为以汉语为专业的外国学生及其他汉语水平较高的学习者准备的。

(1) “附录”汉字表的筛选。附录的选字范围与三级字的选字范围相同。也就是说，从选完三级 900 字后剩下的 816 个字中选取附录字。其中姓氏、地名常用字 30 个（均为前三级字中未选收的字），其他次常用字 270 个，共 300 个。

(2) “附录”音节表的筛选。在选取附录汉字表时，既要考虑汉字的常用度，又要注意音节的频率，尽量扩大音节的覆盖数量。

(3) “附录”词汇表的筛选。①筛选的主要依据仍然是 5 个大型频度统计词表，与前三级相同。②筛选的基本方法。一、二、三级各 900 字，加上附录 300 字，共 3000 字，分别与 5 个词表的前 15000 词进行交集，保留那些完全由这 3000 字组成的词。将这 5 个词表进行比对，从比对结果中去掉一、二、三级词，剩下的词就是附录词的主要选词范围。“附录”词汇表收入了较多的习用语和四字格词语。主要依据有 3 个：一是当代大型动态语料频度统计数据（2005~2008）；二是《汉语水平等级标准与语法等级大纲》（1996）收录的习用语和四字格词语；三是《汉语教材常用词表及常用字表统计分析报告》（2008）的相关资料。

三 性质、用途和基本框架

3.1 性质

汉语国际教育“面向母语非汉语者的汉语教育、教学，包括世界各地的国际汉语教学和中国国内的对外汉语教学”。

《等级划分》是面向全球汉语国际教育的国家标准，是汉语国际教育新兴学科的基本建设，是一种标准化、系统化、规范化、精密化的等级水平划分。

《等级划分》是教育部社科司的重大研究课题，是国家汉办/孔子学院总部的重要科研项目，它具有多种用途和广泛的适应性。

3.2 主要用途

(1) 世界各地的国际汉语教学和中国国内的对外汉语教学进行总体设计、教材编写、课堂教学和课程测试的重要依据。这是最主要、最基本的用途。

(2) 《汉语国际教育用中国国家级计算机辅助汉语口语水平考试》（简称为 HKK）最重要的命题依据。HKK 分为普及化水平（初级水平）、中级水平和高级水平三级九等。

(3) 新型中国国家级汉语水平考试（修订版）初级水平、中级水平和高级水平的重要命题依据与参照。

(4) 中国少数民族汉语教学以及国内其他普通话教学、中小学语文教学及其他相关标准化、规范化语言考试的重要参照。

(5) 编制汉语国际教育常用字典、词典及计算机音节库、字库和词库的重要参照。

3.3 基本框架

(1) 汉语三要素：音节、汉字、词汇。

(2) 三个等级与三级水平：一级（普及化水平）、二级（中级水平）、三级（高级水平）。

(3) 新标准、新模式、新体系：音节、汉字、词汇三合一新型《等级划分》，统一、协调、完整的一级（普及化等级，也称为初级）、二级（中级）、三级（高级）三个级别的等级划分和普及化水平、中级水平、高级水平三个级别的等级水平及评价体系。详见表 1：

表 1 《等级划分》基本框架^⑤

等级划分与等级水平			音节	汉字	词汇	汉字音节比	词汇汉字比
一级（普及化等级）	普及化	入门等级	272	300	505	1.10	1.68
	水平		608*	900*	2245*	1.48	2.49
二级（中级）	中级水平		300/908	900/1800	3211/5456	1.98	3.03
三级（高级）	高级水平		163/1071	900/2700	4175/9631	2.52	3.57
	高级“附录” ^⑥		39/1110	300/3000	1461/11092	2.70	3.70
总计			1206**	3000	11092		

注：* 此 608 个音节、900 个汉字、2245 个词语含入门等级 272 个音节、300 个汉字、505 个词语。

** “总计”中的音节数 1206 是基本音节表的总数，但是 3000 字覆盖的音节为 1110 个。

四 探寻汉语国际教育新思维

汉语国际教育是一门新兴的学科。自主开发与这个新兴学科相适应的《等级划分》（国家标准），是当务之急，也是必由之路。研发《等级划分》是一项非常复杂的语言系统工程，需要根据时代和学科发展的要求进行认真的反思，需要从汉语自身的特点入手，需要从更多的角度将历史、分析及建议结合在一起，需要将汉语国际教育的战略目标、现代化科技手段、实事求是的精神和更多的中国元素结合在一起。

遵循以上几条原则，我们从 8 个不同的角度和层面关注、把握对外汉语教学出现的“六大转变”（许琳，2006）^⑦，对汉语国际教育新思维进行概括性论述，形成汉语国际教育新思维的初稿。

⑤ 2010 年 2 月至 5 月，课题组使用《等级划分》（征求意见稿）征询了海内外 100 位专家学者的意见后，对《等级划分》中部分音节、汉字、词汇作了必要的调整与修正。因此，这里的数据与 2010 年第 1 期《世界汉语教学》发表的文章中的某些基本数据不尽相同，特此说明。

⑥ 高级“附录”是为以汉语为专业的外国学生及其他汉语水平较高的学习者准备的，是为提高服务的。高级“附录”中包括部分姓氏、地名常用字（这些都是三级字表中未收入的字）、其他次常用字，以及由这些次常用字组成的次常用词、习用语和四字格词语。

⑦ 参见《中国教育报》7 月 6 日第 1 版。“六大转变”最重要的内容有三项：一是发展战略的转变，从对外汉语教学向全方位的汉语国际推广转变。二是工作重心的转变，从“请进来”学习汉语向在“请进来”的基础上，大力发展“走出去”教汉语转变。三是推广理念的转变，从专业汉语教学向在此基础上的大众化、普及化、应用型教学转变。

4.1 倡导通俗化、大众化、普及化

推行“普及化等级水平”（汉字 900 个、“覆盖”音节 608 个、“生成”词语 2245 个，详见表 2）、关注规范口语教学，是这次研发《等级划分》最大的观念跨越，是汉语国际教育实行改革与转型向前迈出的一大步。

表 2 普及化等级水平

		音节	汉字	词汇	汉字音节比	词汇汉字比
普及化等级	入门等级	272	300	505	1.10	1.68
(普及化水平)		608	900	2245*	1.48	2.49

注：* 这 2245 个词分为三个小层次：一①505 个，一②837 个，一③903 个。这是普及化等级常用词筛选的特有现象，意在为使用者提供循序渐进、多层次的选择。

（1）建立“普及化等级水平”。汉语国际教育重在普及，世界各地国际汉语教学的工作重点是普及。《等级划分》特别是普及化等级水平将以全方位的汉语国际教育和中国文化“走出去”为新思路、新目标和新定位，其最主要的目的是实现汉语国际教育的国家战略目标与外国人学习汉语的现实需要广泛、大面积的“对接”，是逐步实现汉语国际教育的通俗化、大众化、普及化。研制、推行普及化等级水平，一定要有全球化视野；一定要从纯“学术系统”、“专业系统”、“语法讲解系统”中解放出来，要从老师“讲课”、学生“听课”的误区中摆脱出来；一定要坚持科学性与实用性的结合与统一；一定要认真推行国家汉办/孔子学院总部提出的“三贴近”原则：贴近外国人的习惯，贴近外国人的思维，贴近外国人的生活。

（2）关注规范口语教学。汉语国际教育重在交际。我们把普通话口语放在第一位，是因为学习语言最主要的目的就是交际，而最直接的交际就是对话，就是口语表达。我们这次研发工作，在继承传统的基础上，特别关注口语，尤其是口语中的词语节律与应用，注重标准口语教学的针对性、交际性和时效性。特别是在“普及化等级”教学中更是如此。我们把口语中使用频度最高、“流通度最高”（刘长征、张普，2008）、最常用、最规范、最适合交际的口语字词筛选出来，编入普及化等级和中级水平。这样的设计与考量体现了普及化等级和中级两级标准以口语为主导的国际汉语教育大众化、普及化的新理念，有利于汉语国际教育的普及与推广。

4.2 设立“最低入门等级”

设立汉语三要素“最低入门等级”是这次研制工作的一个亮点。

（1）所谓“最低入门等级”，是指汉语国际教育中“普及化等级”内的最低入门水平，包括汉字入门等级（300 个）、音节入门等级（272 个）和词汇入门等级（505 个）。这是《等级划分》中的一组“最低量指标组合”，是汉语国际教育的一个新特点。最低入门等级既可在世界各地的国际汉语教学中应用，也可用于国内的对外汉语教学，同时还适用于未来现代化的中国国家级计算机辅助汉语口语水平考试（HKK），通过者可获得“最低入门等级证书”。

(2) 设立最低入门等级,目的是降低起步的门槛,稀释学习的难度,回归语言学习的本质。建构最低入门等级,打开汉语国际教育的大门,关键在于实践。“最低入门等级”仅有 272 个音节、300 个汉字、505 个词语,学会这些音节、汉字、词语,就能学会用汉语进行最简单、最基本的日常交际。此举是向社会各界、学界同仁和全球广大学习者发出的正面信号:学习汉语、达到最低入门等级并不难。这是全世界最广大普通学习者又好又快掌握汉语、提高自信心、激发学习兴趣、迈进汉语国际教育大门的有效途径,是汉语国际教育大众化、普及化、规模化的重要环节和标志。

(3) 汉语国际教育重在应用。研制最低入门等级的基本思路和做法是:①最低入门等级中的所有音节、汉字、词汇,原则上要从一级音节表、汉字表、词汇表中选取。这是最重要的一条原则。这样设计既可保证一级的关联性和应用性,又能突出音节、汉字、词汇三个最低入门等级的基础性和通用性。同时,还要与未来中国汉语口语水平考试(HKK)普及化水平内的入门等级接轨。②最低入门等级特别注重交际。选取最基本、最常用的日常交际用字、用词,最大限度地贴近了外国人最基本的交际需要。它包括三个方面:第一,选取一级水平中最基本的结构、功能字词。这类字词是最基本、最基础的字词,都是封闭性、限定性的。如:表示数目、时间、指代和方位的最基本字词,最常用的副词、介词、助词、连词,最常见的“兼语句、双宾句动词”(史有为,2008)。在这方面,中国和世界 60 年汉语教学积累了非常丰富的经验。第二,选取一级水平中最低需要的日常交际字词。这些字词都是开放性、多样化的,灵活性较大。可以以最基本、最常见的“任务、话题、功能、情景”等为中心、为重点,进行综合筛选与平衡。第三,最低入门等级中选取的音节、汉字、词汇,一方面要与更贴近现实生活的学习内容相一致;另一方面又要注意基础教学中的可操作性,即符合基础教学实际与一般规律;同时还要关注常用音节、汉字、词汇之间的多维平衡与合理搭配。

4.3 汉语音节引入新标准

汉语音节引入新标准是这次研发工作最重要的突破口和创新点。

(1) 汉语国际教育发掘汉语音节潜在能力和引领作用最重要、最理想的领域。音节在汉语里是很重要的语言单位。口头汉语的基本单位是音节,汉语音节是有限的。“一般来说,一个汉字代表汉语一个音节”(林焱、王理嘉,1994),汉语“一音一义,单音成义”(王力,1964),就是说,一个音节,一个汉字。这是汉语最重要的特点之一,掌握了汉语的基本音节,千变万化的语音就能运用自如。总之,普通话音节是汉语国际教育基础的基础。

(2) 《等级划分》第一次将汉语音节作为突破口和创新点融入新模式中,它给“国家标准”注入了新鲜血液,使它拥有广阔的国际化视野和强大的生命力,也更具汉语特色。它把音节的概念明确引进汉语国际教育,重视汉语音节的统计和覆盖率,倡导利用音节整体教学,以逐步化解和突破所谓“汉语难学”的瓶颈。它从音节教学入手,抓住了汉语运用的枢纽,将给零起点的基础汉语教学带来重要变革,对汉语国际教育事业及新时期对外汉语教学的发展产生积极而又深远的影响。

(3) 经过在香港这个“世界的窗口”对普通话教学与香港普通话水平考试(PSK)13 年的深入研究,特别是对汉语国际教育大发展的感悟和求索,我们决定以《现代汉语词典》(第 5 版)的“音节表”作为最重要的依据,将汉语音节融入新模式中,创立汉语国际教育使用的基本音节表。《现代

汉语词典》音节表中包括基本音节 1298 个，轻声音节 40 个，共计 1338 个。我们经过几年的论证和反复的实践，遵照删除方言音节、古汉语音节、拟声词音节及书面上的文言词语音节四项原则，从 1298 个音节中删去 92 个，选取 1206 个音节构成汉语国际教育用《基本音节表》。^⑧基于同样的理由，我们又根据 60 年来全世界多国汉语教学的成功经验和国家语委普通话水平测试（PSC）、香港普通话水平考试（PSK）的亲身体验，综合选择 32 个轻声字与轻声音节作为《汉语国际教育用轻声字与常用轻声音节（表）》；选取 1110 个音节和 3000 个汉字构成《按音序排列的分级同音字表》。

（4）这项研究可以为汉语国际教育提供新的重要依据与参照，可以为建立新型教学体系，特别是普及化等级、中级两个阶段的规范口语教学体系，为建构新型的、富有成效的“多轨教学新模式”，提供一种标准化、规范化、多样化的样本。我们赞同不同国家、不同地区根据不同教学需要，利用以上科研成果，开设结合当地文化特色的大众化、本土化课程，利用音节整体教学，“用汉语拼音教学汉语口语”，“尽快提高汉语口语能力”（柯彼德，2003）。

4.4 调整字词筛选顺序、比例与等级

经过多年的不断探索与研究，我们对汉字和词汇的筛选、优化与分级作出了三项重要的设计改变：

（1）改变“先词汇，后汉字”的筛选原则和顺序，打破“词汇牵着汉字走”的传统做法。首先，把分级常用汉字表的筛选放在最重要、最优先的位置。汉字筛选，特别是一级汉字筛选的主要根据依次是：①多种类型和不同类别的大规模语料，特别是大型对话体口语语料的统计频率；②日常生活和口语会话的常用度；③音节的统计与覆盖率；④与其他汉字搭配的构词能力；同时，适当照顾汉字书写的难易度，参考汉字的中华文化内涵（如“茶、筷、饺、龙”）。其次，用现代化手段对多种具有代表性、针对性、影响较大的频率汉字表和频率词汇表，分级、分层次进行交集、比对和筛选。第三，遵照时代性、针对性、实用性、易联想性、可操作性五项原则，对分级汉字表和词汇表进行必要的群体性专家干预与打磨。如一级 900 字的筛选结果，专家干预的比例很小，不足 5%，是比较理想的。

（2）改变词汇量决定汉字量的顺序和做法，提出“提高汉字与词汇量比例的期待值”的新概念。根据多种类别的同级汉字和同级词汇交集、比对的结果，立足于汉字和词汇的“优化配置”，在确保汉字质量及汉字数量不变的情况下，调整同级汉字与同级词汇的配置比例，循序渐进，适度扩大分级词汇量，提高同级汉字“生成”、“搭配”同级常用词的数量和比例，即提高汉字与词汇量比例的“期待值”。这样做既符合汉语的生成规律，又有益于汉字、词汇教学，同时还可以充分发挥现代大型语料库的作用，提高语言材料之间“质”与“量”的关联度，提高整体教学质量与速度。

^⑧ 1995~2008 年，本文作者之一刘英林先后应张日昇教授、张群显博士和陈瑞端教授的邀请，连续 13 年在香港理工大学中文及双语学系任访问教授，主持香港普通话水平考试（PSK）的研究工作。2003 年 2 月，PSK 研究成果通过了教育部、国家语委语言文字规范（标准）审定委员会的权威审定。审定结论是：香港理工大学 PSK 前“三级六等”与国家普通话水平测试（PSC）“三级六等”的等级水平具有等效性。同时，刘英林还与香港孔子学院常务副院长陈瑞端合作研制了《香港地区普通话教学与测试词表》，本课题组研制的《基本音节表》是在上述词表相关内容的基础上，进一步提炼、转型、升级研制出来的。在此一并说明，并向香港理工大学等有关院校、单位长期的合作伙伴致谢！

(3) 改变词汇与汉字分为四个等级的惯常做法,采用三级划分的统一规划。

关于汉字等级划分:根据汉语国际教育的总体目标和 60 年对外汉语教学、世界汉语教学的经验,《等级划分》三级水平的汉字基本量定为 2700 字。汉字等级划分主要特点是采取均衡性、灵活性和汉字词汇覆盖率相结合的新思路、新手法,具体包括:①均衡性:将 2700 字均衡地分为三个等级,一级汉字 900 字;二级汉字 900 字,一、二级累计为 1800 字;三级汉字 900 字,一、二、三级累计为 2700 字。这三个等级还与未来新型中国汉语口语水平考试(HKK)的三级评价体系相衔接。②灵活性:在普及化水平内设立最低入门等级,汉字为 300 字,占一级 900 字的 1/3;同时,又在高级水平后增补 300 字高级“附录”汉字,汉字总量达到 3000 (2700+300) 个。它使《等级划分》具有更大的灵活性和包容性。③汉字覆盖率:关注、重视大型对话体动态语料及汉字覆盖率是这次研发工作的一个突出特点。国家语言资源监测与研究中心有声媒体语言分中心提供的数据显示,对话体动态语料 300 字覆盖率为 72.80%, 900 字覆盖率达到 92.10%, 1800 字覆盖率达到 98.21%, 2700 字覆盖率达到 99.59%, 3000 字覆盖率达到 99.75%。《等级划分》3000 常用字和次常用字,是经过反复比对、推敲精选出来的,其覆盖率比上述相应级别的覆盖率更高。

关于词汇等级划分:根据世界各地国际汉语教学和国内对外汉语教学的总体战略目标,《等级划分》三级水平的词汇基本量定为 9600 个左右。词汇等级划分重点关注点有两个:①注重包容性、层次性:第一,普及化等级分为三个小层次:一①最低入门等级最常用词 505 个,一②最常用词 837 个,一③常用词 903 个,总计 2245 个。第二,高级水平分为两个小层次:三级(高级)词汇 4175 个,高级“附录”词汇 1461 个,共计 5636 个。这样的精心设计,可为国内外不同教学环境,不同学习目的,使用不同教学法的教学机构、单位和个人多层次、多样化的需求提供科学依据和方便。②注重三级词汇量的配置比例:普及化等级、中级和高级阶段的词汇量分别是 2245 个、3211 个、5636 (4175+1461) 个,词汇总量达到 11092 个;三大级别词汇的配置比例大致为 2:3:5。这样的比例配置符合词汇学习规律,便于指导教学。

无论是汉字等级划分还是词汇等级划分,《等级划分》对于世界各地的国际汉语教学和国内的对外汉语教学来说,都具有较强的实用性、针对性、灵活性和多用途适应性。

(4) 这项研究跳出了惯常模式,对汉字和词汇筛选原则、顺序、比例和等级作出了重大调整,进行了深层次的改革;与此相适应,充分整合各种字频与词频统计,为用现代化手段科学、合理、适度、有效地扩大词汇量找到了新方法,开辟了新途径。如:《等级大纲》甲级字 800 个,甲级词 1033 个,“汉字词汇比”为 1:1.29;《等级划分》普及化等级汉字 900 个,“生成”、“搭配”常用词 2245 个,“汉字词汇比”为 1:2.49。

4.5 创建三维基准新体系

创建三维基准新标准、新模式、新体系是这次研发工作最重要的着力点。

(1) 1992 年出版的《等级大纲》是“汉字跟着词汇走”的二维基准模式。它是“一种规范性的水平大纲”(刘英林、宋绍周,1992),在那一历史时期,它曾被奉为对外汉语教学的“圭臬”(姜德梧,2004),在总体设计、教材编写、课堂教学和水平测试等诸多方面发挥了重要的作用,引起了

众多专家、学者的关注（李英，1997；李晓琪，1997；李清华，1999；李红印，2005；苏新春，2006；张和生，2006；周健、李海霞，2008）。但是，这种模式已经不能适应当前形势发展的要求，必须更新观念与方式，深入进行面向汉语国际教育的汉语要素应用研究，创立一个面向全球的、具有鲜明时代和学科特点的新型《等级划分》（国家标准）来取代它。

（2）《等级划分》主要“新”在五个方面：①《等级划分》依据的是30多亿字次、多家有代表性的最新大型动态语料库和多种有针对性的词典、词表、字表，它们为创制科学、系统、全面的“国家标准”奠定了坚实、可靠的基础。②在汉语国际教育的新形势下，通盘考虑了等级划分、等级水平和等级测试的完整性、均衡性及今后长远发展的连贯性、稳定性，建立了长效机制，将其统一规划、系统调整为三个等级、三级水平及三级评价体系。③立足于从汉语自身特点入手，引入音节，重视汉字，首创汉语三要素——音节、汉字、词汇并行的三维基准新标准、新模式、新体系。④建立现代汉语规范口语主导型的新型“国家标准”。以词语通用性作为最重要的标准：普及化水平和中级水平略倾向于普通话口语，高级水平略倾向于普通话书面语。⑤探究汉语国际教育新理念、新路径、新途径，着力推行普及化水平特别是最低入门等级，适度增加同级汉字构建同级常用词的配比，多方位、多层面地解决汉语国际教育领域中的瓶颈问题，加速汉语的国际化、大众化、普及化、规模化进程。

（3）《等级划分》设计理念成型于2007~2008年。我们经过17年的不断思考、积淀和梳理，经过课题组多学科专家一年多的集思广益和预选试验，决定以创新的理念自主研发新型的“国家标准”。它集音节功能、汉字功能和词汇功能于一身，实现了汉语自身三大要素的融合发展。它运用现代科技手段，充分、合理、有效地利用计算机自动分词系统科研成果，寻求频度最高，流通度最高，最常用、最适合交际和操练的音节、汉字和词汇，注重它们的优选和优化。新体系中汉语三要素是相互贯通、互相依存、不可分割的统一整体，将它们组合成均衡、稳固的“铁三角”，意在系统整合与融通，意在用最经济、关联度最紧密、最实用的语言材料进行组合，实现教学效果和学习效率的最大化、最优化。

（4）《等级划分》是适应汉语国际教育的重大转变，在求真务实的基础上对1992年《等级大纲》综合进行的多角度、多层次、多样化的改革、转型与升级。创立汉字、音节、词汇“三位一体”的《等级划分》新标准、新模式、新体系，是为了使汉语国际教育更符合教育现代化和多样化的要求，更具科学性、时代性、实用性和包容性，更富有中华文化特色。

4.6 探寻多维平衡理念和方法途径

“多维平衡”、“综合协同”的新理念和新方法是这次研制工作关注的焦点之一。

（1）寻求多维平衡的理念和方法。在继承传统与创新发 展之间找平衡；在“请进来”的对外汉语教学与“走出去”的国际汉语教学之间找平衡；在普及化水平、中级水平、高级水平之间找平衡；在现代计算机自动分词系统的研究成果与传统语言研究、语言教学之间找平衡；在规范口语教学与普通话书面语教学之间找平衡；在汉语三要素音节、汉字、词汇之间找平衡；词汇筛选“在语用、语义和句法三种标准”（马清华，2008）之间找平衡。

(2) 音节、汉字、词汇是汉语国际教育的基石。我们认真探寻新时期《等级划分》的汉语三要素(音节、汉字、词汇)、两个层面(宏观与微观)、两个方向(横向与纵向)相结合的多维平衡,综合协同新理念、新方法;我们强调平衡、综合,强调发挥整体合力,强调协同发展,这表现在“质”和“量”两个方面;同时在平衡、综合和协同的过程中,往往会产生新的创意。

(3) 运用多维平衡、综合协同理念和方法途径,最重要的有三点:一是创新,一是务实,一是细节。既要有宏观视野,又要关注细节,因为在很多情况下,细节决定品质。

4.7 运用现代科技手段,吸收常见常用结构

这是这次研发新型《等级划分》的一个鲜明特点。

(1) 运用现代科技手段。汉语国际教育必须走科技之路。《等级划分》中汉字和词汇的优选、优化是无法通过某一种类型的、单一的字频和词频统计数据得到的,必须运用现代科技手段,加强数据源头的质量控制,掌握第一手原始数据,以大规模标准口语对话体动态语料为基础和核心,以多家不同类别的汉字、词汇动态频率统计和多种具有代表性、针对性的词典、词表、字表为重要依据,对汉字和词汇进行多类别、多角度、多层次的交集、比对与分级。这种交集、比对与分级不是一次性完成的,而是根据不同语言要素和不同级别的需要,分层次交叉衔接,多次性完成的。

(2) 运用计算机自动分词系统的科研成果,吸收常见、常用结构。为适应国家汉办/孔子学院总部提出的“学得简单,学得容易,学得快”的要求,从学术思维转为务实思维,从学术系统转为应用系统,尽量把复杂的问题简单化,以多类别的频率统计作为最重要的依据,科学、合理、适度地筛选一些常见、常用结构,收入普及化等级和中级水平,如“下去、回去、上来、出来;收到、回到、遇到、来到;开车、打电话、看病、回家;比如说、实际上、就是说、基本上;是不是、有没有、能不能、是否;请问、家里、随着、在于”,等等,使词语表更贴近语言生活实际。

4.8 关注等级质量,与未来发展接轨

关注等级质量,与未来发展接轨,应该进行认真检讨,科学规划,统筹兼顾。

(1) 等级质量是《等级划分》的命脉。所谓等级质量,主要是指《等级划分》各个级别的内容、范围和等级界定等方面的整体质量,是指每一级等级划分的数据质量控制。我们提出“等级质量”这个新概念,是为了使《等级划分》更加系统化、精密化、多功能化;是为了使“国家标准”更有效地发挥引领和积极导向作用:导方向、导重点、导方法、导规律;是为了更好地“教学生学”;是为了在有限的(音节、汉字、词汇)范围内创造出最大的效益。

(2) 与未来长远发展接轨。我们既强调立足于汉语国际教育的现实需要,又着眼于汉语国际教育的长远持续发展。《等级划分》着眼于与未来汉语国际推广的学校教育和社会教育接轨,着眼于与新形势下的对外汉语教学和新汉语水平考试接轨,着眼于与现代化、信息化、多媒体、网络教育接轨。

根据《等级划分》这一最新研究成果,课题组将在更大范围、更高技术层面上展开“强强”、“跨界”合作与升级,充分发挥自身文化优势和团队精神,在建立《汉语国际教育用分级词语节律语言材料库》的同时,研制《汉语国际教育用多维平衡语言材料和自动评测系统》(拟作为国家标准);

并在此基础上,研发具有自主知识产权的《汉语国际教育用中国国家级计算机辅助汉语口语水平考试》(HKK)。HKK 与三级等级划分和等级水平相适应,分为普及化水平、中级水平和高级水平三级九等。

参考文献

- 白乐桑 张 丽(2008)《欧洲语言共同参考框架》新理念对汉语教学的启示与推动——处于抉择关头的汉语教学,《世界汉语教学》第4期。
- 白乐桑 张朋朋(1997)《汉语语言文字启蒙》1、2,北京:华语教学出版社。
- 陈瑞端 刘英林(2008)《香港地区普通话教学与测试词表》,香港:商务印书馆(香港)有限公司。
- 国家对外汉语教学领导小组办公室(2002)《高等学校外国留学生汉语言专业教学大纲》,北京:北京语言大学出版社。
- 国家对外汉语教学领导小组办公室汉语水平考试部 北京语言大学汉语水平考试中心(刘英林主编)(1992)《汉语水平词汇与汉字等级大纲》,北京:北京语言学院出版社。
- 国家对外汉语教学领导小组办公室汉语水平考试部 北京语言大学汉语水平考试中心(刘英林主编)(1996)《汉语水平等级标准与语法等级大纲》,北京:高等教育出版社。
- 国家汉办/孔子学院总部教材处(2008)《汉语教材常用词表及常用字表统计分析报告》。
- 国家语言文字工作委员会(1988)《现代汉语常用字表》,北京:语文出版社。
- 国家语言文字工作委员会普通话培训测试中心(2008)《普通话水平测试实施纲要》,北京:商务印书馆。
- 姜德梧(2004)关于《汉语水平词汇与汉字等级大纲》的思考,《世界汉语教学》第1期。
- 柯彼德(2003)汉语拼音在国际汉语教学中的地位和运用,《世界汉语教学》第3期。
- 李红印(2005)《汉语水平词汇与汉字等级大纲》收“语”分析,《语言文字应用》第4期。
- 李乐毅(2008)“孔子学院”离不开汉语拼音(纪念《汉语拼音方案》颁布50周年学术研讨会发言),《语言文字应用》第3期。
- 李清华(1999)《汉语水平词汇与汉字等级大纲》的词汇量问题,《语言教学与研究》第1期。
- 李晓琪(1997)《HSK 词汇等级大纲》中形容词和副词的词类标注问题,《汉语学习》第4期。
- 李 英(1997)关于《汉语水平词汇与汉字等级大纲》的几个问题,《中山大学学报论丛》第4期。
- 林 焘 王理嘉(1994)《语音学教程》,北京:北京大学出版社。
- 刘长征(2007)《词汇大纲》与2005媒体高频词语比较研究,《云南师范大学学报》(对外汉语教学与研究版)第3期。
- 刘长征 张 普(2008)对外汉语教学用词表的多元化与动态更新,《语言文字应用》第2期。
- 刘英林主编(1989)《汉语水平考试(HSK)研究》,北京:现代出版社。
- 刘英林主编(1994)《汉语水平考试(HSK)研究》(续集),北京:现代出版社。
- 刘英林(2001)普通话水平考试的理论思考与标准化,《中国语文》第1期。
- 刘英林 宋绍周(1992)汉语常用字词的统计与分级,《中国语文》第3期。
- 吕叔湘主编(1998)《现代汉语八百词》(增订本),北京:商务印书馆。
- 罗青松(2006)《美国21世纪外语学习标准》评析——兼谈《全美中小学中文学习目标》的作用与影响,《世界汉语教学》第1期。
- 马清华(2008)唯频率标准的不足性——论面向汉语国际教育的词汇大纲设计标准,《世界汉语教学》第2期。
- 史有为(2008)对外汉语教学最低量基础词汇试探,《语言教学与研究》第1期。
- 苏新春(2006)对外汉语词汇大纲与两种教材词汇状况的对比研究,《语言文字应用》第2期。