

汉语能力测试的 理论

■ “国民语言文字能力标准与测评体系研究”课题组 编

高等教育出版社

汉语能力测试的理论与实践

HANYU NENGLI CESHI DE LILUN YU SHIJIAN

“国民语言文字能力标准与测评体系研究”课题组 编

高等教育出版社·北京

图书在版编目(CIP)数据

汉语能力测试的理论与实践 / “国民语言文字能力
标准与测评体系研究”课题组编. --北京:高等教育出
版社, 2016. 4

ISBN 978-7-04-045061-3

I. ①汉… II. ①国… III. ①汉语-水平考试-研究
IV. ①H19

中国版本图书馆 CIP 数据核字(2016)第 050952 号

策划编辑 雷旭波 责任编辑 王江媛 封面设计 赵 阳 版式设计 马 云
责任校对 刘 莉 责任印制 耿 轩

出版发行	高等教育出版社	网 址	http://www.hep.edu.cn
社 址	北京市西城区德外大街 4 号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.hepmall.com.cn
印 刷	中国农业出版社印刷厂		http://www.hepmall.com
开 本	880mm×1230mm 1/32		http://www.hepmall.cn
印 张	6.5	版 次	2016 年 4 月第 1 版
字 数	180 千字	印 次	2016 年 4 月第 1 次印刷
购书热线	010-58581118	定 价	15.00 元
咨询电话	400-810-0598		

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换

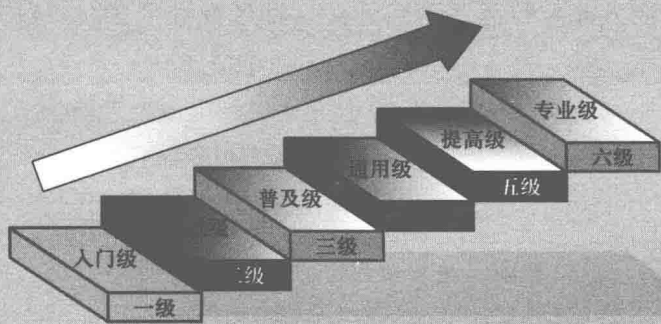
版权所有 侵权必究

物 料 号 45061-00

前言

为了贯彻执行《中华人民共和国国家通用语言文字法》，满足社会对汉语言文字运用水平评价的要求，促进人们汉语应用能力的提高，在中华人民共和国教育部（以下简称教育部）、国家语言文字工作委员会（以下简称国家语委）的指导下，教育部考试中心组织实施了汉语能力测试项目。该测试作为一项国家级语言类标准化水平测试，旨在评估以汉语作为生活、学习、工作基本用语人群的汉语应用能力。本测试可用于鉴定考生的汉语综合应用能力，帮助应试者了解自己的汉语应用能力；可为相关用人单位了解员工的汉语水平提供参考；测试结果还可以为各级各类教育机构开展汉语教育和培训提供参考。

汉语能力测试项目学术委员会和课题组经过调研，依据人们完成语言运用任务时的表现，同时参考使用汉语人群的职业和学历，由低到高设置了六个等级：一级为入门级，二级为基础级，三级为普及级，四级为通用级，五级为提高级，六级为专业级（如下图所示）。



语言运用能力等级图

汉语能力测试作为一个综合性的语言能力评价系统，既包括对



汉语口语和书面语的理解能力(输入)的评价,也包括对这两方面表达能力(输出)的评价。所以,测试共分为口语理解、口语表达、书面语理解和书面语表达四个部分,分别从听、说、读、写四个方面考查考生的汉语应用能力。上述各语言能力要素在不同级别中的权重不尽相同,如低等级(一级和二级)的测试偏重于对考生口语能力的考查;在高等级(三级和以上)的测试中,对书面语能力考查所占的比重则相对更大。

汉语能力测试选用基于计算机的施测方式(computer based testing),这是因为本项目是对汉语能力的综合测试,利用计算机技术可以有效整合各种所需的媒体资源;另一方面,随着信息技术在人类生活中的逐步渗透,它已经成为处理语言文字的主要工具。

本书在“国民语言文字能力标准与测评体系研究”课题组成员的共同努力下完成,其中邱静远、王莉主要负责第一章,华东师范大学的巢宗祺教授撰写了第二章,第三章由孙显福、张均兵、杨卓、赵允溪共同完成,吴瑞林等人承担了第四章和第六章的编撰工作,郭轶舟、刘鹏、鲍恒超负责第五章的内容。全书的框架和定稿由课题组负责人完成。

本书系国家语委“十二五”科研规划 2011 年度重大项目“国民语言文字能力标准与测评体系研究”(课题批准号 ZDA125-1)的成果。

目 录

第一章 国内外语言类测试现状分析	1
1.1 国内主要语言测试比较分析	1
1.1.1 校内汉语测试	1
1.1.2 社会汉语测试	2
1.1.3 现有社会汉语类测试的比较与评析	5
1.2 汉语能力类测验的题型分析	6
1.2.1 考查口语表达能力的题目	7
1.2.2 考查书面语理解能力的题目	8
1.2.3 考查书面语表达能力的题目	12
1.2.4 对题型的讨论	17
1.3 美国 ACT 和 SAT 语言测试及启示	17
1.3.1 ACT 和 SAT 的测试依据和目标	18
1.3.2 ACT 和 SAT 的测试结构和内容	19
1.3.3 ACT 和 SAT 的评价与服务	25
1.3.4 ACT 和 SAT 的启示	26
第二章 汉语能力标准与测试大纲	28
2.1 汉语能力总体描述	28
2.2 汉语能力分级描述	30
2.3 汉语应用能力等级划分的说明	40
2.4 汉语能力测试四级大纲	43
2.4.1 汉语能力测试四级级别标准	43
2.4.2 汉语能力测试四级整体描述	44
2.4.3 汉语能力测试四级试卷结构与说明	44
2.4.4 汉语能力测试四级题型示例及评分标准	45
2.5 汉语能力测试五级大纲	56



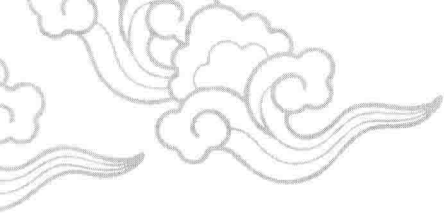
2.5.1	汉语能力测试五级级别标准	56
2.5.2	汉语能力测试五级整体描述	57
2.5.3	汉语能力测试五级试卷结构与说明	57
2.5.4	汉语能力测试五级题型示例及评分标准	58
第三章	面向能力的情境化测试题型	72
3.1	采用计算机情境命题的必要性和挑战	72
3.1.1	汉语能力测试的背景和特点	72
3.1.2	计算机情境命题的必要性和面临的挑战	73
3.2	对剑桥若干考试的借鉴	73
3.2.1	标准与大纲(以 PET 为例)	74
3.2.2	题型对比分析(以 A-level 为例)	77
3.2.3	测试方式与评分	78
3.2.4	分数报告方式(以 A-level 和 GCSE 为例)	80
3.2.5	当前英国考试界对计算机化考试的主要 观点	81
3.3	汉语能力测试命题的思路	82
3.3.1	标准参照考试	82
3.3.2	命题依据	83
3.3.3	汉语能力测试的试题特征	85
第四章	现代测量和评价技术的应用	90
4.1	面向评价的考试与测量	90
4.1.1	评价工作的基本思路	90
4.1.2	汉语能力测试的评价内容与形式	92
4.1.3	汉语能力测试评价的实施	98
4.1.4	汉语能力测试评价工作存在的困难	101
4.2	汉语能力测试中的题组及其效应	102
4.2.1	题组题及其处理	102
4.2.2	题组分析的模型	103
4.2.3	汉语能力测试中的题组题	106



4.2.4	独立模型与题组模型比较	108
4.3	固定路径 CAT 测试技术	113
4.3.1	语言能力测验与 CAT	114
4.3.2	在阅读和听力测验中使用 CAT 技术所面临的 挑战	116
4.3.3	固定路径 CAT 的思想	118
4.3.4	固定路径 CAT 的测量精度	120
4.3.5	固定路径 CAT 的优点	121
第五章	网络化测试平台与考务管理	124
5.1	计算机化考试系统的总体设计	124
5.2	基于计算机网络的考务工作	128
5.2.1	考务实施机构及考务工作	128
5.2.2	安全保密工作	130
5.2.3	报名与准考证的编排和打印	131
5.2.4	考前准备与考试实施	132
5.2.5	评卷、成绩查询和证书管理	136
5.2.6	考试工作日程	137
5.3	汉语能力测试考点设备标准	138
5.3.1	考场条件需求	138
5.3.2	考试用机软硬件环境要求	138
第六章	试测实施的效果和数据分析	140
6.1	试测目标与实施过程	140
6.1.1	试测目标	140
6.1.2	试卷结构	140
6.1.3	测试实施	142
6.2	试测成绩分析	143
6.2.1	四级试测成绩	143
6.2.2	五级试测成绩	145
6.2.3	四级与五级成绩比较	147



6.3 客观题测量学分析	148
6.3.1 四级客观题的经典测量理论分析	149
6.3.2 五级客观题的经典测量理论分析	151
6.3.3 客观题的测量维度分析	152
6.3.4 四级与五级客观题的难度比较	154
6.3.5 单选题与多选题的难度比较	156
6.4 主观题评分情况	157
6.4.1 主观题评分过程	157
6.4.2 主观题评分标准	158
6.4.3 主观题评分结果	163
6.4.4 主观题评分维度的分析	166
6.5 结论与建议	169
6.5.1 四级与五级试卷间的难度区分不明显	170
6.5.2 应加大对考题类型和考核点的有关研究	170
6.5.3 与传统语文考试的区别可以进一步突出	170
6.5.4 亟须指导和规范命题工作的有关文档	171
6.5.5 考生的作答动机对试测结果影响较大	171
第七章 媒体关注与各界声音	172
7.1 领导对试点工作的指示	172
7.2 主要媒体报道汇总	173
7.3 代表性观点呈现	175
7.3.1 “汉语四六级”,在争议中起步	176
7.3.2 汉语测试“教”我们如何教语文	179
7.3.3 为什么要进行汉语能力测试	181
7.3.4 汉语考试考倒了谁? 国人需要一场母语 考试吗?	183
7.3.5 中国人测汉语能力,究竟怎么考?	186
主要参考文献	192



第一章

国内外语言类测试现状分析

1.1 国内主要语言测试比较分析

国内汉语类测试可以分为两大类:一为校内汉语测试,例如某校一年级的语文期末考试;二为社会汉语测试,例如汉语水平考试(HSK)。

1.1.1 校内汉语测试

校内汉语测试以语文高考为典型。这是因为在我国,高考语文学科目既检测考生该科目的学习情况,又为高等学校的招生选拔提供依据,还担负着指导中学语文教学的使命。2003年,教育部颁布《普通高中语文课程标准(实验)》作为指导语文教学与学习的纲领性文件,该文件也是高考语文考试大纲制定的依据。以2011年高考语文全国卷考试大纲为研究对象可知,高考中对语文能力的考查分为五个层级,即识记、理解、分析综合、表达应用和鉴赏评价。考查内容包括语言知识和语言表达、文学常识和名句名篇、古代诗文阅读、现代文阅读、写作。考试形式为闭卷笔试。试题类型包括单项选择题、多项选择题、填空题、古文翻译题、简答题和写作题。可以看出,高考中对语文学科目主要考查的是阅读和写作这两项能力。在这种情况下,不少学校的语文教育出现重读写,轻听说;重书面,轻口语;重应试,轻能力的现象。在我国,不同的校内汉语测试其模式却是类似的,而且这方面的研究较多。

1.1.2 社会汉语测试

社会汉语测试因其涉及的考生面广,覆盖人群多样,社会影响力正逐渐增强。根据其测试对象的不同,社会汉语测试可分为面向母语为非汉语人群的测试项目和面向母语为汉语人群的测试项目。

(一) 面向母语为非汉语人群的汉语测试项目

目前,以母语为非汉语人群为主要测试对象的汉语测试项目有汉语水平考试(HSK)、中国少数民族汉语水平等级考试(MHK)、国家汉语口语水平测试(HKC)、商务汉语考试(BCT)、中小学生汉语考试(YCT)和实用中国语水平认定考试(C.TEST)等。(见表1.1)

表 1.1 面向母语为非汉语人群的汉语测试项目一览表

测试名称	主办方	测试对象	标准和大纲	测试内容	测试时间	测试形式	证书颁布
汉语水平考试(HSK)	国家汉语国际推广领导小组办公室	第一语言非汉语的人群	《国际汉语能力标准》 《国际汉语教学通用课程大纲》	听力理解/阅读理解/语法结构/综合填空/综合表达等	120 分钟左右	笔试和口试	分为 HSK 基础, HSK 初、中、高等, HSK 少儿, HSK 商务, HSK 文秘, HSK 旅游等
中国少数民族汉语水平等级考试(MHK)	教育部民族教育司	母语非汉语的中国各少数民族学生	《中国少数民族汉语水平等级考试大纲》	听力理解/阅读理解/书面表达	110 分钟	笔试和口试	1—4 级

资料来源:根据教育部、国家语委《汉语水平考试(HSK)考试大纲》(2003)整理。



续表

测试名称	主办方	测试对象	标准和大纲	测试内容	测试时间	测试形式	证书颁布
国家汉语口语水平测试(HKC)	国家语言文字工作委员会	第一语言非汉语的人群	《汉语口语水平等级标准及测试大纲》	读词/ 读句/ 提问/ 设问/ 复述/ 描述/ 说话	30 分钟	口试	三等九级
商务汉语考试(BCT)	国家汉语国际推广领导小组办公室	第一语言非汉语,从事商务活动的人群	《商务汉语考试大纲》	听说/ 读写	150 分钟	笔试和口试	1—5 级,考生可得到成绩单,合格者可获得证书
中小学生汉语考试(YCT)	国家汉语国际推广领导小组办公室	15 岁以下,第一语言非汉语的人群	《新中小学生汉语考试大纲》	听力/ 阅读/ 写作	45~60 分钟	笔试	1—3 级
实用中国语水平认定考试(C.TEST)	北京语言大学汉语水平考试中心	第一语言非汉语的人群		听力理解/文法/ 阅读/综合运用	115~160 分钟	笔试	初级(E—F 级)、中高级(A—D 级)

(二) 面向母语为汉语人群的汉语测试项目

目前,以母语是汉语人群为主要测试对象的汉语测试项目有汉字应用水平等级测试(HZC)、普通话水平测试(PSC)和国家职业汉语能



力测试(ZHC)等。(见表 1.2)

表 1.2 面向母语为汉语人群的汉语测试项目一览表

测试名称	主办方	测试对象	标准和大纲	测试内容	测试时间	测试形式	证书颁布
汉字应用水平等级测试(HZC)	国家语言文字工作委员会	中等以上受教育人群	《汉字应用水平测试字表》 《汉字应用水平测试大纲》	书写填空题/ 选择题/ 选择填空题	80 分钟	纸笔测试	分为三级六等, 统一发放成绩单
普通话水平测试(PSC)	国家语言文字工作委员会	教师、播音员、节目主持人、公务员、学习的专业与口语表达关系密切的学生	《普通话水平测试等级标准(试行)》	100 个单音节字词/50 个双音节词语/判断测试/作品朗读/说话	15 分钟	口试	三级六等, 证书由国家语言文字工作委员会统一印制
国家职业汉语能力测试(ZHC)	人力资源和社会保障部职业技能鉴定中心	职场人群	《国家职业汉语能力测试大纲》	阅读理解/书面表达/作文	150 分钟	纸笔测试	初、中、高三个等级, 成绩合格者可获得相应等级的证书



1.1.3 现有社会汉语类测试的比较与评析

从表 1.1 和表 1.2 可以看出,国内汉语类测试项目种类繁多,多头管理,标准不一,且各有侧重。

从主办及承办方来看,有国家语言文字工作委员会(简称“国家语委”)和国家汉语国际推广领导小组办公室(简称“国家汉办”)这样的政府专职机构,有北京语言大学这样在国内专门从事对外汉语教育的高等院校,也有以职业培训、资格认定为导向的人力资源和社会保障部职业技能鉴定中心。目前国内社会汉语类测试项目的承办方呈现多头主体的特点,而且承办方之间并没有明显的隶属关系。

从测试对象来看,在汉语为非母语人群的汉语测试中,有覆盖面广泛的汉语水平考试(HSK),也有针对从事商务活动人群的商务汉语考试(BCT),但是在面向母语人群的汉语测试中,尚缺乏面对所有母语为汉语人群的综合测试与评价体系。

从标准和考试大纲来看,各个机构研发的汉语测试项目都制定了各自的标准,使得目前社会上的汉语测试无论测试对象是母语人群还是非母语人群都缺乏统一的指导性纲领文件。

从测试内容来看,以汉语水平考试(HSK)为代表的非母语测试覆盖了听说读写四项语言能力。但反观母语测试,目前的汉字应用水平等级测试(HZC)、普通话水平测试(PSC)、国家职业汉语能力测试(ZHC)只偏重于测试听说读写四项语言能力中的一项或者两项,有各自明确的目标内容。

从测试方式来看,对汉语为非母语人群的测试更加注重口语沟通与交际能力的测评,在汉语为母语人群的测试中除普通话水平测试(PSC)对应试人在运用普通话口语进行表达过程中所表现的语音、词汇、语法规程程度进行评定外,尚无对口语沟通与交际能力的测评。

另外,目前国内社会汉语类测试的标准不一导致考试结果不具备任何的可比性,给使用测试成绩的机构和学习者自身带来很多不便。如果要使不同的社会汉语类测试具有可比性,就必须制定统一的测试评价体系,不同的考试机构都在这个共同的测试评价体系内进行测试

研发,并以它为依据来报告测试的结果。

1.2 汉语能力类测验的题型分析

汉语是我国的通用语言文字,是我国不同地区、不同民族之间沟通的桥梁。然而随着世界经济一体化的发展和我国的对外开放,国内“外语热、汉语冷”的现象造成国人汉语应用能力水平的下降。这也引起了语言文字工作管理机构和社会各界的关注。

本节将对国家职业汉语能力测试(ZHC)、普通话水平测试(PSC)、汉字应用水平等级测试(HZC)和国家公务员考试这四种将汉语作为第一语言并从不同角度考查其应用能力的考试的样题和真题进行分析,以期对新的汉语能力类测试项目的开发有所帮助。

在语言学领域,对语言能力的研究始于艾弗拉姆·诺姆·乔姆斯基(Avram Noam Chomsky),他在著名的转换—生成理论中提出了“语言能力”的概念。他认为语言的功能在于充当认知增长的媒介,而不是交际的工具。1990年,美国应用语言学家巴克曼(Bachman)在综合各家之说的基础上提出了交际语言能力(Communicative Language Ability)的观点。他认为,语言交际能力包括语言能力、策略能力和心理、生理运动机制三部分。其中,语言能力包括组织能力和语用能力。交际语言能力观是目前影响力最大的语言能力理论。

语言能力分为两个方面:一是语言理解能力,主要指对他人语言的把握能力,包括口语理解、书面语理解两部分。口语理解包括对语音的辨认、语义理解等;书面语理解包括对文字的辨识,对字、词、句、篇章等意义的理解,对感情色彩的把握等。二是语言表达能力,主要指运用语言进行交流的能力,即让他人明白自己所说的能力,包括口语表达与书面语表达两个部分。在社会实践中,人们习惯将语言活动根据语言交际的方式和过程分为听、说、读、写四部分。与此对应,语言能力也可分为口语理解能力、口语表达能力、书面语理解能力和书面语表达能力。据此标准,上述国家职业汉语能力测试、普通话水平测试、汉字应用水



平测试以及国家公务员考试皆没有涉及口语理解能力的题目,下面将从口语表达、书面语理解和书面语表达三个方面展开比较。

1.2.1 考查口语表达能力的题目

普通话水平测试是上述这四项考试中唯一针对口语表达能力进行考查的项目,汉字应用水平测试中也包含部分对字词发音测试的题目,可以被算作考查口语表达能力的试题。根据普通话水平测试大纲,普通话水平测试考查应试人的普通话规范程度、熟练程度,认定其普通话水平等级,属于标准参照性考试。测试以机试的方式进行。根据考查方式划分,它的题型包括:

(一) 读字词

包括读单音节字词和多音节词语。

例题(选自普通话水平测试样题):

澎(单音节词) 拐带(双音节词)

上例中,普通话水平测试要求考生直接朗读字词,而下例中汉字应用水平测试则不同。如下面两道例题所示,它通过为字词加注音,或者判断两个汉字读音是否相同的形式考查考生对汉字普通话读音的掌握情况。

例题(选自汉字应用水平测试样题):在下列各选项的四个字中,请找出加点字注音错误的一项。

A. 一丘之貉(hé)

B. 寡廉鲜耻(xiǎn)

C. 引吭高歌(kàng)

D. 人头攒动(cuán)

参考答案:C

例题(选自汉字应用水平测试样题):请找出下面4组词语中,加点字读音不同的一组。

A. 暴露——露骨

B. 参差——差错

C. 称职——对称

D. 处罚——处理

参考答案:B

(二) 朗读短文

给定一段文字进行朗读,要求语音标准,语调自然,表达流畅。题



库中共 60 篇短文,已在大纲中公布,考生可提前练习、做准备。

例题(选自普通话水平测试样题):

盼望着,盼望着,东风来了,春天的脚步近了。

一切都像刚睡醒的样子,欣欣然张开了眼,山朗润起来了,水涨起来了,太阳的脸红起来了。

.....

春天像健壮的青年,有铁一般的胳膊和腰脚,领着我们上前去。

(三) 命题说话

要求考生在给定的 30 个话题中抽取 2 个话题,选择其中一个进行为时 3 分钟的说话,其目的是考查考生在无文字凭借情况下说普通话的规范程度。

例题(选自普通话水平测试样题):我喜爱的动物

从以上题型可看出,普通话水平测试的题型虽为主观题,却在一定程度上具有客观性。读字词和朗读短文的答案有明显的客观性。虽然考生可对命题说话自由发挥,但考核的重点是普通话的规范程度,而非考生表达的内容和思想。因此,上述四项考试中并没有真正重点考查汉语口语表达的题目。

1.2.2 考查书面语理解能力的题目

在上述四项汉语能力类考试中,国家公务员考试和国家职业汉语能力测试有针对书面语理解能力的考查。对书面语理解能力的考查主要集中在语段和篇章上,包括独立题(Independent Item)和题组题(Testlet Item)。独立题是指一段语料带一个问题,公务员考试和国家职业汉语能力测试中都有这类题目出现。题组题是指一段语料带两个或两个以上的问题,这些问题都依附于同一段语料,有时问题间还存在某种联系。国家职业汉语能力测试设计了题组题,语料长度为 500 到 1000 字,每篇语料后带有 3 至 5 个问题。从表现形式上看,国家公务员考试和国家职业汉语能力测试都采用了选择题的形式,属于客观性试题。

无论题组题,还是独立题目,对阅读理解能力的考查都可以分为以