

基于《国际中文教育中文水平等级标准》 的字词关系资源库构建研究

Research on the Construction of Resource Database of the Relationship Between Zi and Ci Based on *the Chinese Proficiency Grading Standards for International Chinese Language Education*

白小丽¹

BAI Xiaoli

西南大学国际学院

International College, Southwest University

E-mail: bxl80@163.com

朱海涵²

ZHU Haihan

西南大学国际学院

International College, Southwest University

E-mail: haihanzhu6@gmail.com

摘 要 汉字和词汇是《国际中文教育中文水平等级标准》中的重要组成部分，而《等级标准》中字与词的对应关系存在一对一、一对多、多对一等多种情况，复杂的字词对应关系导致了汉字难、汉语难。为解决汉语二语学习者这一困难，本文提出以汉字为出发点，以汉字取用理论为指导，将《等级标准》中所列汉字和词语作为整理对象，提出构建字词关系资源库的具体思路及方法，以期建设一个集字形、结构、记录词语职能为一体的资源库，为国际中文教师提供可靠的汉语教学资源，同时助力汉语学习者对汉语汉字、词汇的深度学习。

关键词 《等级标准》；字词关系；资源库构建

收稿日期：2022 年 10 月 15 日

作者简介：¹ 白小丽，女，中国四川内江人，西南大学国际学院副教授，硕士生导师。主要研究方向为文字学、国际中文教育。

² 朱海涵，女，中国山东日照人，西南大学汉语国际教育硕士。主要研究方向为国际中文教育。

基金项目：重庆市研究生教育教学改革研究一般项目——新形势下国际中文教育汉语语言要素课程资源建设（项目编号：yjg223026）。

Abstract Chinese characters and vocabulary are important parts of *the Chinese Proficiency Grading Standards for International Chinese Language Education (Grading Standards for short)*, and are also an important content for Chinese learners to learn Chinese. However, the correspondence between Chinese characters and words in the Grading Standards exists in many cases, such as one-to-one, one-to-many, many-to-one, etc. The complex correspondence between Zi and Ci leads to difficulty in learning Chinese characters and Chinese. In order to solve the problem mentioned above, this paper proposes to take Chinese characters as the starting point. We will take the theory of Function and Usage of Chinese Characters as a guide and take the Chinese characters and words listed in the Grade Standards as the sorting object, to propose specific methods for constructing a Zi and Ci relationship resource database. We expect to construct a resource database that integrates glyphs, structures, and the functions of recording words, provides reliable Chinese teaching resources for international Chinese teachers, and helps Chinese learners to learn Chinese characters and vocabulary in depth.

Keywords *Grading Standards*; The relationship between Zi and Ci; The construction of resource database

《国际中文教育中文水平等级标准》(下文简称《等级标准》)是中华人民共和国教育部国家语言文字工作委员会于2021年3月24日发布的最新标准,为国际中文教育的学习、教学、测试与评估设立了标准、提供了参考。该标准具有创新性,重新划分了等级分类,分为“三等九级”,可以更准确地判断学习者的汉语水平。该标准也具有前瞻性,在原有六级基础上添加了七一九级,为汉语高级水平的学习者提供了新的学习内容。

鉴于《等级标准》新增大量汉字与词语,为减轻学习者负担和方便教师教学,有必要对零落散乱的字词关系进行统一梳理,建立字与词的有机联系,创建一个字词关系清晰的资源库。

一、字词关系资源库建设的必要性



其一,新增汉字量、词汇量较大,提高了学习要求。《等级标准》具有全面性,不止新增了七至九级的语言要素,也对前六级的内容做了调整,特别是汉字和词汇,数量增加较多。《等级标准》一至九级要求掌握的汉字总量为3000个,比《HSK考试大纲》增加了337个;在不同级别需掌握的词汇量上,一级词汇增加350个,二级增加972个,三级增加1645个,四级增加2045个,五级增加1816个,六级增加456个,到七至九级需掌握11092个词汇,词汇总量比《HSK考试大纲》增加了6092个。

国家语委发布的《汉语国际教育用音节汉字词汇等级划分》(中国国家对外汉语教学领导小组办公室,2010)(下文简称《等级划分》)“是汉语国际教育的第一个国家标准,设立最低入门等级,普及化等级为先导的等级划分框架体系,为汉语教学和测试提供参考依据。”(李方延,2011)《等级标准》中规定的字词量与作为“国家标准”的《等级划分》中规定的字词量一致,说明我们越来越重视中文教育,对学习者的中文学习提出了更高的要求。

学习者要想达到更高的水平必须学习更多的汉字和词语。如何让学生在有限的时间内掌握更多的字和词,这是需要汉语教学及研究者思考的问题。

其二, 字词关系是汉语教学中具有一定难度的教学内容, 对其掌握程度决定着汉语学习者的习得水平。

汉字是记录汉语的符号系统, 而现代汉字最基本的职能就是词中的语素。同一个汉字可以记录不同的语素, 这些语素义之间有何关联, 涉及到词与词的关系, 语素与词的关系。

因此, 理解这个字在作为语素参与构词时的职能非常重要。然而, 字、词的历时演变导致二者在现代汉语中多已丧失其固有的联系, 理据不明晰。字、词之间似乎成了两个孤立的系统, 字与字之间, 甚至词与词之间也难以直接显示其联系。如“敌”, 简体已看不出其声符的表音作用, 已经记号化。如果还原成繁体, 写作“敵”。表示敌人意思的{敌}一词在金文中写作, 读 dí, 由于字符本身还有其他职能, 表示仅、只, 故《说文解字》中小篆分化出“𡗗”字, 用来专门表示{敌}。后经过隶变, 字符写作“敵”。由于部件“商”是由“啻”演变而来, 因此凡是含有部件“商”的形声字多从帝声, 如“滴、嘀、鎬、嫡、適”。而记录“敵”义的字符进一步简化成“敌”, 从而丧失了该词语音与字符之间的显性联系。所以有必要对现代汉语中基础要素的字与词进行梳理, 以构建起有机的字词网络, 帮助海内外汉语教学者进一步掌握基础的汉语字词知识。

由于汉字使用职能历时演变的复杂性和曲折性, 使得汉字的习得对汉语二语学习者来说成了“老大难”问题。汉字是用来记录汉语的书写符号, 通常情况下, 一个汉字记录一个语素。而词汇中大多词语是采用词根构词法, 即由不同语素构成。随着社会的发展, 词语的变化速度较快, 相较而言, 作为书写符号的汉字系统的发展变化却相对较为稳定, 这使得我们总是在用相对有限的汉字去记录无限生成的词, 二者之间的关系就会变得日趋复杂, 可以是一字一词, 也可以是一字多词, 或是一词多字。而现代汉语中, 很多字词之间的关系变得不够明晰, 甚至看不出有任何联系。如“天然气”为什么不是“天燃气”? “然”与“燃”有什么关系? “方便”的“便”为什么又可以用在“大便”这一词中。要弄清这些问题, 必须要对字词关系进行梳理。

其三, 从汉语学习者习得的实际情况来看, 不能准确把握字词关系是二语学习者产生偏误的重要原因。

如同字异音的问题, 学生往往先入为主, 先习得哪个读音, 就会在所有词中读这个音。“音乐”和“快乐”, 汉字“乐”在二词中读音不同。再如“银行”读成 hěnxíng “很行”, 这不仅是对同声符形声字认识的模糊, 也是对多音字认识的不全面。根据对外国学生汉字习得偏误已有的调查研究, 汉字习得的偏误主要存在以下三种: ①同音字导致汉字使用的偏误, 如“在”写成“再”, “纹”写成“问”; ②形近字导致汉字使用的偏误, 如“你也可以现左(在)去学校找他。”(陈勇, 2016); ③不理解部件的意义而导致汉字使用偏误, 如“哭”的下半部写成“大”, “睡”的左边写成“日”等(胡勇, 2013)。不论是形体书写上或是汉字使用职能的偏误, 归根结底, 造成这些偏误的原因, 实则就是将字与词割裂开来, 没有理清字与词之间的关系, 完全不明白字在不同词语中的使用职能, 或是将字的职用混淆, 或是囫圇习得, 或是死记硬背, 完全不懂汉字的字理。加之, 《等级标准》中字、词的安排顺序也并非完全按照这个字用法在历时演变中出现的先后顺序编排的。特别是对于音异且意义有联系的多音多义字, 如“差”, 首先在一级词汇中出现的“差”, 读音为 chà, 读音相同的“差不多”是出现在二级中的词汇; 而读“chā”的词

“差别”则在五级词汇中才出现。意思是学生可能先习得的是 chā, 而后习得 chā。但考察“差”的历时演变, 会发现, 读音 chā, 实则早于 chā。而学生先学了“chā”, 可能在后面遇到“差”时都会产生先入为主的偏见, 都读作“chā”。并且, 在《等级标准》的词汇表中, 我们统计了“差”参与组合的词语总计16个, 而读“chā”的词语就有12个, 占绝对优势。实则, 字符“差”在历史上使用时间更早的“chā”在《等级标准》中构词能力也强于使用时间较晚的“chā”, 建议《等级标准》在编排词汇顺序时可以将字词关系梳理中反映出来的构词能力作为考量因素。所以, 对《等级标准》字词关系进行梳理也便于指导调整教学顺序, 可以先出现 chā, 再出现构词能力弱的读音 chā。

其四, 目前汉字的二语教学及研究陷入困境, 这一瓶颈亟待突破。汉字、词汇是汉语学习的基础。但是一直以来, 专门的汉字课多开设在初级阶段, 以汉字认读和书写为教学重点。到中、高级阶段, 专门的汉字教学课减少, 甚至忽略汉字的教学, 或者说认为汉字教学已无甚可教。并且, 目前疫情背景下, 线上教学与线上考试的普遍推行, 更是大大地削弱了汉字教学的重要性。而曾经作为教学重点的汉字书写对二语学习者来说几乎是不太必要的, 因此汉字教学及其研究陷入困境。从海外汉语教学研究中, 我们也可以看到, 非目的语环境下汉字教学的状况更是堪忧。信息时代和疫情背景下, 该如何打破汉字教学面临的困局, 如何设置汉字课程、确定汉字教学的重点, 这些都值得我们深思。李运富先生的汉字三平面理论, 尤其是其中的汉字职用理论可以为我们在汉字教学上提供新的思考方向, 也让我们进一步反思当下学生学习汉语时面临的最大的问题应如何去解决。

其五, 已建汉语语言要素相关资源多为显性的教学资源, 不能完全满足教学的需求, 需对隐性的知识与能力类资源多加挖掘。目前, 涉及到汉语语音、词汇、汉字、语法等要素教学的文字材料或网络多媒体等基本形态的显性资源较为丰富, 但隐于教学中所需的隐性的知识与能力资源的建设还较匮乏, 而这些知识与能力也是国际中文教师所必须要具备的基本素养。需要加强专门的关于语言要素本体知识系统的资源建设。

其六, 国内关于字词关系理论的研究已渐趋成熟, 可为面向国际中文教育的字词资源库的建设提供理论指导。近年, 学者们在研究汉字历时演变的过程中涉及到字记录词的演变情况, 尤其是对于古文字的字词关系研究成果较为丰富, 在此基础上, 李运富先生提出的汉字学“三平面理论”, 将汉字分成形体、结构、职用三大维度, 其中职用理论即涉及到字记录什么词以及词用什么字来表达的问题, 也就是讨论的字词关系。这一理论的提出为国际中文教育的汉字教学指引了方向, 也提供了教学的思路。汉字教学虽已从只重视形体、结构的教学到逐步开始重视字词关系的教学, 但针对国际汉语的字词关系教学的系统性还不够, 教师也缺乏可以利用的专门针对国际汉语教学的全面系统的字词关系资源。同时, 现代汉字教学也存在一些问题, 首先, 在广泛应用的“随文识字”的教学中, 汉字教学依附于词汇学习, 并不进行专门的汉字教学, 导致学生学习汉语时字词分离, 如“康”在“健康”一词中认识, 可能到了“康复”中就不认识了。其次, 对于现代汉语来说, 字与词之间关系不甚明晰, 学生难以建立二者之间的联系。对于一些多音同义字, 如血, 有“xiě”和“xuè”两读, 但二者有其使用环境的区别, 前者多用于口语中表示单音

节词，如“流了点儿血”“鸡血”等；而后者多用于表示书面语中双音节词的语素，如“流血”、“血压”等，但实际上有的汉语教师也是混淆不清的。这就需要加强现代汉字中字词关系的整理与研究，为二语学习者和国际中文教师提供便利的可利用的教学资源，助力学生深度学习和教师日常教学。

综上，字词关系资源库的建设对国际中文教学来说十分必要。不但可以丰富汉语语言要素教学甚至其他相关课程的教学资源，提升教学质量，如果对汉语字词关系掌握比较牢固，在教授汉字、词汇方面都可以充分利用汉字的字理知识进行教学。尤其是在新的时代背景下，国际汉语教学亟待对国际中文教育教学中的核心内容汉语语言要素进行课程资源的开发与建设，尤其是针对《等级标准》中涉及的字、词进行有针对性的系统的资源建设。

二、《等级标准》字词关系对应的基本类型

此处所指的字词关系对应，并非是指一个汉字对应某个个体的词语，而是指汉字与对应汉语的语言单位之间的关系，现代汉语中大多表现为一个汉字对应一个语素，对应表达的就是语素义，古代汉语以单音节词为主，汉字大多对应表达的就是一个词，本文以现代汉字和现代汉语为出发点，但在讨论个别字符与所记录词音、词义历时演变时，也会涉及到古文字及古汉语词。而且，“字词关系”这一术语也是目前学界在讨论汉字与汉语对应关系时的一种习惯性表达，故本文依然沿用这一术语。李运富先生将一项带有词音的意义称为“词项”，不同的词义就是不同的词项。同时，他也认为字符单位跟语符单位的对应关系不能简单地都看作是“字”与“词”的关系，也可能是“字”与“词音”的关系，也可能是“字”与“词项”“词位”甚至“词族”的关系（李运富，2020）。本文将借用“词项”来表示一项词义（语素义）。此外，下文所用例字均在括号内注明了该字《等级标准》所在等级，该字作为语素在现代汉语词汇中的释义综合参考了《现代汉语词典》（第7版）和《新华字典》（第12版）概括而来。

根据对《等级标准》中字词对应关系初步的整理，大致存在以下三种对应情况：

（一）一字对应一词

如以下例子：

蛇（五级）：{蛇 1}¹画蛇添足。指一种陆地上无脚的爬行动物。

丘（七至九级）：{丘 1}丘陵。指小土山。

拳（七至九级）：{拳 1}拳头、太极拳，指紧握的手。

可见，以上三字在《等级标准》的不同词中表示的词项是单一的，皆是一字一词对应的关系。从汉字使用职能的角度来看，皆是本用。

但也有一类情况，文字造字之初的本义早已消失，该字的本用不复存在，而是一直被借做表示其他的词，在文献中只保留了借用的职能。如“我”，本义是表示一种武器，但早在甲骨文时

¹ 符号{字 1}表示该字的第一个词项，有几项词义就用相应的数字标示。

代就一直借来表示第一人称,并且沿用至今。因此,“我”也就只保留了一个义项,该字与词的关系也是一一对应的关系。

(二) 一字对应多词

一个汉字可以用来表示多个词项。


1. 同音同形,意义有联系。

取(二级):

{取1}领取、取款、窃取、索取、取经,表示拿;

{取2}取消、取缔,表示消除;

{取3}采取、吸取,表示采用。

“取”,甲骨文作 (《合集》296),象手持耳朵之形,《说文解字》:“捕取也。从又从耳”。由该字形可直接关联出的几个相关的义项,从施事一方,可以表示“拿取”“夺取”之义,而从受事一方,可以表示失去、去除之义。所以二者皆可视为本用。在表示“拿取”的基础上又引申出“采用”的意思。“取”在《等级标准》中参与构成的词语有24个,但词项是有限的。如“取款、领取”等词语中都同属一个词义,表示拿;在“取消、取缔”中表示消除;在“采取、吸取”中都表示采用。这些词项在意义上都有联系,并且使用了同一字来表示,读音也相同。这些词项皆是从本义引申而来的引申义,直接与字形构意相关的{取1}{取2}皆可视为本用,而由{取1}引申而来的{取3}可视为兼用。

2. 同音同形,意义无联系。


如以下例子:

花(一级):

{花1}花园、花瓶、开花、梅花、奇花异草,表示植物的繁殖器官或有这种器官的植物,是名词。

{花2}五花八门,眼花,表示杂乱的,模糊不清的,是形容词。


{花3}花钱、零花钱,表示消费,支出,是动词。

“花”,西周金文作 (《字形表》236页),与“华(華)”使用同一字符,本义为一棵树上繁花盛开的样子。由此字形可直接关联出“表示植物繁殖器官或某种植物”的义项,此义项为本用。植物的花色彩繁多,形态各异,由此引申出杂乱、模糊不清的义项。但是“花”作动词,表示“消费、支出”的义项时,与前面两个义项毫无关联,只是借用其字形和字音。“花”在《等级标准》中参与构成的词语有22个,涵盖了以上列举的三个义项。如“花园、开花、鲜花、桃花”等词语中的“花”都属于同一个词义,表示植物的繁殖器官或有这种器官的植物;在“五花八门”中表示杂乱的;在“零花钱”中表示消费、支出。可以通过字形直接解释出来的{花1}是本用,而{花2}是从{花1}引申而来,可视为兼用。{花3}与前面的意义都没有联系,是借用来表示动作意义的消费。

升(三级):

{升 1} 毫升，表示量器升斗，也用于表示容量单位，作名词和量词。

{升 2} 升高、升级、提升，表示向上，提高，是动词。

“升”，甲骨文作（《合集》36123），本义为一带柄的量具。由该字形得出表示量器升斗的义项，后引申为表示容量单位，因此{升 1}为本义，可视为本用。而{升 2}是假借“升”字来表示提高，向上的意思。因“升”职能过多，历史上后又另造新字，用“昇、𡵓”表示向上之意。但在汉字简化时，除了人名、地名保留这一写法外，分化字“昇、𡵓”又简化为原来的“升”，因此升又身兼数职，本用、借用同时存在。《等级标准》中“升”参与构成的词语有 13 个，但表示的义项有限。如“毫升”中的“升”表示容量单位，在“升值、飙升、直升机”等词语中表示向上，提高。这两类词语中的升只是字音、字形相同，意义之间没有联系。因为{升 1}与字形直接关联，可视为本用；{升 2}则是借来表示其他的义项，为借用。






3. 异音同形，意义有联系。

如以下例子：

朝（三级）：

{朝 1} zhāo——朝气蓬勃、朝三暮四、朝夕相处，表示早上。

{朝 2} cháo——朝着，表示向着；朝代，表示某一帝王世代的统治时期。


“朝”，甲骨文作（《合集》33130），由“日、月、艸”组成，字形为残月当空，太阳隐于草丛中。表示太阳出来而月亮还未退去的时候，意为早上，读作 zhāo。金文中因常借用来表示早上的潮水，因此字形发生改变，“月”旁变为表示水流的部件“𠂔”，金文写作（《集成》2837）、（《集成》9901），隶书定作淖。战国文字延续了金文的写法，但右边又改成了表示舟的部件，写作（《集成》11182），小篆写作。隶书以后，右边又写作“月”。字符“朝”既用来表示“早上”，又兼用来表示“潮汐”。西汉以后“潮”“淖”二字合流，义为朝夕、朝见，但是“cháo”的读音保留了下来。清代《隶辨》后以“潮”专门表示潮汐义。由此可见，朝的本义为早上。古代官员早上朝见天子以议政事，朝见时需面向天子上奏政事，由此引申出“会见、向着”之义，同时也以另一读音 cháo 来表示；听政是皇帝在固定的时间和处所进行的活动，由此引申出表示某一帝王的统治时期之义。“朝”在《等级标准》中参与构成的词有 6 个，在“朝夕相处、朝三暮四、朝气蓬勃、有朝一日”中读作“zhāo”，表示时间；在“朝着”中表示向着，在“朝代”中表示某一帝王的统治时期，读作“cháo”。这两类词语中的“朝”虽然读音不同，但字形相同，且字义之间有联系，因此{朝 1}表示本义，为本用；{朝 2}是由{朝 1}引申而来的，可视为兼用。

恶（四级）：

{恶 1} è, 恶劣、凶恶，表示很坏的、令人厌恶的行为。

{恶 2} wù, 可恶，表示讨厌、憎恨。

{恶 3} ě, 恶心，表示有呕吐的感觉或厌恶某人某事。

“恶”，小篆写作，从“心”“亚”声。“亚”像墓葬形，一般指墓葬的四角，“亚”在此不仅表音也带有一定的意义，因此“恶”的意义含有恶象，本义为罪过、罪恶。恶意的行为会遭人

唾弃、厌恶,由此引申出讨厌、憎恨的义项;一旦厌恶从心理转到生理,就会产生想要呕吐的感觉。“恶”在《等级标准》中参与构成的词语共有11个,在“罪恶、丑恶、恶化”等词中表示坏的、不好的意义;在“可恶”中表示讨厌、憎恨的意义;在“恶心”中表示生理上的呕吐感。这三个义项虽然读音不同,但是各个义项之间都存在联系,{恶 2}和{恶 3}都是由{恶 1}引申而来的。{恶 1}直接与字形构义相关,为本用;{恶 2}{恶 3}可视为兼用。


4. 异音同形,意义无联系。

如以下例子:

省(二级):

{省 1} xǐng, 反省, 表示自我检查、审视。



{省 2} shěng, 节省、省钱、省略, 表示节省。

“省”, 甲骨文作 (合集 9641), 甲骨文从目, 生省声, 本义为视察、察看。由本义引申出反省义, 读作 xǐng。后音变为“shěng”, 表示减省、区划、姓氏等义项。“省”在《等级标准》中参与构成的词语共有5个, 在“省钱、省事、节省、省略”中表示减省义, 在“反省”中表示自我审视的意义。这两个义项只是用同一个字形表示, 读音不同且二者之间没有联系,{省 1}由本义引申而来, 可视为兼用; {省 2}与本义之间无联系, 为借用。

斗(四级):

{斗 1} dòu, 泰斗, 表示北斗。

{斗 2} dòu, 战斗、斗争、搏斗, 表示对打、斗争。

“斗”, 甲骨文作 (《合集》21348), 由字形可看出其本义为挹酒水的勺子。本义直接关联出表示北斗的义项。表示{斗 2}的字原本写作 (《合集》152 正), 两人互相撕扯头发打斗, 应为鬥字。后来实行汉字简化方案, 借用已有同音字形“斗”表示打斗这一义项。“斗”在《等级标准》中参与构建的词语共有7个, 在“斗志、搏斗、艰苦奋斗”等词汇中表示打斗义; 在“泰斗”中表示北斗义。这两个义项之间没有联系, 后世为了认读书写方便选择了笔画少的“斗”兼具这两个意义。因为{斗 1}是从本义中扩展出来的, 可看为本用; {斗 2}假借“斗”的字形表示与本义无关的意义, 为借用。

(三) 多字对应一词

多字对应一词的情况较为复杂。一种就是本字与新造分化字的词义在现代汉语中同时保留下来。由于本字的使用范围越来越广, 本字的职能过大, 于是新造一字分化其本义的职能, 文字上虽已分化, 但所构成的词不易消失, 本字在个别词语中仍然保留了其本义的职能, 并一直得以延续。所以就造成了我们在共时的材料中还时不时发现本字本用的情况。

《等级标准》虽然规范了多字记录一词的异体字的情况, 然而存在本字与新造字, 只不过本字与新造字在使用职能上已经剥离, 新造字分化出了本字的功能, 而本字则用来另表其他词项, 通常情况下不会同时使用, 但可能在某些词中本字还保留了其本义的用法, 偶有本字和新造字同时表示本义的情况。如表示“燃烧”意义的本字“然”, 后另造字“燃”, “然”就用来表示其他

意义。但在词语“天然气”中仍然保留了“然”的本义“燃烧”。所以《等级标准》中“然”和“燃”二字都用来表示“燃烧”这一词义。而从职用上来看，二者在表示“燃烧”这一义项时，皆是本用。

在字词关系整理中，不仅要整理字与词的关系，同时也要关注到字与字的关系。如表示“水多”意义的“益”与“溢”的关系，就是本字与新造字的关系。

多字对应一词，还有一种情况，就是同音义近的字。如同样表示“大”这一义项的，对应的汉字有宏、洪、弘，它们都读作 *hóng*，那么三者是什么关系？为什么都可以表示这一意义，而且经常出现混用的情况。如“弘扬”，也可写作“宏扬”。“弘”本义指拉弓发出的巨大声响，后引申为大。《说文解字注》中明确指出：“经传多段此篆为宏大字”。而“宏”，据《汉字字源》解释：“金文中‘宏’字不从‘宀’，而与‘函’字相近。从字形上看，应该是藏放弓的皮囊。小篆以后从‘宀’。本义早已不存，古籍中多用于‘广大’、‘宽博’、‘发扬’等义。”所以《说文解字》解释为：“屋深响也。从宀宏声。户萌切。”可见，“弘”与“宏”是被借字与借字的关系，表示广大的“宏”属于借用。而“洪”本义表示大水，后也可引申表示大，如“洪大、洪钟”等。而“洪大”与“宏大”虽皆可表示大，但二者所指对象有所区别，前者指声音大，而后者指规模、场面大，或者是抽象的主观表现。在使用上略有区别。因此，表示“大”这一义项，洪是本用，而宏是借用，弘则是兼用，它们的职用不同。

综合以上三类字词对应情况来看，无论是哪种对应，涉及到汉字的问题，实则表现为汉字的使用职能上的变化，即汉字的职用在不断地做出调整，所以要整理字词关系，其基础就是要理清每个汉字在不同词中的使用职能。

三、关于汉字职用理论

所谓字词对应关系，无外乎三种情况：一字一词，一字多词，多字一词。无论哪种对应情况，都是以字为基础，是字在作为语素后参与构词形成的复杂关系。针对这种复杂的字词关系，2012年李运富先生出版《汉字学新论》，从形体、结构、职能三维角度考察汉字，从字的角度提出汉字的三种使用职能，即汉字的本用、兼用、借用。本文将借鉴职用理论对《等级标准》的字词关系进行梳理。

汉字是用来记录汉语的书写符号系统，人们视觉上能看到的物质外壳是字形，而字形用来表达的读音和意义就赋予了它音和义，所以汉字也就不是一个平面结构，而是一个三维立体结构。

由于字符的数量是有限的，而汉语中词汇的数量在不断递增，有限的字符去表达无限多的词义，势必会导致一个字符在使用职能上变得不再单一。如“门”，本是表示房屋的门，这是本用；而进入不同的门后，就分属于不同的类，“门”又引申出了表示不同门类的意义，“分门别类”中的“门”的意义是由表示房屋的门的直接引申而来，故也可视为本义。但其实义也可继续虚化，用来表示某一类别的量词，如“一门学科，一门大炮”等，这是由表示不同类别的意义再引申而来，属于间接引申。在我们现实生活中，“门”的使用还在继续虚化，甚至有逐渐由词根向词缀

发展的趋势。吴恩锋指出“美国共和党政府在 1972 年总统竞选活动中暴露的政治丑闻‘水门事件’使得汉字‘门’受英语的影响而产生了‘丑闻’这一含义。”(吴恩锋, 2007)现实生活中一些不好的事件往往被称为“XX 门”,如“艳照门”“合约门”“移动门”等,“门”的功能和意义逐渐虚化,成为了一个类词缀。“门”虚化以后做类词缀,抑或是做量词,都属于兼用。而在“所罗门群岛”这样的译名中,“门”只是音同,与汉语中本来的“门”并无意义联系,所以是借用。

像这样根据汉字使用职能的变化分析出汉字的本用、兼用和借用就是汉字职用理论(李运富, 2012)。以下对三种职用的概念作简要介绍:

1.本用:指用来表示字形构造能直接呈现出来的意义或密切相关的意义。李运富先生认为引申义和本义皆可视为广义的本用(李运富, 2021),但这样的范围太宽泛,起不到细化分类理解的作用。所以根据对现代汉字整理的需要,我们将与字形构造直接相关的意义才界定为本用,间接引申的意义视为兼用。

例如:

百(一级),本义表示计数单位,即 100。词语“百分点”中的“百”就是表示这一意义,是本用。

益(四级),本义表示水多,后另造新字“溢”表示水多。该字又由本义引申表示“多”,如词语“日益”。

2.兼用:指由本义派生出来表示其他意义的现象。表示为一个字既可用来表示本义,也可用来表示其他派生出来的意义。

可以用来表示派生出与本义相关的意义,读音不变;也可以表示与本义有联系,而读音发生了变化的词。

前者如:百(一级),在“百货”“千方百计”等词语中表示多样的,多种的,从表示具体数量的数词引申而来,是兼用。

从(一级),在“跟从、随从、从众”等词语中表示本义,动词,是本用;但“自从、从早到晚”中“从”就虚化为介词了,读音不变,但意义已经虚化,表示另一个词,视为兼用。

后者如:少(一级),本义表示不多,读作“shǎo”,“少数、多少、少量”中“少”都是本用;而“年少、少年、老少”中“少”读“shào”,表示年龄不大,由本义“不多”引申而来,这就是兼用。意义有联系的多音多义字,读音不同,但大多是源词与派生词共用一字,派生词中字的使用即为兼用。

3.借用:该字形被借去记录一个与自身意义无关的音同或音近的词。包含两种情况:一种是无本字的假借。例如莫(七至九级),借用本表示黄昏义的字形来表示否定义“没有、不要”的副词,如“莫要”“莫吃”。另一种是有本字的假借。如现代汉语中的“干燥”一词,其中“干(一级)”是“乾”的简化字,在简化时借用音同的“干”这一字符来表示干燥之义。

四、资源库构建的目标、思路、方法及主要内容

（一）资源库构建的目标

资源库拟以汉字为出发点，建立字与词之间的联系，字与字之间的联系，创建一个集汉字形体、结构、记录词语职能于一体的完备的字词关系资料库。试图解决以下问题：1.解决同一个字形表示多个词的问题；2.解决词的本义、引申义无序的问题；3.有共同形符和声符与意义、语音之间的联系问题。

（二）资源库构建的思路

资源库建设的具体思路：整体遵循“一明三定”的思路。以字为纲，系联各等级的词语，考察字在各词中不同的使用职能。首先将《等级标准》中的汉字按其现代常用义进行基础分类，常用义分类可参考《现代汉语分类词典》（苏新春，2013），便于后期对同一义类的汉字进行整理说明；然后查阅《字源》（李学勤，2012）《说文新证》（季旭昇，2010）等工具书考察汉字造字之初的形体，同时结合最新的古文字考释成果，以确定汉字的本义，这是整个资源库建设中最重要也是难度最大的一环，必须要做到绝对精确、有理有据。接下来，结合汉字字形的演变和对现代汉字结构的分析，确定该字在《等级标准》中对应词语中所表示语素的意义，明确汉字在实际使用中的意义，同时梳理该字的本义及其其他义项的关系；最后将汉字作为语素使用时的意义与汉字的本义进行对比分析，确定汉字在词语中的使用职能。

如图1所示：

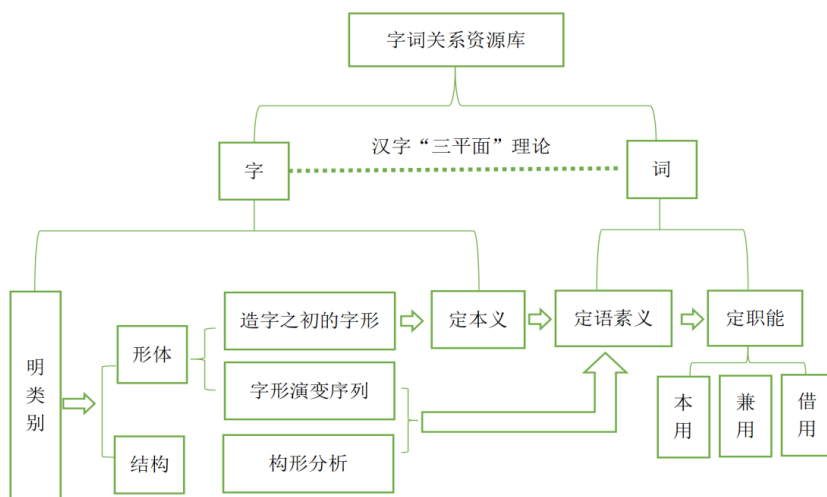


图1 资源库建设具体思路

（三）采取的方法

1.字形溯源法

利用现有出土文献和传世文献资料库，为《等级标准》中的每个现代汉字对应找到其造字之初的字形，通过字形分析及文例证实其字形构意，从而分析该字的本义。

2. 构形分析法

将《等级标准》中的所有汉字进行构形分析,分析其现代汉字的结构特征。

3. 语义对比法

通过本义与该字在《等级标准》中的构词中的语素义进行对比,对该字在该词中的使用职能进行分析归类,按照李运富先生的“汉字职用学”理论,分为本用、兼用、借用三类,明确字在词中的使用职能。

(四) 资源库构建的主要内容

基于《等级标准》,对面向国际中文教育中语言要素最基本的字、词及相互间关系进行详尽系统地梳理,创建《等级标准》字词关系资源库。

本资源库实则包含三大数据库:汉字形体数据库、汉字构件分析数据库、汉字使用职能数据库。

1. 汉字形体数据库









汉字形体是分析字词关系中最为基础的要素。现代汉字是简化字,而简化字是经由各阶段汉字形体发展演变最后整理规范而来的,大多简化字已符号化,很难直接从字形上去判断该字最初造字时的意义,也就很难理解其本义及现代常用义的由来。要弄清楚这些问题,必须追本溯源,了解其形体演变的轨迹。如“沉(四级)”,甲骨文作 (《合集》780)、 (《合集》484),像牛或羊浸没水中,这是古代的祭礼,后世多用其“沉没”的意思,读 chén。金文作 (《集成》4330),改为从水、允 yǐn 声。《说文》小篆作,隶书作)→金文()→小篆()→隶书()→楷书(沉)。因此,本数据库以现代简化字为字头,并附上各简化字的汉语拼音,按照《等级标准》各等级汉字顺序排列,展示字形演变图。字形来源主要参考工具书为《字源》和《说文新证》。具体建设内容如下表:

表1 汉字形体数据库建设内容简介

序号	主要内容	建设描述
1	序号	与《等级标准》中汉字总表的序号保持一致。
2	简化字形	以现代汉字简化字字形为字头,出现一个简化字字形记录多个意义之间无关联的语素或词时,则分别记作字1、字2,依次类推。
3	繁体字形	在对应的简化字后同时展示繁体字形,以便分析其字形构意。
4	字形演变图	根据现有楷体简化字,追溯每个字的形体演变,以绘制各字的字形演变图,以便于厘清字形来源,同时帮助理解字形构意及其本义。
5	字形演变动态视频	收集可靠正确的已有字形演变动态视频,随时补充更新。

2. 汉字构件分析数据库

对汉字形体演变有了系统全面的认识还远远不够，对于分析字形、字义，分析其构件是必须要完成的重要环节。汉字形体的演变，其很重要的变化除了笔画的变化外，构件的变化也是非常重要的，而构件的分析能让我们对汉字整字有更明晰的认识。构件是衔接笔画与整字的中间层级，构件构成汉字的数量庞大，无疑给汉语二语学习者带来困难。而汉字的构件数量却相对有限，能理解这些构件的基本功能，也就能快速地掌握更多的汉字。所以，构件分析数据库不仅要现代汉字逐字拆解为构件，分析其组合关系及层次，还须对拆解出来的构件进行功能分析，构件功能分析时重点标注出有共同形符和声符与意义、语音之间的联系。对失去理据的构件、隐性状态的构件或构件关系重组的，加以重点说明。如“残（六级）”，可直接拆分为两个构件：歹+戈，而此处的构件1“歹”实际应是“𠂔（è）”，表示骨架，表形功能，象骨架之形；构件2“戈”，通常认为表音，但通过分析字形演变可以发现，“戈”实则是“𠂔”，表示两戈相向，会残害义，也是“残”的初文，可见构件“戈”在合体字“残”中，既表音又表义，并非只表音，所以在功能分析时需要重点说明这一特殊情况。构件分析以王宁先生（2021）的“汉字构形学”为理论基础，同时结合字形演变图，展示每个汉字的构件层次及其功能剖析图，便于教学取用。具体建设内容如表2所示。

表2 汉字构件分析数据库建设内容简介

序号	主要内容	建设描述
1	序号	与《等级标准》中汉字表的序号保持一致。
2	简化字形	以现代汉字简化字字形为字头，出现一个简化字字形记录多个意义之间无关联的语素或词时，则分别记作字1、字2，依次类推。
3	结构分析	对该字是独体还是合体字进行准确判断。
4	构件拆分	将每个简化字进行有理据的构件拆分，明确其结构次序，是平面结构、层次结构还是综合结构。
5	构件功能	对拆分出来的每个构件的功能进行判断。对丧失理据的构件、构件理据不明显或构件关系重组的加以重点说明。
6	构件层次及功能剖析图	基于上述内容的建设，最终为每个汉字描绘一个完整的构件层次及功能剖析图。

3. 汉字使用职能数据库

汉字形体与构件分析数据库的创建最终都是为汉字使用职能分析服务的，资源库构建的终极目标是要理清字与词之间的关系，也就是汉字记录词语的职能。该数据库以字为纲，系联《等级标准》中含有该字的所有的词，并对该字在这些词中的语素义进行解释。本库重点标注该字参与构词后的本义、引申义，并且要厘清各引申义之间的关系。如“金（三级）”现代常用的意义指黄金或金属，但都不是本义。“金”的本义为青铜，故“金文”可以指铸刻在各种铜器上的文字。后引申指金属的总称。黄金的意义是非常晚才有的。同时，资源库要标注出这些字与词（语素）的对应关系，关联字与词，标出字的本用、兼用与借用，重点需要分辨清楚假借字和同形字的关系。如“干（一级）”实际上表示四个不同的词。“里（一级）”也表示两个不同的词。不厘清同形字表示不同的词，不能解释清楚一个字的不同义项间的关系。

表3 汉字使用职能数据库建设内容简介





序号	主要内容	建设描述
1	序号	与《等级标准》中汉字表的序号保持一致。
2	简化字形	以现代汉字简化字形为字头,出现一个简化字形记录多个意义之间无关联的语素或词时,则分别记作字1、字2,依次类推。
3	汉字读音	标明每个字的读音,如果存在一字多音的情况,则分别列出。
4	最早字形	找到每个汉字最早的形体,并注明出处。
5	字形构意	对最早字形进行构意分析。
6	本义	分析现代汉语中每个字的本义,明确本义是否与字形构意一致。
7	现代常用义	参考《新华字典》(第12版)所列第一条义项。
8	所属类别	根据常用义,参考苏新春《现代汉语分类词典》的分类标准为其归类,以便从意义上对现代汉字进行归类处理和理解。
9	词语	列出《等级标准》中参与组合的词语及其所在等级,便于各等级教学或相关教材编写使用。
10	语素义	参考《现代汉语词典》(第七版)解释该字在词中的语素义。
11	使用职能	参照字形构意,结合该字在各词中的语素义,进一步分析该字在不同词中的使用职能。
12	字词对应关系图	基于上述各项内容的分析整理,发现各字不同语素义之间的联系,准确判断本义、引申义、假借义之间的前后关联,进而绘制字词对应关系图

4. 三大数据库之间的互联互通

形体数据库、构件分析数据库以及职能数据库可以互联互通:由一个字形,可以查询其相应构件、所表词语及语素义;由一个构件,可以查询由这个构件所构成的所有字及该构件在这个字中的作用;由一个语素义,可以查询所有具有该语素义的字及表示该语素义的构件。

五、资源库建设的重点问题讨论

(一) 对汉字本义的考察

汉字本义的考察是汉字研究中的难点,充分借鉴古文字研究的成果。如“丁”,甲骨文作,金文作,象钉头的俯视形,古玺文作,象钉子侧视形。本义表“钉”,甲骨卜辞中经常假借表示天干“丁”。本义{钉}后来加上“钅”作“钉”。又如“升”,现代常用义是提高、向上的意思,但这个字甲骨文作,本义为一带柄的量具。由该字形得出表示量器升斗的义项,后引申为表示容量单位。而提高,向上的意思的{升}是假借用法。因“升”职能过多,后来又另造新字,用“昇、𪔐”表示向上之意。但在汉字简化时,除了人名、地名保留这一写法外,分化字“昇、𪔐”又简化为原来的“升”,因此“升”又身兼数职,本用、借用同时存在。

(二) 对汉字职能的判定

用有限的字符表达无限多的语义,势必会导致一个字符在使用职能上变得不再单一,需要准确判定这些字符表达相应语词的职能。如“门”,字形造意是表示房屋的门,这是本用;而进入

不同的门后，就分属于不同的类，“门”又引申出了表示不同门类的意义，“分门别类”中的“门”的意义是由表示房屋的门的引申而来，故可视为兼用。但其实义也可继续虚化，用来表示某一类别的量词，如“一门学科，一门大炮”等，这是由表示不同类别的意义再引申而来，属于间接引申，视为兼用。现实生活中一些不好的事件往往被称为“XX门”，如“艳照门”“合约门”“移动门”等，“门”的功能和意义逐渐虚化，成为了一个类词缀，也视为兼用。而在“所罗门群岛”这样的译名中，“门”只是音同，与汉语中本来的“门”并无意义联系，所以是借用。

（三）对多音多义字演变层次的考察

音异且意义有联系的多音多义字，演变过程较为复杂，如汉字“被”，其本义是被子。用作动词为覆盖。又引申为披在身上或穿在身上，读音变为 pī。这个意义后来写成“披”。又引申为施加，施及。进一步引申为蒙受，遭受。再引申为虚词表示被动。类似这样的字，需要重点考察。

（四）现代简化汉字与本有汉字的关系处理

简化汉字与本有汉字之间的关系可以从宏观上分为两类：一是原有汉字与简化后的汉字意义无关。如“后”是“後”的简化字：

後：遲也。《说文新证》

李学勤《字源》注明“後”的本义为行走迟后。

后：继体君也。象人之形，施令以告四方，故厂之从一口，发号者君后也。《说文新证》

“后”本义为君主，后来引申为君主之后、皇后之义（曹星辰，2019）。“后”比“後”出现得早，宋代之后出现了“后”与“後”的混用，如：

以俟后来之惠，全元元之命。（元·郝经《郝氏续後汉书》）

这里假借“后”来表示后来之义，是有本字的借用。

二是原有汉字与简化后的汉字意义相同。如：

號：老虎號叫，引申为一切號叫。《说文新证》

号：呼號，號的简体字。《说文新证》

1964年，中国文字改革委员会编印的《简化字总表》中，把“號”简化为“号”，二者意义无差别。

综合上述两种情况，在处理简化汉字的职用时，应立足于简化字本身，根据简化字的本义来确定其职用。如“后”的本义为具有最高权力的统治者，那“後”所携带的表示先后的意义就是“后”的借用；“号”的本义为大声喊叫，也就是“號”所表示的意义，至于“号”分化出的去声读音所表示的意义，则要根据其与本义的关系来确定是本用、兼用还是借用。

（五）汉字读轻声时应归为何种职用

汉字作为语素组成词后有轻声化的现象，并且轻声后与非轻声意义有别。如“地道”一词，当“道”读去声时作名词，意义跟本义一样，指道路；当“道”读轻声时，“地道”为形容词，而“道”的意义也发生了变化。由历史追溯上去，“道”在唐代用来表示行政区域，从宋代开始

与“地”连用,组成“道地”,即“地方与区域”,表示物产的地域;后引申为真正有名产地产出的,真正的、纯粹的。在明清时期“地道”出现,多在北方使用,新中国成立之后,随着普通话的普及,“地道”逐渐占据使用的主导地位(王浩垒,2016)。“地道”和“道地”意义相同,只是语素的排列次序不同,是联合型的合成词。“地道”中的“道”是由“行政区域”义转化而来的,而能构成行政区域的必然有相互贯通的多条道路,这时“道”与本义“道路”是有关联的,因此也应划为本用。轻声字只是声调受语流音变的影响变成的又轻又短的另外的调子,而不是有固定调值的声调,不能看作是分化出新的读音;但是读轻声时意义变化较为复杂,如“爸爸”中的加点字读轻声时保留了本义,“大部分”的“分”则跟本义无关。总的来说,当汉字在词汇中读轻声时,应按照汉字的意义来区分职能,而不能按照读音一概而论。

六、结 论

综上所述,资源库中字词关系的整理可分为两个层面,第一层面是字的整理,第二层面是字表示的语素参与构词的整理。字层面的整理实则包含两大步骤:明类别,定本义。字表示的语素在构词层面的整理则包含:定语素义和定职用两个步骤。

在分析汉字职能的时候,本义为最根本的立足点,抓住本义才能有序展开分析。对《等级标准》中的汉字进行职用整理是重要性、必要性和可行性三性并存的。一方面能提高教师汉字教学的效率和水平。在汉字整理的基础上制定出一套汉字职用教学的方案,使所有一线教师都能理解且易于操作,能最大化地调整汉字教学状态,使汉字教学走向一个新的阶段。另一方面,能真正满足学生的实际需求,提高汉语水平。汉字学习的重点在于会用,由于计算机的普及,很多学习者没有了书写汉字的需求,但是会认和会读仍是所有汉语学习者的必修内容。从职用角度学习汉字不仅真正贴合了学习者的学习需求,新的教学方式也使学习者学到的知识更加扎实。比如利用声旁认读字音,利用形旁区分字义,有理有据且不易混淆。最后,目前学界对汉字使用职能整理的研究还不够充分,需进一步深化落实,使资源库的建设得以发挥其价值。

注

文中涉及到的古文字字形主要来自以下工具书:

- 1.中国社会科学院历史研究所.甲骨文合集[M].北京:中华书局,1978-1982.文中简称《合集》,后附字形编号。
- 2.中国社会科学院考古研究所.殷周金文集成[M].北京:中华书局,2007.文中简称《集成》,后附字形编号。
- 3.徐中舒.汉语古文字字形表[M].北京:中华书局,2010.文中简称《字形表》,后附字形所在页码。

参考文献

- 曹星晨.浅谈“后”与“後”[J].汉字文化,2019(21):122-123.
- 陈 勇.“拼打”识字偏误分析及教学策略——基于智能电子产品的零起点汉字教学[J].汉语学习,2016(01):84-94.
- 胡 勇.留学生汉字书写偏误与原因分析[J].语文建设,2013(30):17-18.
- 季旭昇.说文新证[M].福州:福建人民出版社,2010.
- 李方延.《汉语国际教育用音节汉字词汇等级划分》出版[J].世界汉语教学,2011.25(01):119.

李学勤. 字源[M]. 天津：天津古籍出版社，2012.

李运富. 汉语字词关系研究之检讨[J]. 温州大学学报（社会科学版），2020(1):7.

李运富. 汉字的三维属性与汉字教学[M]. 南宁：广西教育出版社，2021.

教育部中外语言合作交流中心. 国际中文教育中文水平等级标准[M]. 北京：北京语言大学出版社，2021.

苏新春. 现代汉语分类词典[M]. 北京：商务印书馆，2013.

吴恩锋. 此“门”非彼“门”[J]. 语文建设，2007(09):54-55.

王浩垒. 地道·道地[J]. 语文建设，2016(01):69-70.

王 宁. 汉字构形学导论[M]. 北京：商务印书馆，2021.

中国国家对外汉语教学领导小组办公室. 汉语国际教育用音节汉字词汇等级划分[M]. 北京：北京语言大学出版社，2010.