



教育部人文社会科学百所重点研究基地
西南大学西南民族教育与心理研究中心

问学西南丛书 第2辑

张诗亚 ◎ 主编

基于字源的汉字 生成系统构建与 教学应用研究

Research on the Construction and Teaching
Application of Chinese Character Generation
System Based on Etymology

刘翔◎著



科学出版社



教育部人文社会科学百所重点研究基地
西南大学西南民族教育与心理研究中心

问学西南丛书 第2辑

张诗亚 © 主编



基于字源的汉字 生成系统构建与 教学应用研究

Research on the Construction and Teaching
Application of Chinese Character Generation
System Based on Etymology

刘翔 © 著

科学出版社

北京

内 容 简 介

汉字作为一种意音文字，字形是汉字的本体，以形见义是其本质特点，因而汉字教学的根本是汉字字形。当下的各种汉字教学方法从本质来讲大多是基于拼音的机械重复。这些方法无法利用汉字自身由具象到抽象的演变过程，无法以契合学生由具象到抽象的认知方式进行教学，导致了汉字与其发展历史、汉字与其所表述的事物的脱离。这对于学习、理解和运用汉字来讲，当然极其不利。

本书研究了汉字自身特点及其演变的规律，结合学习者学习汉字的规律，构建了基于字源的汉字生成系统的理论；在该理论的指导下，构建了字源汉字教学资源平台，生成了汉字教学需要的各种资源，在各级各类学校开展了一些教学实验，取得了不错的效果。

本书融合了基于字源的相关理论，以汉字教学实践为导向，具有较强的实践指导意义，适合汉字理论研究者、汉字教学研究和一般教学人员参阅。

图书在版编目（CIP）数据

基于字源的汉字生成系统构建与教学应用研究 / 刘翔著. —北京：科学出版社，2017.12

（问学西南丛书 / 张诗亚主编. 第2辑）

ISBN 978-7-03-053502-3

I. ①基… II. ①刘… III. ①汉字-教学研究 IV. ①H12

中国版本图书馆 CIP 数据核字（2017）第 135813 号

策划编辑：付 艳 汪旭婷

责任编辑：朱丽娜 苏利德 / 责任校对：韩 杨

责任印制：张欣秀 / 整体设计：楠竹文化

编辑部电话：010-64033934

E-mail: edu_psy@mail.sciencep.com

科学出版社出版

北京东黄城根北街 16 号

邮政编码：100717

<http://www.sciencep.com>

北京九州迅驰传媒文化有限公司印刷

科学出版社发行 各地新华书店经销

*

2017 年 12 月第 一 版 开本：720×1000 B5

2017 年 12 月第一次印刷 印张：24 1/2

字数：440 000

定价：128.00 元

（如有印装质量问题，我社负责调换）



中国西南，自司马迁《史记·西南夷列传》始，便把其作为一个多民族的地区，以黄河文化开始的中华文化多地多元并存的局面，也慢慢整合为以中原黄河文明为主线的多民族文化共同体。西南，在这一进程中一方面不断与中原文化碰撞、交融，另一方面又保有其原有的文化生态。这种既共生又各具特色的局面从古至今基本未变，只是随着中原王朝向西南扩张的高潮而呈现出差异：自西汉始，通身毒国道，西南纳入中原版图，这一进程直到东汉文翁治蜀基本形成。东汉之后，随着三国格局的形成，蜀开始经略西南。隋唐时期是西南文化与中原文化的又一次大交融，此时广义的西南包括两广地区，是柳宗元、韩愈、刘禹锡等唐代文人被贬西南的缘故，使得儒学在西南深入根植。两宋时期，西南更是以其独特的地理人文优势成为南宋的大后方，支撑着南北对峙中的南宋王朝。元明清时期，自忽必烈完成云南行省，清政府派驻藏大臣之后，整个西南在政区上完全成了中原的一个不可再分的整体。抗日战争之际，西南因陪都的地位，更使之成为世界反法西斯的东方中心，其政治作用已为国之中流砥柱，又因大量教育、文化机构的西迁，西南的文化发展便有了翻天覆地的变化。从此，西南学兴矣！

此时，西南虽然完成了政治、经济、文化与中原的一体，却因其交通的不便、地理的阻隔，以及民族支系的磅礴，而始终保

有多元的文化格局。这些不同的文明在人类整体文明进程中所处的阶段是不一致的——一些仍处在母系社会末期，另一些却已完全融入现代大都市——可以说，由母系社会一直到现代文明形成了一个阶梯状的发展系统。然而，这样的文化又必须在统一的国家里实行统一的义务教育，因此，如何根据不同民族文化的情况推行国家统一的义务教育，既要照顾文化的特质，又要考虑到学生的认知、文化的特点和背景的差异来推进义务教育，成为西南民族教育发展中的一个不可回避的问题。在这种情况下，“问学西南丛书”应时而生。该丛书力求从不同民族的文化特色，不同学生的认知特点，不同学校的校园文化资源等诸多方面考虑，寻求一条切合“因地制宜、民族发展”实际的教育现代化道路。这是一个巨大的实践，其工程量之大，面积之广，涉及的民族之多，文化类型之繁是新的挑战，这也必将使得试图通过简单照搬西方、国内发达地区的教育发展模式、理论方法来推进西南民族地区教育的发展路径遭遇阻抗。对于西方、国内发达地区的教育发展模式和理论方法，一方面要借鉴参考，另一方面又必须把它化为与当地的特点、实际相结合的教育发展之途。这一任务摆在了西南学子面前。

“问学西南”，这一问是探求，是寻访，是深入实践，是 field work（田野考察），是从活的教育实践中找到活的、具体的问题，继而求其学、求其理、求其道。如此，既结合实际对西南民族教育的实际推进、事业发展有所裨益，又对中国特色教育理论体系的建构有所帮助。更重要的是，使得我们的学生、教师能脚踏实地面向生活，深入到活的西南教育发展的实际中去亲试自己的能力，整合自己的知识，继而形成自己的研究平台、事业基础。只有如此，才能把学生的培养、教师的教学与学理的探究、研究结合在一起，这两者是一个整体，其源都在活的西南之中。

来自五湖四海的学生，其培养体系和学科背景十分狭隘。很多学生对西南最初的了解都是旅游式、观感式的肤浅认识。没有多次的深入、长久的沁入，不可能找到真实的问题，也不可能对其真正理解，更不可能成为自己学术研究的基础、沃土。这个过程对每个学子都是挑战，可以说，我们的学生都经历了这么一个并非轻松的过程。甚至，有的学生说过“开题总通不过，简直脱了几层皮”之类的话。诸如此类的话颇能反映实际。是要脱皮，脱掉从书本中来再到书本中去的做学问的“皮”；脱掉从单一学科、狭隘的视野获得的价值观，居高临下的指导，自以为优越的“支教”的“皮”。这层皮不脱，不能成为真正的研究者，也不能培养出脚踏实地的实践研究者。春蚕吐丝，是要经过若干次蜕皮的，这个过程不是简单的，会有痛苦，但是我们的坚持，我们的主张，我们的执著，我们长期以此为特色培养学生、发展队伍、开展研究……相信只要坚持下去，这一“问学西南”会有成果的，我们的学生也会实实在在地提升自己的能力。当然，我们现在做的还不够，离我们的目标还有很大的距离，但是，坚持下去一定会有收益。这一感受我们的学生都有。无论是来自西南以外地区，或是本身就在西南生长；无论是其他民族，或是西南诸多民族中的一员，其感受都是实在的、研究都是诚实的。尽管还很不够，甚至有的地方还很浅薄，但是，我们的“问学西南”开始了，第一步迈出了，我们相信只要实实在在地走下去、再走下去，就可以至千里。

是以为序。

张诗亚

壬申岁末于说乎斋序



如果以文字的发明作为一个文明的标志，那么世界上大家公认的便有五大文明。其中三种文明的文字，即距今 6000 多年的两河流域的楔形文字，距今 5500 多年到 6000 年的埃及的象形文字，以及距今 5000 年到 5500 年的恒河流域的图章文字，均已成为不再使用的死文字，这一现象引起了学界的高度重视。对楔形文字和埃及象形文字的解读，已得到了长足进步，而印度恒河流域的图章文字至今未得到解读。这些文字所象征的、代表的古文明，绝不同于现在居住在两河流域、埃及和印度恒河流域的民族文化。而另外两种文字，一种是腓尼基人的字母，被希腊人改变之后，又传到了古罗马成为拉丁语；而拉丁语又生发出了英语、法语、德语、斯拉夫语等若干世界的拼音文字。腓尼基字母虽不能说已完全死亡，但它的传承已发生极大的变异，现今的英语、法语、德语、斯拉夫语等拼音文字，绝不同于古代腓尼基字母，除了拼音相似外，其语义、字形，均发生了天翻地覆的变化。但中国的甲骨文则不同，其从发明^①一直延传至今，字形、字义、字音，均未发生巨大变化，可以说是唯一流传下来的活的文字，换言之，华夏文明也是一直没有间断的一个活的文明。探究其中的道理，便可发现文字的生命力何以如此之强。

同样是象形文字的古埃及文字，为何却未得到延续？深究一下，会有很多发人深省的东西。古埃及的象形文字，虽然象形，但几乎停

^① 一般公认为 3500 年到 4000 年前，在殷墟出土。

留在图画文字阶段，并未完成由象形向抽象过渡的演变。而中国的甲骨文则不同，它从完全具象到逐渐完成形而上的变化，这一转变打通了从具象文字向抽象文字转意的、变化的通道。

人类的思维常分为两大类：一是形象思维，二便是抽象思维。如何完成从形象到抽象，涵盖思维发展的主要类型，同时又使其能够由具象到抽象进而完成一个全环节的转换？这一转换至关重要，任何人类的符号，俱是从形象到抽象的过程。拼音因其拼、其形（除字母之本形之外），所表达的意蕴早已没有丝毫象形的含义；而甲骨文则不同，其既具象又抽象，这就使大脑的左右两半球得以均衡发展，得以连通聚合，进而形成全脑的深层思维。拼音字母的字义形象割裂、去掉以后，便只停留在大脑的左半部，这就使得常说的裂脑人、半脑人出现；这也是西方频频出现失读症的关键，而这种现象在使用汉字的人群中几乎是没的。究其原因便是拼音字母使用者能拼出声音，却不知晓其背后的意义。背后的意义，正是由拼出来的抽象能贯通意蕴的形象，而这一点恰好是拼音文字根本欠缺的。

中国的文字改革往前追溯，直到20世纪20年代、30年代、40年代，汉字要走拉丁化道路、要改革成为强势运动。而认为中国落后是因汉字造成的这一结论，显然没有对症下药或是下错了药，因此，以改革汉字来解决中国进步问题实际上是南辕北辙。还好，也幸好，汉字因使用人群众多及内在机理的合理、文化积淀的深厚，才没有被一次又一次的汉字简化、汉字拼音、汉字注音等运动而改变。类似这样的情况也发生在越南。越南、朝鲜、日本等以前都用汉字，当他们因国小、没有深厚的文化积淀，去掉汉字后，便迅速实现了拉丁化；而中国地大物博，文化积淀深厚、历史悠久，汉字背后又极具合理性，能全面发展人的思维，于是在我国，汉字没有被拉丁化运动所擅改，应该说这是幸事，无须赘述。

汉字也因笔画的繁缛、字形变化不显著等，导致学习汉字存在诸

多困难。如何应对这一挑战，使汉字便于学习，同时又能传承自己的文化、担负起将古老的华夏文明带入 21 世纪，以其特色推进民族发展便是问题的关键。

多媒体、互联网时代的到来，使这一挑战有了应对的利器。我们开展的一系列汉字语境原生境生成的实验和研究便是试图应对这一挑战的尝试。这一工作虽不能说已大见成效，但可喜的是这一尝试、这一努力已经有了一些可喜的、可资鼓舞的苗头，虽还幼小，但其长势喜人，可以预言其生命力强大，其发展势头是不可阻挡的。多媒体使其可以完成从形象到抽象的全环节转换，使其创生的形象的语境实现到抽象的升华，可以有一个换挡不脱节的形象展示过程。也就是说，可以形象地呈现生动的字义语境产生的情境，迅速的动漫变化使其实现符号化、抽象化的过渡。这一工作要由系统电脑程序编程来完成，不是一件易事。

刘翔的博士论文为从事这一研究奠定了一定的基础。刘翔开始全然不知甲骨文、不知汉字生成，后来以之作为博士论文选题，并展开研究继而完成博士论文，并形成汉字生成系统初步设计的成果，可资欣慰。这一工作不能说已经到头，应该说远未到头，还早得很；但这个头开了，这个基础奠定了，这个发展就有望了。因为这不仅仅是一个系统设计，也不仅仅是一个程序问题，这涉及了一个庞大的文字及其背后所承担的文明的继承、发展和弘扬、升华、扬弃的重任。汉字是一种极为成熟的、传承了几千年的文字，要改变大家的认知习惯，使之适合新形势下的学习，并非易事。这一工作既要完成系统本身的设计、应用的转换、链接，同时又要适应大家的习惯。要获得众口一词的赞誉，便需要实验，便需要反复提升、反复纠正。我相信只要坚持下去，这些都不是问题。

现在多媒体的应用、互联网的普及可谓日新月异，古老的汉字、悠久的华夏文明正面临这样一个千载难逢的机遇。能不能抓住它、能

不能使之成为让汉字、让华夏文明得以跃升的新的机遇，便是我们所面临的挑战。我相信在我们的努力中、在我们的失败中、在我们的探索中、在我们的搏击中、在我们的创新中、在我们的开拓中，这个过程会取得成功的结果。

此书产生不易，从博士论文奠定基础，一直到博士毕业后的数年研究，最终在大洋彼岸的美国访学过程中得以完成。难得的是，刘翔身边还有一个可爱的女儿，小含章为刘翔的访学添了不少辛劳、麻烦，但更重要的是，她常常成为这一系统的第一个品鉴者、实验者、赞助者、支持者、批评者，这一点，应使刘翔辛苦之余颇感欣慰。

当然，这一切有待检验，有待有识之士、习用汉字的广大人群，乃至全球的华人来鉴定、评价。我相信在大家的关注中，它会成长。

是以为序。

张诗亚

2016年7月6日

于旧金山伯克利



汉字教学是一个长期困扰教育界的问题。汉字教学最浅层次的目标就是达成汉字形音义的准确联结，即见形而知音义、听音而知形义。为此，汉字教学主要分为两大类：一类据形，一类据音。以“三百千”为代表的古代蒙学教材，发挥韵文朗朗上口、好读易记的特点，在较短时间内让儿童机械记忆大量汉字，并快速掌握大量关于自然、社会等基本常识。这种方法在学习内容以文言为主而相对单调的时代的确有效且可能。但在知识爆炸的近现代，却越来越不能适应时代的要求。于是，汉字教学开始了以儿童为本位、以培养听说能力为目的的改革。总的来说，虽然汉字全盘拼音化的思路最终被抛弃，但拼音作为汉字认知最重要的桥梁的地位却已成既定事实。

汉字认知从象形、象意、形声为主体转向拼音系统，表面上，通过拼音认识了汉字的读音似乎就“认识”了汉字，实现了“化繁为简”，方便了学习者学习汉字。实质上，这是一种认知方式的彻底改变。它使汉字认知由符合汉字的生成与人的认知特点的、从具象到抽象的方式，转向了按拼音规则进行的、由声而非形的抽象方式。这一转变，造成了汉字认知与汉字的本体——汉字字形的脱离，从根本上导致汉字与它所承载的文化脱离。具体来讲，就是汉字与物、汉字与历史、汉字与日常生活等关系的“三个脱离”。也正是由于这个原因，汉字教学费时多、周期长、见效慢、效果差。从本质上讲，是因为汉字教学只注重汉字的工具理性而忽视了其价值理性的结果。

如何解决这一问题,使对汉字的认知重新回到“三个联系”上来?只能从汉字本身以形为本体的结构特点出发,根据汉字的生成方式,重新建构汉字知识体系^①和汉字教学体系^②,实现符合汉字特点的汉字教学。本书正是基于此,拟探索结合拼音的汉字形音义生成^③系统的认知建构,以从根本上解决“三个脱离”的问题。

要实现这一构想,不仅需要理论研究,更重要的是要完成从理论观念(核心是认知和汉字生成理论建构)到技术实现,再到应用设计,最后通过实验设计进行验证的全环节研究。为此,本书构筑了汉字生成系统理论,完成了基于此理论的平台设计,提出了基于此理论的应用模式,并有针对性地设计和实施了相关实验,以验证该理论及其实践的科学性和有效性。

本书的工作集中在四个方面:一是汉字生成系统理论的构建;二是字源汉字资源平台的设计与实现;三是字源识字教学资源生成;四是基于字源的汉字教学应用研究。从学科领域来看,本书是一个综合多学科交叉研究,涉及教育学、心理学、计算机科学、语言学等。从教育技术学范畴来看,本书是一个教育技术领域的全环节研究。

在理论研究方面,本书着力于在整合相关理论的基础上,构建汉字生成系统理论;在应用实践方面,以相关理论研究为指导,以开发为取向,设计和实现了字源识字基础平台,并充分发挥信息技术的优势,生成用于支持字源识字的教学资源;同时,本书还将在汉字生成系统理论的指导下,提出字源识字的微观和宏观应用模式,并以多种研究方法初步论证其可行性和合理性。

就本书的两个重点而言,本书将首先在综合其他领域的相关理论

① 本书中体现为字源识字基础平台,特别是字源识字资源平台。

② 本书中体现为字源识字教学实践。

③ 这里的“生成”当然不是指汉字是由这个理论“产生”的,而是指这套理论试图反映汉字创造和发展变化的途径,并揭示构成汉字的部件与整字的关系,以“动态生成”的观点来看汉字的创造与发展演变。或者说,它试图模拟一个汉字如何产生、如何演变到现在的过程,以便使学习者以发展的观点看待汉字,而不是把它视为僵化的知识。

的基础上，构建汉字生成系统理论，并以此理论为指导，完成字源识字基础平台的设计与实现，使字源识字既有坚实的理论支撑，又有实践环节的支持。

就“字源识字教学资源的生成”这一难点而言，研究的主要工作体现在三个方面：一是如何选择适用于字源识字的材料；二是如何将这些材料数字化，使之成为教学、研究材料；三是这些教学资源应该如何为教学和研究提供支持。就“基于字源的汉字生成系统构建与教学应用研究”这一难点而言，本书将从微观和宏观两个方面研究信息技术支持下的字源识字模式，然后综合运用各种研究方法进行初步研究。

本书是综合研究，力图在字源识字综合理论的指导下，开发支持系统、生成系统资源，并将其融入教学实践，研究教师和学生如何利用系统进行教和学，以满足字源识字的微观和宏观应用模式的要求。



总序（张诗亚）	
序（张诗亚）	
前言	

第一章	汉字生成系统与汉字教学	1
第一节	当前汉字教学领域中存在的问题	1
第二节	相关理论研究	4
第三节	汉字生成系统理论的基础研究	42
第四节	汉字生成系统理论与汉字教学	47
第二章	面向字源汉字教学的汉字生成系统的设计与实现	72
第一节	汉字生成系统总体设计	72
第二节	汉字基础信息收集与整理	81
第三节	汉字字源信息的收集与整理	83
第四节	汉字生成信息设计与实现	100
第五节	字源汉字教学资源平台设计与实现	124
第六节	关于汉字字形统计特点的研究	147
第七节	平台资源的教学转换	159
第三章	学前儿童汉字教学实践	170
第一节	研究背景与研究设计	170
第二节	学前汉字教学实施	176
第三节	效果测试	197

第四节	教学反思	204
第五节	总结	212
第六节	基于绘本的学前汉字教学探索	214
第四章	汉族地区小学教学实践	236
第一节	实验研究与实施	238
第二节	结果分析	243
第三节	字源识字典型应用研究	271
第四节	字源识字应用实例	295
第五章	藏族地区小学汉字教学实践研究	312
第一节	理论探讨	312
第二节	藏族地区小学汉字教学设计	323
第三节	藏族地区小学汉字教学效果研究	331
参考文献		364
后记		374

第一章 汉字生成系统与汉字教学



第一节 当前汉字教学领域中存在的问题

从相关领域研究者的文献和笔者的亲身体验都可以看到，不管是在国内小学还是在对外汉语教学中，识字教学普遍存在诸多问题^①。

笔者从2008年10月开始，采用文献研究、与教师和学生访谈、课堂观察与分析等方法进行了识字教学方面的调研。结果显示，由于存在教师、教材、配套资源、教学管理等诸多原因，汉字教学普遍存在以下问题：教学理念比较落后，基本上只关注认知方面，即看学生是否认识和识记了汉字，而对于文化渗透和思维训练等基本上无从顾及；教学方法比较单一，基本上还是沿用数笔画、“加一加”“减一减”“换一换”、分析偏旁、编故事、做游戏这些传统的教学法；教学资源贫乏，基本限于教材及教学参考书，很少扩展。

一、教学理念方面的问题

教学理念方面的第一个问题是教学中没有科学地体现汉字自身的特征，忽视汉字形音义之间的联系，将汉字当成纯粹抽象的符号。

由于汉字是形音义的综合体，我们一般认为识字教学的任务就是要统一地掌握字的形音义，即要求认清和正确书写字形，读准字音，理解字（词）义；

^① 例如，王宁在《汉字教学的原理与各类教学方法的科学运用（上）》中针对识字教学中存在的问题指出：由于关于汉字宏观理论的研究未能及早介入汉字教学，作为表意文字的汉字所具有的规律未能引起识字教学领域研究者的关注，随着教育工作的进展，在识字教学领域也产生了几个普遍的问题，有些地方似乎进入了误区。

在大脑中建立字的形音义的牢固联系，达到看见字形，能读出字音、明了字（词）义，听到字音、知道字（词）义，能写出字形的熟练程度。我们需要注意两点：一是这种达成汉字形音义固定联系的方法是多次重复读和写，即机械地重复；二是它所形成的汉字形音义的联系是一种机械的、无意义的联系。这种汉字教学把汉字当作纯粹的抽象符号，其形音义之间并不存在任何联系。

教学理念方面的第二个问题是忽视汉字教学目的的层次性，仅注意到了汉字教学目的中的“识字”功能。据笔者的访谈，很多小学语文教师有意无意地将“识字教学”等同为“认字教学”，没有意识到识字与学生文化素质的提升、观察能力的培养、思维水平的提高之间的关系；他们更多地关注识字教学工具层面的问题，而忽视了通过识字教学让学生了解中华文化、培养学生思维能力等深层次的问题。识字教学理念的落后导致识字教学定位层次太低，其后果是长期以来形成的机械重复式的汉字教学大行其道。

识字教学理念方面的第三个问题是忽视了汉字教学的系统性，特别是汉字形义的系统性。这一点与以随文识字为主流的汉字教学方法有很大的关系。虽然很多教师在教学中也注意形近字的辨析、同声旁字的归类，但还是缺乏有意识地利用汉字形义关系，将汉字放到一个系统中去组织教学的意识。

总之，识字教学理念的落后是当前识字教学中存在诸多问题的最根本的原因。

二、教学方法方面的问题

在汉字教学方法方面，最大的问题就是采用机械教学。教师由于缺乏文字学知识，难以有效地指导学生基于汉字本身的生成规律来识字。教师往往将汉字简单地视为抽象符号，按照传统的教学经验，沿用反复读准字音、数笔画、“加一加”“减一减”等简单枯燥、缺乏理据的教学法，要求学生简单机械地朗读、听写、默写。课后布置大量反反复复的抄写与听写来巩固，不仅加重了学生的学习负担，使学生易于产生厌学心理，同时也加重了家长的负担，更重要的是这种教学方法将鲜活的汉字硬生生地变成了毫无生气的东西，使原本自然、有趣、充满探索乐趣的汉字学习变成了呆板、乏味、痛苦的重复劳动。

虽然有一些老师也在尝试字族识字、字理识字、字源识字等方法，却无从改变机械记忆法大行其道的主流。从实践来看，一线教师很少会有意识地反思