



上海市学术著作出版基金

殷商甲骨文字形系统再研究

陈婷珠 著



上海世纪出版集团



上海市学术著作出版基金

博士文库

殷商甲骨文字形系统再研究

陈婷珠 著

世纪出版集团 上海人民出版社

图书在版编目 (C I P) 数据

殷商甲骨文字形系统再研究/陈婷珠著. —上海: 上海人民出版社, 2010

ISBN 978 - 7 - 208 - 08326 - 4

I. 殷… II. 陈… III. 甲骨文—研究 IV. K877.14

中国版本图书馆 CIP 数据核字 (2008) 第 199398 号

责任编辑 曹勇庆

殷商甲骨文字形系统再研究

陈婷珠 著

世纪出版集团

上海人民出版社出版

(200001 上海福建中路 193 号 www.ewen.cc)

世纪出版集团发行中心发行

上海商务联西印刷有限公司印刷

开本 635 × 965 1/16 印张 40.5 插页 4 字数 484,000

2010 年 10 月第 1 版 2010 年 10 月第 1 次印刷

ISBN 978 - 7 - 208 - 08326 - 4/K · 1518

定价 68.00 元

获 2008 年全国“百篇”优秀博士学位论文提名奖

获 2008 年上海市优秀博士毕业论文奖

上海市教育委员会和上海市教育发展基金会 2009 年“晨光计划”项目

(项目编号: 09CG09)

上海交通大学国际教育学院创新计划青年研究项目

(课题编号: 10YQN02)

序一

黄天树

2007年，我蒙陈婷珠女士的导师刘志基教授的邀请，到华东师范大学参加了陈女士的博士学位论文答辩。我认为，作者在短短的数年时间里，对殷商甲骨文字形系统作了全面的整理和研究，不但重视新出土的甲骨材料，注意吸收学术界最新的研究成果，还有自己的一些新见解，实属不易。例如，殷墟甲骨文不重复的单字的字数究竟是多少？这是一个看似简单实际却是非常复杂的问题。作者经过研究，得出一个阶段性成果，她说：“我们最后完成了甲骨文字形数量的清理，得到的甲骨文字形数为6232个。”又如独体字的字数究竟是多少？作者经过研究之后说：“在甲骨文独体字中，甲骨文的字头数为1025个，占到甲骨文单字字头数的26.32%。这个比例比以往研究要来得高。……从历时比较可以看出，在越早期的文字中，独体字的比重越高。”上面这些数字或结论都是阶段性的研究成果。其结论“越早期的文字中独体字的比重越高”无疑是非常正确的。至于统计数字还有待今后作进一步的修正和补充（经答辩与修改，统计结果有所改动）。但是，从这些数字的背后可以看出，作者对甲骨文资料的搜集与整理是下了很大的工夫的。现在，陈婷珠女士来信告诉我，她的博士学位论文《殷商甲骨文字形系统再研究》即将出版，要我写一篇序。我想乘

此机会，谈谈我个人关于甲骨文字构造的一个新看法，即所谓“二书说”。

殷墟出土的甲骨文是我国目前所知最早的成系统的古汉字资料。它有15万片左右，总字数超过100万字。这对研究商代文字的构造提供了丰富的资料。它是了解古汉字形体构造的活化石。甲骨文字的形体构造，即许慎《说文》所说的“六书”。若以“六书”去观察甲骨文，可以说都具备了。这说明甲骨文已经是一种相当成熟的文字。商代甲骨文是古汉字的开端，秦代的小篆是古汉字的终端。通过商代甲骨文字形系统跟秦代小篆字形系统构造的比较，可以知道，其构造具有各自的特色。

汉字的结构是非常错综复杂的。古人很早就注意到汉字的构造。在传世文献中保留有很多关于说解文字的记载。如《左传》记载“止戈为武”、“反正为乏”、“皿虫为蛊”之类，说明古人很早就懂得汉字是有结构的。“六书”一语最早见于战国时所作的《周礼》。《周礼·地官·保氏》列举了周代用来教育贵族子弟的“六艺”的项目，其中有“六书”。但是，《周礼》并没有具体说明“六书”的内容。到了汉代，学者把“六书”解释为关于汉字构造的六种结构类型，即象形、指事、形声、会意、转注、假借，称为“六书”。这是非常了不起的事情。汉代对六书的诠释主要有三家。其中，以许慎的说法最为详备。许慎《说文》所收的字以秦汉的篆文为主，兼录西周的籀文和战国时六国的古文等。许慎在分析了大量的篆文（包括古文和籀文等）结构之后，在《说文·叙》中分别给六书下了定义，并举了例字。这是最早关于汉字构造的系统理论。

汉代学者提出六书说是功劳的，为学术界所普遍采用。1899年，甲骨文的发现，对原始汉字形体结构的实际了解已大大超过了许慎，不但可以纠正他对一些汉字构形分析上的具体失误，而且有助于对汉

字构造的重新认识。学者逐渐认识到“六书说”并不像以往所想的那样尽善尽美。有鉴于此，唐兰、陈梦家、裘锡圭先生先后提出了关于汉字构造的新的理论“三书说”。1934年，唐兰在《古文字学导论》一书中说：“象形、象意、形声，叫做三书，足以范围一切中国文字。”1956年，陈梦家先生在《殷虚卜辞综述》一书中认为假借字必须列为汉字的基本类型之一，象形、象意应该合并为象形，所以他的三书是象形、假借、形声。1988年，裘锡圭先生在《文字学概要》一书中认为陈氏的三书说基本上是合理的，只是象形应该改为表意（指用意符造字），所以他的三书是表意、形声和假借三类。我们认为裘锡圭先生的三书说最为合理。

近年来，我利用“六书”或“三书”来分析经过规范的《说文》小篆等的字形结构，是很切合的。但是，要用它来分析比小篆早一千多年的甲骨文的字形结构，其间难免有扞格不合之处。比如，“两声字”等字形结构就很难纳入六书或三书的框架之内。为了克服这一困难，我创用“有声字”这一名称（详黄天树：《殷墟甲骨文“有声字”的构造》，《史语所集刊》76本2分），以是否含有声符（或“表声成分”）作为分类的标准，把甲骨文分为“有声符的字”（简称“有声字”）和“无声符的字”（简称“无声字”）两大类。凡是字形结构中含有声符的字，统称为“有声字”。有了“有声字”这个名称，就可以把殷墟甲骨文中含有声符的字全部纳入其中，包括“假借字”。这样一来，我们认为：在汉字的构造中，声符是枢纽，最为关键。以声符作为分类的标准，可以把甲骨文的构造分为“无声字”和“有声字”两大类。这就是所谓“二书说”。“二书说”包括“无声字”和“有声字”两大类。“无声字”有三种构造类别：（一）象形字；（二）指事字；（三）会意字。“有声字”有五种构造类别：（一）假借字；（二）独体形声字（参看《甲骨文字释林·释具有部分表音的独体象形字》）；（三）附划因

声指事字（参看《甲骨文字释林·释古文字中附划因声指事字的一例》一文）；（四）两声字；（五）形声字。“有声字”的构造类别跟小篆相比，可以知道，（二）至（四）为《说文》小篆所淘汰，而（一）、（五）被继承下来。由此说明，汉字的构造类别因时代的不同而不断更新。

2008年3月28日于北京

序二

刘志基

上世纪七八十年代以来，随着中文信息化的突破性进展，数字化研究方法对传统的汉语言文字研究方法和思维模式提出了颠覆性的挑战。应该说，这种变革是首先发生在现代汉语领域的——基于数字化手段的数理统计、自动化分析、数据库语言等已频繁用于现代汉语和现代汉字的研究。相形之下，古文献语言文字，尤其是古文字研究，总体上说依然徘徊在数字化的大门之外。

古文字的数字化研究，或者说利用计算机来研究古文字，目前似乎每每会遭遇一些诸如“取巧”、“图省力”的评介。平心而论，这是实实在在的误解。古文字研究一旦被纳入数字化的研究模式（并非是指那种仅仅利用计算机做一些辅助性工作的研究），作为研究对象的任何一种材料都转化成一种数据，只有这种数据符合了数字化处理要求的精确性，才能取得被计算机处理的资格，进而实现数字化处理的既定目标。采用数字化研究模式的研究者，除了必须完成传统研究模式下所做的一切工作以外，还需要以数字化处理的严格标准对原始材料进行全面整理，以保证字符的唯一性和字符集对原始材料的覆盖性，继而造字并生成字体，构建数据库，运用程序处理……这一切都意味着需要付出更多的时间和精力，其中的艰辛只能是冷暖自知，很难为

非亲历者道。

种种情况表明，在古文字范围内，甲骨文将首先全方位实现研究方式的数字化转型。甲骨文从被发现到现在虽然只有百余年的时间，但由于其作为迄今可见最早的成系统的汉字材料的特殊地位，相对于其他一些更早被发现的古文字类型而言，吸引了人们更多的关注，相关的考释与研究也后来居上，进而形成了甲骨学这样一门堪称世界性的学问。古文字数字化的开发，是需要以相关考释研究为基础的。也许正是因为这种研究基础的存在，在近年来相关国际标准化组织开始启动的汉字古文字进入国际标准电脑字符集的工作中，甲骨文成为首先进行编码的古文字种类。

甲骨学研究虽然已经取得了丰硕的成果，但是一旦要向数字化转型，首先要做的就是重新审视文字系统的最基本要素——甲骨文字的文字单位及其各种类型、构成元素。因为这才是数字化处理的直接对象，是整个甲骨学在数字化模式下运作的基本载体。如果这个基本载体不够完善，不够精确，那么这种偏差就会在其他层面上被数字化无限放大，数字化的意义也将化为乌有。

2005年，华东师范大学中国文字研究与应用中心承担了教育部重点研究基地重大项目“古文字属性库与先秦汉字发展定量研究”，而本人则忝任项目的具体负责人。而此时也正是陈博士确定学位论文选题之时。作为我的博士研究生和此项目的参与者，再加上她硕士阶段相关知识的积累，她就选择了项目中的甲骨文部分作为其学位论文的研究内容。

陈婷珠博士的这部专著（也就是她的博士学位论文）《殷商甲骨文字形系统再研究》正是在这样的背景下确定选题，撰写完成的。也正因为如此，这部专著方法和目标两个层面都贯穿着数字化基调。

她的这篇博士论文得到了以甲骨学专家黄天树教授为主席的答

辩委员会的充分肯定。论文的答辩决议做出了如下评价:

陈婷珠《殷商甲骨文字形系统再研究》一文在全面掌握前人关于甲骨文字形系统研究成果的基础上,补充了材料,重新进行清理勘误,得出了新的甲骨文字形总表,统计出了甲骨文字形总数为 6 232 个(经答辩与修改,该统计结果有所改动);进而深入探讨了甲骨文字数、字量、构件系统、结构方式、字际关系等问题,对甲骨文异体字、结构、构件、字频进行了全面、深入、多层次的研究与统计,对殷商甲骨文的结构属性类型及其分布特点作了定量分析与定性论证,提供了一个甲骨文字形系统。作为系统汉字的源头,甲骨文构形属性的研究必为汉字发展史特别是汉字构形史的研究所重视。因而,本文具有较高的学术价值。

论文材料翔实,逻辑严密,言而有据。论文运用了形义学方法、全面归纳材料比勘拟定方法,行之有效……尤其可贵的是一切从原始材料出发,对前人的成就既尊重又不盲从,文中纠正了《新编甲骨文字形总表》误增误摹字 200 多个,这是要通过逐一对比原著录拓片才能发现的,若仅依靠二手材料和进行跳跃式浏览,是决不可能有此创获的。

陈婷珠博士,虽然也算是 80 后新人,但自 1998 年进入华东师范大学中文系本科就读以来,即潜心文字之学,至今已近十载。在这十年中,她以异乎寻常的刻苦钻研精神从同龄人中脱颖而出,2002 年直升硕士研究生在前,2004 年提前攻读博士研究生在后。她的博士论文先是得到了答辩委员会的肯定,后又获得上海市学术著作出版基金的出版资助。作为她的硕士和博士导师,我深知在这看似一帆风顺的背后,是同样不同寻常的坚持和付出。

去年拙稿《铁砚斋学字杂缀》出版时，编辑要求写一段最能表达个人学术思想的文字。我记得当时未加思索就写下了下面一句话：

在博大精深的汉字大厦中，文字学研究尚在艰难蹒跚，期待着登堂入室的辉煌。

现在回想起来，这个想法，其实在很大程度上就是发端于本人近10年来专注于古文字数字化研发实践中的切身感受。对于古文字研究这样一个传统学科来说，新材料的发现固然重要，但是研究模式和方法论方面的突破将会引起这个学科质的飞跃。也正是在这个意义上讲，我们所做的虽然只是在古文字数字化研究的万里征途上跨出了一小步，但是对整个学科的发展来说其意义却是重大的。这样一个想法，当然也适用于陈婷珠博士的这部专著。

陈婷珠博士邀请我为她的处女作写几句话，借此机会我对她为“古文字属性库与先秦汉字发展定量研究”项目所作的工作表示衷心的感谢，更希望她能够以这样一部力作作为起点，继续古文字数字化开发研究的事业。

戊子正月于铁砚斋

全书概述

甲骨文是目前可见最早的成系统汉字材料。只有弄清其字形系统（包括字数、字量、构件系统、结构方式、字际关系等），汉字发展史的描述才有了起点。这是甲骨学研究的基础工作。随着《殷墟花园庄东地甲骨》等新材料的发现，以及数字化手段在甲骨文研究中的介入，既有的殷商甲骨文字形系统的研究成果有了重新认识的必要。这对今后的甲骨文研究以及其它古文字学科建设而言，都有着构建基础平台的意义。

本书分为上下两编，共六章。上编分为两章，主要目标是精确统计现有甲骨卜辞中的字形数量。

第一章主要是通过核对甲骨拓片，对迄今关于甲骨文字形数量研究的最新成果——《新编甲骨文字形总表》进行整理，对该书收字范围内的甲骨文字形数进行全面清理。

此章分为三节：第一节，取消《新编甲骨文字形总表》的甲骨文误增字，依次细分为字形重出、误增字形、查无该字和误收伪刻四类。第二节，核实原拓，对比辞例，修订《新编甲骨文字形总表》的误摹字。第三节，增补应当收入甲骨文字表里的字形。

在《新编甲骨文字形总表》的收字范围内，本章共取消 157 个误增字，修订 72 个误摹字，增补了 126 个字。这是完成现有甲骨卜辞中

字形数量精确的清理工作的第一步。

第二章主要是从甲骨文字形系统出发，广泛吸收已有研究成果，全面清理《殷墟花园庄东地甲骨》、《瑞典斯德哥尔摩远东古物博物馆藏甲骨文字》等新出甲骨文材料中的新出字形。

此章分为两节：第一节，根据《殷墟花园庄东地甲骨》已有的研究状况，分为“既有研究成果认定的新字”、“我们认定的新字”和“似是而非的伪新字”三类进行全面清理。其中，“既有研究成果认定的新字”是指我们对《殷墟花园庄东地甲骨》新出字既有的研究成果的整理；“我们认定的新字”是指目前学界尚未论及，我们认为应予以增补的新字；“似是而非的伪新字”是指以前以为是《殷墟花园庄东地甲骨》新见字，但经研究却不是新字的情况，对这部分字我们予以排除。第二节，对《瑞典斯德哥尔摩远东古物博物馆藏甲骨文字》的新出字形进行清理。

在新出材料的范围内，我们共增补了 228 个字，其中出自《殷墟花园庄东地甲骨》的有 223 个字，出自《瑞典斯德哥尔摩远东古物博物馆藏甲骨文字》的有 5 个字。这是完成现有甲骨卜辞中字形数量精确的清理工作的第二步，将为更好地定量描述最新的甲骨文字形系统奠定坚实的基础。

经过上编的清理，我们最后得到的甲骨文字形总数为 6 243 个。

下编分为四章，主要目标是在上编结论的基础上，从异体字、结构、构件和字频等角度，对甲骨文字形系统进行重新考察。

第三章重新考察了甲骨文中的异体字现象。本章的主要创获有四：第一，经过整理与增补，确定了最新的甲骨文异体字数量。我们认为，甲骨文中严格意义上的异体字为 1 045 组，其字形总数 3 118 个；第二，完成了甲骨文单字的异体构形的数量分级。单字的异体字数量保持在 2-3 个的处于绝对优势；4-8 个的其次；10 个以上的最少。我

们还将其与西周金文异体情况相对比,分析了古文字异体字的层级分布规律。第三,系统分析了甲骨文合文的异体情况。我们认为,合文异体共71组,其字形总数为186个,内容多为先公、先王、先妣名。第四,经过对甲骨文单字与合文异体情况的整理,我们认为,甲骨文中的异体情况共为1116组,3304个字形。甲骨文的字头数实际上为4055个。

第四章主要从结构的角对甲骨文字形系统进行全面、深入的探讨。本章的主要创获有三:第一,在完整的材料平台上,对甲骨文材料进行穷尽性的结构方面的研究,得出最新统计数据,并引入西周金文等材料进行深入对比研究。从自然结构角度分析,甲骨文中独体字的字形总数为1674个,约占甲骨文单字字形总数的28.11%;其相应的独体字字头数为1032个,约占甲骨文单字字头数的26.58%。进入平面组合分析的甲骨文字形总数为4281个,约占甲骨文单字字形总数的71.89%;其相应的字头数为2910个,约占甲骨文单字字头总数的74.96%。属于甲骨文层次组合的字形总数共1658个,约占甲骨文字形总数的27.84%;其相应字头数共1167个,约占甲骨文字头总数的30.06%。第二,确立了20种甲骨文平面组合类型,补正了对甲骨文结构类型的已有认识,揭示了甲骨文字形散点分布的状况,反映出甲骨文作为早期文字的图画性特点。第三,甲骨文层次组合中2层的情形最多,4层的极少,与后世文字系统相比,甲骨文字形系统在结构层次上还比较简单。这也反映出甲骨文字整体发展水平尚低。第四,从表词结构的角对分析,甲骨文能进行表词结构分析的单字字形总数共2134个,占甲骨文单字字形总数的35.84%。其中会意字最多,象形字与会意字的比例相当接近,而形声字较少,指事字居最末位。同一字头下的异体字大部分属于同一种结构方式,结构方式有变化的属于少数情况。

第五章主要从甲骨文构件的层级、构字能力、方位、表词功能等不同属性探讨甲骨文构件系统。本章的主要创获有四：第一，本书首次对甲骨文构件系统做了全面的多层次的定量统计。我们从5955个甲骨文单字字形中共拆分出15168个构件，无差别构件为2169个，独体构件为1685个。第二，从构件层级上看，甲骨文构件系统总体上层级较少，最高层级为4。上位构件共2160个，占到甲骨文无差别构件总数的99.59%。下位构件为291个，占到甲骨文无差别构件总数的13.42%。其中，二级构件为291个，三级构件为57个，四级构件为6个。第三，从构件方位看，甲骨文在方位布局上所体现出了对称性规律，表明了甲骨文字已体现出汉字特有的块体结构。但作为早期文字，甲骨文构件虽受到构件自身意义一定程度上的约束，大部分构件组合的位置灵活多变。第四、具有表词功能的甲骨文构件共608个。其中，象形构件为314个，指事构件为58个，会意构件为356个，形声构件为260个。具备两个及两个以上表词功能的构件为263个。甲骨文构件总体的表词功能类型比较齐备，但单个构件的功能类型较为单一，使用效率不高。第五，甲骨文构件系统存在的两极分化特点：a. 低频构件比重大和多层构件的数量少；b. 个体构件方位分布自由与整体方位分布趋于高度统一。

第六章主要是从字频角度探讨甲骨文字在甲骨文字形系统中的不同地位和作用。本章的主要创获有三：第一，在《汉英对照甲骨文今译类检》、《小屯南地甲骨》和《殷墟花园庄东地甲骨》的材料范围内，甲骨文总字量为56743个，单字量1425个。从字频级别上看，常用字共366个，次常用字共500个，罕用字共559个。甲骨文文字系统在字频分布上存在着极不平衡性，整体呈两端分布的趋势。在甲骨文常用字一端上，存在着常用字尤其高频字高度集中的现象。在甲骨文罕用字一端上，存在着罕用字单字量最多而使用量最低的现象。

第二，与西周金文和现代汉字对比，从集中程度看，甲骨文常用字的集中程度比西周金文与现代汉字都要高，体现出了甲骨文字形系统的不成熟性。第三，以保留前辞等为甲骨文字频分析的版本一，以除去前辞等为甲骨文字频分析的版本二，两者相对比。版本二中，甲骨文总字量为 32 848 个，单字量 1 386 个。字频级别上，常用字共 327 个，次常用字共 500 个，罕用字共 559 个。与上文版本一的结果作对比，两者在单字数上仅差别 39 个，而总字量上却差别近 24 000 个。这说明甲骨文体式对甲骨文用字特点有着强烈的影响，也在很大程度上影响了甲骨文字频分析。