

# 上编 史前汉字符号研究



# 第一章 对以文字起源为核心的汉字理论研究的反思

## 第一节 以许慎为代表的传统文字观及其不足

许慎是著名的文字学家，《说文解字》是中国文字学的宝典，许慎的文字观代表了中国古代传统文字研究的最高峰。要想对传统的文字研究理论有所认识，就必须从许慎入手，以《说文》为对象学习借鉴剖判分析。

许慎的文字研究以汉字的起源为核心，所以我们也首先探讨许慎的文字形成观。

### 一、许慎的文字形成观和变异发展观

#### (一) 形成观

许慎云：

古者庖牺氏之王天下也，仰则观象于天，俯则观法于地，视鸟兽之文与地之宜，近取诸身，远取诸物，于是始作易八卦，以垂宪象。及神农氏，结绳为治而统其事。

庶业其繁，饰伪萌生。黄帝之史仓颉，见鸟兽蹄远之迹，知分理之可相别异也，初造书契。百工以义，万品以察，盖取诸夬，“夬，扬于王庭”，言文者宣教明化于王者朝廷，君子所以施禄及下，居德则忌也。仓颉之初作书，盖依类象形，故谓之文。其后形声相益，即谓之字。字者，言孳乳而浸多也。著于竹帛谓之书，书者，如也。以迄五帝三王之世，改易殊体，封于泰山者七十有二代，靡有同焉。（《说文解字序》）

许慎的这段话是对周秦以来有关汉字形成和变异发展问题的总结。一般理解它包含如下几层意思：八卦和结绳是汉字产生前的原始记事符号，是汉字的萌芽；汉字的发明者是一位圣人，他就是黄帝时的史官仓颉；汉字产生的途径是由简易到繁难，先“依类象形”造出“文”，其后“形声相益”造出“字”；“文字”产生之后书体有很大的变异，以至于封于泰山的七十二家封禅文字没有一家是相同的。

这样的理解表面上看起来有理有据，其实掩盖了许多矛盾。矛盾之一是：八卦、结绳的原始记事方法与文字的产生到底有没有关系？按照大多数人的看法，八卦和结绳是文字产生前的记事方法，他们与文字产生无关。那么，我们该如何理解许慎把八卦、结绳当成文字前身的用意呢？依照许慎的叙述，伏羲造八卦使用了“观象取物”的手法，仓颉造“文”采用了“依类象形”的手段，这两者是不是一回事呢？如若不是一回事，那我们又该如何解释先秦易学中以八卦类比山川方位与人身器官的观念？进一步说，我们该怎样分别易学中以人体五官类比卦名和文字的近取其身现象？

矛盾之二，仓颉造“文”使用了“依类象形”的办法，这种

手法即便与“观象取物”不完全相同，其中的差异也不会是质的区别，它们大概跟美术中的写生和白描差不多。我们都知道，写生和白描这样的画图活动不需要语言的直接参与依照许慎的说法，造“文”之初的情形正是这样的。其后才有了“形声相益”，但这“其后”到底是多长时间呢？许慎没说，我们也不得而知。这当中的问题其实很大，如果与造“文”基本同时，那就意味着许慎心中文字生成的手段可以分成两类：即“依类象形”、“形声相益”。这也正是多数人的看法。但是“依类象形”和“形声相益”这两种本质不同的造字方法凭什么统一于同一个文字范畴之内？要不要区分呢？回答当然是要得。因为前者使用图画的手法，排除了语言的因素；后者文字与语言参合，文字呈现语言，成为“符号的符号”但是如果用“形声相益”之法所造的“字”与“依类象形”的“文”相隔较久，那又如何理解“依类象形”的“文”的起源及其与仓颉的关系？换句话说，我们又从哪里去找文字起源的原点呢？

该如何解决这两大矛盾呢？我们以为：

首先，汉字的起源应该被看成是一种过程，而且也只能看作过程。因此，我们在本书中凡是谈到“文字起源”问题时都使用“形成”这一术语，而避免使用“起源”和“创造”。我们认为，文字的形成要从内外两方面去考虑。从文字自身的内因条件讲，文字是在吸收其他艺术手法的过程中逐渐形成自己的体系的。正因为如此，八卦、结绳等原始记事符号虽然因为尚未能与语言结合并记录语言，因而不被承认为文字，但由于它们（尤其是八卦）在表达手法上与文字（尤其是文）有异曲同工之处，因而被许慎当成文字的萌芽。就“观象

取物”而言，可以说八卦的构成方式与仓颉初文的构造方式并无本质的区别。许慎的文字观是总体观照的文字观，综合的文字观。虽然从已有的材料说明文字演变的全过程有点力不从心，而且在一定程度上也违背了他的相对理性的思维特征；但是，在他的观念中，他的确想从整体上把握汉字的起源并为之找到合理的源头，想要为一统天下的经学找到坚实的基础，进一步说，作为古文经学大师的许慎，他要奠定古文经的正统地位。为此，他必须为文字找到合理的源头，那当然就是由圣人创造的八卦和结绳了。从文字形成的外部因素来看，王宫的统治需要、社会交际的需求、生产力的发展、技术条件的成熟等等都是应该具备的因素。

第二，讨论文字的形成，应该注意区分“文字个体”与“文字体系”的概念。我们认为：文字个体的形成大体上与“观象取物”、“依类象形”等图形绘制的方法有关，与结绳记事、刻符记事的记号手法有关，但跟语言中的声音无关，与记录语言无关；文字体系的形成是文字与语言结合的产物，是文字用于记录语言的自然结果。文字个体的数量虽然不少，但与成体系的字符相比，实在是沧海之一粟。文字个体必然经历了长期的累积过程，逐渐由量变引起质变。之所以产生飞跃，那还跟记录语言的需求有极大的关系。

第三，“记录语言与否”是衡量文字体系形成的最根本标志。单个的文字个体可以是一幅画、一种图腾形象、一个记号，它可能是原始人敬畏、崇拜的对象，可能是原始人寄托情感、表达心愿的工具，也可能是他们区分彼此，标明所有权的符号。这种对象、工具、符号都可以表情达意，但并没有跟语言中的意义直接挂钩，更不是表达语言中的语词。文字体系

是复杂的巨系统，成分不纯，来源不一，能详尽而无透露的记录语言中的词，有相对稳定的排列次序，能够按照语言的规则表达语言。两者是根本不同的。单个的文字个体上升为成系统的文字体系之成员，是一次质的飞跃。它由巫师通神的载体，民众崇拜的对象摇身一变成为社会人群交际的工具。正是在这一变动中，文字完成了由表意到记事、记言的转变，日益紧密的与语言结合在一块，成为语义延宕的最好承载者。当然，在文字个体与文字体系阶段还存在相当长的亦此亦彼、彼此混淆的中间阶段。

第四，所谓“仓颉造字”并非是白手起家，而是有所承继的。《荀子·解蔽篇》云：“好书者众矣，而仓颉独传者一也。”章太炎《检论·论文·论造字之起源》亦云：“文字源流极久，而仓颉独传者，仓颉其人能划一之，统一之。”正是这个意思。归于仓颉名下的“文”，恐怕大部分为前代所创，仓颉只起了整齐归类并颁布实行的作用。把某一种重大发明归功于某一个特殊人物或者由代表官方的皇帝指派某个人物去完成，这种情况在中国传统文化中是屡见不鲜的现象。如黄帝时期的大桡作甲子、容成造历、伶伦造律吕、隶首作算术；后代的蔡伦造纸等等。我们推想：仓颉对前代无数巫师所造之文（文字画）大概是熟悉的，因此，他具备了继承前代之“文”的条件。许慎对这一点是坚信不疑的，所以他才对俗儒鄙夫淆乱仓颉古文的做法表现出极大的愤慨。章太炎对这一点也是坚信不疑的，所以他有《文始》有所谓“初文”、“准初文”的说法。只是他们二人对造字传说的理解未免拘泥，而且都想找到文字产生的唯一源头，以便于梳理文字变异之流，多少有点理想化。我们认为，如果真有仓颉其人其事，那么，仓颉

所造之“字”，大体上是对史前文字意象表达手法的肯定和对史前视觉符号与有声语言日益结合事实的确认，他的工作重心大概会放在整理已有之文的图形并确定他们的读音上。

## （二）变异发展观

许慎指出：不论文字的形成还是变异，都与社会发展的阶段相联系。

在文字的形成期：从远古的伏羲、神农发明八卦、结绳的原始记事方法，到黄帝时期仓颉造字，那是因为“庶业其繁，饰为萌生”，即社会发展到了文明发生阶段的缘故；从仓颉初文到孳乳成字，再著之于书帛，文字体系从它形成之初就蕴涵着变异的因素，所以才会出现“改易殊体……靡有同焉”的情况。

在文字的变异期：先秦，随着社会的激烈动荡，文字系统的内部经受着变异与规范的双重压迫，加上社会“以趣约易”需求的强大动力，最终导致整个系统改变性质，“古文由此绝矣”，单一的变异主流（文—字—史籀—大篆—古文—小篆）格局受到挑战，在隶书通行的同时，出现了所谓“八体”、“六书”。

不论是字体还是字形，文字及其系统的发展变化始终都与社会同步。而作为人造符号的文字，与语言相比，它所受到的社会的影响，尤其是受政治权势的干预程度，要比语言大得多，对这一点，许慎是有着深刻认识的。依照许慎的看法，政权对文字的影响有四个关键的阶段：1. 形成期在“饰伪萌生”的情况下，由黄帝的史官仓颉初造文字，用来统治百官，察明万事，宣布号令，分发财物，增修其德。2. 变异发展期 战国时代，由于诸侯割据，出现“文字异形”的情况，秦始

皇统一天下，由丞相李斯“罢其不与秦文合者”，制定小篆，颁行天下。3. 秦代 统治者“大发吏卒 兴戍役”致使“官狱职务繁”，为了使文字简便易行，下层书吏于是改造小篆，发明隶书，至汉代而大行，仓颉以来流传之古文由此断绝。4. 隶书通行以后 古文虽然未能通行于社会，但文人学士对前代古文的研究并未中断，随着经古文学地位的日益提高，朝廷对古文字的重视也不断升温。在汉代，由官方出面组织的学术讨论就有多次，其成果还被黄门侍郎扬雄编成《训纂篇》；此外，有汉一代，古文不但是学童发蒙教育的主课，还是官定的入仕必考的内容。汉代的文字研究之所以称为“小学”正是这个原因。

另一方面，许慎还认识到，文字对文化制度的许多方面也会产生很大的影响。首先，文字是民族文化之根，按照许慎的说法 文字是“经艺之本 王政之始”是“前人所以垂后，后人所以识古”的对象，“经艺”指以儒家六经为核心的古代文化、礼仪制度，汉字对汉民族文化尊崇古人、重视礼仪思维定势的形成给予了根本性的影响；文字的兴衰，不但事关学统的正变，而且与政统的强弱有关。太平盛世，文字在统性、规范性上都贯彻得比较好；相反，则会歧异百出。正因为如此，历代的政治家上台伊始，总把文字的规范放在重要的地位；许多思想家也都以订正文字作为实行自己政治主张的开端。其次，文字对士人的心态也产生了重要的影响。在经学昌明的古代，识字是读经的最基本要求。用隶书记录的古书虽然也是圣贤的言论，毕竟不如地下挖出来的文字简古，内容可靠，血脉正统。一个经学家，如果没有文字学的根底，就不可能被同道认可，他对经学问题做出的判断也就不能令

人信服。所以，培养合格人才的第一步，就是学习怎样写好字。这不仅仅是对学童的要求，也是对初为官吏者定期举行的专门考试，如果书体不够端正，官职就会因此而罢免。尤其是对科举考试的答卷来说，书体工整端正是最基本的要求，做不到这一点，那就只能被弃置脑后，无人理睬。再次，汉字不仅仅是作为汉文化的载体而存在的，它本身也是人们崇拜的对象。有些时代，汉字被蒙上一层神秘的面纱，许慎揭露的以《仓颉篇》中的“幼子承诏”为“神仙之术”的荒诞，汉代官吏根据文字任意解释法律条文的怪异，时至今日，民间迷信中流行的猜字把戏，以及“敬惜字纸”的说法等等，其实质都跟远古文字崇拜的习俗有着密切的关系。

许慎关于文字变异发展的观点无疑是正确的。但是，经他指出的从仓颉古文—籀文（古文）—小篆—隶书的书体变异的单线演进途径，则有点过于理想化了。对这一点，许慎自己也是有所认识的，他在叙述文字形成过程时将伏羲画卦，神农结绳记事放在仓颉造字之前，交代“以迄五帝三王之世，……靡有同焉”的状况，在叙述文字变异发展过程时涉及“八体”“六书”等等，都是对单线演进方式的补充。但是，由于许慎以维护文字正统为追求目标，以尊古崇古为己任，以古为纯正，这就使他不得不采用选取一点而不及其余的做法，为文字起源确立了唯一的原点，并千方百计的把他和文字的变异发展联系在一起，因而构成了中国文字研究史上影响深远地过于理想化的单线演进模式。

我们认为，文字不是从某一个原点诞生出来的自为性符号，不是某一个圣王在某个早上突发奇想的创造物，而是由数不清的各个氏族的文化精英们在漫长时期的艰苦卓绝的

劳动中逐渐构思并加以完善，通过战争和部落的联合通行于更大的社会集团，而逐渐形成的象征集体意识和权力意志的符号。这样看来，文字的形成因子绝不会是单个的、唯一的，相反，是多元共存的。只有了解这一点，我们才能正确认识“自源型”象形文字体系内部构成的极端复杂性。当然，多元共存并不等于杂居混淆，分清主次也还是很有必要的。一般说来，文字个体的形成多与图画性的因子有关；文字体系形成以后则相对靠近语言中的声音因素；介于两者的中间状态需要有相当长的过渡，这种过渡首先得使用口说的方法辅助图形，然后借助语音固化简单的符号并使该符号具备用于他处的能力。

## 二、许慎的形体结构观

许慎对汉字小篆形体结构的研究是非常深透的。这主要体现在他对“六书”的继承、界定和解释以及对汉字部首的分析归并两个方面。

### （一）六书说

战国以迄秦汉，经古文学家在研究了出土于各地的古文字形体构造的基础上，提出了有关汉字形体结构的规律——“六书”，企图从字形上揭开汉字起源之迷。最有代表性的说法是：

六书：象形、会意、转注、处事、假借、谐声也。（郑众《周礼·地官·保氏》）

谓象形、象事、象意、象声、转注、假借，造字之本也。（班固《汉书·艺文志》）

指事、象形、会意、形声、转注、假借。（许慎《说文解

字序》)

以上三家在“六书”名称、次序方面有一定的分歧关于名称:许慎的说法由于充分考虑了“象”、“事”、“意”的差别,正确揭示了三种结构方法的概念内涵,具有相对的合理性,因而被多数人所认可。相对而言,班固对“前四书”都用了“象”,不大符合理性思维的习惯,“象”就是使用图画的方法模拟自然万物,物象可以模拟,声音也可以模仿,但“事”与“意”怎么个“象”法呢?这是理性思维所不能回答的。但是,如果我们换一个角度,我们试着把文字的形成与前逻辑思维相联系,我们就不得不为班固的英明击掌叫绝了。一个“象”字,恰当地透露出古人“观象取物”造字方法的无限生机关于次序:由于班固意识到了创造文字时由易到难、由具象到抽象的思维特点,从理性的观点看,具有一定的合理性,许慎把“指事”置于“象形”之前,忽略了一部分指事字要以象形为基础构件的事实;郑众有意按形、义、音的次序排列,这有他的道理,许慎的次序和《说文》的说解也正是按着这样的顺序,两人在这里是有共识的。但郑众把“处事”的位置放错了,许慎把它和象形列在一起,则是基本正确的。郑众的排列还有一大优点,就是他将“假借”与“谐声”的次序颠倒,似乎揣摸到形声字之所以出现的真正原因。语言中假借方法的使用,应当与文字体系的形成有相当的关系,而汉字之所以能构成复杂的巨系统,能够完整地记录语言,当与形声字的加盟有关。

研究“六书”的大多数学者赞同许慎的名称、班固的次序。

接下来便是“六书”的概念问题,三家中只有许慎有论

述，他说：

周礼：保氏教国子，先以六书。一曰指事，指事者，视而可识，察而见意，上下是也。二曰象形，象形者，画成其物，随体诘屈，日月是也。三曰形声，形声者，以事为名，取譬相成，江河是也。四曰会意，会意者，比类合谊，以见指撝，武信是也。五曰转注，转注者，建类一首，同意相授，考老是也。六曰假借，假借者，本无其事，依声托事，令长是也。

对许慎的六书定义，大部分学者是基本认可的。但是对其中存在的问题，大家也不回避。首先是对指事字的界定：定义似乎清楚，例字也还算明白，问题是《说文》全书九千多字，许慎自己标明“指事”者，只有几处，与“象形”“会意”“形声”的义例几乎不成比例。说是“指事字”少吧，与事实不符；说是后人窜乱吧，又缺乏证据。其次是对会意字和假借字的例字，多数人不恰当。“武”“信”虽然属于会意，但“止戈为武”，“人言为信”的说法显然是周秦以来的伦理观念，用后起的观念解释字的本义，当然不合适。至于“令”“长”，由“发号施令者”到“县令”，由“长发老人”到“官长”，那都是词义的引申，不是声音的借用。所以朱骏声《说文通训定声》改“令长”为“朋来”。再其次是对转注的理解，千多年来，简直是“斩不断，理还乱”的疑团。

对以上三个问题，我们的初步回答是：“指事”是将已有之图形与已有之记号相互嫁接，通过图形与记号的关系呈现透露物象涵义的结构方法。由于原始的可用图形呈现的物象数量有限，而且要满足“察而见意”的条件，所以构造能力

有限。至于许慎没有明确标明的原因，我们还不得而知。许慎所举的会意和假借的例字的确是不恰当的，但问题的关键似乎不在这里。我们以为：会意是把已有之若干图形叠加在一起，构成极富刺激味的新意象，以便获得心灵震撼的一种结构方法。指事也好，会意也好，都是原始构字手法之一，其实质都是“观象取物”。至于假借，本质上属于听觉符号，是为语言符号常用的“借音表义”的手段，它的出现应该与文字体系的形成有关。把假借的手法运用于视觉图形中，是史前文明向有序阶段过渡的关键步骤，它促成了语言跟图画 的因缘，由此形成了文字体系，并成为人们超时空传达经验，表达感情的有力工具。

许慎关于六书的定义，既规定了汉字形成及结构研究的方向，又带来了许多新的矛盾。大体上说来，清代以前，由于字本位的观念占着主导地位，人们把六书看作“造字条例”，并认为依照六书原则可以造出一切汉字。到了清代，随着《说文》研究的不断深入，语言学家发现：“六书”并不能涵盖所有的汉字；与字形相比，有时候语音的作用更大，解决实际问题的能力更强。这可以说是抓住了语言研究的关键。用这样的眼光看问题，“六书”里所说的六种造字条例就绝不是一个平面上的事情。语言学家戴震首先提出了“四体\_用”的观点，认为“六书”中的前四种是“造字之法”，后两种是“用字之法”。换句话说就是：按照前四书可以造出新字来，后两书则不会。这种观点得到大多数语言学家的认同。因为它既看到许慎定义的不足，又提出了解决的办法，最为关键的是它维护了语言的纯洁，肯定了理性思维的可贵。但是，如果我们能正视文字形成的背景，能看到前逻辑思维在文字符

号形成过程中的影响，能正确区分史前单个文字与有史文字体系，就不会盲目乐观了。

事实上，继戴震提出“四体二用”之后，针对“六书”的层面问题，文字学家们也提出了不少新的创见。如唐兰的“书”是指象形、象意、形声三类。依唐先生的看法，六书中的“指事”这个名目，只是前人因一部分文字无法解释而立的……在文字史上，根本就没有发生过指事文字”（《古文字学导论》87页，齐鲁书社1981年版）。既然六书中的指事不成立，转注、假借也不是造字法，那就只有象形、象意和形声了。唐先生的“三书说”很有见地，但也有不足：首先是象形和象意不大好区分，操作起来较困难；其次是他将假借排除在“三书”之外，不承认假借也是造字之法，这样，就使文字体系的形成变成了无源之水。陈梦家1956年出版的《殷墟卜辞综述》对唐兰的“三书说”作了一些补充，同时提出了自己的“新三书”：即象形、假借、形声。这一主张还得到了裘锡圭先生的响应，不过，裘先生也作了小小的改动，将“象形”改为“表意”。在他的《古文字学概要》（商务印书馆，1988年版）里，裘先生正是按照这样的分类分别讨论汉字的结构。同样是主张“三书”刘又辛先生在他的《汉字发展史（二）》（《语文建设》1998、12）里也主张象形、假借、形声。与陈梦家的三书不谋而合。刘先生认为：六书中的象形、指事、会意应合并为一类，理由是他们都表形。象形用写实手法表形，用于具体事物如：人、马、竹、木、山、水之类；指事用象征手法表形，如上、下、一、二、三等；会意是用较繁复的象形法表形，如：武、趾、林、森等。这类字是从远古的记事图画演变而成，其共同的特点都是用表形的手法记录语词，因而应合并为一类，名

为象形。再加上用其音不用其形的假借，兼表形和音的形声，汉字的造字法一共有三类：象形、假借、形声。何九盈先生指出：陈与唐的分歧，其实质是分类标准的不同。唐兰讲的是汉字的结构，陈梦家指的是汉字的基本类型。如果从结构的角度看，表意、假借、形声根本不属于一个层面；只有从汉字的类型角度入手，才可能将表义、假借、形声纳入同一个层面中加以讨论（见《汉字文化学》，辽宁人民出版社 2000 年版 180 页）。何先生不但赞成三书说，而且从理论上进一步深化，他说：“我所谓的‘三书’是建立在二元化的表达机制之上的。即造字表达与借字表达。造字表达又分为表意类、形声类。（同上 181 页）他认为：“由二元化表达机制构成的‘三书说’，既吸收传统六书说的优点，也区别了造字表达机制与借字表达机制的不同。（同上 181 页）“造字表达属于结构机制，借字表达属于转换机制。这两种机制上的特点构成了汉字的系统性、完整性、灵活性。（同上 181 页）”

在我们看来，“六书”既不属于汉字的结构机制，也不属于汉字的造字机制，而是汉字的表达手段。采用相对写实的图形表达心智就是象形，用相对写实的图形加抽象记号就是指事，整合几个图形就是会意，以上三书都属于象意的范畴。为什么用“象意”而不用“象形”呢？这是因为有史以来的象形字大部分并不是对客观事物的直接描摹，而是对史前流传下来的意象符号的摹写，因此，所谓象形字本质上是写意而非写实的，所以应该沿用唐兰先生的术语称之为“象意”。形声是语言中的声音因素介入文字体系以后的结果，是写意形的汉字体系发展的最理想境界。形声文字的构成借助了“假借”的手段，假借为形声之因，形声为假借之

果。假借和形声共同构成了语言和文字的因缘，促成了从表达心智的史前文字向记录语言的文字体系的过渡。所以我们把假借和形声统称为“象声”。（参见第五章第三节汉字的结构方式）

我们认为，讨论文字的形成问题要有以下几方面的共识 第一 应该以“文字形成”的概念取代“起源”的含混概念。“形成”是二元或多元互补的结果，“起源”则是事物的单个的源头，我们应该破除起源问题上的一元论的思维模式，采用二元或多元对立的研究格局。第二，应该以“表达方式”取代“结构方式”的概念，充分认识文字符号系统形成的多元互补、开放共生性质，破除单线演进式的自然进化观。在这样的背景下，我们认为：假借肯定是汉字的重要表达方式，就连转注，也是一种修辞意义上的表达手段，是用隐喻联系意义的方式。汉字之所以有诗化的特点，跟转注的表达方式很有关系。靠了它，图形起了联想，生成新的含义，象征的作用扩张，美的意味加强。如果我们把转注看成是图符之间的借用，那么，转注也许还是开启假借之门的一把钥匙。既然图符之间可以借用，那么音符之间为什么不能借用呢？由此看来，汉字的表达方式可以也应该是多元的和分层的：说多元是指图形表达、审美表达、声音表达等等；说分层是指文字单体形成时期和文字体系形成阶段。把两者分开可以更充分的讨论汉字的“起源”，尽可能避免不必要的纠葛。实际上文字及其体系的形成当然要借助多种表达方式，而多种表达方式的使用也不见得是平均分派。大体上说来，在史前文字阶段，图形的、象征的、联想的、审美的表达恐怕是主要的方式；在有史文字体系形成阶段，声音的、记号的或者声音兼图