

民国学术文化名著

文字学概论

刘大白 著

文字学是一种科学，是说明文字底性质和效用，以及它的形态意义和声韵底发生和演变的科学。但是文字是因为国家或民族底不同而不同的，文字学所研究的对象，自然也跟着不同。现在所研究的，是被称为中国国文的文字学，所以可以称为国文文字学。

民国学术文化名著

文字学概论

刘大白 著

岳麓书社

图书在版编目 (CIP) 数据

文字学概论 / 刘大白著. — 长沙: 岳麓书社, 2011.6

(民国学术文化名著丛书)

ISBN 978-7-80761-653-5

I. ①文… II. ①刘… III. ①汉字—文字学—概论 IV. ①H12

中国版本图书馆 CIP 数据核字 (2011) 第 103862 号

文字学概论

作 者: 刘大白

责任编辑: 张卫国 靳龙龙

封面设计: 肖睿子

岳麓书社出版发行

地址: 湖南省长沙市爱民路 47 号

电话: 0731—88885616 (邮购)

邮编: 410006

网址: www.yueluhistory.com

2011 年 7 月第 1 版第 1 次印刷

开本: 640×960 1/16

印张: 5.75

印数: 1—6,000

ISBN 978-7-80761-653-5/G·993

定价: 12.00 元

承印: 长沙鸿发印务有限公司

如有印装质量问题, 请与本社印务部联系

电话: 0731-88884129

整理说明

一、丛书着力于“学术”与“文化”两方面，所收著作或为学术上开新之作，或为文化上奠基之作。

二、丛书之收书范围，原则上起于民国建立，迄于中华人民共和国建立。然某些著作之成形，可追溯至民元之前若干年，因其有重要地位，亦酌情收入。

三、文、史、哲之分，原系西洋通则，本就不太适用于中国学术，故丛书不按学科分类，而是根据整理进度，顺次出版。

四、丛书所收诸书，原版均为繁体竖排，在其流布过程中，亦有版本差异、文字错讹等现象，为方便读者，此次做如下整理工作：

1. 繁体字改为通行之简体，竖排改为横排（原书中一般“右表”、“左表”、“右文”、“左文”均改为“上表”、“下表”、“上文”、“下文”），但为充分尊重原著，原书中专名（人名、地名、书名等）及其译名皆一仍其旧，凡底本脱、衍、讹、倒之处，除个别讹错明显且影响文意阅读者稍作改动外，皆一仍其旧。

2. 凡排印误刻者，如日曰、己巳巳、戊戌戌之类，均径改，不出

校记。

3. 为方便当代读者阅读，标点符号按现代汉语使用规范作了处理。

4. 丛书中多本有作者原注，原书以夹注出之，此次整理皆排入正文，并以楷体小字以为区分。

5. 各书附“后记”一篇，说明著者爵里、版本流布、各界评论等情况，以期为读者提供阅读指南。

古人云：“校书如扫落叶，旋扫旋生。”吾人虽勉力为之，而乖漏难免，还祈方家教正。

目 录

一 解 题	1
二 文字底性质和效用	2
三 文字底三要素	7
四 文字底起源	9
五 文字底构成	12
六 文字发生和演变的次第	19
七 文字底形态	23
八 文字底声韵	56
后 记	83

一 解 题

文字学是一种科学，是说明文字底性质和效用，以及它的形态意义和声韵底发生和演变的科学。但是文字是因为国家或民族底不同而不同的，文字学所研究的对象，自然也跟着不同。现在所研究的，是被称为中国国文的文字学，所以可以称为国文文字学。或许有人说，“中华民族，是合汉满蒙回藏五族而成的。这五族都有文字，那末，中华民国底国文，共计有五种；国文底名称，不应该由汉族所用的文字来独占”。这话似乎也颇有道理。但是，（一）汉族在五族里面，人数占得独多，文化也处于最高的地位；（二）满蒙回各族，差不多都能懂得汉文，尤其是满族，因为被汉族文化所征服，连懂得本族文字的，也寥寥无几了；（三）汉文在习惯上向来被称为国文，将来的趋势，也一定成为事实上的国文：所以称研究汉族底文字的文字学为国文文字学，并非过分。至于本编所说明的，只是国文文字学底概略，所以称为国文文字学概论；而把文字的形态意义和声韵三部分，详细地分别研究起来，还应该另作各论。

二 文字底性质和效用

大家知道，文字是记录语言的，而语言是发表思想的。所以语言就是发表思想的工具，文字就是记录语言的工具，也就是发表思想的工具；不过语言是用声音作符号而直接表出思想的，文字是用形态作符号，而直接记下语言，间接表出思想的。人类在不曾创造文字的时候，早已有了语言；所以人和人间，能把意思互相传达，感情互相联络，知识互相教导，而组织比较复杂，常常演进的团体。但是声音的符号，是时间性的，是跟着时间底过去而消灭的；一经消灭，便只能在听者底回忆上再现；而且在空间上也受着限制，不能传到较远的地方。时间上既不能久留，空间上也不能远达，都要凭着口耳相传；所以在只有语言的时候，人类不能组织较大的团体，也不能有分明的历史。到了发明了文字，就把这种困难解除了。‘因为形态的符号，在时间上可以留得很久，在空间上可以传得很远；所以社会底组织能够扩大，人群底文化也能够加速地增进。人类所以不同于其它动物，自然有许多原因；但能够用语言，是最要的原因之一。文明人所以不同于野蛮人，自然也有许多原因；但能够用文字，也是最要的原因之一。

文字既然是记录语言的，所以文字底性质，合语言底性质有密切的关系。世界上的各种语言，语言学者依着语言形态底特质，大体上分为三类：（一）孤立语，（二）黏著语，（三）曲折语。

（一）孤立语 语词底形态上，没有表示语法上的关系的变化，不能表出语词相互间的联络关系。中国底语言，是属于这一类的。例如“在止于至善”，善字是名词。“善善恶恶”，第一个善字是动词，第二个善字是名词。“其言也善”，善字是形容词。“善搏虎”，善字是副词。如果只取善的一语，是名词，是动词，是形容词，是副词，不能知道；一定要从文句底位置上才能明白它底品词的性质。所以中国语又称为列位语，就是从语底排列的位置上显出语底品词的性质。

（二）黏著语 把没有独立的资格的音或音节——就是表示语法上的关系的成分，黏附在表示意义的成分——就是语底本体——底前面或中间或后面，而表示出语法底关系来；可是这些黏附上去的成分，合语底本体，不相混和，两者的区别是判然的。日本语和朝鲜语，是属于这一类的。

（三）曲折语 表示意义的成分，合表示语法上的关系的成分，完全混和而成为一，或是这两者非常紧密地相结附的。为了表示语法上的形式的缘故，语根底内部的音变化着，再在语根上附加上接辞（上述黏著语底黏附上去的成分，黏附在前面的称为接头辞，在中间的称为插入辞，在后面的称为接尾辞，总称为接辞）的也有。有些人称这一类语言为曲尾语，以为只限于语尾底变化；其实不然。从语尾部分底变化，表示出语法上的形式的固然很多，但是语词内部的母音变化的也有；以为只限于语尾底变化，是不对的。印度日耳曼语，是属于这一类的。

至于世界上的各种文字，文字学者大体上把它们分为表意文字和表音文字两类。

表意文字一类，更分为图画文字和象形文字两种。所谓图画文字，不是纯粹的图画，不过是图画的成分比较显著地表出的文字罢了。图画

文字和图画底区别，就是图画是事物的，不可读的，图画文字是语言的，合语言底声音结合起来而可以读的。从图画文字更进一步，就是象形文字。例如中国文字中的“山”和“川”，本来从图画而出的；但是楷书不消说，即使是篆文隶书，它们底图画的成分，在表面上也并不显然地表示出来。这就是所谓象形文字。不过纯粹的图画和图画文字，不能区划很分明的界线；而图画文字和象形文字之间，也是这样。

表音文字一类，又分为单节音词表音文字，多节音词表音文字和多节音词表音分文字三种。单节音词表音文字，是用一个文字表出一个单节音词底一个音的。本来，在语言和文字底关系上，文字能够表出语言底单位，是一句极限的话。语言底第一次的单位就是句子，但是表出一个句子的文字，是不存在的。这理由是很简单的。因为即使不论怎样简单的国语，可以有无数的句子存在；如果有能表出一个句子的文字，那末，文字不得不是无数的了，到底是难供实用的。这表出单节音词底一个音的文字，中国的汉字就是一个例。照普通讲，汉字是表意文字，是象形文字，但这是简单的分类，事实上并不如此，而是包含着种种的性质的。用作助词的假借字“耳”和“焉”，从字体底起源上讲，“耳”是耳朵底象形，“焉”是鸛鸟底象形；可是用作助词的时候，是单用它底音的，是属于表音文字底性质的。又如形声字“江”和“河”，也可以看作属于表音文字的部类。因为“江”和“河”底“工”和“可”，不过是表出音的。偏旁的“氵”，是水底象形底变化；所以它们是象形的——就是表意文字的——成分，和表音的——就是表音文字的——成分结合起来的。但是汉字底这一类的偏旁或上下，不过是一种识别的限定成分，是象形的；而它们底显著的特征，可以说是在表出一个单节音词底一个音的性质。多节音词表音文字，是用一个文字表出一个多节音词底可以分析的一个音节的。日本底平假名和片假名，就是这种文字底代表。例如日本人称你为“アナタ”，是一个多节音的词。“ア”和“ナ”和“タ”，都是一个多节音词中可以分析的一个音节，是每一个音

节都用一个文字表出它的。从多节音词表音文字更进一步，就是多节音词表音分文字，所谓音分，是指一个音节底成分而言，就是构成一个音节的母音或子音。这种文字所代表的，是分解语言而得到的最终的成分，就是构成一个音节的成分。用在语言上的一切不同的音节底成分，各各用不同的符号记出的，是理想的表音分的文字；但是这种理想的文字，是实际上不存在的，像欧洲各国所用的字母，可以算是表音分文字底代表的文字。不过这种文字，也有种种的差异。只就它们底主要点上区别起来看，有像欧洲底字母，把子音和母音作对等看的文字组织以及像印度底梵文，把母音分作独立的母音和音节构成的母音，而音节构成的母音是认为从属于子音，而用别种的符号表出的文字组织。以上的分类，只是从大体上观察文字的结果：如果细说起来，兼有多节音词表音文字和多节音词表音分文字的两种性质的也有。

照前边所说，中国底汉字，是单节音词表音文字。所谓词，就是说话的时候表示思想中一个意象——就是观念——的语词。所谓单节音词，就是一个语词，只有一个音节。拿一个音节表出一个观念，就拿一个符号来表出一个音节，所以叫作单节音词表音文字。但是这不过是大体上可以这样说，或者从前造字的时候是这样，而现在已经不如此。例如蝴蝶、蟋蟀、小孩子、便宜、分付等，都不能认为单节音词了。不过大体上说汉字是单节音词表音文字，也是不妨的。

一般的文字学者，虽然把世界上的文字，只分为表意文字和表音文字两类；但是这种分类法，实在是不精确的。构成文字的要素，有形态，意义和声韵三者；而形态这一个要素，本来都是从象实物底形体的象形文字上递嬗出来的。所以文字底种类，应该有表形文字、表意文字、表音文字三类；而中国底汉字，就是兼具这三类的。他们所谓表意文字中的象形文字，实在就是表形文字；而中国底汉字，在表形文字以外，还另有表意文字。不过表音文字，固然是表音的符号，而表意文字，既用这个符号表意，同时也就用这个符号表音；表形文字，既用这

个符号表形，同时也就用这个符号表意而且表音。从表音方面看，表形文字和表意文字都是表音的符号；而中国底汉字，又多数是拿一个音节表出一个观念，就是拿一个符号表出一个音节的；所以说它是表音文字，是单节音词表音文字，也是可以的罢了。

从语言底分类上看，中国底汉语，是孤立语；从文字底分类上看，中国底汉字，是单节音词表音文字；那末，这两者底密切的关系，可以看出来了。因为语言是孤立语，所以不需要多节音；因为没有语尾，没有语尾底变化，没有语词内部母音底变化，所以只消拿一个符号表出一个音节，而成为单节音词表音文字。

三 文字底三要素

前边说过，语言和文字，都是发表思想的工具；所以语言和文字底效用，都在乎表出意义。不过语言是用声音作符号而表出意义的，而文字是用形态作符号而记下语言底声音，表出语言底意义的。文字合语言底不同，就是多了一个形态的符号；所以形态是文字底特征，而意义却是合语言共同的第一要素。没有意义的声音，不成为语言；没有意义的形态，也不成为文字。用形态表出意义，不只是文字如此。例如图画，也是用形态表出意义的；没有文字的时候的所谓结绳，也是用形态表出意义的。不过它们都缺了一种要素，就是没有声音，是不可以依着语言底声音读出的。所以声音是文字和语言所共有的第二要素。没有声音的意义，只有思想；没有声音的形态，只有图画或记号。至于第三要素，当然是形态了。只有声音和意义而没有形态，便只是语言而不是文字。不论何种文字，都备具这三种要素；而因为意义和声音，都仗着形态表出，所以形态虽然是第三要素，却是文字底最重要的要素，也可以说是最表面的要素。但是这所谓形态，只是符号底形态，不是专指象形文字底形。象形文字底形，是指所象的实物底形体而言，是形态底一部分。

符号底形态，表出实物底形体的，是表形文字，就是形符；表出一个具体的或抽象的观念的，是表意文字，就是意符；单表出一个声音的，是表音文字，就是音符。形符底使用，是最早的；其次是意符，最后是音符。中国底汉字，是这三种兼用，而且错综起来用的，所以汉字底造字原则，有所谓六书。

四 文字底起源

中国向来的传说，都以为汉字是仓颉所造；可是好笑得很，仓颉到底是什么时候的人，至今没有人弄得清楚。其实，这种无聊的问题，咱们可以丢开不管；因为文字是民族的社会的产物，是自然地偶然地产生的，决不是某一个人制造出来的。咱们所要讲的文字底起源，并不是说明文字是起源于某一个人底制作，而是说明文字是从某一种东西演变出来的。

向来的见解，都以为文字是起于象形的；换句话说，就是文字是起源于图画的。然而文字底发生，实在有记号和图画的二源；而且记号底起来，似乎更早于图画。试看《周易·系辞》下篇说：

上古结绳而治，后世圣人易之以书契。

汉代许慎《说文解字叙》说：

……神农氏结绳为治……黄帝之史仓颉，见鸟兽蹄迹之

迹，知分理之可相别异也，初造书契。……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。文者物象之本，字者言孳乳而寔多也。著于竹帛谓之书；书者如也。……

所谓结绳，就是一种记号。书就是著，也就是刻画底画；契就是契，也就是刻画底刻。这不但中国上古时代如此，现在未开化的野蛮民族，还有实行结绳以作记号的。可见未有文字的时代，先用结绳以作记号；后来渐渐进化了，就把这种结绳的记号，刻画在甲骨金石竹木上，便成为文字。这就是所谓“易之以书契”。记得蔡元培先生曾经说过：现代某种民族，把绳子一条一条地结在木橛上以一条表示一，二条表示二，三条表示三。因此推知中国篆文“一”“二”“三”三个字底古文作“弋”“弌”“弌”，反比篆文更繁复；就是不但把所结的绳子刻画下来，而且把木橛也刻画下来，而以弋象木橛的。后来把木橛省去不刻画了，便成为篆文底“一”“二”“三”（原文如何，因为一时无从检寻，已经记不清楚了，这只是大意如此）。这话虽然只是一种类推，而所谓古文，也并不一定早于篆文；然而“一”“二”“三”等字，发生较早，只是一种简单浑沦的记号，而不是图画，是很明白的。文字最初的发生，是把结绳的简单浑沦的记号，刻画到甲骨金石竹木上来，也是比较地可以相信的。至于图画，是比较繁复分明的工作，比用结绳的方法作简单浑括的记号难得多；要等到发明了刻画记号的方法以后，更进一步，才能发明图画的方法。先有记号，后有图画，正合从简单到繁复，从浑沦到分明的进化通则相合。所以我们可以知道，文字是发生于记号和图画二源的，而记号一源，更早于图画。换句话说，也可以说图画是由记号演进的。因为把简单浑沦的记号，渐渐地繁复分明起来，就渐渐成为图画；而把结绳的记号刻画在甲骨金石竹木上，广义地说，也可以说是最初的图画。

关于许氏所说，结绳是不是神农氏时代，仓颉是不是黄帝时代，书

契是不是仓颉所造，乃至神农氏、黄帝和仓颉等，中国古代有没有这个人，当然都是疑问。但是古人见鸟兽蹄迒之迹，知分理之可相别异，因而造成书契，或许有这么一回事。因为人类当初只知道结绳记事；后来看见鸟兽底蹄，还能印出迒迹来，人类底手，难道不能吗？鸟兽蹄迒之迹，还有分理可相别异，人类底手刻画起来，难道不能吗？因此仿效鸟兽蹄迒之迹，用手刻画起来，造成书契，以代结绳了。然而所仿效的，只是鸟兽蹄迒之迹，所以初作书的时候，只是把鸟兽蹄迒之迹，照样——就是“书者如也”的“如”——地刻画在甲骨金石竹木上，而只是些简单浑沦的几何学点线面式的记号。并且这些记号，还只是一个记号，只是有形态有意义的记号，而不是有声音可读的文字。后来因为既有形态，渐渐地把它合语言底声音结合起来，成为有声音可读的记号文字，是合由图画而成为有声音可读的图画文字的过程一样的。

照前边所说，文字是起源于记号和图画两者；由记号而演成记号文字，由图画而演成图画文字和象形文字。记号文字是形意兼表的，就是六书中所谓指事字；图画文字和象形文字是表形的，就是六书中所谓象形字。