

# 第 1 章

## 汉字与符号学

### 第一节 语言、文字与符号

#### 一 符号

符号学 (semiotics, semiology) 的研究可以追溯到遥远的古希腊时代, 医学家希波克拉底 Hippocrates 把病人的症状看做符号, 世称“符号学之父”。然而作为一门现代科学, 符号学研究的是更为一般的符号理论, 奠基于 20 世纪初。人们公认瑞士语言学家索绪尔 (Saussure, F. D. 1857 ~ 1913) 和美国哲学家皮尔斯 (Peirce, C. S. 1839 ~ 1914) 是现代符号学的两大先驱。索绪尔认为符号学是研究“社会生活中符号生命的科学”<sup>①</sup>; 语言学不过是这门一般科学的一部分”, 而皮尔斯则把符号学看做是更广泛意义上的逻辑学的代名词。在 20 世纪, 符号学很快发展为一门跨学科的新兴科学, 它从哲学、逻辑学、语言学迅速地发展到社会学、人类学、生物学、美学乃至建筑理论、工业设计等各个领域。

“符号”(sign)一词的使用虽然渊源已久,然而它的含义却一直模糊不清。人们通常是把符号简单地看成是某些图形或图像之类的东西,甚至在经典著作家们那里,对“符号”一词也往往有不同的理解。直到索绪尔把语言符号解释为由能指(signifier)和所指(signified)组成的统一体时,人们才算对“符号”的含义有了比较一致的解释。在《普通语言学教程》一书里,索绪尔所说的“能指”指的是语言符号的“音响形象”,“所指”是它所表示的概念。他还把能指和所指比作一张纸,思想、概念是纸的正面,声音是纸的反面,它们永远处在不可分离的统一体中。所谓符号的“意指”或意指作用,实际就是把能指和所指结合成一体的行为过程,其结果就是符号。值得注意的是,人们通常所说的“符号”,一般仅仅是指符号的能指,而不是能指和所指的结合体。符号有两个层面,能指只是表达面,即形式或符形;所指为内容面,是对能指的解释,也可以称之为“符释”。所以,当我们说某个图像或事物是一个“符号”时,它的含义就是一个能指和所指的结合体,而不仅仅是它的能指或符形。当然,人们出于习惯或者为了方便,在不造成误解的情况下,也可以把符形直接说成“符号”。

符号在本质上是一种关系。索绪尔把符号解释为能指和所指的二元关系,而皮尔斯的符号概念则是三元关系。在皮尔斯的著作里,符号除符形和符释以外,还有符号所指示的客观事物,即符号对象。符号总是用甲来表征乙,并且传达一定的讯息。甲即是符形(能指),也就是人们习惯上常说的“符号”,有人称之为“符号自身”或“符号形体”;乙是甲所表征的符号对象;甲所传达的讯息即所指,皮尔斯称为符号解释(符释)或解释项。把符号定义为三元关系,是皮尔斯对符号学作出的最重要的贡献,也是皮尔斯符号学的精髓。

皮尔斯对符号学的另一重要贡献,是关于符号的分类。其中最有价值的是把符号分为图像符号(icon)、指索符号(index)和象征符号(symbol)三类。在图像符号中,符形与符号对象之间的关系,表现为某种肖似性。如照片、图形、模型、比喻、方程式等,都属于图像

符号。在指索符号中，符形与对象之间具有因果的或时空上的联系。如烟作为火的指索标记，风标作为风向的指索标记，都是指索符号。象征符号的符形与对象之间无直接联系，它们的关系建立在社会约定的基础之上。如鸽子是和平的象征，狮子是强大的象征，国旗是一个国家的象征等等。皮尔斯这一符号分类，在符号学研究中产生了深远的影响。

符号是人类使用的一种普遍工具，我们每个人都生活在符号的世界之中。符号有人工符号和自然符号的区别，人们大量使用的是人工符号。例如交通路口的红绿灯，战场上飞升的信号弹，篮球裁判员的手势 朋友见面时的握手 数学符号、逻辑符号等人工语言等等，它们都是能指和所指的统一体，都属于人工制定的符号。自然物也可以充当符号。例如 预示雷鸣的闪电；“一叶落而知秋”的落叶，都是自然符号。并非自然物都是符号，只有当自然物用以表征另一事物并且传达一定讯息的时候，它们才是符号。比如月晕和础润都是自然现象 不能算是符号 但是“月晕而风 础润而雨”当人们把它们作为风雨先兆的时候，赋予它们所指的含义，于是成了符号。

其实 古代符号学的研究 并非只有西方。在东方 中国就是一个具有悠久的符号学研究传统的国家。古代中国虽然没有系统而完整的符号学理论，但是远在先秦时期，诸子百家的著作中确实有许多关于符号学的讨论，其中名实关系即是这种讨论的重要内容。那时候“名实相怨 处士横议”，一场名实关系的大讨论热闹非凡。“名”即是能指或符形，“实”为符号对象，名实关系即符号的能指和指示对象之间的关系。由于“名实”又同“言意”相联系 所以也涉及符号的所指或符释的问题。

在诸子百家中，名家的符号学研究是最具代表性的，《公孙龙子·指物论》即是一篇精彩的符号学论文。《指物论》开宗明义提出“物莫非指 而指非指”的基本命题 然而由于费解 学术界至今仍然争论不已。其实 这里的“指”就是符号。中国当代学者周礼全在翻

译莫理斯的符号学著作时就把“sign”翻译为“指号”。《指物论》的作者一次次说到的“指”是在不同层次上使用了“符号”的含义。“物莫非指”当代有的学者解释为“所有事物都可以充当符号（“莫非”双否表示肯定）”，这个“指”是符号的本义，具有能指和所指的二元关系或符号形体、符号对象和符号解释的三元关系。“而指非指”表面上是个矛盾命题，所以最使人困惑。但是可以分析为：前面的“指”是符号的能指，后面的“指”是符号对象，即事物。也就是说，符号能指有别于符号对象，即甲表征乙而不是乙。在当时，有人认为事物有了“名”就用不着“指”（符号）了。在公孙龙看来，“物不可谓无指也”，“名”同“指”是不能混淆的。“名”只是能指，是符号的一个因素，“名”有异于符号，所以公孙龙仍然坚持“物莫非指”的基本命题。这些分析自然是很有道理的，只是在多大程度上符合公孙龙的原意，那就不得而知了。公孙龙毕竟是古人，他并没有真正弄清楚符号的二元或三元关系，更没有对符号的含义作比较清晰的描述。但他能够意识到“指”有异于“名”并在不同意义上使用“指”的含义，这已经难能可贵了。

符号学的研究经历了漫长的岁月，虽然至今还没有一个公认的学科体系，但它已经唤起了人们广泛的兴趣，成为众多学科关注的对象。正如法国哲学家马里坦所说：“没有什么问题像与符号有关的问题那样对人与文明的关系如此复杂和如此基本的了。符号与人类知识和生活的整个领域相关，它是人类世界的一个普遍工具，正像物理自然世界中的运动一样。”

## 二 语言符号

现代符号学思想的首要来源是语言学。在人类的符号世界中，语言符号占有特殊重要的地位，这是因为语言符号具有特殊的透义性。语言符号的能指和所指，能够在自然状态下形成一个特定的有机整体，因而在交际过程中，只要不发生某种障碍，比如说不是说话人的口音难懂，人们的注意力总是集中在了解对方所传达的思想

上,甚至完全不注意声音的物理性质,这就是透义性的表现。语言符号的透义性使得它在交际的实践中,无需译解为其他符号就能为人们所理解。而其他符号则需要译解为语言符号,如把作为赠品的玫瑰花译解为“爱情的象征”“交通路口的红灯译解为“车辆停驶”等等。因此其他符号只具有语言符号的导出的性质或者形象地说,是以语言符号“反射的光发亮的”。语言符号还由于透义性使得它不像其他符号受到材料的限制,因而具有无可比拟的抽象能力,达到了高度精确化的水平。

索绪尔最先提出语言符号为能指和所指的双面体,并且由此类推到一般符号上面使“符号”一词有了比较确定的含义。皮尔斯的符号学说是以符号形体、符号对象和符号解释的三元关系为特征的,而在索绪尔看来,语言理论所需考察的只是符号的能指与所指之间的关系,而不是符号能指与它所表示的事物之间的关系。所指不等于事物,它只是同事物相关的思想或观念。符号的能指当然与它所表示的客观事物相联系,但这不是符号意指理论考虑的主要内容。这也就是说,索绪尔的符号的二元关系学说同皮尔斯的符号的三元关系学说在理论上虽然有所不同但大体是相一致的只是侧重点有所不同罢了。索绪尔这种语言符号学思想,后来为丹麦学派的叶尔姆斯列夫(Hjelmslev, L.)所发展。叶尔姆斯列夫把能指和所指的关系看成表达面和内容面的关系,并且认为符号就是由表达项和内容项构成的一种函数,即符号函数。符号函数是一种倚变关系,在叶尔姆斯列夫的理论中,还对表达项和内容项作了进一步的层次分析。

索绪尔的语言符号学是一种“双面”、“双轴”的关系学说。语言结构除能指和所指双面关系以外还有组合轴和聚合轴的“双轴”关系,即组合关系和聚合关系(后者在索绪尔的术语里为“联想关系”)。组合关系指成分与成分作线性排列并按一定的方式组合起来而形成的关系,是语言片段中成分与成分组合的一种模式。如“桌子”这个词里,“桌”与“子”形成附加关系附加关系是构词平面

的组合关系；“教育事业”这个短语里，“教育”和“事业”形成偏正关系，偏正关系是句法平面的组合关系。聚合关系指可以在某一组合中出现的所有成分之间的关系。如“桌子”的“桌”可用“篮”来替换成“篮子”，“桌”与“篮”是语素平面的聚合关系；“教育事业”的“教育”可用“旅游”来替换成“旅游事业”，“教育”和“旅游”是词平面的聚合关系。索绪尔认为“组合和聚合的‘双轴’关系是语言符号系统中极为重要的内容。

语言是一个复杂的符号系统，除双面、双轴以外，还是一种多平面的系统。所谓“多平面”主要是指音位、词、句和文本四个平面。音位是语言系统中最为基础的平面。音位是在语言中能够区别意义的最简单的语音形式，音位本身与意义无涉，属于前意义平面。由于音位不受意义的“纠缠”，可以进行纯粹的符形描述。由于它是前意义平面，能够同体现意义的词平面相联系，又成为体现意义的不可或缺的物质实体，对于研究能指和所指的关系具有重要的意义。处于音位平面之上的是词平面。词是最小的能够独立运用的有意义的语言形式。因为词是最小的语言形式，说明它是一个整体，不能再分解。如果再进行分解，它就不是有意义的语言形式了。因为词可以独立地运用，它可以成为句平面的组成成分。词是把意义引入语言形式的最为基础的平面，是用以联结音位平面和句平面的一个中间层次。句子是能够表达完整意义的语言形式，句平面是组合关系轴系上的完整单位，因而也是语言系统的核心层次。句子的意义是由组成句子的词的意义和词与词结合的语法意义共同体现的，句平面的研究包括语法、语义和语用三个方面。然而句平面也还是表达完整意义的基础平面，更为完整的意义表达则是文本（篇章）平面。文本是由句子组成的话语单位。文本成立的基本条件是讯息的连续性，更多地涉及说话主体和环境世界，在形式的规定上具有较为松弛的性质，判断“规范文本”的标准不像句法那样“正确”还是“错误”，而是“适当”或者“不适当”。比起高层次的句平面，文本平面是更高层次的复合单位。语言的符号系统，主要由音

位、词、句和文本四大平面构成，但是相邻的两个平面之间，还往往存在某种中间层次（平面），如音位与词之间有语素，词与句之间有短语，句与文本之间有句群、语段、筹等。

在语言学的研究中，有些语言学家，例如结构主义语言学派，只研究能指面的联结方式，而不关心它们的意义问题。生成语法学派的创始人、美国语言学家乔姆斯基（Chomsky, N.）在《句法结构》一书中也曾经说过：“几乎没有什么证据可以证明‘关于意义的直观’在实际研究语言形式时会有什么用处。”（中译本第 95 页。本书引文夹注凡有简而未详者，均参见书后所附‘主要参考书目’），虽然他后来也注意到意义的研究。然而从符号学的观点来看，语言符号学则必须研究能指和所指之间的关系，亦即关于语言的意义或意指作用的问题。现代语言学的语义学研究，倡导者是法国语言学家布勒阿尔（Breal, M.），他把语义学看成是一门研究词义演变的历史性科学。而美国哲学家莫里斯（Morris, C.）把符号学分为语形学（*syntactics*）、语义学（*semantics*）和语用学（*peagmatics*）三个部分，在语言符号学研究中产生了更为深远的影响。莫里斯在 1938 年《符号学一般理论》的著作中，把语形学定义为研究“符号相互间的形式关系”，语义学研究“符号与其指示的对象之间的关系”，语用学研究“符号与解释者之间的关系”。在 1946 年的《符号、语言和行为》一书中，他又重新定义了这三个术语：语形学研究“符号的种种联合”，语义学研究“符号所具有的各种方式的意谓”，语用学是“在符号出现的行为中，研究符号的起源、应用和效果”。在这里，除语形学研究符号的能指以外，语义学和语用学都着眼于符号的意义、所指研究。关于语义学和语用学的关系问题多有争议，在争论中有三种不同的观点：（1）语义原则；把语用学归入语义学；（2）语用原则；把语义学归入语用学；（3）互补原则。英国语言学家利奇（Leech, G.）赞同互补的观点，认为语义学研究“*sense*”是语言片段的抽象和字面的意义，语用学研究“*force*”是语言片段在特定语境中的意义。因此，语义学和语用学是互不相同但又相互补充的研究领域。

### 三 文字符号

文字是记录口头语言的书写符号系统，是书面语言的载体。作为一种符号，文字是记录语言、口头语言符号的符号，通常称之为二级符号。关于这一点，早在古希腊亚里士多德 Aristotle 的《工具论》里就有明确的论述。他说：“口语是内心经验的符号，文字是口语的符号。”（《解释篇》[1]16a）中国古籍《尚书》的注释者说：“言者，意之声；书者，言之记。”同样明确地指出了文字作为口语的书写符号的这一特征。然而值得注意的是，语言为听觉符号，文字在记录语言的时候，把听觉符号转换为视觉符号。由此可以看出，文字符号与语言符号又是物理性质完全不同的两个符号系统。

作为符号，语言有能指和所指两个“面”：能指面是“音响形象”（即符形，所指是它所表示的思想、概念）即符释。作为符号，文字既然也是由能指和所指构成的双面体，那么它的能指是什么呢？它的所指又是什么呢？文字符号的能指和所指，是由“文字是记录语言符号的符号”这一本质特征所决定的。文字是形（字形）、音（字音）、义（字义）三者的统一体。作为符号，能指即是字形，亦即文字符号的符形；所指是字音和字义，亦即文字符号的符释。从这里可以看出，语言符号同文字符号的能指和所指之间存在一定的对应关系。语言作为听觉符号，能指是“音响形象”，文字作为视觉符号，能指是字形，亦即文字符号的符形。语言符号的所指是它所表示的思想；文字符号采用诉诸视觉的方法记录下语言符号，包括语言符号的能指和所指，所以音和义都成了文字符号的所指，其中字音记录的是语言符号的能指，字义则是语言符号原来的所指。当然，文字记录语言并非总是一一对应，形、音、义之间往往呈现错综复杂的关系。

文字符号是记录语言符号的符号，并不意味着文字符号缺乏应有的独立性。索绪尔在《普通语言学教程》一书中，对于文字与语言之间的关系以及文字的符号体系，都有比较具体的论述。索绪尔明

白无误地指出：“语言和文字是两种不同的符号系统 后者惟一的存在理由是在于表现前者。”（第 47 页）这当然是说 文字是记录语言符号的符号 但它又是一个独立的符号系统。索绪尔认为，“语言有一种不依赖于文字的口耳相传的传统 这种传统并且是很稳固的”然而文字的威望却超过了语言。为此，索绪尔找出了文字凌驾于口语形式的四点原因：（1）词的书写形象使人突出地感到它是永恒和稳固的，比语音更适宜于经久地构成语言的统一性。（2）在大多数人的脑子里 视觉印象比音响印象更为清晰和持久 因此更重视前者。（3）语言有一种人人必须遵守的法规——正字法 而文学语言却可以有自己的词典和语法，使得文字语言更增强了文字不应该有的重要性。（4）当语言和正字法发生龃龉的时候，语言学家没有发言权，结果差不多总是书写形式占了上风。为了确认文字对于语言的特定的地位，索绪尔提出了研究“文字表现语言”这个题目的必要性。

文字符号的体系，索绪尔认为不外乎两种：表音体系和表意体系。表音文字是把词中一连串连续的声音摹写出来。表意文字的符号同词赖以构成的声音无关，而是间接地与它所表达的观念发生关系。表意文字体系的古典例子就是汉字。在《普通语言学教程》一书中 索绪尔的研究只限于表音文字体系。他在《语言的价值》一章中，从能指的角度讨论了文字系统的一些重要特征：（1）文字的符号是任意的 例如字母 t 和它所表示的声音之间没有任何关系。（2）字母的价值纯粹是消极的和表示差别的，例如同一个人可以把 t 写成好些变体 惟一要紧的是 在他的笔下 这个符号不能跟 l、d 等等相混。（3）文字的价值只靠它们在某一个由一定数目的字母构成的系统中互相对立而起作用。因为书写符号是任意的，它的形式并不重要，或者毋宁说，只在系统所规定的限度内才是重要的。（4）文字符号是怎样产生的，这完全无关轻重，因为它们与系统没有关系。人们把字母写成白的或黑的，凹的或凸的，用钢笔还是用凿子 对它们的意义来说并不重要。

在索绪尔关于语言和文字关系的论述中，文字只是为了“表现”语言，处于从属的地位，但是不会因此而有损于文字的积极作用。在人类的发展史上，语言的功用无需多说，语言的出现使人类走出了动物群，语言是人区别于其他动物的本质特征之一，是人类最重要的交际工具和思维工具。然而，声音一发即逝，人们说话要受到时间和空间的限制。为了克服这种限制，经过漫长的岁月，人们发明了文字。我国清代学者陈澧在《东塾读书记》中说：“声不能传于异地，留于异时，于是乎书之为文字。文字者，所以为意与声之迹也。”原来文字就是作为语言的“意与声之迹”而产生的，亦即文字是记录语言的书写符号。文字的产生，是人类由野蛮社会进入文明社会的一个重要标志。自从有了文字，人类除了拥有听觉符号的口头语言之外，又拥有作为视觉符号的书面语言。由于书写出来的文字可供人们反复地推敲和琢磨，因而可以使语言更加缜密、丰富和规范化。我国汉代卓越的语言学家许慎，曾在《说文解字·叙》中说过：“盖文字者，经艺之本，王政之始，前人所以垂后，后人所以识古，故曰本立而道生，知天下之赜而不可乱也。”文字的使用，大大地有利于民族文化的不断传承、不断创造、不断发展。

## 第二节 汉语符号与汉字符号

### 一 汉语符号

汉语是汉民族的语言，属汉藏语系，是世界上最丰富最发达的语言之一。汉语还是世界上使用人口最多的语言，当代说汉语的人数约占世界人口的五分之一。

汉语有古汉语和现代汉语的区别。汉语的历史也同汉民族的

历史一样悠久，最早的有史时期汉语称为古汉语。汉民族在历史上长期用‘文言’作为统一的书面语。这种书面语最初必定是建立在口语基础上的。后来它跟口语的距离越来越远，到唐宋时代，又产生了一种接近口语的书面语言——白话。白话是在北方方言的基础上形成的。现代汉语由古汉语发展而来。现代汉民族的共同语是以北京语音为标准音，以北方方言为基础方言，以典型的现代白话文著作作为语法规范的普通话。

汉语符号具有语言符号的一切特征。汉语符号也是这样的语言符号系统：由能指和所指构成双面体，由组合轴和聚合轴构成双轴关系，由音位一词一句一文本等构成多平面的符号系统。但是作为汉藏语系的主要语言，它与其他语系的语言，如印欧语系的英语、法语、德语等相比，汉语符号存在一系列明显的特征。汉语符号的主要特征是：

在音位平面上，汉语普通话共有 10 个元音音位：[a]、[o]、[ɤ]、[ɛ]、[i]、[u]、[y]、[ɿ]、[ʅ]、[ɔ]；22 个辅音音位：[p]、[pʰ]、[m]、[f]、[t]、[tʰ]、[n]、[l]、[k]、[kʰ]、[ŋ]、[x]、[tɕ]、[tɕʰ]、[ç]、[tʂ]、[tʂʰ]、[ʂ]、[z]、[ts]、[tsʰ]、[s]。元音都有音位变体；辅音大多数也有音位变体。普通话还有 4 个声调音位：阴平、阳平、上声和去声。除阴平、阳平以外，上声和去声都有音位变体。元音音位之多和存在声调音位，是汉语音位平面的重要特点。

在词平面上，古汉语里绝大多数是单音节词，现代汉语一般是单音节的实语素与实语素构成的合成词。它们不像印欧语那样，很多词是由词根加上前缀、后缀之类构成的。汉语有些词，如“老虎”、“桌子”之类，像是有前缀、后缀，但是为数不多。汉语的词缺乏印欧语里名词、动词、形容词那些性、数、格、时、人称的变化。

在句平面上，汉语不像印欧语那样，词类和句子成分之间有一种简单的一一对应关系。如印欧语里的名词同主语、宾语对应，形容词同定语对应；而汉语的名词可以充当主语和宾语，也可以充当其他成分，形容词可以充当定语，同样可以充当其他成分。印欧语

里句子和短语有不同的构造；而汉语的主谓短语，既可以充当句子的一个成分，也可以独立成句。汉语的语序是表达句法意义的重要手段。

在文本平面上，汉语符号学的研究刚刚起步，汉语文本的符号学特征还有待于学者们的概括和总结。

## 二 汉字符号

汉字是记录汉民族口语的书写符号系统。汉字是世界上历史最悠久的文字，也是当今世界使用人口最多的文字。作为符号系统，汉字还是世界上独一无二的一种文字符号类型。汉字符号的特征，主要有以下几点：

汉字符号最明显的特征是方块形。汉字同英文、法文、德文等印欧语的文字相比，后者为线性的拼音文字，而汉字则是方块状的“拼形”文字。拼音文字与“拼形”文字在能指上有显著的区别。拼音文字的能指，是把听觉上声音的前后相继转换为视觉上字母的线条形状；而汉字属于非字母书写系统，它的能指是由笔画按一定规则组合成的方块形状，一字一方块。汉字符号系统的初始符号是笔画，由笔画组合成独体字或偏旁，偏旁再按一定的规则组成合体字。笔画、偏旁和它们的组合规则，构成了汉字符形的结构系统。汉字符号长时期保持方块形，这与汉语符号的特点有一定的关系。在古代，汉语的词以单音节占优势，没有形态变化，缺乏词缀等附加成分，这些特点有利于保持一词一字的书写体系。现代汉语里复音词大量涌现，如“衣服”、“妻子”、“自行车”、“艺术性”等等，可以用两个或三个方块字表示，保留方块字并没有遇到多少困难。现代汉语里“子”、“儿”、“头”等附加成分同样可以用方块汉字表示。

汉字符号最本质的特征是它属于表意的文字体系。方块形并非汉字所独有，例如朝鲜文字也呈方块形，但它是方块表音文字。所以严格地说，方块汉字应是方块表意文字。表音文字的特点是把词中一连串的声音摹写出来，它有时是音节的，有时是字母的。而

汉字不然。作为表意文字，汉字与语音不发生直接的联系，每一个字的能指只表示一个音节，不能明确地记录声音，但它能够表示一个意思。这里所说的是“汉字与语音不发生直接的联系”而不是说它们没有任何联系。汉字首先是能读，通过读音确定所表示的相应的词，这样才谈得到字义问题。但是汉字的读音不是来自对语音的直接摹写，所以不能算是表音文字。汉字有形声字，形声字还确有“表音”部分，称做“声符”。其实形声字的“声符”本身原来也是个表意符号。如“晴”、“昭”、“晨”等形声字，其声符“青”、“召”、“仄”等在离开它们的意符“日”的情况下，仍然是个表意字，归根结蒂，还是与语音不发生直接的联系。有人说汉字是“意音文字”，虽有道理，但仍是不够确切的。

汉字符号的第三个特征是非任意性和任意性的统一。表音文字的符号是任意性的，如 t、w、l 等字母形体，的确同它们所表示的声音之间没有任何关系，只要大家承认，完全可以写成另外的样子。但是，作为表意符号的汉字符号就不同了。汉字的每一个符形都有它的来源，符形如此而不如彼，并不是那么随意的。例如“旦”字，上部是“日”，下部是地面或者水面，从地面或水面上升起太阳，是表示早晨的意思。比如说，把“日”换成“月”或其他，那么它就起不到表示“早晨”的作用了。当然，汉字作为表意文字，其符形也不是无任何任意性可言，就是说，汉字的符形与它所表之意的关系，也不是必然地如此而不如彼。比如“旦”字，为什么用地面或水面上升起太阳来表示早晨呢？也许这样地表意最为简单，但不会是惟一的。又如“男”字，从田从力，“言男用力于田也”，这当然有道理，但也不具有必然性或惟一性。所以说，汉字符号具有一定的非任意性特点，但也具有相对的任意性，是非任意性和任意性的统一。

汉字符号的另一个特征，是具有极强的稳定性。任何一种文字都有一定的稳定性，但是汉字更加突出。汉字的这一特征是同汉字的表意性质相关的。表音文字记录的是语言的声音，说和写相距不会太远，语音变了，拼写法也往往随之变化。例如古代拉丁语发展

到现代的法语、西班牙语、意大利语等语言 作为文字的拉丁文也随之发展为法文、西班牙文、意大利文等等。汉字符号与它们有些不同。汉字作为表意文字，字形不与语音发生直接的联系，尽管语音变了，字形却可以不变。从古汉语到现代汉语，语音发生了很大的变化，可是方块汉字并没有改变。虽然汉字的形体经历了从甲骨文到篆书、隶书、楷书等诸多变化 但仍然是同一个字。不同地方的人认识同一个汉字 读音可以不同 甚至可以用方音去读汉字的古籍，人们因此说汉字具有超方言的性质。索绪尔说：“对汉人来说 表意字和口说的词都是观念的符号 在他们看来 文字就是第二语言。”（第 51 页）这段话需要讨论。汉语里一个口说的词同它相应的方块汉字，它们的所指是同一观念，但汉字仍然是记录汉语的符号，或者说 是记录汉语符号的符号。汉字只是具有超方言的特点 不是“第二语言”。汉字符号这种超越语音的特点，还曾经被境内少数民族以及朝鲜、越南、日本等邻国借去记录他们本民族的语言 在历史上发挥了特殊的文字符号作用。其他文字中虽然也有某种超越语音现象 但像汉字如此突出而且长期稳定 实为罕见。

最后，汉字符号还有一个显而易见的特征，就是它的符形繁杂而造成学习和使用上的许多困难。汉字符号的这一特征，也同它的表意特点相关。一是难读。汉字中的纯表意字，没有表音成分，必须一个一个地去认去记 形声字虽有表音成分 也由于语音的演变，相当一部分已经不能准确地表示读音。二是难写。汉字笔画繁多，结构复杂，不容易书写。相当数量的汉字在十画以上，有的多达几十画。这些繁简各异的符形，要求正确而又匀称地写在同样大小的方块里 绝不是一件容易的事情。三是难记。汉字难读、难写 加之数目庞大 也就必然难记。正如清末学者梁启超所说：“中国文字奥博 字多至于数万 通儒不能遍识。”

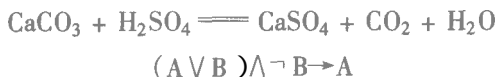
符号是“人类世界的一个普遍工具”，汉字符号自然也是“工具”。一个民族发展到一定的文明程度，必然地要选择某种文字符号工具，但是选择什么样的文字符号为工具并非是必然的，虽然都

有它们各自的历史文化原因。方块形表意的汉字符号，是汉民族历史和文化的选择。

### 第三节 独特的汉字编码

#### 一 编码

“编码”是符号学的重要术语，频繁地出现在现代符号学的论著之中。直观地看，“编码”一词不难理解。比如说，下面的化学方程式和数理逻辑公式是编码，当无问题：



前者属于化学编码，公式表示把大理石（碳酸钙）和硫酸混合在一起，将得到硫酸钙、二氧化碳和水；后者属于逻辑学编码，是一个析取式推理公式。这样的符号公式，在皮埃尔·吉罗《符号学概论》的《逻辑编码》一章中，称之为“形象化编码”，属于图像符号的编码。

“编码”(code)一词，从词源上说，是将法律条文按一定分类和次序排列在一起，以免彼此冲突或重复。以后这个词逐渐抽象为“规则或系统”的意思。然而，在应用中，“编码”除名词的用法以外，还有动名词或动词的用法。即使同为名词用法，有时候还有相关而又有所区别的含义。具体说来，编码问题可以作以下一些理解：

编码是一种符号行为。在通常的理解中，人们容易把符号简单地看成符号的能指，而在实际上，符号是由能指和所指组成的双面对体，能指是表达面，所指为内容面。人们也容易把编码看成只属于能指或表达面的事情。实际上，编码是在特定的符号对象领域里，

应用某种编码规则把能指和所指结合起来，并在能指和所指的关系上体现符号的意指作用。在这里，规则是至关重要的，没有规则便没有编码。符号行为按照规则编织出来的“码”不妨称之为“代码”或“符码”当然也可以简单地说成“编码”如果不被误解的话。代码或符码是符号的符形实体或序列。

编码是一个系统。编码是关于符号单元以及符号之间联结规则的系统，例如语言符号的编码系统和非语言符号的编码系统等。符号必须组成系统才能产生意指作用，而系统又必须借助于编码才能成立。因此可以说，符号的意指作用就是建立在编码的基础之上。在系统中，符号间相互依存。有不少符号的组合难以看得出系统，因之在表面上为非系统的符号组合中确定其系统的存在，就是符号学的一项重要任务。

编码是一个过程。编码过程是符号使用者之间的一种约定，这些使用者承认能指与所指之间的关系并在使用中遵守这种关系的规则。由于编码是一个过程，因而都有一定的程序，不同符号系统有不同的编码程序。例如汉语编码和汉字编码、汉字编码和电脑的汉字编码、汉字字符编码和汉字书符编码等等，它们的编码程序各不相同。编码的过程也可以称之为“码化”过程。

不同的编码程序决定了不同的编码方式。编码可以根据不同的方式分为不同的编码类型，例如语言编码和非语言编码，认知编码和情感编码，逻辑编码和美学编码等等。不同编码方式的确定性程度不同，它们编码化的程度也不相同。如逻辑编码化的程度高，而美学的编码化程度就比较低。

同编码相关的还有术语“解码”。解码是编码的逆过程，没有编码，就无所谓解码。如果说编码属于表达，那么解码则属于理解。表达者运用符号的编码传达某一讯息，理解者接收这一讯息，必须对其解码，也就是根据符号的编码重建意指。编码者根据规则，把能指和所指结合成符号或符号系统，所指即是“释义”。解码者根据符号的能指重建符号的所指，这也是释义。但是这编码中的释义并

不等同于解码中的释义，它可能同符号本来的所指相一致，但也可能不相一致。解读一篇文艺作品或者欣赏一幅书法作品，都是解码所谓“仁者见仁 智者见智”不同解码者所释之“义”有时候可能相距甚远。

## 二 汉字编码的特征

文字作为一种符号系统，有自己的编码方式。按照索绪尔的观点，文字惟一存在的理由只是表现语言。然而文字和语言毕竟是两个不同的符号系统，语言属于听觉符号系统，而文字属于视觉符号系统。它们的编码方式也因此有所不同：文字属于非语言编码。

汉字是一种独特的表意文字，作为一种非语言编码，有自己独一无二的编码系统。汉字编码系统有以下几个主要特征：

汉字编码是以形表意的编码。法国语言学家马丁内 (Martinet, A. 提出语言的双层分节理论 第一层可以切分出意义单位 为词素 第二层切分出语音单位 即音位。表音的字母文字的编码 作为分节语言的替代，其作用是当分节语言的使用受到时间或空间限制的时候代替分节语言。它们是分节语言的实录。而表意的汉字编码 却不是分节语言的简单替代 而是以形表意。汉字“六书”中的象形、指事和会意编码，都是直接的以形表意编码；形声字如前所述 除“形符”表意外；“声符”原来也是个以形表意字。汉字独特的编码方式使得汉字符号具有颇为突出的超语音的特点，使得使用汉字的人们不管天南地北，都可以通过汉字符号来彼此交流思想，但未必能够听得懂彼此的方言。

汉字编码是语音和图形的双重编码。如前所述，作为表意的方块汉字，同语音没有直接联系但不是没有任何联系。一个汉字首先必须能读 即通过读音来确定它所表之“意”。例如“人”字尽管提示了人的图形 但还是通过“rén”的读音同它的意义发生关系。同拼音文字相比 那就是汉字“人”的编码除读“rén”这个声音以外 还提示了“人”的形象 虽然汉字的读音编码只是间接而非直接的。在形