

# 现代汉字学纲要

苏培成 著

北京大学出版社

北 京

# 新登字(京)159号

## 图书在版编目(CIP)数据

现代汉字学纲要/苏培成编著.-北京:北京大学出版社,1994.

10

ISBN 7-301-02597-1

现... 苏... 汉字学-概论 .H12

出版者地址:北京大学校内

邮政编码:100871

排版者:北京大学印刷厂激光照排车间

印刷者:北京大学印刷厂

发行者:北京大学出版社

经销者:新华书店

850×1168毫米 32开本 6.75印张 165千字

1994年10月第1版 1994年10月第1次印刷

定 价:6.75元

# 前 言

近十多年来,汉字学的一个新分支——现代汉字学悄然兴起,人们似乎突然认识到这是一块大有开发前途的处女地。现代汉字学的研究对象是现代汉字。查一查语言学文献,早在1952年8月号的《中国语文》上就有丁西林先生的《现代汉字及其改革的途径》,这篇文章中已经使用了“现代汉字”这个术语。可是在其后的很长时间里,现代汉字的研究并没有引起人们的注意。直到1980年,周有光先生在《语文现代化》丛刊第二辑上发表了《现代汉字学发凡》之后,随着改革开放大潮的迅速发展和汉字信息处理的迅速发展,现代汉字学才找到了适宜的土壤,正如周先生所说的那样“嫩黄新绿渐见于今日”。1991年8月,在北京举行了“海峡两岸汉字学术交流会”,我提交给会议的论文是《现代汉字研究简述》,意在向与会的海峡两岸的学者介绍现代汉字学这个新生儿。这几年,研究现代汉字的论文多起来了,大学用的“现代汉字学”专业教材也出版了,真是令人高兴。不久前我读到台湾大学黄沛荣教授的新著《汉字的整理与统合》,在这本研究报告中三次出现了“现代汉字”这个术语。黄教授参加了1991年的“海峡两岸汉字学术交流会”和1992年在北京举行的“海峡两岸汉字学术座谈会”。这说明现代汉字的研究已经引起了他的注意。

什么是现代汉字?有人理解为目前正在使用的这套汉字,也有人理解为记录现代汉语用到的字。这两种解释不完全相同,但是都表明它和我们的现实文字生活密不可分。随便想一想就可以提出一连串的问题:出版印刷要用多少字,都是哪些字?九年义务教育要教多少字,都是哪些字?汉字的笔画有多有少,哪几种笔画

负载的汉字数量最多,为什么?汉字在学习和书写时,要对字形做出分析,应该怎么分析?多数汉字是由几个小的构字单位组成的,组合时有没有理据可言?要把成千上万个汉字排成一个系列,根据什么原则,是字音、是字形、还是什么别的?用什么办法可以比较方便地把汉字输入电脑?如何评价汉字简化的利弊得失?如何使汉字变得更加合理更加规范,使学习和使用都比较容易?汉字的前途是什么,要不要改为拼音文字?等等,等等。现代汉字学要对这些问题做出科学的回答。只要现实生活中这些问题还存在,现代汉字学就有存在的价值,就有用武之地。现代汉字学的研究已经做了许多工作,可是还很不够,因为它太年轻,不成熟。和传统汉字学相比,它的底子很薄,基本理论和基本方法还不很完备,需要有更多的学者关心它扶植它。相信它一定有远大的前途。

我于1988年秋开始为北京大学中文系汉语专业讲授“现代汉字学”这门课程,其后又讲了几次,每讲一次都要对讲稿做些补充和调整。1993年下半年,用了几个月的时间把历年来所用的讲稿做了全面的补充和改写,就成为现在这本教材。

我在讲课和写作这本教材时,参考了许多前辈的著作,并从中引用了一些观点和材料,在此向这些先生们表示深深的感谢。我在讲课和写作这本教材时,还得到许多师友的关心和帮助,在此向他们致以诚挚的敬意。

这本书讨论现代汉字,涉及已经停用的繁体字、异体字和旧字形以及古字,为排版增加了不少麻烦。这里特向北京大学印刷厂激光照排车间的同志表示感谢。

我自己水平不高,对现代汉字的研究刚刚起步。书中定有不妥之处,敬请读者批评指正。书末所附的参考文献疏漏也很多,敬希各位专家见谅。

作者

1993年岁末于北大

# 目 录

第一章 汉字概说.....	1
1.1 汉字的性质 .....	1
1.2 汉字的特点 .....	4
1.3 汉字和汉语的关系 .....	7
1.4 汉字的总字数 .....	9
1.5 汉字和汉文化.....	10
1.5.1 汉字是汉文化的一部分 .....	10
1.5.2 汉字具有丰富的文化内涵.....	11
1.5.3 以汉字为本体产生的文化现象 .....	13
1.6 汉字文化圈在过去和现在.....	13
1.6.1 汉字文化圈在过去 .....	13
1.6.2 汉字文化圈的现在 .....	16
第二章 现代汉字和现代汉字学 .....	19
2.1 现代汉字.....	19
2.2 现代汉字学.....	22
2.2.1 研究现代汉字的意义 .....	22
2.2.2 现代汉字研究简史 .....	23
2.2.3 现代汉字学研究的内容 .....	25
第三章 现代汉字的字量 .....	28
3.1 现代汉字的字频统计与分析.....	28
3.1.1 《语体文应用字汇》 .....	28
3.1.2 《汉字频度表》和《汉字频度统计》 .....	29

3.1.3	《现代汉语频率词典》中的《汉字频率表》	31
3.1.4	《现代汉语字频统计表》	32
3.1.5	汉字效用递减率	33
3.1.6	常用字笔画趋简率	35
3.2	现代汉字的总字数	35
3.2.1	《现代汉字用字全表》的研制	35
3.2.2	现代汉字的来源	38
3.3	通用字和常用字	39
3.3.1	通用字,常用字,罕用字	39
3.3.2	选取通用字和常用字的标准	40
3.3.3	通用字和常用字的字量	41
3.3.4	《现代汉语常用字表》	43
3.3.5	《现代汉语通用字表》	43
3.4	专门用字	44
3.4.1	扫盲用字表	44
3.4.2	小学用字表	45
3.4.3	汉语水平·汉字等级表	46
3.5	专业用字	46
3.6	汉字构词能力的分析	47
3.7	限制和减少汉字字数的研究	48
3.7.1	洪深的“基本汉字”	49
3.7.2	翟健雄的“音标字”	50
3.7.3	《1469个精简汉字表》	50
3.7.4	郭沫若的《日本的汉字改革和文字机械化》	52
3.7.5	林汉达的《一万个通用汉字能不能砍去一半》	53
3.7.6	周有光的《限制和减少现代汉语用字》	53
3.7.7	“标准汉字”的设计	53
3.7.8	郑林曦的《精简汉字字数的理论和实践》	54

第四章 现代汉字的字形 .....	55
4.1 现代汉字字形分析 .....	55
4.1.1 部件分析法 .....	55
4.1.2 笔画 .....	55
4.1.3 部件 .....	63
4.1.4 整字 .....	69
4.2 现代汉字的构字法 .....	72
4.2.1 意符, 音符, 记号 .....	72
4.2.2 独体表意字 .....	73
4.2.3 会意字 .....	74
4.2.4 形声字 .....	76
4.2.5 半意符半记号字 .....	77
4.2.6 半音符半记号字 .....	78
4.2.7 独体记号字 .....	79
4.2.8 合体记号字 .....	80
4.3 现代汉字的理据性 .....	81
4.3.1 文字的理据性 .....	81
4.3.2 现代汉字的理据性 .....	82
4.3.3 理据度及其计算 .....	82
4.3.4 理据的溯源分析 .....	83
4.4 现代汉字的简化 .....	84
4.4.1 汉字发展中的繁化与简化 .....	84
4.4.2 汉字简化运动的兴起及发展 .....	86
4.4.3 中华人民共和国建立后的汉字简化工作 .....	89
4.4.4 汉字简化的方针、方法和效果 .....	91
4.5 现代汉字的整理 .....	94
4.5.1 字形整理的要求 .....	94
4.5.2 整理异体字 .....	95

4.5.3	整理印刷铅字字形 .....	98
4.5.4	改换生僻地名用字 .....	100
4.5.5	统一部分计量单位用字 .....	101
4.5.6	汉字的进一步整理 .....	102
<b>第五章 现代汉字的字音</b> .....		<b>104</b>
5.1	字音的确定 .....	104
5.1.1	字音的性质 .....	104
5.1.2	以北京语音为标准音 .....	105
5.1.3	和古反切对应的传统读音与现代实际口语读音 .....	106
5.2	异读字 .....	108
5.3	同音字 .....	110
5.4	多音字义字 .....	113
5.4.1	多音多义字的数量 .....	113
5.4.2	多音多义字的音级 .....	113
5.4.3	多音多义字的来源 .....	114
5.4.4	多音多义字的读法 .....	115
5.4.5	减少多音字的两种方法 .....	116
<b>第六章 现代汉字的字义</b> .....		<b>119</b>
6.1	字义与语素义 .....	119
6.2	字义的增加与减少 .....	121
6.3	字形和字义的关系 .....	123
6.4	单义字与多义字 .....	126
6.4.1	单义字和多义字的数量 .....	126
6.4.2	单义字的主要类型 .....	126
6.4.3	多义字的主要类型 .....	127
6.4.4	多义字的使用 .....	128

6.5	同义字 .....	128
6.6	字义和词义 .....	130
第七章 现代汉字的字序.....		132
7.1	古代的字序法 .....	132
7.2	部首法 .....	134
7.2.1	立部和归部 .....	134
7.2.2	《新华字典》的部首 .....	135
7.2.3	《辞海》(1979年)的250部 .....	135
7.2.4	《汉字统一部首表》的201部 .....	136
7.2.5	《汉语大字典》和《汉语大词典》的200部 .....	137
7.2.6	部首法的标准化 .....	137
7.3	音序法 .....	139
7.4	号码法 .....	141
7.5	笔画笔形法 .....	142
第八章 汉字的熵和多余度.....		143
8.1	信息量 .....	143
8.2	汉字的熵 .....	144
8.3	汉字的多余度 .....	146
第九章 现代汉字的应用.....		149
9.1	汉字应用两个界面 .....	149
9.2	人际界面的主要应用——记录现代汉语 .....	149
9.3	人际界面的其他应用 .....	150
9.3.1	书法与篆刻 .....	150
9.3.2	印刷,打字,电报 .....	151
9.4	现代汉字的正字法 .....	151

9.5	纠正社会用字的混乱现象 .....	152
9.6	现代汉字的标准化 .....	155
<b>第十章</b>	<b>汉字和信息处理.....</b>	<b>159</b>
10.1	中文信息处理.....	159
10.2	汉字的输入.....	160
10.3	汉字标准交换码.....	163
10.4	汉字点阵.....	164
<b>第十一章</b>	<b>台湾地区使用汉字的情况.....</b>	<b>166</b>
11.1	字量.....	166
11.2	字形.....	167
11.2.1	印刷用标准字形 .....	167
11.2.2	海峡两岸现行汉字字形的比较 .....	169
11.2.3	关于简化汉字的讨论 .....	170
11.2.4	《标准行书范本》 .....	172
11.3	字音.....	173
11.3.1	海峡两岸字音的比较 .....	173
11.3.2	台湾使用的注音工具 .....	173
11.4	字序.....	174
11.5	计算机用字.....	174
11.6	增强了解,促进统一 .....	175
<b>第十二章</b>	<b>汉字的评价和前途.....</b>	<b>176</b>
12.1	汉字的评价.....	176
12.1.1	古代对汉字的评价 .....	176
12.1.2	近百年来对汉字的评价 .....	176
12.1.3	“文革”后有关汉字评价的讨论 .....	183

12.1.4	科学地评价汉字 .....	185
12.2	新中国的汉字改革.....	186
12.2.1	汉字改革分两步走的设想 .....	186
12.2.2	新时期语言文字工作的方针和任务 .....	189
12.3	对拼音化的三种理解.....	192
12.3.1	汉语拼音作为辅助汉字的工具 .....	192
12.3.2	汉字和汉语拼音并存并用 .....	192
12.3.3	用拼音文字代替汉字 .....	193
12.4	对汉字前途的设想.....	195
参考文献.....		196

# 第一章 汉字概说

## 1.1 汉字的性质

文字是记录语言的符号体系,这是世界上一切文字的共性,它正确地说明了文字和语言的关系。某种具体文字的性质,比如汉字的性质,指的是该种文字区别于别种文字的本质特征,是文字类型学要研究的问题。正确认识一种文字的性质,是正确解决有关这种文字种种问题的基础。

文字是用来记录语言的,所以要联系语言来研究文字。确定文字性质的标准只有一个,就是看这种文字记录的是什么样的语言单位。

语言是个层级体系。它的下层是音位和音节,上层是音义结合的语言符号列序。这一层又分为若干级,第一级是语素,第二级是词,第三级是句子。世界上的文字有许多种,从记录的语言单位来看主要有:记录音位的音位文字,记录音节的音节文字,记录语素的语素文字。

音位文字和音节文字都是拼音文字。拼音文字的基本单位是字母,语素文字的基本单位是单字。拼音文字的字母和语素文字的单字是很不相同东西。

音位文字有许多种,如国外的英文、俄文,国内的藏文、壮文。一种语言的音位数目是有限的,不过几十个,音位文序的字母数量因而也是有限的,也只有几十个。这些代表音位的字母,按照一定

的规则拼合起来,就拼出该语言的词的声音。阅读用这种文字写的读物,按照拼出来的声音,就知道它表示的是语言里的哪个词。拼音字母所表示的音位,就是这个字母的音值。每个字母都有确定的音值,而没有意义。一种历史较久的音位文字,按照字母组合拼出来的读音,有时和实际语言中的读音有些差别。这是由于拼写法有保守性和实际语言脱节而造成的,如现代的英文和藏文。

语素文字的代表是汉字,汉字记录的语言单位是语素。语素也叫词素,是“最小的语音语义结合体”。从语素和词的关系看,语素分为成词语素和不成词语素两大类。成词语素独立应用时就是词。作为语素文字的汉字,它用不同的形体表示不同的语素,不但能区分读音不同的语素,而且还能区分读音相同的语素。例如“蝗、徨、惶”都读 huáng,但是意义不同,是三个不同的语素,就写成三个不同的汉字。当某个汉字被用来表示某个语素时,该语素的音和义就成了这个汉字的音和义。因为汉字的形体和它所表示的语素的音和义紧密结合在一起,所以人们常说汉字是形音义的统一体。有些汉字不记录语素,只记录语素的一部分,这样的汉字只有形体和读音,没有意义,不是形音义的统一体,如“葡、萄、徘、徊、蟋、蟀”。这样的字数量很少,不反映汉字的本质。这样一些字的存在不影响汉字是语素文字的论断。

汉字是语素文字,这是许多学者经过多年探索得出来的科学结论。赵元任把汉字叫做词素文字。他说:“在世界上通行的能写全部语言的文字当中,所用的单位最大的文字,不是写句、写短语的,是拿文字一个单位,写一个词素,例如我们单独写一个‘毒’的字形,来写‘毒’这个词素。”“用一个文字单位写一个词素,中国文字是一个典型的最重要的例子。”他跟世界多数其他文字的不同,

---

吕叔湘《汉语语法分析问题》,《汉语语法论文集》(增订本)商务印书馆 1984 年版第 489 页,第 491 至 492 页。

不是标义标音的不同,乃是所标的语言单位的尺寸不同。”赵元任说的“所标的语言单位的尺寸不同”,也就是所记录的语言单位大小不同。赵元任讲的词素,为英语 morpheme 的意译。近几十年中,中国大陆许多学者把 morpheme 译为语素,所以赵元任的词素文字也就是我们说的语素文字。

吕叔湘说:“世界上的文字,它的形式是多种多样的,但是按照一定的原则来分类,也就是按照文字代表语言的方式来分类,可以分成三类。一类是音素文字,一个字母代表一个音素(又叫做音位)。英语、法语等等所用的拉丁字母(罗马字母),俄语、保加利亚语所用的斯拉夫字母,都是音素文字。第二类是音节文字,一个字母代表一个音节,就是辅音和元音的结合体。日语的字母(假名)、阿拉伯语的字母,都属于这一类。音素文字和音节文字都是拼音文字,拼音文字的字母原则上都是没有意义的,有意义是偶然的例外。第三类文字是语素文字,它的单位是字,不是字母,字是有意义的。汉字是这种文字的代表,也是唯一的代表。汉字以外的文字都只是形和音的结合,只有汉字是形、音、义三结合。”吕先生说的音素文字也就是我们说的音位文字。

朱德熙说:“文字是记录语言的。就汉字跟它所要记录的对象汉语之间的关系来看,汉字代表的是汉语里的语素。例如 [tàn k u] 的 [k u]、[k u iào] 的 [k u] 和 [tiào k u] 的 [k u] 读音相同,意思不一样,是三个不同的语素,分别由三个不同的汉字‘糕、膏、高、’(蛋糕、膏药、跳高)来表示。从这个角度看,汉字可以说是一种语素文字。”

---

赵元任《语言问题》,商务印书馆 1980 年版第 142、144 页。

吕叔湘《汉语文的特点和当前的语文问题》,《语文近著》上海教育出版社 1987 年版第 142 页。

朱德熙《汉语》,《语法丛稿》上海教育出版社 1990 年版第 198 页。

持这种观点的还有别的学者,这里不再征引。

与语素文字说相近的有语素—音节文字说。叶蜚声、徐通锵说:“汉字”是一种语素—音节文字,即每一个汉字基本上记录语言中的一个单音语素;少数语素不止一个音节,只能用几个字表示,但每个字记录一个音节,如‘玻’‘璃’‘徬’‘徨’等。”语素—音节文字也可为叫做音节—语素文字,两个名称的含意是相同的。

## 1.2 汉字的特点

汉字的特点指的是汉字区别于别种文字的不同之点,主要是区别于拼音文字的不同之点。文字的性质和特点的关系,是纲和目的关系。文字的性质是纲,而特点是由性质派生出来的目。抓住了文字性质这个纲,就可以说明该种文字的特点这个目。

汉字的特点可以从不同的角度去认识。比如,从书法的角度,汉字是可以成为书法艺术的文字;从神经心理学角度,有人提出汉字是复脑文字。这些看法都有价值。下面从作为记录汉语的符号体系的角度提出它的几个主要特点。

第一,数量繁多。一种语言的语素总要有相当的数量,因此语素文字的基本单位也必然要有相当的数量。国家语言文字工作委员会和中华人民共和国新闻出版署联合发布的《现代汉语通用字表》,收通用字 7000 个。这只是现代汉字通用字的数量,现代汉字总字数自然要比这个数目多。如果说的是自古以来历代积累下来的汉字的总字数(包括数量很多的异体字),已经达到六万。而拼音文字一般只有几十个字母,比汉字的数量少得多。此外,使用拼音文字的语言,词增加了可是字母数并不增加。汉字则不然,语素

---

叶蜚声、徐通锵《语言学纲要》,北京大学出版社 1981 年版第 164 页。

尹斌庸《给汉字“正名”》,《中国语文通讯》1983.6。

增加了有时就要增加新字。例如:化学家发现了氢的三种同位素,相应地也增加了三个新字,就是“氕”“氘”“氚”。

第二,结构复杂。这和数量繁多是连在一起的。成千上万个汉字在形体上都要有区别,要有一定形差度,结构必然要复杂。现代汉字结构的复杂性,表现在两个方面。一方面是结构单位多:(1)笔形多。据研究,汉字的基本笔形有6种,派生笔形有25种。如果从书写上来看,全部笔形大约要有80多种。(2)末级部件多。据统计,《辞海》(1979年)的11834个汉字共有末级部件648个。第二方面,组合方式多。例如,由相离的两画组成的部件就有“二 FDA# 卩 八儿 一 ; 刁”等多种形式。又如,合体字的结构类型多。从大的类型看,虽然可以概括为上下、左右、内外等几类,实际上每一类中又有许多细微的不同。“古苦营薑森垒”虽然都是上下结构,但又有差别。据统计8075个通用字就有250种形式。

第三,使用多种构字法。古代学者总结了古汉字的造字规律,提出了六书理论,隶楷以来的汉字在造字规律上又有新的发展。现代汉字的一级部件可以分为意符、音符和记号三大类,把这三大类部件进行组合,可以归纳出七种构字方式。这和拼音文字很不相同。拼音文字只有字母组合的方式,没有像汉字这样的部件组合方式(拼音文字也没有部件这个结构单位)。汉字在构字上意符和音符有理据性,而拼音文字只有音位组合的理据性。

第四,是二维的平面文字。从现在能看到的最早的汉字直到现代汉字,都是二维的平面。汉字的各种结构成分都是通过横向

---

张静贤《现代汉字笔形论》,《第二届国际汉语教学讨论会论文选》北京语言学院出版社1988年版第499页。

杜子劲《汉字在书写上的缺点》,《中国语文》1954.4。

傅永和《汉字结构及其构成成分的分析 and 统计》,《中国语文》1985.4。

杜定友《方块字的怪组织》,《中国语文》1954.12。

和纵向两个座标来展开的,这就使笔画的组合样式更为丰富。“举例来说,两横两竖四个笔画,如果作线性排列,只可能有一一丨丨等六种组合形式;如果作双向组合,则可以有十十、 等难以计数的组合形式。这是因为双向组合比单向组合多了一种分布信息的渠道。”正因为这样,就给众多汉字的构形开辟了广阔的天地。”而音位文字绝大多数是线性的,也有呈方块形的,如朝鲜的谚文,但是汉字的方块和谚文的方块又有不同。谚文方块里是拼音字母,而汉字方块里的是笔画和部件。汉字的这个特点,使汉字的排列可横可竖,使汉字的书写可以成为一种艺术——书法。

第五,具有一定的超时空性。朱德熙说:“汉字最大的长处就是能够超越空间和时间的限制。古今汉语字音的差别很大。但由于字义的变化比较小,而且两千年来字形相当稳定,没有太大的变化,所以先秦两汉的古书今天一般人还能部分看懂。如果古书是用拼音文字写的,现代人就根本无法理解了。有些方言语音差别很大,彼此不能交谈,可是写成汉字,就能互相了解,道理也是一样的。”

在超越时间方面,朱先生只是说“一般人还能部分看懂”,并没有说全能读懂,因为要全懂,还要具备别的条件,还要经过专门的学习。在超越空间方面,操不同方言的人不能口谈,但是可以笔谈。这是因为在识字时已经学习了用汉字写下来的共同语。如果用汉字写的是对方不懂的方言,笔谈还是有困难。例如,《官场现形记》第八回有这样一段话:“阿姨,耐说格阿是老爷?前埭老爷屋里做生日,叫倪格堂差,屋里向几几化化红顶子,才勒浪拜生日,阿要显焕!”不懂得吴方言的人尽管字都认识,可是意思还不懂。

---

范可育《我对现代汉字字形的认识》,《汉字问题学术讨论会论文集》语文出版社1988年版第52至53页。

朱德熙《汉语》,《语法丛稿》第201页。