

5

A12-53
SP1



XIANDAI HANZIXUE CANKAO ZILIAO

现代汉字学 参考资料

苏培成 选编

北京大学出版社

图书在版编目(CIP)数据

现代汉字学参考资料/苏培成编. —北京:北京大学出版社,
2002. 1

ISBN 7-301-00281-5

I. 现… II. 苏… III. 汉字-文字学-文集 IV. H12-53

中国版本图书馆 CIP 数据核字(2001)第 096093 号

书 名: 现代汉字学参考资料

著作责任者: 苏培成 选编

责任编辑: 胡双宝

标准书号: ISBN 7-301-00281-5/H·0028

出版者: 北京大学出版社

地 址: 北京市海淀区中关村北京大学校内 100871

网 址: <http://cbs.pku.edu.cn>

电 话: 出版部 62752015 发行部 62754140 编辑部 62752028

电子邮箱: zpup@pup.pku.edu.cn

排 版 者: 北京大学印刷厂

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 毫米×1168 毫米 32 开本 11 印张 276 千字

2001 年 12 月第 1 版 2001 年 12 月第 1 次印刷

定 价: 16.50 元

前 言

苏培成

一、什么是现代汉字？

现代汉字有两种含义：1. 现代出版物上用到的字，不论常用字或不常用字，不论白话用字或文言用字，都算现代汉字。2. 书写规范化现代汉语所必须用到的字，才算现代汉字。文言用字、古器物用字，即使写进今天的文章，也不算现代汉字。

本书编者采用上述的第二种含义，现代汉字就是现代汉语用字，也就是现代白话文用字。

二、现代汉字有什么用？

现代汉字的使用有两个界面，就是人际界面和人机界面。在人际界面，人们用现代汉字记录现代汉语，把口语变成书面语。现代汉语书面语是当今中国社会一时也不能离开的交际工具。在人机界面，人们用各种办法把用汉字书写的现代汉语书面语输入到计算机去，由计算机进行加工处理，完成各种工作。这两个界面都要用现代汉字，现代汉字在使用中出现的种种问题，都必须得到妥善解决。

三、什么是现代汉字学？

以现代汉字为研究对象建立起来的汉字学的分支学科，就是

现代汉字学。现代汉字学是现代化时期新语文建设的组成部分,它主要研究现代汉字的属性和应用。现代汉字学是“现代汉字|学”,不是“现代|汉字学”。

四、现代汉字学有哪些通论性著作?

笔者见到的有以下七种:

1. 张静贤著《现代汉字教程》,现代出版社 1992 年版。
2. 高家莺、范可育、费锦昌合著《现代汉字学》,高等教育出版社 1993 年版。
3. 苏培成著《现代汉字学纲要》,北京大学出版社 1994 年版。
4. 尹斌庸和 John S. Rohsenow (罗圣豪)合著《现代汉字》(*MODERN CHINESE CHARACTERS*),华语教学出版社 1994 年版。
5. 李禄兴著《现代汉字学要略》,文津出版社 1998 年版。
6. 杨润陆著《现代汉字学通论》,长城出版社 2000 年版。
7. 苏培成著《一门新学科:现代汉字学》,语文出版社 2000 年版。

五、为什么要编选这本论文选?

这本论文选主要是配合我编著的《现代汉字学纲要》这本教材,为学习“现代汉字学”的大学生作为参考用的。大学中文专业开设现代汉字学课程的日见增多,学员学习这门课程只读一本薄薄的教材是不够的,必须阅读适量的参考文献,独立思考,深入钻研,才能真有所得。在《现代汉字学纲要》(增订本)的每章后面,都列出最低限度的参考论文目录,这本就是这些论文的汇编。这本论文选是与《现代汉字学纲要》(增订本)配合使用的,凡是学习现

代汉字学的学生都应该认真阅读这些论文。有些人虽然不是在校的学生,但是对现代汉字学有兴趣,或者在工作中有需要,也可以用这本论文选作为学习现代汉字学的参考。

六、这本论文选是怎么编排的?

这本论文选收入比较重要的现代汉字学论文 30 篇,其中不乏名家名篇。这些论文按照《现代汉字学纲要》(增订本)里的章节排列先后顺序。

七、感谢各位论文作者为写作这些论文付出的艰辛劳动,感谢北京大学出版社对我们的教学给予的大力支持,预祝各位学员学习进步,获得丰收。选编中的不妥之处,敬请指正。

八、所收各篇,散见于各种刊物、文集,体例格式不加统一。

目 录

在“汉字问题学术讨论会”开幕式上的发言	朱德熙(1)
汉字的性质	苏培成(7)
现代汉字学发凡	周有光(24)
建立现代汉字学刍议	高家莺 范可育(34)
三种字频统计资料的比较	李兆麟(43)
频率与汉字的简化	王凤阳(52)
现代汉语用字的定量问题	周有光(69)
字形规范化的重要依据	高更生(81)
现代汉字笔画规范刍议	费锦昌(90)
汉字部件分解的原则	范可育(102)
现代汉字独体与合体的再认识	晓 东(111)
字符的分类和归类	钱乃荣等(118)
现代汉语形声字形符研究	康加深(126)
现代汉语形声字声符研究	李 燕 康加深(141)
关于当前文字改革工作的讲话	胡乔木(155)
谈谈汉字整理工作中可以参考的某些 历史经验	裘锡圭(167)
论审音原则	王 力(174)
学习和贯彻《审音表》的若干问题	曹先擢(189)
现代汉字中的多音字问题	周有光(198)
《四角号码检字法》序	胡 适(208)
部首查字法的回顾与改进	曹乃木(220)
关于汉字统计特征的几个问题	林联合(227)

“多余度”与文字优劣·····	尹斌庸(242)
汉字在计算机中如何表示·····	顾小凤(253)
步入信息社会的汉语和汉字·····	张 普(260)
汉字编码与汉字键盘输入系统·····	陈一凡(287)
汉字规范化琐议·····	高更生(299)
海峡两岸用字比较·····	许长安(308)
汉字简化问题在台湾·····	仇志群(321)
汉语文的特点和当前的语文问题·····	吕叔湘(329)

在“汉字问题学术讨论会” 开幕式上的发言

朱德熙

今天这个会叫“汉字问题学术讨论会”，“汉字问题”归哪一门学问研究？照传统的讲法应该归文字学，可是文字学这个名称好像不能概括我们今天要讨论的所有的问题。裘锡圭同志写了本书，要在商务印书馆出版，他和我商量书名。他本来想叫“汉字学概要”，我说“汉字学”听起来很陌生，不如还是按照传统的说法叫“文字学”。他听信了我的话，这本书不久就要出版，书名就是《文字学概要》。用“文字学”确实有毛病，因为没有点明是中国文字，是汉字。文字学按照传统的用法，有一定的含义。今天讨论的汉字问题超出了传统的文字学的范围。叫文字学概括不了，所以我现在倾向于采用汉字学这个名称。

汉字学包括的内容很广，今天我只想就其中四个方面说一点粗浅的想法。

一是汉字的性质；二是对汉字功能的估价；三是汉字改革问题；四是关于汉字历史的研究。

关于汉字的性质，历来讨论得很多。要弄清楚汉字的性质，似乎应该区别两个不同的平面。一是汉字作为语言的符号，另外一个平面是汉字本身使用的符号。过去有些问题搞不清楚，恐怕是没有把这两个平面分清的缘故。这一点裘锡圭同志在《中国语文》

上的一篇文章里讲得很清楚。作为语言的符号,汉字是代表语素的。从表面上看,每个汉字一个音节,好像汉字是代表音节的。可是相同的音节有时用好几个不同的汉字表示,意义都不一样。可见汉字代表的不是音节而是语素。所以汉字可以说是一种语素文字。这是赵元任先生第一个提出来的。当然有的汉字不代表语素,只代表音节,像大家常举的“葡萄”的“葡”和“萄”,但这究竟是少数,绝大部分汉字都是代表语素的。以上是就汉字代表的语言单位的大小说的。从汉字本身的构造看,汉字是由表意、表音的偏旁(形旁、声旁)和既不表意也不表音的记号组成的文字体系。

汉字起源于图画。在汉字产生的早期阶段,象形字的字形跟它所代表的语素的意义直接发生联系。虽然每个字都有自己固定的读音,但是字形本身不是表音的符号,跟拼音文字的字母的性质不同。象形字的读音是它所代表的语素转嫁给它的。随着字形的演变,象形字变得越来越不象形。结果是字形跟它所代表的语素在意义上也失去了原有的联系。这个时候,字形本身既不表音,也不表义,变成了抽象的记号。如果所有的语素都是由这种既不表音也不表义的记号代表的,那么汉字可以说是一种纯记号文字。不过事实不是如此。汉字有独体字与合体字的区别。只有独体字才是纯粹的记号文字。合体字是由独体字组合成的。组成合体字的独体字本身虽然也是记号,可是当它作为合体字的组成成分时,它是以有音有义的“字”的身份参加的。

讨论汉字的性质,不能离开它所代表的语言——汉语。过去我们对汉字与汉语的关系注意得不够。过去研究文字学的人只讲字形,讲六书,对语言不感兴趣。这是传统文字学很大的弱点。我们现在研究汉字学,要突破这个框框。字形当然要研究,但尤其要研究汉字和汉语的关系。这不仅对研究汉字的人来说重要,对研究汉语的人来说也很重要。研究汉语不关心汉字是不对的。汉语是用汉字记录的,研究汉语不能离开汉字。我们研究先秦汉语,无

论研究音韵、训诂,还是语法,都要通过汉字。当然我不是说文字就是语言本身,但只有通过文字才能看到先秦汉语。研究后代的语言也一样。例如敦煌变文,里边有好多错别字,通过错别字可以发现很多问题。错别字是很重要的语言资料,它告诉你关于当时语言的许多信息。明清小说里许多俗体字也反映当时的语言。在这方面,李荣先生最近写了一些很有意思的文章。总之,我觉得过去研究语言的人恐怕对汉字的重要性估计不足。当然也有另外一种倾向:就是完全相信汉字,把文字记录当作惟一的资料。例如完全拿书面材料来研究现代语法。这是要上当的。因为有很多东西汉字没有记录下来,它给你很多假象。所以这个问题非常复杂。我们一方面要充分利用汉字给我们的信息,另一方面要排除汉字给我们造成的假象。

第二是关于对汉字功能的估价,就是讨论汉字的优缺点问题。过去似乎对汉字的缺点讲得多些,这两年有人做翻案文章,说汉字有很多优点。其实汉字本来就有很多优点,把汉字全盘否定掉是不行的。关于汉字和拼音文字的得失刚才吕叔湘先生做了很详细的比较。近年来国内外有些研究自然科学的人,认为汉字好得不得了。有人说汉字最大好处就是意义和字形直接挂钩,不通过声音。就好像现在国际上通用的一些符号:画枝烟卷再打个叉就表示禁止吸烟。不管是说哪种语言的人都能看懂,用不着学。我曾经碰到过一个意大利经济学家,他说他要创造一种文字,原理和汉字一样,全世界的人都能看懂。这不是世界语,而是“世界文”。说不同语言的人都能看懂,可是读起来当然完全不同。从理论上说,这似乎是可能的。王士元先生就做过这样的试验。他教他的小女儿认汉字,发的音却是英文,例如“火”字她念成“fire”,“刀”念成“knife”。我很怀疑那个意大利经济学家所说的文字体系是否能建立起来,因为这个问题不像他想的那么简单,我们中国人创造象形字很有经验。有些东西是画不出来的。譬如说虚字怎么办?不

过我没有研究过这个问题,所以也不敢轻易下结论。

有的心理学家通过实验,指出儿童学习汉字似乎比学习拼音文字还容易些,至少不比学拼音文字难。西方有些儿童先天就有不能拼音的缺陷。学汉字不会发生这种问题。这方面的研究还刚刚开始,目前还难以得出明确的结论。不过有一点是清楚的。讨论这个问题,要把认汉字、写汉字、用汉字三者区别开,不能混为一谈。三者之中,认最容易,写就比较难。因为汉字的特点很明显,只要抓住某一方面的特征就能认识它。例如繁体字“龜”和“龍”特点鲜明,不难认,可是写起来就难多了。因为写的时候要求掌握全部细节,差一点儿也不行,而且还有笔顺问题。总之,写汉字要比认汉字难。至于难到什么程度,那就要去做实验来测定。用汉字要比认和写都难得多。所谓会用,就是要学会区别同音字,知道哪种场合用哪个。例如“唯、惟、维”三个字同音。“维持”“维护”只能写“维”,不能写“惟”,而“思维”也可以写成“思惟”。“惟独”“惟恐”可以写“惟”,也可以写“唯”;可是“唯心论”“唯物论”又只能写“唯”,不能写“惟”。

所以讨论学汉字到底难不难,得把这三个问题分开,不能笼统地讲。

汉字最大的优点是能够通古今、通四方。古今字音差别很大。但由于字义的变化比较小,而且两千年来汉字字形相当稳定,没有太多的变化,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,汉字的这个特点对于文化的延续和继承,对于说不同方言的人之间的交际和理解,具有重要的作用。

汉字的另一个明显的优点是能够区别同音词。所谓区别同音词,严格地说就是能确定不同的语素。当我们听到一个不熟悉的科学术语的时候,往往不知道指的什么,可是写出汉字来,就会对

我们理解它提供一点启发或线索。我觉得从小就跟汉字打交道的中国知识分子已经养成了一种对汉语作语素分析的习惯。你告诉他一个他不熟悉的词儿,他就要问汉字是怎么写的。对于外来词,我们喜欢意译,不喜欢音译,就是这个道理。

第三个问题是汉字改革。这个问题太大、太复杂、不好谈。我想现在提出在一个时期内让汉字稳定下来,这是很对的,是人心所向,符合群众愿望的。稳定包含两方面意思,一是拼音化,即使对,也不可能马上开始实行,现在只能研究;二是汉字简化也要搁一下。有的同志说,有的简化字不合理,能不能再改一改。我想汉字里不合理的東西多得很,挑了这个还有那个,永远改不完。我主张将错就错,现在已经够乱了,千万不要再随便改动了。我们那么大的国家,那么多的人口,那么多的书,哪怕改动一个字,牵涉的面也很大。而且出现一个新的写法,旧的并不能废弃。结果是增加总字数,给孩子们增加负担。汉字改革问题我们可以研究,但不忙于实行。现在谁也说不出来怎么样就一定对,怎么样就一定错。我想应该多做些实验,有了事实才好说话。

第四個问题是汉字的历史研究。汉字有悠久的历史,这方面的研究很重要。但是过去太着重于古文字的研究,总认为后代文字变化不大,没有什么好研究的。从宏观上讲,可以这样说,如果汉字要划分历史时期的话,从甲骨文字一直到小篆是一个阶段。小篆还属于古文字阶段。汉隶以下直至今天就是近代文字,从宏观上讲没有太大变化,但仔细一看,变化还是不小。就说从汉朝到现在,许多字都经历了很复杂的演变过程,这里面有很多东西值得研究。古文字的研究不是不重要,但近代文字的研究尤其重要。因为它与我们关系更密切,就像古代史跟我们的关系不如近代史密切一样。所以我们应花足够的力量去研究近代文字的历史。现在的常用字里有许多字的历史我们不清楚,譬如说“抛弃”的“抛”,40年代唐兰先生曾写文章专门考证这个字的来历。这种字过去

人不研究是不对的。我们应提倡近代文字的研究,俗字的研究。现在我们对古文字知道得比较多,近代文字反而知道得比较少,有点厚古薄今,这是不对的。

最后说说古文字的研究。在这方面我只想提一件事,就是 70 年代我国南方出土的大量秦汉时代的竹简和帛书。例如长沙马王堆三号汉墓帛书、临沂银雀山汉简、云梦睡虎地秦简等。这些资料加在一起估计在 20 万字以上。这几批古文献的出土,举世瞩目,它们在学术上的重大价值,用不着我在这里说。为了对这批资料进行研究,有一项重要的基础工作要做,就是应该把这大批的竹简和帛书里的字编一个字汇,就像《金文编》和《甲骨文编》一样。如果能同时再编一个引得,那就更好了。但这两项工作不是一个人能做的,必须要由集体做,这就需要有人做组织工作。原来北大打算牵头做这个事,由于种种原因,现在看来是做不成了。我在这里提这件事,是希望今后有合适的人能带头来做这件事。

(《汉字问题学术讨论会论文集》,语文出版社 1988 年版)

汉字的性质

苏培成

第一节 民国时期对汉字性质的研究(略)

第二节 语素文字说

一、从表意文字到语素文字

索绪尔说:“文字表现语言”^①,因此我们必须联系语言来研究文字。确定一种文字的性质所采用的标准,就是看这种文字的基本单位记录的是什么样的语言单位。

(一)索绪尔的表意文字说。对汉字的性质最早做出明确论断的是现代语言学的开创者和奠基者瑞士语言学家德·索绪尔,他在《普通语言学教程》中认为汉字是表意文字。他说“只有两种文字体系:(1)表意体系。一个词只用一个符号表示,而这个符号却与词赖以构成的声音无关。这个符号和整个词发生关系,因此也就间接地和它所表达的观念发生关系。这种体系的古典例子就是汉字。(2)通常所说的‘表音’体系。它的目的是要把词中一连串连续的声音模写出来。表音文字有时是音节的,有时是字母的,即以言语中不能再缩减的要素为基础的。”^②

索绪尔的这段话给了我们很多启示。他认为确定文字的性质要看文字和它所记录的语言的关系,而不是看文字内部的构造。根据这个原理,他把世界文字分为表意文字和表音文字两大类,阐

明了这两大类文字的主要区别。他说:表意体系的文字,“一个词只用一个符号表示”。这就是说这种文字符号表示的是“词”,而不是脱离词去表示观念。表音文字是模写词中一连串连续的声音,而表意文字不做这样的模写。就这个意义说,表意文字的“符号却与词赖以构成的声音无关”。自然,表意文字并不是和观念毫无关系,因为“这个符号和整个词发生关系,因此也就间接地和它所表达的观念发生关系”。索绪尔高屋建瓴,为汉字性质的研究指明了方向。

(二) 布龙菲尔德的表词文字说。美国语言学家列·布龙菲尔德发展了索绪尔的思想。他说:“从表面上看,词(words)显然是首先用符号表现在文字里的语言单位。用一个符号代表口语里的每个词,这样的文字体系就是所谓表意文字(ideographic writing),这是一个很容易引起误会的名称。文字的重要特点恰恰就是,字并不是代表实际世界的特征(“观念”),而是代表写字人的语言的特征;所以不如叫作表词文字或言词文字(word-writing 或 logographic writing)。”^③

布龙菲尔德和索绪尔的观点是相同的,只是感到“表意文字”这个名称容易引起误会,容易误会为文字直接表示观念,其实文字直接表示的是语言里的词。因此他认为最好把“表意文字”改为“表词文字”或“言词文字”。

(三) 赵元任的语素文字说。语素(morpheme)理论产生以后,国内外都有学者认为汉字是语素文字。最早提出汉字是语素文字的是赵元任,不过他把 morpheme 叫做词素,而不叫做语素。赵元任说:“用文字来写语言,可以取语言里头各等不同尺寸的单位来写。……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。……以上是讲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,不是标义标音的不同,乃是所标的语言单位的尺寸不同。”^④

赵元任的词素文字(语素文字)说和索绪尔、布龙菲尔德的观点一脉相承,用的术语虽然不同,但是实质完全相同。在国外,认为汉字是语素文字的还有希尔(Archibald A. Hill)、桑普森(Geoffrey Sampson)、德范克(John DeFrancis)等学者。^⑤

(四) 吕叔湘等赞同语素文字说。吕叔湘对汉字是语素文字的理论做了明白透彻的阐述。他说:“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,有意义的是偶然的例外。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,也是惟一的代表。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。”^⑥

朱德熙赞同语素文字说。他说:“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 kāu]的[kāu]、[kāu iàu]的[kāu]和[t' iàu kāu]的[kāu]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。”^⑦

李荣也持语素文字说。他说:“英文是音素文字。假名是音节文字。汉字是语素文字(logogram)。”(《汉字的演变与汉字的将来》,《中国语文》1986年5期328页)他还说:“汉字是语素文字

(logogram),有形有音有义。”^⑧

(五) 语素—音节文字说和音节—语素文字说。这两种理论和语素文字说在实质上是一致的,不过突出了汉字的特点:一个汉字是一个音节。叶蜚声、徐通锵说:“汉字由于种种原因始终维持着意音文字的格局。它是一种语素—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’‘彷徨’等。”^⑨尹斌庸说:“一个汉字基本上代表一个语素。从语音上来说,一个汉字又表示一个音节。因此,综合上述理由,我们建议把汉字定名为音节—语素文字,或简称为语素文字。这一名称较好地反映了汉字的本质特点。”^⑩

二、反对汉字是语素文字的意见

(一) 郑林曦的意见。郑先生说:“汉字根本不是什么‘语素文字’;在古代,它主要是记写单音节词的文字;在现代,它是记写单音节的词和词素以及音节的文字。”^⑪这句话前后是矛盾的。现代语言学所说的“语素”,相当于古代的单音节词,和现代的单音节的词和词素。语素是最小的语音语义结合体,是语法的最小单位。语素可以分为成词语素和不成词语素两大类。在语素的基础上来描写语法,有许多方便的地方。把语素理论运用到汉语语法研究里面,古代汉语的单音节词和现代汉语的单音节的词和词素,和汉语语素所指的范围大体相当。至于在多音节的语素里面,一个汉字记录的只是一个音节,主张汉字是语素文字的学者和郑先生的看法也没有分歧,因为这个事实并不影响汉字是语素文字的论断。

(二) 裘锡圭的意见。裘先生说:“有的人是因为看到汉字里一个字通常代表一个语素,称汉字为语素文字的。像这样撇开字符的性质,仅仅根据文字书写的基本单位所代表的语言成分的性质,来给文字体系定名,也是不妥当的(这里所说的文字书写的基