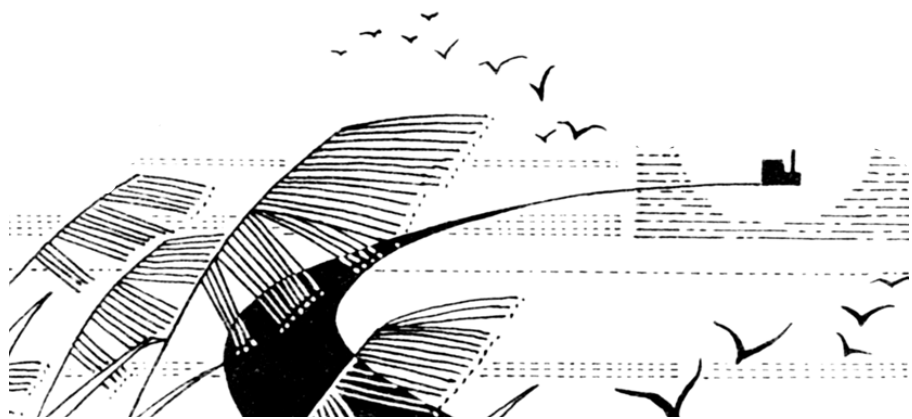


中国语言文字论丛
(二)

马敏 主编



目 录

汉字部件系统研究	1
十五、部件的归并——原形部件和变形部件	54
十六、部件的书写笔顺	60
十七、部件与正字	62
语言的发展与变化问题释疑	71
一、文字有什么样的其作用？	71
二、文字与语言的关系怎样理解？	73
三、怎样理解意音文字、形声字、音位文字等概念？	74
四、为什么说社会的发展是语言发展的基本条件	75
五、语言系统中各种因素的相互影响是怎样促进语言发展的	77
六、怎样认识语言发展的特点	78
七、社会方言和地域方言有何异同	80
八、什么是共同语？现代汉语共同语是什么？	82
九、什么是借词？为什么说意译词、仿译词不是借词？	82
十、什么是双语现象？双语现象的结果怎么样？	83
语言的社会功能	84
一、怎样认识理解语言和说话的关系	84
二、为什么说语言是一种社会现象	87
三、怎样理解语言是特殊的社会现象	94
四、为什么说语言是人类最重要的交际工具	98
五、怎样认识语言和思维的关系	101
语言学的类别和流派	106
一、语言学的类别的要求从哪些角度去认识？	106
二、语言学流派主要有哪些，要求掌握哪些内容？	114
如何理解词义的组合问题	121

一、词义组合的特点	122
二、词义组合的条件	123
三、词义与语境	127
谈谈语法的分析	131
一、怎样认识句子	131
二、怎样认识语素	133
三、怎样认识合成词	135
四、怎样认识语法意义和语法形式	138
五、怎样分析句子的结构层次	141
六、怎样理解语法范畴	144

汉字部件系统研究

一、部件研究的意义及回顾

清末，西方科学和文化潮水般涌入我国，出现了一批教育救国论者，他们认为，西方之所以强大是因为工业发达，工业之所以发达是因为科学发展，科学之所以发展是因为教育普及，教育之所以普及是因为他们用的是拼音文字。再三推理的结果，得出了中国之所以落后是因为汉字是一种落后的文字的结论。于是就把所有的怨气都发泄在汉字身上。这种认识上的偏差使我们在近一百多年来犯了一个大错误，即是企图以拼音化去改造汉字。如果拼音化真的成功了，一个直接的后果是：不会拼音的老一代都成了文盲，所有的文化记载就都成了“天书”，这就叫做文化断层。

现在，有的人则一反常态，认为汉字伟大得了不得，宣称“二十一世纪将是汉字的世纪”。汉字一字都不能动，拆分汉字就犯了“导向错误”；部首检索是天经地义。有人还认为部首，当然就是部件。那就不需要研究部件了！

大家都已经知道，即使终生使用拼音文字的西方人，也觉得他们的文字并不十全十美：拉丁字母是外来的，不是西方人自己的语素，这种文字无法完全反映他们的语言实际。所以，西方也有一些语言学家在呼吁改革。但问题在于语言文字的应用首在约定俗成，没有约定的东西叫得再响也没人理睬。

尽管拼音文字有众多的优点，但拼形汉字也同样有自己的优点。文字是用来表达意义的（即传递信息），不

经常使用的词语，虽然能够读出来，人们却不懂它的意义。拼音文字五百年前的著作，非专家就读而不懂。而我们中国人，一二千年前的诗文，如“床前明月光，疑是地上霜”，连小孩子都能理解。当然，我们也不必过分夸大汉字的优点。汉字的多符性、无序性确实使汉字的信息应用成为难题。与拼音文字一样，有缺点，并不就是落后。如果我们能够使汉字带有一套直接、有序的符号系统，不仅解决了当前信息应用的要求，也解决了历史上遗留的直接排序问题；更有意思的是一百多年来，中国人迫切追求提高识字速度的梦想，能以此得到实现；同时能在基础教育中统一书写规范，消灭错别字、异体字，可以很方便地在计算机上使用。汉字拼音化的热心人并未发觉，拼音文字在计算机上使用方便并不因为它是拼音的，而有三个原因：其一，它的字形是“拼”出来的；其二，它是有序的；其三，它的“部件”是按线性排列的。因此，汉字只要有序性，按线性排列，能“拼”，也就能很方便地在计算机上使用。

计算机的应用，使信息交换的速度得到几倍以至几万倍地提高。中华民族正面临一个全新的信息化时代。这个即将到来的时代使汉字处于信息革命的中心，汉字实现一次自我完善的自我革命已成为历史的必然。

汉字自我革命、自我完善的目标是拥有一套直接的、有序的符号系统，使汉字实现直接排序检索。编码方案设计的发展，使我们看到了解决这个问题的曙光：把汉字拆分为部件，把部件分为几十个类别，每一个类别借用一个有序的键盘符号作为代码。汉字中的部件，虽然没有象拼音文字那样，字母排列有明显的线性，但汉字的书写笔画也可以认为是隐藏着线性的。利用部件笔画

出现的先后就可以使它们排成线性形式。部件是随着字形编码的设计而出现的新概念，我们要使用它，就必须对它进行认真的探索和研究。

1983年，张普先生发表了《汉字部件分析的方法和理论》，总结了几十种编码方案，归纳出它们命名、部件数、部件归并方法、使用部件数量、编码方法等多种不同。张普先生的分析，说明编码方案的不统一，根子在部件。现在我们的教育部门要想统一编码，不应该寄希望于“选码”、寄希望于“规范”。实践证明：没有以科学为基础的“选码”和“规范”，都不会获得良好的效果。我们应该发动与开展汉字编码基础理论的研究与争鸣，以及它的方向、方法和部件的系统研究。

陈爱文先生在其《汉字编码的理论与实践》一书中一针见血地指出：“汉字编码的困难在哪里？——在部件的处理。”一语道破天机。所以，他设计出的表形码，才能突破字形编码的障碍，使设计走向成熟。汉字编码是一门刚形成并不断成熟的设计科学；作为这门设计科学的基础，它的唯一的基本材料——汉字部件的内涵及其外延，正在被逐步认识；解剖部件，分析编码方案的结构，才能真正揭露部件的应用规律。实践表明，只有在对部件规律形成共识的基础上，才能使编码学术界免除彷徨和疑惑。

二、部件的命名

陈爱文、陈朱鹤先生修改了表形码，拿出了一个新版本《表形码和电脑操作》，把编码的基本单元改称字根，据说是为了适合市场要求。是不是市场要求买“字根”而不愿意购“部件”？非也！市场要的是一种易学易记编码及其输入方法，而不管它称为“字根”或“部件”。

五笔字型以其适天时和经营手段取得市场,而不是以“字根”的命名取得市场的。

汉字部件的命名始于字形编码。曾有许多专家探讨过,有人称部件为字根、字符、部首等,但许多人觉得称“部件”比较合适,不过没有统一,因为拿不出有说服力的理由。

命名的确切性是很重要的。汉字合形、音、义为一体。字音可以分解为元素——拼音字母,它们能够拼出所有汉字的读音。利用拼音输入,我们已经有了好多的方案。但拼音输入不能拼出汉字的字形,大量的同音字和同音词以及数万个罕用字无法使用,影响字音编码普及的广度和深度。随着计算机中文应用的逐步铺开,一个全社会、全范围统一使用汉字编码的前景已经明朗。字形编码的实践,证明了它的可行性。字形编码需要拆分汉字,需要得到形码方案的基本材料。这个基本材料,必须有比较准确的命名。名正才能言顺。于是,“部件”这一新概念的语词载体应运而生。

我们认为,“部件”这个命名之所以确切,是因为:

- (一) 它指的是中间层次;
- (二) 它指明是可拆卸的;
- (三) 它表明是相对独立的;
- (四) 它示意拆——拼应该是可逆的;
- (五) 它摆脱了字义、字音整体性的干扰。

这五种属性,正是我们设计字形编码所需要的。王凤阳先生曾断言汉字不能适应机械化、信息化的要求,有趣的是我们却从机械学中把“部件”这个名称移植过来,使汉字适应了机械化、信息化的应用要求。

三、部件定义的设计

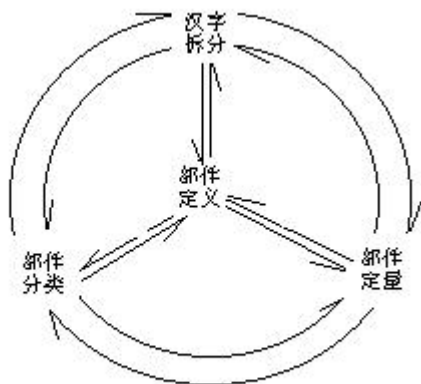
汉字形码方案的根本是部件的定义。不给部件下定义,汉字的拆分就会成为难题。如果我们认为任何人不能也无权随意拆分汉字,就应该首先统一部件的定义。

张普先生很明确地指出:“汉字部件的拆--拼过程,就是分解--合成的过程,汉字部件的拆字法就是汉字造字法的逆顺序。”问题在于汉字该如何拆,如何拼。陈代于先生说:“汉字部件的定义规定的是汉字中什么地方不能拆开。实际上,部件定义和汉字拆分规则是同一个问题的正反面。定义了什么地方不拆,就等于定义了什么地方要拆。”这些话的共同含义是:汉字拆分的前提是部件的定义。

汉字字形编码方案的设计是一门新出现的设计科学,对它的基础理论,学术界未作深入和有权威的探讨。有的专家认为每个形码方案的设计人,都有自己的部件定义。即认为大家都可以按自己的思想给部件下定义,或者认为现在给部件下定义是多余之举。这种思想,实质是否认部件的客观性(它导致部件产生可以“因人而异”的设计思想)。有的专家认为单个笔画不能算部件。在新华字典里,单个笔画如横、竖、撇、折、点,都可算部首;在汉字集中,单个笔画如一、乙也可以作汉字。所以,认为单个笔画不能作部件,不符合汉字结构的实际。诚然,由两个以上笔画构成的部件占部件总量的绝大部分,但我们也不应忽视一个笔画构成的这小部分。有的专家认为拼字量多的才能算部件。拼几个才算多?没有标准!“民”,可以分为两个部件“尸、氏”,拼字量很低,使用频度很高,算不算部件?“拜”的右半部“𠂇”,是一个很明显该分离出来的部分,构字频度、使用频度都很低,算不算部件?如果不算,怎么办?部件,是汉

字的一部分。汉字中拆出的各个部分，都可以称为部件。一个汉字集中的汉字，要拆出一张部件清单；这张清单中的所有部件，可以拼合这个汉字集的所有汉字。

部件是设计形码方案的基本材料，如果不能定性定量，设计就无据可依。部件的定义，就是定性；部件清单，就是定量。给部件定义是迈出科学地设计编码方案的第一步。



汉字部件定义设计图

提供一个汉字字符集的汉字部件清单，记住这张清单中的几百个部件和每个部件类的代码，我们就会知道任何一个汉字的代码，就可以避免逐个拆分时的二义性，和逐个记忆的不便。有人认为记住几百个部件不易，不如主观挑选一些使用频度高的，作为设计编码方案的基本单元，以降低识记材料的数量。我们认为不能以牺牲部件定义的科学性来解决这个问题，而应该从科学的分类上展开思路。几百个部件如果是逐个记忆，当然很困难。但如能将它们构建为互相关联的分类系统，利用系统记忆，引起联想，一类带出一类，一个带出一个，记忆量就会大幅度下降。

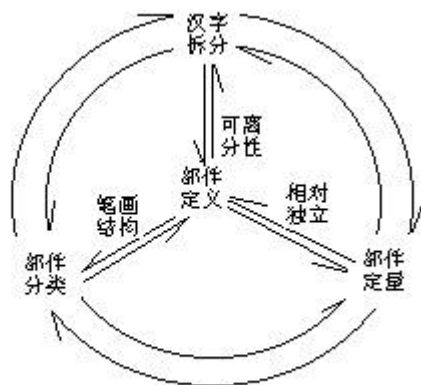
并不是每个方案的设计者都可随意给部件下定义的，拆分汉字、部件定量、部件分类，都制约部件定义的产生。因此，这个定义，不是文字学的，而是汉字字形编码设计学的。设计示意图如右：

右图表明给部件定义的要求：要导出拆分的原则，定位（定量的依据）的界限和分类的方法。这三者必须制约部件定义的制订。该图还表明：拆分、定量、分类，三者构成了一个以拆分决定定量，以定量完善分类，以分类制约拆分的系统结构。

说明白一点：没有拆分，有多少部件就不能确定（即定量）；没有定量，就不能进行分类；没有分类，就不能制约拆分。这是一个相互资助，相互制约的循环图。拆分、定量、分类，这三者既生于定义，又制约定义的产生。

通过这个设计图，我们就可以“按图索骥”，根据上图给部件下定义。

四、部件定义



汉字部件定义验收图

根据上节分析，我们把部件定义为：拼构汉字的基本单元；汉字中、相对独立的、可分离的笔画结构块或基本

笔画称为部件。上述定义可见部件的本质属性有二：一是相对独立性；二是可分离性。相对独立性又受制于笔画结构。现在，我们根据设计要求，对照定义进行验收。图示如右：

右图说明，部件定义所包含的拆分、定量、分类都已经产生，并形成了一个相互制约的系统。该定义已经

实现设计要求。

（一）可分离性。可分离性是部件的动态属性：部件与部件之间因拼合汉字而产生间隙，这种间隙就是拆分的依据。

间隙可以分为显性间隙和隐性间隙。

显性间隙是指：部件与部件之间，有明显的可以一目了然地看出来的间隔。

隐性间隙所指有六：

（1）两个不同的成字部件之间，如“古”的“十、口”之间；

（2）两个笔画走向不同的笔画结构块之间，如“示”的“二、小”之间；

（3）两个不同的笔画结构块之间，如“早”的“日（包围）、十（交叉）”之间；

（4）一个成字部件与另一个笔画结构块之间，如“占”的“卜、口”之间；

（5）一个撇笔与一个笔画结构块之间，如“自、干”的“丿”（撇笔）与“目、十”之间；

（6）与“戈”字类的横笔相接或相交，如“戢、我们”。这里事实是分开的两个部件，但它们的笔画走势和位置在同一条线上，为了书写（或刻字）方便，连在一起了，现在我们需要拆分它们，必须恢复它们原来的面貌。

这里成字部件是指含有字义的部件，如“大、木、口、日”等。笔画走向是指笔画书写的方向，就是自上向下、自左向右。

（二）相对独立性。相对独立性是部件的静态属性：它既维护了基本笔画结构块不再被继续拆分为单个笔

画；又保证了符合条件的单个笔画作为部件的“合法地位”。

笔画结构块是指笔画之间构成最基本的结构关系，它可以分为三种基本类型和两种特殊类型：

- (1) 笔画与笔画相离，称相离型；
- (2) 笔画与笔画相交，称相交型；
- (3) 笔画与笔画相接，称相接型；
- (4) 笔画与笔画相接后形成包围，称包围型；
- (5) 两种基本笔画结构混合构成的，以某种基本特征为架子的结构关系，称字架型。

字架的条件是每一个小类都以一个共同的特征为架子，然后增加笔画或改变笔形，形成的类别。

“块”是指由两个或两个以上笔画为基本结构构成的部件。在汉字中，它是可分离的；对汉字中的其他部件来说，它是相对独立的。

笔画与笔画因构造部件产生的间隔称为结构关系。结构关系具有内聚性，是制约拆分的依据。为了防止将部件（基本笔画结构块）拆到笔画一级，必须确认：基本笔画结构不能拆分的原则，因为，它们是相对独立的。

五、部件定义的补充解释

如果部件只有可分离性，拆分必定要拆到笔画，正因为如此，有的人才会认为汉字的基本单元是笔画，只有笔画，才不能再拆下去了。这是它的动态属性所决定的。可是，汉字检索和计算机的应用的实践告诉我们，在汉字与笔画之间，还应该有一个中间层次——部件。缺了这个层次，不宜应用。这个中间层次，必须是相对独立的。只有相对独立，才能制约它的可分离性。因此，相对独立是它的静态属性。

然而，相对独立的概念，还是模糊的。这是指它的笔画结构块和基本笔画。陈爱文先生说：“部件是一个‘笔画结构块’那么它应该有两个以上笔画，而且应该形成块状”。“一、相交叉的笔画是一个结构块。二、相粘连的笔画，一般属于一个结构块。三、布局匀称的相分离的笔画是一个结构块。四、封闭框内部的笔画如果跟外框是粘连的，则合起来算一个结构块。框内的点笔一般附属于外框。五、单独的点笔一般附属于它附近的结构块。”这里的解释已基本清楚。

例如“艾”，这是两个相交叉的笔画结构块；“从”，这是两个相粘连的笔画结构块；“泣”，这是两个相分离的笔画结构块；“昌”，这是两个包围状的笔画结构块。两个块既可很方便地分离开来，又是相对独立的。所谓相对独立，有一个意思就是它保持自己的完整，让我们一眼看出来。例如“泣”，左边是“氵”，右边是“立”，两边都可以独立组字。又如“旦”，“日”和一个横笔。“日”是独立的“块”，“一”这个横笔怎么办？明明一眼看出是相对独立的，不称它为部件，就无法解决。

有明显间隙的部件，可以一目了然地看出来。对那些没有明显间隙的、连在一起构字的部件，我们提出隐性间隙的概念。隐性间隙也符合认知心理。我们在认字的过程中，先都是接触一些笔画结构简单的字。这些字，有的自己就是一个“结构块”，例如“十”（交叉结构）、“口”（包围结构）。它们拼起来是个“古”字，虽然粘在一起，但拆分它在心理上是能够接受的；再如“二”与“小”，拼成“示”，这是两种不同的笔画走势，拆分开来，也可以接受。如果拆为“一、丁、丶”，心理上反而是不容易接受的。正因为这样，张普先生才认为“相

离应该优于相接”。

“相离优于相接”的拆分原则，用于“辛”，很容易接受，而用于“亓、亏、产”，却有困难，影响直观原则，心理上有障碍。这是什么道理？北京计算机学院的徐劳力副教授有一个很好的量化计算方法：成字部件=3；笔块部件=2；单笔部件=1。“辛”，拆为“立、十”，两个都是成字部件，合起来计数为6；拆为“丿、干”，只能是5。“亓、亏、产”，如拆为“一、丁”，“一、彳”，“立、厂”，比拆为“二、丿”，“二、彳”，“立、丿”，计数大，故心理上容易接受。因为，在我们的认知心理中，对汉字拆分为部件的接受，成字部件>多笔部件>单笔部件。在汉字中，“一”是特殊的。它身兼“三职”，既是汉字，又是部件，又是笔画。运用以上计数方法的时候，“一”必须在单独处于分离状态的时候才能作为成字部件。

运用定义拆分汉字，不要受定义的束缚，例如有的人认为“止”应该算一个“块”，因为，这是从“止”演变过来的，它代表人的脚。如果这么讲，“虎”变成了“虍、七、几”也不能拆分为三个“块”了，“邑”只能作为“阝”，也不能拆开了。我们离开字义对拆分的干扰，必须把它们看成为在现在的情况下，都是一些笔画结构符号。把汉字拆分为部件，要符合当前的认知心理，这是最重要的。为什么要说“当前的”？例如“足”，上面的“口”和下面的“人（人的变体）”都应算成字部件，那么，把当中的“卜”拆出来，也就容易接受。

运用定义分解汉字，个别地方会有一些困难，例如“习”还可以再拆为“勹、冫”，“𠃉”还可以拆为“レ、彳”，这些东西，拆，还是不拆？有的人借此强调汉字的拆分没有规律，就可以任意，这不是道理。因为，这些

是汉字中极个别的,而且,可以根据利弊关系进行评估,再确定拆还是不拆。

六、部件清单

清代文字学家王筠先生说:“人之不识字也,病于不可分。苟能分一字为数字,则点画必不可缺,易学而难忘矣!”这是一个文字学家追求汉字规范和易学的梦想。计算机中中文应用的客观要求,汉字编码设计的发展,许多编码专家的日夜努力,已经可以使这个梦想变成现实。

在这里我们还必须再次强调部件定义的重要性和必要性,没有它,就没有拆分的界限和分类的整体体系。表形码之所以能够得到突破,就是它有一个科学的部件定义,并以此为据整理出一张部件清单和以笔画结构为基础的整体分类系统。陈代于先生说:“掌握了定义的用户即使从未见过部件总清单,也能鉴别其中的大多数部件,避免了对部件的死记硬背。这种特点叫作‘部件的可预知性。’”专家和蒋文钦副教授对表形码的386个部件清单进行了一次清理,把部件总数压缩为368个(现再合并为352个),并为教学方便,将五个笔画结构类型试分为七个。下附汉字部件清单:

汉字部件清单(一)

部件类型	代码	部 件
单 笔 型	横(提)	A 一(一)
	竖(竖钩)	I 丨(丨)
	撇	J 丿
	点、捺	N 丶(丶)
	横折	Z 乙(乙) 冫(冫) 冫(冫)
	竖折	L 乚(乚) 乚(乚)
相 离 型	二笔相离	B 二(二) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	三笔相离	C 三(三) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	四笔相离	D 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	五笔相离	E 立
	对称笔块相离	X 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
相 交 型	点竖(i形)相离	I 丶(丶) 丨(丨) 丶(丶) 丨(丨)
	单交叉	X 十(十) 十(十) 十(十) 十(十)
	横式多交叉	H 卅(卅) 卅(卅) 卅(卅) 卅(卅)
	竖式多交叉	F 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	双竖交叉	S 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	横折交叉	Z 子(子) 子(子) 子(子)
	竖折交叉	L 乚(乚) 乚(乚) 乚(乚) 乚(乚) 乚(乚)
	斜勾交叉	S 乚(乚) 乚(乚) 乚(乚) 乚(乚)
	J形	J 丿(丿) 丿(丿) 丿(丿) 丿(丿)
	K形	K 丿(丿) 丿(丿) 丿(丿) 丿(丿)
相 接 型	Q形	Q 彳(彳) 彳(彳) 彳(彳)
	V形	V 人(人) 人(人) 人(人)
	W形	W 彳(彳) 彳(彳) 彳(彳) 彳(彳)
	S形	S 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇) 𠂇(𠂇)
	Z形	Z 乙(乙) 乙(乙) 乙(乙)

