

# 如何界定汉字的理论单位

——试论汉字的性质及同一性问题

朱 一 星

〈要約〉

筆者は2011年の年初に「国際漢字単位」の概念を初めて打ち出した。その後、さらに「漢字単位説」の明確化を目指すべく、「デフォルト読み」という、漢字に絶対音価のない事実と、漢字の「標意文字」性質などを手掛かりに、日本語における中国人漢字名読みの問題について議論した。筆者の考えでは、現代漢字の理論はつねに現実の文字現象で検証可能であり、難問の解決に役立つ方法論を提供してくれるものでなければならない。本稿では、議論の更なる深化を目指し、現代漢字理論が立脚する漢字の文字論的性質とその基本単位の定義についての検証を試みる。

〈キーワード〉

标意文字 文字功能取向 默认读音 同一性 国际汉字单位

本文借鉴近年脑神经科学，认知神经科学的研究成果，尝试以结构主义语言学，功能语言观的立场来解释汉字——这一语言视觉符号系统，讨论其中若干本质性的问题，通过这些讨论希望打破长期困扰汉字研究的僵局，试图为解决一系列悬而未决的难题打开道路。

## 一、汉字对应语言的语意切分

不少语言学家经常重复这样一个观点：即文字不等于语言本身，文字只能是语音的书面表达形式，即符号的符号。这个观点阻断了语言学界许多汉字研究者的思路，直到今天，仍然认为汉字也和字母文字一样，必定是通过（语音的）语言才能实现其符号功能的“不可动摇”的观点在语言学界十分有市场。

这一观点的持有者经常引用索绪尔（F. de Saussure）《普通语言学教程》中的如下一段话：“后者（指文字符号——引用者）唯一的存在理由是在于表现前者（指语言符号——引用者）。语言学的对象不是书写的词和口说的词的结合，而是由后者单独构成的”<sup>1)</sup>。

笔者其实也无意否定索绪尔的观点，然而一百多年前，普通语言学诞生时主要观察研究的对象，是以字母文字为其书写形式的欧洲语言。站在这一立场上来描述字母文字体系和语言的关系是可以接受的，但是今天当我们思考这一观点对于汉字体系是否得当时，就很值得商榷了。

汉字和字母文字有着本质上的区别，进一步说：汉字对应语言“双层切分”的第一层，而字母文字则对应语言“双层切分”的第二层。大家知道，在马丁内（André Martinet）“双层切分”思想中，语言切分第一层的最小单位是一系列有着“两张脸”（语意和语音）的“符号素(monème)”。

第二层则失去语意，只存在“一张脸”，即只对应音节或音素的分节。第一层的语意分节对区别性要求高，各异的分节数量十分庞大；第二层分节则不超过人类发音器官的生理极限，一般对区别性

要求不高，各异的分节数量也不会很多。马尔丁内认为，双层切分机制符合语言的经济原则：用（绝对数量少的）语音符号间接地代表（绝对数量多的）语意符号。显然，第二层的语音符号就不可能越位对应第一层的语意分节，这就是（字母）文字不可能是语言符号的实质性含义<sup>2)</sup>。

然而，汉字则未必然。

同样套用马尔丁内的语言“双层切分”思想，就应该说：汉字（个体）恰恰对应的是语言的第一层（语意带语音的）切分，这才顺理成章。打比方说，单个汉字能够对应的是语言的“分子颗粒”，保留着最小的语言功能。而字母文字对应的就不是语言的“分子颗粒”了。犹如物质世界构成分子的原子不再具备分子的物理性质一般，字母文字对应的基本是不具备语言功能的语音成分。

汉字作为语言的书写符号体系，它的单位个体（符号能指）不仅能够在理论上对应语意（符号所指），而且事实上的确不必由语音中介，也能对应语意分节<sup>3)</sup>。

这一认识得到当今认知神经科学研究成果的支持。读者可以参阅彭聃龄，陈宝国等在2001年以后多次的研究报告<sup>4)</sup>。这些报告中的实验结果说明，大脑在汉字识读过程中，字义和字音呈现同时激活，或者字义的激活发生在字音激活之前（因汉字使用频率的高低不同）。这一现象在该领域同行中也得到了确认。

汉语常用字里有为数不少的“多音多义”现象，比如像“都”[dū/dōu]、“还”[hái/huán]、“行”[xíng/háng]之类的字。经验告诉我们：当你念书时，必须在确认了字义之后才可能读准字音，比如“长头发”这个短语如果没有语境，可能让人无从开口。对于这些多音汉字，是很难想象必须先念出它的声音后才能懂得字义那样的理解机制的。所以，上述实验结果符合我们日常生活中“见字知义”的实际体会，甚至并不让人感到意外。值得注意的是，一系列实验并没有报告测试对象的汉字因其属于所谓象形、指事或是形声而有差异，说明汉字的造字理据并不构成对汉字功能的影响。

日语是目前世界上仅存的，同时使用汉字和字母文字（假名）的语言。长期以来，日本的脑神经科医生就十分注意对大脑损伤者失读，失写病症状进行临床观察和记录。医生和研究人员积累了丰富的第一手资料，有力地说明了日语母语的失读或失写患者对于汉字和假名——这两种不同性质的文字呈现不同的临床症状。最早发现这一事实的，据说是一名叫三浦谨之助的医生。他在1901年第一次记录了一名念不出假名但却能念出汉字的失语症患者的病例。1975年，脑神经科医生山鸟重（Yamadori Atsushi）首次确认了一名左角回脑损伤后患有失读失写症的患者虽读不出假名，却仍保持完好的汉字念诵能力。这一发现促使岩田诚（Iwata Makoto）等进一步做了大量临床调查，并在1984年提出“汉字·假名读写双通道”假说。经过多年的临床采样证明，这些假说如今已经成为该领域的常识性见解<sup>5)</sup>。

尤其给人留下深刻印象的，是1996年东京电机大学的小谷诚（Kotani Makoto）和NTT基础研究所的今田俊明（Imada Toshiaki）共同领导的一个实验：使用122通道全头型脑磁图扫描系统记录了日语单词识别时的脑磁图。6名男性受试者的测试结果表明，假名单词在大脑呈现活跃的语音活动，而汉字单词的语音活动相对不明显。在激活日语词义的测试中，汉字词语所费的时间比假名词语更短，只需要假名的三分之一时间。对于这一结果，小谷诚认为是因为假名词语必须分两步完成：先激活语音，然后通过语音才能通达词义，因此耗费一些时间。而汉字词语则一步完成，由视

觉信号直接通达词义，不必以语音为中介，所以速度快得多。这一实验结果在当时就受到了媒体的关注，被认为证明了汉字能直接表达语意<sup>6)</sup>。

## 二. 文字语言的时代

语言学的讨论对象只能是（语音）语言这一“正统”观念还从另一个侧面受到挑战。手语，作为一种“纯正”的视觉语言已经在国际上广泛被确认为属于“自然语言”之一。这也是经过脑神经科学以及脑功能障碍研究证明的<sup>7)</sup>：尽管听障者依靠手、指、眼神、表情等多种视觉手段进行交流，但是他或她赖以思维的大脑深层区位却是与依靠听说进行交流时大脑所使用的区位一致的。也就是说，听障者和非听障者虽然使用不同的知觉系统（听觉和视觉），却都是使用相同的大脑语言中枢接受、记忆、和思维，并进行“语言”交流的。

比如 Howard Poizner 的研究团队就曾经详尽地描述了几例以美国手语（ASL）为母语的听障者大脑受损伤后并发失语症的临床观察，证实美国手语母语患者同普通英语母语患者具有许多相同的失语症病理机制。

据说当乔姆斯基（A. N. Chomsky）定义语言“是语音和语意的固有对应”时，曾有人问他手语是否被排除在语言之外？对此，乔姆斯基回答说：“我所说的语音指的是‘信号’，应该说（语言）是信号和语意的对应。语音是否具有决定性意义是个悬而未决的问题。语音也许是决定性的，但支持这一论断的佐证尚不充分。”<sup>8)</sup>

如果说汉字有别于手语的话，那就只能是汉字比起手语更靠近语音语言。在这里，语言的符号能指不只限于语音。这与我们在第一节讨论过的，汉字对应语言的第一层切分的观点紧密关联。由此看来，将汉字视为语言的视觉符号，象索绪尔言及的“对汉人来说，……文字就是第二语言”<sup>9)</sup>的观点，是否也应该如同手语那样，成为语言学界的共识？

在此，我们不可能回避一个汉字争论中十分初级的，但却往往非常迷人的问题：汉字对于使用者来说，虽然是可以“见字知义”的。但是对于非使用者就不可能“看出”个所以然了。不用说对于非汉字圈阅读者如此，即使对于汉字使用者来说，从未学过的汉字也是“看”不出意义的（根据上下文推测不在此列）。即便是对于所谓“象形”或“指事”的汉字，也不会特殊照顾。人们通常说汉字“承载”或“表现”语意，这种说法只是在作为比喻时才说得过去。不论这种观点的持有者是否悟到这一点，汉字的偏旁部件既不是，也无法成为准确描写语意的手段。这一点，在任何时候都是不可改变的既定事实。

当伍铁平再版他主编的《普通语言学概要》（第二版）时，终于撤销了把汉字描写成“表意文字”的习惯说法。他在“第二版后记”中对此专门做了论述：

本书不采用“表意文字”这一术语……

当代中外许多语言学家反对用“表意文字”给“汉字”命名，还因为“表意文字”易被人利用来宣传一种错误观点：汉字可以见形知义，无须通过语言。

虽然“第二版后记”中论述的许多理由，直至某些基本观点与笔者并不一致（本文正在讨论之中），然而和“表意”挥手告别这一点，笔者认为合理的，科学的，面向未来的。

然而，恐怕连伍铁平也未必认识到的是，“表意”这个说法之误，误在一个“表”字，而不是“意”（或“义”）字。因为如果我们承认汉字为符号系统（这一点估计不会有人愿意反对），那么按照索绪尔的符号理论，我们就必须承认符号的第一原则是它的任意性。汉字对应语言第一层切分，从符号理论的观点来说，是不具有任何必然性的。汉字对应语意的具体方法，可能会有许多种。但是任何一种都不会是无从选择的，绝对的，或永恒的。刚才我们讨论过，汉字对于从未学过的人来说，字义是怎么也“看”不出来的，就是这个道理。我们能够认读某个汉字，除了通过课堂传授或字典说明等手段以外别无途径。学习过程，就是将汉字同我们的某个“语意心像”（大脑意识到的知觉活动）建立对应关系<sup>10)</sup>，当你下一次看到同样的汉字时，就能够凭借字形激活（想起）该汉字所对应的“语意心像”。如果同一个汉字对应几个截然不同的“语意心像”的话，学习的过程也毫无疑问地必须分别进行。

在世界上有不少学者认为阿拉伯数字(0~9)才是名符其实的“表意”文字，理由是阿拉伯数字似乎不依靠语言，能够超越语言“以形示义”。这里姑且不论阿拉伯数字能不能属于严格意义上的“文字”，更重要的是这些数字本身亦无“表”可言。它们必定是经过学习才得以理解的，只不过有的人忘记了这一幼年时代的事实而已。

如此看来，符号的任意性原则决定汉字对应语意的方法不是“表示”（=自身具有决定权），而是“标示”（=自身不具有决定权，符号能指和符号所指的对应关系来自于约定俗成）。字形和字义的关系恰如语音和语意的关系一样，不是天然的，而是被赋予的；不是积极的，而是消极的。

说明这一点的最有力的旁证，就是现代中国大陆和日本实行的汉字简化。两个政府对汉字分别实施的大胆改革，历来被谨慎的文字学家们视为大大削弱了汉字的“理据”，有悖于汉字的表意性。然而，多年的汉字运用实践表明，字形上的改变，虽然抹去了一些字形中隐含着的最久远的记忆（喜好问题权且不论），并没有在实质性意义上削弱汉字的语言符号功能。

笔者赞成把汉字称做“标意文字”，因为这既体现了汉字对应语意的语言符号性质，又能避免把人们引入认为汉字能够通过不通过学习就能“见形知义”的误区。

细心的学者也许能够感觉到，在近几十年的汉字讨论中已经有不少学者不同程度地意识到了这一问题的“症结”所在。要不然为什么那么多论文作者有意无意地避开“表示”而使用“标示”的说法呢。可是，大胆地直称汉字就是“标意文字”的，据调查还不多，陆锡兴和张树铮或许是论及这一核心问题的仅有的两位<sup>11)</sup>。张树铮对此明确地指出：

如果往深一步说的话，那“表意文字”也是约定的一种符号，……换句话说，表意文字和表音文字一样，字形都是“标”而不是“表”语言中的义或音。因此，准确的称名应该是“标音文字”和“标意文字”。<sup>12)</sup>

此话言之有理，但同时又不得不指出的是：上述两位探索者都未能在文字和语意对应方式的原理层面上追根刨底，因此，也就未能将“标意”之说的理论意义最大化<sup>13)</sup>。

应当明确地断言：汉字，在成为记录语言符号的那一刻起，作为符号的“本质功能”就是“标意”的。这样，有关汉字的定义，就可以描述为：