

试论我国汉字的源起


陈光宇

甲骨文在世界古文字中的地位

目前存世的三种主要文字系统中，历史最早的文字是（1）两河流域苏美人（Sumer）用的楔形文字（cuniform language），（2）尼罗河畔埃及人用的象形文字（hieroglyphic language），及（3）黄河中游中国殷商时代的甲骨文字（oracle bone language）。前二者的考古证据可以上溯到公元前三千年。甲骨文字最早的考古证据，是盘庚迁殷武丁为帝时的遗物，相当于公元前一千四百年左右。

苏美人的楔形文字主要刻在泥版上。中东气候乾燥，大量泥版因而得以保存。例如英国人 Henry Austin Layard 在十九世纪于 Tigris 河畔挖掘出一座公元前七百年的大图书馆，内藏的图书便是数以千计的泥块。

古埃及文多刻在石壁（金字塔中）、石碑上，也有写在纸莎草纸上的。最古的一块石碑叫 Namer Palette，距今有五千年，刻的是古埃及第一朝代的创立者 Menes 征服尼罗河下游的情形，画面上有与象形文字近似的文字符号。

商代文字存世的，目前有刻在龟甲兽骨上的“甲骨文”，刻在青铜器上的“金文”，还有为数极少的刻在陶器上的陶文。可以肯定的是，商代文字一定还有书写在其他材料上的，例如竹子、丝帛等。甲骨文中的“册”字写作，绝类似将竹片以绳子系成大片，其上可书写文字。最近国内接二连三地在战国时代、秦代、汉代的古墓中发现大量的竹简，便是很好的实例（如云梦秦简、信阳长台关的竹书、银雀山竹简等）。虽然竹、帛不易久贮，但将来随着殷商考古的发展，也极有可能发现书写在竹帛上的“甲骨文”出土的。

甲骨文字在世界各种古文字之中，最突出的一点就是它是唯一不需要被“解谜”（decipher）的古文字。其他各种古文字（楔形文字，埃及文字，玛雅文字……）都已失传数百年或数千年，只有甲骨文，下接周朝的金文，而篆，而隶，而楷，一脉传下，未曾间断。世界上的文字能直溯三千五百年而未中断的，只有中国的汉字了。^①

（二）甲骨文在我国文字演进史中的地位

商代的甲骨文字，目前已知的有四千多个单字。确定可识的字将近两千，其他不识的字，多为人名、地名。所以一般而言，甲骨片上的文句，通读不成问题，因之，提供了不少有关商代社会的第一手史料。从中国文字史上来看，甲骨文已是相当进步的文字，与周代的金文颇为近似。

① 通常说的中国字，实在是汉字。中国为一多民族国家，汉族占 94%，使用汉字；其他少数民族也有使用汉字的（如羌族，撒拉族，回族，满族等），也有使用本族文字的（如藏族，蒙古族，维吾尔族等）。

最近李孝定先生利用六书的原理，分析了1226个已识的甲骨文字，得到如下的数据：在这1226个甲骨文字中，象形字占23%，指事字占2%，会意字占32%，假借字占10%，形声字占27%。②换言之，在商代，甲骨文字基本上已经具备了中国文字所有的构成要素了。

我们知道，文字的制作，脱离不了形、音、义三个元素。表形的文字最先产生，次为表义，到了表音的文字出现，我们才能说这个文字脱离了“图画文字”的阶段。换言之，文字要在与语言相结合之后，才能发展与丰富。按照表形、表义、表音三个文字发展的阶段来看，我们可以说，一种文字的成熟程度可以用它所含表音成份的程度多少来表示。中国汉代文字学者许慎提出的“六书”，实际上也是分析文字中形、义、音三个元素的关系与比重。六书与这三个元素的关系简列如下：

| | | |
|-----------|-----------|-----------|
| 象形——表形为主； | 指事——表义为主； | 会意——表义为主； |
| 形声——表音表义； | 假借——表音； | 转注——表音。 |

必须指出，我国汉字从甲骨文字发展到今天，是世界上唯一的形、音、义三元素密切结合的文字，而其他国家、民族的文字都已发展成为纯表音的文字。根本的原因是汉语是单音节语言，相应的文字发展便走上了与其他文字很不同的道路。虽然如此，汉字的成熟度仍然可以从文字的表音成份增加表示出来。更具体明确地说，汉字中，形声字与假借字的比例增加，可以看作汉字发展与成熟的一种指标。

(三) 汉字的线性发展

上节提到用六书的原理分析甲骨文字，表音字大概占甲骨文字的三分之一强。至于利用六书原理分析历史上的汉字的，则有如：

- (1) 清代朱骏声所著的《六书爻例》，分析了东汉时的汉字9,475个。
- (2) 宋代郑樵所著的《书略》，分析了宋代通用的汉字24,235个。
- (3) 现代文字学家李孝定所著的《中国文字的原始与演变》，分析了商代1,226个甲骨文字。

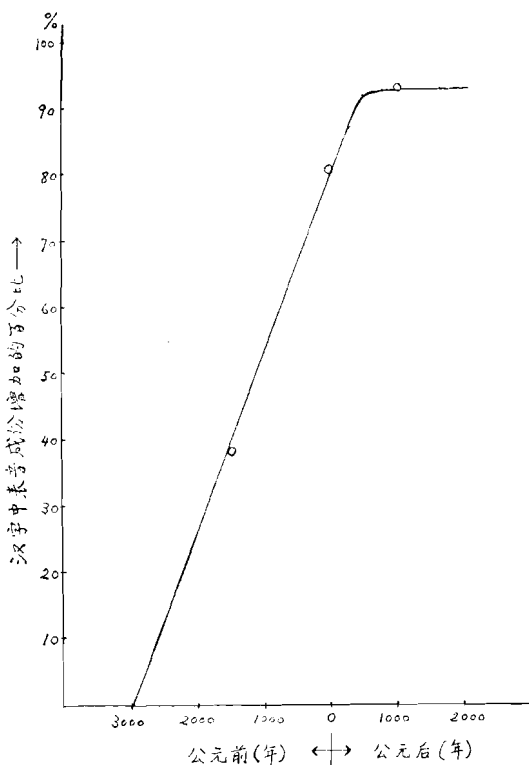
对这三个不同时代汉字的分析所得的结果，可以表列如下：

| 朝代 六书百分比 | 商代 公元前1300年 | 东汉 公元前100年 | 宋代 公元前1000年 |
|-------------|----------------|---------------|----------------|
| 象形 | 23 | 4 | 3 |
| 指事 | 2 | 1 | 0.4 |
| 会意 | 32 | 12 | 3 |
| 转注 | 0 | 0 | 1.5 |
| 假借 | 11 | 1 | 2.5 |
| 形声 | 27 | 81 | 90 |

② 李孝定，《中国文字的原始与演变》，中央研究院史语所，1971年。

如果将这些数据画在坐标图上，以纵坐标表示含表音成份文字的比例（表中形声、假借、转注三项的总和），以横坐标表示时间，我们得到如右图所示的结果。从右图我们大致可得出下列数点：

- (1) 中国文字的演进是线性发展的。
- (2) 甲骨文在中国文字演进史上，决不是处于“原始”阶段，而是一种已经相当高度发展的文字。
- (3) 中国文字发展到公元五百年，已基本上定型（换言之，即六书的比例没有显著的变动）。
- (4) 中国文字的起源约在公元前三千年左右。③



（四）关于汉字的起源

最近国内接二连三出土的陶文，例如西安半坡村的陶文④，河南偃师二里头的陶文⑤，大汶口陶文⑥，江西清江吴城遗址的陶文⑦等，引起了国内外考古学家、文字学家极大的兴趣。特别是有关汉字起源的问题，很多学者利用这些陶文对汉字起源作了探究或推测。其中何炳棣先生认为，半坡的陶文使中国文字至少可以上溯到公元前四千年。⑧中国知名的文字学家唐兰先生过去认为，“中国的象形文字至少有一万年以上的历史，象形象意文字的完备至迟也在五千年以前，而形声字的发轫至迟在三千五百年前。”⑨最近他根据新的考古发现的资料说，“由于大汶口文化，陵阳河遗址和前寨遗址中陶器文字的发现，我国商周以前的图画文字体系，至少可以推到五

③ 由于数据有限，我们不能计算图中的误差。因而，由外推法所得汉字起源的年代公元三千年，很可能有正负两三百年的误差。

④ 《西安半坡考古报告》，中国科学院考古研究所，1963年。

⑤ 《河南偃师二里头遗址发掘简报》，《考古》，1965年第5期。

⑥ 《大汶口》，文物出版社，北京，1974年。

⑦ 唐兰，《关于江西吴城文化遗址与文字初步探索》，《文物》，1975年第7期。

⑧ Ping-ti Ho (何炳棣)，《The Cradle of the East》，The Chinese University of Hong Kong & University of Chicago Press, pp. 223—267, 1975。

⑨ 唐兰，《古文字学导论》，1935年。

千年以前。”^①李孝定先生比较了各期的陶文，大致同意唐兰先生最近的看法。目前出土的陶文多半是单字，半坡遗址的陶文则只有数目字。根据国内以放射性碳—14测定年代的数据，半坡遗址是一个公元前四千年到三千六百年的村落，因而可以肯定的是，在公元前四千年左右，汉字至少已经在图画文字阶段，至于有没有表音成份，目前还不清楚。

虽然目前已知的数据有限，但是由附图我们可以看出，汉字线性发展的特性是很明显的。如果利用数学上的外推法（extrapolation method），可以得到一个结论，即汉字脱离图画文字的时间（亦即加入表音成份的时候）约在公元前三千年左右。^③这个年限与古埃及文、苏美文启萌发生的时间一样。笔者认为，这三种人类史上最古老的文字，一个在尼罗河畔，一个在两河流域，一个在黄河流域，经过一段长期的图画文字阶段（一千年到两千年），一直到公元前三千年，因为种种其他文化因素的刺激（例如生产技术的进步，商业行为的发达等），不约而同地开始脱离图画文字的阶段，与语言结合而正式发展成为流传通用的文字。事实上，国内最近发掘出来的新石器时代遗址像夏县西阴村遗址、半坡遗址等提供了丰富的考古资料，说明了在公元前四千年，定居的村落居民已经以农业为主要生产方式。因此像农业知识的传递，以及定居后人與人之间语言的交通等，都是促使文字脱离图画阶段（包括结绳记事）进而迅速发展的可能因素。

（五）总 结

（1）对于文字起源的观念，我们应该将一般所说的文字与图画文字加以区分。最初人类是用图画文字（包括结绳及种种符号）表达有限的意义，此时“文字”同语言没有很密切的关系。经过一段长时间的“酝酿期”，“图画文字”中开始加入了表音成份（形声、假借）或者音符化（phonetization），其后才成为今日我们所称的文字。因而，我们所谓的文字起源，实际上是指文字开始同语言结合的时候，或者说文字脱离图画、符号阶段的时候。利用汉字线性发展的事实（见图）及数学上的外推法，汉字的起源当可定在公元前三千年左右，其时正当中原处于仰韶文化时期，传说中的黄帝时代也在这时期左右。这个对汉字起源的定点，基本上和最近发现的史前陶文，以及放射性碳—14测定年代所得的数据相符合。

（2）文字的发展是许多文化因素的函数（包括语言、风俗、宗教、科学、战争、商业……）。从附图看来，汉字的发展是线性的。这个现象在人类学及文字学上的意义，还有待进一步的思考与研究。另一方面，关于汉字线性发展的理论还有待更多的数据来验证，换言之还需要做更多的统计分析工作，例如：

- （a）期待从史前陶文到甲骨文中间的“空白”时期的文字能早日出土。
- （b）对于周代金文有必要做系统分析来填补图中由公元前一千年到公元元年左右的空白。
- （c）最近大量出土的汉唐时代的帛书竹简^④是良好的第一手文字素材，有助于补充我们对当时文字的分析。

（3）世界上的文字，无论是古文字或现在仍在使用的，能够持续四千年以上未曾间断过的，只有中国的汉字。研究汉字的起源及其演变，不但有助于了解我国早期的社会发展，更可以帮助我们合理地改造与简化汉字，使它更臻完美。

美国耶鲁大学

1977年2月

^① 《马王堆二、三号汉墓发掘的主要收获》，《考古》，1975年第1期。

《马王堆汉墓出土帛书〈战国策〉释文》，《文物》，1975年第4期。