

基于四本汉语教材的汉字字源研究

刘 曼

摘 要：在中国表意文字的成熟过程中，随着字汇数量的增加，形声字造字作为唯一模式彻底规则化、理性化。通过对四本不同时期法国汉语教材中字汇表的六书统计，我们发现：《汉语札记》《汉文启蒙》这两部教材中所体现的形声字高比例显示出对中国语言文字自身规律的自然接受。形声字在数量上的绝对优势并没有在我们所研究的两本当代汉语教材中明显体现出来。《汉语言文字启蒙》第一册字表中，形声字和会意字比例几乎持平，这两个比例说明了对中国文字自身特点的尊重，同时，对现代汉语高频字的倾斜与会意字比例的上升有关。由汪德迈教授和金丝燕教授共同编写的古汉语教材《古文言读本》，会意字比例比形声字比例要高。后两本供初学者使用的当代教材中大量会意字的呈现是合理的，这些会意字会引导他们更加快速地发现汉字的源起以及这种语言自身的逻辑。

关键词：《汉语札记》《汉文启蒙》《汉语言文字启蒙》《古文言读本》 六书 形声字

第一本教材题为《汉语札记》(*Notitia Linguae Sinicae*, 1831)^①，由法国耶稣会传教士马若瑟(Joseph-Henri-Marie de Prémare, 1666—1736) 1728年在广东编写完成，近百年后被雷慕沙(Abel-Rémusat Jean-Pierre, 1788—1832)在国王图书馆东方手稿部找到，最终于1831

^① Joseph-Henri-Marie de Prémare, *Notitia Linguae Sinicae*, Malacca: Cura-Academia Anglo-Sinensis, 1831. 法国国家图书馆手稿部，编号：Chinois 9259；文学与艺术部，X-3003；NUMM-316242. Trad. BRIDGMAN James G., *The Notitia Linguae Sinicae of Prémare*, Canton, The office of Chinese repository, 1847.

年在马六甲出版。^①

第二本题为《汉文启蒙》(*Éléments de la grammaire chinoise*, 1822)^②, 1822年首版, 1857年再版。作者雷慕沙曾担任法兰西学院(Collège de France) 1814年首创“汉文与鞑靼文、满文语言文学讲座”教授教席(chaire de Langue et littérature chinoises et tartares-mandchoues)的第一任主讲教授。

《汉语札记》由拉丁语撰写, 绪论和正文部分共262页, 后附28页由雷慕沙添加的索引。这本教材同时涵盖了汉语的典雅文风(*la langue savante*)和通俗文风(*la langue vulgaire*), 雷慕沙曾评价“马若瑟是唯一将汉语的两种语体风格进行区分的人, 其他语法学家都是将两者混杂而谈。”^③很显然, 这部教材套用了那个时代为印欧语系而构建的框架, 西方的词类划分, 以及通过欧洲人所熟知且难以摆脱的形式来介绍规则。作者在文风和修辞的介绍上着墨甚多, 期待以此引领学习者感受汉语的真实风格。另外, 大量引自中国古代典籍的例证是这本教材的特色之一。

《汉文启蒙》由法文撰写, 1822年首版共214页, 1857年再版共240页。此教材效仿马若瑟的《汉语札记》, 正文分为“古文”(*kou wen ou style antique*)和“官话”(*kouan-hoa ou style moderne*)两部分,

① 此书完成于1728年, 马若瑟与时任法国皇家铭文与美文学院(Académie des Inscriptions et Belles-Lettres)院士的傅尔蒙(Étienne Fourmont, 1680—1745)长久的书信往来并没能为此书的出版提供便利, 反而遭受了否定和批评, 以致其手稿辗转沉寂近百年, 在19世纪初才由雷慕沙于法国国王图书馆的东方手稿部找到。后由英国新教传教士马礼逊(Robert Morrison, 1782—1834)和金斯博鲁勋爵(Lord Kingsborough, 1795—1837)共同努力完成了这本汉语教材的出版。

② Jean-Pierre Abel-Rémusat, *Éléments de la grammaire chinoise ou principes généraux du kou-wen ou style antique et du kouan-hoa, c'est-à-dire de la langue commune généralement usitée dans l'empire chinois*, Paris: Imprimerie Royale, 1822. 法国国家图书馆手稿部, 编号: Chinois 9257; X-16595; Z RENAN-6226; 8-IMPR OR-73; 8-BL-2051 < Ex. 1 >. Nouvelle édition, Paris: Maisonneuve, 1857. BnF. MFICHE X-3001; X-3001; NUMM-5720892. Paris: Ala productions, 1987.

③ “Le P. Prémare est le seul qui ait tracé nettement la division des deux styles chinois, que tous les autres grammairiens ont confondus.” In ABEL-RÉMUSAT Jean-Pierre, *Éléments de la grammaire chinoise* (Paris: Maisonneuve, 1857), “Préface”, pp. ix-x.

采用相同的模式，以词类划分为框架描述汉语中的名词（Substantif）、形容词（Adjectif）、专有名词（Noms propres）^①、数词（Noms de nombre）、量词（Numérales）、代词（Pronoms）、动词（Verbe）、副词（Adverbes）、介词（Prépositions）、连词（Conjonctions）、叹词（Interjections）和虚词（Particules），以及一些句法规则。此外，教材中出现的大量引自经典文学或古白话著作的例证，同样也是借鉴了马若瑟的教材和其他传教士前辈撰写的著作。^②雷慕沙虽然从之前的西方研究中，尤其是马若瑟的《汉语札记》，汲取了很多思想，但他真正考虑到了中国语言的自身特点，远离了熟知的印欧语言框架。因此，《汉文启蒙》被认为是欧洲“第一部科学地从普通语言学的角度论述汉语语法的学术性著作”^③。（参见表格1）

表1 《汉语札记》和《汉文启蒙》基本信息表

	<i>Notitia linguae sinicae</i> 《汉语札记》	<i>Éléments de la grammaire chinoise</i> 《汉文启蒙》
年代	1728年完稿，1831年出版	1822首版；1857年再版
作者	耶稣会传教士	医学博士、汉学家
页数	262页+28页索引	214页
框架结构	第一编 论通俗文学中的汉语 第二编 论典雅文学作品中的汉语	第一部分 古文或古体风格 第二部分 官话
涉及对象	语音、汉字、语法、虚词、修辞	语音、汉字、语法、虚词、句法
写作语言	拉丁语	法语
引证编写 体例	注音—汉字—拉丁文释义（少量 法语释义）	汉字—注音、单字拉丁文释 义—法文整体释义

① 专有名词只在“古文”部分作了介绍。

② “On ne fait nulle difficulté d’avouer que plusieurs exemples qu’on trouvera rapporté dans ce volume, ont été empruntés, soit à l’ouvrage du P. Prémare, soit aux autres dont on vient de parler...” In ABEL-RÉMUSAT Jean-Pierre, *Éléments de la grammaire chinoise* (Paris: Maisonneuve, 1857), “Préface”, p. xix.

③ [法] Paul Demieville: 《法国汉学研究史》, [法] Jean-pierre Drege编: 《法国当代中国学》, 耿昇译, 北京: 中国社会科学出版社, 2004年, 第27页。

接下来两本是当代汉语教材。其一是由首任法国国家教育部汉语总督学白乐桑（Joël Bellassen，1950—）教授编写的《汉语语言文字启蒙》第一册（*Méthode d'initiation à la langue et à l'écriture chinoises*，1989），此教材1989年首版，2008年再版。其二是由法国铭文与美文学院通讯院士汪德迈（Léon Vandermeersch，1928—）教授和法国阿尔多瓦大学特级教授金丝燕共同编写的《古文言读本》。这本古汉语教材先后于2017年和2020年出版法文版和中文版。

《汉语语言文字启蒙》是现代汉语初级教材，旨在通过对话体和叙述体的口语来提高学习者的交际能力。此教材采用简体字编写印刷，共21课内容，提出一个汉字门槛“400字表”，课文中所用汉字均取自此表^①，作者所选的这400字可满足认知日常阅读中66.27%的汉字^②。

《古文言读本》采用繁体字编写印刷，共22课，课文全部由中国古代典籍节选构成，书后附424字字汇表。这本教材“建立在以四个目录范畴（经、史、子、集）划分的古典典籍的传统系统之上”^③，“22课内容涉及古汉语所承载的历史、文化、文学知识，可以学习追溯到中国古老时光的不同类型文学作品中古典写作的规则”^④。

《汉语语言文字启蒙》课文中的口语对话和叙述文本由作者自创，而《古文言读本》课文中的文本都是取自文言典籍的短小片段，例如《周易》、《礼记》、《诗经》、《道德经》等等。

这两本教材都注重汉字字源的解释。《汉语语言文字启蒙》在“汉字记忆法（mnémotechnique）”部分列出每个汉字的原始字形和部件解

① “该教材的所有文字都是四百字表中选出来的”，参见〔法〕白乐桑：《一元论抑或二元论：汉语二语教学本体认识论的根本分歧与障碍》，《华文教学与研究》，2018年第4期，第8页。

② “Les 400 caractères utilisés dans les textes l'auteur été choisis de telle manière qu'ils permettent de reconnaître 66,27% de l'ensemble des caractères des lectures courantes.” In Joël Bellassen, *Méthode d'initiation à la langue et à l'écriture chinoises* (Paris: Les éditions la compagnie, 2008), Tome 1, introduction, p. 5.

③ “est basé sur le système traditionnel des études anciennes constitué des quatre catégories bibliographiques (*jing, shi, zi, ji*) .” In Léon Vandermeersch, JIN Siyan éd., *Manuel de chinois Classique* (Paris: Edition You Feng, 2017), introduction, p.6.

④ Ibid.

构。《古文言读本》中每个字汇都根据“六书”逐一标注出字源意义和字形。(参见表2)

表2 《汉语语言文字启蒙》和《古文言读本》基本信息表

	<i>Méthode d'initiation à la langue et à l'écriture chinoises</i> 《汉语语言文字启蒙》	<i>Manuel de chinois classique</i> 《古文言读本》
课数	21	22
字数	400	424
汉字字体	简体字	繁体字
文本形式	对话、叙述	古代典籍文本
语言形式	现代汉语/书面口语	文言/书面语
字源解释	汉字原文; 部件解构	六书类别; 字源详解

通过对上述四本教材中汉字字源的分析^①, 我们发现:

《汉语札记》绪论部分的第三章附录了一份31页的“音节/汉字表”, 此表统计字数共计1324个^②, 其中象形字128个, 占9.67%; 指事字18个, 占1.36%; 会意字263个, 占19.86%; 形声字915个, 占69.11%; 没有转注字和假借字。(参见表3)

表3 《汉语札记》音节/汉字表六书统计

六书类别	汉字数量	百分比
象形字 Pictogramme	128	9.67%
指事字 Déictogramme	18	1.36%
会意字 Syllogigramme	263	19.86%
形声字 Morphophonogramme	915	69.11%

① 字源统计分析以《古文言读本》和《说文解字》为依据。

② 统计说明如下: (1) “下、洒、殺、差、嗟、些、好、叔、揉、雨、去、女、鑽、選、當、強、媵、更”此18字重复出现, 只统计一次。(2) 音节表1334个音节中只有17个注音旁标有两个汉字, 情况不一: 受(授)、摸(模)、泄(洩)、杓(勺)、蘇(甦)、睹(觀)、版(板)、溫(嗚)、噴(諷)、唱(倡)、咱(咱家)、呵(打)、(一)個、娜(婀)、唆(哄)、番(一)、頃(刻)。其中用于构成词组的字不作统计, 其他一个音节两个汉字的情况两个汉字都作统计。

《汉文启蒙》后附一份汉字表，按214部首排序（Table des Caractères chinois employé dans ce volume, et dans l'Édition du *Tchoung-young*, arrangés suivant l'ordre des 214 Clefs.）。此表共1168个字，其中取自《中庸》^①的有333个字，选自本教材的有835个字，除去重复出现的五个字^②，统计字数总记为830个。我们对这830个字作了六书比例统计，其中，象形字有83个，占10%；指事字有22个，占2.7%；会意字有259个，占31.2%；形声字有465个，占56%；假借字有1个，占0.1%；没有转注字。（参见表4）这个结果完美地回应了作者在绪论中关于形声字和会意字数量的阐述：“会意字数量很多。”；“文字中至少有一半是由形声造字法构成。”^③

表4 《汉文启蒙》字表六书统计

六书类别	汉字数量	百分比
象形字 Pictogramme	83	10%
指事字 Déictogramme	22	2.7%
会意字 Syllogramme	259	31.2%
形声字 Morphophonogramme	465	56%
假借字 Emprunt	1	0.1%

《汉语语言文字启蒙》400字表中，象形字有97个，占比24.25%；指事字有10个，占比2.5%；会意字有146个，占比36.5%；形声字有147个，占比36.75%；没有转注字和假借字。（参见表5）

① 字表中选自《中庸》的汉字下均标有大写字母“T”，所标页码来源于雷慕沙1817年《中庸》译本。参见 Jean-Pierre Abel-Rémusat, *Éléments de la grammaire chinoise*, Nouvelle édition, (Paris: Maisonneuve, 1822), p. 181.

② “来”、“併”、“的”、“眾”、“厦 / 廈”五个字重复出现，只统计一次。

③ “ces sorte de mots composés, dont le nombre est très-grand, sont appelés (会意) *hoèi-i*”; “Ces sortes de caractères, qu'on nomme (形声) *hing-ching*... forment au moins la moitié de la langue écrite” Jean-Pierre Abel-Rémusat, *Éléments de la grammaire chinoise* (Paris: Maisonneuve, 1857), prolégomènes, pp. 1-4.

表5 《汉语语言文字启蒙》400字表六书统计

六书类别	汉字数量	百分比
象形字 Pictogramme	97	24.25%
指事字 Déictogramme	10	2.5%
会意字 Syllogigramme	146	36.5%
形声字 Morphophonogramme	147	36.75%

《古文言读本》的词汇表中共424个字，其中象形字有74个，占比17.45%；指事字有11个，占比2.6%；会意字有183个，占比43.16%；形声字有153个，占比36.08%；假借字有3个，占比0.71%；没有转注字。（参见表6）

表6 《古文言读本》词汇表六书统计

六书类别	汉字数量	百分比
象形字 Pictogramme	74	17.45%
指事字 Déictogramme	11	2.6%
会意字 Syllogigramme	183	43.16%
形声字 Morphophonogramme	153	36.08%
假借字 Emprunt	3	0.71%

上面出现的数字可总结到一张图表。（参见表7）我们可以看到：前两本教材中，形声字所占比例最大（分别为69.11%和56%），而且其数量远超排在第二位的会意字。在《汉语语言文字启蒙》中，形声字比例也是最高的（36.75%），但其数量与会意字的数量基本持平，也就是说，形声字和会意字所占比例几乎一样。《古文言读本》中，会意字所占比例最高（43.16%），形声字所占比例排第二位（36.08%）。这些百分比数字是否意味着一种普遍规律或者一种选择标准？我们有必要回到中国表意文字的源起和演变发展的不同历史阶段。

表7 四本教材中汉字六书比例统计表

六书类别	<i>Notitia linguae sinicae</i> 《汉语札记》	<i>Élemens de la grammaire chinoise</i> 《汉文启蒙》	<i>Méthode d'initiation à la langue et à l'écriture chinois</i> 《汉语语言文 字启蒙》	<i>Manuel de chinois classique</i> 《古文言读 本》
象形字 Pictogramme	9.67%	10%	24.25%	17.45%
指事字 Déictogramme	1.36%	2.7%	2.5%	2.6%
会意字 Syllogigramme	19.86%	31.2%	36.5%	43.16%
形声字 Morphophonogramme	69.11%	56%	36.75%	36.08%
假借字 Emprunt	—	0.1%	—	0.71%

我们选取了四个数据库作为参考框架，分别是：殷商时期（约前1570—约前1045年）的甲骨文；西周时期（前1046—前771年）的金文；战国时期（公元前3—公元前5世纪）的郭店楚简；东汉时期（25—220年）许慎编撰的《说文解字》中的汉字。^①这四个不同历史时期汉字的“六书”比例分别是：李孝定^②对1266个从殷墟出土的已确认的甲骨文字形进行了六书分类统计，其中会意字最多，有396个，占比32.30%；形声字次之，有334个，占比27.24%；象形字227个、

① 这四个历史时期汉字数据库的选择参考了黄德宽的研究方法。参见黄德宽：《汉字构形方式的动态分析》，《安徽大学学报》，2003年第4期，第1—8页。

② 李孝定：《汉字史话》，台北：联经出版公司，1977年，第68页。

占22.59%；假借字129个，占10.52%；指事字20个，占1.63%；分类不详的字有70个，占5.71%。江学旺^①在其博士论文中对《殷周金文集成》所收西周金文中能确认的1706个字作了穷尽性结构类型分析，因有47个字兼具两种结构，因此他的统计数量总和为1753个字，其中，形声字有1051个，占比59.95%；会意字有333个，占比19%；象形字有224个，占比12.78%；指事字有57个，占比3.25%；未详者有88个，占比5.02%。张静^②在其博士论文中对郭店楚简中的1303个既成文字做作了构形分布研究，其中，形声字909个，占69.76%；会意字148个，占11.36%；象形字118个，占9.06%；指事字24个，占1.84%；构形不详者104个，占8%。清代朱骏声的《说文通训定声·六书爻列》对《说文解字》所收录的9475个字的统计结果是：形声字有7697个，占比81.24%；其次是会意字，有1167个，占比12.31%；象形字有364个，占比3.84%；指事字有125个，占比1.32%；假借字有115个，占比1.21%；转注字有7个，占比0.07%。^③（参见表8）

表8 不同时期六书比例参考框架

时间	殷商时期 (约前1570—约前1045年)	西周时期 (前1046—前771年)	战国时期 (公元前5—公元前3世纪)	东汉时期 (25—220年)
语料库	Jiaguwen 甲骨文	Jinwen 金文	Guodian Chu jian 郭店楚简	Shuowen jiezi 《说文解字》
汉字总数	1226	1753	1303	9475
象形	277	224	118	364
	22.59%	12.78%	9.06%	3.84%
指事	20	57	24	125
	1.63%	3.25%	1.84%	1.32%

① 江学旺：《西周金文研究》，南京大学博士学位论文，2001年。

② 张静：《郭店楚简文字研究》，安徽大学博士学位论文，2002年。

③ 李孝定：《汉字史话》，第68页。

续表

时间	殷商时期 (约前1570—约 前1045年)	西周时期 (前1046—前 771年)	战国时期 (公元前5—公元 前3世纪)	东汉时期 (25—220年)
会意	396	333	148	1167
	32.3%	19%	11.36%	12.31%
形声	334	1051	909	7697
	27.24%	59.95%	69.76%	81.24%
转注	0	0	0	7
	0	0	0	0.07%
假借	129	0	0	115
	10.52%	0	0	1.21%
未详	70	88	104	0
	5.71%	5.02%	8%	0

在上述四个历史时期汉字的演变发展中，形声字比重的强势增长是清晰可见的，甲骨文时期形声字占27.24%，西周金文中形声字增长到59.95%，战国时期郭店楚简中形声字继续增长为69.76%，《说文解字》中形声字达到81.24%。此外，1716年官方出版的《康熙字典》里，47021个表意字中，形声字有42000多个^①。郑德坤认为：“如果把近两三个世纪的新词算上，汉字的字汇百分之九十五以上由形声字组成。”^②唐兰在其《中国文字学》一书中也指出，“中国文字在近古期以后，几乎全是形声字”^③。这些数据层层递进地展示出，在中国表意文字的成熟过程中，随着字汇数量的增加，形声字造字作为唯一模式彻底规则化、

① [法] Léon Vandermeersch:《中国思想的两种理性: 占卜与表意》, [法] 金丝燕译, 北京: 北京大学出版社, 2017年, 第52页。

② 郑德坤:《中华民族文化史论》, 香港: 香港商务印书馆, 1978年, 第85页。转引自 [法] Léon Vandermeersch:《中国思想的两种理性: 占卜与表意》, [法] 金丝燕译, 第52—53页。

③ 唐兰:《中国文字学》, 上海: 上海古籍出版社, 2001年, 第89页。

理性化^①。

因此，我们可以说，形声字在数量上的绝对优势是中国表意文字的特点之一。那么，前两本教材中高比例的形声字体现了对汉语自身规律的尊重。《汉语札记》中形声字比例（69.11%）接近战国时期郭店楚简中形声字的比例（69.76%）。《汉文启蒙》中的形声字比例（56%）接近西周金文古文字中形声字的比例（59.95%）。两本教材中形声字比例都低于《说文解字》中形声字比例，而远高于甲骨文时期形声字比例。来自18世纪初法国传教士编写的教材比19世纪初法国本土汉学教授编写的教材对形声造字的接受更自然，后者已表现出对会意字的有意识关注。

但形声字的强势比例这一特点并没有在后两本当代汉语教材中明显体现出来。《汉语语言文字启蒙》中形声字比例36.75%和会意字比例36.5%几乎持平，只有一个字的数量差。《古文言读本》中会意字比例占43.16%，比形声字比例36.08%要高。也即是说，在两本当代教材中，会意字比形声字数量多或数量相当。（参见表9）

表9 四本汉语教材与不同历史时期形声字比例对比

四个历史时期	形声字比例	四本汉语教材	形声字比例
殷商时期 (约前1570—约前1045年)	27.24%	<i>Notitia linguae sinicae</i> 《汉语札记》	69.11%
西周时期 (前1046—前771年)	59.95%	<i>Éléments de la grammaire chinoise</i> 《汉文启蒙》	56%
战国时期 (公元前5—公元前3世纪)	69.76%	<i>Méthode d'initiation à la langue et à l'écriture chinoises</i> 《汉语语言文字启蒙》	36.75%
东汉时期 (25—220年)	81.24%	<i>Manuel de chinois classique</i> 《古文言读本》	36.08%

① 此句为作者在金丝燕教授译文基础上，对汪德迈教授的思想的总结。参见〔法〕Léon Vandermeersch:《中国思想的两种理性：占卜与表意》，〔法〕金丝燕译，第52—55页。

作为一本现代汉语教材，使用频率是《汉语语言文字启蒙》中所选400字的一个重要标准。通过对教材中所选400字和《现代汉语单字频率列表》^①中使用频率最高的400字的六书比例对比，我们发现：教材400字表中高频字的覆盖率为61.75%。而在这400个高频字中，象形字有88个，占比22%；指事字有15个，占比3.75%；会意字有166个，占比41.5%；形声字有131个，占比32.75%，没有转注字和假借字。（参见表10）这显示出了，在400个高频字中，会意字比例高于形声字比例的特点。因此，《汉语语言文字启蒙》尊重中国文字自身特点，对现代汉语高频字的倾斜与会意字比例的上升有关。

表10 《现代汉语单字频率列表》前400字六书统计

六书类别	汉字数量	百分比
象形字 Pictogramme	88	22%
指事字 Déictogramme	15	3.75%
会意字 Syllogigramme	166	41.5%
形声字 Morphophonogramme	131	32.75%

事实上，“会意字的产生，确能回应中国表意文字的本质，由此造出最初的字……形声造字法的出现，是为减弱会意造字的强烈限制……即便是形声，也保留了表意文字的原则，即对象征含义的寻求更胜于对完美同声字的选择”^②。因此，后两本供初学者使用的当代教材中大量会意字的呈现是合理的，这些会意字会引导他们更加快速地发现汉字的源起以及这种语言自身的逻辑。

最后，我们以问题的形式结束这篇文章。

《汉语札记》和《汉文启蒙》这两本在19世纪初出版的教材都同

① 依据美国中田纳西州立大学宣骏教授的《现代汉语单字频率列表》。白乐桑教授在统计教材中的人名频率时也是依据此列表。参见〔法〕白乐桑：《一元论抑或二元论：汉语二语教学本体认识论的根本分歧与障碍》，《华文教学与研究》，2018年第4期，第9页。

② 〔法〕Léon Vandermeersch：《中国思想的两种理性：占卜与表意》，〔法〕金丝燕译，第53—54页。

时兼顾了古汉语和通俗汉语。但“文言”的弃用随之带来了古汉语教材的逐渐消失。然而，现如今的法国，从中小学到大学，有差不多五万名的汉语注册学员。那么，法国汉语教学的未来，是否会出现一种趋势，即，对古汉语教材给予和现代汉语教材同等的重视，以引导汉语学习者探究中国思想的理性？形声字的强势比例作为中国表意文字的特点之一，是否应该体现在所有汉语教材中？



商務印書館
The Commercial Press



商務印書館
The Commercial Press