

严鲁生 ◎ 著

# 拉丁汉语

## ——未来的语言文字

Latinized Chinese — A Language of the Future

这是一本专题性很强的书，在语言学家和文字学家看来也许不够专业，但却提出了他们至今很少触及的许多问题。比如：为什么西方产生表现语音的字母而中国产生表达概念的汉字？汉字对人类语言文字的主要贡献是什么？汉语语素能否用拉丁字母有规律地表达？从语言文字发展的轨迹和未来信息社会的要求来看，人类语言文字的发展趋势是什么？在知识爆炸性增长的时代，人类语言文字适合什么样的结构形式？本书从理论和实践上对这些问题的解答作了有意义的探索，也以这本书告诉外国人：汉语很好学，中国很容易了解。



严鲁生 ◎ 著

# 拉丁汉语

——未来的语言文字

Latinized Chinese —A Language of the Future

 云南大学出版社  
YUNNAN UNIVERSITY PRESS

## 图书在版编目 (CIP) 数据

拉丁汉语：未来的语言文字/严鲁生著. —昆明：  
云南大学出版社，2013

ISBN 978 - 7 - 5482 - 1540 - 0

I. ①拉… II. ①严… III. ①汉语—拉丁化—研究  
IV. ①H1②H024

中国版本图书馆 CIP 数据核字 (2013) 第 097506 号

## 拉丁汉语

——未来的语言文字

严鲁生 著

---

责任编辑：叶枫红

封面设计：周 旻

出版发行：云南大学出版社

印 装：昆明市五华区教育委员会印刷厂

开 本：787mm × 1092mm 1/16

印 张：13.125

字 数：235 千

版 次：2013 年 6 月第 1 版

印 次：2013 年 6 月第 1 次印刷

书 号：ISBN 978 - 7 - 5482 - 1540 - 0

定 价：36.00 元

---

社 址：昆明市翠湖北路 2 号云南大学英华园内

电 话：(0871) 65031071 65033244

网 址：<http://www.ynup.com>.

E - mail : [market@ynup.com](mailto:market@ynup.com)

志在千里  
志在千里

魯生先生雅存癸巳年友毛北生書



一种文化工具，只要易学使用，适合时代需要，它本身就会自动传播，不胫而走。

——周有光《汉语拼音文化津梁》

# 前 言

西方世界一直梦寐以求一种语素组义语言，“能够列数人类的一切思想概念，并将它们按系统次序分类，从而使它们分外明了和简单。而这一切又取决于每一个人的感知对所有固有概念的发现，并由这些概念组成人们所想的一切”（商务印书馆《世界共通语史》第38页）。在今天，这已是人类语言文字应对知识爆炸性增长的信息时代的新课题，也是《拉丁汉语——未来的语言文字》一书的主题。笔者认为，数千年不间断文明提炼出来的汉语基本语素（词根）代表的基本概念（2300多个），以及语素组义表达复杂概念的方法，就是这种科学的思想表达模式。拉丁字母是迄今为止最科学的语音符号系统，两种科学形式的结合——用拉丁字母给基本概念按汉语语素的语音（包括声调）、词性、词义进行编码，再按基本概念分类组合表达人类思想，就产生了拉丁汉语。

本书是专著《汉语速成之路——拼音汉字创新》的姊妹篇。相对于前书，本书着重阐述了拉丁汉语（拼音汉字形式）创制的理论基础和现实意义，并提出了拉丁汉语语音结构的表义方式。由于在拼式的结构上借鉴了阿拉伯语省写元音的技巧，在技术方法上有新的改进，使得拼式更加简洁。汉语语素的拉丁字母拼写形式，使汉字结构向下延伸到一个更基本的层次——用音素字母按规则拼写汉字，使汉语与印欧语同为音素制的同时，又比印欧语系多了一个字（语素）的层次。这是一本对汉语的认识有新突破的书，它打破了一个误区，开创了五个第一。

本书打破了人们普遍认为汉语不适合于拼音文字、用拉丁字母无法表现汉语古典文学的误区。书中用大量的实例说明了拼音文字对汉语的适配性，用大量翔实材料证实了拉丁字母对汉语古典文学的表现力。

第一次系统阐述了汉语被忽视的四条规律：

（1）用2315个常用汉字所代表的汉语基本语素能够表达99%的汉语语义。

(2) 汉语在语音与语义的联系上有分类表述的规律，也就是音近义通规律。

(3) 汉语的韵母可以归结为7个元音系列。

(4) 表达汉语语素只能利用其固有的“语音（包括声调）、词性、词义”形态。

第一次用一个符号系统完整表现汉语语素的“语音（包括声调）、词性、词义”三大要素。

第一次设计出了汉语语音结构的表义分类系统。

第一次把汉语语素组义特征提高到按基本概念分类表达人类思想的高度，提高到人类语言如何应对知识爆炸性增长的高度，并把它与拉丁字母一起并列为世界共通语的两大基础。

第一次提出了汉字的重要贡献是固化人类基本概念分类，发展语素组义。

本书还致力于用拉丁字母而非方块汉字向中外读者介绍汉语语素所表达的基本概念及其表达汉语的方法。这对于汉字的信息处理将带来极大帮助，用电脑把汉字文件转换成拉丁汉语文件就可以使汉语的信息处理变得像英语一样。我国需要提升在世界的文化影响力，但由于汉字难以在世界流通，世界需要一种表达汉语的标准符号系统帮助汉语在国际流通，我们自己也需要与普通话配套的汉语语素的拉丁字母拼写标准（而不仅仅是语音标准）应对全球化、信息化的挑战。由于现行《汉语拼音方案》无法承担这一任务，客观上需要研制准确表达汉语的“拼音汉字”来承担这一任务。拉丁汉语就是在新世纪解决这一问题的全新的技术方案。

# 目 录

|                                            |      |
|--------------------------------------------|------|
| 前 言 .....                                  | (1)  |
| 第一部分 五个基本问题的讨论奠定了拉丁汉语的理论基础 .....           | (1)  |
| 用 2315 个常用汉字所代表的汉语基本语素能否表达 99% 的汉语语义 ..... | (1)  |
| 汉语构词的语素组义方式是否没有英语表达精确 .....                | (21) |
| 汉语语素的拉丁字母拼写形式是否符合语音学 .....                 | (37) |
| 词性调符对于表达汉语语素的意义 .....                      | (51) |
| 汉语韵母的规律 .....                              | (62) |
| 第二部分 从笛卡儿、莱布尼茨的共通语构想到拉丁汉语 .....            | (66) |
| 语言的共性及语言文字的发展趋势 .....                      | (66) |
| 如何理解笛卡儿“应该给所有的思想概念建立一种方法体系” .....          | (70) |
| 人类文字适合什么样的结构形式 .....                       | (71) |
| 文字是语言的视觉编码 .....                           | (73) |
| 汉语的特性及其传承 .....                            | (74) |
| 方块汉字对我们走向信息社会有没有影响 .....                   | (78) |
| 汉字与中国人的思维 .....                            | (80) |
| 西方学者称汉字影响中国“创造力” .....                     | (82) |
| 只有拼音文字才能做到早期的阅读和写作 .....                   | (85) |
| 理想的义音合一文字 .....                            | (87) |
| 拉丁汉语比世界语更实用 .....                          | (88) |
| 拉丁汉语与日语拉丁化 .....                           | (89) |



|                           |       |
|---------------------------|-------|
| 信息时代的中国梦 .....            | (90)  |
| 中国下一轮经济增长的爆发点 .....       | (91)  |
| 第三部分 拉丁汉语 .....           | (94)  |
| 拉丁汉语构成 .....              | (94)  |
| 拉丁汉语语音结构的表义方式 .....       | (100) |
| 第四部分 阅读材料 .....           | (107) |
| 汉语学习的革命 .....             | (107) |
| 汉字改革势在必行 .....            | (111) |
| 国民文化素质与民主 .....           | (115) |
| 中国静默,向遇难者致哀 .....         | (117) |
| 给伦敦的借鉴 .....              | (119) |
| 奥巴马的获胜演说 .....            | (123) |
| 巴菲特:“买美股吧!我就在买” .....     | (133) |
| 人生感悟(2007) .....          | (137) |
| 商业共和国 .....               | (139) |
| 电子书市场热起来 .....            | (143) |
| 失去的阅读艺术 .....             | (148) |
| 拐点临近? .....               | (152) |
| 奥巴马就职演说 .....             | (156) |
| 成功秘诀:需要一点儿失败 .....        | (169) |
| 春意盎然 春风在履 .....           | (174) |
| 蹉跎岁月 新鲜感受 .....           | (179) |
| 古诗词欣赏 .....               | (184) |
| 第五部分 拉丁汉语(汉英)词典(范例) ..... | (190) |
| 参考文献 .....                | (196) |
| 后    记 .....              | (199) |

# 第一部分 五个基本问题的讨论奠定了 拉丁汉语的理论基础

在北大中文论坛和南方中文论坛以及本人的新浪博客中进行的关于如下五个基本问题的讨论，为拉丁汉语的创立奠定了理论基础。

## 用 2315 个常用汉字所代表的汉语基本语素能否 表达 99% 的汉语语义

### 正方

如何解读国家语委调查数据

同样的生活经历比如知青生活，作家从中看到了人生百味，写出小说来反映时代，而普通人只留下回忆。同样的数据，许多人看不出其中的变化。国家语委《2005 年中国语言生活状况报告》调查了平面媒体、有声媒体、网络媒体、文本文件 890 多万个，共 9 亿多字次，其中汉字 7 亿多字次。上述三类媒体共用汉字 5607 个，但仅 581 个汉字就可以覆盖其中 80%，2315 个汉字可覆盖 99%，一个人掌握 1 万多条词语，900 多个汉字就可阅读 90% 的出版物（《北京青年报》2005 年 5 月 23 日）。这些资料说明，汉语从用字形区分同音语素（如使用越来越多的汉字字种）向用语音区分同音语素（更多地使用常用字组词表达语义）的方向发展，基本语素（词根）的作用越来越突出。那些很少用或基本不用的几万个汉字所代表的语义不是被废弃，而是被常用字的表达所代替。

### 反方

“2315 个汉字可覆盖 99%”指的是覆盖文本的范围，并不是覆盖语义的范围。它只不过是常用语义的 1/5 到 1/4。

“用 2300 多个词根字能表达 99% 以上的汉语语义”的说法毫无道理！

“覆盖率”百分比跟“语义”含量毫无关系！只有丝毫不懂统计学的人才会如此荒谬地讨论严肃的问题。2000 个汉字大约是 2000 个语素（忽略一字多义），多 1 个汉字就是多 1 个语素，就是 2001 个语素；多 8000 个字，大约是 10000 个语素。语义是由语素的个数说明其广泛程度的。

### 正方

反方对汉语语素表现力高度集中的特点一无所知。汉语演变至现代汉语，许多语素已由现代词汇表现，按你的说法 10000 字 10000 个语素，那么汉字总数超过 50000，一般人只认识 5000 字，我们只掌握了 1/10 的语素，大家都是文盲！其实要驳倒你的论点很简单，我请你写出一篇文章，统计一下不是由国家语委统计的 2315 个常用字表达的语义，看其是否占到 4/5，尽管对于 890 多万个统计文本来说，一篇文章只占千万分之一，但只要你能写出来也算你沾点边。

### 反方

看你一个帖子，大约 200 字，不能证明你只知道 200 字。我写一本书，即便十万八万的篇幅，也许使用了 5000 字，也未必证明我只知道 5000 字。汉字也许有十万八万，今天少数人用得着的也就是两万三万，多数人恐怕只有一万两万了。尽管如此，那十万八万字依旧还活着，随时都在待用之列，谁又能够限制它们呢？这跟我们任何人能够掌握多少个汉字毫无关系。说到根本的一点，认为常用汉字 2300 多个在多少的大规模统计中覆盖了大约 99%，然后就说那 2300 多个字就是语义的 99%，这绝对是错误的。

### 正方

我有必要明确一下，我们辩论的问题是：用国家语委统计的 2315 个常用汉字，能否表达 99% 的汉语语义，还是只能表达 20% 的汉语语义。什么叫“用汉字表达汉语语义”，你不会不理解吧？就如用拉丁字母表达英语一样，从你写的这个 300 来字的帖子看，没有一个语义不是用这 2300 多个常用汉字表达的，这也算我用你的语料举的一个例子吧！哪些语义是常用字表达不了的？既然你说此种情况占 75% 以上，总该举得出几个例子吧，它必须满足语义不能用常用字定义。

### 反方

动物和植物分类名称使用的汉字在最常用的 2315 字范围的不多，难道动植物名词的语义区分作用不重要吗？难怪动物园和植物园中的许多文字说明必须标记它们的读音，这样做就是为了普及文化科学知识。我这里只是举了

动植物为例说明问题,《现代汉语词典》包含的汉字远远超过了 2315 字的范围,那么 2315 字是词典中上万通用字的几分之几?

### 正方

《分瓣臀凹盾蚧 (*phenacaspis kentiae kuwana*) 防治指标的研究》(王海林、严鲁生、吴琳、黄志勇、杨建国,载《西南林学院学报》1992 年第 1 期)是我和他人一起发表的论文,动物和植物分类名称是你不熟悉的领域,动物和植物分类名称正规来讲应该使用拉丁文,动物园和植物园标注的是它们的拉丁文学名而不是标记读音。就算用中文表述,以上例来看,6 个字中只有“臀”“蚧”是非常用字,只占 1/3,离你的 4/5 还差得远呢,何况“臀”“蚧”的语义也是由常用字定义的。

### 反方

请你不要否定动植物中文术语的学术价值,6 个字中就有 2 个是非常用字,这个比例还低吗?你懂不懂统计规模(基数)是什么意思?我告诉你,6 个字的统计规模太小了,根本就无法提供有价值的置信度。没有置信度还妄谈什么统计学?尽管如此,1/3 的欠缺就是 2300 个汉字范围的语义的欠缺,跟 1%~99% 中 1% 的欠缺完全不相关。

### 正方

小样本是置信度偏低,但相对于反方的无样本来说,应是无穷大。你每写一篇帖子都在证明我的论点:99% 的汉语语义都是由常用字表达。那么,我请你写的证明你的论点的文章在哪里?(注意:必须有占 4/5 的语义不是由常用字表达)

### 反方

你一辈子只看过一篇文章,也只写过一篇文章,所以 100% 的语义全部在那里了。现在小学生的书包即使只放一本电子书,也可以博览群书。十年寒窗不可能只是一本《三字经》的经历。你永远不可能明白的,大概就是博览群书与语义的关系。如果大量文章的 99% 都是常用字词,那么 1% 的非常用字词散布在千百万册图书中,偏偏没有重复出现,进不了常用的 99% 的领域,所以属于非常用字范围。然而,千百万册图书中散见的 1% 记录到词典中去,必须一项一项单列,它们何止 2/3、3/4、4/5、5/6、6/7。这些比例恰恰就是常用字分布率中的剩余数字 1% 说明的语义事实。因此,古人可以收集到十万八万个汉字,现代人发现依旧有三五万字是中国人现代生活不可缺少的。把那三五万字写到一篇文章中去,除了字典、词典,除了《三字

经》《百家姓》一类读物，其他就没有了。遗憾的是，你从来没有看过字典、词典，那里每一词条的解释都离不开不同的语义。

### 正方

你所谓的另外 4/5 的汉语语义难道是用这 2315 个常用汉字之外的字表达的？你怎么解释国家语委的调查？汉字对汉语语义的覆盖不是雨滴似的随机覆盖，它是组成语义的表达形式——词、句来覆盖的。2315 个常用汉字已经覆盖了 99%。100 字的文中有一个非常用字它就带走了 4/5 的汉语语义，这就是你的统计学？请问，你说的字典、词典中的非常用字是不是用常用字解释、表达的？比如“膂”现在说“脊骨”，“脔”现在细分成“肉片、肉丝、剁肉”，“日薄崦嵫”现在说“日薄西山”。哪些语义不能用常用字表达？憋了半天想出个“百家姓”来交作业。请问姓氏代表什么语义，比如姓“沈”代表什么语义我愿领教，声称“博览群书”就是这样理解汉语语义的吗？现代中国的人才 99% 都是读现代白话文书籍成才的，没有几个懂你说的那几万个几乎不用的汉字，不懂“膂”不等于不知道脊骨。翻开《现代汉语词典》，所有非常用字代表的语素、所要表达的语义都是用常用字解释、定义的。

### 反方

就正方的 2300 字来说，占到语义的 99%，是指频率，还是概率？当然是指频率。是动态的还是静态的？如果是指概率，朱镕基的“镕”，它的概率是不变的，而频率是可变的。朱镕基出山前，镕是冷僻字；朱镕基出山后，成为非常用字，后来又成为常用字；退休后回归为冷僻字了。作为概率来说，6763 个汉字，都是  $1/6763$ 。其中任何两字的语义不可互相替代，所以 6763 个汉字的语义的概率，也都是  $1/6763$ 。先生的 2300 字的语义的概率（不是频率）是  $2300/6763 = 34\%$ 。这不奇怪，你用不到，不等于别的地方用不到。囧字就是很好的例子，蟑螂、蝴蝶、蚱蜢、臀、蚧之必不可少正在于此。

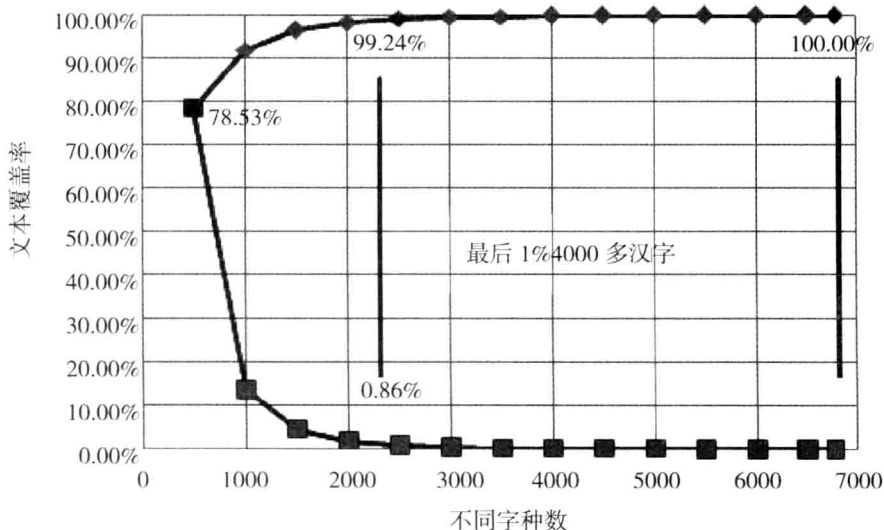
### 正方

首先纠正一下，2315 个常用汉字不是“占到语义的 99%”，而是表达语义的 99%。因为这是中文论坛，我本不愿在这里讨论应用数学，既然有人提出来，看来有必要在理论上厘清楚。就我们辩论的题目来说，在中国语言生活中所出现的“不能为 2315 个常用字表达的汉语语义”为随机事件 A，国家语委调查的“890 多万个文本文件包含的汉语语义”为试验总体 n。890 多万个文本文件包含多少个汉语语义，我目前无法统计，也就是说 n 很大，它并不等同于反方理解的汉字字数，它至少应包括汉语词汇所代表的汉语语义，

按 2008 年语委的统计为词种数 226 万多个。 $m$ （在  $n$  中  $A$  出现的次数）应该是很小的，我就举不出多少例子，我好不容易想出一个“簪”，但它又被常用字“脊骨”表达。反方认为  $m/n = 4/5$ ，我自然请他举一个随机事件  $A$  发生的例子，他最后举了一个“百家姓”，但姓氏不属于汉语语义，到现在为止他的  $m=0$ 。我们姑且用 226 万个词种数代表在中国语言生活中被使用的语义（实际要比这多得多），那么就算 1% 也要举出 22600 个在中国语言生活中使用的而不能用常用字表达的语义，如果按反方说的  $4/5$ ，他就得举出 180 万个这样的语义，就算他能举出几个，能在多大范围内被使用几次恐怕也很难说清楚，占  $4/5$  肯定是瞎说。由此可见，我说反方对“汉语的表达能力高度集中于汉语常用字所代表的基本语素”这一汉语特点一无所知并非武断。

### 反方

这是最荒唐的谬问！那个 1% 覆盖的是常用字中剩余的数千字（ $5607 - 2315 = 3292$ ），代表 1 万通用汉字剩余的近 8000 字（ $10000 - 2315 = 7685$ ）。假设 1 个汉字主要代表 1 个义项，那么剩余的义项大约是 2315 字的 4 倍！也就是说，覆盖率 99% : 1% 的语义比率为  $1/5 : 4/5 = 1 : 4$ 。



图中假设通用汉字 6763 字（100%）。曲线示意每 500 字的覆盖比率，也即覆盖率增量。

请注意，最后 1% 包含的汉字空间，它们是由数千汉字微小百分数累加

得到1%覆盖面的。

## 正方

### 评反方的坐标图

反方坐标图的横轴是按使用频率从高到低排列的7000个汉字（按500字分组），纵坐标是它们各自对应的使用频率，以百分比为单位，最大值是前500字的统计频率78.53%。下面一条曲线是这7000字分组统计频率值的连线，曲线与坐标轴所围成的区域表示这7000字的覆盖面，可以看出第2315个汉字以前已覆盖了绝大部分区域。上面一条曲线是频率的依次累计值曲线，它对应的坐标系的纵坐标表示频率累计值，以百分比为单位最大值是100。把两个有着不同度量不同含义的纵坐标的坐标系叠加在一起，正如把一个数列的第 $n$ 项和其前 $n$ 项和叠加一样荒谬，用这两条曲线包围的区域来表示汉语语义的空间（还煞有介事地画了两条竖线），犹如“痴人说梦”。如此不严谨的治学态度，不知误人多少子弟？对这一主题的数学理解我曾在北大中文论坛贴过，被反方封了，它应该是：在由全部汉字的所有有意义搭配组成的语义表达空间中，99%的汉语语义都落在了由常用汉字搭配组成的语义表达空间（其中重叠少部分由常用汉字与非常用汉字搭配的语义表达空间）。汉字是表达汉语语素的，而语义是由语素组成的，在语义表达空间中汉字频率的密度分布反映了语义的密度分布。反方真想搞清问题是很容易的，守着北大这座知识宝库还会知识贫困？请把你的坐标图和我的评论拿到北大数学系随便找一位教师问问就清楚了，中文系教授问数学题不丢人。如果你有足够的学术公正，请把你问的结果公布在论坛上。

## 反方

历时多年的统计涉及的汉字数大约7000~10000字。其中对前3500个汉字的5次统计中，有3358个汉字是相同的，表明常用字和次常用字是稳定的。由此可以推论，超出这一范围频繁复现的汉字并非稳定不变。如果大大扩充统计规模，所涉及的汉字将达到数万字的规模，必定与现行字典、词典工具书趋向一致。所谓语义量是由统计规模总体说明的，与字义或词义没有直接关系。汉字的事实说明中华文化宏伟发达，深不可测，我们千万不要相信2300来个汉字就可以代表所谓99%语义的荒谬结论！

## 正方

注意：这里反方已经把“表达”偷换成“代表”！而论题是说用2315个常用汉字来“表达”而不是“代表”99%的汉语语义，如同用26个拉丁字

母可以表达而不是代表 99% 的英语语义一样。

反方似是而非的统计图表只能糊弄那些不懂数学的人。现在讲的是常用字的累计频率与语义表达空间的关系，你那个坐标系中代表语义的变量是哪个轴？你以为你画的那两根线之间就是未表达的语义空间吗？你说的“1 个汉字主要代表 1 个义项”本身就是谬论（大多数汉字都不止一两个义项），汉语的语义不仅仅是由单字表示，主要是由它们组成的词、句表示，单字不过几万，而词种类超过 226 万个，比如“马虎”“马上”“知音”……都组成了不同于原单字表达的语义，而构词能力最强的正是这 2315 个常用字，那些非常用字构成的语义就很少，比如那个“炯”字除了形容目光有神以外，还能构成什么语义？如果都像那个“炯”字一样“1 个汉字主要代表 1 个义项”，汉语语义只局限于几万个汉字义项，那 226 万个词所表达的语义怎么来？非常用字除了构词能力低、表达的语义少外，由于认识的人少，人们经常用常用字来表示它们的词义，比如人们写不来“炯炯有神”就写“明亮的眼神”，比如“饷女”用“新婚后娘家送食物”表达。就是非常用字偶尔出现一下也多半由常用字支撑来表达语义，如“挥霍”中的“霍”就不单独表达语义，它们如何能表达反方说的 4/5 的语义？语义的数量主要不是与字种数挂钩，而是与字的使用频率和有意义的组合数挂钩。比如“马”“上”这两个字就可以表达“马、上、马上、上马”10 多个语义，仅“上”就有“上面、上学、上楼、上瘾、上策、上回”等许多语义，其中“马上”表示“立即”和“马背上”两个语义，“上马”表示“骑上马背”和“启动项目”两个语义。若增加两个字种“霍”“炯”，又能增加多少个语义呢？

#### 字频覆盖率与语义覆盖率的关系

反方认为，汉字使用的频率与表达语义没有关系。这又是他知识的一个盲点，其实高频汉字正是由于其在表达汉语语义中起到重要作用才会多次出现。比如“的”，相当多的词汇加上“的”后语义就发生了变化，如我的、红的、前面的、创造的、创造性的……又如“不”，几乎任何一个动词或形容词加一个“不”字意义就发生了相反的变化，这些变化难道没有表示不同的语义吗？表达许多语义都需要用到它们，何止成千上万。“炯”字有那么广泛的语义表达力吗？又如“痘”能表达多少语义呢？用“1 个汉字主要代表 1 个义项”大概能概括非常用字，但能概括得了常用字吗？那些表达语义有限的非常用字有许多表达的语义已被常用字的表达代替了，自然出现频率低了。这就是字频覆盖率与语义覆盖率之间的正相关关系。反方说了半天，就是不愿用实证、举实例来证明他们的观点。关于这一主题的讨论帖不下万



字了吧，除了我举了几个非常用字外，反方都没有人举非常用字表达常用字不能表达的语义的例子。这些帖包含了多少语义，不论怎么统计，常用字表达的语义在论坛上也就超过了99%。举不出实证、实例，光靠空讲，总不是个事吧？

### 反方

你拿一个偷换概念得到的伪命题来说事，真的没有一点觉悟么？

### 正方

关于“字频覆盖率与语义覆盖率的关系”，你看不懂我这个帖吗？什么“伪命题”，你举出实例推翻这一命题了吗？

### 反方

你使用字频覆盖率来等效取代语义覆盖率，这个是你的猜想而已。没有经过证明（或证实）就拿来到处说事，是不是有欺骗善良大众之嫌呢？

### 正方

据国家语委《2005年中国语言生活状况报告》调查，平面媒体、有声媒体、网络媒体、文本文件890多万个，用汉字共7亿多字次，2315个汉字可覆盖99%。这些汉字当然是通过组合来表达汉语语义的。我解读为“2315个常用汉字可以表达99%在中国语言生活中使用的语义”，并不等于你说的“用字频覆盖率来等效取代语义覆盖率”。别人不同意我的解读很正常，但我希望他们用实例来推翻这一解读，而不是用误解和指责“伪命题”这样的方法来推翻。从论坛的意见看，“汉语是词根表现力高度集中浓缩的语言”这一特征已为不少人认可。

### 再论字频覆盖率与语义覆盖率的关系

地质学中有一种伴生元素现象，就是由于某种内在原因使得A、B两种元素总是伴随出现，含量成正相关关系。由于B元素不易测量，人们通过测量A元素的含量和分布来测算B元素的含量和分布。这一方法在云南稀有元素矿床的勘测中屡见成效。要否定这种方法必须提出反例来切断A、B元素的这种正相关关系。把这种方法运用到语言文字的研究上，对于不易测算的汉语语义分布利用与其相关的汉字的频率分布来进行测算，关键是要确定汉字出现的频率与其表达的语义之间是否存在正相关关系。我前面已对这种正相关关系进行了阐述，并在论坛上多次邀请坛友举出反例以切断这种正相关关系，可是没有一人能举出一个低频非常用字表达语义比高频常用字多的例子，许多人只凭着主观感觉发言。