
俞 敏

语 言 学

论 文 集

俞 敏 著

黑龙江人民出版社

俞敏语言学论文集

俞 敏 著

黑龙江人民出版社

1989年·哈尔滨

责任编辑：任国绪

封面设计：安振家

俞敏语言学论文集

Yumin Yuycnxue Lunwenji

俞 敏 著

黑龙江人民出版社出版发行

(哈尔滨市道里地段街179号)

牡丹江书刊印刷厂印刷

开本850×1168毫米1/32·印张13 6/16·字数870,000

1989年9月第1版 - 1989年9月第1次印刷

印数 1—1,000

ISBN7-207-01228-4/H·28 定价：6.00元

序 言

这个集子里收的是我历年来发表过的论文的大部分。除了84年6月，日本光生馆印行的，包括《北京口语轻音节的调值》、《永明运动的表里》、《古汉语派生新词的模式》、《后汉三国梵汉对音谱》、《汉藏虚字比较研究》等十八篇的《中国语文学论文选》以外的全部文章；有些原本遗失，或者没什么多大用处的不算在里头，外加一两篇没发表过的。就这么些了！

记得1937年春天，我刚看完G.O.Russel的《元音》(The Vowel)，莘田师让我写篇书评。我说：“我还正吃着桑叶，还不到吐丝的时候儿哪！”这是谦虚还是自视过高，我说不上来了。反正从那以后，我写文章较比慎重(当然，不保证词句不尖酸)。从1939年到如今，转眼四十六年了。回头看看，禁不住发点儿感慨：这些年就凑了这么点儿东西呀！这当然不都怨我，里头有十八年空白，跟王宝钏住寒窑一样长了。我自己也得负一部分责任。1976年到1979年，我研究后汉三国梵汉对音，遇上昙果在《中本起经》里把(梵)yaçah 耶舍译成“蛇蛇”、康孟祥译《兴起行经》把deva·(调达)译成“地婆”，总是“百思不得其解”。直到63岁，听人说了点儿孟加拉文概况才懂得印度方言真有把所有梵文y都念j，所有v都念b的，才后悔自己不用功——这些早就在《不列颠百科全书》14版里写着哪！别看围棋帮我“入静”，“入静”帮我好赖从动乱的十年活过来了，到了这儿还是太浪费时间了。

从去年下半年起，门人谢纪锋同志就劝我把这些文章结集。我腾不出手来，就由纪锋同志慨然担当起来了。现在这本集子就

是他们贤伉俪编起来的。这也非说一句不可。

“集之为言杂也”。这些文章涉及的方面不少，不好照论文选那样大致分类，再由外籍朋友那样加上标题。只按发表年代顺着排下来就是了。我想也没什么不方便的。谁要说这简直是叫人东撞一头西撞一头，我就说这可以训练人的耐性：能从第一页看到末一页。看书只看头五十页的人海了去了。真发现了不方便咱再改。

贯串着这些文章有两条线。一条是上推汉藏语的亲属关系，不过尽可能“言必有徵”，少摆七巧板。用汉朝人的话说叫“通经”。另一条是拥护拼音字，希望扫除愚昧，免得人们因为无知上别人的当。这叫“致用”。

用老话儿说，我已经到了“古稀之年”了。但愿我能看见拼音文字——一种真正供广大群众使，不由少数儿人摆弄的求知工具推行开了。到那个时候，人们能发现我不光是个咬文嚼字的人。

最后还要谢谢我的旧学生任国绪同志，要不是走到路上碰上，提起找稿子的事，这批稿子到现在还得在纪锋同志的书桌上躺着呢！

作者 1985 年 5 月 2 日

士林硯兄賜正

俞叔學

3/12



作者近照

目 录

序言	1
万斤重的小毛锥	1
语言与文字的失真性与独立性	10
认识 and 表达	18
古汉语里面的连音变读(sandhi)现象	31
古汉语里的俚俗语源	51
释甥	69
评《中国语言》(弗瑞斯特)	78
评《语言与文化》(罗常培)	86
什么叫一个词?	89
评《北京话单音词汇》(陆志韦)	107
说“跟”跟“和”	112
使动词	117
“了”跟“着”的用法	120
口语里“写不出”的字	126
语法札记(一)	129
语法札记(二)	134
汉语动词的形态	138
语音和语法	149
谈民族标准语问题	155
汉语的爱称和憎称的来源和区别	160
民族语言、文学语言跟地域方言	167

形态变化和语法环境	173
评《普通语言学》上册(高名凯)	181
汉语的句子	188
广播语言艺术的欣赏	200
我怎么看拼音方案草案修正案	210
主语和谓语	214
六书献疑	235
白话文的兴起、过去和将来	244
我听着别扭	251
叶斯柏森	255
汉藏两族人和话同源探索	262
古诗今读议	276
古四声平议	279
说“百家语”	284
与人论“六书”书	287
倒句探源	288
古汉语的人称代词	295
《诗》“薄言”解平议	300
谈谈动宾关系	309
古汉语“风”字确实有过象“字纛”的音	312
北京口语音位的出现频率	317
北京口语里的“给”字	327
佛教词语小议	332
等韵溯源	338
北京话的语音、词汇和语法特点	357
《国故论衡·成均图》注	377
李汝珍《音鉴》里的入声字	414

音轨和语素的出现频率	433
说到怎么分出义项	439
论词源学(译文)	443
○化石语素	464
不分开词会给读者添麻烦	471
哥儿 哥儿俩 哥儿们	476
北京音系的成长和它受的周围影响	480
王删郝疏议	492
《尚书·洪范》“土爰稼穡”解	500
从“看不见”说到“说不上来”	515

万斤重的小毛锥

一、半夜醒来，觉着右胳膊有些酸疼，在半醒半睡的状态中自己下了个断语——压的。好！翻过身来！这似乎该重新入睡了吧。然而不然，压是不压了，而酸疼如故。再说手腕子决压不着，偏是他疼得上紧，把胳膊伸到被外边去，冷空气带来了清醒的头脑——这都是那两三千誊写板小楷所赐，当时颇觉得不无愤慨，“××的！”我想，“从四五岁弄到三十岁，整天笔不离手，而写字还是我一个重担！如果生在印度日耳曼语系的国度里，打字机一响，反正要不了两个钟头。我亦不是大夫，亦说不出来写这些字得消耗多少加洛里的热，反正凭那分虽非九牛二虎可亦够瞧的劲儿来说，决非半个馒头所能补偿是可以断言，把一天精力的几分之一浪费在上面！每天半个馒头，一年二百来个，吃下肚去，变成力气，足顶四十个人的力量，五匹马力！而且，如果用打字机，省下一点钟来作些有益于国的事，譬如说罢，学甘地老先生的纺棉花，足可以纺个一两半两的。一年十斤，每斤换回十块钱的利权，全国识字的，以百分之二十计，八千万，好劲，八十万万！文盲扫净之后加五倍，五八四……不行，再算就别打算睡了。况且，中国文字本身就是扫文盲的一大障碍，不先扫他，又能扫得了谁？这方块字真不是个玩艺儿！可又说回来了，这方块字倒底算个甚么玩艺儿。”

二、据洋人说：他们的文字是拼音文字，我们的方块字则是、或曾经是表意文字(ideograph)。这话根本欠通，文字是记载语言

的符号，语言是传达意义的符号，识字的人看书报，都是先把文字翻译成语言——熟练的在脑子里翻，不熟练的采取“口中念念有词”的方式——再从语言中翻出语义来，世上那有直接表意的文字？不错，红印度人有一种连续的图画传达意义的办法，如先画二帐篷，再画一人右手放在嘴上，再画一人双垂手，来表示“家中食绝”之类，但是那是画儿啊！那么我们的方块字究竟是什么呢？这就须要一番历史的稍稍专门化的检讨了。世上没有不明病人身体现状就开方子的大夫，似乎亦不该有不研究对象就大叫改革的改革家，是不是？

三、从汉朝到清朝的小学家，都说中国字是仓颉开始造的，造字的法则是象形、指事、会意、形声、转注、假借六书。仓颉这人和他的事迹，我们无从考证，亦犯不上那么浪费精力去考证。六书的解释亦极多，转注一辞就有好几十种说，我们更犯不着去翻那个二千年陈粪坑，只就古文字之存于现代的——安阳出土的甲骨文和一些殷器铭文——来推测，文字起源是由于图画的。最早的文字是“画成其物，随体诘屈”的。例如𠂇字之为人犬的侧面轮廓是也。这也算是画，如果我们愿意沿用旧名词，便可以把这些字叫什么“象形”字。这些虽然同最幼稚的图画差不多，但是与图画有一个异点：在看一副图画的时候，我们可以得到“人”“立”“侧”“伸臂”几种概念，而看𠂇字时我们却不约而同地不注意后三者，这是为什么呢？因为文字是个社会的工具，如果没有这个“不约而同”或“共诺”，他就不会存在的缘故。但是这个共诺那里产生的呢？答案是：由于大家都把𠂇这个符号代表“人”这个语词，这组语言的习贯而来。即便𠂇这个字由某人造的，当他教给他的第一个弟子的时候，换言之，当𠂇不再是一个私人的产品而是一个社会的工具的时候，他已经和“人”这组语音牢结不解了，换言之，最原始的象形文字亦有是“表音”的。

四、到后来，经过社会的使用，每一个“字”——图形——常有代表一个以上的同音“词”——同样成分构成的两个语音组——的事。例如𠂇这个字本来在周初的器铭里是代表“在”这个语词的，但是由字形看来他却是为“才”而造。及至社会上觉得这种办法失之混乱的时候，便在旁边加上一个偏旁，来分辨一下，所以在西周中叶的孟鼎里就有了加土的“在”了。这种用一字表原词以外的词的办法，便是一部分学者如朱骏声等所认为“假借”但是并非造字之法。必得加偏旁的方法通行之候，我们才又有了一种新的造字法：“形声”。此法既行，随便取得两体一配，便得一新字，所以字就孳乳而孔多了。

五、此外还有些字是用二字配合，连续以成义的，如卫恒所云“止戈为武”“人言为信”之类。这两例是说文所举“会意”的例字，但却不确实：信本是从言千声，武字之止不过是人走的路图，正如𠂇字里一样，实在是画“武步”的概念而已。真正的会意字如“二手下”为拜，则是秦汉人所作，为数又极少。还有用现成的字予以增，如人加二为仁；减，如木减头为丕；变，如反从为比，倒大为𠂇等，则是一种间接的造字法，数又极少。所以我们说古代的造字法主要只有“象形”“形声”二法，是没有大语病的。

六、经过书写工具习惯的演变，中国字现在已经是僵化了的传统的工具，谁亦说不出来“鸟”为什么会是四条腿的动物，“月”形状有如梯子，“米”颇令人想起十字街头；至于“肉”则更无从辨其为肥瘦儿为五花儿与夫里脊了。形声字多到五六万，同一种偏旁的，意思可以差到无限，如“矫”和“矮”，“蝎”和“风”等等。又有些部首，因为字多，意义可以复杂到惊人，如口部里的“𠂇”“𠂇”是无论如何不易了解的，而清末地图上的“𠂇猓猓”恐亦非我盟友不列颠所愿闻者也。号称主谐字的声旁和全字的声音全无半点相象，如“其”之于“斯”，“奴”之于“帑”是也。象形既不象，谐

声亦不谐，于是绝大多数字之构成失去理性，在这种状态之下，有何意之可表？所以我们说：现代的中国文字，只是一套书写的单位。每一单位的结构是由传统来规定的，对非言语学者，毫无意义与道理可言。每一个字，大体是用来代表一组语言和发这一串音的时候的腔调的。因为现代中国语的非复合语词绝大多数是单音缀的 (monosyllabic)，而一字所表之音亦只有一音缀，所以我们可以说“每字代表一语词”是不大错的。中国语里的同音语词颇多，经瑞典珂罗伦教授统计，北平语的不同结构的音缀约有四百，用四种不同的腔调说起来，约一千五。但是报纸上常用的字足有五千或更多。因此我们凭着传统的赏赐，竟能大体上供应每一词一字了。现在我们回答“方块字是甚么玩艺儿”这个问题就比较容易了。我们可以说“现代中国字是以形异来分别同音异义词的音缀音标”。这才是他们的定义！梵藏文所用音标，都是每音缀一个符号。

七、现在我们可以进一步来讨论他们的优点和缺点了。先说优点。一班替他们辩护的人常说中国字好处是，即便是黑龙江人和广东人相遇，虽然言语不通，但是可以借笔谈达意。这个优点诚然不小。不过我们应当看看这好处来自那里。我们为了明白起见，可以先由别的语言借个鉴镜。比方说，拉丁语系吧。现在的意、法、西、葡和罗马尼亚人都是古罗马人的子孙，他们的语言都是拉丁语的子孙。经过长久的分居，这些兄弟之邦的语言各有变化，弄得彼此不能互相了解了。但是随便由那一种语言里取出一个字来，去和二千年前的母语拉丁语的原字相较，却总有几分相象。所以虽然意、法等人彼此不能了解，可是如果把他们的共同老祖宗拉丁人请一位来谈，大约他们都可以听懂一部甚至于大部。同样，取一个拉丁字来给他们看，他们亦大半会说出这个字怎么讲，在他本处的方言念什么来。再进一层，如果他们现在仍

然写的是他们的老祖宗所用的字，——这自然不可能，因为他们使用的主要是拼音文字——而用现在的意西等语音来读的话，那末每一个拉丁字在意、法、西、葡、罗人念起来自然是各不同，但是他们都能了解，因为他们所读的正是同一个语词的变形的原故。那么这类的例子究竟有没有呢？居然有几个。古罗马人用 I II III IV 等字来记数目，——请注意 I II III 和一二三相类似，这几个亦非拼音文字。——他们的后人现在还用着。所以当葡萄牙人写 IX 的时候，法国人、意国人、西班牙人、罗马尼亚人都能懂，虽然他们念得满不一样，甚至彼此听不懂，至于拉丁语的 IX 原是怎么念，他们甬说不知道，简直的就管不着！现在我们用这把镜子照照自己罢！我们都是炎黄的子孙。我们所说的话都是古汉语（Archaic Chinese）的子孙。我们沿用殷周时代的文字。我们用同字代表同语词的各种变型。广东人写一个“三”，黑龙江人、甘肃人都能懂，虽然他们念的满不一样，甚至于彼此不了解。至于殷周人把他怎么读，我们不但管不着，简直亦无从断定！回到本题。究竟这优点从何而来呢？曰：从“我们的语言出于共同的母语”和“我们使用同一套非拼音文字”这两件事而来。何尝是“中国字的优点”？阿拉伯人的数目字，普天下都懂——亦非拼音字母——那是因为他们结构异常简单的原故。我常想：假使由语义学家和教育家依 Roget 的 *Theasurus of the English Language* 规定几千抽象的概念，加上草木虫鱼鸟兽器物之名，共得若干，由图案画家配上几千极简单不混乱的符号，发动几千万同道，传给全世界人，那么一个人就环游全球永不会有不能达意的痛苦了。这真是一套极有用的工具，无量的功德，天下一家的门经，如果中国字真是这套东西，退一步说和阿拉伯数字同样简单的东西，我亦就不写这篇文章而去从事于以夏变夷的丰功伟业去了。无疑，洋人亦会放弃他们一部分文字而采用些中国字，如同他们对于阿拉伯数

字一样——他们并不管拼音不拼音，他们只问省事不省事——无如其不是何！

八、现在我们再来看看中国字的缺点。第一个缺点就是书写困难。我们的文字实在是太难写了。别国的文字，尤其是印度日耳曼系的，大体上只有几种用笔的方法：我们随便取一个英文字写写看，全字的总趋势是由“自左下向右上”形的线来规定的，保持着一个总的和谐。部分的形态则由许多曲线、直、和点三种基本笔法构成。我国的文字则基本的部分形态既繁——如王云五所分的十类号码所代表的有几十种，至少有“永”字之八法——而每一字用笔，忽左忽右，忽自上而下，忽自下而上，更是不可捉摸。用这种线条组织成的字写起来有许多麻烦，第一是时时变换用笔方向，例如“必”字虽然止有五笔，而需要至少六次变方向。第二是每一笔自成起讫，虽然写在纸上是一笔，如果从第一笔末到第二笔首的手在空中的活动算进去，那至少是共十九次动作，这九次空中动作，纯粹是精力的浪费，换言之，写中国字须浪费一半精力在空中动作上！第三是每字所占的空间大致相同，二十九笔——五十七次的动作——的“鬱”字和“一”占同样大的格，对于初学写字的儿童实在是一种苛酷的要求，注意力的浪费！——事实上我们并不大感觉此苦，正有如小脚妇女忘了初缠足之苦一样。除写殿试卷的超人们以外，连王右军亦不能把这个要求彻底实现。——有了这三种困难，再加上毛笔和墨，无怪乎老百姓不写，不学写，不想学写，不敢想学写中国字了！古人常以“如椽”来形容“大手笔”，其实对于老百姓，一枝小毛锥何止如椽简直重逾万斤！中国果然工业化了，小毛锥自会为自来水笔代替，我们现在用不着讨论这死刑已宣、不打自倒的工具。但是想清扫文盲的话，文字本身的麻烦非取消不可。

九、第二个缺点是字数太多，我先请读者不要误会，我所谓

字，是书写的单位，并非语言中表义的语言单位，世人对此颇有误解，如云英语有三五十万字而汉语只有几千字之类，英语之三十万，乃是语词，其书写单位只有二十六，外加 I II & @ 等非拼音字，共计不过九十。汉语之词，复合者颇多，如朋友、交情之类，假如以北平语的一千五百单词，配成两音缀、三音缀（口香糖）、四音缀（话剧演员，宋四六文）的复合词，其最多的数可以到下列式：

$$1500 \times 1499 + 1500 \times 1499 \times 1498 + 1500 \times 1499 \\ \times 1498 \times 1497 = ?$$

如果重叠字（轻轻、悄悄）亦算在内的话，数目还要大些，我们当然不能用这许多字来记这许多可能的词，我们只能供给每一音缀三五个字。请看，我们省下了一个十三位数之多的书写单位！不过以五千御若干万万总还不如以二三十御若干万的省事吧！北平话四百音缀不过是十七个辅音和十三个元音组成的，如果把这三十个音，每一个给他一个符号，此外再加上四个调号，那末千言万语无不可写了。假使记元音用“每个音位一符号”的原则（one letter per phoneme）那就还可以省去五个，我们何必要粘滞在五七千上呢？

十、第三个缺点是字数太少。这话乍看好象和（九）正矛盾，我的真意是说我们并没有能真正供给每词以一字：我们有好多“有音无字”的词，这使我们不能写出生动的口语，而只可迁就用些文言成分，这些词有一部分，本来是有字的，不过因为读音有些变了；如北平话之“别扭”实在即是“悖谬”。或语义引申了。如“小大姐儿裁絨子”之“絨”，小儿尿布，实即“藉”（磨）字。或是根本从未见人用过，所以未学会。如“拐弯抹角”之“抹”即释名之“陌”，常会变成“有音无字”。有一部份言语学者把他们的应有的写法找出来再用古代的字典来加以证明，亦获得了若干成功，这种

工作是他们称为“考本字”的工夫之一部，最多得成功是章太炎先生的“新方言”，经此法探知的字自然该复其本位。但是往往却得不到社会的通行。除去这些以外，就得靠着新造，最著者，化学之氧氢是也。但是仍然还有些能写不能写的，如鲁南人谓谈话为“/la kuar\”，却是无古字可考、又无新造，只好认为“不登大雅之堂”，不以入文了。所以中国文学内容之枯窘亦有一部是由字之缺点促成的，好多人之不敢为他的女仆写家信，职是之故也。

十一、第四个缺点是排列困难，自从小学开始发达以来，对于中国字之分类与排列，就是一个恼人的问题，许慎、第一位字典学者，从史游的急就里采取了“分别部居不杂厕”的办法，用五百四十个部首来统率九千三百五十三文，可以说是巨大的成功，不过即使在说文中亦有不规则的字，非部首所可统摄的，例如“豌”在“豆”部，就与全部的字冲突，还有从蜀益声的“蠲”字放在“虫”部亦是无可奈何的事情，这都显示出来一个事实：中国形声字的形旁异常复杂，多亏许君的天才与三四十年的功力，才勉强从一团稀糟里整理出头绪来，但是这五百四十部的排列又是个大问题，许氏的解决方式是“据形系联”，这一点他亦勉强作到了，但并不十分圆满，因为这五百四十部根本并不能安排成一个“甲象乙，乙象丙……”的次序，所以段玉裁注部首第时常说：“无所蒙”，其意若曰“联不上了”而已。一个非专家查说文的时候，极不易知道某在那一篇，这是他的一个大短处，到明末耶稣教士金尼阁又发明了一套部首为康熙字典所沿用，直至今日，这一套部首的短处亦颇多，详见王云五和吾师洪煊莲先生的《中国字皮瓣》的解释文字里。王的四角法的分类与排列据洪先生说尚有三十余字号相同之病，而洪先生之皮瓣亦还是有十余字同号。所以中国字的排列和分类的困难可以说还没有彻底解决，至于他所给予编查的字典的人，收藏图书和卡片的人的麻烦，实在是不可胜计的。