

◎汉语字本位研究丛书
徐通锵 主编 潘文国 副主编

汉字信息语法学

Hanzi Xinxì Yafaxue

鲁川 王玉菊 著

山东教育出版社

图书在版编目(CIP)数据

汉字信息语法学/鲁川,王玉菊著. — 济南:山东教育出版社,2008

(汉语字本位研究丛书)

ISBN 978-7-5328-5914-6

I. 汉… II. ①鲁…②王… III. 汉语—语法—研究
IV. H14

中国版本图书馆 CIP 数据核字(2008)第 033548 号

汉语字本位研究丛书

徐通锵主编 潘文国副主编

汉字信息语法学

鲁川 王玉菊 著

主 管: 山东出版集团
出版者: 山东教育出版社
(济南市纬一路 321 号 邮编:250001)
电 话: (0531)82092663 传真: (0531)82092661
网 址: <http://www.sjs.com.cn>
发 行 者: 山东教育出版社
印 刷: 山东新华印刷厂临沂厂
版 次: 2008 年 3 月第 1 版第 1 次印刷
印 数: 1—2000
规 格: 650mm×970mm 16 开本
印 张: 29.25 印张
字 数: 366 千字
书 号: ISBN 978-7-5328-5914-6
定 价: 48.00 元

(如印装质量有问题,请与印刷厂联系调换)

(电话:0539—2925659)

本书根据下列科研项目最后成果写成

国家语言文字应用“十五”科研项目
《信息处理用规范汉字字义统计和造词模式》
(YB105—48)

并引用了下列科研项目的中间成果

2004 年度国家社会科学基金项目
《信息处理与对外汉语教学的句子语序模式》
(04BYY010)

因字而生句。

——刘勰 501《文心雕龙》

说文等于语法 (grammar)；

音韵等于语音 (phonetics)；

训诂等于语汇 (vocabulary)。

——林语堂 1969《林语堂选集》

内容提要

《汉字信息语法学》是适应信息时代需求的一部语法学专著。工业时代实现体力劳动机械化,信息时代将实现脑力劳动机械化。在信息时代的知识经济中最重要的是知识。知识的主要载体是语言。电脑要处理知识,先要理解语言。于是语言学成为不同领域科学家们共同关注的焦点,出现多种科学与语言学相结合的认知语言学、对比语言学、计算语言学以及“信息语言学”。信息语言学的主要分支是“信息语法学”。本书只讨论语言的结构规则,不涉及信息处理的技术,通俗易懂。

英语的基本结构单位是“词”,汉语的基本结构单位是“字”。

英语是“形态语言”,语法学分为“词法学”和“句法学”。

汉语是“意合语言”,语法学分为“字法学”和“句法学”。

汉语和英语的差异都取决于各自的“认知模式”。

汉族人的认知模式侧重于:整体感知、意象思维、类比推理。

英国人的认知模式侧重于:细部感知、抽象思维、形式推理。

“整体感知”使汉语语音敏感单位是“音节”,导致“字本位”,而“细部感知”使英语语音敏感单位是“音位”,导致“词本位”。

“意象思维”使汉语的文字成为注重视觉信息的二维图形,而“抽象思维”使英语的文字是记录听觉信息的一维字母串。

“类比推理”使汉语造词法主要是偏正式的“特征+类别”,而“形式推理”使英语造词法主要是名词和动词的“词形变化”。

汉语用单音节和方块形结合的一个“字”来表达“基元概念”,用多个“字”来表达“复合概念”,这使“概念系统”简明清晰。造词模式“特征字+类别字”符合认知世界“概念系统”的公式:“下位概念=个性特征+上位概念”。

信息时代新知识新概念急剧增长,作为西方语言基本结构单位的

“词”也相应地急剧增长。但作为汉语基本结构单位的“字”却并不增长,因为可以用有限的熟“字”造成无限的新“字组”来表达急剧增长的新知识。

上卷字法学一至四章:字形学、字音学、字义学、字汇学。

字形学研究字形演变、字根及造字模式、检字法和输入法等。

字音学研究字音(音节)的音·韵·调、声符的示音四边形等。

字义学研究字义和概念、意符的类义功能、字义是框架义等。

字汇学研究词是最小固定字组、造词法、字汇的运用技巧等。

下卷句法学五至八章:句构学、句意学、句模学、句序学。

句构学研究汉语句法结构特点、词类功能、汉英句法对比等。

句意学研究句子的语意网络、语言的主观信息和依附标记等。

句模学研究语言模型、谓词配价、基本句模和特殊句式等。

句序学研究语序的预想论、语块的词序和句子的语块序等。

汉语语法研究有两个重点:①“字”,②“序”。

“字”是按“序”组成“汉语”的基元性单位。

“序”是用“字”组成“汉语”的理据性排列。

如果把“语言”比喻为“建筑物”,那么本书的编写体系是:

上卷“字法学”研究的是:作为“建筑材料”的“字”。

下卷“句法学”研究的是:作为“建筑法规”的“序”。

本书是作者根据二十年从事中文信息处理、机器翻译、对外汉语教学的经验和所主持的国家级科研项目研究成果编写而成的。对汉语语言专业和有关中文信息处理的科学技术专业的师生以及关心上述领域的广大读者有重要的参考价值。

《汉字信息语法学》是我国第一部按“字法·句法”体系编的汉语语法学专著。作为一次尝试,本书必有不成熟之处,欢迎广大读者和专家给予批评指正。

ABSTRACT

A Sinogram-based Informative Grammar of the Chinese Language is a treatise to meet the needs of the era of information. Mechanization of physical labor was realized in the industrial era, and mechanization of mental labor will be realized in the era of information. In the intellectual economy of information era, knowledge becomes the most important factor. The major carrier of knowledge is language. In order to process knowledge, first of all, computers have to understand language. Therefore linguistics becomes the focus concerned by scientists in different domains, and many cross-disciplinary sciences, such as Cognitive Linguistics, Contrastive Linguistics, Computational Linguistics and Informative Linguistics. The main branch of Informative Linguistics is *Informative grammar*. This book only deals with the structural rules of the Chinese language without touching the technology of information processing. So this book is easy to understand.

The basic structural unit of English is *word*, and the basic structural unit of Chinese is *Sinogram*.

English is a morphological language, and its grammar is composed by *Morphology* and *Syntactics*. Chinese is a semotactic language, and its grammar is composed by *Grammology* and *Syntactics*.

The differences between Chinese and English depend on their cognitive models.

The Chinese cognitive models lay particular stress on “*perceiving the objective as a whole*”, “*thinking in images*” and “*analogical inference*”. The English cognitive models lay particular

stress on “*perceiving the objective in details*”, “*abstract thinking*” and “*formal inference*”.

“*Perceiving the objective as a whole*” results in the fact that the sensitive phonetic unit of Chinese is *syllable* and that leads to the Sinogram as the basic unit of Chinese, on the other hand, “*perceiving the objective in details*” results in the fact that the sensitive phonetic unit of English is *phoneme* and that leads to the word as the basic unit.

“*Thinking in images*” results in the fact that the scripts of Chinese is *two-dimensional figures*, but “*abstract thinking*” results in the fact that the scripts of English is *one-dimensional string of letters*.

“*Analogical inference*” makes the word-building of Chinese mainly characterized as “*feature + category*”, and “*formal inference*” makes the morphology of English mainly characterized as *inflection* such as *declension* and *conjugation*.

In Chinese, *the primal concept* is expressed by *a Sinogram* which is the combination of monosyllable and square-shape, and *a compound concept* is expressed by *several Sinograms*. This makes the concept system simple and distinguishable. Word-building model “*feature + category*” conforms to formula of concept system in the cognitive world:

Hyponymy concept = Individual feature + Hypernymy concept.

The new knowledge and new concepts in the information era are rapidly increasing, correspondingly, the words as basic units of Western languages are also increasing rapidly. But the Sinograms as basic units of Chinese do not increase, because the limited number of Sinograms can be assembled to create numerous new gram-groups to express rapid-increasing new knowledge.

Volume I *Grammology* consists of four chapters:

Chapter 1 ***Graphology of Sinograms*** studies Evolution of Sinogram-shapes, Gram-roots and Sinogram-building models, Indexing system for Sinograms, Keyboard-input of Sinograms, etc.

Chapter 2 ***Phonology of Sinograms*** studies Structure of Chinese syllable, Rhymes and rhyming tracks, Sound-suggesting quadrilateral of the phonological symbols, etc.

Chapter 3 ***Semology of Sinograms*** studies Sinogram-meaning and concept, Functions of distinguishing semoclass of semological symbols, Cognitive frame as the basis of gram-meaning, etc.

Chapter 4 ***Lexicology of Sinograms*** studies Words in Chinese as fixed gram-groups, Word-building of Chinese, Equations of word-meaning, Applying Sinograms in a skillful way, etc.

Volume II ***Syntactics*** consists of four chapters:

Chapter 5 ***Syntactic structure of sentences*** discusses Peculiarities of syntactic structures in Chinese, Grammatical functions of word-classes, Contrast between Chinese syntactics and English syntactics, etc.

Chapter 6 ***Semotactic networks of sentences*** discusses Semotactic roles and semotactic relations, Pivot roles and periphery roles in collocational relation, Subjective information of language and attachment marks, etc.

Chapter 7 ***Semotactic models of sentences*** discusses Language models, Valency and essential argument of verbals, Basic sentence models of Chinese, Special sentence constructions of Chinese, etc.

Chapter 8 ***Chunk sequence of sentences*** discusses *Theory of preconception* in studies on language sequence, Word sequence in a periphery chunk, Chunk sequence in the conventional sentence, Chunk movement in the communicative sentence, etc.

There are two main focuses in the studies of Chinese grammar:

① ***Sinograms***, ② ***Sequence***.

The *Sinograms* are the basic primal units assembling Chinese according to the *Sequence*.

The *Sequence* is the reasonable permutation assembling Chinese with the *Sinograms*.

If we compare the language to a building, then the written system of this book is as following:

Volume I *Grammology* studies the *Sinograms* as the *building materials*.

Volume II *Syntactics* studies the *Sequence* as the *building regulation*.

This book was written on the bases of the authors' twenty years of experiences in *Chinese information processing*, *Machine translation*, *Teaching Chinese as a second language*, and research results of National scientific research projects directed by the authors. It will serve as a valuable reference material for teachers and students majoring in Chinese linguistics, Information science and technology, *as well as those with great interests in the above disciplines mentioned*.

A Sinogram-based Informative Grammar of the Chinese Language is the first Chinese grammar book based on the system of *Grammology* and *Syntactics*. As the result of an tentative study, there must be some space for improvement in this book. Criticism and suggestions from the readers are sincerely welcome.

“字本位”和语言研究

——《汉语字本位研究丛书》总序

徐通锵

人们将以字为汉语基本结构单位的语言研究思路称为“字本位”。这一说法不一定很准确，但简单明了，因而日渐流行，我们这些倡导这种研究思路的人自然也乐于接受。

“字本位”这个概念最早见于郭绍虞写于1938年的《中国语词的弹性作用》一文，以“字本位的书面语”与“词本位的口头语”相对。此后人们对此鲜有讨论。近20年来，由于人们越来越感到现行的语言理论难以有效地分析汉语的结构，于是探究原因，重新提出“字本位”的问题，不过它不限于书面语，而是指称以字为汉语基本结构单位的一种新的研究思路。山东教育出版社认为“字本位”是一种有发展前景的语言理论，值得倡导和鼓励，因而决定出版一套《汉语字本位研究丛书》，由徐通锵、潘文国任正、副主编。2005年4月1~2日，山东教育出版社在济南东方大厦召开了《汉语字本位研究丛书》编写会，应邀与会的有戴汝潜、鲁川、吕必松、孟华、潘文国、汪平、杨自俭和我8人，一起讨论《丛书》的编写事宜。这几位先生都是“字本位”理论的鼓吹者和实践者，在语言基础理论研究、对外汉语教学、中文信息处理、汉语方言研究、对比语言学、符号学和中小学语文课程教学等领域都已进行过有相当深度的研究，确信“字本位”是适合汉语研究的一种理论和方法。

本丛书的主题,顾名思义,自然是“字本位”,就是把“字”看成汉语的基本结构单位,并以此为基础探索汉语的结构规律、演变规律、习得机制、学习规律和运用规律。本丛书就是对这些规律的探索。以“字”为汉语基本结构单位的“字本位”理论与现在通行的以“词”为汉语基本结构单位的语言理论是两种不同的理论体系,因而必然会产生尖锐的冲突。“词”是印欧系语言的一种基本结构单位,汉语的“词”是从西方语言学中输入的一种语言结构单位,其内涵与“字”有重要的区别。基本结构单位是统率语言结构的“纲”,纲举目张,所以“词”与“字”虽然仅仅是一字之差,但与两个不同的“纲”相联系,体现语言理论研究的不同走向。我们为什么要弃“词”而选“字”,倡导“字本位”呢?就是由于“词”是一种舶来品,在汉语中没有“根”,而形、音、义三位一体的字是汉语的载体,而且也是汉文化的“根”,因而需要以“字”这个“纲”为基础探索汉语的结构规律、演变规律、习得机制、学习规律和运用规律,不然就难以有效地实现语言研究的预期目标,找到普遍有效的规律。百年来汉语研究的实践已为此积累了丰富的经验和教训。

鉴于“字”与“词”是和两种不同理论体系相联系的基本结构单位,我们自然需要在这里对“字本位”理论思路的来龙去脉进行一些必要的梳理,申明立论的缘由,为探索汉语研究的新途径做一点鸣锣开道的工作。

西学东渐和汉语研究传统的中断

语言研究有三大发源地,形成三大语文研究传统,即一般所说的希腊-罗马、印度和中国。鸦片战争后,西学东渐,敲开了中国学术研究的封闭的大门,语言学也以《马氏文通》为标志走上了与西方语言学的研究传统相结合的道路。希腊-

罗马和印度两大传统诞生于印欧系语言的研究。这类语言富有形态变化,以“词”和“句”为基本结构单位,两大传统研究侧重点虽然有别,但核心与方向一致,都是围绕着词与句的关系展开它们的语法研究,探索名、动、形的词类划分和它们与语句结构成分之间的有规律的对应关系,建立语法理论。语法,这是两大传统的核心,或者说,是印欧语研究传统的核心。汉语没有形态变化,没有印欧语类型的那种“词”和“句”,因而也就没有印欧语类型的语法;它以“字”为基本结构单位,从文字、音韵、训诂三个方面研究字的构造规则,突出语义,形成汉语特有的研究传统。西学东渐,发生了印欧语传统与汉语传统的接触和碰撞,从而促使中国语言学家去探索这两大传统相互结合的途径和方法。当时的社会思潮是“一边倒”学习西方,汉语的研究自然不会游离于这一思潮之外,因而也是师法印欧语的研究路子,建立中国的现代语言学,试图实现与印欧语研究传统的结合。

不同传统的结合是推进语言学发展的强大力量,印度传统和希腊-罗马传统的结合催生历史比较语言学,使语言研究摆脱哲学、历史学等学科的附庸地位而成为一门独立的学科,就是一个有力的证明。但是,汉语研究传统与印欧语研究传统的结合碰到了以往不曾碰到的困难,这主要是:一是基本结构单位不同,汉语是字,印欧语是词和句;二是研究的核心内容有别,汉语重语义的研究,主观的因素参与语义的解释,而印欧语重语法的研究,讲求对形态变化规则的把握,排斥主观因素的参与;三是汉语重视觉的“形”的研究,汉字在汉语研究中占有重要的地位,而印欧语重听觉的“音”的研究,遵循索绪尔的“语音中心主义”原则,排斥文字在语言研究中的地位。面对这些矛盾,汉语研究的应对策略比较简单,就是接受印欧语的理论和方法,用它来解释汉语的结构。

说得具体一点,就是以“词”和“句”为汉语的基本结构单位,否定“字”的语言学性质,用印欧语的理论、方法建立汉语的语法体系。这里的核心是语言基本结构单位的认定和把握,用“词”代替“字”做汉语的基本结构单位,实际上也就放弃了汉语结构的“纲”,将汉语纳入印欧语研究的理论框架中去分析。所以,汉语研究采用这样的应对策略去迎接西学东渐的思潮,实现中西语言学的结合,只能是利弊并行。这就是:一方面诞生了中国现代语言学,从传统走向现代,但另一方面是使中国语言学自觉地中断了汉语悠久的历史传统,以《马氏文通》为界,“现代”和“传统”脱节。这种利弊相生的现象说明我们对中西语言学的结合的应对策略出现了偏误,因为以中断自己的悠久传统为代价去接受另外一种传统,这不是不同传统的结合,而是“张冠李戴”式的替代。结合,应该是学术思想的相互渗透,从不同传统中吸取自己所欠缺的理论和方法,融会贯通,从中找出驾驭不同传统的语言共性结构原理,建立新的理论,既能解释不同传统的异同的成因,又能开创新的研究前景,推进语言学的发展,犹如希腊-罗马传统与印度传统的结合催生历史比较语言学那样。所以,结合,是一种更高层次的理论创新,汉语研究没有因西学东渐而实现这样一种理论创新,说明中国现代语言学还没有找到汉语研究传统和印欧语研究传统相结合的正确途径。

面对不同传统的接触、碰撞和结合,汉语研究的应对策略为什么会产生这类偏误?最重要的原因就是放弃和否定了“字”的汉语基本结构单位的性质,而把印欧语的“词”和“句”作为自己的基本结构单位。语言理论和研究方法都是在基本结构单位的研究基础上提炼、总结出来的,否定了“字”的语言结构单位的性质,自然也就中断和否定了以字的研究为基础的汉语研究传统,使不同传统的结合走上了歧

路。传统,它凝聚着历代学者对语言特点的认识,只能继承、完善和发扬,绝不能中断,更不能否定;中断和否定了自己的传统,也就只能接受别人的学术教条,丧失自己独立的开创性研究的能力。《马氏文通》以来,中国语言学的许多学者都在进行汉语语法的研究,用印欧语的语法理论来解释汉语的结构,著述尽管很多,但进展不大,原因就是“汉语语法研究不应该模仿印欧语,古汉语也一样。中国本来没有语法,对外来的东西可以学习、吸收、借鉴,这并不错。问题是你老跟别人走,看不到自己的特点,就不对”,“我们的汉语有我们自己的规律”,但由于“受印欧语的影响,不知不觉的影响。这个东西使得我们不能往前走,问题早就提出来,但摆脱不了,这是因为先入为主。各个学科都有这个问题。科学最可怕的是一种教条,或者是框框……总觉得这是大家这样说的,不应该有问题呀!其实,问题就出在这儿。过去荒谬的东西,现在都变成了真理。我们语言学也不例外”(朱德熙1982)。印欧语的理论、方法所以会成为束缚我们前进的学术教条,最根本的原因就是我们否定了汉语自己的基本结构单位“字”,由此也就不能不遵从这种教条来研究汉语,中断自己的研究传统。本丛书编纂的目的是改变这种结合的思路,重新确立“字”的基本结构单位的性质,并以此为基础,吸取西方语言学理论研究的思路,也就是不拘泥于它的现成结论,而是参照它如何在具体语言材料的基础上升华、提炼相应的理论的精神,展开相关的研究,探索汉语研究的途径,使“现代”与“传统”相衔接,实现中西语言学的结合。这种想法不自我们今日始,前辈学者早就提出过相关的问题,我们只是想在他们的基础上前进一步。

汉语特点的直觉感悟和独立研究意识的觉醒

《马氏文通》以后的中国现代语言学基本上都是根据印

欧语的理论框架,以“词”为汉语的基本结构单位,探索“词”与“句”的关系,寻求名、动、形的词类划分和它们与句子结构成分的对应性关系。但是,由于这种思路远离汉语的“字”这一基本结构单位,因而成效不大,虽经若干次周期性的大论争,也没有找到一条摆脱困境的出路。实践的挫折使一些有经验的语言学家把目光转向“字”的地位和作用。赵元任、吕叔湘、王力三位先生是中国现代语言学的开创者,他们在研究实践中已清楚意识到“字”与“词”的矛盾,因而都以不同的方式强调“字”的重要性。这里不嫌啰嗦,将三位先生的几段相关论述引述于下:

汉语中没有表示“word”的词。最贴近的是“字”,翻译过来是“word”,但实际意义是“音节”或“音节成分”。——赵元任(1940)

“如果我们观察用某一种语言说出的大量话语,例如英语,考虑一下这些话语中小片段的情况,并拿它们跟汉语中同样的小片段作个比较,我想,‘字’这个名称(这样说是因为我希望先避免把 word 这个词用于汉语)将和 word 这个词在英语中的角色相当,在说英语的人谈到 word 的大多数场合,说汉语的人说到的是‘字’。这样说绝不意味着‘字’的结构特性与英语的 word 相同,甚至连近于相同也谈不上。‘字’和 word 的关系就好比通常用‘橘子’对译英语的 orange,其实橘子在构造上属红橘(tangerine),与 orange(甜橙)是不同的植物”,“汉语中没有词但有不同类型的词概念”,“按西方语言学家的眼光来分析汉语并确定像结构词这样的单

位可能有用……但这不是汉人想问题的方式,汉语是不计词的,至少直到最近还是如此。在中国人的观念中,‘字’是中心主题,‘词’则在许多不同的意义上都是辅助性的副题,节奏给汉语裁定了这一样式。”——赵元任(1975)

汉语词汇的基本单位还是一个一个的单字。应该把现代汉语中最有活力的两千来个字(估计不超过此数)给学生讲清楚。不能把汉字只看成符号,像对待外国语的字母那样。——吕叔湘(1962)

“词”在欧洲语言里是现成的,语言学家的任务是从词分析语素。他们遇到的是 reduce(缩减),deduce(推断),produce(生产)这些词里有两个语素还是只有一个语素的问题。汉语恰恰相反,现成的是“字”,语言学家的课题是研究哪些字群是词,哪些是词组。汉语里的“词”之所以不容易归纳出一个令人满意的定义,就是因为本来没有这样一种现成的东西。其实啊,讲汉语语法也不一定非有“词”不可。那末为什么还一定要设法把它规定下来呢?原来“词”有两面,它既是语法结构的单位,又是组成语汇的单位,这两方面不是永远一致,而是有时候要闹矛盾的。讲汉语语法,也许“词”不是绝对必要,可是从语汇的角度看,现代汉语的语汇显然不能再以字为单位。用汉字写汉语,这个问题还不十分显露;如果改用拼音文字,这个问题就非常突出了。所以汉语里的“词”的问题还是得解决,可是只有把它当作主要是语汇的问题来处理,而不专门在