

文献·理论·信息*

——试论中医古籍语言系统构建的三个重要维度

□朱玲** 崔蒙** (中国中医科学院中医药信息研究所 北京 100700)

摘要:本文从文献、理论、信息3个不同的方面分析了构建中医古籍语言系统的重要维度,提出文献是基础,理论决定深度,信息是关键的论点,认为在语言系统构建的过程中,虽不能离开信息技术的支撑,但对文献中概念内涵的准确把握、对中医理论的深刻理解是不容忽视的重要方面。

关键词:文献 理论 信息 语言系统 中医古籍

一、相关研究工作背景介绍

随着信息技术日新月异的发展,其应用领域越发宽广,医学领域亦不例外。美国早于1986年便已开始建立一体化医学语言系统,对于医学的发展起了重大的推动作用,并于1999年^[1]被介绍引入国内,进行中医药一体化语言系统(TCMLS)的设计,2002年开始进行具体词条的加工。中医学作为有中国特色的传统医学体系,是我国医疗力量的重要部分,中西医并重是我国医学发展的重要政策。就一体化医学语言系统而言,能否以及如何处理中医学知识体系,是一个极其重大的学术及实践问题。目前已经实现了TCMLS的远程发布与浏览,实现了多版本关联查询与检索,同时为数据加工与改造、存储等设计了良好的环境。

经过数年的运行,这个系统当然也不可避免地存在一些问题或争议之处,有待完善。至于这些问题,究其原因,可能与当时收词范围、信息技术方法的局限有关。从另一方面来看,不能忽视的是中医学特殊的学科性质,尽管在进行中医药一体化语言系统构建时已经对若干语义类型和关系进行增删改动,如删减了“练习”;“展示”两个语义关系,增加了“开窍于”;“与…相表里”;“概念上对应”;“与…类似概念”;“与…相互依存”;“与…相互制约”;“与…相互转化”等7个语义关系^[2],但是还不能完全匹配古籍语言系统的需求。

中医学是具有强烈的学术传承性质的学科,历代的古籍是中医学的主体,也是后世发展的源头活水。而中医古籍又是以古汉语为表述形式的。中医理论是以概念术语为核心构建的。因此,可以说,没有中医古籍内容参与的中医药一体化语言系统是不能完全或真实地反映中医学体系实质的,也就不能

收稿日期:2009-07-24

修回日期:2009-08-09

* 科技部基础性工作专项(2005DKA32405):中医药学科科学数据中心-科学数据共享工程医药卫生科学数据共享网,负责人:刘保延,崔蒙;中国中医科学院基本科研业务费自主选题项目(Z02039):中医古籍语言系统框架构建的探索性研究,负责人:朱玲。

** 联系人:朱玲,博士,主要研究方向:中医药信息学,Tel: 010-64014411-3228, E-mail: jjzhuling@163.com;崔蒙,本刊编委,博士生导师,主要研究方向:中医药信息学的学科建设,中医药信息数据库与网络建设,中医药信息数据的挖掘与利用研究,以及中医药软科学战略研究, Tel: 010-64013395, E-mail: cui@mail.cintcm.ac.cn。

发挥它的应有作用。因此,对中医古籍进行语言系统构建的研究是相当必要和紧迫的。从目前学界的研究现状来看,这方面的系统尚未进行,只有杨继红^[3]对《千金方》进行了初步研究,但还不够全面,不能真正反映古籍和中医学术的特点。鉴于此,我们提出进行中医古籍语言系统构建的研究工作,并得到了中国中医科学院自主选题项目的资助,目前研究工作按计划顺利开展。本文将针对前一阶段的研究工作,阐述我们对于中医古籍语言系统构建的若干进一步的深入思考,希望得到前辈同道的批评指正,共同促进此研究的顺利进行。

中医古籍语言系统的建设究其本质属性,具有强烈的交叉学科性质,包含中医学与信息学的特点,在进行该系统的构建过程中必然要充分考虑这种特点。本研究认为,文献、理论、信息是中医古籍语言系统构建的3个重要维度,下文将试从这3方面逐一论述。

二、文献之维

众所周知,中医古籍文献浩如烟海,体系庞大,门类繁多,是中医学术理论传承的载体,是中医古籍语言系统构建的基础和原始处理对象,也是首要环节。换言之,轻视或忽略文献的特点将会影响该系统的构建。

进行任何古籍文献的研究,首要的问题便是文献版本,其优劣决定着文献的真实程度。最大程度上可信而真实的文献,是任何研究工作的前提和基础,对于中医古籍语言系统的构建也不例外。在此坚实文献基础上构建的语言系统才能真正展现和深刻揭示其相互关联,从而为学习、理解、运用中医学理论知识提供帮助。

除了一般所具有的古文献的特点,中医文献还具有其自身的特点。比如,本草文献的流传便具备极其鲜明的特点,即后世本草书在《神农本草经》的基础逐渐添加主治功用,进行了所谓的“滚雪球”式的文献积累。明堂文献的流传便与此不同,其特点是,相互传抄,但时常不注明其来源,有的甚至是抄自以前的数家文献,类似于“杂烩”式。这样的弊端是,后人易于不明某些文献的来源,造成认识上的混乱。因而,找出源头文献,梳理出清晰的文献传承脉络,厘清各类古籍文献之间的学术关联,对于中医古籍语言系统的构建是极其重要的基础工作。再比如经典

注释性的文献,《内经》《难经》《伤寒》《金匱》等,这些是中医学术的源头活水,是学术的主干,后世医家对之有不少的注释文献。这个注释系统也存在一定的特点。

中医古籍分类在某种程度上是语言系统构建的前提,要对纷繁复杂的中医古籍进行分类,离不开对文献本身的深入考察和理解,对历代古籍分类情况的把握以及对古籍本身蕴含的知识全面的领悟。考察现有的分类,无论是《中图法》还是相对比较适合古籍的《联目》分类法,均不能较好的解决构建中医古籍语言系统所需要的分类体系。因其或注重学术流派,或关注图书本身,而对文献的内容关注较少。语言系统整体框架内古籍部分的分类显然不能局限于文献图书本身,而必须立足于文献的内容,方能更好的分析把握其含义及相关关系。这也就是所谓的知识分类法,即通过不同的类型将中医古籍中所蕴含的知识尽可能完全呈现,以利于现代人方便、快捷地获得这些知识,甚至是借助强大的类表谱系树来进行知识挖掘,凸显隐含在古籍文本形式中的知识体系之间的相互关联。

三、理论之维

语言系统构建的目的在于,展示同一概念在不同时代,不同书籍中的涵义变迁,以及概念与概念之间的语义关系,并最终建立古籍的语义网络。而概念术语的语义类型、关系等又是基于其涵义的正确而富有历史感的理解而达到的,那么如何才能达致这样的理解呢?因为有很大一部分概念术语的涵义不是自己呈现的,而是需要通过考证、对比来揭示的,这依赖于文献和理论研究两方面的保证。中医古籍或概念本身具有强烈的古代文献属性,而古代文献是有其自身的流传特点的,其中包括历史时代性的特点、学科属性的特点(即某一类型的文献流传特点,如针灸类、明堂类、注家类等),忽略这些特点便难以把握其文献内容的本质与相互关系。

目前的语言系统构建基本是以教科书或某些工具书的概念术语的释义作为蓝本的,但是我们深知,无论是教材还是工具书均有自身的缺点。如教材的性质更倾向于前后统一、协调,而中医理论内部是相当复杂的,因此就会舍弃一些内容或观点而达致人为的统一,以这样的教材为蓝本的语言系统显然是不能反映中医学术理论的本质的。例如,教材所谓的

经脉理论即是《灵枢·经脉》篇所载,实际上《内经》中包含多种形式的经脉理论(十一脉形式经脉理论、十二脉向心形式的经脉理论、足六经理论等),但由于《经脉》篇十二经脉循环理论在后世获得了经典性的地位而被广泛认同,如果只是按照教材的经脉理论,必然会遮蔽或遗漏早期的其他形式的经脉理论,这样丰富多彩的经脉理论形式便不能在语言系统中得以呈现。工具书虽然也是学习中医理论知识的重要帮助,对概念术语的解释也尽可能包含了较多的义项,但应当注意到的是,这些义项都是一条条独立呈现的,而从理论发展的历史进程来看,这些义项的解释都是在理论的历史传承中迁衍而来的,相互之间的关系极为密切。再者,概念与概念之间的关系也是相当复杂的,而这些在一般工具书中并没有体现,这也不利于语言系统的构建。上述这些内容,无论是概念丰富的涵义,还是概念之间的相互交错的关系,甚至学术理论本身演变的轨迹,都需要通过理论来深入体现。但这也并不意味着可以抛弃教材与工具书,毕竟它们有着相对成熟的形态和广泛接受的基础,对于当前的中医古籍语言系统构建来说还是相当重要和必须的。与此同时,我们应当重视从学术理论的角度去挖掘概念术语涵义及其相互关系,至少应当密切关注这方面的最新研究成果,或采取与这方面权威专家协作的模式,为中医古籍语言系统构建提供扎实的学术基础。

四、信息之维

从信息学的角度来看,古籍语言系统构建有两种可选的途径,其一,在中医药语言系统之外,单独构建中医古籍语言系统;其二,在已有的系统内,进行古籍语言系统的建设。前者有利于古籍自身特点的全面展现,相对独立的系统也符合中医古籍在整个中医学学术中的重要地位,但后者则更利于古籍与现代文献之间的兼容与互补,且可避免系统重建的繁复浪费。权衡现有的条件,本研究选用后者,即在现有的中医药语言系统的内部进行古籍语言系统的试加工。

本研究的具体步骤大致如下,首先在现有的中医药一体化语言系统 16 个大类中增加古籍术语一项,下设内经、伤寒论、金匱要略 3 个大类;以《内经词典》等几本词典类工具书作为加工的最初概念来源,每一个词条在相应的章节目录下进行加工,以章

节名作为暂时的框架进行加工,但不是最终古籍框架的展现。每一个概念都必须添加其上位词,在完成这 3 本词典的加工后,通过对上位词的提取,进而建立古籍语言系统的最初框架,替代原有的章节名框架,并在此基础上进行修订,最终完成古籍语言系统框架的搭建,为今后的加工建立更好的平台。

古籍概念的加工不但要建立古籍概念自身与上下文的概念词之间的联系,而且要完成古籍概念与 TCMLS 中现有概念的链接与兼容,如在进行概念添加的时候要先搜索 TCMLS 中是否有该词。如果有,就在原有的概念上进行编辑修改,如此可以避免重复及系统的过于冗杂。

五、小结

以往的中医药一体化语言系统设计当初不是针对古籍的,尽管其中也不可避免地涉及不少古籍文献的内容,但毕竟未能真实、全部地揭示古籍所承载的学术。需要注意的是,古籍自身的学术内容是不会主动呈现的,因而以文献学研究为基础,揭示学术理论内涵,并以信息学理论和技术方法将之充分地展示出来,是中医古籍语言系统构建的初衷。同时也应重视的是,如何将中医古籍语言系统构建工作与现有的中医药一体化语言系统进行更好的融合与匹配,在更大范围内实现中医学体系的信息学式的展示,是今后一段时间内的工作内容,也是本研究的深化与发展。

此外,从研究模式而言,在当前的中医药语言系统研究或更大范围的信息学研究工作中,尽管注意或强调中医学的特点是心知肚明、不言而喻的,但在实际工作中或多或少地还存在个别未能认真贯彻落实的情形。因而,本研究希望通过中医古籍语言系统构建工作,进一步强化关于这一点的认识,切实地将中医特点与信息学理论技术充分结合。只有基于扎实的文献理论研究工作的中医信息学研究,才能真正反映中医的学术特点,也才能为中医的发展拓展出信息学的向度。进一步的来说,将文献、理论研究模式与目前成熟的中医信息学研究模式相结合,应当是今后中医信息学发展的方向。本研究希望能够为这种新的研究模式提供一定的借鉴。

总之,就中医古籍语言系统构建而言,我们认为,文献是基础,理论是深度的体现,信息是关键,三者缺一不可,这也是本研究的交叉学科性质所决

定的,我们在实际工作中也应当关注并着力体现这一点。

参考文献

1 曾召,王小平.UMLS 与中医药一体化语言系统的建立.中华医学图

书情报杂志,2006,15(3):1-3.

2 贾李蓉. 中医药学语言系统语义关系初探. 中医药发展与人类健康,2005:180~183.

3 杨继红.基于本体的中医古籍叙词表构建方法研究.中国中医科学院博士论文,2008.

Literature, theory and information, three important aspects for the establishment of the language system for ancient TCM books

Zhu Ling , Cui Meng

(Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China)

Abstraction: This paper points out that literature, theory and information are three important aspects to build up a language system for the ancient TCM books. Literature is the basis, the theory decides the depth, and information is the key. With the support of information technique, the precise and deep understanding of the TCM recordings and theories plays a critical role in the construction of such a language system.

Keywords: literature; theory; information; language system; TCM ancient books

(责任编辑:李沙沙,责任译审:张立崴)

科学家找到一种艾滋病治疗全新方法

加拿大和美国的研究人员找到了一种治疗艾滋病全新方法。新方法将定向化疗方法与目前较为普遍使用的高效抗逆转录病毒疗法结合使用,为治疗艾滋病提供了一个新的可能。

据参与研究工作的科学家介绍,这种新的治疗方法既可以杀灭在人体内四处游荡的病毒,还能够杀灭隐藏在免疫细胞内的病毒。该项研究成果的价值在于它为全世界研究艾滋病治疗的科学家指明了前进的方向,为找到一种全新的具有创新性的艾滋病治疗方法奠定了基础。相关文章发表在6月21日出版的《自然医学》杂志上。

迄今为止,抗艾滋病治疗的疗效一直不甚理想,原因是病毒隐藏在免疫系统细胞里,而目前使用的高效抗逆转录病毒疗法对躲藏在免疫细胞内的病毒无计可施。加拿大和美国的研究人员成功地识别出那些躲藏有病毒的细胞,找到了病毒可以逃避现有治疗的“隐身”机理。

参与研究工作的麦吉尔大学血液学家基恩·罗迪教授认为,该项研究成果首次证明所谓的艾滋病病毒潜伏池现象并

不是因为抗逆转录药物缺乏效力,而是由于病毒隐藏在两种不同类型的长寿命 CD4 记忆免疫细胞内。艾滋病病毒潜伏池有几种类型,每一种类型潜伏池都需要不同的治疗方法来消除它们。一旦病毒隐藏在这些潜伏池细胞里,它们就会对其产生依赖。如果细胞存活,病毒就可以存活,但如果细胞死亡,则病毒也无法存活。因此,消灭这些免疫细胞就能够杀灭这些隐藏其中的病毒。现存的高效抗逆转录病毒疗法可以杀灭在体内流传的病毒,但是无法对付隐藏在潜伏池细胞内的病毒。

主持研究工作的蒙特利尔大学塞卡里教授表示,该研究结果与用于治疗白血病采取的策略相似,即同样是定向化疗,并同时与定向免疫治疗相结合。这种方法既可以杀灭细胞中的病毒,同时也为免疫系统留出时间再生健康细胞。

研究人员认为,目前所得到的成果为开发出全新的抗艾滋病病毒治疗方法提供了一个重要的选项,但要进入实际应用还需要若干年进一步的研究工作。

(文 摘)