

国外情报检索语言理论与实践对我国的影响

ABSTRACT The study and application of Chinese retrieval language have all along drawn on and introduced foreign theories and practical experiences. CRL, especially DD, SRL and the integration of classification and subject have given certain impetus to the research work concerned.

SUBJECT TERMS Information retrieval language - Introductions
Subject retrieval language - China

CLASS NUMBER G254

1 引言

情报检索语言(Information Retrieval Language,以下简称 IRL)在我国起源可以追溯到两千年前。及至近代,伴随西方科学文化大量涌入的国外 IRL 成果,对我国传统 IRL 产生了深刻影响。从此,我国 IRL 的发展就与国外研究相联系,在借鉴过程中不断前进。

2 对国外 IRL 成果的初步引进阶段

2.1 以 DDC 为代表的国外 IRL 对我国的影响时期(1840~1949)

鸦片战争后,西学东渐,翻译书籍日益增多,局限于传统学术思想的四部分类法无法处理,从国外引进新法成为一种需要。译介引进既有主题检索语言(Subject Retrieval Language 简称 SRL)成果,又有分类检索语言(Classification Retrieval language,简称 CRL)成果。其中 CRL 成果又包括了当时西

方流行的十进法体系、展开法体系和主题法体系。国内图书馆学家对 DDC 积极推广与使用,对其先进之处吸取较充分完备,引发了我国 CRL 的首次变革,使得发展本来缓慢而艰难的中国体系分类法有了飞速的进步和完善。随 DDC 引入的西方理论也为我国 IRL 理论的丰富发展提供了新的源泉。

此时期我国吸取国外成果不足之处有二。一是学习的低层次性,对具体分类法的介绍推广使用远超过对国外理论成果的学习研究,兼之影响巨大的 DDC 本身就重于实用,从而使得国内实践成果斐然,理论上却进展不大。其二是借鉴的盲目性和局限性。由于偏重于 DDC,同期引进的 EC、SC 及标题语言尽管也有着种种优点,甚至有些理论实践价值高于 DDC,却在盲目的热潮中被冷落一旁。可以说过分过深地局限于 DDC 的结果,一定程度上妨碍了我国 IRL 对国外成果的全面学习和吸取。

2.2 前苏联分类理论与实践为主导的国外 IRL 对我国的影响时期(1950~1966)

新中国的建立,文化事业的发展,为本时

期 IRL 发展注入了新的活力,在 CRL、分面理论、SRL 等方面均有国外成果的引进借鉴。社会主义紧邻苏联的理论和实践经验最早也最深地对我国产生了影响,如分类思想、指导原则、类法结构、标记制度等,并一直延续到其后相当长的时间。

此外,阮氏分面理论及 CC、UDC 于 1957 年引入,从理论实践两方面引起了国内重视;尤其结合 UDC 的推广产生了一定的成果。而对 SRL 各种类型的译介打破了国内主题法研究应用长期沉闷的局面,使得各种先进的主题检索方法为国内所知。

对国外成果的消化吸收促进了国内 IRL 实践的发展,理论方面也明显有所进展。刘国钧于 1964 年完成的《现代西方主要图书分类法评述》一书,对国外分类法及分类理论作了系统阐述和比较研究。科技情报系统则将国外成果与国内实践结合,在检索语言的微观特性、使用方法、检索效率方面形成了一定的见解。但由于社会原因,引进借鉴政治性强调过多而科学实用性打了折扣。如对 CC、UDC 的态度不够中肯客观,对单元词法等的研究没能深入下去。这一切又限制了吸取后的自我发展,使得国外先进理论技术未能在国内得到广泛应用和发展。

2.3 对国外 IRL 成果的批判及引进的中断时期(1967~1976)

此时期由于国内动乱,引进交流几成空白。而盲目地批判、错误地对待国外成果则对我国 IRL 的发展产生了消极影响。后期稍有起色,在几大类法、词表的修订编制中注意了对外国成果的学习和应用。

3 综合吸取国外 IRL 成果及我国 IRL 理论与实践繁荣发展的新阶段

80 年代以来,随着改革开放而来的新一轮引进交流促进了我国 IRL 的发展。在全面综合吸取国外最新成果的同时,IRL 学科理

论及学科建设开始起步。

3.1 CRL 继续发展——体系分类法的分面组配化

80 年代,随着文献资料的增长,学科之间的交叉渗透,国外分面分类法的有关论著开始译介出版,DDC、BC 新版的分面改造经验被介绍到国内。体系分类法增加组配成分而分面化很快成为大多数 IRL 研究者的共识,其直接成果便是《中图法》、《资料法》的二、三版修订。

两法的第二版在组配技术采用上迈开了步伐,不仅增设了一批专用复分表及专用组配符号,还大大扩大了冒号组配使用的范围。1985 年铜陵会议、1990 年天津会议又为第三版修订进一步分面组配化奠定了理论基础,作了技术准备。这样,90 年代出版的《中图法》第三版更多运用了复分、仿分和组配编号法,而《资料法》在此基础上突破了在规定类目下进行组配的限定,使用更趋灵活。除尝试在几处运用观点区分组配外,在《使用手册》中还规定了一些概念组配的方式,可以说《资料法》正在演变为一部半分面分类法。

3.2 SRL 异军突起——IRL 跨入叙词语言时代

SRL 在我国尽管很早就有所应用和研究,但其发展却一直徘徊不前,迟迟未能与体系分类语言的发展并驾齐驱。

1975 年 7 月,以国内首部供机检用综合词表——《汉语主题词表》的编制为标志,SRL 的研究、创制和推广在我国真正开展起来。为了编制该表,短时间内编印了 5 部国外叙词表参考资料、有关编制标准及具体词表。但这部大型词表对比当时国外的不少词表,水准却相当不尽人意。未能借鉴国外最新成果,尤其是 70 年代以来的分类主题一体化技术是原因之一。该表完成后,对国外 SRL 主要是其中的叙词语言的成果引进日趋全面系统。译介了国际标准、理论专著、词表编制手册及新进展、新动态,有力促进了我国 SRL

体系的建立及叙词表编制模式的形成;在国内数据库建设中,叙词语言亦成为所使用的主要语言。

3.3 分面分析理论技术的学习应用

分面分析技术的产生最早应归功于奥地利的拉芳田和奥特勒,后来印度阮冈纳赞在CC中系统提出并加以运用。

我国自70年代以来先后发表10余篇介绍阮氏生平、理论的文章及有关专著的中译本,高等院校图书情报专业教材中也都安排了阮氏分面理论或CC的内容。同时引进了阮氏以后国外分面分析的发展成果,尤以英国分类法研究小组(CRG)的成果译介最多。80年代以来,分面理论及技术应用的文章日渐增多,涉及到了分类法、叙词法、标引检索、文摘编制等方面;对引用次序(即标引模式)的研究也明显受到了阮氏PMEST公式及CRG成果的影响。《中图法》、《科图法》也进行了分面组配化的改造。

3.4 分类主题一体化理论与实践成果的借鉴

早在60年代,国内进行主题法、分类法的比较研究时就对分类主题一体化趋势有所察觉。刘国钧得出结论:“合其双美,离则两伤”;杜定友则试图从实践上进行尝试。然而,直至80年代,国内未再有新的突破,而此时国外的一体化技术已发展得相当成熟。1980年一种结构简单、性能优越、形式新颖的一体化词表形式——分面叙词表被介绍到国内,国内进入一体化研究的高潮期。

1983年,完全以英国《分面叙词表》的结构为模型编制而成的首部分面词表问世,几年后《教育分面叙词表》印行。该表成功运用了分面分析技术和计算机技术,是国内在实践层次上的一次飞跃,并为其后几部大型词表的编制提供了范例和经验。

与此同时,国内也开始尝试一体化的另一种形式——对应索引。1986年开始着手编制《中图法》与《汉语主题词表》的对应索引

——《中国分类主题词表》。围绕此表进行的对应试验及编制工作具有较大的实用价值,对我国一体化研究的深入及词表编制必将产生积极影响。

经过对国外一体化成果的吸取和实践,我国已形成分面叙词表,体系分类法、叙词表双向对应索引和专业叙词表分面改造三种形式,并编制了一批一体化词表。

3.5 计算机技术对IRL的影响

国外自50年代以来,计算机技术开始应用于图书情报领域,IRL适应变化随之进行自我改造和创新,如词表类法机编、机读、机管化,文献标引和索引编制自动化,自然语言检索的实现等。国内70年代末80年代初逐步采用计算机技术,国外成果的引入很快对我国IRL研究产生了重大影响。

自动标引在国外起步早,但引入国内首先遇到了汉字切分问题,国内研究很快就集中于此。经过10余年,汉语分词技术取得了很大进展,并在此基础上形成了各种汉语自动标引方案。国外的PRECIS、NEPHIS、POPSI等机编索引系统自80年代也被译介到国内。目前,与汉语PRECIS配套的软件已在微机上研制成功,循环嵌套主题索引及其机编系统也通过了鉴定。国内已有100多个计算机系统有了机读词表。全文检索、关键词法、后控词表法等传入国内后也在汉语应用方面取得了不少成果。

第二阶段的成果引进源自欧美,以英美为甚。英国以CRG的分面分析成果及一体化实践为主导,引入我国后导致了传统体系分类法的分面改造及一体化词表的产生、发展。美国则以其叙词语言的完备及计算机技术的广泛应用对我国IRL发展起了积极促进作用,并为国内学者提供了许多值得研究的新领域。

本阶段国内的理论研究也取得了突破性进展。一方面国外IRL理论著作的翻译出版开拓了理论研究的视野,如 (下转74页)

5 《三国志》

《三国志》，6卷，240段(节)。书题首卷作“新刻按鉴演义京本三国英雄志传”，但其他各卷及板心俱刻“三国志”，故以之为题名。首有玉屏山人《三国志小引》，次为《新刻三国志目录》，又次图像二十四幅。正文半叶15行，行32字。封面黄纸，题“绣像三国志传”(中)、“毛声山先生原本”(右上)、“宝华楼藏板”(左下)。岑桺先生据此推断，此书虽为清初毛评本《三国志演义》通行以后宝华楼重印，但其原刻当在明末坊间。

六卷本《三国志》总目录除卷三脱漏一目外，皆为整齐划一的七言句式，文字多近嘉靖元年本，但与正文段目差异甚大。正文段目除

末卷脱漏一目外，分别为六、七、八言句三种形式，文字与嘉靖元年本、双峰堂本《批评三国志传》皆有异同，相比之下，前者似更俚俗。又，六卷本基本情节大体与他本相差无几，但文字节略之处较多。

近年来，编辑与出版古版文献成为出版界的热门选题。事实上，这是一项繁琐甚至枯燥的文献整理工作，需要编者付出非常细致和艰苦的努力。从《三国志演义古版丛刊》的首辑看，它再现了古版文献重新整理工作的学术价值。编委会的态度是认真的，工作也是富有成效的，愿这些为海内外学人所共知。

厚艳芬 1988年杭州大学语言文学专业毕业。

通讯地址：北京王府井大街36号，邮编：100710。

(来稿时间：1995—07—05。编发者：刘喜申。)

(上接67页)

切尔内《情报检索理论概述》(1980)、兰开斯特《情报检索词汇规范化》(1982)；另一方面，引进结合所取得的实践成果推动了理论研究向纵深发展。1980年张琪玉《情报检索语言》一书完成，提出了以检索效率为中心的观点，并开始了IRL 在我国的正式研究、教育和普及。1987年，该书改名《情报语言学基础》修订出版，进而将IRL 作为一门学科提出，我国情报语言学学科建设随之起步。这不仅将我国IRL 理论研究推向一个新高度，就世界范围而言也是首创。

4 结 语

回顾中国IRL 这些年的发展，可以清晰地看到国外成果对我国IRL 理论与实践的影响。

在这几个时期，国内的引进借鉴从单一化到多元化，从狭窄到广泛，从浅层到深入，从机械照搬到消化吸收直至开拓创新，走过了一段并不平坦的道路。在这个过程中国内

IRL 研究从宏观微观、定量定性、理论实践或有偏颇发展到力求二者的平衡并重。这种日渐理性的趋势体现了我国IRL 借鉴手段及自身发展的逐步成熟。80年代以后IRL 理论尤其是情报语言学学科在国内的创立发展，更体现了在引进后的结合，学习后的出新，也标志着与国外先进水平的迅速缩小。

在成果喜人的同时教训亦不少。总体而言仍是模仿照搬有余，探索创新不足。有以下几点今后须注意：(1)借鉴成果要科学客观，一分为二，既反对一切批判，以偏概全，又反对一切拿来，滥而无用。(2)对国外研究的译介要迅速及时，充分利用。(3)引进吸取要密切联系国内现状及特点，有的放矢，提高有效性。(4)注重理论与实践结合。国外理论要用之于行，在实践中加深认识。对国外技术也要总结吸收。唯其如此，引进学习方可深入。

朱 丽 1994年上海空军政治学院国情专业毕业。通讯地址：上海五角场，邮编：200433。

(来稿时间：1994—12—25。编发者：黄文田。)