

当代国外语言学与应用语言学文库



Dictionary of Lexicography

词典学词典

R. R. K. Hartmann

Gregory James

外语教学与研究出版社

Foreign Language Teaching and Research Press

劳特利奇出版社

Routledge

·语言学与应用语言学文库

Dictionary of Lexicography

词典学词典

R. R. K. Hartmann and Gregory James 著

张柏然 导读

外语教学与研究出版社
劳特利奇出版社

(京)新登字 155 号

京权图字: 01 - 2000 - 2409

图书在版编目(CIP)数据

词典学词典/(英)哈德曼(Hartmann, R. R. K.)等著;张柏然导读. - 北京:
外语教学与研究出版社, 2000.8
ISBN 7-5600-2034-8

I. 词… II. ①哈… ②张… III. 词典学-词典-英文 IV. H06-61

中国版本图书馆 CIP 数据核字(2000)第 69492 号

©R. R. K. Hartmann and Gregory James 1998

All rights reserved.

This edition is published in China by arrangement with Taylor & Francis Limited, 11 New
Feter Lane, London EC4P 4EE, UK. It is for sale in the mainland territory of the People's
Republic of China only, excluding Hong Kong and Taiwan.

本书由泰勒·弗朗西斯出版集团授权外语教学与研究出版社出版

版权所有 翻印必究

词典学词典

R. R. K. Hartmann and Gregory James 著

张柏然 导读

* * *

责任编辑:霍庆文

出版发行:外语教学与研究出版社

社 址:北京市西三环北路 19 号 (100089)

网 址: <http://www.fltrp.com.cn>

印 刷:北京市鑫鑫印刷厂

开 本:650×980 1/16

印 张:13

版 次:2000 年 8 月第 1 版 2000 年 8 月第 1 次印刷

印 数:1—5000 册

书 号:ISBN 7-5600-2034-8/G·1095

定 价:16.90 元

* * *

如有印刷、装订质量问题出版社负责调换

Preface by Halliday

Foreign Language Teaching & Research Press is to be congratulated on its initiative in making these publications in linguistics available to foreign language teachers and postgraduate students of linguistics in China.

The books are a representative selection of up-to-date writings on the most important branches of linguistic studies, by scholars who are recognized as leading authorities in their fields.

The availability of such a broad range of materials in linguistics will greatly help individual teachers and students to build up their own knowledge and understanding of the subject. At the same time, it will also contribute to the development of linguistics as a discipline in Chinese universities and colleges, helping to overcome the divisions into “English linguistics”, “Chinese linguistics” and so on which hinder the progress of linguistics as a unified science.

The series is to be highly commended for what it offers to all those wanting to gain insight into the nature of language, whether from a theoretical point of view or in application to their professional activities as language teachers. It is being launched at a time when there are increasing opportunities in China for pursuing linguistic studies, and I am confident that it will succeed in meeting these new requirements.

M. A. K. Halliday
Emeritus Professor
University of Sydney

王宗炎序

近年来，国际交往日益频繁，国际贸易急速发展，出现了一种前所未有的现象：学外语、教外语、用外语的人多了；研究语言学和应用语言学的人多了；开设这方面专业的高校也多了，语言学硕士生和博士生也多了。就是不以此为专业，学习语言学和应用语言学的也不乏其人。为了给从事这个专业的师生提供便利，同时又帮助一般外语教师、涉外工作者以及汉语研究者开阔思路，扩大视野，提高效率，我们献上这套内容崭新而丰富的丛书——英文版《当代国外语言学与应用语言学文库》。

文库首批推出 54 部外国英文原著，它覆盖了语言学与应用语言学 28 个分支学科。这批书是我们与各地有关专家教授反复研究之后精选出来的。出版这样大规模的语言学与应用语言学丛书，这在我国语言学界和外语教学界是破天荒第一次。

我们这样做，抱着什么希望呢？总的说来，是遵循教育部关于加强一级学科教育的指示，在世纪之交，推出一套书来给中国的外语教育领航，同时也给一般外语工作者和汉语研究者提供信息，拓宽思路。

我们希望这个文库能成为进一步带动外语教学改革和科研的发动机；我们希望它能成为运载当代外国语言学理论、语言研究方法和语言教学方法来到中国的特快列车；我们希望，有了这套书，语言学与应用语言学专业师生就能顺利地进行工作；我们希望，通过读这套书，青年外语教师和外语、汉语研究者能迅速把能力提高，把队伍不断扩大。

以上是我们的愿望，可是从广大读者看来，这个文库是否真的有出台的必要性呢？我们想，只要大家看一下今天的客观情况，就知道这套书有填空补缺的作用，是让大家更上一层楼的扶梯。

我们跟许多人一样，认为国内的外语教学和语言学与应用语言学研究是成绩斐然的，但是某些不足之处也无庸讳言。

在语言研究方面，有大量工作还等着大家去做。汉语语法研究，过去由于结构主义的启示，已经成绩卓著，可是现在虽则引进了功能主义，还看不出什么出色的成果。语料语言学是新兴学科，在我国刚刚起步，机器翻译从 50 年代就有人搞，然而其进展至今不能令人满意。

在语言理论方面，我们不时听到一些片面的、所见不全的论调。有人说，1957 年前西方根本没有什么理论语言学，其创始者是 Chomsky；也有人说，语言纯属社会文化范畴；还有人说，搞语言研究只有量化方法才是科学方法，定性方法不值得一提。

谈到外语教学，某些看法做法是分明不值得赞许的。有人以为交际教学只管听说，不管读写，也有人以为教精读课就是教阅读，不管口语。在分析课文时老师满堂灌，学生开口不得，是常见的；教听力课时老师只管放录音，对学生不给半点提示点拨，也并非罕有现象。

上述这些缺点，我们早有所知，现在我们更加明白，必须力图改进，再也不能安于现状了。为了改进，我们就得参考国外的先进理论，借鉴国外的有效措施。眼前这个文库，就是我们上下求索的结果。

在编辑这个文库时，我们在两方面下了功夫。

一方面，在选书时，我们求全，求新，求有代表性和前瞻性。我们不偏爱一家之言，也不只收一家外国出版社之书。语言学与应用语言学的主干学科固然受到了应有的重视，分支学科可也不忽视。语料语言学、语言统计学是新兴学科，我们收入了专著；句法学、语义学久已有人研究，我们也找到了有关的最新著作。

另一方面，我们邀请了国内知名的博士生导师、硕士生导师为各书撰文导读，为读者铺平道路。语言学和应用语言学专著包罗宏富，初学者读起来可能觉得茫无头绪。为了助他们一臂之力，本文库中每一种书我们都请专家写了一万字左右的导读材料。哪怕书中

内容比较陌生，谁只要在读书前看一下导读材料，读书后把材料再看一遍，一定能弄清脉络，掌握要点。

在结束本文时，我们想向爱好泛读的人们提个建议。语言和社会生活息息相关；我们靠语言与他人协作；通过语言继承传统文化，接受外国先进思想和科学知识；利用语言来教育下一代，帮助他们创造美好的未来；语言又反过来表达着我们的个性和我们充当的各种角色。学一点语言学和应用语言学，有助于增强我们的语言意识，对我们的工作和生活都是有利的。我们不妨把此事作为一个项目，列入自己的日程。持之以恒，必有所获。

王宗炎

中山大学教授
博士生导师

导 读

在人类认识世界、改造世界的诸多活动中，词典是不可或缺的工具之一。据史料考证，早在 3000 多年前人类便开始编纂词典。词典的发展是一条奔腾不息的长河，但真正意义上的词典理论研究则始于 20 世纪 50 年代。自此以降，有关中外词典理论研究的论著陆续问世，其数量虽不能说是汗牛充栋，却也林林总总，不胜枚举。但是，以全球的眼光和从历史的角度来进行综述的著作极其少见。由英国埃克塞特大学词典研究中心主任哈特曼（R. R. K. Hartmann）和香港科技大学语言中心主任格雷戈里·詹姆斯（Gregory James）两人编著的《词典学词典》（*Dictionary of Lexicography*, 1998）一书，填补了这一方面的空缺。通读全书，从这部书中得到的最重要的启示便是：眼光、积累、情怀。编著者的全球眼光和历史眼光，他们深厚的学养和长年积累，使这部著作充满了冷静的观察、哲理的思考和人文精神。

出版前言开宗明义地指出，本词典是以阐述词典编纂的理论与实践以及两者之间的关系为己任的。词典收录 2000 余个条目，全部按字母顺序编排。条目涉及的主要范畴有：词典学理论与实践、词典史、词典类型研究、词典质量标准与批评、词目的释义与编辑、词典的结构与使用、语言资料的收集、词典用户展望、语料库技术、词典编纂计算机化，以及国际词书组织和刊物名录及其互联网址，等等。所收条目典型、全面、新颖，概括地反映了词典学发展的脉络与最新动态；词目释义取短语模式，用词浅显易懂，言简意赅；词目内容编排清晰有序且举例适当，词条末所附参见条别具一格，简明直观、信息量大；每个条目自成一体，但整本词典的条目内容又构成对词典编纂的理论与实践的历史回顾、现状扫描、未来展望和整体宏观描述。

全书近 30 万言，时间跨度上下千百年，以其丰富的史实和翔实的资料，在回顾历史、扫描现状、展望未来的过程中，阐述了许多颇有新意的独到见解。现择其要者，分述如下：

一、从历史的角度及现实的发展阐明了词典学不是语言学的附庸，而是一门独立的综合性学科（*sui generis*），廓清了对词典学的学科性质的种种模糊认识。该词典首先区分“词典编纂”与“词典学”这两个不同的概念，认为前者是技艺，后者是科学；并提出对词典编纂活动可以进行科学的描写这一

命题,其真正含义是指有这么一门学问可对这项活动实施描写,而活动本身无自我科学描写可言,它只能是科学描写的对象。这种科学描写,就是词典学的任务。20世纪50年代以来,科学技术尤其是电子信息技术迅猛发展,大大拓展了词典学的学科疆域,极大地丰富了词典学的研究内容。该词典收录了一系列与 lexicography (词典学) 有关的条目,如: academic lexicography (学术性词典学)、automatic lexicography (自动化词典学)、bilingual lexicography (双语词典学)、biographical lexicography (传记词典学)、comparative lexicography (比较词典学)、computer lexicography (计算机词典学)、corpus-oriented lexicography (语料库技术词典学)、cultural lexicography (文化词典学)、EFL lexicography (英语作为外语的学习词典学)、electronic lexicography (电子词典学)、ELT lexicography (英语教学词典学)、historical lexicography (历史性词典学)、ESP lexicography (专业英语词典学)、machine-aided lexicography (机助词典学)、monolingual lexicography (单语词典学)、multilingual lexicography (多语词典学)、normative lexicography (规范性词典学)、pedagogical lexicography (教学词典学)、phraseological lexicography (短语词典学)、prescriptive lexicography (规定性词典学)、special lexicography (专科词典学)、text-specific lexicography (专门文本词典学)、theatre lexicography (戏剧词典学)、thesaurus lexicography (类义词典学)、usage lexicography (用法词典学) 等等,足以说明这一点。词典学虽说主要同语言打交道,但它还需要综合利用信息科学、计算机科学技术、哲学、美学、文学、历史学、人类文化学、社会学、教育学、心理学、历史语言学、比较(或对比)语言学、应用语言学以及翻译与教学等方面的知识。正由于如此,在当今社会文化语境下,词典学正日益成为更具全球性的学术领域。

二、从历史的角度和现实的发展提出了词典分类的新思维和新方法。自从1940年前苏联词典学家谢尔巴(Л. Щерба)基于词典多相特征对词典类型进行理论探讨以来,世界上不少词典学家陆续写出了这方面的论文和著作,提出了自己的分类法。该词典的编著者在研究人类信息交流发展史特别是信息载体的演变史的基础之上,结合词典编纂与出版的现实状况,从中抽象出如下四条基本分类准则: 1. 现象分类法(phenomenological typology): 指按词典的篇幅(如袖珍词典、节编本词典、简明词典等)或内容涵盖范围(如普通词典、专科词典等)等形式特征进行分类; 2. 表述(或构造)分类法(presentational or tectonic typology): 即按词典的版式(如按字母顺序编排的词典、分类词典、意念词典)或载体(如手稿本、印刷本、电子本词典等)

来进行划分；3. 功能分类法 (functional typology): 一种根据词典的信息类别 (如正音词典、拼写词典、词源词典等) 或编纂方法 (如详解词典、教学词典、术语词典等) 的划分法；4. 语言分类法 (linguistic typology): 主要依据词典涉及的语言多寡来进行分类 (如单语词典、双解词典、双语词典、多语词典等)。这样的分类法简要清晰, 以全面的分类体系出现, 概括程度较高; 且为容纳尚未问世的词典留下了余地; 还有助于从分类的角度完整地描写一部具体的词典, 比传统的或多或少建立在对立面的基础上的分类法要大大前进了一步, 更符合词典的生成规律, 为词典分类学研究提供了一个新视角。

三、该词典的编著者从接受理论中汲取营养, 辩证地论述词典文本 (通过编者) 与词典使用者之间的互动关系, 强调词典使用者的参与作用。对于词典编纂者来说, 词典是一个阐释文本, 旨在完成原语言功能; 对于词典使用者来说, 他 (或她) 的参与需要词典的引发。为了构建意义和达到词典的解惑释疑效果就需要词典的编著者和使用者双方的共同努力。换言之, 词典的编纂者只是提供意义 (或所指) 产生的可能, 而要使这种可能变为现实, 获得接近所指的意蕴则是词典使用者的任务。这种对词典使用者的关注和对词典使用者的作用的抬升, 正是近年来词典学研究中的一个重要课题。这一课题的深入研究对提高词典使用者的素质, 对确立词典为使用者服务的观念, 对改进并提高词典词目释义的质量, 以及对提高语言教学的水平, 无疑将起到重大的推动作用。

四、在悠悠 3000 余年的历史长河中, 自从华夏祖先发明纸和印刷术以来, 词典出版的媒质和载体大致经历了四个阶段: 第一阶段是以“手”与“纸”为工具的手工书写阶段; 第二阶段是“火”与“铅”为特色的印刷阶段; 第三阶段是以“光”与“电”为标志的计算机时代; 第四阶段是以“网”与“天”为特征的网络时代。科学技术的迅猛发展, 尤其是电子信息技术、信息数字化技术、语料库技术和网络技术等的日新月异, 使得人们还来不及与“火与铅”、“光与电”告别, 便一步跨进了网络时代。于是, 各种电子词典, 光盘只读词典, 甚至网上在线词典等应运而生。媒质和载体的转变, 不仅对词典编纂和出版的各个环节产生革命性的影响, 而且对改变人类在查询系统方面的观念和处理方法起着巨大的作用。该词典的编著者敏锐地看到这一点, 并作了较为详细的研讨。有关这方面的理论与实践还有待于人们作进一步的研究。

Dictionary of Lexicography

Anyone who has ever handled a dictionary will have wondered how it was put together, where the information has come from, and how and why it can benefit so many of its users. The *Dictionary of Lexicography* addresses all these issues.

The *Dictionary of Lexicography* examines both the theoretical and practical aspects of its subject, and how they are related. In the realm of dictionary research the authors highlight the history, criticism, typology, structures and use of dictionaries. They consider the subjects of data-collection and corpus technology, definition-writing and editing, presentation and publishing in relation to dictionary-making. English lexicography is the main focus of the work, but the wide range of lexicographical compilations in other cultures also features.

The *Dictionary* gives a comprehensive overview of the current state of lexicography and all its possibilities in an interdisciplinary context. The representative literature has been included and an alphabetically arranged appendix lists all bibliographical references given in the more than 2,000 entries, which also provide examples of relevant dictionaries and other reference works.

The authors have specialised in various aspects of the field and have contributed significantly to its astonishing development in recent years. **Dr R. R. K. Hartmann** is Director of the Dictionary Research Centre at the University of Exeter, and has founded the European Association for Lexicography and pioneered postgraduate training in the field. **Dr Gregory James** is Director of the Language Centre at the Hong Kong University of Science and Technology, where he has done research into what separates and unites European and Asian lexicography.

Acknowledgements

We owe a deep debt of gratitude to the very many people who have helped us along our way in the compilation of this dictionary.

The University of Exeter and the Hong Kong University of Science and Technology provided congenial environments within which we were able to work. We benefited greatly from the study leave granted to R. R. K. H., which allowed him an extended period in Hong Kong, sponsored by the British Council and by the Language Centre at HKUST, where we were able to work together face to face.

We are extremely fortunate to have had excellent office back-up both at Exeter and HKUST. We should like to thank Cathy Pantry for secretarial assistance at the Dictionary Research Centre at Exeter; and at the Language Centre, HKUST, Ada Fan Shui-fun for secretarial assistance, Candice Poon Wai-yi for executive assistance, and Forward Hui Kam-chun and Terence Chow Chi-wai for technical assistance.

Our professional colleagues have helped us in a variety of ways, and we should like to record our appreciation especially to: (at Exeter) Steven McGill for the database design and transfer, Alan Wakelam for the initial design of the panels, Trevor Learmouth for assistance with 'library' matters and entries, John Wilkins for references for 'Classics'; (at HKUST) John Milton for assistance with 'computer' matters and entries, and the staff of the University Library for general helpfulness.

We should like to thank the postgraduate students at Exeter, especially the 1994-5 cohort, who helped us in our initial item search; in particular Martin Stark made a great contribution to our wordlist by his very thorough sifting of available sources. Li Lan and her fellow doctoral students helped us greatly in our search for references.

We are grateful to colleagues far and wide for their willingness to answer our queries: especially A. D. de V. Cluver on terminology, Edward Gates on the DSN, Andrew Taylor on textbooks; others commented very substantially on our text, in particular Adelia Carstens, Amy Chi Man-lai, Turki Diab, Tom McArthur and Daan Prinsloo. We ought to mention, too, the anonymous reviewers of our first draft: their constructive observations on our entries were extremely helpful to us.

We have drawn upon the ideas of many colleagues, both from their published work and from informal personal contacts; we should like to thank, in this regard, Henning Bergenholtz, Rufus Gouws, Colin Yallop and Jacques van Keymeulen, as well as several Visiting Scholars at Exeter.

We appreciate very much the assistance of Rosalind Fergusson, our perspicacious copy-editor, who made significant contributions to the final form of the text. Our editors at Routledge deserve our thanks too, for their care and attention to our work.

We should like to thank the many publishers who responded to our requests for catalogues and other information; and the colleagues and students who answered our initial questionnaire, which persuaded us to see our proposed project to fruition.

We hope that all those who made contributions, large and small, to the final product will be forgiving of our sins of commission and omission, for which we alone remain responsible.

Finally, we should like to thank our families for putting up with our prolonged hibernation in our offices, and general absent-mindedness at home while entries were germinating in our heads.

R. R. K. Hartmann
Exeter

Gregory James
Kowloon

Introduction

Theory and practice

A dictionary of lexicography? A dictionary of dictionary-making? Indeed, 'lexicography' is commonly understood in a narrow sense as the theory and practice, or the 'art and craft' (Landau 1984), of dictionary-making, the compilation and preparation of reference texts for publication. However, in the latter half of this century, the importance of a second strand to the discipline, namely the scholarly field of dictionary research, has become increasingly recognised. With the very rapid advances in the technologies and formats of reference materials, particularly as a result of the development of electronic media, the horizons of lexicography have extended, and we are witnessing the emergence of a more global academic domain of 'reference science' in the modern context of the widespread availability of, and ever-increasing needs for rapid access to, all forms of information.

Lexicography, often misconceived as a branch of linguistics, is *sui generis*, a field whose endeavours are informed by the theories and practices of information science, literature, publishing, philosophy, and historical, comparative and applied linguistics. Sister disciplines, such as terminology, lexicology, encyclopedia work, bibliography, terminography, indexing, information technology, librarianship, media studies, translation and teaching, as well as the neighbouring disciplines of history, education and anthropology, provide the wider setting within which lexicographers have defined and developed their field.

The boundaries between the professional activity and the academic field of lexicography are fluid, and several associations, e.g. the DSN (the Dictionary Society of North America, est. 1975), the Lexicographical Society of India (est. 1975), EURALEX (the European Association for Lexicography, est. 1983), AUSTRALEX (the Austral-

asian Association for Lexicography, est. 1990), the Lexicographical Society of China (est. 1992), AFRILEX (the African Association for Lexicography, est. 1995), and ASIALEX (the Asian Association for Lexicography, est. 1997), have been formed to act as bridges between dictionary makers and academic lexicographers. Professional training, regular national and international conferences, seminars and workshops, and academic publications have served to mature lexicography into an independent field with its own principles and practices, purposefully making use of, and qualifying, the findings of other disciplines.

Dictionary typology

The core material of lexicography is the 'dictionary', or 'wordbook', the commonest variety of reference work, at once the subject of lexicographical theory (dictionary research) and the product of lexicographical practice (dictionary-making). The dictionary *qua* dictionary encompasses many genres, or types, which can be structurally classified. For example, a *phenomenological typology*, or categorisation based on formal features, might take account of such compositional characteristics as size ('pocket dictionary', 'abridged dictionary', 'concise dictionary' etc.) or coverage of the content of the work ('general dictionary', 'specialised dictionary' etc.). A *presentational*, or *tectonic*, typology would focus on the format ('alphabetical', 'classified', 'thematic' etc.) or medium ('manuscript', 'print', 'electronic' etc.) of the dictionary. A *functional typology*, or categorisation based on the contextual uses of the dictionary, would focus on the information categories provided ('pronunciation', 'spelling', 'etymological' etc.), and the ways these are presented ('explanatory', 'pedagogical', 'terminological' etc.) within the perspective of the target user ('scholarly', 'learner's', 'translator's')

etc.). A *linguistic typology* would be based on the language(s) of the dictionary ('monolingual', 'bilingual', 'bilingualised' etc.).

These different types of dictionary have developed over a considerable period, in response to linguistic and cultural demands, and as a result of changes in the use and availability of communicative media. McArthur (1986a:4ff.) discusses four major 'shifts' in the process of the development of human communication: (i) the consolidation of speech and gesture; (ii) the development of writing; (iii) the advent of print technology; and (iv) electronic computation. In several lexicographical traditions, e.g. Arabic, Persian and Sanskrit, many early manifestations of lexical reference works were in metrical format. For example, the thematically-arranged synonymies of the Sanskrit tradition (called *nighaṇṭus*) emerged at the time of advanced development of pre-literate speech and the beginning of the development of writing. These reference tools were at first unwritten, and were passed from generation to generation in oral mode. To facilitate memorisation, they incorporated mnemonic devices such as rhythm, rhyme and perhaps music, and some were extremely sophisticated in their structure. The later emergence in the same tradition of thesauruses with letter-order arrangement demonstrates the second 'communicative shift', from total reliance on speech to a formalised and exclusive codification in the written medium, and involving a tactical reformulation of the dictionary consultation process. This shift of the basis of the process of dictionary compilation from spoken medium to script has implications for, and implies alterations in, at least four 'factors of dictionary use': (i) *information*: considerations of the written medium – ordering of items, spelling, legibility, development and preservation of writing materials – now supersede those of sound; (ii) *operations*: the nature of the consultation of the dictionary, i.e. reading not memorising, will be extended as the information is offered in differently accessible formats and literacy will assume enhanced social and referential functions; (iii) *users*: the users will be those who are functionally literate, and the oral traditions of memorisation and recitation will become inadequate; and (iv) *purposes*: the contexts of dictionary use will be extended as their content and format change, for example they may serve as repositories of knowledge, and develop a role as arbiters in matters of orthography and other areas of usage.

From print to computerisation

The third of McArthur's 'communicative shifts', that from writing to print, permitted the mass production of dictionaries, and a change in users' attitudes towards the social and educational purpose of the dictionary, which because of its stability in print format assumed the role of an authority and a judge in language. All of the major European languages can claim their 'own' dictionaries, many of them monumental pioneering scholastic enterprises, either as individual initiatives or under the auspices or sponsorship of an academy, and all of which served, in various ways, to codify their respective languages: such examples are the *Vocabolario degli Accademici della Crusca* (Venice, 1612) for Italian, the *Dictionnaire de l'Académie française* (Paris, 1694) for French, the Spanish Academy's *Diccionario de la lengua Castellana* (Madrid, 1726–39) for Spanish, Jacob and Wilhelm Grimm's *Deutsches Wörterbuch* (Leipzig, 1854–1961) for German, the *Tolkovi slovar' zhivago velikorusskogo yazyka* (St Petersburg, 1863–66) by Vladimir Ivanovich Dal' for Russian, James Murray *et al.*'s *New English Dictionary on Historical Principles*, which became the *Oxford English Dictionary* (Oxford, 1884–1933) for English, and the *Novo Dicionário da Língua Portuguesa* (Lisbon, 1899) by Cândido de Figueiredo for Portuguese.

The widespread use of these and other European languages, and the historical and political circumstances of their dissemination, have meant that the lexicographical traditions of languages in other parts of the world have often been somewhat neglected outside their own spheres. In China, for example, lexicography has a distinguished pedigree. The first thesaurus made its appearance during the early Qin dynasty (221–207 BC), and the 49,000-entry character dictionary, *Kāngxī zidiān*, commissioned by the Qing Emperor Kangxi in the early eighteenth century, eclipses anything produced in Europe up to that period.

The fourth 'communicative shift' proposed by McArthur, that of electronic computation, has revolutionised lexicography not only with respect to the stages of the dictionary-making process (e.g. the automatic retrieval of lexical evidence from large corpora of texts, and on-screen text processing), but also in the conception and treatment of reference systems (e.g. in the use of multimedia, and the varied possibilities of search and access methods). From the hand-held

electronic spelling dictionary to the multimedia encyclopedia on CD-ROM, lexicography is an area of academic and commercial enterprise that has turned computerisation to account in a variety of novel ways. In particular the development of the relational database as a design feature has served to corrode traditional organisational access structures, such as alphabetisation. In print format, the structure and composition of the material collected by the lexicographer determined the access path taken by the user: alphabetically organised headwords would be looked up alphabetically, the ordering of senses determined a sequence of search strategies to locate the sense sought, etc.

Computerisation has radically transformed these relationships. The output can now be specified by the user in a variety of ways, which are not necessarily determined by the internal organisation of the dictionary or the structure of the input. For example, material may be accessed and retrieved in alphabetical order, in reverse alphabetical order, by chronology, by definition keywords, by grammatical function, or by a wide variety of different configurations of information categories, irrespective of the way they have been organised as input to the database. Thus it is the user who can stipulate the limits of the information sought, and retrieve that information according to a designated reference scheme, avoiding redundancies. The emphasis is less on following a predetermined pathway through the dictionary structure, more on navigating relationships across and within entries, via a choice of, often invisible, programmed links.

The user perspective

One of the principal advances in lexicography in recent years has been the focus on the user perspective, that is the realisation that different users have different reasons for using a dictionary, and that the dictionary can, and should, respond to these. Whilst dictionaries have usually tried to satisfy the overall perceived needs of large classes of users, there has often been little attention to the specific needs of smaller groups or individuals, or to the fact that users differ in their needs depending on the immediate context of dictionary use. Users may not, for example, have adequate reference skills to follow the structure of a printed dictionary entry, and may need explicit assistance to locate and extract the information they seek. One very common reason for recourse to a dictionary

is for translation from one language to another, and this need has spawned an abundance of bilingual and multilingual dictionaries of all sorts and sizes, over many hundreds of years. A speaker of English and a speaker of French, for example, translating a non-specialist English text into French, might use a general English–French dictionary. It is only recently that there has emerged a general recognition of the fact that the needs of these users will be different, and that a simple ‘English–French’ compilation may not respond to both equally well. User-oriented research takes account of the profiles of dictionary users; the various contexts of dictionary use; the functions of the dictionary in its many situations of use; and the skills necessary for, and brought to bear on, dictionary use. There now exists a body of investigatory techniques, which are constantly being improved upon, and a considerable cumulation of findings of enquiries carried out in many different language communities around the world, which provide valuable information for the incorporation of a user perspective into dictionary conception and design.

The response to the potential needs of users has implications for the teaching of dictionary use, an area of education that was somehow taken for granted, and either neglected or, at best, skimpily treated, in the past. Specifically, more attention is now being paid to the abilities needed, and the behaviours associated with those abilities (or lack of those abilities), on the part of the dictionary user to locate the information being sought. The complex operations involved in the dictionary consultation process are being investigated and analysed, and there is a conscious attempt in many dictionaries to relate the conventions of the dictionary text (e.g. abbreviations, codes, labels) to these operations. One example is the increasing use of full-sentence definitions, pioneered by the COBUILD dictionaries, in learners’ dictionaries. Nowadays we are also witnessing the growth and extension of explicit instruction in dictionary use, in the form of the inclusion of ‘dictionary skills’ in school syllabuses and ‘the teaching of dictionary skills’ in teacher-training syllabuses, all directly informed by the findings of research.

The computer provides the latest user-determined type of retrieval system, but even very early dictionaries in different cultures were compiled with specific purposes in mind. For example, one of the oldest dictionary traditions in China – where dictionaries have been known for over 3,000 years

– was that of examination preparation. The development of dictionaries of Chinese was closely tied to the civil service examination system. Indeed, most of the dictionaries before the establishment of the Republic in 1911 were compiled by members of the élite scholar–official class associated in one way or another with these examinations, which tested literary erudition and written composition through mastery of intricate rhyme patterns and a knowledge of Chinese characters. In India, on the other hand, dictionaries developed within an oral tradition, the earliest being cumulative synonymies composed in poetical format. The metrical structure of these dictionaries facilitated memorisation and oral recitation. The dictionaries were composed as aids to oral composition of poetry, and for the teaching and understanding of religious texts. The very different types of dictionary which developed within these two traditions – China and India – were a function of the different cultural milieux, and the different purposes for which they were required. But in both traditions, there was an authoritarian perspective. The dictionary was a resource of what was considered to be the ‘best’ or ‘correct’ language. This perspective has carried on to our own day, and there is a very deep-seated public attitude, shared among many language communities across the globe, that ‘the dictionary’ is supposed to represent some form of final authority in matters of lexical meaning and use. The academy dictionaries typically exert considerable influence – sometimes upheld by legal sanction – in protecting a language from what are perceived as unacceptable or corrupting pressures, for example, excessive borrowing from other languages. The acceptance of the authority of the dictionary has also often been brought to bear in litigation, where meanings and judicial interpretations of words have been determined or elucidated by reference to ‘standard’ dictionaries.

The dictionary and linguistic evidence

In 1961, the compilers of the third edition of *Webster's New International Dictionary of the English Language* adopted a new criterion by explicitly breaking with the authoritarian tradition. In its indications of pronunciations, in its definitions, and in its selection of words, *Webster's Third* abandoned any claim to be a lexical canon, and discarded all puristic tendencies (cf. Bolinger 1975:585). But in its shattering the image of the

dictionary as a linguistic arbiter, this lexicographical *cause célèbre* engendered a controversy which lasted for several years. The public, even some literary editors, did not want a dictionary that recorded even careless (albeit generalised) pronunciations, or meanings felt to be ‘incorrect’. It was not, in the view of some, the dictionary’s place to describe, but to prescribe. However, the replacement in *Webster's Third* of a normative, attitudinal approach to dictionary-making by one based on objectively observed facts about language was in tune with the currents of linguistic thought of the time, and indeed within a tradition that had begun with Samuel Johnson’s *Dictionary of the English Language* (1755). Johnson’s initial aim had been to compile a definitive dictionary for English along the prescriptive model of the Académie française’s national French dictionary. However, realising that change was inherent in language, and that ‘fixing’ English in this way was an unrealistic proposition, he sought to show sufficient grounds for his definitions by supplementing them with illustrative citations from ‘the best writers’. Thus, he hoped to temper the process of what was at the time perceived as degeneration of the language, rather than attempt a stabilisation of an unrealisable ‘ideal’. Johnson’s appeal to usage rather than etymology in defining words was as innovative as it was controversial. His methodology of using citations to corroborate definitions, however, has provided the model for the dictionaries of our day. Nevertheless, Johnson was not entirely objective, especially in his selection of words to be defined, and he omitted those he considered would offend the sensibilities of the polite society of his day. In a similar way, conscious choices are sometimes made in our own time. In the context of the new South Africa, for example, a decision was made by the compilers of the *Woordeboek van die Afrikaanse Taal* to omit racial slurs, both as headwords and in illustrative citations.

Whilst the dictionary, therefore, can properly reflect change, it thereby also legitimises it. The ready availability of huge banks of electronic data have provided hitherto unobtainable evidence of language use, indicating subtle discriminations only suspected before. For example, many dictionaries of English have defined (and still define) *big* in terms of *large*, and *vice versa*. The evidence from concordances of occurrences of these adjectives in extensive modern text archives, however, demonstrates convincingly that although they do

overlap in some contexts, they are differentiated systematically in many others, and that an important concomitant of these differentiations lies in their different collocational patterns, an area of usage that is now attracting greater attention of lexicographers than has been the case in the past.

The compilation of large-scale (not *big-scale*!) corpora can enable the recording and documentation of language change in progress. For example, the language of computing has given us back-formations such as *input* and *output* as verbs. In the Hong Kong University of Science and Technology's (HKUST) 1,000,000-word corpus of English in Computer Science (James *et al.* 1994), verbal usage constitutes about five per cent of the total occurrences of these forms. However, the evidence is that there is variability in the formation of the past tenses (*input/output* or *inputted/outputted*?). Nevertheless, the derivational influence remains strong, and the weight is still towards *input/output* (the only forms accepted by, for example, the revised third edition of the *Collins Concise Dictionary*, which advertises itself as 'the authority on current English'). But the use of *input/output* as the past form can cause confusion, as the participial-adjectival usage of *input/output* can thus have two senses, proactive (e.g. *input data* = 'data to be input') and retroactive ('data which has been input'). This example demonstrates the uncertainty that can occur when a novel creation conflicts with the language system. New verbal formations in English characteristically attract the morphological increments of weak verbs (*-ed*); in the case of strong verb compounds, largely confined to 'special language' contexts, the resultant instability of competing forms can create linguistic insecurity. This can be mitigated by corpus data, which can thus be brought to bear on the process as well as the product of language change, and documented in dictionaries in a way that has hitherto been impossible.

Interestingly, whereas in the name of descriptive lexicography, Webster's *Third* reported variant pronunciations for words – even some considered 'unacceptable' by certain speakers – a similar liberality was not exercised with respect to spelling, and the *Third* remains, as does every other dictionary of English, as authoritarian as ever in this regard. This is a reflection of the mood of the English language communities: speakers of English tolerate wide variations in pronunciation, but variations in spelling are codified (*centre* and *center*, for example, are accepted in different contexts, largely determined by regional prove-

nance, but *accomodation* for *accommodation*, or *principle* in the sense of *principal*, are considered errors despite their widespread 'careless' use). A diversification in pronunciation is an indicator of coefficient of individuals' identities; standardisation in spelling is what unites the language across the different communities of its speakers.

Dictionary compilation

How, then, is the dictionary compiled? Dictionary-making *per se* comprises those operations which contribute to the production of dictionaries. In addition to preliminary activities such as surveying the potential market, planning the dictionary, training and recruiting the staff etc., three principal stages in the dictionary-making process may be isolated: data gathering, editing and publishing.

Computerisation has motivated the creation of dictionary databases by publishers, to assist in dictionary compilation. These databases are of two kinds: the *language resource database* (or *citation database*) and the *entry database*. The former is used as the authority for the dictionary entries, and ideally consists of corpora, text archives, individual citations etc. from a very wide range of published and unpublished (e.g. spoken) sources. Creating such a database was once a very onerous task – the compilation of the citation files of the *Oxford English Dictionary*, through a reading programme which involved many people in all walks of life, took some 22 years – but nowadays computerisation and the rapid access to electronic texts have meant that very large resources, extensive in scope, can be established relatively quickly. The *British National Corpus*, for example, consists of some 100,000,000 words, ten per cent of which are transcribed from natural spoken language, and the *Bank of English*, which is used for the CO-BUILD dictionary series, comprises in excess of 320,000,000 words, and is continually being extended. Linguistic archives can also incorporate the findings of fieldwork, that is, the collection and documentation of information obtained by interviews, and, indeed, this form of data-gathering is essential in the compilation of dialect, slang and other special-purpose dictionaries.

The entry database includes the headwords, definitions and other information categories (e.g. pronunciation, syntactic functions, usage labels). This database also includes formatting codes, page layout instructions, cross-references and links across entries, which enable rapid updating of information, and easy extraction of specific entries