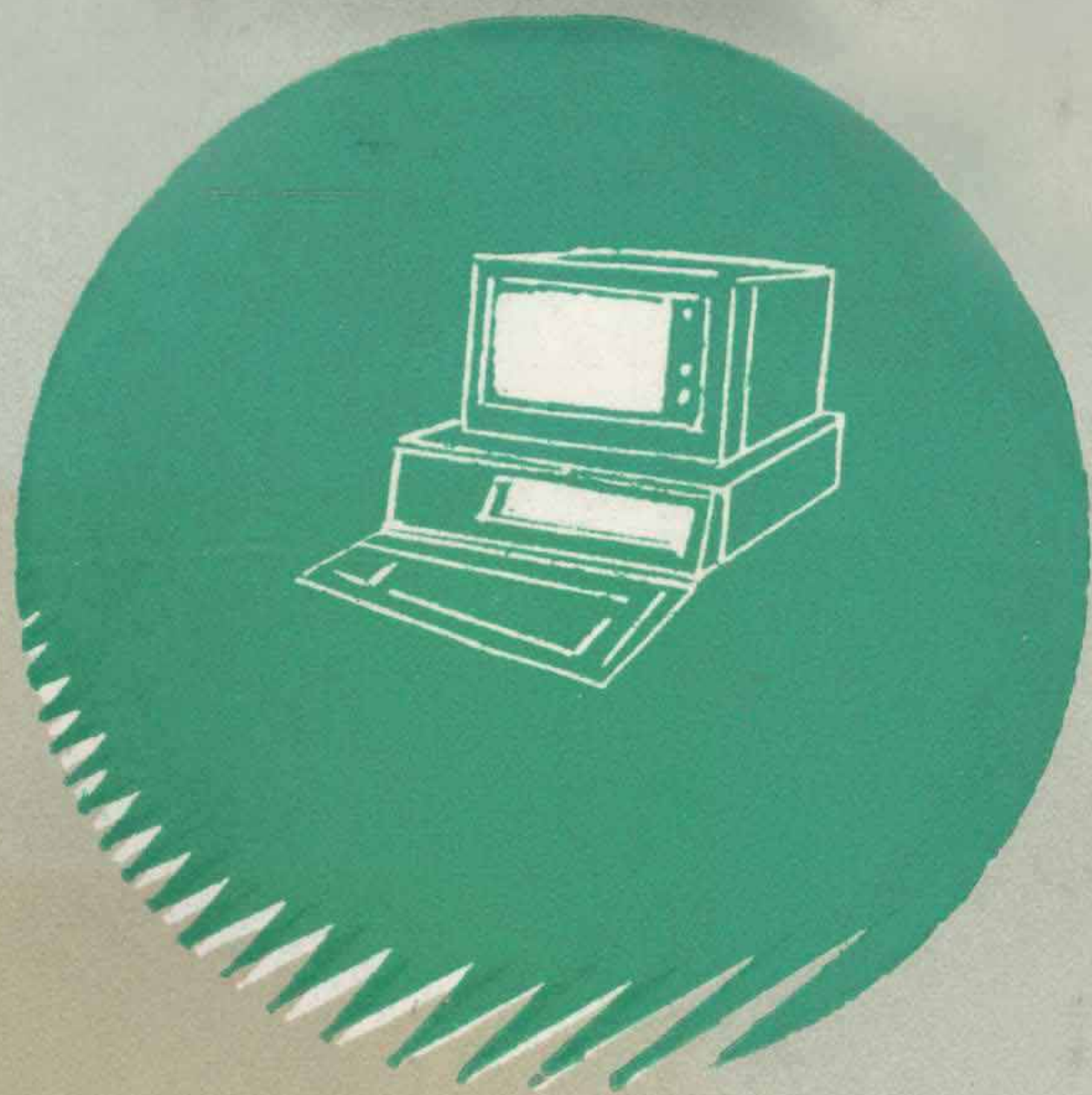


自然科学基金资助项目



中国各民族文字与 电脑信息处理

戴庆厦 许寿椿 高喜奎 主编

中央民族学院语言研究所 中央民族学院计算机系 中央民族学院计算机系

中央民族学院出版社

自然科学基金资助项目

中国各民族文字 与 电脑信息处理

主 编

戴 庆 厦

许 寿 椿

中央民族学院语言研究所

中央民族学院计算机系

高 喜 奎

中央民族学院计算机系

中央民族学院出版社

一九九一年·北京

责任编辑：吉甫
封面设计：张万隆

内容简介

本书是语言文字学家与电脑信息处理专家合作的产物。编写者面对关心电脑信息技术的文字学家及从事文字信息处理的电脑专家以及未来可能从事这一事业的广大青年学生(相关文科及理科的本科生、研究生)。

本书分总论、各论和附录。总论综述了中国各民族文字、东亚汉字式文字及世界文字概况，并简介中文信息处理的进展及有关大事记。各论分别介绍了全部现用文字和部分古文字及相应电脑信息的处理情况。附录包括各民族情况、有关文字工作文件及文献及技术标准简介。

中国各民族文字与电脑信息处理

ZHONG GUO GE MIN ZU WEN ZI YU DIAN NAO XIN XI CHU LI

戴庆厦 许寿椿 高喜奎 编著

*

中央民族学院出版社出版

(北京西郊白桥路 27 号)

(邮政编码：100081)

新华书店北京发行所发行

张家口地区印刷厂印刷

787×1092 毫米 16 开 25.5 印张 587 千字

1991 年 9 月第 1 版 1991 年 9 月第 1 次印刷

印数：1—5000 册

ISBN7-81001-218-5/H·14

登记证号：184 定价：9.25 元

编 者 的 话

本书是中央民族学院计算机系与民族语言研究所(民语三系)在共同承担国家自然科学基金项目*的初期合作编写的。本书的内容并不是项目的基本内容,也没有列入原项目申请计划。由于两专业(语言文字学与计算机信息处理技术)的专家开始承担这个交叉学科的项目时,都甚感十分需要一本基础资料,这资料应该面向两类专业人员,向他们介绍这个交叉学科领域的一些基本概况,以建立最基本的共同语言。这样,共同编写本书就成为合作研究的一个起点。本书的书稿大部分都是在1988年上半年就编写好了的,由于出版方面的原而拖延下来。直到1991年初,课题组独立开发的与汉英兼容的国际音标编辑软件增配到轻印刷系统上之后,课题组才有了无需多少资金便可自行排版的方便,本书的出版才落实下来。

本书面向两个专业(语言文字学与计算机信息技术)的读者,内容都是很基本的资料,编写时都回避了较为专深的内容,使初在这个交叉领域工作的人不会有专业上的困难。这与合作研究初起阶段的状况相适应。鉴于这种合作反映了一种方向,类似的合作许多单位都在进行,我们想这本书或许对这类读者群能有所帮助。

本书正文部分(指总论和各论两部分),除周有光先生的两篇是转载以外,其余的都是首次发表。附录A和附录B是文件或文献的摘录,附录C是作者应本书编者邀而撰写的。付印前,凡时间性较强的都作了些补充修改,如附录A收录了1990年人口统计中关于各民族人口的数据,总论中大事记和附录B、C都增加了1989和1990年两年的新材料。

本书由课题组用轻印刷系统排版。按项目要求设计的与汉英兼容的国际音标编辑软件由计算机系教师蒙家国承担,本书的录入和具体排版由计算机应用专业88级学生陈友在课余进行。本书出版还得到中央民院科研处的帮助与支持。在这里,编者仅代表本书作者向所有于本书出版给予关心、支持的同志深表谢意。

* 项目名称:“多文种微机句词语料库及言语比较”,项目申请人:许寿椿、戴庆厦。

迎接计算机文字信息处理 多文种化的潮流

(代序言)

在拉丁文字圈诸发达国家,计算机文字信息处理已应用于各行各业,传统上由手工或机械方式进行的文字处理已普遍地实现了电脑化,电脑信息时代已经到来。第一台数字电子计算机诞生后马上就开始了对于文字字符采用二进制编码表示的计算机的研制。1951年第一批这种机器用于美国人口统计获得巨大成功。短短的时间内,计算机文字信息处理技术和应用服务迅速扩展到各种领域。随着70年代微型机的诞生,计算机的日益微型化、日益增强的多功能综合处理能力和日益降低的价格又给文字信息处理技术和应用以强大推动。企事业办公室中公文的制作、存贮、管理、联机通信,书、报、刊的排版印刷,图书情报的检索,飞机、轮船票位的预定发售,银行款项的存兑等等,无一不实现了以计算机为核心的自动化、智能化。计算机已经从少数科学家、工程师的实验仪器设备变成普通行政人员办公和学习的用具,变成一般人家中的家用电器和学生学习用品。每一个人,不限于其本职业务工作,包括日常生活,无一不时时刻刻享用着文字信息处理自动化、智能化的成果。诸如银行信用卡的使用,超级购物中心货物管理、货款计算,以及桌上型、膝上型乃至掌上型电脑的种种应用。拉丁文字圈率先实现的这种技术变革,不像它率先实现机械化后那样给其它文字圈造成致命的打击,而是为各类文字的自动化、智能化处理准备了技术条件。

于80年代形成高潮的计算机软件及系统的汉化是文字信息处理技术发展中举世瞩目的事件。汉化,是指对原有英文软件或系统进行改造和扩充,使其具有处理汉字能力的活动或技术。这种汉化工作的基地主要在东亚汉字文化圈,特别是在中国和日本,但不限于此,北美、澳大利亚和西欧众多计算机公司厂商对此也有浓厚兴趣,投入了相当的财力和人力。单纯以或主要以增加软件复杂性的办法使复杂的英文计算机系统具有汉字处理能力是一件有划时代意义的事件。在工业化时代,英文率先实现处理机械化之后,以增加设备的机械复杂性为手段根本无法克服汉字的许多困难。这实际上是汉字不断被批判、被鞭挞、被咒骂的一种社会技术因素。80年代,单纯以或主要以一种无型的软件技术实现的汉化获得了丰硕成果。汉字信息的自动化、智能化处理在许多重要领域获得突破性进展。在并不增加体积和重量,不改变视觉性状的条件下,一个英文系统令常人看来颇为“莫明其妙地”具有了汉字处理功能。这种种成果,极大地推动了计算机在中国的普及应用,属性迥异的汉、英两类文字融于一体兼容处理的成功,也为世界上多文种计算机技术产生了积极而广泛影响,同时,也促使世界范围内重新认识和评价汉字。

汉化不是一个孤立事件，是国际范围内计算机民族文字支撑及国际化潮流的一部分。由于当今计算机应用中数值计算已占极小比重，非数值应用中，文字信息处理是专业范围最广泛、用户人群最众多、数据量极为宠大的一个领域。计算机在非英语国家的应用，都面对扩充当地民族文字处理能力的强烈要求。在英文信息处理获得明显成功之后，在社会各领域获得巨大效益和革命性变革之后，各民族文字支撑就开始形成国际性潮流。任何一个非英语国家，要在企事业办公自动化、书报刊排版、情报检索等领域应用现代信息技术，都自然要求计算机扩充本民族文字支撑能力。任何一个面向世界市场的计算机公司、厂商，如果不能解决这种用户要求，它就必然要失去某些市场，必然削弱自己的竞争力。面向世界诸多文种的民族文字支撑就形成了计算机软件及系统的国际化。民族文字化和国际化正是同一个大潮的两个侧面。从个别国家、民族、地区的视角观察是民族文字化，从综观国际信息界的视角观察则是国际化或多文种化。汉化正是这种历史必然发展的一环，是大潮中的一股流。

中国少数民族文字信息处理于80年代也获得喜人进展，为世界多文种计算机信息处理提供了有益的素材和经验。中国是一个多民族的大家庭，各民族使用着众多的语言和文字。所用文字涉及世界五大文字圈中的所有类型。中国在各民族地区实行民族平等的双语或多语政策，这为中国少数民族文字信息处理的现代化提供了基本的政治条件。民族地区计算机应用同样要求有民族文字处理能力，特别是与汉字、英文兼容的处理能力。办公文书的编写、书刊的排版、为影视屏幕迭加民族文字字幕、民族古籍及语言文字资料的研究整理是一些需求迫切的领域。80年代初期以来，首先在内蒙和新疆出现了蒙文化和维文化的软件。这项事业得到国家有关机构的支持，也吸引了一些关心民族地区四化建设的各民族同志参与。短短的几年内，与汉英兼容的蒙文、满文、维文、哈文、柯文、藏文、朝文、彝文、傣文等计算机操作系统相继诞生。轻印刷、精密激光照排、字幕机及打字机都推出了商品化产品。少数民族文字信息处理有关标准也相继制定。

计算机多文种化潮流空前广阔地拓广了文字学研究的视野。类型、属性迥然不同的多种文字兼容并蓄于一体的要求，推动了人们对现存各类文字的统一观察、比较、研究。这种活动产生了两类结果，一类是具有多文种处理能力计算机系统和软件，一类是对文字类型与属性的新的认识和理解。

计算机语言文字信息处理也推动语言文字学研究的不断深入。计算机信息处理，已经不限于字符的简单编辑加工，还包括词语处理、句子和篇章分析；已经不限于词法，还包括词义和语义；不仅限于字形，还涉及字音、字义；不仅限于字符的录入存贮，还包括识别、理解、转换或翻译。这种不同方向不同层次的发展提出了一系列理论问题和技术问题，这种进展也严肃而又较少主观性的审视已有的理论认识和模式，开始了种种新的探索。一批研究项目具有鲜明的文理结合的性质。具有定量化和工程性的特点，新的突破和变革在酝酿中。

时代呼唤传统语言文字学家与计算机信息处理专家联合起来。计算机文字信息处理技术的发展推动了一大批信息处理专家学习、了解、研究语言文字学。这些人使用的工具和方法，这些人的着眼点和着重点都与传统语文学家有所不同。这种潮流也推动广大语文学家利用计算机为工具从事自己的研究，也影响了一大批语文学家从新的视角去重新观察、思考自己的研究对象。国家正开展着一批不同层次级别的有两个专业专家合作参与的项目。这类项目进展的成功与否，课题成果的深度和质量都强烈依赖于两个领域专家的合理结构、良好的组织和融洽的结合。本书就是中央民族学院这个小角落中两类专业人员首次合作初起阶段的一个尝试产品。它是有些新意的，也带着初生的雅气；它是综合的，有一定的广度，又远非令人满意的丰满完全。它期待两个领域专家的关心指导。赘语数言，仅代为序。

许寿椿

1991. 5.

目 录

编者的话

代序

一 . 总 论

1 中国各民族文字综述	戴庆厦
4 文字的国际分布和历史演变	周有光
8 汉字文化圈的文字演变	周有光
(一) 18 种文字和 4 个演变阶段	8
(二) 东亚的国际文字	9
(三) 南方民族的汉字式文字	10
1. 壮字	11
2. 喃字	11
3. 瑶字	12
4. 布依字	12
5. 侗字	13
6. 白字	13
7. 哈尼字	14
8. 水字	15
(四) 北方民族的汉字式文字	16
1. 契丹字	16
2. 女真字	17
3. 西夏字	18
(五) 汉字式字母文字	19
1. 江永妇女字	19
2. 彝字	19
3. 傈僳字	20
4. 日本假名	21
5. 契丹小字	22
6. 朝鲜语文	22
7. 注音字母	23
(六) 汉字式文字的系统观	24

28 中文信息处理进展

许寿椿

(一) 三个时代·····	28
(二) 古代, 中文信息处理的辉煌成就·····	28
(三) 近代, 汉字失去了机械化·····	29
(四) 现代, 汉字步入新时代的阵痛·····	30
(五) 八十年代, 步入电脑时代的十年·····	31
1. 汉英兼容技术, 多文种兼容技术·····	32
2. 汉字键盘输入·····	33
3. 汉字识别·····	35
4. 汉语语音识别·····	35
5. 汉语语音合成·····	36
6. 汉字、电子排版技术·····	37
7. 机器翻译与自然语言理解·····	38
8. 少数民族文字信息处理和多文种系统·····	39
9. 企事业管理及办公自动化·····	40
10. 语言统计、语料库等文字工程·····	41

42 中国文字工作大事记

许寿椿

二. 各 论

52 汉字	许寿椿
70 蒙文	宝玉柱(蒙古族)
75 蒙文信息处理概况	高喜奎(朝鲜族)
77 藏文	仁增旺姆(藏族)
80 藏文信息处理概况	许寿椿
83 维吾尔文	张铁山
86 哈萨克文	张铁山
88 柯尔克孜文	张铁山
92 维吾尔文、哈萨克文、柯尔克孜文信息处理概况	高喜奎(朝鲜族)
95 苗文	李炳泽(苗族)
100 彝文	朱文旭(彝族)
104 彝文信息处理概况	高喜奎(朝鲜族)
106 壮族文字	覃晓航(壮族)
121 布依文	王 伟(布依族)
127 朝文	高喜奎(朝鲜族)
130 朝文信息处理概况	高喜奎(朝鲜族)
135 满文	白立元(满族)

140	侗文	郑国乔
146	瑶文	邓方贵(瑶族)
153	白文	李绍尼(白族)
157	哈尼文	王尔松(哈尼族)
161	傣文	张公瑾、孟尊贤
179	黎文	文明英(黎族)
185	傈僳文	陈嘉瑛
191	佤文	赵福荣(佤族)
196	拉祜文	杨 春(拉祜族)
200	水文	刘日荣
207	纳西文	陈嘉瑛
213	景颇文	刘菊黄
216	载瓦文	傅爱兰
220	锡伯文	白立元(满族)
224	俄文	张铁山
226	古突厥文	张铁山
229	回鹘文	张铁山
231	察合台文	张铁山
235	八思巴文	白玉柱(蒙古族)
238	回鹘蒙文	白玉柱(蒙古族)
241	西夏文	马忠健
244	契丹文	白玉柱(蒙古族)
245	女真文	白玉柱(蒙古族)

三. 附 录

附录 A	关于各民族情况的一些材料	许寿椿 编
A.1	中国各民族名称的罗马字母拼写法和代码.....	246
A.2	中国少数民族情况简表.....	249
A.3	中国少数民族人口比例综合表.....	255
A.4	中国各民族语言系属简表.....	256
A.5	中国语种代码.....	257
A.6	少数民族文字书、报、刊出版统计四则.....	260
A.6.1	1949—1984 年少数民族文字书、报、刊统计.....	260
A.6.2	1986 年少数民族文字报纸种数及发行份数统计.....	260
A.6.3	中国当代期刊总览 1988 年统计	261

A.6.4	1990 年全国报刊征订简目统计	261
-------	------------------------	-----

附录 B 有关文字工作文件及文献简介 许寿椿 编

B.0	关于语言文字工作方针	262
B.0.1	周恩来：当前文字改革的任务（节录）	262
B.0.2	新时期语言文字工作方针及任务	262
	（1986 年全国语言文字工作会议摘要节录）	
B.1	汉语拼音方案	263
B.2	汉语手指字母方案	266
B.3	手旗通信和汉语拼音方案（鲁屏）	271
B.4	中国人名汉语拼音字母拼写法	276
B.5	中国地名汉语拼音字母拼写法	277
B.6	联合国秘书处关于采用“汉语拼音”的通知	278
B.7	国际标准化组织采用汉语拼音方案作为拼写汉语的国际标准	279
B.8	汉语拼音正词法基本规则	280
B.9	关于《普通话异读词审音表》的通知	289
B.10	关于简化字总表	291
B.10.1	关于重新发表《简化字总表》的说明	291
B.10.2	关于简化字的联合通知	291
B.10.3	《简化字总表》说明	292
B.11	关于统一汉字铅字字形的联合通知及《印刷通用汉字字形表》的说明	294
B.12	关于出版物上数字用法的试行规定	296
B.13	关于部分计量单位名称统一用字的通知	298
B.14	现代汉语常用字表说明	300
B.15	现代汉语通用字表说明	303
B.16	标点符号用法（1990 年）	306
B.17	关于少数民族文字方案中设计字母的几项原则	312
B.18	少数民族语地名汉语拼音字母音译转写法	313
B.19	蒙古语地名译音表及说明	317
B.20	维吾尔语地名译音表及说明	320
B.21	藏语（拉萨话）地名译音表及说明	324
B.22	新疆维吾尔自治区民族语言文字使用管理暂行规定	328
B.23	延边朝鲜族自治州朝鲜语文工作条例	332
B.24	西藏自治区学习使用和发展藏语文的若干规定	336
B.25	中华人民共和国国家标准信息处理用现代汉语分词规范	338

附录 C	文字信息处理有关标准简介	黄伟敏 编写
C.0	引言	350
C.1	我国有关主要国家标准简介	350
C.1.1	汉字编码标准	350
	1. 基本集	
	2. 第一辅助集	
	3. 第二辅助集	
	4. 第三辅助集	
	5. 第四辅助集	
	6. 第六辅助集	
C.1.2	其它民族文字编码标准	353
	1. 蒙文	
	2. 朝鲜文	
	3. 维吾尔文	
	4. 彝文	
C.1.3	汉字字形标准	354
	1. 15×16 点阵 (宋体) 字形	
	2. 24×24 点阵 (宋体) 字形	
	3. 32×32 点阵 (宋体) 字形	
C.1.4	其它民族文字字形标准	356
	1. 蒙古文 16 点阵字形	
	2. 蒙古文 24 点阵字形	
	3. 维吾尔文 点阵字形	
	4. 彝文 点阵字形	
C.1.5	键盘盘面排列标准	357
	1. 七位键盘盘面排列	
	2. 汉字整字键盘盘面排列	
	3. 蒙古文键盘盘面排列	
	4. 维吾尔文键盘盘面排列	
C.1.6	其它有关文字信息处理的基础标准	359
	1. 西文交换码	
	2. 代码扩充技术	
	3. 八位代码结构	
	4. 控制功能	
	5. 字符集的登记	
	6. 汉语信息处理词汇	

C.2	相关国际标准简介.....	362
C.2.1	ISO—646.....	362
C.2.2	ISO—2022	362
C.2.3	ISO—4893	363
C.2.4	ISO—6429	363
C.2.5	ISO—2395	364
C.2.6	ISO/DIS 10646	364
C.3	其它有关汉字标准简介	365
C.3.1	日本的汉字字符集标准.....	365
C.3.2	南朝鲜字符集标准.....	365
C.3.3	中国台湾省汉字交换码、内码、字形标准概况.....	365
	1. 交换码	
	2. 内码	
	3. 字形	

附录 D 部分少数民族文字信息处理交换用编码图形字符集 许寿椿 编

D.1	信息处理交换用蒙古文七位编码图形字符集.....	368
D.2	信息处理交换用蒙古文八位编码图形字符集.....	369
D.3	信息处理交换用维吾尔文七位编码图形字符集基本集.....	370
D.4	信息处理交换用维吾尔文七位编码图形字符集辅助集.....	371
D.5	信息处理交换用维吾尔文八位编码图形字符集.....	372
D.6	信息处理交换用朝文编码图形字符分区表.....	373
D.7	信息处理交换用彝文编码图形字符分区表.....	383

中国各民族文字综述

中国是一个多民族国家，不仅民族成分多，而且语言文字的种类丰富复杂。在长期的历史发展过程中，各民族语言是各民族交流思想、传播知识的重要工具；各民族文字记载了各民族在生产生活、社会生活中积累的知识和经验，保存了人们创造的历史文化和科学技术。在今天的四化建设中，各民族语言文字已更显示出其重要作用，与发展文化教育、科学技术密切相关。搞好四化建设，必须重视研究各民族语言文字的特点及其发展规律。

一. 中国各民族语言

文字是在语言的基础上产生的，是书写语言的符号系统。所以，认识文字必须先认识语言。

中国共有 56 个民族。55 个少数民族除回、满等民族已全部转用汉语外，都有自己通用的语言。少数民族语言是少数民族重要特征之一。由于有的民族使用一种以上的语言，所以中国的语言总数至少在 60 种以上。这些语言分别属于五个语系——汉藏语系、阿尔泰语系、南亚语系、南岛语系和印欧语系。

汉藏语系又可分为汉语和藏缅、苗瑶、壮侗三个语族。属于藏缅语族的有藏、嘉戎、门巴、珞巴、羌、普米、独龙、景颇、傈僳、哈尼、拉祜、纳西、基诺、怒、土家、载瓦、阿昌等语言。属于苗瑶语族的有苗、布努、勉、畲等语言。属于壮侗语族的有壮、布依、傣、侗、水、仫佬、毛南、拉珈、黎、仡佬等语言。

阿尔泰语系又可分为蒙古、突厥、满一通古斯三个语族。属于蒙古语族的有蒙古、达斡尔、东乡、东部裕固、土和保安等语言。属于突厥语族的语言有维吾尔、哈萨克、柯尔克孜、乌孜别克、塔塔尔、撒拉、西部裕固、图瓦等语言。属于满一通古斯语族的有满、锡伯、赫哲、鄂温克、鄂伦春等语言。

属于南亚语系的有佤、德昂、布朗、克木等语言，都属于这个语系的孟—高棉语族。

属南岛语系的是高山族诸语言，有阿眉斯、排湾、布嫩、泰雅尔等语言，都属这个语系的印度尼西亚语族。

属于印欧语系只有属斯拉夫语族的俄语和属印度—伊朗语族的塔吉克语。

此外，朝鲜语和京语的系属未定。朝鲜语的系属，多数认为属阿尔泰语系；京语的系属有的认为属汉藏语系，有的认为属南亚语系。

经过五十年代以来的语言调查，语言工作者又发现了一些新语言，其结论有待进一步确定。这些语言是：属于藏缅语族的木雅、多续、尔龚、尔苏、纳木义等语言，属于壮侗语族的临高、佯僇、莫等语言。

已经消亡的古代语言，有属于藏缅语族的西夏语，属于印欧语系的粟特语、于阗语、焉耆—龟兹语等。

我国各民族语言使用情况复杂，发展不平衡。总的看来，有以下几个特点：（1）在使用人数上，有的相差很大。汉语使用人口至少在十亿以上。少数民族中使用人数最多的是壮语，约有 1500 万；最少的是赫哲语，仅 4200 余人。（2）阿尔泰语系和印欧语系的语言

主要分布于我国的北部；其它几个语系的语言主要分布于我国的南部。(3) 在语言与民族的界限上，出现了许多交叉现象。有的民族使用一种以上的语言，如景颇族使用景颇语、载瓦语等不同语言，瑶族使用勉、布努、拉珈等三种语言，裕固族使用东部裕固语和西部裕固语。也有不同民族使用相同的语言，如回族、满族都使用汉语。(4) 语言内部差异存在不同的特点：有的语言，方言差别不大，如维吾尔语、哈萨克语等；有的语言，方言差别很大，互相不能通话，如彝语、苗语等。有的方言差异比其中一个方言与另一语言的差异大，如壮语南北方言的差别就比北部方言与布依语的差别大。(5) 有的语言有文字，有的语言没有文字。

二、中国现行文字

目前，有 25 个民族有代表自己语言的文字。由于有的民族使用一种以上的文字，如傣语使用四种文字，景颇族使用两种文字，所以 25 个民族共使用 33 种文字。列表如下：

民族名称	使用文字的名称
汉族	汉文
蒙古族	蒙古文
藏族	藏文
维吾尔族	维吾尔文
苗族	黔东南苗文，湘西苗文，川黔滇苗文，滇东北苗文（均为新创）
彝族	彝文（规范音节文字）
壮族	壮文（新创）
布依族	布依文（新创）
朝鲜族	朝鲜文
侗族	侗文（新创）
瑶族	瑶文（新创）
白族	白文（新创）
哈尼族	哈尼文（新创）
哈萨克族	哈萨克文
傣族	傣仂文，傣哪文，傣绷文，金平傣文
黎族	黎文（新创）
傈僳族	傈僳文，老傈僳文
佤族	佤文（新创）
拉祜族	拉祜文
纳西族	纳西文（新创）
景颇族	景颇文，载瓦文（新创）
柯尔克孜族	柯尔克孜文
土族	土文（新创）

锡伯族

锡伯文

俄罗斯族

俄罗斯文

此外，还有一些文字未能通行，只被少数人使用。这些文字有：水书、方块壮字、方块白字、方块瑶字、傣傣竹书、方块拼音苗文等。

三、中国古文字

历史上，我国还有一些文字，曾使用过一个时期，后来停止使用，成为死文字。这些文字有：突厥文、回鹘文、察合台文、于阗文、焉耆——龟兹文、粟特文、八思巴字、契丹大字、契丹小字、西夏文、女真文、满文、东巴图画文字，东巴象形文字、哥巴文、老彝文、尔苏沙巴图画文字。

四、中国文字的结构类型

中国文字结构特点丰富。文字发展各阶段的特点，在中国文字里都能见到。从结构特点上划分，中国文字大致可分为以下几个类型。

(一)、图画文字：东巴图画文字、沙巴图画文字。

(二)、象形文字：汉字，东巴象形文字，契丹大字，西夏文，女真文，水书，方块壮字，方块白字，方块瑶字。

(三)、音节文字：彝文，哥巴文，傣傣竹书。

(四)、拼音文字：按字母形式分拉丁字母和非拉丁字母两类。

1. 拉丁字母：黔东南苗文，湘西苗文，川黔滇苗文，滇东北苗文，壮文，布依文，侗文，瑶文，白文，哈尼文，黎文，傣傣文，佤文，拉祜文，纳西文，景颇文，载瓦文，土文。老傣傣文是采用大写拉丁字母及其颠倒形式。

2. 非拉丁字母：种类较多，主要有以下几类。

(1) 印度字母：藏文，八思巴字，于阗文，焉耆——龟兹文，傣仂文，傣哪文，傣绷文，金平傣文。

(2) 叙利亚字母：粟特文，回鹘文，蒙古文，满文，锡伯文。

(3) 阿拉伯字母：察合台文，维吾尔文，哈萨克文，柯尔克孜文。

(4) 方块型拼音文字：受汉字字形影响，一个音节的字母拼成一个方块。属这一类的有：朝鲜文，契丹小字，方块苗文。

(5) 斯拉夫字母：俄罗斯文。

(6) 如尼字母：突厥文。

本文作者：中央民族学院 戴庆厦

文字的国际分布和历史演变

语言使人类别于禽兽。文字使文明别于野蛮。教育使进步别于落后。

有了文字记录才开始历史时期，在这以前是“史前”时期。在信息化社会，作为信息的主要承载体的语言和文字将发挥更神奇的作用。

可是，鱼在水中不知水。人人会说话，很少人了解语言的性质和结构。天天看书写字，不一定知道文字发展的事实和规律。

汉族原来只用汉字，现在增加了一套字母，叫做“汉语拼音字母”。字母已经是我们日常生活中大有用处的工具。这象征着东西方文化的汇流。

字母是人类的最平凡的又是最伟大的发明。人类文字史中的“字母的故事”成为现代青年感兴趣的知识。

让我们看一看今天世界各地的文字分布。

我们看到，有的地区用汉字，有的地区用字母。用汉字作为正式文字的国家有中国、日本（汉字假名混合体）、朝鲜（北方全用谚文，南方用汉字谚文混合体）和新加坡（同时以英文、马来文和泰米尔文为官方文字）。

我们又看到，字母有多国通用的，有一国独用的。三种多国通用的字母占去了地球的一大半。

分布最广泛的是拉丁（罗马）字母。欧洲的大部分、美洲和澳洲的全部、非洲的大部分和亚洲的小部分，都以拉丁字母作正式文字。欧洲有一条字母分界线，沿着苏联和保加利亚的西面边境，穿过南斯拉夫的中心；分界线以西是历史上的罗马天主教国家，用拉丁字母；分界线以东是东正教国家，用斯拉夫字母。非洲也有一条字母分界线，在北非一排阿拉伯国家的南面边境；分界线以北用阿拉伯字母，以南的大半个非洲用拉丁字母，主要是新独立的国家大都沿用原宗主国的文字（英文、法文、葡萄牙文等）。也有当地语言写成拉丁字母的，例如在南非有阿非利根斯文，在东非有斯瓦希里文等。亚洲用拉丁字母作为正式文字的国家有土耳其、越南、印尼、马来西亚、菲律宾和新加坡（以英文为主要文字）。

阿拉伯字母的分布仅次于拉丁字母。它是北非和西亚十几个阿拉伯语国家的正式文字；此外一些伊斯兰教国家和地区也用阿拉伯语文字作为正式文字，如伊朗、阿富汗、巴基斯坦、中国新疆等。

第三种多国通用文字是斯拉夫字母。它是苏联俄罗斯民族和少数民族的文字（波罗的海东岸三个加盟共和国爱沙尼亚、拉脱维亚、立陶宛除外）。保加利亚、半个南斯拉夫（另一半用拉丁文字母）、以及蒙古人民共和国也以斯拉夫字母作为正式文字。

印度字母系统包含多种字母，同出一源而形态各异，不能彼此通用。这些字母用于印度各邦（有主要的14种文字）、印度四邻（斯里兰卡、孟加拉、尼泊尔、不丹、中国西藏）、印度支那半岛（缅甸、泰国、老挝、柬埔寨）。此外，一国或一地区独用字母主要有：希腊字母（希腊）、希伯来字母（以色列）、假名字母（日本）、谚文字母（朝鲜）、阿姆哈拉字母