

与时俱进、开拓创新，做好信息时代的语言文字工作

——从学习、贯彻党的十六大文件精神谈起

王铁琨

(国家教育部 语言文字信息管理司，北京 100860)

摘要：当前的语言文字工作深受信息化、标准化、法制化和四部大开发四个方面的影响。语言文字工作要做到与时俱进。开拓创新，须做好以下几项工作：加强语言文字规范标准建设；抓好对语言文字信息化处理具有推进意义的基础性工作；抓好语言文字社会应用管理。

关键词：语言文字工作；信息时代；与时俱进

中图分类号：H102

文献标识码：B

一

学习、贯彻党的十六大精神，必须紧紧把握住江泽民同志报告的精髓，就是坚持“解放思想、实事求是、与时俱进”的思想路线。“解放思想、实事求是”是我们党一贯倡导的思想路线，而加上“与时俱进”四个字，则深刻反映了时代变革的主题。

“与时俱进”是近两年来报刊上出现频率很高。人们谈论比较多的一个词语，指一种先进的思想理论必须做到与时代共同进步，才能保持其旺盛的生命力。“与时俱进”命题的提出，是根据国际形势的新变化，根据我国改革开放和现代化建设面临的新问题和任务，根据我们党所肩负的历史使命以及党自身建设的需要提出来的新论断，反映了党顺应时代变革积极进取的精神风貌和社会发展的要求。“与时俱进”是对党的思想理论发展的科学总结，是继续创新的指南，具有划时代的意义。大力弘扬与时俱进精神，始终保持与时俱进的精神状态；才能在思想上不断有新解放，理论上不断有新发展，实践上不断有新创造，也才能真正做到“发展要有新思路，改革要有新突破，开放要有新局面，各项工作要有新举措”。坚持贯彻“解放思想、实事求是、与时俱进”的思想路线，对于我们在新世纪不断拓展、深化语言文字工作具有特别重要的指导意义。

实现“与时俱进”，要有改革的意识和创新的思维。通读十六大报告，印象最深刻、最受启发的就是全文突出了一个“新”字。近两天一边认真学习十六大报告，一边也作了一点统计分析。报告中总共出现了多少个“新”字呢？根据我的统计是110多个，“排重”后也有近70处提到了“新”。请看：新路、新局面、新世纪、新形势、新阶段、新进展、新步伐、新成效、新任务、新秩序、新思路、新要求、新境界、新经验、新概括、新视野、新力量、新活力、新解放、新发展、新创造、新突破、新举措、新高潮、新情况、新问题、新本领、新知识、新特点、新中国、新鲜经验、新兴市场、新型社区、新的阶段、新的表现、新的实践、新的里程碑、新民主主义、新的强大生机、新的世界大战、新的就业岗位、新的理论篇章、新的伟大工程、新的发展阶段、新的明显成效、新的历史性突破、新的经济增长点、新的历史条件下（以上为“新”字打头的词语）；日新月异、崭新局面、高新技术、高新技术产业、浦东新区、创新、创新人才、创新体系、创新体制、鼓励创新、锐意创新、开拓创新、自主创新

能力、理论创新、制度创新、科技创新、文化创新、教育创新、技术创新、管理创新、农村经营体制创新、其它各方面的创新，等等。细读这一大串带有“新”字的词语；确实有耳目一新之感。这些个“新”，充分体现了改革、创新的精神。

创新是一个民族的灵魂，是一个国家、一项工作兴旺发达的不竭动力，也是一个政党、一项事业永谋生机的源泉。创新就是不断地解放思想、实事求是、与时俱进。世界风云变化，中国的改革开放和现代化建设正在不断前进，实践在发展，实践没有止境，创新也没有止境。这就要求我们“自觉地把思想认识从那些不适宜的观点、做法和体制的束缚中解放出来。从对马克思主义的错误的和教条式的理解中解放出来，从主观主义和形而上学的桎梏中解放出来。”我们必须以马克思主义的理论勇气，以实事求是的思想作风，在理论上不断拓宽视野，通过理论创新不断推动制度创新、科技创新、文化创新、教育创新。包括语言文字工作创新。时代在前进，实践在发展，语言文字工作也应该努力跟上时代前进的脚步，根据实践的要求拓展工作领域和范围，与时俱进，勇于和善于不断创新。

二

语言文字工作要与时俱进、开拓创新，首先需要对新时代的形势特点有一个比较准确的把握。根据我们的观察思考，当前对语言文字工作已经和正在发生重大影响的因素，主要有信息化、标准化、法制化和西部大开发这四个方面，机遇和挑战并存，这是一把双刃剑。

（一）信息化

人类文明的历史已经走过了农业时代、工业时代，现在正处在广泛采用数字技术处理信息而引起高新科技迅猛发展的信息时代。高新科技中发展前途远大的是信息技术、生物技术、纳米技术、材料技术和能源技术，而所有高新科技及其相关产业的发展，都无例外地依赖于信息科技的发展，特别是计算机产业及其应用技术的普及。今天计算机和每个人的关系实在太密切了，已经成为工作、生活中须臾离不开的工具。现实昭示我们，21世纪高新科技的发展必将以信息技术的发展为先导、为龙头。正因为此；党的十六大报告根据世界经济科技发展新趋势和我国经济发展新阶段的要求，提出本世纪头20年经济建设和改革的主要任务，即“完善社会主义市场经济体制，推动经济结构战略性调整，基本实现工业化，大力推进信息化，加速建设现代化；保持国民经济持续快速健康发展，不断提高人民生活水平。”并指出“信息化是我国加快实现工业化和现代化的必然选择”，要“优先发展信息产业，在经济和社会领域广泛应用信息技术”。

在这样的形势、任务面前，信息化、信息技术同语言文字工作有没有关系，有什么关系呢？我们说，不但有关系，而且关系还很密切。迎接数字化信息时代的到来，我国要以信息化带动工业化，实现跨越式发展，就必须加速发展信息产业，同时以工业化促进信息化。而发展信息产业；必须研究解决制约信息技术发展的“瓶颈”问题，这些“瓶颈”有很多是由于语言文字问题没有解决或解决得不够好引起的。

信息时代一个重要任务就是要对各类信息进行各种各样的加工处理，根据专家研究估计，语言文字所负载的信息占整个信息量的80%以上。显然，语言文字是信息处理的主要对象，语言文字信息处理是声音信息处理、图像信息处理所不能替代的。过去语言研究主要是解决人与人交际的问题，而到了信息时代；语言学者不得不更多面对人与计算机交际的种种问题，当然“人机交际”本质上还是人与人的交际；只不

过表现形式变了。要让计算机高质量地理解语言文字所负载的信息内容，必须首先让计算机能够像人一样“听懂”人发出的语音、“看懂”人写的文字，而且能够直接“发出”语音、“写出”文字，能够拥有一定的词汇量并学会从语流和字符串当中识别词语。中文和西文不同，汉语句子是由一个一个等距排列的汉字组成的，而不像西文是由一个一个自然隔开的单词组成的句子。所以，中文信息处理必须让计算机“学会”掌握一定的汉语语法、语用规则和语义常识乃至百科知识；才能真正理解汉语。

经过20多年的艰苦努力，我们花费很大代价基本解决了现代汉语“字处理”问题，正在着手解决“词处理”问题，并且向着“句处理”“篇章语义处理”乃至大规模真实文本的自动处理方向进发。应该说，真正的语言信息处理才刚起步，实现真正意义的“人机对话”任重而道远。

自然语言的自动处理（人机对话），并不是轻而易举就能够实现的。由于汉语本身一些独有的特点，进入计算机后就形成了不同于西文的一些与生俱来的难点：汉语没有词的界限标志，需要解决自动分词问题；汉语没有形态变化标志，计算机难于区分词与词之间的句法、语义关系；汉语词类划分和兼类情况复杂，词性自动判别和标注十分困难（语料库语料经机器切分并标注词性后，还要做大量人工校对、复核工作，才能成为“熟”语料）；汉语的句法和语义分析更是困难重重。这些“瓶颈”，严重制约着汉语自然语言自动处理技术的发展。

首先是自动分词。汉语进入计算机后仍然是一个一个等距排列的汉字，而自然语言理解首先要个词，也可以切分成“学生 / 会”。再比如“白天鹅”可以切分成“白 / 天鹅”，也可以切分成“白天 / 鹅”。不同的切分结果带来语义上很大的差异。我们常举“乒乓球拍卖了”这个句子为例，说明不同的切分结果会引起语义歧解（究竟卖的是“乒乓球”还是“乒乓球拍”？），这个例子很生动、也很能说明问题。提高自动分词的切分精度，关键在于处理好交集型歧义字段。分词以后还要进行词性标注，教育部（国家语委）语用所正在进行大型现代汉语均衡语料库的分词和词性标注工作；首先把核心语料库 2000 万字的生活语料加工成熟语料，在此基础上研制出《信息处理用现代汉语分词词表》和《信息处理用现代汉语词汇标记集规范》。这将是两个很重要的信息处理标准。发布后将会对我国的自然语言理解研究、语料库建设规范、语言文字规范标准制定等发挥促进作用。

到了句处理和语义处理，需要解决的问题还要多一些。比如：“你吃大碗”“我吃小碗”两个句子，“吃大（小）碗”是“用大（小）碗吃”的意思，不是真的要“吃大（小）个儿的碗”。再比如：“这种做法我有意见”“对这种做法我有意见”“我对这种做法有意见”和“巴金有剧本吗”“巴金有没有剧本”“巴金有剧本没有”这样两组句子，每组三句话尽管表层结构不同，深层语义却是相同的。所以；要让计算机理解汉语，就需要对深层语义完全相同的汉语句子进行穷尽的研究，分析到句子的深层语义结构，在找到基本句式以后再把它格式化，惟其如此，计算机才能实现对这组深层语义完全相同的句子的理解。穷尽的研究是自然语言理解研究的一个特点，它不像面对人的语法研究那样，遇到例外可以回避或者干脆“存疑”，而面对机器的研究是不能绕开难点走的。这说明，进行句处理、篇章语义处理和海量文本的自动处理，我们要走的路还很长。这些“瓶颈”的突破，涉及多个学科。不仅需要语言学特别是面向计算机的应用语言学研究成果，需要计算机硬件、软件的技术支撑，而且也需要数学、声学、医学、民俗学、心理学、社会学、仿生学、统筹学、神经语言学研究乃至自动化技术等的支持。

那么，搞技术的同志如何看待这个问题呢？一位多年从事语言信息处理研究的老专家以亲身体会对我说：分析语言的目的是什么？数字化的目标又是什么？不只是为了理

解语言，也不只是为了变成电子文本；还要研究解决信息检索和挖掘等问题，能够检索的信息才方便查找、提取和挖掘，也才更有用。王选教授 80 年代搞电子排版系统，引发了出版界一场“告别铅与火的革命”。但今天，据说《大英百科全书》已经不再出纸张版；国际标准 ISO / IEC10646 过去卖纸张版，现在则只出光盘版……现代科技发展的速度实在太快了。世界正进入告别纸张的时代。必将带来一场更为深刻的革命！在这样的大发展、大飞跃面前，信息的贯通，跨平台、跨语言的信息交换和资源共享进一步提上日程。研究解决这些问题，一要引入高科技，做好技术与语言学两方面的准备，避免低水平的简单重复，二是搞技术的和搞语言学的人一定要相互结合、相互学习。信息技术的发展；要求软件工程师也要掌握一定的语言学知识，千万不能脱离实际；一个软件工程师如果只凭借技术，只凭借软件的技巧来进行研发，很容易走进死胡同。这番话出自一位技术专家之口，正好帮助语言学者换个角度思考问题，而且可能想得更全面、立体一些。有鉴于此，有关主管部门急需组织力量，加强多学科协作攻关，深入开展基于语言信息处理的跨学科研究，以帮助汉语插上信息化的翅膀，为中国的现代化和全面建设小康社会更好地服务。

（二）标准化

标准化是当今世界的潮流，已成为国际竞争的重要手段。在现代科技迅猛发展和日益激烈的国际竞争中，标准化不仅渗透到现代科技发展的前沿，形成新的生产力，而且突破了传统的标准化领域。随着我国经济结构的战略性调整 and 社会主义市场经济体制的逐步完善，标准化工作将发挥不可替代的重要作用。

加入世界贸易组织，是改革开放进程中具有历史意义的一件大事，必将对我国未来发展产生巨大而深远的影响；同时也为标准化工作更好地发挥作用提供了发展机遇。“入世”以后，利用关税等手段保护民族产业的作用大为减弱，质检标准将起越来越重要的作用，形成“第二海关”。但是必须看到，如果质量检查所依据的标准都是发达国家制定的，对于发展中国家而言仍然是极其不利的。所以“入世”后，我们更要坚定不移地走国家标准化和国际标准化这条路，加紧制定既符合世界贸易组织规则同时又切合中国实际、对我有利的质检标准，形成“第三海关”（许嘉璐先生语）。

科技、经济活动中，我们吃忽视标准“亏”的事例很不少。远的不说，近一年多来就发生过几起震动颇大的事件。温州的打火机出口欧盟，由于产品与当地新出台的某些标准不尽相符而受限，从而使我国的打火机出口蒙受损失。又如“炒”了一年多的 DVD 之争，我国企业每生产一台 DVD 就要向有关行业联盟（以东芝为代表的 6 家 DVD 核心企业，又称“6C”）交纳专利费若干；此事虽经多次谈判，解决得仍不是很理想。实际上，就国际市场而言，中国产品受制于人的远不止打火机、DVD 两个行业，PC、手机等也都属于拾人牙慧的典型。三流的公司卖产品，二流的公司卖服务，一流的公司卖标准。中国的 DVD 企业之所以在有关谈判中始终处于劣势；一个最核心的因素就在于作为一个组装生产厂商，自己一直处在产业链条的最下端。这个事例，似乎在某种程度上契合了中国的一条古训：“劳心者治人，劳力者治于人。”由此可见，制定和掌握标准，逐步从产业链条的低端走向高端，对中国企业来说是何等重要和急迫！

可喜的是，现在这种状况正悄悄改变。昨天在来上海的飞机上看到《中国民航报》的一则消息说，为了不再吃无专利的亏，中国的 DVD 厂家正与海外知名厂家联手，制定下一代的 DVD 新标准，并在新一代 DVD 国际标准中享有中国的专利（见该报 2002 年 11 月 22 日第 6 版）。这无疑是个好消息，我们的 DVD 厂家终于意识到标准的重要性了。

由于管理体制和运行机制等原因，我国标准化工作虽然取得一定成绩，但在采用国

际标准方面，与发达国家相比仍有不小差距。我国较少参与国际标准化工作，除中文编码字符集外，由我国主导制定的国际标准寥寥无几。即便在中文信息处理领域，伴随信息技术发展带来的标准、市场、人才的竞争也都一直在潜移默化地进行着。信息技术产品中涉及语言文字的很多，外国一些大公司（如微软、摩托罗拉、IBM、富士通等）都已看到中文信息处理潜在的发展前景和巨大的市场，正大规模地涌入中国，在我国设立专门研发机构；用高薪聘请我研究人员。他们这样做的目的，不光是争夺市场，争夺人才，更重要的在于和我们争夺中文信息处理发展的制高点 and 主动权。严酷的事实告诉我们，中文信息处理的发展不仅涉及国家利益，更成为一个敏感的文化问题，甚至涉及到我国的信息安全 and 国家主权。技术标准说到底是一种游戏规则，谁掌握了标准的制定权，谁的技术成为了标准；谁也就掌握了市场的主动权。应对“入世”，我国各行各业都在制定、修订相关的标准（规范），可以预见，在这个标准化大潮中，面向中文信息处理的语言文字规范标准建设将更为急迫。必须大力加强中文信息处理领域语言文字规范标准的制定工作，适时出台数以百计甚至更多这方面的规范标准。这需要语言学界和计算机界紧密配合；联合攻关，在主管部门的组织领导、密切协调下，联合主要应用部队有计划地制定发布相应的国家标准、行业标准（规范）等，继续积极参与有关国际标准的制定工作，以维护和保持我国在中文标准化和中文计算机软件方面仅存的优势。

（三）法制化

法制化的标志是《中华人民共和国国家通用语言文字法》的出台。《国家通用语言文字法》是根据宪法制定的，它体现了国家关于语言文字工作的方针和重要政策；科学地总结了清末以来前贤们在语文革新运动中的探索实践、特别是新中国成立 50 多年来开展语言文字工作的经验、教训，反映了人民的呼声、时代和现代化的呼唤以及几代语文工作者的夙愿。《国家通用语言文字法》确立了普通话和规范汉字作为国家通用语言文字的法律地位，对普通话和规范汉字在国家机关、学校、新闻出版、广播影视、公共服务行业以及公共场所、信息技术产品、广告、招牌、企业事业组织名称和在境内销售的商品的包装、说明等方面的使用等，作出了规定。

《国家通用语言文字法》的颁行，不是语言文字规范化工作的终结，而是新的起点、新的开始；带来了新的前所未有的发展机遇。我们要主动适应法制化的形势特点和依法行政、依法管理的要求，抓住这个难得的发展机遇，乘势而上，一手抓建设，一手抓管理，依法做好信息时代的语言文字工作。当前特别要在“重在建设”上下工夫，力争在国家通用语言文字的法规体系建设、规范标准体系建设、面向信息处理的语言文字基础工程建设、应用语言学学科体系建设和语言文字管理机制建设等五个方面取得突破性进展。

（四）西部大开发

西部大开发是一个关乎中华民族长远发展和国家统一、边疆稳定大局的重要战略。开发西部不仅需要人才、物资和资金的支持，更需要信息的流通。我国西部是一个多民族、多语言、多文种、多方言的地区，目前一些少数民族语言文字的信息处理还是建立在非常落后的 DOS 平台上，无法与其它文种、图片混排。前不久，由教育部（国家语委）、信息产业部、国家民委、国家标准化管理委员会等部委联合组成的调查组在西藏自治区调研时；看到西藏日报社的报纸印刷还是落后的 DOS 系统，还在使用 286、386 微机；同志们心里很不是滋味。我国信息化发展的不平衡于此可见一斑。目前西藏自治区以少数民族文字为载体的网上信息，还要通过照相粘贴的方式才能实现。运行在 DOS 平台上的民族文系统，还不能实现图文混排及与汉文的混排。大多数藏文网站是境外办的；有的少数民族文字还无法上网。语言、方言的隔阂严重制约着西部地区的信息流通

和西部地区的发展。此外，我国一些少数民族语言文字存在着多国、多地区使用的情况，境外某些势力正在资助我国少数民族文字同英文兼容的信息处理平台的研制，甚至将他们研发的软件在我国民族地区低价销售或免费发放。有学者反映，号称支持世界上所有已成为国际标准的文种的某著名外国公司研发的信息处理平台，偏偏不支持早在 1997 年就已成为国际标准的藏文，这是极不正常的情况。

上述种种现象都值得我们高度警觉。所以，配合西部大开发战略，在西部地区加强国家通用语言文字的推广，积极制定完善少数民族文字的国家标准、国际标准，筹措资金尽快打造少数民族语言文字信息处理的 WINDOWS 平台等，将直接关系到西部地区的发展和国家的长治久安。而谋化和做好上述工作，语言文字工作部门、民族工作部门和其它有关部门负有不可推卸的责任。

综上所述，语言文字工作只有把握住形势，主动应对信息化、标准化、法制化和西部大开发带给我们的机遇和挑战，才能真正做到与时俱进，开拓创新。

新世纪语言文字工作要有大的作为，语言文字工作者首先需要固本强体。语言文字工作要与时俱进、开拓创新，为全面建设小康社会做出应有的贡献，我们就需要站得更高一点、视野更开阔一些，对语言文字工作所面临的形势有一个比较全面、准确的把握和定位；语言文字工作者就应该变“一条腿走路”为“两条腿走路”，既要做好规范标准的推行工作，又要参与规范标准的制定工作。因为哪些标准需要研制、哪些课题需要立项，第一线的语委干部最有发言权，标准的研制和完善也应该充分听取他们的意见。面对信息时代新的形势，我们要研究改进管理工作，不断拓宽管理层面，提升管理水平。这次“十五”科研立项，在各地语委的精心组织下，全国大约三千多人次申报课题 400 多项，人气很旺，热情很高。经过严格的专家函评和会议评审，目前第一批课题立项已基本结束，除少数民族人名罗马字母转写规范和普通话水平测试研究的子课题尚未确定外，其余都已基本落实了课题承担单位。我们准备将立项情况书面通知各省、市、自治区，请各地语委参与到科研管理中来，加强对本省“中标”项目的管理和中期检查、验收工作。

适应信息时代语言文字工作发展的需要，我们还应该像贵仁同志要求的那样，善于学习和思考，主动更新自己的知识结构，包括掌握计算机等现代化手段，提高自身素质。语委干部要善于归纳总结，讲政策、顾大局，不断创新，使实践升华为理论。我们每一位同志都要不断地学习新东西，掌握新本领，力争成为内行，成为某一方面的专家。

面对信息化、标准化、法制化和西部大开发的挑战，语言文字工作者还必须采取切实可行的对策。当前和今后一个时期，主要做好以下几方面的工作：

（一）加强语言文字规范标准建设，使管理工作有科学、明确的依据。

要把语言文字规范标准建设同信息技术的发展紧密结合起来，通过调查研究，摸清情况，探索语言文字规范标准建设的新思路。争取用五年左右的时间，构建起规范标准的体系框架。规范标准建设是语言文字工作的基础，是社会推行的依据。语言文字规范标准建设的思路是：政府宏观指导，社会科研支撑，及时维护更新。要确立以科研促进工作发展的观念，因为实现语言文字规范化，一要靠制定科学、适用的语言文字规范标准，二要靠社会对这些规范标准的大力推行，三要靠运用现代化的科技手段来处理语言文字。而这些都离不开科学研究。可以说，语言文字及其应用的科研水平，直接关系到我国语言文字规范化、标准化的水平。因此，必须大力加强科研管理，提倡开门办科研，组织多学科联合攻关，充分发挥地方作用，鼓励社会力量以多种方式开展科学研究，逐步形成中央和地方既分工又合作的语言文字科研体制。当前要抓紧社会急需的语言文字规范标准的研制工作，如《规范汉字表》《信息处理系统语言文字评测标准》等的制定。

“汉语普通话水平测试研究”也是一个重点课题，涉及普通话水平测试的理论与实践、普通话水平测试标准、计算机辅助普通话水平测试、测试的现代化管理及信息化建设、基于网络的普通话培训测试研究等五个子项目，其中包括了测试标准研制的子项目，我们已委托国家语委普通话培训测试中心，整合各地语委、测试中心和有关高校的力量来进行，希望大家多支持这项工作。

在文字标准化方面，我们的基本想法是一边抓《规范汉字表》的研制（该字表中的一、二级字，自然应当成为“CJK 基本国际汉字子集”的组成部分），一边会同有关部门，积极参与并主导国际标准“信息技术通用多八位编码字符集”（ISO / IEC 10646）的制定工作。

《规范汉字表》研制是教育部、国家语委去年 5 月启动的“十五”重大项目，预计今年底可以拿出字表初稿。新中国成立以来，政府在汉字的整理和简化方面做了大量工作，有力地促进了汉字社会应用的规范化、标准化，但以往所颁布的规范标准也存在字量不足、字表与字表之间系统性不够等问题，给使用带来不便。进入信息时代，特别是《国家通用语言文字法》颁行后，迫切需要一份字量合适、收字合理、字形规范、排序科学的《规范汉字表》以确保《国家通用语言文字法》的贯彻执行。这是该项目能够提早启动的直接原因。为使字表研制具有扎实的学术基础，课题组在北京、上海、湖南等地先后听取了语言文字学界、语文教学界、中文信息处理界、辞书出版界和书法界等的意见；分别召开了汉字规范问题学术研讨会（2001 年 12 月，上海）、异体字问题学术研讨会（2002 年 5 月，井冈山）、简化字问题学术研讨会（2002 年 6 月，合肥）、汉字印刷字形问题学术研讨会（2002 年 8 月，烟台）等系列学术会议，就有关专题集思广益、深入研讨；在全汉字库暨国际音标问题商讨会（2002 年 7 月，哈尔滨）和两岸四地中文数字化工作研讨会（2002 年 10 月，福州）上，商谈的内容也涉及到《规范汉字表》研制的一些情况。

在广泛征询意见的基础上，课题组经报请有关部门同意，确定了《规范汉字表》的研制原则和总体框架。研制原则共 7 条：①兼顾一般社会使用和信息处理的需要，使字表和信息处理用字符集趋于一致；②坚持科学严谨的作风，不回避、掩盖矛盾；③力求稳定，尽量不出现或少出现新造的字形；④处理问题时把汉字的形、音、义结合起来考虑；⑤着眼于现实应用，兼及汉字的源流演变和字理；⑥充分尊重教师、语言文字学家、辞书专家、书法家和其它汉字使用者的合理意见；⑦尽可能兼顾海外华人的用字习惯，缩小与使用汉字的国家和地区现行汉字字形的差异，为最终实现汉字文化圈的“书同文”打下基础。根据上述原则，《规范汉字表》拟收规范字头 30000 字左右，字表按使用频度分级，其中：一级字为高频字，3500 个左右；大致相当于教育基本用字表的范围；二级字为次高频率，4500 个左右，连同一级字共 8000 字上下，大致相当于现代汉语通用字的范围；三级字为低频字，约收 4000 个，大致相当于专门用字表的范围，一、二级字中不能包含的姓氏、人名、地名和科技术语用字在三级字内解决；四级字为最低频字，大约 18000 个，均为罕用字，主要满足古籍整理、出版印刷使用字量较大时的特殊需要。《规范汉字表》的编纂体例：以规范字为字头，已简化的字后面标注繁体；有异体的字在选用字（正体）后面标注相应的异体，新旧字形有明显差异的标出旧印刷字形；对于简、繁体不是一一对应关系，正、异体之间有读音或意义差别的，另加注释说明，以方便读者使用。这个字表的总字量及分级设计是否合适，最终还要经过实践检验。字表初稿拿出来后，我们将送请有关专家、各类“用字大户”及各地语委审阅。反复征求意见以后再确定。

《规范汉字表》主要是解决当前汉字的使用问题（特别是其中的一、二级字）；而国际标准“信息技术通用多八位编码字符集”（ISO/IEC 10646）则是为汉字文化典籍的

电子保存、数字化图书馆和各种数据库建设、使用汉字的国家和地区的信息交换和资源共享服务的。ISO / IEC 10646 “信息技术通用多八位编码字符集”是一项全球所有文种统一编码、使计算机系统实现多文种相互切换和交流的国际标准，其基本功能是在全球范围内建立起实时、无障碍的信息交换模式。积极支持该标准的制定工作，对于推动中华汉字典籍的数字化，建立数字化图书馆、在互联网上交换和共享中文文献、向世界展示中华文明，具有重要意义。可以说，这项最基础、影响面最大、最具战略性的语言文字信息处理标准的制定工作，关系到中国的长远利益和信息化建设，关系到使用汉字的国家、地区的相互沟通和资源共享，也关系到中国境内少数民族文字的发展和现代化。中国是汉字的故乡，理所当然应该在该项标准研制中发挥主导作用。多年来，在有关政府部门、语言文字和信息处理专家的共同努力下，我国积极参与并主导“信息技术通用多八位编码字符集”的研制工作，取得较大成绩。尽管收入该字符集的很多字并不常用，但也有个上“户口”、占码位的问题，否则将无法实现信息交换和资源共享；因此我们万不可小视这项工作。要努力使没有进入该字符集扩充 B 和扩充 CI 的人名地名用字、科技术语用字、出版物用字、中医药用字和汉字系少数民族文字尽快进入该标准。为此，教育部、国家语委会同有关部委已经做了大量沟通、协调工作。包括召开会议，组织征集字样、进行“排重”和按照规定的文本格式提交等，取得了明显效果。此外，古汉字、少数民族文字的信息化问题以及记录中国语言（含汉语方言和少数民族语言）的国际音标的扩充问题等，也已提上日程；教育部、国家语委拟启动《中华大字符集》工程，统筹解决上述问题。最近，我们还邀请香港、澳门特别行政区和台湾地区的代表，就中文数字化的有关问题进行研讨、协调立场，会议在两岸四地共同努力下开得很务实、很成功，形成了“坚定不移地走国际标准化道路，坚持 ISO / IEC 10646 的方向”的共识和“建立 CJK 基本国际汉字子集”的共同建议，并组成两岸四地中文数字化合作论坛，定期开会研讨中文数字化应用的共同课题（详见《两岸四地中文数字化合作研讨会会议纪要》）。

（二）以信息化为背景，抓好对语言文字信息处理具有推进意义的基础性工作。

①搭建语言文字政策制定、基础研究、应用研究和规范标准建设的坚实平台，抓紧进行面向信息处理的语言文字基础工程建设，包括建设国家级现代汉语均衡语料库、建立语言文字应用信息库和专家库，开通中国语言文字网等，合称“三库一网”。现代汉语语料库正在边建设边利用，即将面向社会发布使用，以充分发挥其在语言文字应用研究、规范标准研制和语言文字信息处理中的作用。语言文字信息库正在建设中，尚需各级语委、有关高校、科研单位和语文爱好者一起努力，最终建成语言文字工作的数字化图书馆和反映百年语文现代化进程的数字化档案馆。语言文字专家库将语言文字学界、中文信息处理界和其它有关学术领域的专家分门别类建库，初步发挥了查找、调用人才的功能。中国语言文字网经过一年多的建设和试运行，已于今年 6 月正式开通，在宣传国家语言文字工作方针政策、面向社会提供咨询服务、开展社会热点问题讨论、促进学术交流和引导社会正确使用祖国语言文字等方面初步发挥了作用，取得良好的社会效益。目前“三库一网”仍在建设和完善中。

②有针对性地开展调研，加强信息时代急需的“两栖型”高层次人才的培养。调研内容包括信息技术和信息技术产品中的语言文字问题及制定管理条例的必要性、可行性，“两栖型”高层次人才的需求情况，大规模现代汉语均衡语料库的知识产权问题，如何加强语言文字应用研究成果的认证评测和信息技术产品的市场监测问题等。制约语言文字信息处理发展的关键问题是人才短缺，特别是兼通语言学和信息科学的高层次“两栖型”人才奇缺，我们要努力创造一种有利于创新型、“两栖型”甚至“多栖型”人才成长的机制，尝试建设应用语言学人才培养基地，以应对当前语言文字工作

所面临的紧迫形势。

（三）抓好语言文字社会应用管理工作，为国家信息化、现代化建设和全面建设小康社会营造良好的语言文字环境。

《中华人民共和国国家通用语言文字法》颁行后，我国语言文字工作的开展和深化进入了最好的时期。要抓住机遇，紧紧围绕《国家通用语言文字法》的宣传贯彻，做好语言文字规范标准的推行和社会应用管理工作。要坚持“一个中心四个重点三项基本措施”，即以城市为中心，以学校为基础，以党政机关为龙头，以新闻出版、广播影视等媒体为榜样，以公共服务行业为窗口，通过开展城市语言文字工作评估、普通话水平测试和推广普通话宣传周等措施，大力推广普通话、推行规范汉字，促使语言文字管理工作迈上新台阶。

以上从学习、贯彻党的十六大文件精神入手，侧重分析了信息时代语言文字工作面临的形势，本着“与时俱进、开拓创新”的思想，简要介绍了因应对策。

**To keep abreast with the times, to exploit and innovate, striving for
better language work in this Information era**

Wangtiekun

收稿日期：2002-12-28

作者简介：王铁琨 男，教授。教育部语信司副司长、中国文字学会副会长兼秘书长。