

SWCL2002

**1st Students' Workshop on
Computational Linguistics
Proceedings**

August 20 – 23, 2002

Initiated by

Chinese Information Processing Society of China

Organized by

Institute of Computational Linguistics
Peking University

信息时代的语言问题（代序）

教育部语言文字信息管理司 李宇明

一、信息时代的数字鸿沟

（一）信息化时代

我们正在跨入信息时代。对信息时代可以从不同的方面进行描述或定义，就技术角度来看，可以将信息时代定义为用数字化技术对信息进行处理的时代。这里的信息处理，指的是利用计算机和互联网对信息进行收集、整理、储存、交换、检索以及各种再加工。

人类历史始终伴随着信息技术的发明与改善，每一次新的信息技术的出现，都会带来人类社会的进步，甚至是翻天覆地的变化。近期多学科的研究表明，语言并不是与人类一起产生的，语言产生之前人类经历了漫长的“前语言阶段”。在前语言阶段，人类的信息交际使用的是各种身势和叫喊声，大约与黑猩猩等猿类的交际差不多。^①大约在距今四五万年之时，人类进化出了语言，^②信息加工和交换的能力发生了质的变化，语言使人类最终脱离动物界成为万物灵长。文字的出现帮助语言打破时空限制，促进了信息加工的精密化和信息传递的效率，使人类跨入文明时代。之后，文字书写、印刷技术、传声通讯技术的一步步发展，电话电视等模拟信号系统的相继问世，带来了一次又一次的信息革命浪潮，使人类快步进入现代化。

数字技术的发明，使大数量、快速度、高质量地处理信息逐渐成为现实，信息将逐渐成为生活的凭借和生产的资本，对人类的生活与生产活动将发挥空前的作用。因此，信息时代，也可以说是利用计算机和网络对信息进行处理的“数字化时代”。在数字化时代，信息处理的水平常常代表着一个国家经济发展的水平。以信息产业为代表的高新技术产业，跃居国民经济中的首要地位。

（二）数字鸿沟

由于各种历史原因，世界各国在政治、经济、科技、教育等方面的发展很不平衡，国际间严重的贫富差异，引发了许许多多国际性的社会问题，威胁着人类的安宁甚至生存。在信息化时代，世界发展的不平衡突出表现为日益扩大的“数字鸿沟”。

1993年9月，克林顿政府宣布将实施“信息高速公路”计划。1994年1月25日，克林顿总统在“国情咨文”中雄心勃勃地指出，要在2000年以前把美国所有的教室、医院和诊疗所、图书馆通过网络联为一体。1994年2月，欧洲委员会不甘示弱，也宣布兴建“信息高速公路”。接着全球掀起了“修路”热。^③1995年，国际互联网开始商业化并带来了信息爆炸。1998年2月26日，克林顿总统发表了被称为“网络新政”的重要演说。1999年12月，欧洲委员会提出建立“网络欧洲”的“政治倡议”。2000年，日本森喜朗内阁提出“日本式IT社会”的构想。

数字鸿沟的外在表现，是国际互联网用户数量分布的不平衡。胡鞍钢、周绍杰（2002）指出，全球收入最高国家中的1/5人口拥有国际互联网用户的93%，而收入最低的1/5人口只拥有国际互联网用户的0.2%。2000年8月，北美地区每千人拥有互联网主机为168.68台，而非洲地区每千人拥有互联网主机仅为0.31台，是北美地区的1/540。2000年10月，经济合作与发

^① 关于黑猩猩等的交际手段，请参见李宇明主编（2000，P191~192）。

^② 参见胡明扬（1985，P17~23）。

^③ 见姚亚平（1997，P34~35）。

^④ 见胡鞍钢、周绍杰（2002，P34）。

展组织（OECD）的国家拥有国际互联网主机总量的 95.6%，而非经合组织国家只拥有国际互联网主机总量的 4.4%。显然，一些国家或地区是现代信息技术的“数字富有”者，而一些国家或地区沦为“数字赤贫”者。

数字鸿沟，绝不仅仅是互联网使用者的数量悬殊和多少人用不上互联网的问题，由于信息时代的特点，数字鸿沟会带来社会、经济、科技、教育等方方面面的鸿沟。数字鸿沟意味着知识资源配置严重失衡，使没有机会使用信息网络的国家或个人，处在经济全球化和信息革命的边缘地带，数字鸿沟带来社会鸿沟。信息产业是当前世界最为重要的经济增长点，数字鸿沟带来经济鸿沟。信息技术成为当前的领先技术和从事科学研究必需的现代化手段，数字鸿沟带来科技鸿沟。联合国教科文组织认为，新的信息技术为改进学习和教学、增加学习机会、提高教育质量、改善管理和办学提供了极大潜力，数字鸿沟加深了知识鸿沟。^①就此而言，数字鸿沟应定义为由于信息技术发展不平衡而带来的社会方方面面的巨大差别。

最早提出数字鸿沟概念的是美国远程通讯和信息管理局（NTIA）。1995 年 7 月至 2000 年 10 月，该机构四次发布了美国国内的数字鸿沟状况，将数字鸿沟列为美国首要的经济问题和人权问题，目的是要全部美国人融入网络社会。^②近些年来，国际社会极为关注数字鸿沟带给国际社会的严重问题。例如，2001 年 8 月 23 日，第四届“九个人口大国全民教育部长会议”通过的《北京宣言》，提出“通过挖掘信息与传播技术在取得全民教育目标方面的潜能，缩小数字鸿沟”。^③2001 年 9 月 5 日至 8 日，联合国教科文组织第 46 届国际教育会议在日内瓦召开，会议将《北京宣言》列入教科文组织文件，并热烈讨论了缩小教育领域的数字鸿沟问题。^④

二、数字鸿沟的语言效应

数字鸿沟并不仅仅影响政治、经济、科技、教育，对人类语言生活也发生了有史以来最为严重的影响。这种影响主要表现在两个方面：

（一）语言更为不公平

语言之间本应当是平等的，但是由于操不同语言的民族在人口、经济、文化、军事等方面的发展不平衡，产生了事实上的语言不平等。绝大多数语言只为极少数人在日常交际中使用，发挥的作用很小；而有些语言成为国家、地区乃至世界的共同交际工具，成为语言中的庞然大物。例如现代国际社会使用较多的语言只有英语、法语、俄语、西班牙语、阿拉伯语等数十种，英语是超级的庞然大物。

英语在 1600 年前后还只是一个小语种。拉丁语是西方古代的“超级语言”，直到中世纪，拉丁语还是欧洲西部的通用语。17 世纪法语取代拉丁语成为外交用语。到第一次世界大战结束时英语才有与法语一决雌雄的资格。^⑤二战之后，美国成为军事强国和技术盟主，使得英语超过法语成为外交、贸易、科技、教育的第一大语言。20 世纪 50 年代以后，美国、英国倾力向世界推行英语，英语的势力急剧扩张，成为当今的超级语言。^⑥

美国是新信息技术的遥遥领先者。1998 年 7 月，美国千人拥有互联网主机超过 179 台，而 OECD 国家平均只有 62 台。2000 年 7 月，美国千人拥有互联网主机超过 230 台，而 OECD 国家平均仅为 88 台，欧盟各国平均仅为 42 台。^⑦美国在信息技术方面的发展，使英语如虎添翼，几乎独霸了网络用语。据统计，2001 年，78% 的网上语言是英语，世界其他语言只占有网络语言的 22%。

^① 见中国联合国教科文组织全国委员会秘书处（2001，P33）。

^② 见胡鞍钢、周绍杰（2002，P35）。

^③ 见中国联合国教科文组织全国委员会秘书处（2001，P23）。

^④ 见中国联合国教科文组织全国委员会秘书处（2001，P40、P33）。

^⑤ 见斯坦利·利伯森（1982，P608~609）。

^⑥ 参见罗伯特·菲利普斯（1994）。

^⑦ 转引自胡鞍钢、周绍杰（2002，P39）。

语言不公平历来存在，但是到了 20 世纪末，竟然发展到了国际大语言之间也差别悬殊的地步。语言不公平必然会带来语言歧视，语言歧视又必然产生民族和文化歧视。联合国教科文组织第 46 届国际教育会议呼吁：“世界文化是一个建设与更新的进程，不能由某一国或某一地区的文化所主宰。各地区和各民族的文化应通过融入世界文化而获得生命力。”^①

语言是金钱，语言不公平也是财富分配的不公平。非英语的国家和民族，必须花费大量时间和金钱学习英语这种权势语言。如果用两万元的学习费用才能达到掌握 5000 英文单词的听说读写水平的话，那么，一个英文单词的价格就是 4 元钱。计算机和网络等现代信息技术设备，必须适应英语及英语国家的设备标准。如此等等，不一而足。非英语国家学习语言的费用和技术设备的投入等等，有相当一部分流向了英语国家，从而加大了经济上的贫富鸿沟。

（二）大量语言发生生存危机

世界上存在或存在过 5000 多种语言，其中有几百种已经消亡，特别是美洲、澳洲的土著语言。近两个世纪，随着工业化的发展以及地区、国际市场的形成，语言消亡的速度明显加快。目前世界上又有上百种语言濒临灭亡。周庆生（2000，P86）指出，在中国，赫哲族、仫佬族、土家族、畲族、满族，现在还能讲本族语的占本民族总人口的比例分别为 14.78%、10.36%、5.27%、0.11%、0.01%，它们都是濒危语言。

语言消亡的原因，据斯蒂芬·武尔姆（1991）的研究，主要有使用语言的群体的死亡，语言生态环境的变化，文化接触与文化碰撞，政治影响与征服等。斯蒂芬·武尔姆发表论文的时候，数字鸿沟还未引起注意。在信息化时代，数字化带来的语言的极度不公平，加快了语言消亡的速度，扩大了语言消亡的数量。今天，信息化已经成为语言消亡的一个因素，甚至是重要因素。

我们生活在现实空间和由互联网塑造的虚拟空间中。随着信息化的发展，虚拟空间越来越重要，许多见微知著的贤达人士，预测将来人类社会是以网络为主导的社会。现实空间中的语言状况，是历史自然发展的结果，而虚拟空间的语言使用，一开始就是以语言霸权的方式出现的，信息化发展较快的国家控制着语言霸权，不断强化着语言霸权。可以预见，世界上绝大多数语言都没有机会成为网络语言，有机会成为网络语言的一些语种，也因其没有掌握网络语言霸权而只能成为网络的点缀。在虚拟空间中，俄语、法语、德语、西班牙语等大语种，也都发生了重大的生存危机。

虚拟空间与现实空间并不是相互隔绝的，而是时时相通交互发展的。一方面，虚拟空间的发展依赖现实空间的经济、信息技术水平及其普及状况，另一方面，现实空间的运行越来越倚重虚拟空间，越来越受到虚拟空间发展的制约。因此，信息化不仅在现实空间影响语言的存活，而且由它建构的虚拟空间的语言状况，对现实空间语言状况的影响也越来越大。不能成为网络语言的语种，其生命力将会逐渐丧失；在网络上处于劣势的语种也会逐渐衰落，甚或消亡。

斯蒂芬·武尔姆（1991，P89）同意一些语言学家的看法，“语言是文化和社会的内在部分，它反映了人们看待周围物质世界和精神世界的方式。……每种语言都反映一个言语社区在世界上解决问题、形成自己思维与哲学体系以及理解周围世界的方式。一旦这种语言死亡……，我们就永远失去了某种不可替代的关于人类思维方式和世界观的知识。”国际社会对信息时代“单一市场、单一语言、单一文化”的发展趋势，表现出极大的普遍的忧虑。许多国家在制定应对信息化的战略时，也在对本国语言的发展进行战略性的思考。2001 年 10 月在巴黎召开的联合国教科文组织第 31 届大会，审议并通过的《教科文组织文化多样性宣言》，^②便是国际社会这种忧虑的反映。

^① 见中国联合国教科文组织全国委员会秘书处（2001，P31）。

^② 见中国联合国教科文组织全国委员会秘书处（2001，P47~48）。

三、信息时代的中国语言对策

(一) 中国信息化的发展

我国一直关注国际信息化发展的动向，积极推进国家的信息化建设。下面，以互联网为例来看我国信息化的发展。

1987年9月20日，钱天白教授发出我国第一封电子邮件“越过长城，通向世界”，揭开了中国人使用互联网的序幕。1993年3月2日，中国科学院高能物理研究所租用AT&T公司的国际卫星信道接入美国斯坦福线性加速器中心(SLAC)的64K专线，这是我国部分连入互联网的第一根专线。1994年4月20日，中关村地区教育与科研示范网络(NCFC)通过美国Sprint公司连入互联网，自此我国被国际上正式承认为有互联网的国家。

上文提到，1993年9月，克林顿政府宣布实施“信息高速公路”计划，而我国则在1993年3月12日，时任副总理的朱镕基就主持会议，提出和部署了建设国家公用经济信息通信网(简称“金桥工程”)工作。同年8月27日，政府决定用300万美元总理预备金支持启动金桥前期工程建设。1994年6月8日，金桥前期工程建设全面展开。1996年2月11日，国务院第195号令发布了《中华人民共和国计算机信息网络国际联网管理暂行规定》，以促进我国互联网的有序发展。^①2001年的《中华人民共和国国民经济和社会发展的第十个五年计划纲要》，制定了“加速发展信息产业，大力推进信息化”的战略。^②北京、上海等地纷纷提出建设“数字化城市”的要求。

1997年10月到2002年7月，中国互联网络信息中心(CNNIC)联合四个互联网络单位，每半年发布了一次“中国互联网络发展状况统计报告”，现已发布10次。六年来，中国互联网在上网计算机数、上网用户数发展情况如下表所示：

年月	1997-10	1998-6	1998-12	1999-6	1999-12	2000-6	2000-12	2001-6	2001-12	2002-6
A	29.9	54.2	74.7	146	350	650	892	1002	1254	1613
B	62	117.5	210	400	890	1690	2250	2650	3370	4580

说明：1.年月是材料统计时截止的时间。2.A代表上网计算机数，B代表上网用户数，统计单位为“万”。^③

从上表可以看到，中国互联网的发展相当迅猛，2002年6月底上网的主机数和网民数量，就绝对量来看，正在接近发达国家水平。中国互联网和其他信息技术的发展，为汉语在虚拟空间的发展提供了良好条件。网络上的汉语使用量逐年增长。中国的信息化正在快步追赶世界先进水平。

(二) 汉语的优势与不利因素

汉语是对当今世界发生影响越来越大的语言，在信息时代，汉语存在着优势，也有相当大的危机。

汉语的优势首先是人口优势，据估计，海内外以汉语为母语的人大约有15亿，是世界上使用人口最多的语言之一。其次，汉语具有文化优势，它负载着中国悠久的对世界充满魅力的优秀文化。再次，汉语还具有潜在的经济优势，近些年中国经济持续高速度发展，综合国力和国际地位明显提高，美国、韩国、日本等国家，纷纷把汉语作为主要外语纳入国民教育体系。

汉语的不利因素表现如下一些方面：

汉语的使用人口虽然众多，但是一致性差。内部方言分歧严重，汉字又有繁体字与简化字之别。普通话和规范汉字的普及程度近些年有重大进展，但还相当不理想，在公务活动、教学

^① 根据中国互联网络信息中心 (<http://www.cnnic.net.cn/>) 网上的资料整理。

^② 见朱镕基 (2001)。

^③ 中国互联网发展的材料，根据中国互联网络信息中心 (<http://www.cnnic.net.cn/>) 网上的资料整理。

活动中，在公共场合的交际中，不使用普通话和规范字的情况还相当多。不少调查显示，大陆年轻汉族人的的母语水平在急剧下降。少数民族聚集区域，国家通用的语言文字并非到处通用。在升学、晋职晋级、就业等领域，事实上存在着轻汉语重外语（主要是英语）的政策规定或心理倾向。社会的语言规范意识薄弱，大众传媒用语用字不规范不讲究的现象还相当常见。对外汉语教育、少数民族的汉语教育和汉语的母语教育三种汉语教育，不仅各自的教育理论和教学方法都不成熟，而且在管理和研究上也没有建立有机高效的协调机制。

对现代汉语现代汉字的研究还不够深入，偌大的中国没有专门的普通话研究室和现代汉字研究室，全国用于语言学的研究经费十分有限。汉语汉字的标准不健全，例如，没有汉语的音节全表和与汉语地位相称的语法长编，只在常用的字量的范围内有印刷汉字的一些规范标准等。

中文信息处理的水平离发达国家还有不小差距，特别是在核心技术方面拥有自主知识产权的成果还不多。网络的对公众开放的汉语数据或包括汉语在内的多语种数据库还很少，导致网络的利用率低、实用价值低。

汉语在国际语言生活中也不占优势。将汉语作为国家通用语言或工作语言的国家和地区不多；联合国虽然将汉语列入工作语言，但是地区性或国际性的组织或会议真正使用汉语的也不多；汉语在地区或国际的重要交际领域，如外交、贸易、科技、教育等，使用十分有限；非汉语区域的华人后代，保持汉语相当艰难，放弃汉语的人并不在少数；非华人的外国人学习汉语的人数正在增加，但真正了解汉语或在社会生活中使用汉语的，却是凤毛麟角。

这些因素的存在，不利于汉语的固本与优化，不利于国家通用语言文字的国内推广与国际传播，不利于提高汉语的声望及其在国内外的地位。

（三）汉语发展对策

为在信息化时代提高汉语的功能与地位，加速汉语的传播，应主要做好以下几个方面的事情：

1. 大力推广普通话

早在清末之际，就有人提出要把语言统一看作“立国之要素”，^①认为统一语言可以“合四方为一心，联万方为一气”，^②“助团体之凝结，增长爱国心”，^③“保全国粹”^④。语言统一首先是方便交际，但它的确也是增强民族凝聚力的重要因素。今天看来，它还是促进信息化发展、让汉语走向世界的基础。

提升大众的语言民族意识、培养对普通话的认同和讲普通话的自豪感、使普通话具有较高的语言声望，实乃当务之急。这就需要教育、广播电视领域、领导干部带头讲好普通话，为社会树立榜样。要丰富说话人的语言技能库，使越来越多的人能够讲普通话，并在各种语言场合切题合意地使用普通话。当前，除了做好城市和公共服务部门的推广普通话工作之外，应注意通过市场职业认证制度抓好公共服务行业的普通话培训与测试；要研究小城镇和农村、以及参加城市建设的农民工的语言问题；要努力在少数民族地区推广国家通用语言；要加强海内外华人社区语言的沟通，支持海外华人及其子弟为语言保持而进行的汉语教育。

2. 加快汉语汉字规范化建设

成熟的优秀的有影响的语言，必须在语言结构、文字系统和语言文字的社会应用等方面，建立一系列科学、管用的规范和标准，以便于语言的教学与使用，便于语言文字的信息处理，并有利于语言声望的提高。

除了考虑语言的基本规范（如规范字表等）之外，应特别注意制定社会语言生活中的各种

^① 庆福等（1910）。

^② 林谔存（1898）。

^③ 吴汝纶（1902）。

^④ 卢慧章（1906）。

应用性的规范标准，使语言生活健康有序。例如不同教育类型（扫盲教育、学校母语教育、少数民族汉语教育、对外汉语教育）和不同教育层次（幼儿园、小学、初中等）用字用词规范，人名地名用字和翻译转写规范，低点阵汉字屏幕显示的省简规范，计算机字库、词库的建设规范，网络出版的语言文字规范，面向网络的专有名词的缩略规范，等等。

3. 充分发挥汉语拼音的作用

汉语拼音是经过百余年的社会实践^①、在上千种方案中优选出来的国家法定方案，其作用是拼写汉语和为汉字注音，并在汉字不便使用或不能使用的场合发挥作用。汉语拼音是汉字的得力助手，特别是在中外交流和人与机器的交流中，其作用非常重要、不可缺少。

汉语拼音方案为汉字注音，已经非常称职，但是，如果行使拼写汉语的职能和在汉字不便使用或不能使用的场合发挥作用，汉语拼音方案就还有许多需要研究和完善的地方。除了同音词区分、大小写规则、分词连写规则等传统课题之外，还有许多应用问题，特别是教育、信息处理和国际交流中的问题，需要花费较多的精力来解决。

对汉语拼音方案的应用研究，非常薄弱；目前在小学教育中出现了“淡化”汉语拼音教学的倾向；成人对汉语拼音的应用能力非常之低；社会对汉语拼音的认识和应用还比较混乱。汉语拼音所应当发挥的作用远没有发挥出来。这些问题应当引起足够的重视。

4. 加快中国语言文字信息处理的研究

语言文字的信息处理，是我国信息化的“瓶颈”，而且也是我国信息化的必决之役、必胜之战。中国语言文字的信息处理正在由字词处理向语言处理发展，由文字处理向语音处理的方向攻坚。当前，首先应当实施人才战略，集结语言学和信息技术等多学科人才联合攻关，并联合培养综合性的语言信息处理人才，特别是要培养一批学科的领军人物。其次是实施资源战略，要集中力量建设各种具有自主知识产权的语言资源和软件资源，如具有语法和语义标记的书面语语料库和口语语料库，各种信息词典和知识词典等，并通过明晰知识产权和统一规范标准，增加资源的共享与互补。第三实行开发战略，及时将科研成果转化为产品占领市场，发挥社会效益和经济效益，并通过产品效益支撑科学研究。争取用五到十年的时间，使我国的语言信息处理跨上新的台阶。

5. 全方位建设基于网络的数据库

网络是“路”，网页是“车”，数据库便是车载之“货”。“货物贫乏”是网络效益低的最大问题。没有足够的数据库，特别是没有足够的面向公众的数据库，电子政务、电子商务、电子教育和电子娱乐就无法顺利进行。数据库的语言基本上应使用汉语，既便于国人使用，增加更多的网民，也能迅速增加虚拟空间的汉语的比重。同时，还要注意开发以汉语为中心的多语种数据库，方便外国人访问我国的网站。虚拟空间也遵循“拥有者占有”的公理，建立一个领先的功能强大的数据库，就等于拥有了一个领域、一个学科。网络世界的分割刚刚开始，占领更多的网络世界的领地，依赖于各种各样的数据库。

中国文化是人类的宝贵财富。建立数字化图书馆等中国文化的数据库，以使中国文化以数字化的新的媒体方式永世长存，且便于中外研究者运用数据库开掘中国文化的宝藏。要建设中国文化的数据库，必须将负载中华文化的古今全汉字、古今民族文字和各种有用的符号进入 ISO 空间，以保证将中国文化以其本来面貌保存、开发及在国际互联网上的无障碍交流。

6. 研究汉语的国际传播战略

美国、英国、法国、西班牙、德国^②、日本^③等国，多年来都在有计划地实施英语、法语、

^① 如果将自明代传教士利玛窦起历代西方教会为汉语设计的教会罗马字，汉语拼音的社会实践年限还要长。

^② 见乌尔里希·阿蒙（1992）。

^③ 见平高文也（1992）。

西班牙语、德语、日语等国际战略。美英等国主要是通过贸易、媒体、教育、文化等途径向世界倾销英语，试图建立“语言领域的帝国主义”。^①法国努力协调法语拥有特殊地位的34个国家和3个国家的部分地区的所谓的法语区（Francophonie）的语言问题，任命一位部长主管法语区的语言问题。^②西班牙在利用“西班牙语世界”（Hispanidad）这一概念向世界进行语言传播（Language Spread）。^③

在信息化时代，汉语的国际传播应当引起重视，应有明确的思路和行之有效的措施。比如，从人群和地区来看可以划分三个战略区域：海外华人社区，传统的汉字文化圈（朝鲜、韩国、日本、越南等），辐射圈（世界其他国家或地区）。可以首先加强海外华人社区的语言协调，并与汉字文化圈以文化为基底结成语言同盟，再尽力向辐射圈辐射。比如可以建立这样的辐射步骤：先争取国际会议的汉语地位，再逐步争取国际组织以汉语作为工作语言的地位，争取汉语在汉字文化圈和辐射圈的主要外语的地位。当前主要的传播手段，可以定位于文化、贸易、教育以及在中国举行的大型国际活动（如2008年奥运会、2010年世博会等）等。从组织形式上看，可以以民间的国际汉语传播组织为主，官方成立有权威的高效率的协调机构。在汉语国际传播战略的指导下，通过各种具体操作，促进汉语尽快走向世界。

四、结语

信息化的飞跃发展在世界范围内形成了数字鸿沟，数字鸿沟带来了语言间更大的不公平，并威胁着许多语言的生存。我国的信息化正在快步追赶世界先进水平，这为汉语在信息化时代的生存与发展提供了良好基础。

在信息化时代的语言竞争中，汉语有着人口、文化等方面的天然优势，而且由于中国经济的发展和国际地位的提高，汉语逐渐以外语的身份进入了一些国家的国民教育体系。但是，汉语的劣势也是相当明显的，例如：汉语方言分歧，书写系统不一致，语言的规范标准不健全；普通话推广任务艰巨；网络上的汉语使用十分有限；汉语在国际语言生活中的作用较小等等。

面对数字鸿沟的挑战，中国应有明确的语言对策：通过推广普通话、完善汉语汉字的规范标准、推行和完善汉语拼音方案等措施，使汉语固本强体；通过语言信息处理和网络数据库的快速建设，争夺虚拟空间的汉语地位；通过一系列汉语国际传播的措施，使汉语在国际语言生活中发挥更大的作用。

参考文献

- [法国]阿尔米·科莱内当 1992 语言传播政策和法语区：法国采取的语言行动，周庆生主编《国外语言政策与语言规划进程》，语文出版社，2001年。
- [西班牙]阿吉利诺·桑切斯 1992 西班牙语的传播政策，周庆生主编《国外语言政策与语言规划进程》，语文出版社，2001年。
- 胡鞍钢、周绍杰 2002 新的全球贫富差距：日益扩大的数字鸿沟，《中国社会科学》第3期。
- 胡明扬 1985 《语言与语言学》，湖北教育出版社。
- 李宇明主编 2000 《语言学概论》 高等教育出版社。
- 李宇明 2001a 规范语言文字，推进信息化进程，《中国教育报》5月7日第1版。
- 李宇明 2001b 信息时代需要更高水平的语言文字规范，《术语标准化与信息技术》第3期。
- 林辂存 1898 上都察院书，《清末文字改革文集》，文字改革出版社，1958年。
- [丹麦]罗伯特·菲利普森 1994 英语传播政策，周庆生主编《国外语言政策与语言规划进程》，语文出版社，2001年。

¹ 参见罗伯特·菲利普森（1992）。

² 见阿尔米·科莱内当（1992）。

³ 见阿吉利诺·桑切斯（1992）。

-
- [丹麦] 罗伯特·菲利普森 1992 *Linguistic Imperialism* (语言领域的帝国主义), 上海外语教育出版社, 2000年。
- 卢懋章 1906 颁行切音字书之益,《清末文字改革文集》,文字改革出版社,1958年。
- [日本]平高文也 1992 日本的语言传播政策,周庆生主编《国外语言政策与语言规划进程》,语文出版社,2001年。
- 庆福等 1910 陈请资政院颁行官话简字说帖,《清末文字改革文集》,文字改革出版社,1958年。
- [澳大利亚]斯蒂芬·武尔姆 1991 语言的消亡与消失:原因与环境,周庆生主编《国外语言政策与语言规划进程》,语文出版社,2001年。
- [美国]斯坦利·利伯森 1982 语言传播面面观,周庆生主编《国外语言政策与语言规划进程》,语文出版社,2001年。
- [德国]乌尔里希·阿蒙 1992 联邦德国传播德语的政策,周庆生主编《国外语言政策与语言规划进程》,语文出版社,2001年。
- 吴汝纶 1902 《东游丛录·与伊泽修二谈话》,《清末文字改革文集》,文字改革出版社,1958年。
- 文字改革出版社编 1958 《清末文字改革文集》,文字改革出版社。
- 姚亚平 1997 《中国计算语言学》,江西科学技术出版社。
- 中国联合国教科文组织全国委员会秘书处 2001 《我国参加联合国教科文组织重要活动文件汇编》。
- 周庆生 2000 《语言与人类》,中央民族大学出版社。
- 周庆生主编 2001 《国外语言政策与语言规划进程》,语文出版社。
- 朱镕基 2001 《关于国民经济和社会发展第十个五年计划纲要的报告》,人民出版社。

编辑说明

第一届学生计算语言学研讨会 (First Student Workshop on Computational Linguistics, SWCL2002) 是由中国中文信息学会主办的首次以学生为主体的计算语言学学术会议。召开学生计算语言学研讨会的倡议是在第 6 届中国全国计算语言学联合学术会议 (JSCL2001, 太原) 上提出的。北京大学计算语言学研究所请缨承办。

学生是新世纪的生力军, 是未来社会发展的希望, 也是计算语言学研究的希望。这一次学生计算语言学研讨会充分体现了以“学生”为主体的特点, 主要表现在以下几方面:

1. 会议录用的所有论文的第一作者都是在读的学生, 所有报告会、讨论会的主席也都是学生;
2. 会议的所有直接组织者, 包括大会主席、程序委员会、组织委员会的所有成员全部由在读的学生担任;
3. 会议论文的评阅人主体也是学生;
4. 为了给与会的学生提供学习的机会, 会议特别安排了多场专题讲座和一次学术参观活动;
5. 为了给与会的学术提供更多的交流机会, 会议还特别安排了自由讨论、系统演示, 并在会间和会后安排了形式活泼、多样的交流活动;
6. 会议不收任何费用, 而且还对确实有困难的学生提供资助。

可以看到, 这是一次不折不扣的“学生”会议, 我们希望这次会议不仅为大家提供一个学习、交流的场所, 而且希望能够吸引更多的同学投身到这个研究领域中来, 为计算语言学的发展提供源源不断的动力。

这次会议共收到论文 75 篇 (其中中国大陆 72 篇, 台湾地区、新加坡、澳大利亚各 1 篇), 经过评审, 录用论文 72 篇, 其中 51 篇全文录用, 21 篇摘要录用。

录用论文覆盖的研究领域以及分布情况如下:

1. 语法与语义研究: 8 篇 (6 篇全文+2 篇摘要)
2. 词法分析与句法分析: 13 篇 (11 篇全文+2 篇摘要)
3. 信息检索与信息过滤: 9 篇 (5 篇全文+4 篇摘要)
4. 信息提取与问答系统: 6 篇 (5 篇全文+1 篇摘要)
5. 机器翻译: 8 篇 (6 篇全文+2 篇摘要)
6. 术语学与计算词典学: 11 篇 (10 篇全文+1 篇摘要)
7. 语料库语言学: 6 篇 (5 篇全文+1 篇摘要)
8. 语音技术: 6 篇 (1 篇全文+5 篇摘要)
9. 其他: 5 篇 (2 篇全文+3 篇摘要)

会议录用的论文既有一些是博士生多年从事计算语言学研究的的心得, 也有很多是初出茅庐者的习作, 其中不乏包含新想法、新思路、新认识的佳作。读后定会有不少收益和启发。

参加本次会议论文评审工作的, 除程序委员会的成员外, 也包括一些热心关注本次会议的教授、副教授、博士和博士生。每篇论文至少送给两位评审人进行了评审。从返回的评审结果来看, 绝大部分论文评审意见基本上是一致的。对个别评审结果反差较大的稿件以及部分有争议的稿件, 会议又专门邀请相关研究方向上的专家进行了复审。

会议成立专门的优秀论文评审委员会, 将从所有投稿论文中评选出一些优秀论文并颁发奖励, 同时还要推荐给相关的学术刊物发表。

在会议筹备就绪、即将召开的令人激奋之际, 我们学生首先要感谢指导委员会的全体老师, 特别是国家教育部语信司李宇明司长、中文信息学会曹右琦副理事长和北京大学计算语言学研

究所俞士汶教授。学界前辈的大力支持促成了这次学生计算语言学研讨会的成功举行。感谢主办、承办、赞助和支持本次会议的所有单位，他们是本次会议得以成功举办的重要保证。我们还衷心感谢所有参加论文评审以及投稿的老师和同学，正是他们付出的智慧和辛劳保证了本次会议论文的质量。

第一次举办学生计算语言学研讨会，就能够收到这么多的论文，使我们深受鼓舞。从会议网站的留言板上，我们也能够深深感受到同学们对这次会议的热切关心和殷切希望。我们感到责任重大。对我们自己来说，举办这样一次会议，也是一次学习的过程。虽然我们已尽了最大的努力，但是由于学识和经验所限，仍然难免疏忽或不当，希望得到大家的指正和谅解。也希望以后的学生计算语言学研讨会能够继续办下去，越办越好。

我们坚信，计算语言学的明天会更加兴盛、更加辉煌!!

SWCL2002 程序委员会主席 刘群

2002年7月28日

SWCL-2002 组织机构

发起单位：中国中文信息学会

承办单位：北京大学计算语言学研究所

大会主席：陈玉忠（北京大学计算语言学研究所）

副主席：刘 群（中国科学院计算技术研究所）

何婷婷（华中师范大学）

指导委员会

主任：李宇明 教授（国家语委语言文字应用研究所所长）

委员：曹右琦 教授（中国中文信息学会副理事长）

陈群秀 副教授（清华大学计算机系）

姚天顺 教授（东北大学计算机系）

吴立德 教授（复旦大学计算机系）

陈信希 教授（台湾大学）

赖金定 博士（新加坡国立大学）

邹嘉彦 教授（香港城市大学）

俞士汶 教授（北京大学计算语言学研究所）

王厚峰 副教授（北京大学计算语言学研究所）

詹卫东 博士（北京大学中文系）

辛乘胜 总工程师（《人民日报》新闻信息中心）

程序委员会

主席：刘 群（中国科学院计算技术研究所）

副主席：吕雅娟（哈尔滨工业大学）

委员：曲卫民（中国科学院软件研究所）

解国栋（中国科学院自动化研究所）

张克亮（中国科学院声学研究所）

李 芸（北京语言文化大学）

李 鑫（复旦大学）

李保利（北京大学计算语言学研究所）

何婷婷（华中师范大学）

吴云芳（北京大学中文系）

柏晓静（北京大学计算语言学研究所）

组织委员会

主 席：何婷婷（华中师范大学）
副主席：吴云芳（北京大学中文系）
 罗盛芬（清华大学）
委 员：钱揖丽（山西大学）
 陈毅东（厦门大学）
 柏晓静（北京大学计算语言学研究所）
 王树西（中国科学院计算技术研究所）
 刘 扬（北京大学英语系）
 咎红英（北京大学计算语言学研究所）
 苏玉梅（北京大学计算语言学研究所）

优秀论文评选委员会

主 席：李宇明 教授
委 员：曹右琦 教授
 陈群秀 副教授
 俞士汶 教授
 赵 军 副研究员
 邢红兵 副教授
 陈玉忠 博士生，副教授
 刘 群 博士生，副研究员
 吕雅娟 博士生
 何婷婷 博士生，副教授
 吴云芳 博士生
 曲卫民 博士生
 解国栋 博士生，讲师

赞助单位

北京 TRS 信息技术有限公司
东芝（中国）研究开发中心

支持单位

国家自然科学基金委员会信息科学部
全国术语标准化技术委员会
《中文信息学报》编辑部
《语言文字应用》编辑部
微软亚洲研究院

参加审稿的人员名单

(按姓氏拼音排序)

柏晓静, 博士生
常宝宝, 博士, 讲师
陈玉忠, 博士生, 副教授
程 葳, 博士生
池毓焕, 博士生
杜利民, 研究员
何婷婷, 博士生, 副教授
黄萱菁, 博士, 副教授
李保利, 博士生
李素建, 博士生
李 鑫, 博士生
李 颖, 博士生
李 芸, 博士生
刘 群, 博士生, 副研究员
刘 云, 博士生, 讲师
吕雅娟, 博士生
雒自清, 博士生
孟 遥, 博士生
那顺乌日图, 教授
曲卫民, 博士生
宋 柔, 教授
孙 斌, 博士, 讲师
孙广范, 研究员
王 斌, 博士, 副研究员
王厚峰, 博士, 副教授
王强军, 博士生
王树西, 博士生
吴云芳, 博士生
解国栋, 博士生, 讲师
邢红兵, 副教授
杨尔弘, 博士生, 副教授
杨沐昀, 博士生
詹卫东, 博士, 副教授
张克亮, 博士生
张 普, 教授
张 全, 博士, 研究员
张 艳, 博士生
赵 军, 博士, 副研究员
周雅倩, 博士生
宗成庆, 博士, 副研究员

SWCL2002 日程安排(总)

8月20日:

08:00-09:00	注册
09:00-10:10	陈玉忠主持开幕式: 杨芙清院士致辞, 李宇明教授开幕式演说
10:10-10:30	会间休息, 照相
10:30-11:50	大会报告 (A1)
11:50-14:00	招待会
14:00-16:00	分组报告 (B2, C2)
16:00-16:20	会间休息
16:20-18:00	分组报告 (B3, C3)
18:00-19:30	晚餐
19:30-21:30	自由讨论, 系统演示

8月21日:

08:00-12:00	专题讲座 (陆俭明教授, 苏克毅教授, 张景新教授)
12:00-14:00	午餐
14:00-18:00	学术参观 (微软亚洲研究院)
18:00-19:30	晚餐
19:30-21:30	自由讨论, 系统演示

8月22日:

08:00-10:00	分组报告 (B4, C4)
10:00-10:20	会间休息
10:20-12:00	分组报告 (B5, C5)
12:00-14:00	午餐
14:00-15:00	大会报告 (A6)
15:00-15:10	会间休息
15:10-17:10	专题讲座 (王逢鑫教授, 刘群副研究员)
17:10-17:20	会间休息
17:20-18:00	刘群主持闭幕式, 优秀论文颁奖
18:00-19:30	宴会
19:30-21:30	娱乐活动

8月23日上午:

学术参观 (富士通研究开发有限公司) 或自费旅游

SWCL2002 日程安排(分)

场次 A1: 8月20日 10:30—11:50, 主会场

主持人: 何婷婷

10:30—10:50	中文信息处理开放平台的设计 刘群, 张浩, 白硕
10:50—11:10	Finding Names in Chinese Text using a Hybrid Rule Induction Model Jimin Liu, Jing Xiao, Tat-Seng Chua
11:10—11:30	基于双语语料的单个源语词汇和目标语多词单元的对齐 陈博兴, 杜利民
11:30—11:50	Study on Link-Based Approaches for Web IR in TREC Experiments ZHANG Min, MA Shaoping, GAO Jianfeng

场次 B2: 8月20日 14:00—16:00, 第一分会场

主持人: 吴云芳

14:00—14:20	代词“他”的虚指用法及产生原因 ——兼论“V 他+……”结构的特征 张蕾
14:20—14:40	从「悬、挂、吊」看现代汉语近义词的区辨 吴欣达
14:40—15:00	“才”字句的句法语义分析 王楠
15:00—15:20	HNC 理论下对“得”的理解和处理 熊亮
15:20—15:40	动态记忆结构与 HNC 的语境研究 韦向峰
15:40—16:00	基于加权概念网络的用户兴趣建模 许欢庆, 王永成, 孙强

场次 C2: 8月20日 14:00—16:00, 第二分会场

主持人: 李保利

14:00—14:20	结构上下文相关的概率句法分析 张浩, 刘群, 白硕
14:20—14:40	基于混合策略的汉语未登录词整体识别 于传武, 李生, 陈鄞, 赵铁军
14:40—15:00	汉语分词及词性标注自动校验方法研究 钱揖丽, 张虎
15:00—15:20	汉语组合型切分歧义字段消歧方法研究 廉竹钧

15:20—15:40	面向中间语义表示格式的汉语口语解析方法 解国栋, 宗成庆, 徐波
15:40—16:00	基于大规模语料库的英语从句识别 黄玉, 李生, 孟遥, 丁华福

场次 B3: 8 月 20 日 16:20—18:00, 第一分会场

主持人: 李 芸

16:20—16:40	汉英双语短语信息数据库的构建 吴云芳, 常宝宝, 詹卫东
16:40—17:00	迭代策略和词典相结合的机器翻译词典获取 刘晓月, 杨沐昀, 赵铁军
17:00—17:20	基于 TFIDF 的专业领域词汇获取的研究 刘桐菊, 于浩, 杨沐昀
17:20—17:40	因特网语料自动下载分析软件的设计 朱凯, 周杰, 何婷婷
17:40—18:00	《人民日报》1998 年语料库中若干基本语言数据的统计与分析 胡景贺

场次 C3: 8 月 20 日 16:20—18:00, 第二分会场

主持人: 解国栋

16:20—16:40	Experiments on Unsupervised Chinese Word Segmentation and Classification Jin Hu Huang, David Powers
16:40—17:00	中国人名识别中规则抽取的一种基于实例的方法 方晓珊, 盛焕焯
17:00—17:20	词性标注规则的获取和优化 陈文亮, 朱靖波, 吕学强, 姚天顺
17:20—17:40	一个特殊句法结构的英汉句类转换 张克亮
17:40—18:00	语义校对系统中的句子语义骨架 模糊匹配算法 郑逢斌, 姜保庆, 陈志国, 郭拯危, 乔保军

场次 B4: 8 月 22 日 08:00—10:00, 第一分会场

主持人: 张克亮

08:00—08:20	日汉机器翻译系统中的多 Agent 研究 张捷, 陈群秀
08:20—08:40	汉英机译系统 ICENT 中主语省略句的处理 马红妹, 齐璇, 王挺, 陈火旺
08:40—09:00	汉英翻译中主辅变换的类型 雒自清, 张艳红

09:00—09:20	雅信 CAT 和东方快车机器翻译软件的分析及建议 刘彬, 谭外元
09:20—09:40	汉藏翻译系统中的动词处理研究 看卓才旦, 金为勋, 洛智华, 朋毛扎西
09:40—10:00	一种新的针对汉语 TTS 的韵律词自动切分方法 李小明, 易立夫, 杨静, 孙金城

场次 C4: 8 月 22 日 08:00—10:00, 第二分会场

主持人: 李 鑫

08:00—08:20	中文文本聚类的研究与实现 张宝艳, 王庆辉
08:20—08:40	基于遗传算法的定题信息搜索策略 许欢庆, 王永成, 孙强
08:40—09:00	信息过滤技术研究 毛颖, 周源远, 王继成
09:00—09:20	基于兴趣模型的 WEB 信息预测采集过滤方法 李振星, 徐泽平
09:20—09:40	信息技术领域术语提取的初步研究 王强军, 李芸, 张普
09:40—10:00	信息技术领域术语字频、词频及术语长度统计 李芸, 王强军

场次 B5: 8 月 22 日 10:20—12:00, 第一分会场

主持人: 柏晓静

10:20—10:40	语料库的数据管理方式的研究 何婷婷
10:40—11:00	黄曾阳先生语料库思想概述 池毓焕
11:00—11:20	基于数据库的现代汉语新词语缩略语的研究 鲍明凌, 亢世勇
11:20—11:40	基于语料库的现代汉语新词语动词语法特征的研究 颜伟, 亢世勇
11:40—12:00	基于数据库的新造词语的构词法研究 徐艳华, 亢世勇

场次 C5: 8 月 22 日 10:20—12:00, 第二分会场

主持人: 曲卫民

10:20—10:40	基于百科词典的知识获取系统的研究与实现 许勇, 宋柔
-------------	-------------------------------

10:40—11:00	偏重摘要及其实现技术 刘功申, 胡佩华, 岳奕, 王永成
11:00—11:20	红楼梦人物关系问答系统 王树西, 刘群, 白硕
11:20—11:40	一个简单人机对话系统的实现方法 胡凤国
11:40—12:00	一种基于聚类的汉语词语知识的获取方法 李盛, 杨尔弘

场次 A6: 8 月 22 日 14:00—15:00, 主会场

主持人: 吕雅娟

14:00—14:20	基于格助词和接续特征的藏文自动分词方案 陈玉忠, 李保利, 俞士汶, 兰措吉
14:20—14:40	句子相似度计算在 FAQ 中的应用 王洋, 秦兵, 郑实福
14:40—15:00	CCD 构造模型及 VACOL 辅助软件的设计与实现 刘扬, 俞士汶

目 录

语法与语义研究

代词“他”的虚指用法及产生原因——兼论“V他+……”结构的特征	1
张蕾	
从「悬、挂、吊」看现代汉语近义词的区辨	6
吴欣达	
“才”字句的句法语义分析	13
王楠	
HNC 理论下对“得”的理解和处理	23
熊亮	
动态记忆结构与 HNC 的语境研究	29
韦向峰	
基于加权概念网络的用户兴趣建模	35
许欢庆 王永成 孙强	
基于概念的自然语言处理（摘要）	44
马颖华 苏贵洋	
现代汉语“被”字句的 HNC 分析（摘要）	45
柯航	

词法分析与句法分析

结构上下文相关的概率句法分析	46
张浩 刘群 白硕	
基于混合策略的汉语未登录词整体识别	52
于传武 李生 陈鄞 赵铁军	
汉语分词及词性标注自动校验方法研究	59
钱揖丽 张虎	
汉语组合型切分歧义字段消歧方法研究	65
廉竹钧	
面向中间语义表示格式的汉语口语解析方法	71
解国栋 宗成庆 徐波	
基于大规模语料库的英语从句识别	77
黄玉 李生 孟遥 丁华福	
Experiments on Unsupervised Chinese Word Segmentation and Classification	83
Jin Hu Huang David Powers	
中国人名识别中规则抽取的一种基于实例的方法	90
方晓珊 盛焕烨	
基于格助词和接续特征的藏文自动分词方案	98
陈玉忠 李保利 俞士汶 兰措吉	
词性标注规则的获取和优化	106
陈文亮 朱靖波 吕学强 姚天顺	

Finding Names in Chinese Text using a Hybrid Rule Induction Model.....	112
Jimin Liu, Jing Xiao and Tat-Seng Chua	
基于统计和规则的人名识别方法(摘要)	117
吴雪军 朱靖波 陈学耀 卓红霞	
解决汉语自动分词的方法选择(摘要)	118
曲维光	

信息检索与信息过滤

Study on Link-Based Approaches for Web IR	119
ZHANG Min MA Shaoping GAO Jianfeng	
中文文本聚类研究与实现	125
张宝艳 王庆辉	
基于遗传算法的定题信息搜索策略	130
许欢庆 王永成 孙强	
信息过滤技术研究	137
毛颖 周源远 王继成	
基于兴趣模型的WEB信息预测采集过滤方法	143
李振星 徐泽平	
一种基于Web Intelligence的智能信息代理系统初探(摘要)	150
刘洋	
基于英汉词典和翻译软件的跨语言信息检索实验(摘要)	151
王树锋 张永奎	
信息自动获取的结构模型(摘要)	152
苏贵洋 王永成 马颖华	
WWW页面信息中特定内容的过滤研究(摘要)	153
胡熠 郑德权 赵铁军 于浩 王青松	

信息提取与问答系统

基于百科词典的知识获取系统的研究与实现	154
许勇 宋柔	
偏重摘要及其实现技术	161
刘功申 胡佩华 岳奕 王永成	
红楼梦人物关系问答系统	168
王树西 刘群 白硕	
句子相似度计算在FAQ中的应用	175
王洋 秦兵 郑实福	
一个简单人机对话系统的实现方法	182
胡凤国	
面向网络的问答系统研究综述(摘要)	188
张琼 陈群秀	

机器翻译

一个特殊句法结构的英汉句类转换	189
张克亮	
日汉机器翻译系统中的多 Agent 研究	197
张捷 陈群秀	
汉英机译系统 ICENT 中主语省略句的处理	203
马红妹 齐璇 王挺 陈火旺	
汉英翻译中主辅变换的类型	209
雒自清 张艳红	
雅信 CAT 和东方快车机器翻译软件的分析及建议	215
刘彬 谭外元	
汉藏翻译系统中的动词处理研究	222
看卓才旦 金为勋 洛智华 朋毛扎西	
基于 HNC 理论的汉英机器翻译策略研究 (摘要)	231
张克亮	
汉英机器翻译中要素句蜕处理初探 (摘要)	232
李颖	

术语学与计算词典学

基于数据库的现代汉语新词语缩略语的研究	233
鲍明凌 亢世勇	
汉英双语短语信息数据库的构建	240
吴云芳 常宝宝 詹卫东	
信息技术领域术语提取的初步研究	246
王强军 李芸 张普	
CCD 构造模型及 VACOL 辅助软件 的设计与实现	251
刘扬 俞士汶	
迭代策略和词典相结合的机器翻译词典获取	257
刘晓月 杨沐昀 赵铁军	
基于 TFIDF 的专业领域词汇获取的研究	263
刘桐菊 于浩 杨沐昀	
信息技术领域术语字频、词频及术语长度统计	268
李芸 王强军	
一种基于聚类的汉语词语知识的获取方法	275
李盛 杨尔弘	
基于语料库的现代汉语新词语动词语法特征的研究	281
颜伟 亢世勇	
基于数据库的新造词语的构词法研究	286
徐艳华 亢世勇	
试析网络术语及其标准化 (摘要)	292
林纲	

语料库语言学

- 因特网语料自动下载分析软件的设计 293
朱凯 周杰 何婷婷
《人民日报》1998 年语料库中若干基本语言数据的统计与分析 299
胡景贺
语料库的数据管理方式的研究 306
何婷婷
黄曾阳先生语料库思想概述 312
池毓焕
基于双语语料的单个源语词汇和目标语多词单元的对齐 318
陈博兴 杜利民
基于统计的汉英句子对齐研究 328
吕学强 李清隐 任飞亮 姚天顺

语音技术

- 一种新的针对汉语 TTS 的韵律词自动切分方法 329
李小明 易立夫 杨静 孙金城
一种新的基于特殊点的基音提取方法 334
王丽 吕馨 赵铁军 刘占一
自组织混合神经网络在说话人识别中的应用 (摘要) 335
王金甲 王成儒 李静
语音合成技术的现状及发展前景 (摘要) 336
沈晋苑
汉语语音技术 (摘要) 337
郭晨
隐马尔可夫模型实现语音和视频识别 (摘要) 338
林文永 石志国 薛为民 陈锋军

其他

- 中文信息处理开放平台的设计 339
刘群 张浩 白硕
语义校对系统中的句子语义骨架模糊匹配算法 346
郑逢斌 姜保庆 陈志国 郭拯危 乔保军
特定领域中的语义校对系统 (YYJDS) 的设计与实现 (摘要) 351
郑逢斌 姜保庆 郭拯危 邱小平 乔保军
计算语言学领域程序编制的工程问题探讨 (摘要) 352
赵岩 王晓龙 关毅
基于语料库的面向对象网络英语学习系统开发研究 (摘要) 353
尚超