

论用户本位的汉字键盘输入编码设计思想[†]

韩布新 李 娟(中国科学院心理研究所 北京 100101)

摘要 评价汉字键盘输入编码应依据满意原则,而不是最优原则。相应地,其设计思想应从主要依据对汉字属性特征的统计分析的汉字本位,或主要依据计算机软件功能和硬件配置特征的计算机本位,转移至以方便用户使用为指导思想的用户本位。文章据此区分并讨论了键盘输入认知过程分析、心理负荷来源及用户培训等问题。

关键词 汉字键盘输入 编码设计 用户本位

1 前言

汉字输入编码应易学、易记、易用、规范、通用^[1]。第二代输入编码的特点是:以词、句为单位,高度智能化;尽量采用音码输入;具有实时造词、随造随用、词组维护功能;利用高频先见、前后字词相关、字词联想、自动跟踪作者写作习惯等技术智能选词^[2]。随着编码技术的不断成熟,对用户相关指标的考虑显得越来越重要。

早期专业录入员对速度的追求导致了延续至今的思维定势。实际上,以看打或听打为主要方式的专业录入员在用户总体中的比例越来越小,随着计算机的普及,以想打为主要工作方式的一般用户越来越多。用户群体特征的变化对于编码设计思想提出了新的要求。

尽管已有人呼吁汉字编码应“以人为中心”^[3,4],但未对用户工作方式及其相应的需求特点进行认真探讨。近年来,相关研究领域的绝对主题仍然是汉语和计算机软、硬件技术。因此,本文就汉字信息处理系统的三要素(汉语、计算机和用户),探讨汉字键盘输入编码设计中的汉语本位、计算机本位和用户本位思想,及其在当前形式下的相对重要性。

2 汉语本位和计算机本位

2.1 汉语本位

汉语本位思想着重分析汉语材料的形、音、义属性特征,并以此为依据来设计开发汉字输入编码。在开发汉字输入编码的初期,对于汉语特征的统计分析为相关技术的实现、市场开发、用户培训等起到了关键作用。

近年来,对于汉语词属性的研究不断演化,诸如“汉语根词属性库”^[5]等基本建设项目为第二代编码的设计、开发提供了保证。但是,对于汉语认知特点等现代心理学研究成果在编码设计上的应用、英语和汉语和键盘录入的比较工效学研究尚有待进一步开展^[6-8]。

2.2 计算机本位

计算机本位着重考虑如何利用现有的软、硬件技术实现汉字编码输入。计算机技术的发展在处理速度、容量上为高度自适应、智能化汉字输入提供了技术基础,从而分担了过去留给用户的编码记忆负担。然而,计算机软件、硬件的设计逻辑与功能极限能否从根本上适合汉字信息、是否为其留下了足够的空间以应付汉字与英语的不同,仍有待于对汉字计算机的进一步研究。

目前在软、硬件技术上已经完全实现汉字输入,并有新一代思路提出。如:基于现行国际通用标准硬件(如QWERTY小键盘)以及软件平台(如Win95),可开发键入后处理模块,对键盘输入

[†] 本研究获得国家自然科学基金资助(39670255)

代码进行智能化解释,从而减轻用户操作负担^[9];还可利用计算机参与编码设计,即通过建立规范化的汉字特征码信息库,实现自动编码操作。这样既可提高编码效率,又可改善编码性能^[10]。

2.3 汉语本位和计算机本位的局限

根据汉字属性特征和计算机性能,开发汉语输入编码的首要目的是上机实现,其次是各项指标的最优化。单纯以上机实现为目标的编制,常给用户留下了难学、不易用等问题。比如编制形码只要完成分解汉字(按笔划或部件、音元)和布键(根位或音位)两个环节即可达到上机实现^[11];而很多用户宁愿从屏幕选字,也不愿采用为降低重码率而增加的末笔划码等新规则。因为不但要记住规则,而且每次操作时都要根据这个规则和具体文本进行另外一种编码操作。这种双重负担的时间开销和心理负荷比选字上屏的要大得多。

语句级汉语输入比较符合人们的思维和记忆习惯,因其允许连续输入一长串拼音码,再由计算机转换成汉字语句,然而,人可在各类专门和普通知识基础上利用语句流中的语音、语调、停顿、重读等多维信息;计算机则只能利用字、词典等知识库及各种语言推理规则等有限维的信息^[12]。加上自动分词、精确变换等技术的局限,计算机转换的结果仍需要用户检查、确认。

3 用户本位及相关研究

3.1 用户本位的基本含义

用户本位着重分析用户的行为学特征,以方便用户使用作为编码设计的指导原则。这符合良好设计原则^[13]。用户本位的思想基于如下事实:不同用户群的背景、习惯、要求不同;人们不愿意受规则的约束,不喜欢复杂的编码;人类学习、记忆活动遵循认知简约原则;人类行为趋向于减少心理负荷。

决定编码方案性能的三要素分别是:编码特征信息(如字根、笔划、声母、韵母等)、取码规则、特征信息键位分配表。这3个因素都与用户密切相关,编码特征信息决定了基本用户群及其使用情况;取码规则影响学习和掌握的难度;特征信息键位分配表则影响用户接受程度及学习编码的记忆量。

根据用户本位思想,需要进行以下几个方面的研究。一是分析用户的操作特征,以设计更合乎用户习惯的编码方法;二是分析用户的思维、推

理及知识库组织等认知加工过程,为智能化处理技术提供新的思路。三是让用户参与编码设计,以此了解并考虑到不同用户群的特殊需求;四是仔细分析汉字录入编码操作过程中的心理负荷来源,并在编码设计时就提出相应的对策。

3.2 用户心理负荷的来源分析

汉字键盘编码输入操作过程中,用户的心理负荷来源可简单地归属于以下几类。(1)小键盘的工效学合理性和对它的熟悉程度;(2)听打、看打和想打等不同输入情境;(3)汉语水平:对汉字、词的形、音、义规范掌握程度,运用汉语进行思维的能力水平;(4)编码方案的设计要求,比如同时完成对不同层次、不同种类信息的加工;(5)对计算机基础知识及中文平台的熟悉性;(6)过去使用编码习惯的影响,先入为主的情性;(7)计算机焦虑。

3.3 键盘录入的认知过程分析

汉字录入操作过程可简单地做如下划分:读文本、思考或听指令获取加工信息的编码前阶段;对信息进行加工的编码阶段;根据编码击键和选字等编码后阶段。

现有输入法要求人适应编码。比如形码输入过程可简单地分解为:汉语→汉字→拆字(字根)→字根代码+编码规则→汉字编码→汉字。其中除了最后一步用到计算机外,其它步骤全部是用户完成的。用户大脑很忙,而计算机则很闲。通过比较分析英语与汉语文本的阅读和录入过程,可以说明汉语编码输入高心理负荷的本质特征^[14]。

分析键盘录入的认知过程,一个重要任务就是研究编码加工中跨通道转换的负荷问题(比如形托或音托的内部操作过程、时间消耗),进而建立基本数据库,为评价音码/形码/音、形结合码提供必要的量化指标。一个编码方案若既用形码也用音码,则跨通道信息加工使反应时增加;规则增加使记忆负荷增加。虽减少了重码,提高了录入速度,但降低了易学、易记和易用性。

3.4 用户培训

用户对编码的掌握情况无疑将影响输入过程中心理负荷变化和输入效率。一般用户由于不注意自我安排或接受系统培训,此问题比较突出。用户对信息系统的适应性不足而产生的心理张力(计算机焦虑),也会影响其学习和掌握汉字输入编码。

研究汉字录入编码方案的学习过程,旨在结

合记忆组块和记忆广度理论、多重记忆系统理论等现代心理学研究成果,研究短时记忆、情节记忆如何向语义记忆的转化,尤其是编码思维和击键动作的自动化过程(动作记忆的形成)。

3.5 用户群分析

中文输入系统是普及型软件,在设计、开发时要充分考虑到使用者的不同层次、不同背景和共同的典型需求。用户群体可粗分为职业用户、一般用户、新手。用户在不同的使用情境下宜采用不同的编码方法。比如想打以用音码较为合适^[15]。因为,想打必须使输入与思维保持高度一致,才能保证录入行为与思维的效率;想打用户一般是非专业录入员^[16]。

心理学研究表明,类比学习是人类学习新知识的重要途径。因此,新一代中文输入编码应继承上一代系统中已经被用户普遍接受的优点设计,在此基础上发展创新,而不要轻易增加新的规则。系统本身的一套特殊规则,或硬性规定将会使用户望而却步,成为最大的障碍。

3.6 用户本位设计思想的主要优势

若从用户角度来考虑录入行为,将会使纯技术角度的分析建立在实用基础之上。实际上很多技术难题并非想象的那样不可接受,只要仔细分析录入行为过程就可以找到解决办法。对键盘本身的熟悉性及操作熟练程度所带来的问题也常常被不适当地归咎于汉字编码。

4 讨论

4.1 有限合理性理论

美国著名学者、信息加工心理学创始人之一 Herbert Simon 教授曾在经济领域提出了著名的有限合理性理论,认为经济组织的决策应遵循满意原则,而不是最优原则^[17]。该理论是 Simon 教授获得 1978 年经济学诺贝尔奖的主要基础。我们今天来思考汉字输入编码所面临的问题正可以根据有限合理性理论来寻找突破口。

4.2 最优原则与满意原则

汉字本位和计算机本位设计思想的深层意识正是寻求最优,只有用户本位才适用于满意原则。用有限合理性理论指导汉字键盘输入编码设计,可使我们摆脱汉字本位和计算机本位所带来的最优桎梏,将视线转移到用户满意这一根本性的指标上,从而冷静地分析目前我们所面临的问题并思考对策。

追求最优的结果使汉字输入编码设计研究徘徊

在上机实现阶段,而不能进入用户选择阶段。从优化角度来看,这已成为汉字编码技术发展的阻碍。作为产品,汉字编码技术的卖方市场已经相当成熟,应尽早培养买方市场。否则,难以摆脱思维定势,难以找到突破点。由于各项理想指标之间有矛盾(比如笔顺规范和部首规范就有不一致的地方),用户往往必须在这几点上进行抉择或权衡^[18]。所以,不存在绝对的最优。奔腾的“万码”各有其存在的合理性(或在某一指标上的突出表现),又怎能用单一的指标将它们排出高低来?

汉字输入应首先允许人在已有的知识背景下能够实现输入,然后逐渐趋于标准化、规范化,同时承担了一定量的记忆负荷后便可以达到快速输入的需求。因此,在最佳、满意和可接受等不同的标准下,易学、易记、易用、规范和通用指标便有了相对不同的重要性。只有从思想上将着眼点转移至用户身上,将研究重点转移至用户的使用过程中来,我们才能真正明白用户到底需要什么样的编码,才能认真思考我们能搞出什么样的编码以满足社会大众的需求,才能开发出易学、易用的编码,才能使我们的编码有市场。

4.3 3种本位的关系

3种本位实际上代表了可以区分汉字键盘输入的编码设计思想不同维度。在汉语信息处理系统的3个构成成分(汉语、计算机、用户)中,用户应该是核心,是最终评价标准。因此,3种本位取向的区别实际上就是对这3个成分的基本着重点的区别。

3种本位的思想明确了汉字键盘输入编码的发展方向,并可指导更客观地建立定性或定量评价编码方案的指标体系。据此开展对现有输入方法的市场调查、用户分析和评价、工效学分析。这3个方面是解决瓶颈难题的关键,而其核心则是用户的接受度(acceptability)、评价(evaluation)以及编码方案的可用性(usability)。

4.4 汉字键盘编码输入的发展

汉字键盘编码输入已从字、词输入阶段进入句输入阶段,表现出强大的生命力。手写输入、语音输入、扫描输入等在用户本位(尤其是心理负荷量)方面有其先天的优势,但因技术发展方面的局限和方块汉字在教、学、写、看等方面的特点,不可能完全取代键盘编码输入。智能化的小键盘输入编码必然从以码为中心走向以人为中心。在买方市场不断成熟的情况下,以用户为中心的用户本

动晕病机理及预防研究[†]

卢 岚 秦娟娟(天津大学管理学院 300072)

交通工具是现代生活中不可缺少的,不少人乘坐交通工具时经常出现不舒服的症状,如晕眩、恶心、呕吐等。在运动生理学中称这种现象为动晕病,也叫运动病。动晕病的病因是比较复杂的,不仅有生理因素和心理因素,还受个体敏感度的影响,为此,我们于1996年~1997年进行了相关的研究。我们分别对青年和少年两组被试者进行

了动晕病经历的问卷调查,其中青年组45人,全部为天津大学本科生,他们的年龄分布为19~23岁,男生30人,女15人。少年组10人,全部为天津南开中学初中一年级学生,年龄均为13岁,其中男生、女生各5人。

问卷针对5种不同的交通工具——公共汽车、出租车、轮船、火车、飞机,由5个小问题组成。

(1)在过去的一年中,你乘坐过多少次下述交通工具?

[†] 本研究项目与香港科技大学合作进行

位设计思想将是编码取得成功的决定性因素。

5 结束语

汉字输入编码设计涉及的问题较多。对这些问题应区别对待,不能混为一谈。目前的当务之急,是将编码设计思路从汉语本位、计算机本位转移到用户本位上来,以满意原则取代最优原则。

参 考 文 献

- 1 范志强. 一种新的汉字编码优化设计模式. 中文信息学报, 1997, 11(2): 63~ 98
- 2 冯志伟. 汉字的极限熵. 中文信息, 1996, (2): 53~ 56
- 3 章剑卫, 周其力. “通人性”的智能输入将促进电脑大普及. 中文信息, 1996, (4): 57~ 58
- 4 余克艰. 汉字输入的大趋势——从“以码为中心”走向“以人为中心”. 中文信息, 1996, (5): 5~ 8
- 5 叶晓石. 以人为中心的字形图形有效码会使中文软件设计突破性跃进. 中文信息, 1997, (2): 3~ 4
- 6 盛玉麒. 把词语处理与智能化联系起来——信息处理用汉语根词属性库的研究与建立. 中文信息, 1996, (4): 35~ 37
- 7 陈一凡, 胡宣华. 汉字键盘输入技术与理论基础. 清华大学出版社、广西科学技术出版社, 1994
- 8 韩布新. 汉字识别中部件组合频率的作用. 心理学报, 1996, 28(3): 9~ 13
- 9 韩布新. 论心理部件、编码部件与统计部件集合. 计算

机世界, 1997, 11

- 9 姜 珊. 汉字输入的结构层次. 中国中文信息学会汉字编码专业委员会(第七届)、中国计算机学会中文信息技术专业委员会(第五出)学术交流会, 1996, 210~ 214
- 10 陈玉龙. 中文自动编码原理. 中文信息学报, 1998, 12(1): 26~ 34
- 11 蒋辅文. 统一汉字形码便于国际信息交流. 中文信息, 1997, (1): 52~ 54
- 12 章 森, 守成庆, 孙建军. 新一代中文输入系统面临的问题. 中文信息, 1997, (1): 3~ 6
- 13 罗闻泉, 史晓强. 人机界面设计的规范化方法. 中文信息, 1997, (2): 15~ 18
- 14 韩布新, 林仲贤. 论汉字部件与英文字母的认知对应关系. 心理科学, 1995, 18(6): 334~ 336
- 15 王 苏. 思维与语言. 见: 潘 菽, 荆其诚主编. 中国大百科全书·心理学卷. 中国大百科全书出版社, 1991: 356
- 16 严戎庚. “想打”输入需要的汉字编码. 中文信息, 1996, (4): 42~ 44
- 17 荆其诚主编. 简明心理学百科全书. 长沙: 湖南教育出版社, 1991: 532
- 18 周 宪. 再论对“认知码”的重码统计. 中文信息, 1996, (2): 70~ 71

[收稿日期 1998-10-20]