

数量、程度的现代汉语言描述分析

苏朋程

(贵州师范大学, 贵州 贵阳 550001)

摘 要:语言是客观事物、现象、人类思想感情的符号系统再现,然而自然界丰富多彩的动物、植物、微生物、矿物的属性,自然现象的特征,以及人们通过感觉器官对物的视觉、听觉、嗅觉等对物感觉的复杂性,还有人类复杂多变的心情、情感等,这些因素是人类对自然界、经济社会、现象的感知、认知的方式和主要内容,它们的语言描述形式与其自身存在对应关系。文章结合物理学的相对理论、科学符号系统,分析了物质的属性、人类对物质属性的感知、人类情感等数量或程度的现代汉语言描述形式,认为现代汉语言不仅具有模糊性,而且还存在不连续性、相对性。文章最后结合实例,分析了现代汉语言的模糊性、不连续性、相对性。

关键词:科学符号系统;现代汉语言数量、程度描述;不连续性;相对性

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2013)06-0094-04

doi:10.3969/j.issn.1673-629X.2013.06.024

Analysis on Modern Chinese Language Description of Quantity and Degree

SU Peng-cheng

(Guizhou Normal University, Guiyang 550001, China)

Abstract: Language is the symbol system of reproduction which is about the objective things, phenomenon, human thoughts and feelings, etc, but the variety of animal, plant, microbial, mineral properties, natural phenomena, and through sensory organs on the visual, auditory, olfactory and the complex feeling of things, and the complexity of the mood, emotion and so on, these factors are the mankind to nature, society, economy phenomenon perception, cognitive pattern and the main content, their language description form and their own exist corresponding relationship. Combining with the relative theory of physics, scientific notation system, analyze the quantity or degree of modern Chinese language description form, such as material properties, the material properties of human perception, human emotions. In this way, the conclusion is that modern Chinese language is not only obscure, but also discontinuity, relativity. Finally, as far as application examples concerned, analyze the fuzziness, discontinuity and relativity of modern Chinese language.

Key words: scientific notation system; quantity or degree of modern Chinese language description; discontinuity; relativity

0 引言

自然语言包括人类交流用语、科学符号系统,自然语言是人类对客观事物的主观表达或描述,采用交流用语进行量化、程度描述时存在的不足或缺陷问题,可以用科学符号系统予以解决。对客观事物进行描述时,是离不开量化描述的,文献[1]研究了时间名词,当然,还有其它用于描述数量、程度的名词值得研究,例如语句“这里有三只老鼠”,“三只”就是量化描述词语,此类量化描述是确定的、肯定的;有的量化描述是不确定的,即模糊的,如语句“我明天要来学校”,到底是哪一天来学校呢,却不是肯定的,“明天”这个表示时间量的词语是模糊的;还有的量化描述是成对出现

的,而且彼此表达的概念是相对的,它们可分为两类,一类是描述客观事物自身性质的,例如词语“高-矮”、“胖-瘦”、“快-慢”、“深-浅”等;另外一类是描述人类对客观事物的主观感觉程度方面的,如“高兴-悲哀”、“美-丑”、“好-坏”、“简单-困难”、“灵巧-笨拙”等等,类似的还有很多,限于篇幅,在此不一一列举出。由于客观事物的属性是连续的、封闭的,而人类交流用语对其描述做不到一一对应,因而造成模糊性,例如描述物体运动的状态,现代汉语言只能用“快”和“慢”两个词来描述,然而用科学符号系统里的物理量——速度来描述,就精确得多了。因而,客观事物及现象量化的现代汉语言描述存在模糊性,需要解决模糊性,语言

收稿日期:2012-09-06

修回日期:2012-12-11

网络出版时间:2013-03-05

基金项目:贵州省科技基金项目(黔科合J字[2011]2216)

作者简介:苏朋程(1976-),男,硕士,助理研究员,从事自然语言处理、信息管理、信息安全研究。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20130305.0815.016.html>

模糊性的根源是什么呢?文献[2]认为自然语言模糊性的根源之一是客观世界的复杂性,但是没有具体列出哪些内容的描述是模糊的。文中深入分析了客观事物在数量、程度方面的现代汉语言描述形式、存在的问题,以及解决方法,是对语言模糊性的补充研究。

1 数量、程度的分类及类型

语言是人类思维的载体,是对客观事物、现象、感觉、人类思想情感的描述,因此,可以认为“语言是客观事物、现象、人类的思想感情的符号系统再现”,语言与语言符号存在对应关系,正如我国中科院声学所首创的概念层次网络(HNC)语言理解技术一样,该理论认为自然语言理解是一个从自然语言空间到语言概念空间的映射过程^[3],将自然语言抽象为概念,然后用形式化的符号描述概念。客观事物种类繁多,比如说有动物、植物、微生物、矿物,它们存在的形态也是各有千秋,有的是可数的、有的是不可数的,有的是凭人的感官可感知到的、有的则必须借助科学仪器才能感知到;自然现象和社会现象层出不穷,现象是一个动态过程,即从产生到发展、终止,它是一个连续的过程;感觉是由物质的外观、颜色、温度、气味、硬度等对人类的感官刺激后形成的生物反应,此类反应是人类源自于自然物质的物理属性,这些属性是人类的肉眼无法观察到的,可以采取科学符号系统、人类交流用语表示程度的词语进行描述;人类的思想感情复杂多变,是对人、事物、现象的主观反映,是抽象的形式,它们可通过人的眼神、面部表情、谈话等方式间接表现出来,可以采用程度方面的语言描述。

基于上述人类对客观事物、现象的感知、认知形式,以及人类的感觉、思想情感的表现形式,可将那些数量、程度方面的现代汉语言描述形式归纳如下:

I. 数的可数形式。如一、二、三、…、一百、一千、一万、十万、百万、千万、亿、十亿、百亿、千亿、万亿……这样的数是可数的,而且是无限的。

II. 数的不可数形式。二分之一、三分之一、四分之一、五分之二、…、零点一二、零点三六,等等。

III. 物的属性的现代汉语言描述形式。对物体几何尺寸的描述“大-小”、对物体温度的表述“冷-热”、对物体速度的描述“快-慢”、对物体硬度的描述“软-硬”、对物体气味的描述“香-臭”、对食物味道的描述“美味可口-难以下咽”、对声音的描述“清脆-浑厚”、对光线的描述“明亮-昏暗”,等等。这些现代汉语言对物的感觉的程度描述形式,从中可看出,经过人类的视觉、听觉、触觉、嗅觉、味觉等感觉器官对物质的物理性质如几何尺寸、温度、速度、硬度、声音、味道、光的亮度的描述,加上修饰词语,现代汉语言只能用有限可数

的词语来描述,如“冷-有点冷-非常冷-不是一般的冷”,这样的描述对于物体的实际温度是无法一一对应的,但是可以采取科学符号系统予以精确描述,如温度℃或热力学温度 K,所表达温度的范围及精确度比交流用语大得多、精确得多。大多数物质的物理属性可以用科学符号描述,然而对于物体气味、味道的描述,到目前为止,科学符号系统尚不能描述。

IV. 人类主观感觉、情感程度描述形式。此类主观感觉、情感程度的描述,“好-坏”、“高兴-悲哀”、“脆弱-坚强”、“美-丑”、“善良-奸诈”、“友好-敌对”、“坚决-犹豫不决”、“忠诚-背叛”、“支持-反对”、“困难-简单”、“灵巧-笨拙”,等等。这些成对出现且彼此对立的反义词,它们所描述的人类主观感觉、情感程度比较模糊,必须以一定的参考标准才成立,这些对立程度有一个临界点,例如保持中立,既不友好也不敌对,既不支持也不反对,除此以外,找不到具体描述此类细微区别程度的词语,也就是语言描述形式是不连续的。目前,尚无描述此类人类主观感觉、情感程度的科学符号系统。

2 数量、程度的现代汉语言描述形式的特性

2.1 模糊性问题

关于模糊性,国内,曾广平,涂序彦等人建立了云模型,分析自然语言的模糊多属性问题^[4],国外研究人员作了相关研究^[5-7]。导致现代汉语言描述的模糊性的因素是多方面的,一方面是客观事物丰富多彩,它们所对应人类的语言描述形式无法与其自身一一对应,因为人类的语言空间是有限的,而客观事物则相互关联呈链式增长趋势。现代汉语言量化描述的模糊性表现在量的不确定性、不精确性,但可以用一个范围来表示,例如语句“三天左右”、“一个月内”。由于语言的模糊性,会使得人们在决策、判断时存在难度,该如何解决呢?对于语言模糊性的量化处理方面,国家社科基金资助项目(06BJY006)阶段性成果《基于模糊语言的科研基金项目立项评估研究》,针对科研项目评审结果的“优”、“良”、“差”等模糊语言,利用三角模糊数排序公式对“优”、“良”、“差”进行量化处理后,在其设计的家模糊语言评估矩阵进行评估优选立项^[8]。国家自然科学基金资助项目(70873058)的阶段性成果^[9]将语言变量转化为梯形模糊数 $\tilde{A}=(a,b,c,d)$,然后进行量化处理。这些科研成果深入研究了语言模糊性的量化处理技术,是针对语言描述的多因素、多指标的处理,是对语言应用的处理研究,但不是语言的描述形式与客观事物属性的对应关系的研究。对于如“高-矮”、“硬-软”等成对出现的反义词,其理论基础是什

么呢?对于该问题,至今无系统、权威的相关研究。文中重点研究客观事物的属性,以及人类对其属性感知、认知的数量、程度的现代汉语言描述形式。

2.2 不连续性问题

(1)问题的提出。

首先,人类的触觉、嗅觉、听觉等感觉器官对客观事物等刺激形成感知,接着经过人脑的思维形成语言概念,然后将这些语言概念描述成语言符号的形式,这样,语言概念与语言符号就存在映射关系,我国中科院声学所的黄曾阳研究员等人认为语言概念是唯一的,语言符号空间却不唯一^[3,10,11],再根据客观事物的客观实在性,于是会出现同一客观事物具有不同的语言符号描述形式。人们将自己对客观事物形成的概念描述成语言,其语言形式存在不连续性问题,特别是人们将自己对客观事物的性质、特点、属性,以及人类对其形成的感觉等,将它们的变化趋势、程度等描述成交流用语的数量、程度的形式描述,比如说现代汉语言量化形式描述,就存在不连续性。例如在描述人的胖与瘦,现代汉语言描述就只能用“胖”、“很胖”、“瘦”、“很瘦”等有限的词来描述,或者可以用语句“胖得只能躺在床上”、“瘦如干材”来描述,这样的描述还是不够的,仍然不知道胖瘦的程度。类似的还有,“高-矮”、“长-短”、“轻-重”、“冷-热”、“硬-软”、“大-小”、“快-慢”等,它们有一个共同的特点,那就是这些字、词是关于物质的物理性质或个体的长短、高度等。造成人类交流用语量化描述时出现的不连续性问题,其原因主要是物质的物理性质是无限不可数的、连续的、封闭的,而且物理性质是物质运动所表现出来的,例如温度,温度是构成物质分子的运动剧烈程度所表现出来的,到底有多少种温度值,无人能回答,然而人类交流用语却是可数的,可数表现在人类对客观事物及其属性的描述是具体文字表现的、是可计数的。因此,认为是人类交流用语描述的可数性导致其描述客观事物的物理属性时,所表现出不连续性,它们之间不是一一对应的关系。

(2)问题解决方案及技术。

上述现代汉语言描述的量的形式,它们是经过人类的感官器官所看见的,如“高-矮”、“长-短”、“大-小”、“胖-瘦”等;触觉感觉到的,如“冷-热”、“硬-软”等;听觉听到的,如用“清脆-浑厚”等词语来描述声音,当然有的声音是人类的耳朵无法听见的。对于现代汉语言无法精确描述的客观事物的属性及程度等问题,可以采用科学符号系统予以解决,如“胖-瘦”的程度问题,有专家提出将人的身高与体重的比值作为一个衡量胖瘦的标准,于是可以区别是胖、还是瘦。就“胖-瘦”的程度而言,如果仅仅是采用现代汉语言描

述,依然无法辨别胖瘦的程度,倘若采用身高体重比,通过该比值,就可以知道胖与瘦的细微区别了。由此,可以看出,对于胖与瘦,现代汉语言无法描述其细微变化程度,但通过身高体重比的数值,由于数值是连续的,则可知其细微变化程度。对于上述物的属性,它们可以被科学符号——物理量进行精确描述,例如,“高-矮”和“长-短”,可以用测量设备或仪器如卷尺、螺旋测微器、游标卡尺、电子显微镜等进行测量,可以以公里、米、毫米、微米为计量单位,甚至可以精确到纳米等不同量级的、连续的描述,“轻-重”程度可以采用重力来计量,“冷-热”程度可以用温度来描述,“硬-软”可用硬度表示法,“大-小”可用体积来描述,“快-慢”可以采用速度来描述,物理表达式为: $v = \frac{s}{t}$,其中,符号 v 表示速度、 s 表示距离、 t 表示时间,于是,描述“快-慢”可用概念“单位时间内物体移动的距离”,距离和时间均采用统一的国际单位米、秒,那么就可以精确描述“快-慢”的程度了。对于描述声音的词语“清脆”、“浑厚”,可以用频率来表示,即单位时间内完成的震动次数来表述,声音的实质就是由震动发出的,衡量标准是震动的快慢——频率(Hz)。

2.3 相对性问题

(1)相对性原理。

相对性最初是在物理学领域被提出的,运动力学描述的相对性为“必须要有一个参照物或者参照系,即一个物体参照另外一个物体位置的改变”。伽利略提出的相对性原理“在一个惯性系内部所作的任何力学实验都不能够确定该惯性系本身是在静止状态,还是在作匀速直线运动”,其实质是惯性系的状态不能由自身所表现出来的状态所确定^[12];爱因斯坦在相对论中提出一个关于时间的膨胀效应,他指出“时间在不同的惯性系之间是相对存在的,而不是绝对的”^[12]。简洁归纳概括并推广,认为:事物的属性、状态不能由自身的属性、状态决定,而是必须以另外的事物或系统作为参照才能确定,而且这些属性、状态在不同的参照物、参照系,相应具有不同的属性值、状态值,参照物主要指实体物,参照系是指物质、现象等所呈现的属性、状态,以及人类对其评价标准和体系等。

(2)现代汉语言描述的相对性。

先来看一些字、词,特别是成对的反义词,“高-矮”、“长-短”、“轻-重”、“冷-热”、“硬-软”、“大-小”、“快-慢”、“简单-困难”、“有-无”、“热情-冷漠”、“高兴-悲哀”、“兴旺-衰落”、“好-坏”、“美-丑”,等等,这些字、词当中,可分为两类,一类是描述客观事物的性质或状态,如“高-矮”、“长-短”等,另一类是人类对客观事物的感受,如“冷-热”、“高兴-悲

衰”、“兴旺-衰落”、“好-坏”、“美-丑”等。这些字、词是成对出现的,表达的概念是相对的,“高”这个字所描述的概念必须以“矮”这个字所描述的概念为参照,否则就不成立,同样“矮”这个字所描述的概念也必须以“高”这个字所描述的概念为参照,否则它所描述的概念同样不成立,由此可看出,现代汉语言描述的相对性在于两个相对立概念之间彼此互为参照才能表达它们所要描述的概念,跟物理学的相对论类似,比如说一个人在工作单位来说身高是最高的,当他到其它地方后就不一定是最高的,因为他所处的参照系发生了改变。语言描述的相对性是很好理解的,如果每个人都一样高、山峰也是一样高,世界上的万物在几何尺寸、形状都无差别,那也不会出现“高-矮”、“大-小”等这样成对的现代汉语言文字。其它的成对反义词类似。

现代汉语言描述的概念的相对性是一对相互对立的,彼此对立的一方各自向相反方向无限延伸,无限接近所要表达的概念,即只是一个趋势。就像物理科学研究的绝对零度非有限次试验可接近、爱因斯坦论证光的速度是速度极限一样的道理,现代汉语言描述的相对性同样是无限接近,例如“高”与“矮”这一对字所描述的概念,就这两个字所对应的自然物是不存在的,只能是一个趋势,因不同的参考物、参照系发生改变而改变,因此“高-矮”所描述的相对性概念是一对相互对立的,对立的一方是无限接近的、渐进的概念。其它成对的反义词也是如此。到底这样成对的反义词有多少呢,目前尚无完全统计,总之,用现代汉语言描述客观事物的属性、状态时,必须要有一个参照物、参考系、参考点,才能加以描述。为便于形象直观地理解现代汉语言描述的相对性,见图1,该图描述了现代汉语言描述的相对性。



图1 现代汉语言描述的相对性

3 应用分析

综上所述,由于数量、程度的汉语言描述存在模糊性、不连续性、相对性等问题,因此人们在现实生活、工作、经济活动中,交流双方对同一问题、事物的理解难免不存在偏差,甚至会导致经济纠纷、冲突等。模糊性方面,例如在签订合同的落款时间就不用“三天左右”、“一周以内”等模糊的词语。不连续性方面,如“滴水油”、“地沟油”,这两个词语所描述的油的本质是同一的,对于这些油,它们具有不同的成分、其来源地也不一样。对于此类油,汉语言的描述是有限的、不

连续的,然而其油的构成成分却是连续的、无限的。相对性方面,在评价客观事物时,对于同一事物,有的人评价为“好”,有的人则评价为“坏”,导致“好”与“坏”的差别在于不同的人具有不同的评价标准,也就是参照系(标准)不一样。


4 结束语

自然界的动物、植物、微生物、矿物等,种类繁多,它们表现为不同的物理属性、化学属性,它们有不同的化学成分、组织构成、自然现象及社会现象、人类对自身的评价、人类的情感、态度等主观因素,以及人类可用语言描述的一切,这些自然界及主观世界与其人类的语言描述形式存在对应关系,在描述形式上,离不开数量、程度的语言形式化描述,由于自然界的无限性、连续性、封闭性,上述物的属性、人类感觉器官对事物的感觉等所反映出来的数量、程度的汉语言描述形式,两者之间无法做到一一对应,只能是对其近似的描述。

参考文献:

- [1] 易帅,赵克,韦卓.领域自然语言理解篇章中时间名词的研究[J].计算机技术与发展,2008,18(6):5-8.
- [2] 张蔚,杜晔.自然语言模糊性的哲学思考[J].社会科学家,2009(3):154-155.
- [3] 黄曾阳. HNC(概念层次网络)理论[M].北京:清华大学出版社,1998.
- [4] 吴爱燕,于重重,曾广平,等.基于自然语言的模糊多属性云决策方法研究[J].计算机科学,2010,37(11):199-202.
- [5] Alonso S. Group decision making with incomplete fuzzy linguistic preference relations[J]. International Journal of Intelligent Systems, 2009, 24(2):201-222.
- [6] Lee S, Huh Y. Automatic generation of concept hierarchy using WordNet[J]. Expert Systems with Applications, 2008, 35(3):1132-1144.
- [7] Francisco M. An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context[J]. IEEE Transactions on Fuzzy Systems, 2009, 17(2):279-290.
- [8] 王雄,吴庆田.基于模糊语言的科研项目立项评估研究[J].科技进步与对策,2007,24(9):61-63.
- [9] 高岩,周德群,章玲,等.模糊语言群决策影响因素的分析方法[J].武汉理工大学学报(信息与管理工程版),2011,33(4):613-617.
- [10] 晋耀红. HNC(概念层次网络)语言理解技术及其应用[M].北京:科学出版社,2006.
- [11] 黄曾阳.语境表示式与记忆[J].云南师范大学学报(哲学社会科学版),2010,42(4):7-14.
- [12] 舒幼生.力学(物理类)[M].北京:北京大学出版社,2011.

数量、程度的现代汉语言描述分析

作者: 苏朋程, [SU Peng-cheng](#)
作者单位: [贵州师范大学, 贵州贵阳, 550001](#)
刊名: [计算机技术与发展](#) 
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2013, 23(6)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjtz201306024.aspx