

模糊语言学综论

伍铁平

[北京师范大学中文系 教授 北京市 100088]

提要: 模糊语言学是运用模糊理论研究自然语言中模糊现象的学科。近一二十年,国内学人发表了不少有关模糊语言的文章,但有的人常将概括性、多义性、歧义性等概念与模糊性混为一谈,那是不对的,因为模糊性实则是指对象类属边界不确定的特性。模糊语言在现实生活中具有很大的意义,它已发展为认知语言学的重要组成部分,后者又和认知心理学同为当代科学关注的一个中心,因为这方面的研究将有助于理解人的思维过程,对人工智能的发展也大有裨益。

关键词: 模糊性 认识 发展

中国图书资料分类法分类号: H0

现在家用模糊电器风靡全球。电视机的亮度、清晰度、响度,电冰箱、电饭煲、电熨斗、空调器等的温度,洗衣机中衣服的清洁度都是模糊概念。只有用模糊数学对这些电器进行模糊处理,它们的多种模糊度才能得到最佳的控制,才能最节省能源。但是使用模糊电器的人大概很少有人知道,模糊之被人们所承认,经历了一个漫长的历史过程。在西方最早注意到模糊现象(当时没有用模糊这个词)的是生于公元前4世纪的尤布利德斯。他提出了著名的“连锁推理悖论”:麦堆是一个模糊概念,一粒麦子自然不能构成一个麦堆,但是要多少粒麦子才能构成一个麦堆,这是谁也无法判断的事情;因为数字是精确概念,而麦堆却是模糊概念,用精确概念衡量模糊概念,自然是方枘圆凿,格格不入,构成悖论。古老的秃子悖论的性质同麦堆悖论类似。英国著名哲学家罗素在其名作《论模糊性》中说:“假定一个人原来不是秃子。他的头发一根根地脱落,最后才变成秃子。于是有人争辩说,一定有一根头发,由于这根头发的脱落,便使他变成秃子。这种说法自然是荒唐的。秃头是一个模糊概念;有一些人肯定是秃子,有一些人肯定不是秃子;而处于这两者之间的一些人,说他们必须要么是秃子,要么不是,这是不对的。排中律用于精确符号时是正确的;但是当符号是模糊的时候,排中律就不合适了。”^[1]排中律是形式逻辑的一条基本规律,其基本内容是两个相互矛盾的思想,不能同时都假,必须有一真,没有第三种可能性。例如一个数只能或者是正数(甲),或者是负数(乙),这一对矛盾判断中,必有一

真一假,不存在第三种可能性。模糊概念则恰好是既可隶属于甲,又可隶属于乙,如头发脱落很多的人既可能接近秃子,也可能不上秃子,这是对排中律的突破。

恩格斯在《自然辩证法》中说过:“辩证法不知道什么绝对分明的和固定不变的界限,不知道什么无条件的普遍有效的‘非此即彼!’,它使固定的形而上学的差异互相过渡,除了‘非此即彼!’又在适当的地方承认‘亦此亦彼!’,并且使对立互为中介。”^[2]这是辩证逻辑对形式逻辑的排中律的最好的批评,是模糊学的理论基础。

麦堆和秃子这两个例子都是对模糊概念的最好的说明。表示模糊概念的词是模糊词,语言中有不少模糊词。研究模糊词和语言中的其他模糊现象的学科叫模糊语言学。

表示时间的许多词(如早、晚等),表示年龄的词(如青年、中年、老年等),表示颜色的词(如红、绿、蓝等),表示味觉的词(如酸、甜、苦、辣等),表示温度的词(如冷、凉、温、热、烫等)都是模糊词。模糊词有一个特点,在模糊词的连续统(continuum)的中间部分,很少模糊性;模糊性主要出现在连续统的两端的边缘部分(认知语言学将这种边缘现象叫做 border line cases)。例如, $18^{\pm} \sim 30^{\pm}$ 岁是典型的青年, $15^{\pm} \sim 18^{\pm}$ 岁介乎少年与青年之间, $30^{\pm} \sim 40^{\pm}$ 岁介乎中年与壮年之间。随着人们平均寿命的延长,这几个模糊词的上下限可能发生变动。例如在旧社会,年近半百的人往往被视为老年,而现在则往往被看作中年,据《湖北

日报》1995年5月17日报道,1991年世界卫生组织规定:“44岁以下的人为青年,45~49岁的人为中年。”处于中间部分的词被现代认知语言学(cognitive linguistics)称作“原型”(prototype,或译为“典型”),即它们具有青年的特征最多:如身体健壮、精力充沛、思想活跃等等。

一对模糊概念之间常常存在一些介乎两可之间的现象。例如在水果和蔬菜之间,苹果和梨等可称为典型的水果,无花果就不那么典型,西红柿则介乎水果与蔬菜之间。喜鹊、燕子和麻雀等是典型的鸟,企鹅属于鸟的隶属度却只有70%左右。求属度(membership)是模糊数学的创始人扎德提出的术语,指隶属于某一模糊集(fuzzy set)的程度。不过也有人对此提出疑问:如果企鹅只有70%[±]属于鸟,那30%又属于什么呢?而且这百分比中的数字也只是大致的推断,并没有绝对可靠的客观依据。事实上,企鹅在中外词典中仍被称作“海鸟”或“水鸟”。

德国哲学、逻辑和数学家弗雷格曾认为,模糊现象既然违反了排中律,就应视为语言的缺陷,应尽力避免;但是奥地利哲学家路德维希·维特根斯坦则相反,认为模糊性是语言的优点。模糊性究竟是存在于客观现实之中,还是仅存在于人们的意识之中,在西方也曾有过激烈的争论。事实上,语言本身就是一种客观存在的现象,语言中有大量模糊现象,其边界不可能像刀切那样,截然分明。例如邻近方之间、词类之间、词素与词之间、词与短语之间的界限都不是那样容易分辨的。对于词类,有些认知语言学家试图采用维特根斯坦所提出的家族类似(family resemblance)的原则来处理。一个家族的人中有的比较多地像父亲,有的比较多地像母亲,或像父母的程度大致相等,有的比较多地像祖父,有的比较多地像祖母,有的各继承了他们的部分特征;除开双胞胎外,家族中完全相像的人是几乎没有的,但是这些相似的模糊性质并没有影响他们属于同一家族。汉语的形容词中有些比较接近动词,有些则否;汉语的名词有些接近形容词,有些则否;用传统的一刀切的方法区分词类,往往会遇到许多困难。看到词类的模糊特征,用家族类似的原则处理,也许比较符合实际。^[3]英国1985年出版的利奇等四位名教授主编的《当代英语详解语法》彻底改造了传统语法的格局,将模糊性(他们称为“渐变性”)的理论贯穿于许多章节,全面地研究了英语语法中的许多模糊现象。然而,令人感到不解的是有个语文刊物竟发表有违常理的文章,将“少年”“青年”“老年”这些语言中客观存在的模糊词说成

是现行词典“人为规定不精确造成的”结果。由此可见,为了革新我国的语言学概论,乃至整个语言学各门课程的教学内容,密切关注并吸收国外有益的学术研究成果是十分必要的。

我国发表了不少关于模糊语言的文章,但是有些人往往将概括、多义、歧义、双关、含混和话语未了都归属于科学意义上的模糊现象,那是不对的。如人的国籍、民族属性、年龄、性别、高矮迥异,但都可以叫做“人”,这是人的概括性^[4],而非模糊性;但是在人的形成过程中,人和猿人的界限不易一刀两断,这却是模糊现象;“深”既指从外到里的深度(如“庭院深深”),又指从上到下的深度(如水深),这是词的多义性;但多深的水算深水?它同浅水没有固定的界限,这却属模糊现象;“咬死猎人的狗”既可指“咬死了猎人”的狗,也可指“咬死猎人的狗”,这是歧义现象,也不属模糊现象。“促织声尖尖似针”(贾岛)中的“尖”既指声音的高度,又指触觉(“尖似针”),还指织工被主人催迫,那催织声刺痛织工心灵的痛楚,这属多义、双关(三关),也不是模糊语言的研究对象。只有多高的声音才能叫做尖声,多锐利的针才能算尖针,才属模糊现象。“花非花,雾非雾,夜半来,天明去”(白居易),这是含混不清,语义不明(什么东西或谁来了?),不是模糊现象。因为“花”同“雾”是迥然不同的两种事物,中间不存在界限不分明的模糊界限。“潭州城廓在何处,东边一片青模糊”中的“模糊”,是指日常用语中的看不清楚,也不属扎德所提出的模糊现象。有人把《红楼梦》中林黛玉的话“宝玉,宝玉,你好……”这种话语未了的现象,也归属模糊,那就更不对了。只有既有原型又有边缘的现象才是模糊语言学的研究对象。一门新学科刚诞生时,它的研究对象往往不明晰,这是可以理解的。连罗素这样的大哲学家也曾将概括和不清晰(如照片的模糊)等现象同科学意义上的模糊混为一谈。直到美国数学家和计算机科学家扎德1965年发表《模糊集》^[5]等一系列文章,美国语言学家乔治·雷可夫将其理论运用于语言学后,模糊语言学的对象才逐渐明确和专一。

模糊语言在现实生活中有很大的意义。例如,让你到会场上找人,如果告诉你要找的是一个高个的、高鼻梁、大耳朵、身体胖胖的中年人,你比较好找;如果告诉你身材、鼻梁的准确高度,耳朵的准确长度,准确的体重和年龄,你反倒不易找到。然而,追捕罪犯的通缉令中却必须说明准确的高度和年龄等,否则就可能抓错人。

律师特别看重精确语言同模糊语言的区别。例

如有一个冒称“教授”“研究员”的学术骗子,自称“著名语言文字学家”,其论著代表当代汉语言文字学术研究的最髙水平”^[6]。律师在驳斥这个人时,特别注重精确词语“教授”“研究员”“最髙水平”,因为它们只存在“是”和“否”两个答案,不似“著名”这个模糊词有不同等级的隶属性。

文学描写中充满了模糊语言。例如,俄国作家莱蒙托夫在《当代英雄》中对主角皮却林有这样一段描写:“他中等身材。他那匀称纤细的躯干和宽阔的肩膀,表明他生有一副强健的体格,能经受流浪生活的种种艰苦和气候的变化。”如果把这句话改成“他身高 1.7 米,腰围 75 公分,肩宽 50 公分,体重 70 公斤,能经受气候从零下 40℃到零上 35℃的变化”,文学味儿便丧失殆尽,变成索然寡味的体格检查表,而不是对人物的生动描写了。这类例子在文学作品中俯拾即是,所以现代诗歌中的象征派很赞赏词的模糊性。象征派的代表、法国著名诗人马拉美曾说:“过分精确的意义,会使模糊的文学变得面目全非。”

模糊的对立面是精确。同世界上的一切对立着的矛盾面可以在一定条件下相互转化一样,模糊和精确也可以相互转化。例如:“咫尺雪山路,归飞西海隅”(杜甫)和“近在咫尺,远在天边”中,“咫”是古代长度名,周制八寸,合现在市尺六寸二分二厘,“尺”则是十寸,两者本来都是表达精确概念的,但是两个精确概念连在一起却变成了一个同“天边”一样的模糊概念。类似的例子还有“寻常”。古代八尺为寻(如“千寻铁锁沉江底”),倍寻为常,本来都是表达精确概念的词,连在一起却表达模糊概念,意义是“平常”。只在杜甫的《曲江》诗中“酒债寻常行处有,人生七十古来稀”中,还有意用借对,暗含着古代“寻常”的精确的数字意义,与“七十”相对。又如,“千钧一发”中的“钧”本为我国古代的重量单位,合三十斤,现在这个成语喻指极其危险,除非专门研究古代语言的,一般人已不再把“钧”看作表达精确概念的词了。

语言中模糊向精确的转化比精确向模糊的转化更为常见,许多精确的科学术语都是用模糊词素构成的。当然这种精确性是为了科学上的需要,人为地加以规定的,并不一定意味着客观现实本身存在泾渭分明的一刀切的界限。如“高——低”是模糊概念,“高

血压症”在医学上却是精确概念,通常将成年人的血压持续地超过 140/90 毫米水银柱的人定为患有高血压症的病人。医学上如果不作出这种硬性规定,医生就没有办法开诊断书,尽管实际上低压在 89 毫米同 91 毫米的人之间未必真有那么明显的一刀切的界限。这类例子还有医学上的“高烧”(或“高热”)被人为地规定为 39℃以上的体温;体育上的“高栏”,人为地规定栏架高 1.067 米,低栏架高则为 76.5 厘米。“高速切削”则因机床不同,而作出不同的规定:在车床和铣床上每分钟切削 100 米以上才叫“高速切削”,而在龙门刨床上每分钟切削 50 米以上就叫“高速切削”。在不同语言中,模糊性的表现可能各有特点。例如“怀孕”是个精确词,但是在英语口语中却可以说 a little bit pregnant(字面翻译是“有点儿怀孕”,指“肚子有点儿大了”),very pregnant(字面翻译是“很怀孕”,指“肚子很大”),这是因为英语中在这种场合忌讳说 belly, abdomen(肚子)等词,只好采用这种不合模糊逻辑原理(因为对精确词是不能用模糊限制词修饰的)的委婉语。

模糊语言的研究正方兴未艾。它已发展为认知语言学的组成部分,后者又和认知心理学同为当代科学关注的一个中心,因为这方面的研究或许有助于理解人的思维过程,从而解决人工智能这一既有理论意义又有实践意义的重大课题。

注:

[1] 《大洋洲心理学和哲学杂志》,第 1 卷,83—84 页,1923。中译文载《模糊系统与数学》,1990(1)。

[2] 《马克思恩格斯选集》,第 3 卷,535 页,北京,人民出版社,1972。

[3] 袁毓林:《词类范畴的家族相似性》,载《中国社会科学》,1995(1)。这是我所见到的第一篇用家族类似原则分析汉语词类的文章。

[4] 最近的例子是池太宁同志在《模糊概念与模糊词语》(香港《语文建设通讯》第 51 期,1997 年 3 月号)中将“人”包括“男人”“女人”“老人”“小孩”“胖人”“瘦人”等视为模糊性的例证;其实,这是典型的概括性。

[5] 刊美国《信息和控制》,第 8 卷,1965(3)。

[6] 李远明:《“就此打住”应该缓行》,载《汉字文化》,1995(3)。

责任编辑 韩云波