

文章编号: 87

生物医学文献中模糊限制语的辖域标注规则研究 *

张静¹, 亢世勇¹, 杨欢², 周慧巍²

(1. 鲁东大学 文学院, 山东 烟台 264025; 2. 大连理工大学 计算机科学与技术学院, 辽宁 大连 116024)

摘要: 该文通过建立生物医学领域文献的语料库对其中所含模糊限制语的辖域标注规则进行了分析研究。模糊限制语辖域的不同主要是由于词性的不同而引起的句法成分的不同所导致的。因此, 文章将模糊限制语分为形容词、副词、连词、动词、概数词、独立语、量词短语几类, 并逐一做了分析说明。该文提出的分类方法和标注规则经过语料检验证明是比较有效的, 可以促进模糊限制语在该领域的识别研究, 并为其它领域的研究提供借鉴。

关键词: 模糊限制语; 辖域; 标注

中图分类号: H08

文献标识码: A

A Study on Identification Rules of Hedges Scope from Biomedical Literature

ZHANG Jing¹, KANG Shiyong¹, ZHOU Huiwei², YANG Huan²

(1.School of Literature, Ludong University, Yantai, Shandong 264025, China; 2.School of Computer Science and Technology, Dalian University of Technology, Dalian, Liaoning 116024, China)

Abstract: This article analyzed the rules of hedges scope by studying the hedges in the biomedical literature corpus established. The rules of hedges scope were mainly due to the different syntactic constituents which caused by its parts of speech. Hedges, therefore, divided into adjectives, adverbs, conjunctions, verbs, approximates, parenthesis and classifier phrases, has been analyzed one by one. The classification and identification rules proposed in this paper will promote the research on hedges, which is proved to be effective by language materials in the hedges corpus constructed.

Key Words: hedge; scope; annotation

1 引言

1965年, 美国加州大学控制论专家 Zadeh 在《信息与控制》杂志上发表了论文《模糊集》, 最早提出了模糊集的概念。这一论文的发表标志着模糊学的诞生, 在其中 Zadeh 提出了著名的模糊理论^[1]。1972年 Zadeh 在其另一篇文章《对模糊限制成分的模糊集论解释》中提出了 linguistic hedges (模糊限制成分) 的说法^[2]。

*收稿日期: 2015-6-10

定稿日期: 2015-8-10

基金项目: 国家自然科学基金 (NO.61272375)

美国著名的生成语义学家 Lakoff 在 Zadeh 模糊集合理论和模糊限制成分的基础上,首次提出了“模糊限制语(hedges)”这个概念,这也是其第一次单独作为一个语言学术语出现。Lakoff 认为^[3],模糊限制语就是把事物弄得模模糊糊的词语。

1.1 英语模糊限制语的研究概况

模糊限制语诞生后引起了诸多学者的关注,人们从各个方面对其进行探索和研究。

70年代,模糊限制语的研究主要表现在语义学方面。Lakoff^[4]以“从术语上讲”、“严格说来”、“大致说来”等几个模糊限制语为例说明了制定隶属度的依据标准,即定义、主要标准、次要标准和临时特征。McCawley 运用不同公式研究了“很”的不同语义解释。此外,基于实际语言材料,Channell 对模糊量词,Mosteller 和 Youtz 对表示可能或频率的词语做了研究。Prince 及其合作者则认为^[5]模糊限制语可以通过两种不同的方式让词语模糊化,并据此将模糊限制语分为变动型和缓和型两大类。

到了80年代初,模糊限制语的研究重心开始转向语用学方面。在这一领域,学者们主要是把模糊限制语与礼貌原则、合作原则和会话含义等语言学理论结合后,研究其在特定的语境中所产生的交际意义。如:Leech、Nikula 等从话语分析, Franken 等从关联理论, House 和 Kasper、Blum-Kulka 等从言语行为理论等角度出发所做的研究。

模糊限制语最初兴起于逻辑学、语义学、语用学、话语分析等领域,在这些方面的研究取得了一定的成果后开始逐渐转向语篇及科技文体中的运用研究。Panther 对科技文体中的间接言语行为做了专门研究^[6]。Markkanen^[7]探讨了模糊限制语在科技文体中的功能。Meyer^[8]等认为,模糊限制语在科技文体中的功能随读者对象、文体特征及作者的人格的不同而不同。他们还提出,不同领域的不同文体中模糊限制语的使用频率也是不同的。

此外,许多学者还从第二语言习得的角度对模糊限制语做了研究。Kasper^[9]发现,一般外国学生在交流时很少使用语气词,他认为这是因为学生还没有发现语气词也可以作为一种语用交际策略。Clyne^[10]发现,母语为德语,第二语言为英语的学生不能像英语母语者那样熟练地使用模糊限制语,他认为这可能是受母语负迁移或目的语文化的影响。Clyne 还发现通过大量含有模糊限制语的语料的阅读训练后,第二语言学习者会比母语使用者更加擅长使用模糊限制语。

模糊限制语的研究最早兴起于国外,到现在已经有近半个世纪的历史了,因而国外对其的研究也是比较充分。从最初对模糊限制语定义的归纳,到对其功能意义的探讨,对其的研究目前已然涉及逻辑学及语义学、语用学、第二语言习得等语言学的各个领域,研究成果也日益深入,并趋向细致化、系统化。近些年,学者们的眼光开始转向模糊限制语在不同文体中的使用情况的研究,比如公文、科技文、新闻访谈节目等。不同文体中的应用研究促使学者一改以往的研究模式,开始从实际的语言材料出发研究模糊限制语本身,并逐渐走出纯理论式的研究。这为语言学提供了一个新的研究思路和方法,也使得研究成果更加贴近语言本身。

1.2 汉语模糊限制语的研究概况

“模糊限制语”这一概念提出后,也引起了很多国内学者的关注,他们开始将英语的研究成果逐渐介绍到国内。英语模糊限制语研究领域的广泛使得汉语研究从一开始就站在了一个比较高的起点,促进了各学科与模糊限制语更深入地融合。对其的研究也开始由各自的单一

领域走向融合，研究方法也开始多样了，不再局限于一种语言内部的纵向挖掘，也开始了各种语言之间的横向比较。研究成果也由单纯的理论研究转向应用研究，对许多学科理论的发展，如第二语言习得、翻译等起到了推动作用。

80年代初，伍铁平先生最早对模糊限制语按照是否有上限和下限进行了分类。1985年，何自然把模糊限制语分为两类：一类是变动型模糊限制语，属于语义范畴，可以按照实际情况改变对话题的认识；另一类是缓和型模糊限制语，属于语用范畴，不能改变话题内容，只表示说话者对话题所持的猜疑或保留的态度^[11]。张乔对模糊数量限制词从语义学等角度进行了系统研究，并按照词性的不同把模糊词语分为了16类^[12]。何济生、姜晓蕙等根据英语模糊限制语的分类对汉语模糊限制语的语义特征及分类做了探讨。陈林华、李福印^[13]区分了“模糊、含糊、歧义、模糊限制语”这几组概念，并从语义方面对模糊限制语做了分类研究，认为应分为程度、范围、数量、质量和方式准则模糊限制语五种主要类型。黎千驹从是否易于数学刻画、位置形式以及语义类型三个角度对模糊限制语进行了不同的分类^[14]。

肖奚强对模糊限制语中的程度副词和形容词的语义限制作用进行了探讨，并考察了二者在语义限制上的差别^[15]。此外，肖奚强，李敏、王安节均表达了自己对Zadeh散漫化算子¹，即“（比）较、微、稍、略微、有点”一类的模糊限制语的不同观点，并作了相应的分析。

李福印还从模糊限制语与性别、社会地位、职业等社会语言学元素的关系入手，探讨了模糊限制语的社会功能^[16]。

张乔对模糊限制语的语用功能进行了讨论，认为模糊限制语²可以“改变模糊语义的隶属度和使精确的语义模糊化^[12]”。黎千驹对此表示赞同，并认为模糊限制语还有改变明晰中心词或明晰命题的真值条件，使表意具有灵活性及增强或舒缓肯定、否定的语气的功能^[14]。孙建荣，刘宝珠，陈林华、李福印，陈治安等众多学者还从会话含义、合作原则、礼貌原则等语用学角度对模糊限制语在言语交际中的语用功能进行了探讨。

我国也有学者从第二语言习得及外语教学的角度研究了模糊限制语对其的促进意义。黄锦如、陈桦，刘珍，段恩香等认为，老师在教学的过程中通常会忽略模糊限制语在交流方面上的语用功能，如果教师能有意识地引导学生在这方面的学习，不仅可以提高自己的教学水平，还能帮助学生了解文化差异。刘珍^[17]还探讨了模糊限制语在以英语为中介语、汉语为母语的科技文和以英语为母语的科技文使用中存在的差异。她发现，后者模糊限制语的出现频率和种类均小于前者，而且在出现频率方面的差别更大。

近年来，对于模糊限制语的研究主要集中在科技文和中英对比研究领域。在科技文方面，赵英玲最早对其进行了分类，认为根据模糊限制语的功能和限制的范围可以分为倾向“准确”、“作者”和“读者”的模糊限制语三大类^[18]。此外还有许多学者均对科技论文中模糊限制语的表现形式、语用功能以及人际功能做了探讨。值得关注的是，杨慧玲首次利用语料对科技论文中模糊限制语的语篇语用功能进行了探讨，并对模糊限制语进行了多层次分类^[19]。余千华、秦傲松也利用语料库对模糊限制语的使用情况进行了一些数据统计分析，认为中方科技工作者在用英语写作科技论文时要分清众多模糊限制语之间的细微差别，恰当而又多样化地使用。谢巧静，郭美兰、黄月华还根据顺应性理论对科技文中的模糊限制语的语境顺应

¹ 算子，数学中的概念，引入语言学中指具有调整和改变中心词词义作用的限制语。

² 张乔在其著作中称为“模糊限制词”。

性进行了探讨。此外,学者们还探讨了模糊限制语在新闻媒体、政治演说、访谈节目、商业广告、文学作品、商务英语、法律语言、医药卫生、经济贸易等各个领域、各种文体中的应用情况。

由于英语模糊限制语研究比较充分,使得国内学者可以从汉英对比的角度对模糊限制语做进一步的研究。李斌、苏远连、张长明、周红等人对英汉模糊限制语的类别和功能作了分析对比。唐艺菲、何建容等对中英文社科人文类书籍中前言部分涉及到的模糊限制语做了研究。此外,王舟、姜旭、陈萍、王玲、梁云菲、解淑暖等利用语料库对学术论文中摘要部分,夜楠对引言部分的模糊限制语做了对比研究。研究结果普遍表明,英美的科研人员在模糊限制语的使用上较中国人更为灵活多样。在各部分文章中,英语模糊限制语的使用频次均高于汉语,英语中主要以直接缓和性模糊限制语为主,而汉语主要以程度变动型模糊限制语为主,并且明显缺失间接缓和型模糊限制语。

由以上论述我们可以看出,八九十年代,我国对于模糊限制语的研究比较缓慢。但是进入21世纪后,随着互联网网络技术的发展以及中外交流的日益频繁和深入,模糊限制语的研究也迅速发展起来。目前,我国对于模糊限制语的研究已经取得了比较丰富的成果,研究内容从模糊限制语本体研究、功能研究转向了对比研究、应用研究,研究领域也逐渐扩展到社会语言学、语篇分析、跨文化交际、语言教学、语用学等学科,引发了模糊限制语研究的一个高潮。然而不可忽视的缺点是:近些年对于模糊限制语的研究多重复表述,真正有价值的研究内容比较少;对模糊限制语的纯理论的研究比较多,基于实际语料的研究比较少;从单一学科出发的研究比较多,学科融合的研究比较少;从人的理解和思维的角度研究的比较多,从计算机信息处理或者说机器思维的角度研究的比较少。

此外,目前汉语研究者大多把目光集中在修辞学、词源学和词汇学等方面,研究结果往往应用在在心理学、语言教学等相关领域,而对于生物医药领域中的模糊限制语使用情况的研究比较匮乏,这也印证了本研究的必要性。基于以上研究现状,本文拟结合语言学理论和计算机处理方法,从计算机对模糊限制语的信息处理角度对生物医药领域的模糊限制语进行研究,希望能探索出一套比较适合计算机标注学习的规则,从而能够对计算机领域模糊限制语的检测方法等研究提供一定的借鉴。

2 本文的研究意义和研究方法

2.1 本文的研究意义

随着计算机应用的普及和互联网技术的高速发展,网上各种文档资料呈井喷式增长。然而如何从这海量的文档中检索出研究人员真正所需的信息却变得日益困难。不仅是因为信息文本数量的庞大,同时也存在着大量模糊信息使得事实信息淹没其中而难以被检测出来的问题。未经检测的信息将会检索出大量不可靠信息,影响信息抽取结果的准确性。因此,如何区分模糊限制信息和事实信息就成了关键的一步,模糊限制信息检测成为信息检索的首要步骤。开展中文生物医学领域及其他领域的模糊限制信息识别研究,有助于提高抽取信息的真实性和可靠性,对中文事实信息抽取意义重大。

模糊性是人类语言的一种属性,由于自然语言或现实条件的各种局限性,无论在在语言交流还是科技论文写作中,学者们都需常常借助模糊限制语来表达一些不确定或可能性的意义。医学论文写作也不能例外,这种情况导致了生物医学文献中大量模糊限制语的存在。而

生物医药领域与人生命攸关的特殊性又使我们不能忽视。此外，模糊限制语的研究虽然已经引起了人们的重视，但是对于生物医药领域的模糊限制语的研究还比较少。是以，对生物医药领域的模糊限制语进行研究势在必行。

目前对汉语中的模糊限制语的多是从其功能、意义上进行研究，真正研究模糊限制语及其辖域标注的比较少，很难找到相关文献，更不用说生物医药领域的模糊限制语标注研究了。本文对知网上所有的文章进行检索，仅发现 5 篇相关的论文，本研究的必要性可见一斑。

2.2 本文的研究方法

(1) 语料库的方法。它分为语料收集、标注和统计三个步骤。本文构建的汉语模糊限制语语料库覆盖了生物医学领域多类期刊，共约 2 万句语料。生物医学领域语料选自 2011-2013 年间《中国生物医学工程学报》、《介入放射学杂志》、《中国生物化学与分子生物学报》、《中国癌症杂志》和《生物医学工程学杂志》等权威性中文生物医学类期刊刊登的科研论文。

首先，从知网上下载上述要求内的论文，并转换为文本格式。对于转换过程中出现的乱码进行人工校对、修改。

其次，对转换生成好的文本进行摘选处理。本文语料库由 5 个部分组成，分别为论文摘要、内容讨论、实验结果、文章结论以及全文。

第三，对摘选好的文本进行格式和编号处理。给每篇论文、每个句子各一个编号，合起来作为这个句子在语料中的编号，如 S117.8 就是 117 号论文中的第 8 个句子，就在这个句子前面标注<sentence id="S117.8">。然后对该句中的模糊限制语进行检索，不含模糊限制语的句子为肯定性句子，含有模糊限制语的句子即为模糊限制性句子。对模糊限制性句子中的模糊限制语按照其出现顺序分别编号，再加上之前的文章号和句子号，合并后作为模糊限制语的编号，并以<ccue>和</ccue>来分别标注其起始位置和结束位置。

第四，对计算机标注好的最初语料进行人工修正，逐一确定计算机标注出的词语究竟不属于模糊限制语，如果属于，对它的辖域起止位置进行标注。我们请两位语言学专业的研究生对标注好的语料分别进行校对，然后对他们校对好的两份语料进行对比分析，统计出其中不一致的地方进行再标注，直到达成统一的结果为止。

经过统计，语料中共含模糊限制语 708 个，其中摘要部分含 27.6%，内容讨论部分含 46.9%、实验结果部分含 16.3%，文章结论部分含 3.5%，全文部分含 5.7%。而模糊限制性句子的最终标记结果是一万八千余句，其中摘要部分含 23.9%，内容讨论部分含 48.3%、实验结果部分含 23.7%，文章结论部分含 35.6%，全文部分含 37.6%。

(2) 定量与定性相结合。本文基于自建的生物医学论文组成的语料库，以其中的论文作为本文研究的封闭域，在对此语料库中的标注语料进行分类定量统计研究的基础上进行定性分析，对比两份标注语料的不同之处加以分析研究。

(3) 描写与解释相结合。描写与解释这两种方法是相辅相成的，本文先对两位研究者的对比结果作了细致的分析描写，探索形成不同标注结果的深层原因，并借助语言学理论做出科学的解释，进而总结归纳出生物医学文献中汉语模糊限制语的产生条件及标注规则。

3 生物医学文献中模糊限制语辖域的标注规则研究

关于汉语模糊限制语的辖域标注的研究非常少，知网上仅有的几篇论文也是从计算机领

域研究如何利用各种数性特征通过运算等来识别的方法的。在一定程度上我们可以说，汉语模糊限制语辖域的具体标注规则的研究还是很缺乏的。因此，本文拟在对大量语料进行计算机标注的基础上，从语言学与计算机结合的角度，总结出简洁有效的中文模糊限制语标注规则，以期促进这一方面的研究。

本文在对模糊限制语的辖域进行标注时遵循的最基本的原则是最大原则，也就是指模糊限制辖域应该包括它能影响到的最大范围。本文采用<scope id="1">和</scope id="1">来分别标注出模糊限制信息辖域的起始和结束位置。其中属性“id=”后面的数字代表的是该模糊限制语在其所在的句子中的编号：1代表第一个模糊限制语的辖域，2代表第二个模糊限制语的辖域，以此类推。此外，本文认为模糊限制语一定包含于模糊限制信息辖域之中，而且如果一个句子中有两个或两个以上的模糊限制语时，它们的模糊限制信息范围一般是层层嵌套的。

模糊限制语辖域的不同主要是由于其词性的不同而引起的句法成分的不同所导致的。因此，在探讨模糊限制语的辖域时，本文主要按照我国学者苏远连对汉语模糊限制语划分的语法类别进行论述。苏远连认为汉语模糊限制语分为八类^[36]：

- (1) 形容词，如：仔细、稳稳、干干净净、整整齐齐等。
- (2) 副词，如：最、极、太、很、很不、不很、十分、一再、偏偏、大概、也许、格外、稍微、常常、永远、终于、渐渐等。
- (3) 概数词，如：来（如：二十来天）、把（如：个把月）、多（如：五斤多）、分（如：九分熟）、成（如：八成）、只、都、上下、左右、也许、大约、仅仅、部分、总共、统统、全部、或多或少等。
- (4) 能愿动词，如：能、能够、可、可以、会、可能、应、应该、愿意、肯、敢、应当等。
- (5) 独立语，如：我估计、依我看、我认为、我想、就我所知、在我看来、看样子、总而言之、总之、一句话、一言以蔽之等。
- (6) 量词，如：些（如：有些傻）、点（如：快点）等。
- (7) 结构助词“得、地、的”构成的短语，如：红得发紫，马马虎虎地，看的看、说的说等。
- (8) 语气词，如：呃、嗯、噢、哦、啊、唉等。

但是基于生物医药文献正规书面语的特点，该类文章中是不存在语气词及口语式的结构助词构成的短语这两种模糊限制语表现形式的，因此这两类形式的模糊限制语本文是不进行讨论的。此外，在具体的语料标注过程中我们还发现一些其他的表达形式，如动词及量词短语等。基于以上几点，本文将生物医药文献中的模糊限制语归纳为形容词、副词、连词、动词、概数词、独立语、量词短语几类，下面将按照这几种分类逐一阐释模糊限制语辖域的标注规则。

3.1 模糊限制语为形容词或形容词性词组

当模糊限制语为形容词或形容词性词组时，一般而言，其模糊限制辖域开始于模糊限制语，结束于它所在的最上层的名词短语，如下面例句中的“一定的”，此时模糊限制语在句中充当定语。

例句 1: 当小胶质细胞已经向不同的方向极化后, 即使在极化因子撤除的情况下, 极化后的小胶质细胞依然能在<scope id="1"><ccue type="程度变动">一定的</ccue>时间内</scope id="1">保持其极化方向。

形容词或形容词性词组在句中还可以充当表语, 但我们生物医学文献标注中未曾发现此类的模糊限制语。这一类型在英语中比较常见, 它的影响范围包含整个句子, 如例句 2 中的“possible”的辖域就是整个句子。

例句 2: The demonstration of hormone receptor proteins in cells from malignant effusions is <possible>.

3.2 模糊限制语为副词或副词性词组

当模糊限制语为副词或副词性词组时, 其辖域的标注有两种不同的情况。

第一种, 是当该副词性模糊限制语为句子层面的副词时, 即该模糊限制语修饰限制该句的谓语动词时, 其模糊限制范围要包括该谓语所在的分句。如下例中“往往”的辖域开始于其所在分句的主语“病态问题”, 结束于其所在分句的宾语“严重降低”。这种类型的模糊限制语如果出现在单句中, 其范围一般是作用于整个句子的。

例句 3: 多通道欠采样非笛卡尔轨迹数据重建是当前磁共振成像的研究热点, 当欠采样因子<scope id="1"><ccue type="程度变动">比较</ccue>大<scope id="1">时, <scope id="2">病态问题<ccue type="频率变动">往往</ccue>使得敏感度编码 (SENSE) 方法重建图像信噪比<scope id="3"><ccue type="程度变动">严重</ccue>降低</scope id="3" ></scope id="2">, 传统的解决方法是在重建方程中引入 Tikhonov 约束或 TV 约束。

第二种情况是, 当该副词性模糊限制语为除句子层面的副词之外的其他副词时, 其所影响的范围一般开始于该模糊限制语, 结束于它所在的最上层的名词或动词短语。如下例中, 模糊限制语“特别”。

例句 4: 在本研究证明 SOCS 蛋白, <scope id="1"><ccue type="程度变动">特别</ccue>是 SOCS2 和 CIS-1 的表达</scope id="1">, 在 OA 中减少。

3.3 模糊限制语为连词

当模糊限制语为连词时, 其影响范围包括模糊限制语作用的句子中的整个条件和结果或者是其所涵盖的所有同位词。如下例中模糊限制语“若”的范围开始于其所在的条件句, 结束于条件所引起的结果句, 而“或”的辖域则包括了前后两种不同的情况——“接种于壳聚糖纳米纤维支架”和“普通条件下连续培养 7d”。

例句 5: 细胞通过跨膜信号传导实现胞外信息传入或胞间通讯, <scope id="1"><ccue type="条件假设">若</ccue>参与此过程的细胞因子或分子发生变化, 则可能导致肿瘤的发生、发展、侵袭及转移</scope id="1">。

例句 6: 将新鲜分离的猪肝细胞<scope id="1">接种于壳聚糖纳米纤维支架<ccue type="探知结论">或</ccue>普通条件下连续培养 7d</scope id="1">, 分为实验组 (Nano 组) 和对照组 (Hep 组)。

3.4 模糊限制语为动词或动词性词组

当模糊限制语为动词或动词性词组时, 其模糊限制辖域有几种不同的情况。

(1) 一般而言, 当模糊限制语为动词时, 其模糊限制辖域是模糊限制语所在的直接动词短

语，即开始于模糊限制语，结束于其所在的分句。由于动词之前的主语是确定的，其可能发生变化，也就是具有模糊限制性的只是它后面所带的宾语，因此，这一类模糊限制语的辖域不包含其主语。如下面例句中的模糊限制语“可能会”。

例句 7: 由于本次研究条件所限，在测量呼吸频率时采用人工测量，未采用机器测量<scope id="1"><ccue type="探知结论">可能会</ccue>造成测量存在一定误差</scope id="1">。

(2) 当模糊限制语为能愿动词且处于被动句式时，其辖域是包括主语在内的整个句子。如下面例句中的“可能”，它的模糊限制范围本来应该开始于其自身，因为它后面紧接被动语态，所以其辖域的起始位置调整至其前面的主语“它们的轨道”。

例句 8: 等离子体团彼此之间可以有交互作用，<scope id="1">它们的轨道<ccue>可能</ccue>也被互相的扭曲</scope id="1">。

(3) 当模糊限制语为动词且是被动语态时，其辖域是包括主语在内的整个句子，如“被认为”、“被视为”、“被报道”等。下面例句中的“被认为”的模糊限制范围就开始于其所在句子的主语，结束于所在句句末。

例句 9: 高胰岛素血症和胰岛素抵抗是 PCOS 患者的生化特征之一，<scope id="1">胰岛素抵抗<ccue type="探知结论">被认为</ccue>是 MS 发病的中心环节</scope id="1">。

(4) 在英语语法研究中，人们把 seem、appear、turn out 等一类的词称为提升动词(raising verbs)，因为这一类词在充当句子谓语时可以使从句主语提升到主句主语的位置。汉语中也有类似的词，如“表明”、“看做”、“说明”等。

当模糊限制语为这类词时，如果前面直接连接主语，其辖域涵盖其所在的整个分句，包括主语；当这类模糊限制语前面不是直接连接主语，而是其他成分，或者直接连接的主语为代词时，其模糊限制范围开始于该模糊限制语，结束于其所在的分句，不包括主语。

如下面例句中的“表明”，其模糊限制辖域就是从前面的主语“研究结果”开始，到整个分句的宾语结束，而“说明”由于它之前的主语是代词“这”，所以它的模糊限制辖域开始于其自身，结束于整个分句。

例句 10: <scope id="1">研究结果<ccue type="探知结论">表明</ccue>，血管壁外表面所受的应力值最大，内表面次之，中心处的应力值<scope id="2"><ccue type="程度变动">略</ccue>小于</scope id="2">内表面的应力绝对值</scope id="1">，这<scope id="3"><ccue type="探知结论">说明</ccue>血管外表面是最先出现裂纹的</scope id="3">。

3.5 模糊限制语为概数词

当模糊限制语为概数词时，其辖域包含其所在的整个上层短语结构，一般多为名词性词组。这类模糊限制语的辖域表现在具体语料标注中有两种不同的情况：

第一种模糊限制语的辖域开始于该模糊限制语，结束于其所修饰的中心语，如下面例句中的“或多或少”，它的影响范围就是开始于它自己，结束于其所修饰的中心语“变化”。

例句 11: 虽然各朝代正经腧穴的数目都有<scope id="1"><ccue type="范围变动">或多或少</ccue>的变化</scope id="1">，但仍在<scope id="2">365<ccue type="范围变动">左右</ccue></scope id="2">徘徊，而未被归经的奇穴或效应点的数量却迅猛增长并广泛用于临床。

第二种模糊限制语的辖域开始于前面的句法成分，结束于该模糊限制语，如例句 11 中

的“左右”，它的模糊限制范围就涵盖了其所在的方位短语结构。

3.6 模糊限制语为独立语

当模糊限制语为独立语时，其模糊限制范围多开始于该模糊限制语，结束于其所在句子的句末，这一类模糊限制语的辖域多为整个句子。如下面例句中的“从实验结果上看”的辖域均为其所在的整个句子。

例句 12: <scope id="1"><ccue type="探知结论">从实验结果上看</ccue>, 配准效果达到预期目标</scope id="1">。

3.7 模糊限制语为量词短语

当模糊限制语为量词短语时，多在句中充当定语，修饰限定后面的中心语，构成定中短语结构。此时，该模糊限制语的辖域就是其所在的上层短语结构，即开始于该模糊限制语，结束于其所修饰的中心语。如下面例句中的“多个”的辖域就为其所在的定中结构“多个凋亡相关基因”。

例句 13: 结论: TSA 可通过调控<scope id="1"><ccue type="数量变动">多个</ccue>凋亡相关基因</scope id="1">来实现其诱导胃癌细胞凋亡的作用，这种凋亡诱导作用是通过半胱天冬酶非依赖途径进行的。

4 小结

经过实际语料标注我们发现，本文中的上述规则已经能够基本涵盖生物医学领域文献中汉语模糊限制语辖域标注的所有问题了。也就是说这些规则还是比较完善的，完全可以应用于该领域其他文献的标注工作中或者为其他领域语料的标注提供一定的借鉴。然而由于本文基于的语料毕竟是有限的，在该领域其他语料标注中也有可能出现一些以上规则均不适用的情况，这种时候我们暂时默认模糊限制语的辖域标注开始于该模糊限制语，结束于句末。

以上规则是在大量的语料标注实践中总结归纳出的，具有较高的可信度和覆盖率。希望以上这些规则的提出能够对计算机领域模糊限制语辖域标注的研究起到一定的促进作用，同时也可以引起广大学者对汉语模糊限制语的辖域标注规则研究的关注。当然由于受本人能力及本文语料库语料规模所限，文中提出的规则难免有所遗漏或者不足之处，恳请广大学者批评指正。

参考文献

- [1] Zadeh, L. A. Fuzzy sets [J]. *Information and Control*, 1965 (8) :338-353.
- [2] Zadeh, L. A. A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges [J]. *Journal of Cybernetics*, 1972 (3) :4-34.
- [3] Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts [J]. *Chicago Linguistic Society Papers*, 1972 (8) :183-228.
- [4] Lakoff, G. 模糊限制词和语义标准 [J]. 廖东平, 译, *当代语言学*, 1982 (2) :22-27.
- [5] Prince, E. F., Frader, J., Bosk, C. *On hedging in physician-physician discourse*. In R.J. Di Pietro (ed.), *Linguistics and the professions* [M]. Norwood: Ablex, 1982:83-97.
- [6] Panther, K. Eignige typische indirekte sprachliche handlungen [A]. In *wissenschaftlichen diskurs*. *Wissenschaftssprache* [C]. 1981: 231-273.
- [7] Markkanen, R., H. Schroumler. *Hedging and Discourse: Approaches to the Analysis of a Pragmatic*

Phenomenon in Academic Texts [M]. Berlin: Walter de Gruyter, 1997:21-41.

- [8] Meyer, P. G. Hedging strategies in written academic discourse: strengthening the argument by weakening the claim [A]. In R. Markkanen., H. Schroder (Eds) *Hedging and Discourse: Approaches to the Analysis of A Pragmatic Phenomenon in Academic Texts* [C]. Berlin: Walter de Gruyter. 1997:21-41.
- [9] Kasper, G. Communication strategies: modality reduction [J]. *The Interlanguage Studies Bulletin*, 1979(4): 266-283.
- [10] Clyne, M. The socio-cultural dimension: the dilemma of the German-speaking scholar [A]. In Schroder, H. (ed.). *Subject-oriented Texts* [C]. Berlin: de Gruyter, 1991:49-67.
- [11] 何自然. 模糊限制语与言语交际[J]. 上海外国语学院学报, 1985(5):29-31.
- [12] 张乔. 模糊语义学[M]. 北京: 中国社会科学出版社, 1998:156-161.
- [13] 陈林华, 李福印. 交际中的模糊限制语[J]. 外国语(上海外国语大学学报), 1994(5):55-59
- [14] 黎千驹. 模糊语义学导论[M]. 北京: 社会科学文献出版社, 2007:181-195.
- [15] 肖奚强. 模糊限制词语义分析[J]. 安徽师大学报(社会科学版), 1990(2):207-212.
- [16] 李福印. 模糊限制语的社会语言学探讨[J]. 外语研究, 1995(4):9-12.
- [17] 刘珍. 模糊限制语与母语迁移[J]. 宁夏大学学报(人文社会科学版), 2003(3):106-109.
- [18] 赵英玲. 英语科技语体中的模糊限制语[J]. 外语与外语教学, 1999(9):15-17.
- [19] 杨慧玲. 科技论文中的模糊限制语[J]. 四川外语学院学报, 2001(1):84-87.

作者简介:



张静(1989—), 女, 硕士研究生, 主要研究领域为中文信息处理。
Email: zhangjing891010@126.com



亢世勇(1964—), 通讯作者, 男, 博士, 教授, 主要研究领域为中文信息处理、辞书编纂、现代汉语等。
Email: kangsy64@163.com



杨欢(1988—), 女, 硕士研究生, 主要研究领域为生物医学信息挖掘、机器学习和自然语言处理。
Email: yanghuan_dlut@mail.dlut.edu.cn



周惠巍(1969—), 女, 博士, 副教授, 主要研究领域为句法分析、生物医学信息挖掘和自然语言处理。
Email: zhouhuiwei@dlut.edu.cn