

语言学研究

近义量词的语料库量化研究

——以“丝”和“缕”为例

吴淑琼¹ 覃友号²

(1. 四川外国语大学 外国语文研究中心, 重庆 400031; 2. 四川外国语大学 英语学院, 重庆 400031)

摘要:既往对近义量词的研究多以定性分析为主,基于真实语料运用量化统计方法探讨其用法特征差异的研究尚不多见。以BCC语料库为数据来源,采用多重对应分析法,对近义量词“丝”和“缕”的行为特征进行系统分析,结果表明:(1)“丝”和“缕”在实际使用中具有交替性,其用法特征整体上呈现原型—样例连续统的特征,但二者的原型用法特征相对独立,且界限分明;(2)“丝”多与感知类和抽象概念类名词搭配,“缕”则更倾向于搭配空间客体类和具体事物类名词;(3)“丝”和“缕”的边缘范畴样例趋于重合,表明二者在范畴化过程中存在一定的同一性,并具有相似的认知理据。

关键词:近义量词;“丝”;“缕”;语料库;多重对应分析

中图分类号:I-1087 **文献标志码:**A **文章编号:**1674-6414(2026)03-0029-11

DOI:10.26922/j.cnki.waiguoyuwen.2026.03.003

0 引言

量词是语言研究的经典话题。量词在范畴化过程中会表现出一定的相似性,因此它们在特定语境中往往可以互换使用,进而形成近义量词。已有研究对近义量词的句法和语义特征进行了深入探讨。宗守云(2005)分析了“组”和“套”在名词性成分语义选择上的差异;李建平等(2009)考察了“枚”和“个”在历时演变中的兴替和动因;周芍等(2014)探讨了“条、根、道”的语义特征及其与名词的组合机制;冯赫(2018)阐释了“合”与“合(盒)”的语义演变;多塞德洛娃(Aneta Dosedlová)等(2019)对“棵”和“株”的名词搭配特征进行了对比分析;王金海等(2022)分析了“X双/对/副N”量词构式在词项搭配上的异同。然而,

基金项目:重庆市教育委员会人文社会科学研究重点项目“词汇类型学视域下的英汉感知动词历时演变研究”(24SKGH170)的阶段性成果

作者简介:吴淑琼,女,四川外国语大学外国语文研究中心教授,博士生导师,主要从事认知语言学、词汇语义学和语料库语言学研究。

覃友号,男,四川外国语大学英语学院硕士研究生,主要从事认知语言学。

引用格式:吴淑琼,覃友号. 近义量词的语料库量化研究——以“丝”和“缕”为例[J]. 外国语文,2026(3):29-39.



本刊微信公众号

既有研究多以内省思辨为主,鲜有借助语料库量化统计方法对近义量词的用法特征进行系统考察。本文以近义量词“丝”和“缕”为例,运用基于语料库的行为特征分析法(Corpus-based Behavioral Profile Analysis)(Gries et al., 2009; 吴淑琼等, 2020),对近义量词“丝”和“缕”的用法模式进行细致分析,揭示二者用法特征的异同及背后的认知理据。

“丝”和“缕”属于一维形状量词的一个子类,其指代的认知实体具有长度特征,而其宽度和高度则可忽略不计(李勇忠等, 2020)。这类量词所指对象通常具有“细长、轻小且不易观察”等特征。当“丝”和“缕”用作量词时,二者在许多搭配中可以互换,如例(1)所示,然而在某些语境下,它们并不能替换使用,如例(2)所示。

(1)a. 人们看到了蔚蓝的天空中飘着一丝白云^①。

b. 高夫人正在望着,忽然脚下边飘过一缕白云,把她的视线遮住。

(2)a. 当我暂时还没有失却我的最后一丝理智/*一缕理智。

b. 她像一缕幽魂/*一丝幽魂飘出他的视线。

量词“丝”和“缕”的用法特征早已引起学界的关注。何杰等(1996)分析了“缕”由名词语法化为量词的路径和形成机制,认为“缕”最初表示“线”,后转义为具体细状物体的计量单位,并进一步扩展为模糊细状物体和抽象事物的计量单位。猴瑞隆等(2009)探讨了“股、束、缕、绺”四个计量长条状物体的量词搭配及其认知基础,提出量词“缕”由名词通过隐喻、转喻以及意象图式拓展而来,适用计量不定集合量的细线状物。陈虞巧(2020)分析了量词“丝”的语义历时演变过程,认为其最初是名词,本义为“蚕丝”,后由名词虚化为量词;吕敏(2021)梳理了“丝”和“缕”语义演变的历时过程,认为二者均源于语法化过程中的重新分析。上述研究对“丝”和“缕”的句法和语义特征进行了深入探讨,但针对二者用法特征异同的专题研究尚不多见。有鉴于此,本研究拟基于语料库,对近义量词“丝”和“缕”的用法模式进行细颗粒度的对比分析,重点回答以下两个问题:(1)近义量词“丝”和“缕”的行为特征有哪些异同?(2)它们行为特征异同背后的认知理据是什么?

1 研究设计

1.1 语料来源

本研究语料选自北京语言大学 BCC 语料库。首先,在语料库的多语域子库中以“m 丝 n”和“m 缕 n”进行检索,分别获得“丝”的相关语料 11 446 条,“缕”的相关语料 3 615 条。然后随机下载了“丝”的语料 10 000 条^②,“缕”的语料则全部下载,再用 concordance

① 本文中所有例句均来自 BCC 语料库。

② 由于受 BCC 语料库规模和访问权限的限制,本研究仅随机抽取了 1 万条语料进行分析。另外,本研究所选语料均源自旧版 BCC 语料库,新版 BCC2.0 已于 2026 年发布。

randomizer 随机抽取两个目标词各 1 500 条语料作为本研究的分析对象。随后对抽取的语料进行人工筛查,剔除无关语料(比如“双丝鱼”“一缕麻”等),最终筛选出“丝”的语料 1 389 条、“缕”的语料 1 427 条。

1.2 语料标注

名量搭配是量词研究中的重要考察维度(Brems, 2010),因此在设置分类变量时,优先考察与“丝”和“缕”搭配的名词语义特征。通过对具体语料的分析,与量词所搭配的名词可分为“表情、触觉、光线、印迹、火焰”等 30 个语义类别。此外,量词所搭配的名词前通常有形容词加以修饰,形容词的语义韵也是影响量词语义的重要因素(Brems, 2003),比如“一缕柔暖的风”“一丝病态的嫣红”等。因此,本研究同时考察了形容词的语义韵,分为积极、消极和中性。

在句法特征方面,本研究主要考察了“丝”和“缕”所在小句的句式、体标记、极性和句法位置等。其中句式包括陈述句、感叹句、祈使句和疑问句;体标记主要涉及“着、了、过”;句子极性分为肯定与否定。就量词所在的句法位置而言,“丝”和“缕”所构成的表量结构通常出现在主语、宾语、定语和状语等位置。此外,两个量词前的数词可以是单数形式(如“一丝机会、一缕光彩”),也可以是复数形式(如“几缕清风、几丝异国情调、万缕阳光”),因此本研究将其列为变量之一。本研究的具体分类变量和变量水平详见表 1。

表 1 “丝”和“缕”的标注变量

标注变量	分类变量	变量水平
语义	名量搭配	毛发(hair)、丝织物(silk)、风(wind)、云(cloud)、光线(light)、火焰(fire)、间隙(aperture)、气体(gas)、植物(plant)、身体组织(body tissue)、色彩(color)、影子(shadow)、气味(smell)、烟雾(smoke)、液体(liquid)、食品(food)、印迹(imprint)、表情(expression)、声音(sound)、触觉(touch)、味觉(taste)、魂魄(soul)、情感(emotion)、理性(reason)、力量(power)、时间(time)、想象(imagination)、意志力(willpower)、内心感知(inner perception)、其他类(other nouns)
	形容词语义韵	积极(positive)、消极(negative)、中性(neutral)
句法	句式	陈述句(declarative)、感叹句(exclamative)、疑问句(interrogative)、祈使句(imperative)
	体标记	现在时:着(zhe) 过去时:了(le)、过(guo)、已(yi)、其他(others)
	极性	肯定(AFFIRMATIVE)、否定(NEGATIVE)
	句法位置	主语(subject)、宾语(object)、定语(attributive)、状语(adverbial)、其他位置(other positions)
	量词前的数词	单数(single)、复数(plural)

1.3 数据统计

本研究采用多重对应分析法对“丝”和“缕”的行为特征进行统计分析。多重对应分析通过降维将列联表的行变量(具体样例)和列变量(概念变量)之间的关系以点的形式展示为概念空间图。这种方法能够直观呈现变量之间的交互作用及其关系。具体而言,多重对应分析根据二维图上数据点之间的距离,揭示变量之间的关联强度和类型。数据点的距离越远,表示变量之间的关联性越弱;距离越近,表示关联性越强(Glynn, 2010: 251)。多重对应分析已被广泛用于语言研究领域,比如多义词的语义特征分析(Glynn, 2014; Liu, 2023)、近义词的区分(杜静等, 2019)、构式交替的制约因素(张炜炜等, 2017; 张懂, 2021)、构式的历时演变研究(邓云华等, 2025; 吴淑琼等, 2025)等,但目前运用多重对应分析探讨近义量词用法特征的研究尚比较少见。因此,本研究拟采用多重对应分析对近义量词“丝”和“缕”的用法特征进行深入考察。

本文的统计分析在R语言中进行。首先将标注好的语料按照多重对应分析的行变量和列变量格式导入到R语言,并加载相关数据包{FactoMineR}和{factoextra}。通过多重对应分析,将语料转换为多维概念空间图,并对多维空间中的变量进行降维处理,最终生成二维概念空间图。“丝”和“缕”的用法特征变量以数据点的形式分布于二维图中。

2 “丝”和“缕”的行为特征对比分析

2.1 “丝”和“缕”的核心概念变量特征

基于“丝”和“缕”的句法和语义分布特征,本文通过降维处理分析了核心用法变量以及变量之间的关联性(见图1)。根据图1,第一维度解释了4.39%的概念变量水平,第二维度解释了3.71%的概念变量水平。尽管这两个维度值之和偏低,这可能与多重对应分析方法本身的统计特性有关(Glynn et al., 2022)。为进一步验证分析结果,本研究使用R软件中的{ca}数据包及mjca()函数绘制碎石图(scree plot),以检验各维度的解释率(Levshina, 2015: 382; Greenacre, 2017: 152),结果见表2。由表2及其右侧的碎石图可以看出,调整后的第一个维度解释为48.1%的变异,第二维度解释了10.9%的变异,两个维度累计解释了59%的变异。此后各维度解释率明显下降,说明前两个维度已能较充分反映数据结构。因此,本文保留前两个维度,用以分析“丝”和“缕”的概念范畴及其特征。

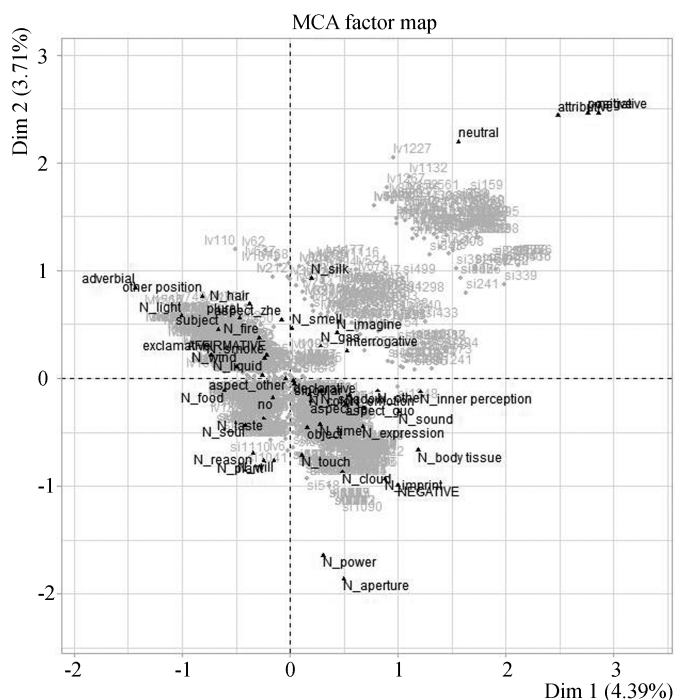


图1 “丝”“缕”基于样例和核心变量的 MCA 图

图1中的黑色三角符号表示“丝”和“缕”核心概念的不同变量水平,而灰色圆点则表示具体的样例。图1中,“丝”和“缕”的核心概念变量包含语义变量(名词搭配和形容词的语义韵)和句法变量(句式、体标记、极性、句法位置、量词前的数词)。为了呈现“丝”和“缕”各自的核心变量特征,本文将“丝”和“缕”作为补充变量,以可视化的形式展现它们变量的分布特征(图2)。从图2中可以看出,“缕”的用法特征主要集中在左上象限。具体来说,与“缕”距离较近的概念变量有体标记“着”,其句法位置多为主语和状语。从量词前数词的单复数来看,“缕”更倾向于与复数数词搭配,且多用于肯定句。就量词的名词搭配来看,与“缕”紧密关联的变量水平有液体(如“一缕春雨”)、风类(如“一缕清风”)、烟类(如“一缕炊烟”)、火焰(如“一缕火苗”)、毛发(如“一缕鬓发”)、光线(如“一缕晨光”)等名词。相比之下,“丝”的用法特征则集中在右下象限。与其距离较近的概念变量包括体标记“了”和“过”,其句法位置多表现为宾语。该量词前的数词主要以单数形式出现,用于否定句的情况比“缕”要多。“丝”的名词搭配特征包括内心感知类(如“一丝快意”)、表情类(如“一丝笑容”)、声音类(听觉)(如“一丝回音”)、情感类(如“一丝乡愁”)、时间类(如“一丝空闲”)、触觉类(如“一丝凉意”)、身体组织类(如“一丝软骨”)、云类(如“一丝白云”)、印迹类(如“一丝足迹”)、力量(如“一丝力气”)、间隙类(如“一丝空隙”)等语义类型。从这些变量的分布特征可以看出,“丝”和“缕”的核心用法特征存在显著差异。其中,名词搭配特征在二者的范畴化过程中起到了关键作用,这也与前人的相关研究结果一致

(如崔健,2010;毛智慧等,2012)。

表 2 碎石图检验结果

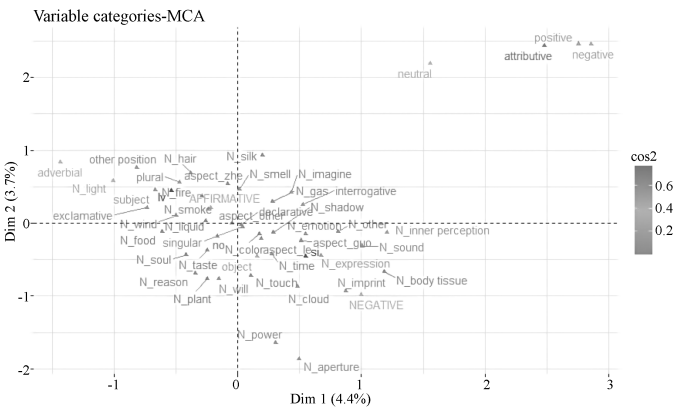
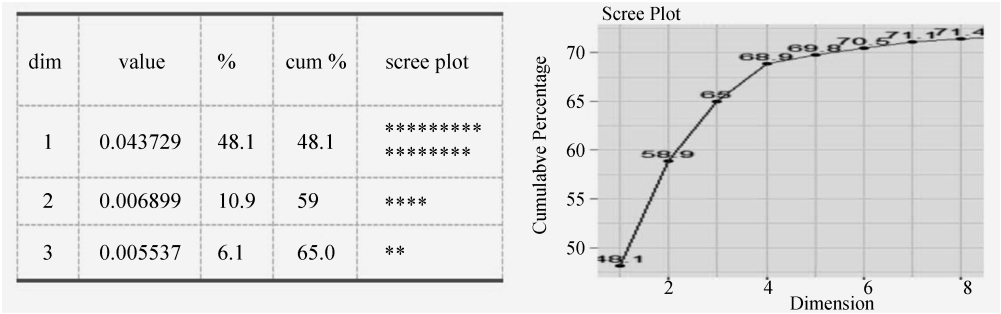


图 2 “丝”和“缕”作为补充变量和概念变量的 MCA 图

范畴化过程通常体现为原型和样例之间的连续相似度 (Taylor,2012),这种连续性也是认知思维在语言中表征范畴化的基本方式。范畴或概念的原型表征是通过对多个同类实例进行抽象和概括的结果 (Divjak et al., 2013: 224)。从图 2 可知,“丝”和“缕”的原型有明确的区分,表明它们在范畴化功能上存在差异。通过分析量词的语义和句法特征,可有效区分它们的范畴化功能。首先,从多个句法特征 (如句式、极性、句法位置等) 可清晰看出“丝”和“缕”之间的差异。其次,从语义范畴的功能来看,图 2 显示,“缕”作为量词时,倾向于与自然客体类名词搭配,如“液体、风、烟、阳光、火焰”等。这类词通常具有扩散性、流动性和动态性等特征,如“一缕水汽/阳光/炊烟/火焰/浮云”。相较于“缕”,“丝”更倾向于与“感知、表情、听觉、情感、触觉、身体组织和力量”等类别的名词搭配,如“一丝内疚/恐惧/缺憾/力气”。“丝”的抽象程度和主观化程度更高,更强调主体的感知体验。根据“丝”的名词搭配特征,可将其分为认知主体对外在事物的感知和反映 (包括视觉、听觉、味觉、嗅觉、触觉等) 和对内在心理状态的感知反映 (如“一丝内疚/恐惧/缺憾/力气”等)。

2.2“丝”和“缕”的范畴原型和边界的可视化分析

根据前述分析,“丝”和“缕”的核心概念变量表现出连续统的分布特征,而在原型搭配用法的范畴化过程中,两个量词的具体样例表现出一定程度的交替性。为进一步探讨“丝”和

“缕”在用法特征上的异同,本研究通过多重对应分析绘制了置信椭圆(confidence ellipses)。首先,围绕“丝”与“缕”的类别中心绘制其 95% 置信椭圆,结果如图 3 所示。图中的小椭圆表示类别中心的 95% 置信区域,圆点表示具体样例。“丝”与“缕”的类别中心分别分布于左上象限和右下象限,显示二者在二维语义空间中形成了相对分离的聚集趋势。这表明,它们在范畴化功能上存在较明显的差异。在此基础上,进一步围绕“丝”和“缕”的 95% 具体样例绘制置信椭圆,以呈现二者具体样例之间的关系(见图 4)。由图可见,两个置信椭圆呈现出部分重叠且相互交叉的分布态势,表明两个量词的用法特征在边界上具有模糊性,并存在互相渗透的现象,这说明二者在实际使用中具有一定程度的交替性。

结合图 2 和图 4,可以发现位于两个置信椭圆重合区域主要包括体标记、句法位置以及名词搭配等句法语义特征。其中,名词搭配特征显示“丝”和“缕”在这些搭配中交替使用的频率较高,涉及的名词类别包括“丝织、液体、味道、食物、想象”等。这种用法特征的交替性反映了它们在边缘范畴中的重叠性。“丝”和“缕”在名词搭配上存在共性,二者均可以与空间类、时间类以及流动性名词搭配,如“一丝/缕阳光”“一丝暖流”“一缕寒气”等。此外,在涉及抽象名词时,二者也可以根据各自的物象特征,将感知、情感 and 态度等抽象事物具象化,如“一丝/缕柔情”“一丝/缕记忆”等。然而,“丝”和“缕”各自仍具有相对独立的范畴区间,表明二者在范畴化功能上存在差异。例如,在名词搭配中,“缕”可以修饰超自然事物,如“一缕魂魄”,而“丝”则未见类似用法;相反,“丝”常用来修饰感知类名词,如“一丝尊严”“一丝安全感”“一丝异样”等,而“缕”则较少出现在这类搭配中。

综上所述,“丝”和“缕”在范畴化过程中既表现出一定的相似性,也展现出明显的差异性,这表明认知主体对二者进行范畴化时所采取的识解方式存在差异。

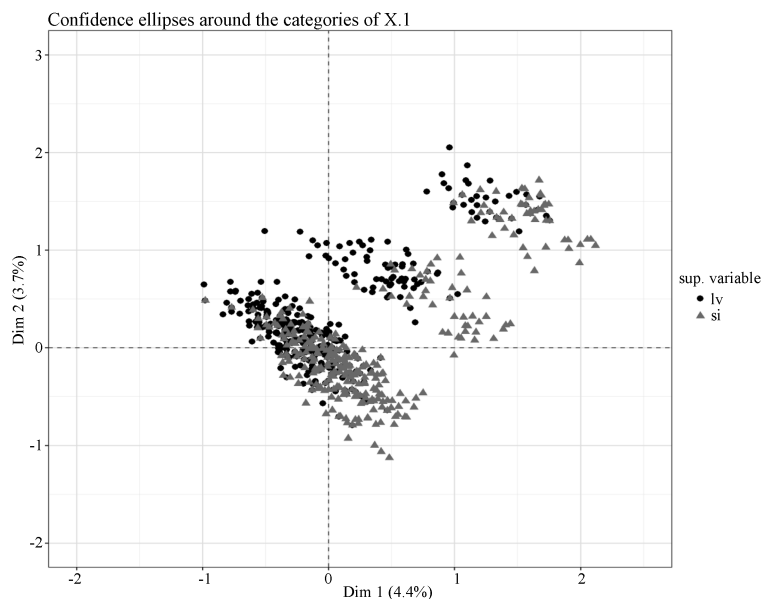


图3 “丝”“缕”基于原型样例的置信区间图

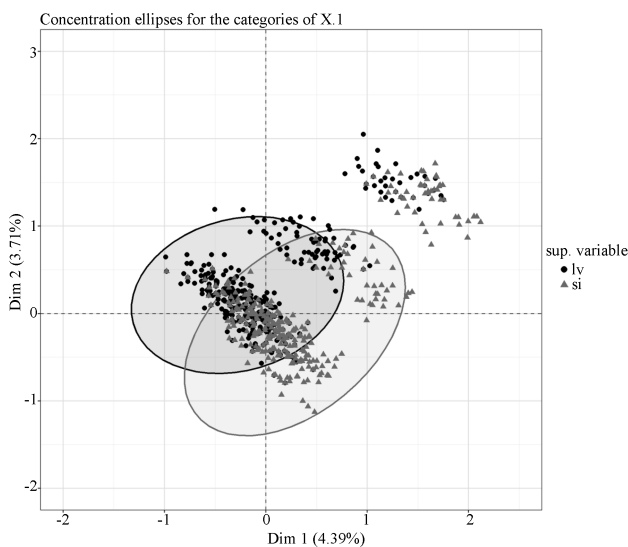


图4 “丝”“缕”基于具体样例的置信椭圆图

3 “丝”和“缕”行为特征异同的认知理据

根据原型范畴理论,范畴具有原型性、边缘性和连续性等特征。在“丝”和“缕”的范畴化过程中,两个量词从描述客观实体,逐渐扩展到涉及主体认知、情感和知觉的抽象表达,表现出名词搭配上的连续渐变特征。这一渐变过程表明了认知主体与客体之间的紧密关联,二者呈现出一种不可分割的连续统关系(邹君等,2021)。一般而言,人们对抽象事物的理解,往往依赖于对具体事物的感知和体验,这一认知过程在“丝”和“缕”的使用中尤为明显。本研究发现,量词“丝”和“缕”从搭配具体名词(如“一丝/缕头发”)到修饰抽象名词(如“一丝/缕线索”)的过程中,经历了明显的范畴化转变。从搭配名词的语义分类来看,“丝”和“缕”的原型范畴主要与具体可见的事物名词搭配,而与抽象名词的搭配则属于范畴化后的边缘用法,虽然二者表现出连续性,但其应用场景有所不同。当“丝”和“缕”与具体事物名词搭配时,主要体现为空间客体性,而在搭配抽象事物名词时,更多反映出认知层面的主体性。在此过程中,量词的功能不再限于单纯的表量,而是进一步参与到认知主体对抽象世界的理解与表征之中。通过对抽象事物的具象化处理,“丝”和“缕”呈现出基于感知体验和想象的意象构建,从而突显它们在情感和态度表达中的重要功能。

为何“丝”和“缕”所搭配的名词会由具体事物扩展至不可计量的抽象事物呢?既有研究从概念隐喻的角度对这一现象作出解释,认为隐喻机制在名量搭配的演变中发挥了一定作用(毛智慧等,2012)。基于此,本研究进一步指出,由于感知、情感、态度等抽象事物难以直接计量或量化,语言使用者往往通过概念隐喻,将具体的源域经验映射到抽象的目标域,从而实现对抽象概念的表达和理解。换言之,理解抽象事物往往需要通过认知加工使其具象化,认知主体通过隐喻将对熟悉事物的认知拓展到陌生领域,实现从具象到抽象的转化。例如,在“一缕思绪”中,量词“缕”搭配抽象名词“思绪”,通过其细长、连续的物象特

征,赋予了“思绪”具象的形态,使思绪呈现出如缕般绵延不断的特性。再如“一丝凉意”,通过隐喻将“丝”的特征映射至“凉意”,借助“丝”的细微来表现凉意的轻微,强化了凉意如丝般纤细的特点,从而达到了表意的效果。

范畴的划分受制于认知主体对事物的感知体验。量词“丝”和“缕”具有相似的真值语义,但根据用法本位观,词汇用法特征的差异最终取决于认知主体在具体语境中的实际使用。由于认知主体的感知方式和识解路径存在差异,近义量词的范畴化过程通常会呈现出一定程度的同质性和异质性。正如宗守云(2007)所言,名量搭配是量词范畴化的重要依据,量词对名词的选择过程也能反映其范畴化特征。“丝”和“缕”均可通过概念隐喻将无形事物具象化,从而实现分类、计量以及表达情感的功能。然而,语料分析表明,“丝”和“缕”的用法存在显著差异,反映出语言使用者对二者认知与理解的不同。在与具体事物名词搭配时,两者都可用于计量细小的条形物质,且具有细、长、轻和离散的语义特征,如“一丝头发/痕迹/细雨”“一缕阳光/青丝”等。但在与抽象名词搭配时,它们更多用于表达认知主体的情感和知觉,如“一缕思绪”“一丝悲哀”等。这表明量词的使用经历了由具体表量到多元表意、表情、表态的动态演变过程。这一过程在原型范畴理论视域下反映出“丝”和“缕”的范畴具有原型性和边缘性之分,其整体用法特征呈现原型-样例连续统的特点。此外,“丝”和“缕”在与名词的搭配中还表现出一定的选择倾向,如“希望”既可以与“丝”搭配形成“一丝希望”,也可以与“缕”搭配形成“一缕希望”。这说明在语义构建过程中,认知主体将“希望”既可感知为微弱而缥缈的情绪,其中“丝”侧重细微、模糊的情感表达,也可将其形象化为柔和的线状意象,其中“缕”凸显可感知的延展特质,从而呈现出情感与意象的双重属性。然而,“理智”通常只能与“丝”搭配,不能与“缕”搭配。这是因为“理智”主要指向内在的心理能力或认知状态,缺乏可被空间化、形象化的具体形态,不易与具有线状延展意象的“缕”搭配。“幽魂”则更适合与“缕”搭配,不宜跟“丝”搭配。这是因为“幽魂”虽属超验性存在,但在认知表征中被赋予了可视化的飘逸形态,与“缕”的线状延展意象相契合。这种差异说明“丝”在表达细微心理状态方面具有较强的适配性,而“缕”在呈现具象化、空间化意象时具有更明显的选择性。由此可见,认知主体在语言使用中不仅依赖物象特征来进行概念化,同时也在量词的选择上表现出基于隐喻映射的认知偏好。

总体而言,“丝”和“缕”的名量搭配体现了从具体客体到抽象概念的范畴化过程。当这两个量词与抽象名词搭配时,认知主体通过概念隐喻将抽象的情感、感知或心理状态具象化,使其获得可感知、可量化的表现形式。这一过程不仅反映了认知主体在语言使用中的主观能动性,也揭示了语言表达中蕴含的深层认知机制。通过类比思维,抽象事物在语言中被转化为具象图式,从而在认知上呈现出可感的空间形态。这种隐喻机制使得抽象事物在大脑中形成一维空间意象,从而展现出形状表量结构中有界与无界范畴的连续性(李勇忠等,2020)。

4 结语

本研究基于语料库,运用多重对应分析对近义量词“丝”和“缕”的行为特征进行了细致探讨。研究表明,“丝”和“缕”在实际使用中表现出一定的交替性,整体上呈现原型一样例的连续性特征。然而,二者的原型用法特征各自独立且边界清晰。具体而言,“丝”更倾向于与感知类和抽象概念类名词搭配,而“缕”则更常与表示空间客体和具体事物的名词搭配。此外,“丝”和“缕”在边缘范畴的样例中有较多重合,这表明二者在范畴化过程中存在某些相似性。

本研究不仅深化了对近义量词“丝”和“缕”的认识,还为相关量词研究提供了新的理论视角和方法论启示。从理论上讲,“丝”和“缕”在原型与边缘样例上的分布差异,揭示了近义量词在语义范畴化过程中所呈现的渐变性和模糊性特征。这一发现与语言学界关于量词类别边界具有动态性和可延展性的论断相互印证,同时进一步彰显了原型理论在汉语量词研究中的解释力。在方法论上,本研究运用语料库量化分析方法和可视化手段,为近义量词的系统性考察提供了新的研究路径。相比于传统的直觉判断或小规模例句分析,基于大规模语料库的数据分析能够更为全面、客观地揭示语言使用中的规律性。然而,需要说明的是,多重对应分析作为一种探索性统计方法,其结论是否具有心理现实性,仍需借助心理实验、神经语言学实验等实证方法加以验证。

参考文献:

- 陈虞巧. 2020. 量词“丝”的产生及其历时演变[J]. 河南理工大学学报(社会科学版)(1): 40-44.
- 崔健. 2010. 量词的功能差异和词类地位[J]. 汉语学习(6): 69-81.
- 杜静,李福印. 2019. “破”类动词的多维语义分析[J]. 外语研究(6): 22-29.
- 冯赫. 2018. 汉语量词“合”与“合(盒)”的历时考察[J]. 汉语学报(3): 81-86.
- 缙瑞隆,黄卓明,刘钦荣. 2009. 示形量词“股、束、缕、绺”的用法及认知基础——对外汉语量词教学个案研究系列之二[J]. 云南师范大学学报(对外汉语教学与研究版)(2): 18-22.
- 何杰,何平. 1996. 现代汉语量词的转义现象[J]. 南开学报(哲学社会科学版)(5): 76-80.
- 李建平,张显成. 2009. 泛指性量词“枚/个”的兴替及其动因——以出土文献为新材料[J]. 古汉语研究(4): 64-72.
- 李勇忠,尹利鹏. 2020. 英汉形状表量结构界限灵活性的认知建构研究[J]. 外语学刊(3): 33-37.
- 吕敏. 2021. 量词“丝”“缕”的语法化及动因探析[J]. 宁夏大学学报(人文社会科学版)(2): 51-57.
- 毛智慧,王文斌. 2012. 汉英名量异常搭配中隐喻性量词的再范畴化认知分析[J]. 外国语文(6): 61-64.
- 王金海,覃修桂,伊小丽. 2022. “X 双/对/副 N”量词构式精细范畴化的实证对比研究[J]. 解放军外国语学院学报(4): 43-51.
- 吴淑琼,江艳艳,刘迪麟. 2025. 空间维度强化副词行为特征的历时演变研究——以 highly/deeply/widely 为例[J]. 现代外语(3): 293-304.
- 张懂. 2021. 英语与格交替原型使用特征的多变量统计建模[J]. 现代外语(1): 49-61.
- 张进凯,金铨哲. 2022. 焦点敏感算子“就”和“才”互换多重对应分析[J]. 语言教学与研究(4): 101-112.
- 张炜炜,王芳. 2017. 从基于样例的概念空间看构式交替——以“让”和“给”的被动用法为例[J]. 外语与外语教学(6): 22-33.
- 周芳,邵敬敏. 2014. 近义量词“条、根、道”的三维解释与组合机制[J]. 语言教学与研究(1): 81-88.

- 宗守云. 2005. 量词“组”和“套”对名词性成分的语义选择[J]. 汉语学习(4): 25-30.
- 宗守云. 2007. 从范畴化过程看量词“副”对名词性成分的选择[J]. 世界汉语教学(4): 21-32.
- 邹君, 成晓光, 于林龙. 2021. 后现代主义一元论视阈下体认语言学的哲学理据再探析[J]. 西安外国语大学学报(3): 44-48.
- Brems, L. 2003. Measure Noun Constructions: An Instance of Semantically-driven Grammaticalization [J]. *International Journal of Corpus Linguistics* (2): 283-312.
- Brems, L. 2010. Size Noun Constructions as Collocationally Constrained Constructions: Lexical and Grammaticalized Uses [J]. *English Language & Linguistics* (1): 83-109.
- Divjak, D. & A. Arppe. 2013. Extracting Prototypes from Exemplars What Can Corpus Data Tell Us about Concept Representation? [J]. *Cognitive Linguistics* (2): 221-274.
- Dosedlová, A. & Wei-lun Lu. 2019. The Near-synonymy of Classifiers and Construal Operation: A Corpus-based Study of 棵 kē and 株 zhū in Chinese [J]. *Review of Cognitive Linguistics* (1): 113-130.
- Glynn, D. 2010. Testing the Hypothesis: Objectivity and Verification in Usage-based Cognitive Semantics [G]//D. Glynn & F. Kerstin. *Quantitative Cognitive Semantics: Corpus-driven Approaches*. Berlin: Mouton de Gruyter, 239-270.
- Glynn, D. 2014. The Many Uses of run [G]// D. Glynn & J. Robinson. *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Amsterdam: John Benjamins, 117-144.
- Glynn, D. & A. Biryukova. 2022. Death, Enemies, and Illness: How English and Russian Metaphorically Conceptualise Boredom [J]. *Yearbook of the German Cognitive Linguistics Association* (1): 33-58.
- Greenacre, M. 2017. *Correspondence Analysis in Practice* (3rd ed.) [M]. Boca Raton: CRC Press.
- Gries, S. T. & D. Divjak. 2009. Behavioral Profiles: A Corpus-based Approach Towards Cognitive Semantic Analysis [G]//E. Vyvyan & S. Pourcel. *New Directions in Cognitive Linguistics*. Amsterdam & Philadelphia: John Benjamins, 57-75.
- Levshina, N. 2015. *How to Do Linguistics with R. Data Exploration and Statistical Analysis* [M]. Amsterdam & Philadelphia: John Benjamins.
- Liu, M. 2023. Towards a Dynamic Behavioral Profile of the Mandarin Chinese Temperature Term re: A Diachronic Semasiological Approach [J]. *Corpus Linguistics and Linguistic Theory* (2): 289-321.
- Taylor, J. 2012. *The Mental Corpus: How Language Is Represented in the Mind* [M]. Oxford: Oxford University Press.

A Corpus-Based Quantitative Study of Near-Synonymous Classifiers: A Case Study of *Si* and *Lü*

WU Shuqiong QIN Youhao

Abstract: Previous studies on near-synonymous classifiers have relied predominantly on qualitative analysis. There are relatively few studies using quantitative statistical methods to explore the differences in their usage based on authentic corpus data. Drawing on data from the BCC corpus and employing multiple correspondences analysis, this study systematically examines the behavioral profiles of the two synonymous classifiers *si* and *lǚ*. The analysis yields the following findings. First, *si* and *lǚ* can be used interchangeably in actual usage, and their usage features demonstrate a prototype-exemplar continuum, though the prototypical usage features of *si* and *lǚ* are relatively distinct with clear boundaries. Second, *si* tends to collocate with nouns related to perception and abstract concepts, while *lǚ* shows a stronger tendency to co-occur with nouns related to spatial objects and concrete entities. Third, the peripheral instances of *si* and *lǚ* tend to converge, indicating some similarities in their categorization and suggesting shared cognitive motivations underlying their usage.

Key words: near-synonymous classifiers; *si*; *lǚ*; corpus; multiple correspondence analysis

责任编辑: 蒋勇军