

量词语言类型学研究的新视野^①

10.1 引言

量词(classifier)是在包括汉语在内的部分语言中,数词量化名词时伴随出现的词类。量词可以根据语义分成两类:个体量词(sortal classifier)及计量量词(mensural classifier)。个体量词选择具有特定语义特征的名词,并以名词指涉的个体作为数词计量的单位。计量量词则指定某个量值作为数词计量的单位。例如(1a)中“根”只能和表长条形物体的名词共现,且它出现时表示数词计算的是该长条形物体的个体数量,因此“根”是个体量词。反之,(1b)中“箱”指定一个箱子的容量作为数词计量的单位,因此是计量量词。

(1) a. 一根香蕉

b. 一箱香蕉

个体量词与计量量词在句法上占据同一个位置,并呈现严格的互补分布,不能同时出现在同一对数词与名词之间,如例(2)所示:

(2) a. *一根箱香蕉

b. *一箱根香蕉

① 本章由何万顺、梁永平撰写。

格林伯格 (Greenberg 1975: 30) 早已观察到同一种语言的个体量词与计量量词的语序相同。该现象不仅出现在现代汉语中, 而且在汉语三千年的历史上一贯如此 (Her 2017b), 在我们所建构包含 439 个量词语言的数据库中也并没有观察到例外 (Her et al. 2015)。因此, 我们可以确定这两者在句法上属同一词类, 可称之为量词。然而, 个体量词与计量量词在语义上的关系却较为复杂, 两者的语义既存在共性, 也有重要的差异。这种复杂的关系反映在语法层面, 可帮助我们辨识量词及区分它的两个次类。我们认为, 这是量词研究的基础。

本章作者所属的研究团队十余年来关注量词的研究, 本章整理了团队部分重要研究成果, 展现在量词类型学上的若干进展。本章的结构如下。10.2 节先列举个体量词与计量量词的语义差异。10.3 节聚焦于个体量词与计量量词的语义共性, 指出量词的语义功能是表达乘法关系中的被乘数。10.4 节讨论量词的语法表现与句法结构。10.5 节介绍我们建构量词语言数据库 (World Atlas of Classifier Languages, WACL) 的过程, 以及从中衍生的类型学研究成果。10.6 节总结本文。

10.2 个体量词与计量量词的语义差异

本节说明个体量词与计量量词的语义差异, 共分为以下四个方面: 本质属性与偶然属性; 语义集合论的差异; 可数/不可数之分; 个体量词独有的彰显功能。

10.2.1 本质属性与偶然属性

戴浩一和王连清 (Tai & Wang 1990: 38) 曾对个体量词与计量量词的语义差异提供了清楚的说明, “个体量词通过挑选出某些明显的感知特征将名词分类, 这些特征可能是物理性或功能性的, 并且是该类名词的指涉对象恒久拥有

的;反之,计量量词没有将名词分类的功能,而是表达名词指涉对象的量”。这也是个体量词又称为“分类词”的原因(何万顺、林昆翰 2015)。何万顺和谢祯田(Her & Hsieh 2010)另外以亚里士多德提出的本质属性(essential property)与偶然属性(accidental property)更清晰地区别个体量词与计量量词的语义:本质属性是指一个物体必然拥有的属性,偶然属性则是一个物体只在特定时空中才获得的属性;个体量词表达了其共现名词之指涉对象的本质属性,计量量词则表达其偶然属性。我们可以从例(3)看出这个差别:(3a)的量词“尾”表达了共现名词“鱼”的指涉对象必然拥有的一个属性——有尾巴,因此“尾”是个体量词。反之,(3b)的量词“桶”表达了鱼的偶然属性:鱼在桶子中或是鱼量是以桶来计算,这些都不是鱼必然拥有的属性,因此“桶”是计量量词。

(3) a. 一尾鱼

b. 一桶鱼

10.2.2 语义集合论

既然个体量词表达了共现名词指涉对象必然拥有的本质属性,计量量词则无此功能,我们就可以运用集合论的概念,更简洁地阐释两者的语义差别。假设一个量词的语义特征的集合是 X,其共现名词的语义特征的集合是 Y,当 X 是 Y 的子集合时,该量词便是个体量词,反之则为计量量词(Her 2012a; Her & Lai 2012)。换言之,个体量词虽有其语义特征,但对其所属的名词短语之语义并没有任何贡献,只是重复或强调了名词的部分语义而已;反之,计量量词必然会对其所属的名词短语增添新的语义。

10.2.3 可数名词与不可数名词之分

个体量词表达名词指涉对象的本质属性这一点和一个争论已久的话题相

关,即量词语言的名词是否有可数/不可数(count/mass)之分。许多研究都观察到个体量词只能跟可数名词共现,即指涉有清楚边界(bounded)的离散性(discrete)物体的名词,而计量量词则没有这个限制(Her & Hsieh 2010),如例(4)所示:

- (4) a. 一根香蕉/*水
b. 一箱香蕉/水

换言之,个体量词只能和可数名词共现。反之,只能跟计量量词共现的是不可数名词。然而,基耶尔基亚(Chierchia 1998)、戴浩一(Tai 2003: 312)却认为在汉语这类量词语言中,名词没有可数/不可数之分,因为量词语言的名词不会标记复数,且必须通过量词才能被数词量化。他们由此认为,量词语言的名词都是不可数名词,只能指涉没有清楚边界的、连续性的物质(substance),需要通过量词在其指涉对象上划出边界以定义出单位,才能被数词量化。

何万顺和谢祯田(Her & Hsieh 2010)指出该理论的相关问题。倘若量词语言的名词都只能指涉不可数的物质,(4a)的“一根香蕉”中的名词“香蕉”应该可以指涉不可数的香蕉泥,被量词“根”塑造成长条形的团块后才能被计算数量。但将“一根香蕉”诠释成一根香蕉泥是违反母语者语感的;对汉语母语者来说,“一根香蕉”只能指一个香蕉个体,具有香蕉皮这个天然边界所带来的离散性以及长条形的形状。量词“根”表达了长条形的形状,这是通过清晰的边界才能成立的属性;而“根”作为个体量词,意味着长条形以及必然伴随之的离散性是其共现名词“香蕉”指涉对象的本质属性。再以“一位老师”的“位”来设想,就更清楚,能与其共现的必然是一个人的个体,且有某种社会地位,因此必然是可数名词。

由此,我们得出结论:量词语言的名词确实有可数/不可数之分,有一部分的名词本身指涉的就是具有个体性的物体,不需要通过量词才能表达个体性,这些名词可以被恰当地称作可数名词。

10.2.4 个体量词作为语义彰显的工具

既然个体量词的语义仅仅是共现名词部分语义的重复而已,它在语义上就是冗余的。然而,这不代表个体量词没有任何功能,否则它们在自然语言中的存在就是一个不经济、不合理的现象。何万顺和谢祯田(Her & Hsieh 2010)、何万顺(Her 2012ab)及何万顺和赖宛君(Her & Lai 2012)指出,从认知语言学的角度来看,个体量词具有彰显(profile, Fillmore 1982; Langacker 1987)名词语义的功能。当名词与个体量词共现时,名词提供了一个背景(base)或框架(frame),个体量词彰显了这框架中的某些特征。以名词“鱼”提供的框架为例,“尾、条、只”虽均为可用的个体量词,但所彰显的特征不同(Her 2012a)。

- (5) a. 一尾鱼
- b. 一条鱼
- c. 一只鱼

“鱼”这个框架在认知中由许多不同的特征组成,它们可以通过不同个体量词彰显出来,如(5a)强调了鱼具有尾巴的特征,(5b)突出了鱼长条形的特征,(5c)则点出了鱼的有生性(animacy)。

个体量词彰显名词语义的功能,解释了为何它们只能选择同一类的名词与之共现:与个体量词共现的名词必须具有个体量词的语义特征,它们才能发挥彰显功能。个体量词因此的确有将可数名词分类的作用,因此也称为分类词,但其主要功能是彰显名词语义,分类仅仅是副作用而已(Her 2012a)。

10.3 量词在乘法上的认知语义共性

虽然个体量词与计量量词有语义差异,但何万顺(Her 2012a)指出,同样作

为量词的成员,两者具有一项相同的认知语义功能:表达数学概念中乘法关系的被乘数(multiplicand),亦即一个单位所包含的量值,而数词则表达乘数(multiplier),亦即单位的数量。这个想法的源头是格林伯格(Greenberg 1990: 172)最初对个体量词的描述:个体量词所表达的都只是“乘一”,又受到兰德曼(Landman 2004)、博雷尔(Borer 2005)和欧阳伟豪(Au Yeung 2005, 2007)等学者的启发。

兰德曼(Landman 2004)认为动词“乘”(time, multiply)能够发挥量词的功能,将名词指涉的总和(sum)包裹(parcel)成可数的单位。例如, four times three boys 这个短语就可以被分析成[four] times [three boys],亦即将三个男孩视为一个单位/包裹,数词 four 是单位/包裹的数量,也就是乘数。兰德曼的贡献在于将量词语义和乘法概念相连,指出量词的功能在于定义计数的单位。然而,何万顺(Her 2012a)指出,若将 four times three boys 翻成汉语“四乘三个男孩”,就可以看出动词 time 其实并不对应量词“个”。他并指出,若将被视为句法单位的 three boys/“三个男孩”替换成语义相同的名词 three-boy group/“三个小组”,便会发现英语中真正对应汉语量词‘个’的不是动词 time,而是复数标记-s,两者都表达“乘一”的概念,亦即以^个一个名词指涉对象为计数的单位。

(6) a. four **times one** three-boy group → four three-boy groups

b. 四乘^个一个三人小组 → 四个^个三人小组

博雷尔(Borer 2005)的观点比前文提到的基耶尔基亚(Chierchia 1998)、戴浩一(Tai 2003: 312)想法更为激进。她认为,不仅是量词语言,而是所有语言的名词都没有可数/不可数之分,名词本身的语义都是不可数的,只有依靠量词或复数标记(简称“复数标记”)发挥分割工具(divider)的功能,将名词指涉的物质分割成离散的单位,才能被数词量化。她指出在英语中典型的不可数名词如 wine 只要加上复数标记-s 就可以获得可数诠释,而典型的可数名词如 dog 只

要不加复数标记就会获得不可数诠释,如: *There is dog on the floor.* 博雷尔的理论已在 10.2.2 节被反驳了: 在区分了个体量词与计量量词后,我们会发现量词语言,或至少汉语名词的可数/不可数之分比英语还要严格: 汉语严格要求只有可数名词可以与个体量词共现,不像英语允许复数标记与不可数名词共现。当然,英语也是有可数/不可数名词之分的语言,因为不可数名词在语境中提供计量单位和将不可数名词诠释成种类——这两种特殊情况下才能加上复数标记 (Scontras 2022), 而可数名词则没有这些限制。但博雷尔的贡献在于指出英语的复数标记发挥了和量词相同的功能, 尽管如何万顺 (Her 2012a) 所说, 博雷尔并没有区分个体量词与计量量词, 所以更准确的说法是个体量词和复数标记具有相同的语法功能, 即: 两者与可数名词共现时, 可以确定以可数名词所指涉的个体为数词计数的单位。

欧阳伟豪 (Au Yeung 2005, 2007) 明确提出量词是乘法数词 (multiplicative numeral)^① 的延伸, 其功能就是表达被乘数。他指出: 汉语的位数词“个”(例如, “三跟五都是个位数”中的“个”) 与可以和多数可数名词共现的通用个体量词“个”(例如, “三个怪物”中的“个”), 有相同的语音形式, 并提出个位数词是个体量词前身的假说。然而, 欧阳伟豪 (Au Yeung 2005) 对量词表达的所谓“被乘数”的内涵, 却给出了一个过于复杂的定义: 他认为量词的语义是“每个单位有 n 个个体 (n token/object per unit)”, 且可以根据 n 的数值区分个体量词与计量量词, 如下例所示:

(7) a. 三个球

b. 三对球

① 表达乘法关系的数词, 可分为表被乘数的位数词 (numeral base) 和表乘数的乘数词, 如汉语数词“三百”表达了“ 3×100 ”的数学关系, 其中“百”是位数词, “三”是乘数词。

(7a)的量词“个”表达了“一个单位有一个个体”的信息,(7b)的量词“对”则表达了“一个单位有两个个体”的信息,两者的数词“三”都表达了“共有三个单位”的信息,数词与量词的信息则共同表达名词“球”的指涉对象的数量。当 $n=1$ 时,该量词为个体量词;当 $n\neq 1$ 时,该量词为计量量词。

欧阳伟豪(Au Yeung 2005,2007)是第一个用明确的数学概念描述量词的学者,但何万顺(Her 2012a)指出:所有量词都表达在一个单位中有多少个个体;换言之,他的理论走向了与博雷尔(Borer 2005)相对的另一极端,认为所有的名词都是可数名词,如“四瓶水”的意思其实是“以一个瓶子包含之数量的水分子为单位,共有四个单位”。何万顺(Her 2012a)认为这与心理语言学的事实不符,使用量词语言的人绝非都有水分子的知识。况且这也无法解释量词的共现名词表达抽象事物的情况,如“四趴报酬”中的“趴”是台湾闽南语中表示百分比(%)的量词,按照欧阳伟豪的理论,“四趴报酬”的意思应该是“以1%包含之数量的报酬个体为单位,共有四个单位”,但报酬并不是由任何个体组成的,因此这句话在欧阳伟豪的理论中是无法解读的。

何万顺(Her 2012a)因此提出最简洁的乘法理论:量词和数词形成乘法关系,共同量化名词;量词表达的是被乘数,其内涵就是一个数学上的值。这个值可以是数值(numerical),即可以用自然数表达,该数值可以是一个特定的自然数,如“个、双、打”分别表达了1、2、12这三个特定的自然数;该数值也可以是一个不特定的自然数,如“群、组、套”表达了某些个体组成的单位,因为这些单位是个体组成的,它们的量值一定可以用自然数表达,在不同的语境下“一群、一组、一套”包含的个体数量不同,同一语境下不同的单位包含的个体数量也可能不同。量词的值也可以是非数值性的,该量值可以是固定的,如标准化的度量衡单位“尺、升、斤”,也可以是不固定的,如“箱、瓶、杯”。

个体量词以个体本身为计量单位,其量值必然是数值,且必然是1。反之,

计量量词可以表达除了 1 以外的任何量值。何万顺、陈盈君和颜乃欣(He & Chen & Yen 2017)将以上的量词分类整理成表 10-1,并分析出以下两个用于分类的参数:量值是否为数值,是否为固定值。

表 10-1 量词分类表(C 指分类词,即个体量词;M 指计量量词)

数 值	固定值	n=1(如,本、颗、条、支)	C
		n=2,(如,对);n=12,(如,打)	M ₁
非数值	不定值	n>1(如,排、帮、叠)	M ₂
	固定值	公斤、公里	M ₃
	不定值	滴、袋、杯	M ₄

表达被乘数,以协助数词量化名词,可以说是量词最根本、最重要的认知语义功能,个体量词与计量量词在数学意义上的差别,也能解释这两者其他的语义差异。10.2.1 小节提及个体量词表达了共现名词指涉对象的本质属性,10.2.2 小节则提到个体量词只能跟可数名词共现。想象一个符合前者但不符合后者的个体量词 A,它表达了一个属性“液体性”,是所有液体名词如“水、酒、果汁”的指涉对象共有的本质属性;因为液体名词是典型的不可数名词,A 只能跟不可数名词共现。A 这样的个体量词是不存在的,正因为个体量词最核心的功能是表达被乘数 1,亦即以个体为计数单位,所以个体量词只能协助量化具有个体性的事物,也就是可数名词的指涉对象,这解释了 10.2.2 小节所描述的个体量词特性。^①

① 表达“液体性”的量词只可能是计量量词,例如“一滴水”的“滴”。但值得特别注意的是,虽然能与不可数名词共现的必然是计量量词,能与可数名词共现的并非都是个体量词。表 1 中的 M₁ 与 M₂ 这两种计量量词,因为与个体量词相同都表数值,因此也只能与可数名词共现,例如“一支筷子、一双筷子、一排筷子”中的三个量词都要求可数名词,但只有“支”是个体量词。然而,只能与不可数名词共现的必然是计量量词,更精准来说,必然是 M₃ 或 M₄。

至于个体量词表达“以个体为计数单位”的方式,就是彰显可数名词指涉的每一个个体必然拥有的、与其个体性有密切关联的属性,让听话者知道说话者欲以具有这个属性的个体而非群体为计数对象,这解释了 10.2.1 小节所描述的个体量词特性,以及 10.2.4 小节中个体量词彰显名词语义框架中某部分的特性。个体量词所表达的本质属性,或与共现名词有重叠的语义特征(见 10.2.3 小节),或彰显名词框架的一部分,都不是随机的,而是跟名词指涉对象的个体性有必然联系的属性。例如(5)提到的三个个体量词中,“尾”是鱼作为个体必有的部位;“条”作为一个形状,是通过清晰的边界才能定义的,而如同我们在 10.2.2 小节提到的,清晰的边界正是离散性、个体性的来源;“只”表达的有生性也是不可分割的一个生物体才有的性质。

最后,通用个体量词“个”更是直接表达了个体性,所以它能与大部分可数名词共现(Her & Hsieh 2010)。个体名词通过表达与个体性紧密链接的属性突出了个体的存在,并借此传达以此个体为数词计数单位的信息,这种运作方式正符合格林伯格(Greenberg 1990: 172)的描述:众多个体量词都只是表达“一”,或更精准地说是“乘以一”的方式而已。

10.4 量词的语法表现与句法结构

本节探讨的是个体量词与计量量词因语义差异而导致语法行为上的不同,以及量词这个词类在认知上可作为被乘数的共性所带来的语法共性,进而探讨类型学中四种语序的[数量名]短语所共有的句法结构。

10.4.1 两种量词的语义差异带来的语法差异

量词表达被乘数的语义功能,导致语法或形式层面产生了许多有趣的现

象,其中一个是个体量词的可省性。因为个体量词表达被乘数 1,这在数学上是一个冗余的信息,因为 1 乘以任何数都等于该数本身,加上个体量词重复名词语义因而在语义上也是冗余的,因此个体量词在许多量词语言中是可省的。“耶稣以五饼二鱼喂饱五千人”这个句子中,展示了三处省略个体量词的情况:该句若将个体量词完全补上,应该表达为“耶稣以五张饼二条鱼喂饱五千个人”(Her 2012b)。在汉语中,不同文体对个体量词省略的接受度不同,文言文的接受度更高(何万顺、林昆翰 2015);官话系统中不同方言的接受度也不同,何万顺和林昆翰(2015)收集电影《非诚勿扰》《非诚勿扰 2》中 282 句可以使用个体量词的台词,发现其中有 74 句省略了个体量词,如例(8)所示,比例高达 26%。最后,量词系统较不发达的语言经常可省略个体量词,如汉藏语系中的藏语和景颇语(蒋颖 2006: 18)。和个体量词相反,计量量词表达不等于 1 的被乘数,这在数学上是重要的信息,且计量量词对其所属的名词短语添加了新的语义,所以计量量词是不可省的,如例(9)所示。

(8) a. 我还真是一坏人。

b. 登一征婚广告。

c. 捐两器官。

d. 办垮过三公司。

(9) a. 一颗苹果 = 一苹果

b. 一箱苹果 \neq 一苹果

个体量词的可省性也意味着个体量词对数词和形容词而言是“透明”(transparent)的,量化个体量词的数词与修饰个体量词的形容词,其量化与修饰范围可以穿透个体量词直至名词,否则省略个体量词就会导致语义的改变;反之,计量量词相对于数词和形容词是“不透明”(opaque)的,量化与修饰计量量词的数词与形容词就只能针对它们起作用,其量化与修饰范围不及于名词

(Her & Hsieh 2010; Her 2012ab)。这导致个体量词与计量量词在语法上的一些差异(Her & Hsieh 2010; Her 2012ab)。第一,个体量词的形容词可以转为名词的形容词,但计量量词不允许这个转换,如(10)所示。第二,同一个名词短语中,个体量词的形容词与名词的形容词不能有语义冲突,但计量量词没有这个限制,如(11)所示。第三,同一个名词不能被两个“数词+个体量词”短语量化,因为倘若这两个短语表达不同的量化信息,便会在名词上发生语义冲突,若两个短语表达相同的量化信息,“数词+个体量词”短语量化的,如(12)所示。不过倘若同一个名词前面有两个个体量词,但其中一个与数词形成短语,另一个与形容词形成短语,这种情况是合语法的,因为这两个个体量词短语分别对名词产生量化与修饰的作用,在语义上既不冲突也不冗余;不过在汉语中,“形容词+个体量词”短语必须后加“的”才能成为名词修饰语,这种修饰语也可以独立于数词存在,且一样可省,省略后形容词转移到名词上,如(13)所示。^①

- (10) a. 一大颗苹果 = 一颗大苹果
b. 一大箱苹果 \neq 一箱大苹果

- (11) a. *一大颗小苹果
b. 一大箱小苹果

- (12) a. *一颗十个苹果
b. *一颗一个苹果

① 何万顺和谢祯田(Her & Hsieh 2010)、何万顺(Her 2012a)另外指出一个有趣的语法现象:在汉语中,若同一个名词前面出现个体量词与计量量词,不论它们前面是数词还是形容词,个体量词一定不可以出现在计量量词前面,亦即个体量词一定要比计量量词贴近名词,如“一箱十颗苹果; *一颗十箱苹果; *十颗一箱苹果”,“一箱大颗的苹果; *一颗大箱的苹果; *大颗的一箱苹果”。何万顺与谢祯田认为这代表计量量词在句法结构的位置上比个体量词高。

- c. 一箱十包苹果
 - d. 一箱十颗苹果 (one box of ten apples)
- (13) a. *一颗大颗苹果
- b. 一颗大颗的苹果
 - c. 大颗的苹果 → 大苹果

另外,个体量词在语义上的冗余性使得它们比较接近功能项(functional item),也具有功能项的另一个典型特征:属于封闭词类(closed class),难以创造新的个体量词。因此,穷举一种语言的个体量词是可能的(Her & Hsieh 2010;何万顺、林昆翰 2015)。反之,计量量词表达除了1以外的无限种可能的被乘数,且它们不需要表达与共现名词重叠的语义,所以它们比较接近词汇项(lexical item),形成一个开放词类(open class),因此创造新的计量量词是非常容易的(Her & Hsieh 2010)。

10.4.2 量词的句法结构

个体量词与计量量词的语义差异导致了两者在语法上的差异,其语义共性也应该在语法及其背后的句法结构上有所反映。一个合理的推测是:既然量词与数词之间有乘法关系,它们在句法结构上应该形成一个短语,在语序上也应该相邻,不被其他词打断。然而,“数词+量词+名词”的句法结构却是句法学界的一大争议,相关理论主要可以分成两类:一类认为这类短语在任何情况下都具有相同的句法结构,称为统一论(unified analysis);另一类认为这类短语在不同情况下会有不同的句法结构,称为分裂论(split analysis)。统一论又可分为两类:一类认为在短语“数词+量词+名词”中,量词先跟数词结合再跟名词结合,形成[[数词 量词]名词]的结构,如图1所示(图中之Num为数词,C/M为量词)(Her 2017b),我们称为左分支论(left-branching analysis);这是比较传

统的理论 (Greenberg 1990; Huang 1982; Li & Thompson 1981: 105 等)。另一类则认为量词先跟名词结合再跟数词结合, 形成 [数词 [量词 名词]] 的结构, 我们称为右分支论 (right-branching analysis), 如图 2 所示; 这是目前在形式句法学界主流的理论 (Cheng & Sybesma 1999; Borer 2005; Tang 2005; Simpson 2005; Huang et al 2009 等)。^①



图 1: 量词左分支结构示意图



图 2: 量词右分支结构示意图

分裂论则更为多元, 例如张宁 (Zhang 2011) 主张量词的两个次类所处的句法结构不同, 属个体量词类的服从右分支论, 属计量量词类的服从左分支论; 李旭平 (X. Li 2011) 则认为不同的句法结构会影响量词的诠释, 同一个有量词的短语 (如“三瓶酒”), 若做计量诠释 (measure/quantity reading), 为左分支结构, 若做计数诠释 (counting reading) 则为右分支结构。一些学者 (Jenks 2010; Little & Moroney & Royer 2022) 则认为, 有些量词语言采用左分支结构, 有些采用右分支结构。我们支持左分支论, 即所有量词语言在所有情况下, 都是量词和数词结合成短语, 再和名词结合成更高的短语。

10.4.2.1 左分支论的证据: 现代汉语

若以现代汉语这一最典型的量词语言来看, 只有左分支论可以解释量词相关的语法现象, 右分支论和分裂论都会遇到问题。分裂论无法解释量词许多稳定的且所有量词都展现的语法特性 (Her 2012b), 包括: 占据同一位置并形成

^① 请注意本章对统一论和分裂论的区分标准是: 是否认为所有量词都服从左分支论或右分支论。倘若有一个理论认为个体量词和计量量词所处的句法结构略有不同, 但都是先跟数词或先跟名词 (如 Cheng & Sybesma 1999) 结合, 该理论会被本章归为统一论。

严格的互补分布,见(2);都允许在正确语境下省略名词,见(14);都允许在正确语境下和名词一起省略,见(15);都允许后插“的”,见(16);都允许后插“半、多”,见(17)等。分裂论难以解释为何在不同句法位置的量词会共享这么多语法特性,只能将其归于偶然。更重要的是,分裂论无法反映数词与量词间稳定的乘法关系,这是在任何情况下都成立的(Her 2012b)。

(14) 他有三箱/本书,我有三本/箱。

(15) 他有一百本/箱书,我有三百。

(16) a. 一箱的书

b. 半颗的苹果(Her & Hsieh 2010)^①

(17) a. 两个半小时

b. 六打多玫瑰

相对于分裂论难以解释量词语法特性的稳定性,右分支论的问题是难以解释一些量词语法特性本身(Her 2012b)。如例(16)显示量词允许后插“的”,这对左分支论而言很好解释,因为[数词 量词]短语是名词的修饰语,因此自然能后插“的”;但大部分右分支论认为量词取名词为补足语(complement)(Her 2017b),因而无法解释为何“的”能插入中心词(head)与其补足语之间。

例(17)中位于量词后的“半、多”的诠释也较支持左分支论:“半、多”都根据前方量词的语义产生一个量值,并和[数词+量词]短语表达的量值形成加法

① 何万顺和谢祯田(Her & Hsieh 2010)发现个体量词与计量量词对后插“的”的接受度不同,计量量词可以自由后插“的”,但个体量词只有在[数词+量词]短语的计算复杂度(computational complexity)高时才可以接受后插“的”。提升[数词+量词]短语计算复杂度的方式包括:提升数词表达的数值,如“五百万只的鸭子”;使用结构较复杂的数词,如“八分之一颗的高丽菜”;使用形容词修饰量词,如“一大颗的苹果”等。[数词+量词]短语的计算复杂度极低时无法接受后插“的”,如“*一颗的苹果”。

关系,整个[[数词+量词]半/多]短语一同量化名词,因此(17a)“[两个半]小时”的意思是 $[(2 \times 1) + (0.5 \times 1)] = 2.5$ 个小时,(17a)“[六打多]玫瑰”的意思可以是 $[(6 \times 12) + (0.5 \times 12)] = 78$ 朵玫瑰^①。反之,右分支论会得出错误的量化结果,即“两[个半小时]”变成 $[2 \times (1 \times 0.5)] = 1$ 个小时(假设“半”与量词形成乘法关系;后面“多”例相同)或 $\{2 \times [1 + (0.5 \times 1)]\} = 3$ 个小时(假设“半”与量词形成加法关系;后面“多”例相同)，“六[打多玫瑰]”变成 $[6 \times (12 \times 0.5)] = 36$ 朵玫瑰或 $\{6 \times [12 + (0.5 \times 12)]\} = 108$ 朵玫瑰。

例(15)中量词可以和名词一起省略,看似支持右分支论,但何万顺(Her 2012b)指出量词只有在数词是乘法数词时才可以省略:

- (18) a. 他有一百本/箱书,我有三百。
b. *他有一百本/箱书,我有三。

倘若量词和名词可以一起省略的原因是因为它们形成短语,那这个省略不应该受数词的影响。何万顺认为量词在数词是乘法数词时可省的原因,跟例(14)中名词在[[数词+量词+名词]短语(简称“数量名短语”)中可省的原因一样:数量名短语具有[[数词×量词]名词]的结构,乘法数词与量词也具有类似的[[乘数词×位数词]量词]的结构。何万顺认为这两个结构中名词和量词都被乘法结构量化,所以都可以发生中心词省略(head-ellipsis),而这个理论是以左分支论为基础的。何万顺(Her 2017b)另外指出数量名短语中名词可以被提取(extraction),但[[量词+名词]不行,这是对右分支论不利的证据:

- (19) a. 书,他有三本 e 。
b. *本书,他有三 e 。

① “多”的意思是多出前方量词表达之量值,但不超过一半,因此“一打多”约指13到17个。

最后,右分支论有和分裂论一样的缺点:无法反映数词与量词间的乘法关系。反之,左分支论认为数词与量词形成一个短语,这恰好对应两者之间的乘法关系。

从上面的讨论来看,左分支论的优势似乎是无疑的。然而,张宁(Zhang 2011)指出左分支论也有两个缺点:第一,在左分支论中量词并没有c统制(c-command)名词,无法解释量词是怎么选择特定种类的名词的。第二,左分支论中数词与量词的形容词也不c统制名词,所以理论上它们是无法量化与修饰名词的,但我们在10.4.1节已知道个体量词的共现数词与形容词是可以量化与修饰名词的。针对这两个问题,何万顺(Her 2012b)在词汇功能语法(Lexical-Functional Grammar, LFG)的架构下提出理论解释左分支结构如何与这两个现象并存。另外,对于第一个问题,李艳惠(Li 2014)指出可以借由延伸投射(extended projection)的概念解决:简而言之,词汇性的中心词(lexical head)如名词可以和与其相关的功能性中心词(functional head)如限定符(determiner)形成延伸投射,这被用于解释为何动词取限定短语为论元,却可以对其内部的名词进行语义选择。总而言之,左分支论的优势是明显的,其劣势也可以在理论上获得解释,因此我们认为支持左分支论是目前最合理的选择。

10.4.2.2 左分支论的证据:量词语序的类型学

以上的论证都建立在现代汉语语料的基础上。从汉语历史以及其他量词语言的量词语序类型学来看,左分支论仍然是解释力最强的理论。贝罗贝(Peyraube 1998)整理出汉语历史上出现过的七种名词短语语序:〔数词+名词〕、〔名词+数词〕、〔名词₁+数词+名词₂〕^①、〔名词+数词+计量量词〕、〔名词+

① 在〔名词₁+数词+名词₂〕短语中,名词₂表达了名词₁所属的类别,如“羌十人”。

数词+个体量词]、[数词+计量量词+名词]、[数词+个体量词+名词] (例：一牛、虎一、羌十人、贝廿朋、马三匹、一杯羹、一株松)。另外，上古汉语也常使用所谓的反响量词(echo classifier)，亦即[名词₁数词 名词₁]的结构(蒋颖 2006: 106)，如“俘牛三百五十五牛，羊二十八羊”(张亚初 2001: 5. 2839)。蒋颖(2006)指出以上结构中名词₂和反响型量词在句法上是不是量词还有争议，何万顺(Her 2017b)便将它们并入[数词 名词]或[名词 数词 量词]结构中，并将以上结构合并为[数词 名词]、[名词 数词]、[数词 量词 名词]、[名词 数词 量词]四种结构。另外，何万顺(Her 2012b)指出[数词 名词]、[名词 数词]可以被分析为[数词 量词 名词]、[名词 数词 量词]省略个体量词的结果，这是现代汉语也允许的现象(见 10. 4. 1)。简而言之，汉语历史上其实只出现过两个数量名短语的语序：[数词 量词 名词]和[名词 数词 量词]。这两个语序的共同特征是数词与量词永远相邻，名词不会插入两者之间，这可以简洁地以左分支论解释：假设数词与量词形成一个数量短语，再由名词作为中心词与数量短语结合形成名词短语，这两个语序其实只是名词短语的中心词参数(head parameter)的两种可能性：中心词置前(head-initial)[名词[数词 量词]]与中心词置后(head-final)[[数词 量词]名词]的体现而已。

汉语数量名短语语序的历史符合格林伯格(Greenberg 1990: 185)经过艾肯瓦尔德(Aikhenvald 2000: 104 – 105)确认的观察：虽然数量名短语在数学上可能的语序有 $3! = 6$ 个，但在量词语言中只发现 4 个而已；除了汉语历史上出现过的[名词[数词 量词]]与[[数词 量词]名词]，数量短语本身也有两种中心词参数的可能性，和名词短语的两种中心词参数可能性形成四种可能语序，如表 2 所示(基于 Her 2017b, 有改动)。

表 10-2 人类语言中存在的数量名短语语序

	名词置于数词后	名词置于数词前
量词置于数词后	(A) [[数词 量词] 名词] 如,汉语、越南语、苗语、乌兹别克语	(B) [名词[数词 量词]] 如,泰语、缅甸语、日语、高棉语、马勒语(Mal, 南亚语系)
量词置于数词前	(C) [[量词 数词] 名词] 如,伊比比奥语(Ibibio, 尼日尔-刚果语系)、基里维纳语(Kiriwina 或 Kilivila, 大洋洲语族)	(D) [名词[量词 数词]] 如,路易西亚德群岛语(Louisiade Archipelago, 大洋洲语族)、景颇语、博多语(Bodo, 汉藏语系)

量词语言中只存在以上四个语序,不存在名词插入数词与量词间的语序,这是对跨语言的左分支论的有力支持。这项规律与格林伯格(Greenberg 1963)提出的共性 20(Universal 20)有关,该共性描述了名词与其三项修饰语:限定符、数词、形容词的语序。加上量词语言不存在名词中插语序的规律是由格林伯格(Greenberg 1990: 185)首先发现的,何万顺(Her 2017a)因而将这项规律命名为格林伯格共性 20A(Greenberg's Universal 20A;简称“20A”)。

(20) 格林伯格共性 20A

20A-1: 数词、量词、名词三者的语序,除了名词介于数词与量词之间的语序,其他语序都存在。

20A-2: 量词在数词后的语言,比量词在数词前的语言多。

20A-2 与格林伯格(Greenberg 1990: 292)发现的以下现象有关:在量词语言中,乘法数词的位数词语与数量短语的量词语序有“和谐”(harmonization)的倾向,亦即位数词在数词后的语言倾向将量词置于数词后,如汉语[[三百]个];位数词在数词前的语言倾向将量词置于数词前,如属尼日尔-刚果语系(Niger-Congo)的伊比比奥语(Ibibio)遵守量词置前语序[量词 数词 名词]

“量化名词”却不会。换言之,(21)其实不应该理解成“两粒橘子籽”,而是“两个是橘子籽的粒(状物)”,亦即所谓的“量词”其实是量化名词,而所谓的“量化名词”其实是量化名词的名词修饰语。类似英语的 *six pieces of paper*, 其中 *pieces* 的功能看似量词,但在句法上因为标复数,所以其实是量化名词(见第三节复数标记与个体量词功能相同的部分),而 *of paper* 则是 *pieces* 的修饰语 (Her 2017a)。何万顺和徐宏欣 (Her & Hsu 2022) 确认了尼日尔—刚果语系中疑似有 [量词+名词+数词] 语序的语言,其“量词”其实都是名词,解决了这些语言对 20A-1 的挑战。

至于以毛南语为代表的壮侗语和南亚语中看似为 20A-1 的反例,毛南语数量名短语的常态语序是 [数词+量词+名词], 如例(23)所示;只有数词 $d\epsilon u^{231}$ “一”可以出现在名词后,且 $d\epsilon u^{231}$ 也只能出现在名词后,如例(24)所示:(转引自 Her 2017a; C 为个体量词)

- (23) a. $ja^{42}/sa: m^{42}$ ai^l $z\epsilon n^l$
 2/3 C person
 ‘2/3 persons’
- b. $*ai^l$ $z\epsilon n^l$ $ja^{42}/sa: m^{42}$
 C person 2/3 (张景霓 2005)
- (24) a. ai^l $z\epsilon n^l$ ($d\epsilon u^{231}/*t\epsilon^{231}/*jit^{55}$)
 C person 1/1/1
 ‘1 persons’
- b. $*d\epsilon u^{231}$ ai^l $z\epsilon n^l$
 1 C person (张景霓 2005)

$d\epsilon u^{231}$ 的数词身份是很可疑的,因为:第一,除了 $d\epsilon u^{231}$ 以外,没有其他数词能出现在名词后。第二, $d\epsilon u^{231}$ 不能出现在其他数词分布的正常位置,即量

词前。第三, $d\epsilon u^{231}$ 是可省的, 如例(24a)所示, 但其他数词都不可省。第四, 毛南语有三个表数量 1 的数词: $d\epsilon u^{231}$ 、 $t\epsilon^{231}$ 、 jit^{55} , 其中 $d\epsilon u^{231}$ 和 $t\epsilon^{231}$ 是原生词, jit^{55} 是汉语借词, 但只有 $d\epsilon u^{231}$ 可以出现在名词后, 如例(24a)所示; 这不只是毛南语的特色, 所有看似违反 20A-1 的亚洲语言都只在表达数量 1 时才出现 [量词 名词 数词] 语序, 且只有使用表达数量 1 的原生数词时才会出现该语序, 而使用汉语借词就不会有这种现象^①(Her 2017a)。基于 $d\epsilon u^{231}$ 这样异于其他数词的特性, 张景霓(2005)和蒋仁萍(2007)都不认为 $d\epsilon u^{231}$ 是数词。张景霓(2005: 303)指出 $d\epsilon u^{231}$ 的词源是表“单独、只有”的形容词, 蒋仁萍(2007)更认定 $d\epsilon u^{231}$ 是形容词。毛南语的形容词确实在名词后, 如例(25)所示, 但何万顺(Her 2017a)指出这个理论无法解释为何 $d\epsilon u^{231}$ 只能出现在其他形容词后, 如例(26)所示。

(25) ai^l $z\epsilon n^l$ $da: i^2$

C person good

‘the good person’

(26) a. ai^l $z\epsilon n^l$ $vo\eta^l$ $d\epsilon u^{231}$

C person good INDEF.SG

‘a tall person’

b. $*ai^l$ $z\epsilon n^l$ $d\epsilon u^{231}$ $vo\eta^l$

*C person INDEF.SG good

何万顺(Her 2017a)认为 $d\epsilon u^{231}$ 其实是不定指限定符(indefinite

① 事实上, 毛南语禁止所有形式的数词“一”出现在常态的 [数词+量词+名词] 语序中。何万顺(Her 2012b)指出在量词语言中, 它有时可省, 理由跟个体量词可省的原因一样: 数词与量词形成乘法关系, 而在数学上不论是乘数 1 还是被乘数 1 都是冗余的(见 10.4.1 节)。何万顺(Her 2017a)另外提出一个类型学分类: 量词语言分为三类, 一类禁止省略(如闽语), 一类允许省略(如官话、粤语、吴语), 一类要求省略, 毛南语便属于最后一类。

determiner), 类似英语的不定冠词 a/an。他指出, $d\epsilon u^{231}$ 跟定指限定符 ka^2 ‘那’和 $na:i$ ‘这’成互补分布, 如例(27)所示, 支持这个假说。

- (27) a. ai^1 $z\partial n^1$ $vo\eta^1$ $ka^2/na:i$
C person tall that/this
‘that/this tall person’
- b. $*ai^1$ $z\partial n^1$ $d\epsilon u^{231}$ $ka^2/na:i$
C person INDEF.SG that/this
- c. $*ai^1$ $z\partial n^1$ ka^2 $d\epsilon u^{231}$
*C person DEF INDEF.SG

既然 $d\epsilon u^{231}$ 作为毛南语中唯一会导致[量词+名词+数词]语序的“数词”其实不是数词, 那该语言就不构成对 20A-1 的挑战了。曹子允(Cao 2023)以何万顺(Her 2017a)的论证检视亚洲所有看似违反 20A-1 的语言, 证实它们与[量词 名词 数词]语序共现的数词“一”都不是数词, 因此都不违反 20A-1。名词不会介于量词与数词间的共性至今仍未被挑战成功, 这是对统一左分支论跨语言普遍性的一大肯定。

10.5 数据库与类型学研究

研究团队在多年的量词研究中, 为了验证各项理论所收集的量词语言的相关数据, 逐步累积成为目前世界上规模最大的量词语言数据库“世界量词语言地图”(WACL: World Atlas of Classifier Languages)(详见: Her & Hammarström & Allasonnière-Tang 2022)。本节简述资料收集的过程及从中衍生的各种类型学研究。

10.3 节提到博雷尔(Borer 2005)认为复数标记与量词的功能相同;她更进一步宣称复数标记与量词不会共现,主张两者占据同样的句法位置(Borer 2005: 6, 10, 95)。何万顺(Her 2012a)欲验证复数标记与个体量词^①成互补分布的类型学主张,参考奇尔(Gil 2013)和哈斯普马特(Haspelmath 2011)的世界语言量词相关类型学调查,以两份研究重叠的 114 种语言为样本,调查其个体量词与复数标记的有无。在这 114 种语言中,有 84(73.7%)种语言只有个体量词或复数标记,符合互补分布;但也有 22(19.3%)种语言两者都有。何万顺(Her 2012a)指出这 22 种语言不一定违反复数标记与个体量词成互补分布的假说,因为它们有可能规定这两者不能在同一个名词短语中共现。然而,他却指出汉语和日语似乎确实允许这种情况,因而转而支持较弱版本的假说:一种语言的分类词系统和复数标记系统的发达程度成反比,违反复数标记与个体量词互补分布的语言很少见,且可以用语言演变解释。以汉语为例,这是一个分类词系统发达的语言,所以其复数标记系统衰弱,只能标人类名词(即“-们”),而且是选择性而非强制性标记;汉语允许分类词与复数标记共现的原因可能是因为它经历从综合语(synthetic language)向分析语(analytic language)演变的过程。另外,何万顺(Her 2012a)也根据量词的乘法理论(见 10.3 节)提出一个类型学假说:有量词的语言一定会有乘法数词。他通过交叉比对科姆里(Comrie 2008)对世界语言数词系统的调查以及奇尔(Gil 2013)的量词相关类型学调查证实这一假说:没有乘法数词的语言都没有量词,有量词的语言都有乘法数词。

何万顺等(Her et al. 2015)与何万顺、唐威洋和李孟璋(Her & Tang & Li 2019)收集了六个位于亚洲且贡献绝大多数量词语言的语言族群:汉语族、苗瑶语族、南亚语系、壮侗语族、藏缅语族、印度—亚利安语族(Sinitic, Miao-Tiao,

^① 何万顺(Her 2012a)强调只有个体量词和复数标记功能相同。见 10.3 节。

Austroasiatic, Tai-Kadai, Tibeto-Burman, Indo-Aryan) 的量词语言数据,建立了 SMATTI 数据库。何万顺等(Her et al. 2015)是第一份以统计学验证 10.4.2.2 节提及的格林伯格(Greenberg 1990: 292)发现之和谐律,即位数词语序和量词语序倾向一致的研究:在其样本的 194 个量词语言之中,189(97.4%)种语言遵守和谐律,只有 5 个违反。另外,何万顺等(Her et al. 2015)也解释了(20)介绍的 20A-2 共性:在 194 个样本语言中有 168(86.6%)种语言采用位数词后置语序,因此根据和谐律,量词后置的语言自然多于量词置前的语言。

何万顺、唐威洋和李孟璋将 SMATTI 数据库中量词语言的数量提升至 219 个(其地理分布参看:Her & Tang & Li 2019: 432, Figure 3),且再度验证了和谐律与 20A-2: 样本中有 213(97.3%)种语言服从和谐律,187(85.4%)种语言和谐后置(位数词和量词都后置)。另外,他们还验证了何万顺(Her 2012a)提出的假说:量词的存在预设乘法数词的存在。科姆里(Comrie 2013)的调查中有 87.8%(172/196)的世界语言有乘法数词,但何万顺、唐威洋和李孟璋(Her & Tang & Li 2019)的量词语言样本全部都有乘法数词,证实量词与乘法数词的存在在统计上显著相关。最后,何万顺、唐威洋和李孟璋指出其样本中藏缅语的特殊之处:所有和谐置前与违反和谐律的语言都是藏缅语,后者在地理分布上位于和谐置前语言与和谐后置语言的交界上(参看:Her & Tang & Li 2019: 440, Figure 5);他们因而猜测藏缅语原本是和谐置前的语言,其样本中所有违反和谐律的藏缅语都是语言接触的结果。

继承何万顺(Her 2012a)的研究,唐威洋和何万顺(Tang & Her 2019)是第一个以大规模量化形式验证复数标记与个体量词成互补分布的类型学假说的研究。最早观察到这现象的是格林伯格(Greenberg 1990)与桑切斯和斯罗宾(Sanches & Slobin 1973),所以唐威洋和何万顺(Tang & Her 2019)称这个假说为格林伯格—桑切斯—斯罗宾通则(Greenberg-Sanches-Slobin Generalization,

GSSG)。他们对格林伯格—桑切斯—斯罗宾通则之研究的一大贡献是参考凯博特和卡贝特(Kibort & Corbett 2008)区分构词语义学(morphosemantic)和构词句法学(morphosyntactic)上的复数标记,后者的一大特征是参与和名词以外其他单位如限定符、形容词、动词的呼应;他们主张只有构词句法学上的复数标记才是格林伯格—桑切斯—斯罗宾通则的描述对象,如此便排除了汉语和日语的复数标记,这两种语言不再是 GSSG 的例外。唐威洋和何万顺(Tang & Her 2019)根据不同语系和地区的语言占世界语言的比例收集了 400 种语言的数据以提升样本的代表性,统计出来的结果是:只有个体量词或构词句法性复数标记的语言有 247(61.8%)种,都没有的语言有 149(37.2%)种,只有 4(1.0%)种两者都有(其地理分布参看:Tang & Her 2019: 314, Map 3)。在 4 个看似违反 GSSG 的语言中,亚美尼亚语和匈牙利语的个体量词与构词句法性复数标记仍成互补分布,因为个体量词一定要跟数词共现,但这些语言的名词在数词出现时都是标单数。只有波斯语和土耳其语的个体量词与构词句法性复数标记使用情况比较复杂,需要进一步研究。

唐威洋和何万顺(Tang & Her 2017)试图证实和谐律可以推广至位数词语序、量词语序、名词语序三者之间。事实上,格林伯格(Greenberg 1990: 292)当初发现的和谐律称为通则 28(Generalization 28),并分为两个部分:第一个部分是对非量词语言,位数词语序倾向与名词相对于数词的语序和谐;第二个部分是对量词语言,他强调位数词语序是倾向与量词语序和谐,而非与名词相对于数量短语的语序和谐。但唐威洋和何万顺(Tang & Her 2017)认为位数词、量词与名词都是其所在短语的中心词,所以在一种语言中应该倾向有相同的中心词语序,并试图通过与唐威洋和何万顺(Tang & Her 2019)共享的 400 种语言样本,来验证该假说。统计的结果是:三个语序间都有高强度的正相关,但量词语序和名词语序的相关性弱于另两项和谐律;400 种语言样本中只有 33

(8.3%)种违反任一和谐律,其中大部分都是违反涉及名词语序的和谐律,如表3所示(基于Tang & Her 2017,有修改)。从三语序和谐置前、和谐置后与违反任一和谐律的语言的地理分布(参看:Tang & Her 2017: 542, Map 5),可以看出大部分违反和谐律的语言都分布在两种和谐语序语言的分布区的交界处,唐威洋和何万顺(Tang & Her 2017)因而推测这些和谐律的反例是由语言接触造成的。

表 10-3 400 种语言的位数词语序、量词语序、名词语序分布情形

		位数词置后	位数词置前	无位数词
量词置后	名词置后	90(22.5%)	0(0.0%)	0(0.0%)
	名词置前	16(4.0%)	0(0.0%)	0(0.0%)
量词置前	名词置后	1(0.3%)	2(0.5%)	0(0.0%)
	名词置前	2(0.5%)	21(5.3%)	0(0.0%)
无量词	名词置后	110(27.5%)	5(1.25%)	44(11.0%)
	名词置前	7(1.8%)	67(16.8%)	35(8.8%)

注:表格中用阴影表示违反和谐律的语言占比。

我们在多年研究中累积的量词语言相关数据,其成果便是本节开头提及的“世界量词语言地图”(WACL)。该数据库依靠乘法理论给定的量词精确定义,收集了3338个世界语言的数据,其中723(21.7%)种语言是量词语言(其地理分布参见:Her & Hammarström & Allasonnière-Tang 2022: 158, Figure 1)。何万顺等(Her & Hammarström & Allasonnière-Tang 2022)通过WACL验证过往文献对量词语言的描述,发现以下两方面的异同点:第一,量词是比较稀少的现象,这与奇尔(Gil 2013)的统计结果大体相合,但其世界语言样本中有35.0%

的量词语言,比例略高于 WACL 的量词语言比例,这有可能是因为奇尔(Gil 2013)的样本偏向量词语言,或因为他对量词的定义与 WACL 不同。第二,量词语言主要分布于亚洲,这符合尼科尔斯(Nichols 1992: 200)的统计结果,他的描述是在亚洲之外,少数**量值语言**分布在环太平洋区域,尤其是北亚沿海、大洋洲及美洲西岸。何万顺等(Her & Hammarström & Allasonnière-Tang 2022)的研究添加了一些细节:欧洲和非洲有少量量词语言;在环太平洋区域,量词语言主要集中于 Glottolog 数据库定义的“巴布尼西亚语区”(Papunesia),在澳洲则罕见;各大洲量词语言的数量排序为:亚洲>美洲>大洋洲>非洲>欧洲,量词语言按比例排序则是亚洲>美洲>欧洲>大洋洲>非洲,亚洲和美洲在这两方面都名列前茅,大洋洲的量词语言比例则很低,这和宣称大洋洲量词语言很常见,胜过其他地区的先行研究不同(Nichols 1992; Sinnemäki 2019)。最后,何万顺等(Her & Hammarström & Allasonnière-Tang 2022)创新性地统计了各语系的量词语言比例:在 WACL 内 203 个语系中,56 个语系有量词语言,其中只有 22 个语系内部有 50% 以上的语言是量词语言;有些语系完全由量词语言构成,它们都分布在亚洲和美洲,合乎前面提及的各大洲排序。

最后,我们在研究过程中收集的资料并不只限于涉及量词,还包括与量词相关的其他类型学特征,如名词分类系统的语法性别和名词类的分布(Allasonnière-Tang et al. 2021),以及乘法数词的位数词语序。后者是我们研究团队目前关注的重点:最新的研究(Her et al. 2024)整理了 4 099 个世界语言的位数词语序(其地理分布参见:Her et al. 2024: 2, Figure 1)。可以发现,世界语言目前主要采取位数词置后的语序,但位数词在**数词**前的语言也不少。引人注目的是,这两种语言的地理分布非常不均衡,后者几乎集中在非洲,其余**几大洲陆地地区的数词语序,几乎都被位数词置后的语序占据**。何万顺和唐威洋(Her & Allasonnière-Tang 2022)提出了一系列假说试图解释这样的分布情

形,涉及乘法数词的产生和演化过程。相信在探究这些假说的过程中,我们能对计数、量化这些基本认知能力在语言中的反映,在人类历史上的发展历程,以及以上两点是如何受到各种自然与社会因素影响的,形成更深刻完整的认识。

10.6 结论

本章综合了第一作者与团队十余年来量词相关研究的成果,主要分成三个部分:第一,在语义上,量词与数词形成乘法关系并表达被乘数,亦即计量单位的量值,和数词一同量化名词。量词可根据所表达的量值区分为个体量词及计量量词,个体量词只能表数值 1,计量量词可有表数值 1 以外所有可能的量值。这一差异导致其他的语义差异,即个体量词通过彰显名词提供的认知框架,突出名词指涉的个体与个体性相连的本质属性,进而表达以个体为计量单位的信息,亦即被乘数 1,所以个体量词的语义是名词语义的一部分,且只能跟可数名名词共现。反之,计量量词没有这些限制,所以它有自己独特的语义,且不限制共现名词的可数性。

第二,在语法上,个体量词具有数学和语义上的冗余性,导致它的可省性及对数词与形容词量化与修饰范围的透明性,计量量词则是不可省的,且会阻断数词与形容词的量化与修饰范围。但这两者都是与数词形成乘法关系的量词,所以必须和数词相邻,在句法上形成短语。这项左分支论涉及句法学界的争议,本章列举了现代汉语的语料、汉语史研究中的量词相关资料、以及提供了跨语言的类型学资料的支撑。

第三,我们为验证相关理论收集了量词语言大量的数据,得以建立起目前世界上最大的量词语言数据库“世界量词语言地图”(WACL)。我们证实了以下类型学事实:个体量词与构词句法性复数标记成互补分布,极少出现在同一

语言中,即使有也基本上不会出现在同一名词短语中,只有零星的有争议的例外。量词的存在,预设乘法数词的存在,没有乘法数词的语言不会有量词。乘法数词中的位数词语序、数量短语中的量词语序,名词短语中的名词语序,这三者在同一种语言中语序倾向相同,此和谐律的例外大多涉及名词语序,且应该是由语言接触导致。量词语言主要分布于亚洲和美洲,这两个大洲不论是量词语言数量还是比例都是最高的,且在世界量词语言地图中,只存在量词语言的语系,也都在这两个洲;论量词语言的数量,大洋洲排名第三,但论量词语言的比例,欧洲反而升至第三,这是过去类型学研究没有发现的。

量词是名词量化的工具之一,要理解量词必须理解数词。我们对这两者的研究已有较多的成果,相信学界对名词量化结构的理解必会更加深刻,我们对量化这一认知能力与语言和其他因素的关联将会有更完整的认识。