

量词语言类型学研究的新视野

何万顺^{1, 2} 梁永平²

¹台湾东海大学外国语文学系 ²台湾政治大学语言学研究所

摘要

在量词语言中，数词量化名词时需要量词作为中介。量词作为句法上一个独立的词类，在语意上可以分成个体量词与计量量词，前者以个体为计量单位，且在语意上彰显名词的某些固有特征，而后者则指定某一量值作为计量单位。作者与研究团队近期在量词的研究上有一些重要的发现，扩展了量词类型学研究的视野。在认知上，量词与数词形成乘法关系，量词为被乘数，可根据其值区分个体量词与计量量词，前者彰显名词指涉个体的某个本质特征，因而只能表达数值 1，后者则可能表达 1 以外的数值或非数值，因此可分为四个次类。数词与量词的乘法关系也反映在句法上，二者先形成短语才与名词结合，这一左分支结构的普遍性有许多语法和类型学上的事实支持，如 [数量名] 短语在自然语言中只存在四种语序，名词不得介入数词与量词之间。为进行量词语言的类型学研究，我们收集了大量的相关信息，建立了目前世界最大的量词语言数据库 World Atlas of Classifier Languages (WACL)，本文简介了在此基础上的部分类型学研究，聚焦于位数词、量词与名词在词序上有趋近一致的特性。最后呈现我们从量词研究扩展至数词研究的初步成果。

一、前言

量词 (classifier) 是在包含汉语的部分语言中，在数词量化名词时伴随出现的词类。量词可以根据语意分成两类：个体量词 (sortal classifier) 及计量量词 (mensural classifier)。个体量词选择具有特定语意特征的名词，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。例如例 (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，例 (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样的位置，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间，如例 (2) 所示：

- (2) a. *一根箱香蕉
- b. *一箱根香蕉

以上的现象不仅出现在现代汉语中，Greenberg (1975:30) 早已观察到同一个语言的个体量词与计量量词之语序会相同，这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，称之「量

词」。然而，个体量词与计量量词在语意上的关系却较为复杂，两者的语意有共性存在，但也有重要的差异。这复杂的关系在语法层面所反映出的各种效应，可帮助我们辨识量词及区分其下的这两个次类。我们认为这个议题是量词研究的基础，在厘清之后才可能进一步拓展视野。

本文作者所属的研究团队十余年来关注量词的研究，本文整理了部分重要的研究成果，展现在量词类型学上的若干进展。本文的结构如下：第二节先列举个体量词与计量量词的语意差异。第三节聚焦于个体量词与计量量词的语意共性，指出量词的语意功能是表达乘法关系中的被乘数。第四节讨论量词的语法表现与句法结构。第五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。第六节总结本文。

二、个体量词与计量量词的语意差异

本节说明个体量词与计量量词的语意差异，分为四个面向：本质属性与偶然属性、语意集合论的差异、可数/不可数之分、个体量词独有的彰显功能。

2.1 本质属性与偶然属性

戴浩一与王连清 (Tai and Wang 1990:38) 曾对个体量词与计量量词的语意差异提供了清楚的说明：「个体量词透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物恒久拥有的；反之计量量词没有将名词分类的功能，而是表达名词指涉物的量。」这也是个体量词又称为「分类词」的原因 (何万顺、林昆翰 2015)。何万顺与谢祯田 (Her and 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意：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物的本质属性，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」的指涉物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

2.2 语意集合论

既然个体量词表达了共现名词指涉物必然拥有的本质属性，计量量词则否，我们就可以运用集合论的概念，更简洁地阐释两者的语意差别：假设一个量词的语意特征的集合是 X，其共现名词的语意特征的集合是 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 and Lai 2012)。换言之，个体量词虽有其语意特征，但对其所属的名词短语之语意并没有任何贡献，

只是重复或强调了名词的部分语意而已；反之，计量量词必然会对其所属的名词短语增添新的语意。

2.3 可数名词与不可数名词之分

个体量词表达名词指涉物的本质属性这一点和一个争论已久的议题相关，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，亦即指涉有清楚边界 (bounded) 之离散性 (discrete) 物体的名词共现，而计量量词则没有这个限制 (Her and 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，Chierchia (1998) 与戴浩一 (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，且必须透过量词才能被数词量化；他们因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)，需要透过量词在其指涉物上划出边界以定义出单位，才能被数词量化。

何万顺与谢祯田 (Her and Hsieh 2010) 指出这个理论的问题：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的语感的；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，这是透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」之指涉物的本质属性。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，有一部份的名词本身指涉的就是具有个体性的物体，不需要透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。

2.4 个体量词作为语意彰显的工具

既然个体量词的语意仅仅是共现名词部分语意的重复而已，它在语意上就是冗余的。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。何万顺与谢祯田 (Her and Hsieh 2010)、何万顺 (Her 2012ab) 及何万顺与赖宛君 (Her and 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) 名词语意的功能。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」这个框架在认知中是由许多不同的特征组成的，这些特征透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) 突显了鱼长条形的特征，(5c) 则点出了鱼的有生性 (animacy)。

个体量词彰显名词语意的功能解释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意特征，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意，分类仅仅是副作用而已 (Her 2012a)。

三、量词在乘法上的认知语意共性

虽然个体量词与计量量词有语意差异，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意功能：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 Greenberg (1990[1972]: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 Landman (2004)、Borer (2005) 和欧阳伟豪 (Au Yeung 2005, 2007) 等学者的启发。

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。Landman 的贡献在于将量词语意和乘法概念相连，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」替换成语意相同的名词 three-boy group/「三人小组」，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 1」的概念，亦即以 一个名词指涉物为计数的单位：

- (6) a. four **times one** three-boy group → four three-boy groups
- b. 四**乘**一个三人小组 → 四**个**三人小组

Borer (2005) 抱持着比 2.2 提到的 Chierchia (1998) 与戴浩一 (Tai 2003:312) 更激进的想法：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，名词本身的语意都是不可数的，只有依靠量词或复数标记（简称复标）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 只要加上复标-s 就可以获得可数诠释，而典型的可数名词如 dog 只要不加复标就会获得不可数诠释，如 There is dog on the floor。Borer 的理论已在 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，不若英语允许复标与不可数名词共现。当然，英语也是有可数/不可数名词之分的语言，因为不可数名词只有在语境清楚暗示计量单位，以及将不可数名词诠释成种类这两种特殊情况下才能加上复标（Scontras 2022），而可数名词则没有这些限制。但 Borer 的贡献在于指出英语的复标发挥了和量词相同的功能，尽管如同何万顺（Her 2012a）所说，Borer 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才和复标发挥相同的功能，即和可数名词共现并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）明确地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」的意涵，却给出了一个过于复杂的定义：他认为量词的语意是「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

(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」的指涉物的数量。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）指出其理论的问题：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 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」是台湾华语中表示

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 $[3 \times 100]$ 的数学关系，其中「百」是位数词，「三」是乘数词。

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,只是在不同语境下「一群、一组、一套」包含的个体数量不同,甚至在同一语境下不同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」。

个体量词是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、陈盈君与颜乃欣(Her, Chen, and Yen 2017)将以上的量词分类整理成以下表格,并分析出两个参数:是否是数值、是否为固定值。

表1:量词分类表(C指分类词,即个体量词;M指计量量词)

Numerical	Fixed	n = 1 e.g., <i>ben</i> (本), <i>ke</i> (顆), <i>tiao</i> (條), <i>zhi</i> (隻)	C
		n = 2 e.g., <i>duel</i> (pair 對); n = 12 e.g., <i>da</i> (dozen 打)	M ₁
	Variable	n > 1 e.g., <i>pai</i> (row 排), <i>bang</i> (gang 幫), <i>die</i> (stack 疊)	M ₂
Non-numerical	Fixed	e.g., <i>gongjin</i> (kilogram 公斤), <i>gongli</i> (kilometer 公里)	M ₃
	Variable	e.g., <i>di</i> (drop 滴), <i>dai</i> (bag 袋), <i>bei</i> (cup 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、最重要的认知语意功能,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意差异。2.1提及个体量词表达了共现名词指涉物的本质属性,2.2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」的指涉物共有的本质属性;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物,这解释了2.2描述的个体量词特性。²

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都必然拥有的、和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2.1所描述的个体量词

² 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「一枝筷子、一双筷子、一排筷子」中的三个量词都要求可数名词,但只有「枝」是个体量词。然而,只能与不可数名词共现的却必然是计量量词,更精准的说,必然是M₃或M₄。

特性，以及 2.4 中个体量词彰显名词语意框架中某部分的特性。个体量词所表达的本质属性，或与共现名词重迭的语意特征（见 2.3），或所彰显名词框架的一部分，都不是随机的，而是跟名词指涉物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」是鱼个体必有的部位；「条」做为一个形状，是透过清晰的边界才能定义的，而如同我们在 2.2 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 and Hsieh 2010）。个体名词透过表达与个体性紧密链接的属性突显了个体的存在，并藉此传达以此个体为数词计数单位的讯息，这种运作方式正符合 Greenberg (1990[1972]:172) 的描述：众多个体量词都只是表达「一」，或更精准的说是「乘一」的方式而已。

四、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意差异而导致在语法行为上的不同，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。

4.1 两种量词的语意差异带来的语法差异

量词表达被乘数的语意功能导致语法或说形式层面许多有趣的现象，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，加上个体量词因重复名词语意而在语意上也是冗余的，因此个体量词在许多量词语言中是可省的。著名的圣经故事「耶稣以五饼二鱼喂饱五千人」展示了三个省略个体量词的例子：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，何万顺与林昆翰（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意，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 ≠ 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意的改变；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作用，其量化与修饰范围不及于名词 (Her and 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 and Hsieh 2010; Her 2012ab)：首先，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。其次，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意冲突，但计量量词没有这个限制，如 (11) 所示。最后，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意冲突，若两个短语表达相同的量化信息，又会产生冗余性；相反地，同一个名词是可以被两个 [数词-计量量词] 短语，或一个 [数词-计量量词] 短语与一个 [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，在语意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。³

- (10) a. 一大颗苹果 = 一颗大苹果
 b. 一大箱苹果 ≠ 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
 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

³ 何万顺与谢祯田 (Her and Hsieh 2010) 与何万顺 (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上的位置比个体量词还高。

c. 大颗的苹果 → 大苹果

另外，个体量词在语意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，穷举一个语言的个体量词是不可能的（Her and 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迭的语意，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 and Hsieh 2010）。

4.2 量词的句法结构

个体量词与计量量词的语意差异导致了两者在语法上的差异，其语意共性也应该在语法以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（Num 为数词，C/M 为量词；图 2 同），我们按照汉语的语序称为左分支论（left-branching analysis）；这是比较传统的理论（Greenberg 1990[1975]; Huang 1982; Li and 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；这是较新且目前在形式句法学界较主流的理论（Cheng and Sybesma 1999; Borer 2005; Tang 2005; Simpson 2005; Huang et al 2009 等）。⁴

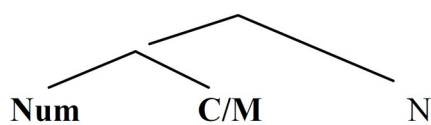


图 1：量词左分支结构示意图

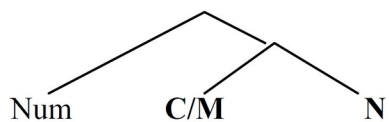


图 2：量词右分支结构示意图

⁴ 请注意本文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 and Sybesma 1999）结合，该理论会被本文归为统一论。

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。Jenks (2010) 与 Little, Moroney, and Royer (2022) 等学者则认为有些量词语言采左分支结构，有些采右分支结构。我们支持统一左分支论，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。

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所有量词都展现的语法特性 (Her 2012b)，包括：占据同一位置并成严格的互补分布 (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 and Hsieh 2010)⁵

(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语 (complement) (Her 2017b)，因而无法解释为何「的」能插入中心词 (head) 与其补语间。

⁵ 何万顺与谢祯田 (Her and 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」。

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意产生一个量值，并和 [数词 量词] 短语表达的量值形成加法关系，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⁶。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 \times 量词] 名词] 的结构，乘法数词与量词也具有类似的 [[乘数词 \times 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*。

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庸置疑的。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 4.1 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 (A. Li 2014) 指出可以藉由延伸投射 (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

⁶ 「多」的意思是「多出不满前方量词表达之量值的量」，(0.5 \times 12)意指多出一打的一半的量，这是合乎「多」语意的情况。

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意选择。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。

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中量词语序的类型学来看, 左分支论仍然是解释力最强的理论 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]⁷、[名词 数词 计量量词]、[名词 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型量词 (echo classifier), 亦即 [名词₁ 数词 名词₁] 的结构 (蒋颖 2006:106), 如「俘生三百五十五牛, 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 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 Greenberg (1990[1972]:185) 经过 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 (Her 2017b)。

表 2: 自然语言中存在的数量名短语语序

⁷ 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」。

	<i>N-final</i>	<i>N-initial</i>
<i>C/M-final</i>	(A) [[Num C/M] N] e.g., Chinese, Vietnamese, Hmong, Miao, Uzbek	(B) [N [Num C/M]] e.g., Thai, Burmese, Japanese, Khmer, Mal (Austro-Asiatic)
<i>C/M-initial</i>	(C) [[C/M Num] N] e.g., Ibibio (Niger-Congo), Kiriwina (a.k.a. Kilivila, Oceanic)	(D) [N [C/M Num]] e.g., Louisiade Archipelago (Oceanic), Jingpho, Bodo (Tibeto-Burman)

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 Greenberg (1963) 提出的普遍律 20 (Universal 20) 有关，该普遍律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 Greenberg (1990[1972]:185) 首先发现的，何万顺 (Her 2017a) 因而将这项规律命名为 Greenberg 普遍律 20A (Greenberg’s Universal 20A；简称 20A)：

(20) Greenberg 普遍律 20A

1. 数词、量词、名词间除了名词介于数词与量词间的语序，其他语序都真实存在。
2. 量词在数词后的语言比量词在数词前的语言多。

20A-2 与 Greenberg (1990[1972]: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序与数量短语的量词语序有「和谐 (harmonization)」的倾向，亦即位数词置后的语言倾向将量词置于数词后，如汉语 [[三百] 个]；位数词置前的语言倾向将量词置于数词前，如属尼日-刚果语系 (Niger-Congo) 的 Ibibio 语遵守量词置前语序 [量词 数词 名词] (Greenberg 1990[1972]:185)，且其乘法数词是位数置前，如 *ikie iba*, *ikie ita* (100×2 , 100×3 ；Chan 2015)。这项「和谐律」也可作为支持跨语言左分支论的证据，因为它可以用数词与量词形成表乘法关系的短语，和乘法数词一样，且一个语言应倾向使用同一个中心词语序来解释 20A-2 与和谐律的关系是，大部分的语言采用位数置后的乘法数词 (Her et al. 2015)，因此和谐律会使大部分量词语言将量词摆在数词后。

20A-1 及和谐律对左分支论的普遍性提供强力的支持。然而，一些研究主张有少数语言确实违反了 20A-1。这些语言似乎存在 [量词 名词 数词] 语序，且可以分成两部分讨论 (Her and Hsu 2022; Cao 2023)：一部分是非洲尼日-刚果语系的语言，以 Ejagham 语为代表；另一部分是亚洲壮侗语系和南亚语系的语言，以毛南语为代表。Watters (1981) 分析 Ejagham 存在 [量词 名词 数词] 语序，如例

(21) 所示：

- | | | |
|-------------------|----------------------|--------------|
| (21) <i>à-múg</i> | <i>i- çkúú</i> | <i>á-bá'</i> |
| 6- CL:small.round | GEN.6 19-orange.seed | 6-two |

‘two orange seeds’ (Watters 1981: 310)⁸

然而，Kihm (2005) 指出 Ejagham 的所谓「量词」在语法上比较像名词，因为它们可以标名词类 (noun class)，如例 (21) 的「量词」à-múg 上有第六名词类的标记 à-，这是非洲语言中常见的类似欧洲语言语法性别的语法标记 (Allasonnière-Tang et al. 2021)。Watters (1981:310) 进一步注意到在 Ejagham 所谓的数量名短语中，数词呼应 (agree) 的是「量词」的名词类而非所谓量化名词的名词类，这在例 (21) 中也可看见。何万顺 (Her 2017a) 认为这意味着 Ejagham 所谓的量词其实才是被数词量化的名词。他指出 Ejagham 和英语一样利用属格标记 (genitive marker) 建构属格结构 (construct genitive construction)，但和英语的语序相反，Ejagham 将所有物名词短语 (possessed NP) 置于属格标记前，所有者名词短语 (possessor NP) 则置后，如例 (22) 所示：

(22) *é-kpìn* *i* *á-tém*
 5-life GEN5 2-father
 ‘fathers’ life’ (Watters 1981:354 [61a])

基于 Ejagham 所谓的数量名短语也有属格标记，何万顺 (Her 2017a) 认为其实该语言的「量词」与名词的句法关系就是属格结构中所有物名词短语和所有者名词短语的关系。Kiebling (2018:51) 也认为 Ejagham 的数量名短语是从属格结构演变来的，虽然他认为该语言的「量词」是真正的量词而非名词。何万顺 (Her 2017a) 根据 Ejagham 的「量词」还是会标记名词类判断它们仍然是名词，并建构了该语言「数量名短语」的句法结构，假设数词是中心词，取属格标记投射的限定短语 (DP) 为补语，属格标记本身又取所有物名词短语为补语，所有者名词短语为定语 (specifier)，如图 3 所示：

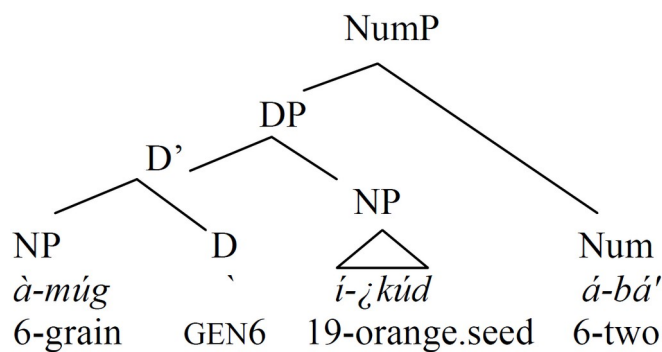


图 3：Ejagham 「数量名短语」句法结构

这样的句法结构能解释 Ejagham 「数量名短语」中名词类呼应的情形：属格标记呼应其补语之中心词的名词类，数词又呼应其补语之中心词的名词类，因此在这个短语中，「量词」、属格标记和数词才会标记同样的名词类，但「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

⁸ 这笔记语的标音符号和逐词翻译批注参考 Kihm (2005:497)。

橘子籽的粒（状物）」，亦即所谓的「量词」其实才是量化名词，而所谓的「量化名词」其实是量化名词的名词修饰语。这例子类似英语的 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（见第三节复标与个体量词功能相同的部分），而 of paper 则是 pieces 的修饰语（Her 2017a）。何万顺与徐宏欣（Her and Hsu 2022）确认了尼日-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 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 1 dɛu²³¹ 可以出现在名词后，且 dɛu²³¹ 也只能出现在名词后，如例（24）所示（张景霓 2005，引自 Her 2017a；C 为个体量词）：

- (23) a. ja⁴²/sa:m⁴² ai^l zən^l
 2/3 C person
 ‘2/3 persons’
- b. *ai^l zən^l ja⁴²/sa:m⁴²
 C person 2/3
- (24) a. ai^l zən^l (dɛu²³¹/*tɔ²³¹/*jit⁵⁵)
 C person 1/1/1
 ‘1 persons’
- b. *dɛu²³¹ ai^l zən^l
 1 C person

dɛu²³¹的数词身分是很可疑的，因为以下的原因：第一，除了 dɛu²³¹ 以外，没有其他数词能出现在名词后。第二，dɛu²³¹ 不能出现在其他数词分布的正常位置，即量词前。第三，dɛu²³¹ 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：dɛu²³¹、tɔ²³¹、jit⁵⁵，其中 dɛu²³¹ 和 tɔ²³¹ 是原生词，jit⁵⁵ 是汉语借词，但只有 dɛu²³¹ 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 的亚洲语言都只在使用数词 1 时才出现[量词 名词 数词] 语序，且只有使用原生数词 1 时才会出现该语序，使用来自汉语的数词 1 时不会有这种现象⁹（Her 2017a）。基于 dɛu²³¹ 这样异于其他数词的特性，张景霓（2005）和蒋仁萍（2007）都不认为 dɛu²³¹ 是数词。张景霓（2005:303）指出 dɛu²³¹ 的词源是表「单独、只有」的形容词，蒋仁萍（2007）更认定 dɛu²³¹ 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 dɛu²³¹ 只能出现在其他形容词后，如例（26）所示。

⁹ 事实上，毛南语禁止所有形式的数词 1 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，数词 1 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 4.1）。何万顺（Her 2017a）另外提出一个类型学分类：量词语言分为三类，一类禁止省略数词 1（如闽语），一类允许省略数词 1（如官话、粤语、吴语），一类要求省略数词 1，毛南语便属于最后一类。

(25) *ai^l* *zən^l* *da:i²*
 C person good
 ‘the good person’

(26) a. *ai^l* *zən^l* *voŋ^l* *dɛu²³¹*
 C person good INDEF.SG
 ‘a tall person’
 b. **ai^l* *zən^l* *dɛu²³¹* *voŋ^l*
 C person INDEF.SG good

何万顺 (Her 2017a) 认为 *dɛu²³¹* 其实是不定指限定符 (indefinite determiner), 类似英语的不定冠词 a/an。他指出 *dɛu²³¹* 跟定指限定符 *ka²* 「那」和 *na:i* 「这」成互补分布, 如例 (27) 所示, 支持这个假说。

(27) a. *ai^l* *zən^l* *voŋ^l* *ka²/na:i*
 C person tall that/this
 ‘that/this tall person’
 b. **ai^l* *zən^l* *dɛu²³¹* *ka²/na:i*
 C person INDEF.SG that/this
 c. **ai^l* *zən^l* *ka²* *dɛu²³¹*
 C person DEF INDEF.SG

既然 *dɛu²³¹* 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 1 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普遍律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。

五、数据库与类型学研究

本研究团队在多年的量词研究中, 为了验证各项理论收集了许多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: World Atlas of Classifier Languages (WACL) (详见 Her, Hammarström, and Allasonnière-Tang 2022)。本节简述资料收集的过程及从中衍生的各样类型学研究。

第三节提到 Borer (2005) 认为复标与量词的功能相同; 她更进一步宣称复标与量词不会共现, 主张两者占据同样的句法位置 (Borer 2005:6, 10, 95)。何万顺 (Her 2012a) 欲验证复标与个体量词¹⁰成互补分布的类型学主张, 参考 Gil (2011) 和 Haspelmath (2011) 的世界语言量词相关类型学调查, 以两份研究重迭的 114 个语言为样本, 调查其个体量词与复目标有无。在这 114 个语言中,

¹⁰ 何万顺 (Her 2012a) 强调只有个体量词和复标功能相同。见第三节。

有 84 (73.7%) 个语言只有个体量词或复标，符合互补分布；但也有 22 (19.3%) 个语言两者都有。何万顺 (Her 2012a) 指出这 22 个语言不一定违反复标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：一个语言的分类词系统和复标系统的发达程度呈反比，违反复标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，所以其复标系统衰弱，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标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 (见第三节) 提出一个类型学假说：有量词的语言一定会有乘法数词。他透过交叉比对 Comrie (2008) 对世界语言数词系统的调查以及 Gil (2011) 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、唐威洋与李孟璋 (Her, Tang and Li 2019) 收集了六个位于亚洲，贡献绝大多数量值语言的语族：汉语族、苗瑶语系、南亚语系、壮侗语系、藏缅语族、印度-亚利安语族 (Sinitic, Miao-Tiao, 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 4.2.2 提及的 Greenberg (1990[1972]: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%) 个语言遵守和谐律，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 普遍律：在 194 个样本语言中有 168 (86.6%) 个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、唐威洋与李孟璋将 SMATTI 数据库中量词语言的数量提升至 219 个 (其地理分布参见 Her, Tang and Li 2019:432, Figure 3)，且再度验证了和谐律与 20A-2：样本中有 213 (97.3%) 个语言服从和谐律，187 (85.4%) 个语言和和谐后置 (位数词和量词都后置)。另外，他们还验证了何万顺 (Her 2012a) 的量词之存在预设乘法数词之存在的假说：Comrie (2013) 的调查中有 87.75% (172/196) 的世界语言有乘法数词，但何万顺、唐威洋与李孟璋 (Her, Tang and 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、唐威洋与李孟璋指出其样本中藏缅语的特殊之处：所有和谐前置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前置语言与和谐后置语言的交界上 (参见 Her, Tang and Li 2019:440, Figure 5)；他们因而猜测藏缅语原本是和谐前置的语言，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，唐威洋与何万顺 (Tang and Her 2019) 是第一个以大规模量化形式验证复标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 Greenberg (1990[1972]) 与 Sanches and Slobin (1973)，

所以唐威洋与何万顺 (Tang and Her 2019) 称这个假说为 Greenberg-Sanches-Slobin 通则 (Greenberg-Sanches-Slobin Generalization, GSSG)。他们对 GSSG 之研究的一大贡献是参考 Kibort and Corbett (2008) 区分构词语意学 (morphosemantic) 和构词句法学 (morphosyntactic) 上的复标, 后者的一大特征是会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标才是 GSSG 的描述对象, 如此便排除了汉语和日语的复标, 这两个语言不再是 GSSG 的例外。唐威洋与何万顺 (Tang and 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 个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学性复目标语言有 247 (61.6%) 个, 都没有的语言有 149 个 (37.3%), 只有 4 (1%) 个两者都有 (其地理分布参见 Tang and 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学性复标仍成互补分布, 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学性复标使用情况比较复杂, 需要进一步研究。

唐威洋与何万顺 (Tang and 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, Greenberg (1990[1972]: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。但唐威洋与何万顺 (Tang and Her 2017) 认为位数词、量词与名词都是其所在短语的中心词, 所以在一个语言中应该倾向有相同的中心词语序, 并试图透过与唐威洋与何万顺 (Tang and Her 2019) 共享的 400 个语言样本验证该假说。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 个语言样本中只有 33 (8.3%) 个违反任一和谐律, 其中大部分都是违反涉及名词语序的和谐律, 如表 3 所示。從三语序和谐置前、和谐置后与违反任一和谐律的语言的地理分布 (参见 Tang and 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交界处, 唐威洋与何万顺 (Tang and Her 2017) 因而推测这些和谐律的反例是由语言接触造成的。

表 3: 400 语言的位数词语序、量词语序、名词语序分布情形

		Base-final	Base-initial	Base-absent
CLF-final	N-final	90 (22.50%)	0 (0.00%)	0 (0.00%)
	N-initial	16 (4.00%)	0 (0.00%)	0 (0.00%)
CLF-initial	N-final	1 (0.25%)	2 (0.50%)	0 (0.00%)
	N-initial	2 (0.50%)	21 (5.25%)	0 (0.00%)
CLF-absent	N-final	110 (27.50%)	5 (1.25%)	44 (11.00%)
	N-initial	7 (1.75%)	67 (16.75%)	35 (8.75%)

在多年的研究中，我们累积了数量可观的量词语言相关数据，终得在 2022 年向学术界推出目前世界上规模最大的量词语言数据库：World Atlas of Classifier Languages (WACL) (Her, Hammarström and Allasonnière-Tang 2022)。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 (22%) 个语言是量词语言 (其地理分布参见 Her, Hammarström and Allasonnière-Tang 2022:158, Figure 1)。Her, Hammarström and Allasonnière-Tang (2022) 透过 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 Gil (2013) 的统计结果大体相合，但其世界语言样本中有 35% 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 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 Nichols (1992:200) 的统计结果，他的描述是在亚洲之外，少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。Her, Hammarström and Allasonnière-Tang (2022) 的研究添加了一些细节：欧洲和非洲还是有少量值语言；在环太平洋区域，量词语言主要集中于新几内亚岛 (Papunesia)，在澳洲则非常罕见；各大洲量词语言的数量排序为亚洲>美洲>大洋洲>非洲>欧洲，量词语言的比例排序则是亚洲>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，Her, Hammarström and 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 个有量词语言，其中只有 22 个语系有达到一半的语言是量词语言；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的数据，还包括在我们的理论中与量词相关的其他类型学特征的数据，如同为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。后者是研究团队目前关注的重点：最新的研究 (Her et al. 2024) 整理了 4099 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)。可以看见世界语言目前主要采取位数词置后语序，但位数词置前语言也不少。引人注目的是，这两种语言的地理分布非常不均衡，几乎所有位数词置前语言都集中在非洲，使其成为世界上唯一以位数词置前语言为主的大陆，其余地区几乎都被位数词置后语言占据。何万顺与唐威洋 (Her and 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。相信在深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人类历史上的发展历程，其在语言中的反映，以及以上两点是如何受到各种自然与社会因素影响有更深刻完整的认识。

六、结论

本文综整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，在语意上，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，计量量词可以表 1 以外所有可能的量值。这一差异导致其他的语意差异，即个体量词透过彰显名词提供的认知框架，突显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，所以个体量词的语意是名词语意的一部份，且只能跟可数名词共现。反之，计量量词没有这些限制，所以它有自己独特的语意，且不限共现名词的可数性。

第二，在语法上，个体量词具有数学和语意上的冗余性，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，本文列举现代汉语的语料和汉语历史及跨语言的类型学数据以提供支持。

第三，我们为验证相关理论收集大量量词语言相关数据，得以建立目前世界上最大的量词语言数据库 World Atlas of Classifier Languages (WACL)。在这过程中我们证实了以下的类型学事实：个体量词与构词句法性复标成互补分布，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三者在一个语言中倾向相同，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 WACL 中所有只有量词语言的语系都在这两个地区；论量词语言数量大洋洲排名第三，但论量词语言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丰硕的成果，相信往后学界对名词量化结构的理解必会更加深刻，让我们对量化这一认知能力与语言和其他因素的关联有更完整的认识。

参考文献

- 何万顺、林昆翰. 2015. 〈分类词与量词的区分——以台湾地区华语为例〉，《汉语学报》2015.4:56-68。
- 张亚初. 2001. 《殷周金文集成引得》。北京：中华书局。
- 张景霓. 2005. 〈毛南语的量词短语〉，李锦方、胡素华主编《汉藏语系量词研究》，296-308。北京：中央民族大学。
- 蒋仁萍. 2007. 《基数词和序数词的类型学研究》，南昌大学中国语言文学系硕士论文。
- 蒋颖. 2006. 《汉藏语系名量词研究》，中央民族大学中国少数民族语言文学学院博士论文。

- Aikhenvald, Alexandra Y. 2000. *Classifiers: a typology of noun categorization devices* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press
- Allasonnière-Tang, Marc, Olof Lundgren, Maja Robbers, Sandra Cronhamn, Filip Larsson, One-Soon Her, Harald Hammarström, and Gerd Carling. 2021. Expansion by migration and diffusion by contact is a source to the global diversity of linguistic nominal categorization systems. *Humanities and Social Sciences Communications* 8(1):1-6.
- Au Yeung, W. H. B. 2005. *An interface program for parameterization of classifiers in Chinese*. Hong Kong: Hong Kong University of Science and Technology dissertation.
- Au Yeung, W. H. B. 2007. Multiplication basis of emergence of classifier. *Language and Linguistics* 8(4):835-861.
- Borer, H. 2005. *Structuring Sense, vol. 1: In Name Only*. Oxford: Oxford University Press.
- Cao, Zi-Yun. 2023. *On the Putative [Classifier Noun Numeral] Construction in Tai-Kadai Languages*. Taipei: National Chengchi University MA thesis.
- Chan, Eugene S. L. 2015. *Numeral system of the world's languages*. Available online at <https://lingweb.eva.mpg.de/channumerals/>
- Cheng, Lisa L.-S., and Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4):509-542.
- Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of “semantic parameter”. *Events and Grammar*, ed. Susan Rothstein, pp.53-103. Dordrecht and Boston: Kluwer.
- Comrie, B. 2008. Numeral bases. In *The World Atlas of Language Structures Online*, eds. M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie, chapter 131. Munich: Max Planck Digital Library. Available online at <http://wals.info/feature/131>. Accessed on 2011/4/23.
- Comrie, Bernard. 2013. Numeral Bases. In *The World Atlas of Language Structures Online*, eds. Matthew S. Dryer, and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/131>. Accessed on 2021/08/06.
- Fillmore, Charles J. 1982. Frame semantics. In *Linguistics in the Morning Calm*, ed. Linguistic Society of Korea, pp.111-137. Seoul: Hanshin.
- Gil, D. 2011. Numeral classifiers. In *The World Atlas of Language Structures Online*, eds. M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie, chapter 55. Munich: Max Planck Digital Library. Available online at <http://wals.info/feature/55>. Accessed on 2011/12/13.
- Gil, D. 2013. Numeral classifiers. In *The World Atlas of Language Structures Online*, eds. Matthew S. Dryer, and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <https://wals.info/chapter/55>.
- Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. J. Greenberg, pp.73-113. Cambridge: MIT Press.

- Greenberg, Joseph. 1975. Dynamic aspects of word order in numeral classifier. In *Word order and word order change*, ed. Charles Li, pp.27-46. Austin: University of Texas Press.
- Greenberg, J. 1990[1972]. Numerical classifiers and substantival number: problems in the genesis of a linguistic type. In *On Language: Selected Writings of Joseph H. Greenberg*, eds. K. Denning, and S. Kemmer, pp. 166-193. Stanford: Stanford University Press. [First published 1972 in *Working Papers on Language Universals* 9, 1-39. Department of Linguistics, Stanford University, Stanford.]
- Greenberg, Joseph H. 1990[1975]. Dynamic aspects of word order in numeral classifier. In *On Language: Selected Writings of Joseph H. Greenberg*, eds. K. Denning, and S. Kemmer, pp. 227-240. Stanford: Stanford University Press. Stanford: Stanford University Press. [First published 1974 in *Word order and word order change*, ed. C. Li, pp.27-46. Austin: University of Texas Press.]
- Haspelmath, M. 2011. Occurrence of nominal plurality. In *The World Atlas of Language Structures Online*, eds. M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie, chapter 34. Munich: Max Planck Digital Library. Available online at <http://wals.info/feature/34>. Accessed on 2011/4/23.
- Her, One-Soon. 2012a. Distinguishing classifiers and measure words: A mathematical perspective and implications. *Lingua* 122(14):1668-1691. <https://doi.org/10.1016/j.lingua.2012.08.012>
- Her, One-Soon. 2012b. Structure of Classifiers and Measure Words: A Lexical Functional Account. *Language and Linguistics* 13(6):1211-1251.
- Her, One-Soon. 2017a. Deriving classifier word order typology, or Greenberg's Universal 20A and Universal 20. *Linguistics* 55(2):265-303.
- Her, One-Soon. 2017b. Structure of numerals and classifiers in Chinese: Historical and typological perspectives and cross-linguistic implications. *Language and Linguistics* 18(1):26-71.
- Her, O. S., and Allasonnière-Tang, M. (2022). *Is It 'three hundred' or 'hundred three'?* *An Exploration of the Worldwide Distribution of Numeral Base Orders in Human Languages and Its Implications*. Research project proposal, National Science and Technology Council, Taiwan, Project No. 111-2410-H-029-009-MY3.
- Her, O., and C. Hsieh. 2010. On the Semantic Distinction between Classifiers and Measure Words in Chinese. *Language and Linguistics* 11(3):527-551.
- Her, One-Soon, Harald Hammarström, and Marc Allasonnière-Tang. 2022. Defining numeral classifiers and identifying classifier languages of the world. *Linguistics Vanguard* 8(1):151-164.
- Her, One-Soon, Hui-Chin Tsai, Kun-Han Lin, Marc Tang, and Meng-Chang Lee. 2015. Unification of numeral bases and numeral classifiers: Evidence from SMATTI. Paper presented at the Workshop on the syntax and semantics of numerals, the 48th Annual Meeting of the Societas Linguistica Europaea, September 2-5, 2015, Leiden University Centre for Linguistics (LUCL), Leiden, The Netherlands.
- Her, One-Soon, and Hung-Hsin Hsu. 2022. Numeral classifiers in Niger-Congo languages revisited. Presented online at the 53rd Annual Conference of African Linguistics (ACAL 53), April 7-9, 2022, University of California, San Diego.

- Her, One-Soon, Marc Tang, and Bing-Tsiong Li. 2019. Word Order of Numeral Classifiers and Numeral Bases: Harmonization by Multiplication. *STUF - Language Typology and Universals* 72(3):421-452.
- Her, One-Soon, and Wan-Jun Lai. 2012. Classifiers: The Many Ways to Profile 'one' — A Case Study of Taiwan Mandarin. *International Journal of Computer Processing of Languages* 24(1):79-94.
- Her, One-Soon, Ying-Chun Chen, and Nai-Shing Yen. 2017. Mathematical values in the processing of Chinese numeral classifiers and measure words. *PLoS One* 12(9): e0185047.
- Her, One-Soon, Yung-Ping Liang, Eugene Chan, Hung-Hsin Hsu, Anthony Chi-Pin Hsu, and Marc Allasonnière-Tang. 2024. Early humans out of Africa had only base-initial numerals. *Humanities and Social Sciences Communications* 11.254:1-7. <https://doi.org/10.1057/s41599-023-02506-z>
- Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge: MIT dissertation.
- Huang, C.-T. James, Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Jenks, Peter. 2010. Evidence for the syntactic diversity of numeral classifiers. Paper presented at the 84th Annual Meeting of the Linguistics Society of America, Jan 7-10, 2010, Baltimore.
- Kießling, R. 2018. Niger-Congo numeral classifiers in a diachronic perspective. In *The Diachrony of Classification Systems*, eds. William B. McGregor, and Søren Wichmann, pp.33-76. Amsterdam: John Benjamin.
- Kihm, Alain. 2005. Noun class, gender, and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages. In *Handbook of comparative syntax*, eds. Guglielmo Cinque, and Richard Kayne, pp.459-512. Oxford: Oxford University Press.
- Landman, F. 2004. *Indefinites and the Type of Sets*. Malden: Blackwell.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Li, Xu-Ping. 2011. *On the semantics of classifiers in Chinese*. Ramat Gan: Bar-Ilan University dissertation.
- Li, Yen-Hui Audrey. 2014. Structure of noun phrases: Left or right? *Taiwan Journal of Linguistics* 12(2):1-32.
- Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Little, Carol Rose, Mary Moroney, and Justin Royer. 2022. Classifiers can be for numerals or nouns: Two strategies for numeral modification. *Glossa: a journal of general linguistics* 7(1):1-35. DOI: <https://doi.org/10.16995/glossa.8437>
- Allasonnière-Tang, Marc, and One-Soon Her. 2017. Numeral base, numeral classifier, and noun: Word order harmonization. *Language and Linguistics* 21(4):511-556.
- Nichols, J. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Sanches, Mary, and Linda Slobin. 1973. Numeral classifiers and plural marking: An implicational universal. *Working Papers in Language Universals* 11:1-22.

- Scontras, Gregory. 2022. On the semantics of number morphology. *Linguistics and Philosophy* 45:1165-1196.
- Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In *The Oxford Handbook of Comparative Syntax*, eds. Guglielmo Cinque, and Richard Kayne, pp.806-838. Oxford: Oxford University Press.
- Sinnemäki, K. 2019. On the distribution and complexity of gender and numeral classifiers. In *Grammatical gender and linguistic complexity*, eds. F. Di Garbo, B. Olsson, and B. Walchli, pp.133-200. Berlin: Language Science Press.
- Tai, James H-Y. 2003. Cognitive relativism: resultative construction in Chinese. *Language and Linguistics* 4.2:301-316.
- Tai, James H-Y., and Lianqing Wang. 1990. A semantic study of the classifier tiao (条). *Journal of the Chinese Language Teachers Association* 25.1:35-56.
- Tang, Chih-Chen Jane. 2005. Nouns or classifiers: a non-movement analysis of classifiers in Chinese. *Language and Linguistics* 6.3:431-472.
- Tang, Marc, and One-Soon Her. 2019. Insights on the Greenberg-Sanches-Slobin Generalization: Quantitative typological data on classifiers and plural markers. *Folia Linguistica* 53(2):297-331.
- Watters, John R. 1981. *A Phonology and Morphology of Ejagham, with Notes on Dialect Variation*. Los Angeles: University of California dissertation.
- Zhang, Niina Ning. 2011. The constituency of classifier constructions in Mandarin Chinese. *Taiwan Journal of Linguistics* 9.1:1-50.

Some New Perspectives in the Typological Study of Numeral Classifier Languages

One-Soon Her^{1,2} Yung-Ping Liang²

¹*Department of Foreign Languages and Literature, Tungshih University, Taiwan*

²*Graduate Institute of Linguistics, Chengchi University, Taiwan*

Abstract

In numeral classifier languages, numerical quantification of a noun requires a numeral classifier as an intermediary. Numeral classifiers constitute an independent syntactic category, which can be divided into two subcategories semantically: sortal classifiers and mensural classifiers. The former quantifies discrete entities and highlights certain inherent features of the noun, while the latter specifies a particular magnitude as the unit of measurement for the noun. In recent years, the authors' research team has made some important discoveries in the study of numeral classifiers, expanding the scope of typological research on classifiers. Cognitively, the classifier and the numeral form a multiplicative relationship, with the classifier as the multiplicand. Based on their mathematical values as multiplicands, classifiers are divided into sortal and mensural classifiers. The former highlights a certain essential feature of the noun referring to an individual and necessarily denotes the numerical value 1, while the latter may express values other than 1 or even non-numerical values and can thus be further subdivided into four types. The multiplicative relationship between numerals and classifiers is also reflected in syntax, as the two must form a constituent before merging with a noun. The universality of this left-branching structure, i.e., [[Numeral Classifier] Noun], is supported by many syntactic and typological facts, e.g., there are only four attested word orders of the [Numeral Classifier Noun] construction, where N never comes between

Numeral and Classifier. To conduct typological research on the world's classifier languages, we have collected a large amount of relevant information and established the World Atlas of Classifier Languages (WACL), the largest classifier language database in the world. This paper provides an overview of some typological studies based on this database, focusing on the tendency for the numeral, classifier, and noun to harmonize in word order. Finally, we present the preliminary findings from our extended research from numeral classifiers to numeral bases in the world's languages.