

· 语言学研究 ·

基于统计的概数助词“来”的动态倾向性考察

张言军

(信阳师范学院 文学院,河南 信阳 464000)

摘要:从句法组配上看,“来”字概数结构有多种的表现形式,但在实际运用中,只有三种格式具有较高的使用频率。在语义表达上,助词“来”的表义极其丰富,可以呈现“略多”“略少”以及“左右”三种状态,但在现实话语理解中,又体现出“指多”的理解倾向性。

关键词:概数助词;来;多样性;动态倾向性

中图分类号:H042 **文献标志码:**A **文章编号:**1003-0964(2016)06-0100-06

一、引言

现代汉语中,概数的表示方法可以分为两种:一种是通过相邻基数词的连用来表示概数,另一种则是通过后加概数助词来表示概数^{[1]248}。前者如:两三个、五六个、七八个;后者如:上下、左右、来、把、多。

对于概数助词“来”,学界已有多篇论文对其做了专题研究,这些研究既有基于共时层面的分布考察,也有基于历时层面的演化路径分析。此外,还有学者从对比的角度将其和“把”“多”等概数助词做了对比分析。中肯地说,前贤时彦对很多问题都已做了较为全面、深入的分析。但由于受研究者研究理念以及研究目的的限制,仍有一些问题,值得我们继续关注。

储泽祥指出语言除了具有符号性、工具性之外,还具有概率性的属性。基于这一语言观,储泽祥提出在汉语语法研究中应加强基于统计基础上的动态倾向性考察,这种研究思路的核心是汉语句法语义结构的优先序列研究,他认为倾向性考察不仅是研究视点的转向,也可能造成一个新的巨大的研究领域^[2]。我们认为储泽祥先生的这一研究理念对于汉语的语法研究是非常有针对性的,因为汉语很多的语法规则并不完全是强制性的,而多表现为一种倾向性。基于上述研究理念,本文将在已有相关研究的基础上,对概数助词“来”在句法、语义等方面所表现出的一些倾向性做详细考察。

二、组合形式的多样性与倾向性

作为一个典型助词,“来”在分布上具有黏着性和定位性^[3]。换言之,作为概数助词,“来”是不能独立使用的,而是必须跟其他词语组合成概数结构才能使用。所谓“来”字概数结构,指“十来(个)、三十来(碗)”之类表示概数的结构^[4]。关于“来”字概数结构的组合形式,学界已有多位学者对此问题做了深入分析。他们的研究,有的侧重于共时的层面,有的侧重于历时的层面。下面我们将立足于现代汉语共时层面,主要讨论以下三个问题。

(一)组合形式上的多样性

现代汉语中,概数助词“来”可以同数词、量词、名词以及形容词组合构成“来”字概数结构,在语表形式上,它们的组配方式也是较为多样化的。杨德峰将“来”字概数结构的组合形式概括为如下3种类型^[5]:

- A. 数 + 来 + 量 + (名)/形
- B. 数 + 量 + 来 + (名)/形
- C. 数 + 来 + 位数(万、亿) + 量 + (名)

杜晓艺将“来”字概数结构的组配方式概括为如下4种类型^[3]:

- A. 数词 + 来 + 量词
- B. 数词₁ + 来 + 数词₂ + 量词
- C. 数词 + 量词 + 来 + 形容词
- D. 数词 + 来 + 二 + 数词 + 量词

我们以国家语委现代汉语平衡语料库(平衡语

收稿日期:2016-10-05

基金项目:2014年度国家社科基金项目(14CYY035);信阳师范学院2014年度青年骨干教师资助计划(2014GGJS-18);信阳师范学院博士科研启动项目资助

作者简介:张言军(1981—),男,河南新乡人,语言学博士,副教授,主要从事汉语语法研究。

料库是指兼顾口语、小说、新闻、法律、说明等语体差别的语料库)做了检索统计,在2 000万字的现代汉语语料中,共检索到“来”字概数结构190条。经过认真地分析,我们认为“来”字概数结构可以概括为以下7种组配方式:

A. 数+来+位数(万、亿)+(量)+(名/形)

(1)二月廿一,正是废历新年“年初三”,百来万香港人还沉醉在歌舞升平的废历新年的美梦里。

这一格式中的“量词”以及“名/形”都可以在表意清楚的情况下隐含不出现。如:

(2)目前从自然界发现的和人工合成的有机物已达数百万种,而无机物却只有十来万种。(这一例是名词的隐含)

(3)有个二十来万。(这一例是量词和名词同时隐含)

虽然在国家语委平衡语料库中并未发现形容词进入该格式的实例,但在网络、报刊这些动态语料库中,我们确实发现了这样的用法。如:

(4)这么一个小小的玉环,却管着十来万斤重的石门,定是一件宝物。(网络小说《血色蜀山》第287章)

B. 数+来+量+名/形

(5)没几分钟,我们便看见十来个孩子进了影院。

(6)这一堆货物都是电缆捆,每一件有七百米斤重。

其他再如:“十来个人、十来棵树、十来丈高、十来斤重”等。

C. 数+来+量

有的学者把这一格式归入B格式中,确切地说,这一格式确实可以看作格式B的变体,即隐含了“名/形”而来的,但考虑到这一类型数量较多,我们认为可以把它看作是独立的一类。如:

(7)新中国邮票问世最长不过40来年。

(8)他戴着一副镜片很大的眼镜,长脸、大门牙,大约四十来岁。

D. 数+来+名

这一格式,理论上也可以看作是格式B的省略,即量词不出现或隐含了,但需要注意的是只有名词前的量词可以省略不出现,形容词前的量词(通常是度量词)是不能省略不出现的,所以我们认为也应独立成一类。如:

(9)原来认识二百五十来字的吕国丁同志,提高到认识一千三百来字了。

(10)一份两百来字的说明书,他整整写了一个礼拜。

E. 数+量+来+名/形

(11)这位天津音乐学院附中毕业的姑娘,刚刚新婚宴尔,家住北仓,来回坐汽车就需要三个来小时。

(12)它离周总理逝世只不过九个来月。

(13)你随便拉一根麦穗量一量,都有五寸来长,真是块块都比往年强。

F. 数₁+数₂+来+(量)+(名/形)

(14)在丰收大队粮仓的屋顶、墙角等处,盘踞着百十来只褐鼠。

(15)他说百姓们已经统统跑散啦,他说日本兵只有百十来个人啊。

当格式是名词时,用于计量的量词跟名词之间的组合已经规约化,也可隐含不出现,格式变化为:数₁+数₂+来+名。如:

(16)当夜,由于百十来人一下子涌到青云店,住处一时难以解决……

当格式是名词时,由于“名词”在语境中是已知信息,所以也可以在表达中隐含,格式形式变化为:数₁+数₂+来+量。如:

(17)我临回来,李书记让我估计一下这回扩社的数字,我说了个百十来户,比眼下多上一倍,你们看行不?

但不管是量词的隐含,还是名词的隐含,从语义理解上来看,它们都是存在的。虽然语料库中未发现形容词进入的例子,但从语感预测这样的用法应该是存在的,我们也在网络这一动态语料库中找到了这样的用例。如:

(18)唐寅正色道:“将军,小人有的是力气,百十来斤重的东西,单手就能提起来。”(网络小说《唐寅在异界》第451章)

G. 特殊形式

这一格式的特殊性表现在有两个概数助词出现,也可以看作是一种强化格式,具体有两种表现形式:

量+把+来+名

(19)个把来月的病魔,其间那老太太更饱受她女儿的虐待。

数+多+来+量+名

(20)孩子们一听,几乎同时伸出了二十多条紫堂堂的手臂。

但上述特殊格式每种类型只发现1例,故合并在一起讨论。

此外,对于“碗来粗”这样的用法,我们认为应该看作是比况助词,不应再看作是概数助词,所以本文不涉及对这一类型的统计与分析。

通过对大规模语料库的调查统计来看,共时层

面中,概数助词“来”的组合形式是多种多样的,从语用的角度看,这也满足了人们多方面的表达需求。同时,这种多样化的分布也说明了助词“来”在使用中的活跃程度。

(二) 动态使用中的倾向性

但是如果对概数助词“来”的认识止步于此的话,那将会忽略一些更为重要的问题。那就是在多

样化的组配方式中,哪一种格式是最为常用的,多种格式从高到低的优先序列又是怎样的?对这一问题的回答,不仅直接有助于我们的汉语教学(特别是对外汉语教学),而且对于汉语应用研究(如中文信息处理)也将是有帮助的。基于“语法优选论”的研究理念^[2],我们对190条语料做了逐一分析,并统计得出每类格式的使用频率,具体数据见表1。

表1 “来”字概数结构不同格式的使用频率

格式	格式 A	格式 B	格式 C	格式 D	格式 E	格式 F	格式 G
数量	4	57	77	7	40	3	2
百分比	2.1%	30%	40.5%	3.7%	21%	1.6%	1.1%

表1的数据表明,在7种组配方式中,排在前三位的分别是格式C、格式B、格式E,它们占到了总数的90%以上,其他4种类型的总和还不到10%。由上述统计数据我们可以得出“来”字概数结构在动态使用中的优先序列为:

格式C > 格式B > 格式E > 格式D > 格式A > 格式F > 格式G

观察上述序列,可以得出以下几点认识。

其一,助词“来”最常见的组配形式是“数+来+量”,名词和形容词并不是构成“来”字概数结构的必有成分。换言之,在言语交际中,当背景信息能够较为清晰地指出计数的对象时,“数+来+量”是最为简练的表达方式,这也符合语言表达中的经济性原则。同时对这一类型的语料做进一步的观察,还可以发现,这一格式最常见的运用背景是用以表达人物的年龄。之所以年龄表达最青睐这一格式,是因为年龄的计数对象“年纪”在表达中是不言而喻的,“数+岁”已经是一种被社会高度规约化的表示方式。

其二,格式B中,“数+来+量+名”组合的使用频率较高,而“数+来+量+形”组合的频率则较低。在57条例句中,54例是“数+来+量+名”组合,仅有3例是“数+来+量+形”组合,两者的比例为18:1,由此可见,格式B是极为偏向名词而排斥形容词的。

其三,格式E中,“数+量+来+形”共出现23例,而“数+量+来+名”则共出现17例,两者的比例为1.35:1,从使用频率上看,“数+量+来+形”组合略占优势。同时,分析语料还可以发现,进入该格式的形容词都只能是正向的,而不能是负向的^[6]。在语料中出现的形容词有“长、深、宽、粗、高、重”,而不见相反义的“短、浅、窄、细、低、轻”的

使用。

其四,格式A、D、F、G合在一起的使用频率还不足10%,由此可以看出,这4种格式的使用是极其受限的。而受限的原因大概有如下几种情况:一是格式A之所以受限,是因为只有数字达到十“万/亿”以上,且前面的数字个位必须是零时才能使用格式A,这也就制约了它的使用频率;二是格式D是“数+来+量+名”省略量词而来的,但在现代汉语中数词和名词之间要使用量词几乎是一条强制性的规则,只在个别带有熟语性或被社会高度规约化的表述中才可以省略量词,这就制约了格式D的使用;三是格式F的受限则更为明显,因为只有“百”“十”两个数词才能进入该格式,这也就制约了它的使用频率;四是格式G,我们认为应该将其看作是一种强化格式,这种格式本身跟格式B是一致的,只是在表达中说话人想从不同角度对事物数量进行估计,结果就造成了概数结构的杂糅,所以这一格式本身就是一种非常规的现象,自然使用频率也就不高。

(三) 构成上的限制

表面上看,“来”字概数结构的构成极其简单,只要在数词的后面添加上一个助词“来”就实现了。实则不然,如“十来个人”可以说,但“三来个人”就不能说。所以,我们认为,搞清楚“来”字概数结构的准入条件,不仅有助于深化对汉语相关语法现象的认识,而且也有利于汉语教学,特别是对外汉语教学。那么“来”字概数结构在构成上究竟有哪些限制呢?杨德峰、李宇明、张谊生、杜晓艺、邢福义等学者在研究中都曾对这一问题做过探讨分析,特别是邢福义先生在已有研究的基础上对“来”字概数结构组合形式上的限制做了更为深入细致的刻画^[4]。结合学者们的已有研究,对不同组配格式的准入条件做如下概括。

第一,能进入格式 A 中数词必须是“ $10 \times X$ (X 为整数)”,且个位不能是零的数词。而进入该格式的量词,张谊生概括为“量词一般只限于‘元’和‘人’”^[7]。而根据实际语料,我们认为这一结论稍有不妥,进入该格式的量词既有针对钱的“元/块”,也有针对人或物的“个”“种”,而且量词、名词有时也可以隐含在语义中不出现。

第二,格式 B 中低于 10 的数词都不能进入,且数词也必须个位为零。进入格式中的量词则根据表义的需要,名量词和度量衡量词都可以进入,但动量词一般不能进入该格式,在语料中也没有发现实际用例。

第三,格式 C 对数词的限制跟格式 B 是一致的,而对于量词,该格式则几乎没有限制,名量词、动量词以及度量衡量词都可以进入,是所有格式中对量词最为宽容的一种。

第四,格式 D 实际是格式 B 中“数+来+量+名”省略量词而来的,所以实际限制跟格式 B 是一样的,不再赘述。

第五,进入格式 E 的数词则跟格式 B 正好相反,不能是大于 10 的整数,但特殊数词“半”也可以出现在这一格式中。而对量词的限制,则只有名量词、度量衡量词可以进入,动量词不能进入该格式。

第六,能进入格式 F 的数词只限于“百十”连用,其他数词皆不能进入,而对于量词则没有明显限制,名量词、度量衡量词以及动量词皆可进入,有时量词、名词也可以隐含在语义中不出现。

第七,格式 G 因为是一种特殊格式,实际可以看作是格式 B、E 的一种变体,所以从理论上看在构成限制上应该跟格式 B、E 是一致的。但因实际语料中只有 2 例,“量+把+来+名”中数词原则上只能是“1”(用于“个把”的前面,同时在实际表达中多省略不出现)。而“数+多+来+量+名”中数词应大于“10”(用在“多来”的前面),且个位必须是零的数词才能进入。同时根据 2 例的情况以及我们的语感推断,名量词、动量词以及度量衡量词应该都可以进入。

三、表义的多样性与理解上的倾向性

研究认为,概数助词“来”是由“以来”缩略而来的,在近代汉语里,助词“来”最初只表示不超过某数,到后来才可兼表比某数略多或略少^[8]。而我们都知,词语的使用并不总是一成不变的,随着社会的发展,词语的意义也多会发生某种变化。那么,现代汉语共时层面上,助词“来”的表义问题又呈现出

怎样的状态?从话语理解的角度来看,受话人听到或看到“来”字概数结构又会做出怎样的理解?对于第一个问题,学界讨论的较多,而第二个问题则还缺乏深入的分析,下面我们分别叙述。

(一)助词“来”表义的多样性

观察已有的研究成果可以发现,学界对助词“来”的表义状态共有三种意见,即大于前面的数值、小于前面的数值以及在数值上下浮动。

大于数词所表示的数值,持这一观点的有姚怀德、靳丹丹等;小于数词所表示的数值,持这一观点的有李行健等;在数词所表示数值的上下浮动,持这一观点的有吕叔湘、张必东、杨德峰、张谊生、王改改、邢福义等。

从我们看到的文献来看,比较多的学者是支持第三种观点的,即“上下浮动”或“左右浮动”。但不可否认的是,仍有一些学者是支持第一、第二种观点的。从学者们的考察分析中我们可以看出,概数助词“来”的表义是较为丰富的,特别是邢福义先生的研究,他在考察了近 16 年《人民日报》刊文情况后指出“大多数用例会‘见仁见智’,不同的人可能有不同的理解”,“有的用例,由于能够找到判定的依据,可以确认为‘略少’”,“有的用例,由于能够找到判定的依据,可以确认为‘略多’”^[4]。

“略多”“略少”以及“左右”的同时存在表明了助词“来”表义的丰富性与多样性,而或许正是这种内涵丰富的特点造成了它在现代汉语中的广泛分布与应用。

(二)理解上的倾向性

助词“来”在表义上有“略多”“略少”以及“左右”三种可能性,那么在话语理解的过程中,人们对助词“来”字概数结构有没有一定的倾向性呢?换言之,在多种可能性中,哪一种是在人们在中性语境中最有可能做出选择的理解呢?关于这一点,以往的研究还较少关注,或虽有论及,但也往往较为简略,缺少详细的论证过程。根据学者们的讨论,他们对助词“来”在理解中的倾向性也存在两种不同的,甚至是相对立的观点。

一类研究认为助词“来”在理解上倾向于“略少”。在表义多样性中,学者们明确支持“略多”“略少”的,实际上已经不存在倾向性的问题,故此他们的观点也不再重新提出讨论,重点是那些认为表义是“左右”的学者的观点。如张谊生指出概数“来”的义域虽可以两头延伸,但向上延伸是受限制,且出现频率也较低的^[7]。由这一论述,我们可以推导出张谊生是倾向于“略少”的理解倾向的。

另一类研究认为助词“来”在理解上倾向于“略多”。如邢福义先生指出现代汉语中“来”字概数结构的含义存在两种情况:第一,无法确认,可能见仁见智;第二,可以认定,但不会一边倒。如果用“>”号表示多于,那么,可以列出如下两个顺序:

顺序 A:总体情况:无法确认的 > 可以确认的。

顺序 B:略多于本数 > 本数左右 > 略小于本数^[4]。

虽然邢先生始终认定“左右”是助词“来”表义的核心,但从邢先生给出的两个排序来看,我们还是可以看出他认为人们的语用心理可能更倾向于“略多”。

而不得不说的是,很多学者观点的得出并没有明确的数据支撑。换句话说,在很大程度上是根据个人语感得出的。只有少数学者做了调查统计,王改改对助词“来”的表义做了调查分析,她根据调查

统计认为:被调查者的方言背景越是远离北京话,就越是倾向于理解为“略多”或“左右”,而来自北京话或官话区的人则多是倾向于理解为“略少”于前面的数值。同时如果从年龄的角度看,则老年人倾向于理解为“少于”,而年轻人则倾向于理解为“多于”或“左右”^[9]。

但对于王改改的调查分析,我们认为还存在两点不足:其一她在调查时只选择了2个句子,例句数量太少,且都是询问年龄的语用环境;其二所选例句属于同一种组配方式,即“数+来+量”。我们认为例句数量过少以及类型上的过于单一都有可能影响调查的结论。所以为了更好地验证概数“来”的理解倾向性,我们又重新做了一次调查。这次调查一是要涵盖尽可能多的组配类型,二是例句数量有所扩大,增加至7个。调查数据如表2。

表2 概数助词“来”理解倾向性调查

例句	例句1	例句2	例句3	例句4	例句5	例句6	例句7	合计
略多	7	8	6	4	6	9	11	51
略少	2	2	1	4	3	2	4	18
左右	9	8	11	10	9	7	3	57

从上面的调查统计来看,基本结论跟邢福义先生的研究还是比较一致的。从单个例句看,“左右”都是优选项,其次是“略多”,而“略少”是最不易被选中的选项。如果从整体上加以量化统计,则选中“左右”有57次,选中“略多”的有51次,选中“略少”的有18次。据此,助词“来”的理解倾向性可以排列为:左右 > 略多 > 略少。

单独比对“略多”“略少”两个选项,从总量上,两者的比例为3:1。换言之,在可以判断多少的表达中,“略多”选项明显占有优势。从我们的调查统计来看,那些认为助词“来”的理解倾向是“略少”的观点是难以成立的。换言之,在中性语境中,“左右”是最优势选项,其次是“略多”,再次才是“略少”。

那么,人们在理解过程中为什么会表现出这一倾向性呢?我们以为这或许跟中国人的传统心理有关。经受了几千年儒家文化思想熏陶的中国人,在言语、行动乃至学术观点等方面都遵循谦虚谨慎的态度。这一文化心理在言语方面的典型表现就是主张实事求是,不夸大其词,具体到数量表达方面而言,常常是就低不就高,亦即把事物的数量往小处说。如果一件东西的价钱不到十万,那么一般会用“九万多”这样的表述方式,而一般不会用“十来万”

这样带有容易使别人产生误解的模糊性词语。所以在这种群体心理的制约下,即使助词“来”最初的表义是不多于,但后来听话人却在语用推理的层面上逐渐向“略多”靠拢,因为既然说话人用了“十来万”,那一定是起码到了“十万”他才会这样说,不然他就是在夸大其词。在这种语用推理的反复使用中,“略多”也就渐渐被突显出来。至于这种变化的进程会在何时结束,我们还难以有一个定论,但我们相信这种演进的趋势是不会改变的。

四、结语

本文基于“语法优选论”的研究理念,对概数助词“来”在句法、语义以及语用等方面的动态倾向性做了全面考察,得出以下几点认识。

句法组合上,概数助词“来”共有7种组配形式,这种组配形式的多样性不仅体现了它较强的活动能力,而且也满足了人们多方面的表达需求。但在具体使用中,又只有3种格式具有较高的使用频率,其他4种格式的使用则较为受限。

从语义表达的角度看,概数助词“来”有“左右”“略多”“略少”3种表义状态,这体现了它表义的多样性与丰富性。而从话语理解的角度看,“来”字概

数结构又呈现出“指多”的理解倾向性。

通过对概数助词“来”句法语义结构优先序列的考察,不仅深化了我们对其动态使用特征的认识,而且通过我们的研究也可以补充前人说法中的一些不足,如有研究者指出当量词为“度(温度)”和“倍”时,“多”适用于“数词+多+类别词/度量单位+名”格式,而“来”不适用^[5]。但考察语言事实,我们认为这一结论有待修正。在人民网中,就可以找到多例“数词+来+度/倍+名”格式的材料。如:

(21) 相差肯定会有,但不可能会一下子相差十来倍。

(22) 从1998年房改政策结束的这十五六年间,石市房改房价升值幅度非常大,总房价一般都上涨了10来倍。

(23) 这些地方风景如画,盛夏时气温也不过二十来度……

(24) 温度会慢慢升起来,估计明天就能达到二十来度了。

参考文献:

- [1] 齐沪扬. 汉语通论[M]. 北京:中央广播电视大学出版社,2004.
- [2] 储泽祥. 在多样性基础上进行倾向性考察的语法研究思路[J]. 华中师范大学学报(人文社会科学版), 2011,(2):90-94.
- [3] 杜晓艺. 概数助词“来”的句法分布[J]. 高等函授学报(哲学社会科学版), 2006,(6):31-32.
- [4] 邢福义. 事实终判:“来”字概数结构形义辩证[J]. 语言研究, 2011,(1):1-10.
- [5] 杨德峰. 表示概数的“多”和“来”的全方位考察[J]. 汉语学习, 1993,(3):10-16.
- [6] 张国宪. 现代汉语形容词的体及形态化历程[J]. 中国语文, 1998,(6):403-413.
- [7] 张谊生. 概数助词“来”和“多”[J]. 徐州师范大学学报(哲学社会科学版), 2001,(3):80-84.
- [8] 江蓝生. 概数“来”的历史考察[J]. 中国语文, 1984,(2):145-153.
- [9] 王改改. 概数词“来”的语义调查和研究[J]. 汉语学习, 2001,(6):46-49.

The Dynamic Tendency of Approximate Number Auxiliary "lai(来)" based on Statistics ZHANG Yanjun

(College of Liberal Arts, Xinyang Normal University, Xinyang 464000, China)

Abstract: From the perspective of the syntactic equipping, there are various forms of combinations about estimates structure "lai(来)". But in practical application, only three kinds of formats have higher use frequency. In terms of semantic expression, its meaning is very rich, showing "a little more", "slightly less", as well as "about". But in the real discourse understanding, they also reflect the tendency of "pointing to a little more".

Key words: approximate number auxiliary; lai(来); diversity; dynamic tendency

(责任编辑:韩大强)