

“联系项平衡假设”的语料库验证

余泽超

(浙江师范大学 外国语学院,浙江 金华 321004)*

摘要:以含有框式关联词的零形下指句为对象,提出汉语零形下指的“联系项平衡假设”:下指句中的小句主题与关联词相互依赖、虚实互补,构成联系项均衡分布。本文以自建语料库为依据,进一步对上述假设进行验证。

关键词:零形下指;框式关联词;“联系项平衡假设”;语料库验证

中图分类号:H109.4 **文献标识码:**A **文章编号:**1001-5035(2018)04-0079-05

一、引言

以 Dik 所提“联系项居中原则”^[1]为依据,国内学界(如刘丹青,储泽祥等,张建)将汉语复句的关联词标记模式(即分布位置)概括为以下优先等级:居中单用 > 成对使用(框式) > 前置单用;^[2-4]典型例子如:“但是/所以/而且₂” > “虽然₁ + 但是₂/因为₁ + 所以₂/不但₁ + 而且₂” > “虽然/因为/不但₁”。^①余泽超以包含框式关联词的汉语零形下指句($\emptyset + P/N$ 式)为依据,提出不同于 Dik 的“联系项平衡假设”,认为下指句内的小句主题兼有连句功能,主题与联词的互动结果为相互依赖、虚实互补,即:主题出现时,同句联词可省去;主题空缺时,同句联词需保留。对照“联系项居中原则”及其对应的优先等级,我们将零形下指句的联系项分布特征,概括为:

(1)“联系项平衡假设”:

鉴于下指句($\emptyset + P/N$)的后句主题(P/N)具有居中联系功能,框式关联词的分布以前置单用或成对使用为主,居中单用次之。

关联词优先等级为:前置单用/成对使用(框式) > 居中单用。

例(2)中三句,在建设本文所用语料库过程中提取,关联词分布特征分别属于前置单用、成对使用(框式)、居中单用;指代关系分别是: $\emptyset + P$ 、 $\emptyset + P + \emptyset$ 、 $\emptyset + \emptyset + P$ 。鉴于除例(3)外所有语例均出自《骆驼祥子》,我们在例句中省去语料出处与编码,同时省略表示指代关系与小句顺序的下标;括号内的关联词由本文添加。

- (2)a. 虽然 \emptyset 不怕打架,(但是)他 \emptyset 不想打架。
b. 虽然 \emptyset 用老婆的钱不大体面,但是他与她的关系既是种有口说不出的关系, \emptyset 也就无可如何了。
c. \emptyset 怎想怎不是味儿,可是 \emptyset 带着那么个大肚子,她 \emptyset 不敢去打架。

本文以自建下指语料库对我们的“联系项平衡假设”进行验证。下文先介绍语料库建构、语料检索与统计过程,然后具体呈现量化结果与实例分析。

* 收稿日期:2018-02-20

作者简介:余泽超(1965—),男,浙江金华人,浙江师范大学外国语学院教授,文学博士。

基金项目:教育部人文社会科学研究项目“第三人称下指英汉对比研究”(12YJA740098)

二、语料库建构及语料统计

本文所用语料库以我们前期建立的下指语料库为基础:我们的原始语料库,用 EXCELL 录制了 1 100 组英汉叙事语篇平行语料,其中汉—英部分的源语语料由《骆驼祥子》和《子夜》构成,分别有 466 和 178 个句对。语料标记与分类的两大项目是:指代关系与关联特征,指代关系以 θ 、P、N 组合表示,关联特征以联系词与下标显示。语料输入与标注,均用手工完成。^[5]

本文拟采用原始语料库的《骆驼祥子》分库而不用《子夜》分库,理由如下:在《子夜》分库中,例(3)类带框式关联词的语料合计只有 8 句,数量不及《骆驼祥子》的十分之一。另外,在(3b)中,虽然“也”“就”两个联词都有与代词互动、构成联系项平衡的作用。但是,两者分列明示主题 N 前后,分别与主题构成下指与回指关系,互动结构式可拆分为“ θ 不 + θ 也不 + $N_{3/4}$ ”“ $N + \theta$ 就_{4/4}}”,^②也就是说,这里的“就_{4/4}}”字不是本文讨论的重点。

(3)a. 一路上 θ 并没说得几句话,他们三位就走到了吴公馆的前面……(《子夜》#zy059)

b. θ 更不等范博文回答, θ 也不招呼他同走,林佩珊旋转身体, θ 很快地就向园子里的大路上跑去。(《子夜》#zy053)

从容量为 466 句组的《骆驼祥子》分库中,我们检索并提取含有框式关联词的语句,合计 113 例,再对框式关联词的分布进行归类、统计,最后做成一个独立的“联系项互动语料库”。

框式关联词的分类问题比较复杂,需要补充说明。我们以邢福义为参照,将下指句按语义关系分为先后、因果、转折三大类:1)先后类包括顺承、递进(表 1 前 6 栏)。“从一/自从一/只要一 + 就”与“(一)直到 + (这)才”,时间本义与强调功能相似,也归入先后类;2)因果类包括因果、假设、条件(7—12 栏);3)转折类只有一个单类(最后一栏)。^[6]时间“先后”与逻辑的“因果”,虽然语义上还有重叠,但是外形不难甄别:前者以“一 + 就/便”为典型,后者以“因为 + 所以”为代表。关

联词的分布特征,按前置单用、居中单用、成对使用/框式分为三类,单用与成对使用不重复计数。部分关联词有复合使用的情况:表 1 第三栏及第一栏中的“刚一 + 就”,都以“一 + 就”为基础,其中“只要一 + 就”看作“一 + 就”与“只要一 + 就”(另有单用/列一例)的复合;倒数第二栏否定关联“既没 + (便)没”是“既 + 便”与“没 + 没”的复合;第七、第九栏的“也就”“也还”也是合用。

因为总量不大,统计并不复杂。

三、结果与讨论

关联词分布的统计结果,见表 1。我们从总体趋势与具体实例两方面来呈现关联词与主题的互动结果。

先看总体趋势。1)从表 1 左列看,共有 12 小类(type)关联词;后三列合计频数(token)为 113。按比例看,前置单用、居中单用、成对使用/框式的频比为 40 : 26 : 47。前后两项合计 87 例,明显大于居中频数(26);2)在成对使用/框式的 47 例中,凡是前置、居中两栏为零值的,表明无法拆分使用,可看作“强制框式”。这样的“强制式”共 15 例,占框式总数近 1/3,以复合关联为主,包括“从一/自从一/只要一 + 就”“越 + (就)越”“既没 + (便)没”等;3)从三大类比例看,以先后与因果为主(各占 66 与 29 例),转折类 18 例。先后类中以“一 + 就/便”居多(表 1 前三栏):合计 44 例,占表 1 总数近四成。转折类中有“轻转折”^③关联词 12 例,因为不带对应前框词看作“强制居中”;4)逆向分布算作居中的有 7 例,占居中分布的一半,因为不便省略也看作“强制居中”;5)如果不看强制式,那么前置单用、成对使用、居中单用的比数为 40 : 32 : 7。换言之,关联词优先等级表现为:前置单用 > 成对使用(框式) > 居中单用,完全符合上文(1)所述“联系项平衡假设”。

综合表 1 联系项特征,可见:成对使用的(47 例)属于显性平衡,前置单用的(40 例)因主题、关联词互动取得平衡,合计频数(87)是居中单用(26 例)的三倍有余。统计结果完全符合上节“平衡假设”。

接下来,我们分类细看一些例句。

先后类。例(4)的前三句,是“一+就/便”类典型例句。(4d)是个复合式:“一+就”“只要+就”复合的结果是省去一个后框词“就₂”。框式关联“一+就/便”主要用于连接短时相继、顺向承

接的真实事件;其象似性理据^[7]既可从四成多的频比看清,又可从后框词(“就₂”)的省略得到验证:(4a)的结构式为“一₁+P(就)₂”,是表1前置单用列和全表频数最高的句式,合计频数27,明显高出对应的框式频数(11)。

表1 下指句关联词分布特征表

关联词	频数		
	前置单用	居中单用	成对使用
一/刚/才/刚一/刚要/没+就/便	27	0	11
一+忽然/猛的/轰的一下	0	0	3
从一/自从一/只要一+就	0	0	3
(一)直到+(这)才	0	2	6
越+(就)越	0	0	5
不但/除了+而且/还;既+又;并非+而是;是+还是	2	2	5
因为/因+所以/也就	5(+3)	1+3	1
如果/要是/假若+就/便	3(+2)	1+2	1
即使/就是+也/还;设若+也还;无论+都	0	1	5
只要+就	0	0	1
既+便	1	0	1
非+不;不+就不;既没+(便)没	0	0	3
虽然+但是/可是/可/可也/反倒	2(+2)	12+2	2
合计	40	26	47

加号后数字(合7例)在“前置单用”一列中放在括号内,表示逻辑上前置的关联词实际在后句出现,即属于逆向分布,算作“居中单用”。

(4)a. 他一停住,他们的汗(就)在背上结成了冰。

b. 他一坐在车上,他的大手便向脚垫下面摸去。

c. 他这么一想,他轰的一下冒了汗!

d. 是的,只要他一索回那笔款来,他就没有不放心的事了!

(5)a. 他既心疼钱,他又恨自己这样的不济,他

居然会被一场雨给激病,他不肯喝那碗苦汁子。

b. 他不但是出了钱,他还亲自去吊祭或庆贺,因为他明白了这些事并非是只为糟蹋钱,而是有些必须尽到的人情。

例(5)两句的结构比较复杂。从主题结构看,(5a)前置三个零形小句($\emptyset\emptyset\emptyset P$),在开头两句的“虚”空位置,分别填补一个单音联词(“既”“又”)与末句主题(P)呼应,显然有利平衡全句。细看语义关系,第二、第三小句之间还有一对更隐蔽的关联“这样的→居然”,衔接功能与英语

“so... that...”相当。从逻辑关系看,前三小句层层递进,与末尾P句构成因果关系。这样的功能分析,显然不能简单依赖句法的“等立/并列”关系。(5b)句的主题结构为 \emptyset PP,内含一个非人称主题(“这些事”),句法、语义关系更为复杂。这是个“释因句”:“因为₂”逆向/居中分布。从认知加工的角度看,逆向关联因违反象似原则更难理解,关联词也就更难省略。综合句中所有联系项——包括从句中“并非+而是”,(5b)同样符合平衡原则。

因果类。成对使用关联的,以“即使+也”“既+便”两类条件句为主,因果句、假设句各占一例。以下是假设关系的一组例句,联系项总体分布比较匀称:

- (6)a. \emptyset 今天要是走不进城去,他想,祥子便算完了;
- b. \emptyset 假若这么活下去,他(便)会有一天成为一个干骨头架子,还是这么大,膛儿里全是空的。
- c. \emptyset 独自抱着壶茶,假若 \emptyset 是赶上在茶馆里,或 \emptyset 独自数着刚挣到的铜子,设若是在车口上,他用尽力量把怒气纳下去。

(6a)是语料库中唯一用了框式关联的假设句,理由似与前后框跨度(有插入语小句相隔)有关;(6b)单用前置“假若₁”,未用后框关联“便₂”,这与周刚及董秀英所说假设句需要合用关联词^[8-9]的结论不太一致;(6c)相对复杂:主题结构为“ $\emptyset\emptyset\emptyset\emptyset$ P”,两个假设关联虽然放在末句“他”前,相对前面的小句却是倒置,所以标作“假若/设若₂”,属“强制居中”。

3)转折类。与前两类不同,转折类关联词以居中单用为主,共14例,占居中总数的大半;又以强制居中的“轻转类”为主(12例)。与框式“重转类”不同,轻转类关联词不带对应的前框词(另加“?”表示)。例(7)前两句与后句差异明显:前框关联的隐、现与句间跨度密切相关,符合语言的象似性原则。单看关联词,“轻转类”的明显特征为居中分布(“可是₂”);结合主题而看,互动的结

果还是平衡分布。

- (7)a. (? 虽然) \emptyset 怎想怎不是味儿,可是 \emptyset 带着那么个大肚子,她不敢去打架。
- b. (? 虽然) \emptyset 跑起来,腿确是不似先前了,可是一股热气支撑着全身,他拚了命!
- c. 虽然 \emptyset 不肯思索, \emptyset 不肯说话, \emptyset 不肯发脾气,但是 \emptyset 心中老堵一块什么,在工作的时候暂时 \emptyset 忘掉,只要 \emptyset 有会儿闲工夫,他就觉出来这块东西——

最后,我们看一下诸如(5b)(6c)(7c)之类含四个以上小句的多重复句。因为语义关系更复杂,多重复句一般需要更多的关联标记(平行结构是个例外),以(8)为例:

- (8)a. \emptyset 摸了摸脸上那块平滑的疤, \emptyset 摸了摸袋中的钱,又 \emptyset 看了一眼角楼上的阳光,他硬把病忘了, \emptyset 把一切都忘了,好似 \emptyset 有点什么心愿,他决定走进城去。
- b. \emptyset 既没顺着山路走,他白天没有可以隐藏起来的机会; \emptyset 要打算白天也照样赶路的话,他必须使人相信他是个“煤黑子”。
- c. 直到 \emptyset (被)逼急了,他的那块疤有点发红了, \emptyset 才说,“发财,妈的我的车哪儿去了?”

(8a)与(7c)相似,由七个小句合成,也包含平行结构。除去平行部分,以上“大长句”的关联词比例明显高出只含两三个小句的“小短句”; (8b)中的关联词还与否定标记构成互动关系:我们把“既没₁+他没有₂”拆分为“既₁+(便₂)”+“没₁+(就)没₂”。换言之,借用居中代词“他₂”,(8b)省去了两个后置联词(便₂、就₂)。事实上,(8b)分号之后也是个下指式,同样省略一个后置联词“就₂”,可写作“要打算一的话₁+(就₂)”。

综上,量化结果与具体例证,都证实了上文(1)所述下指句“联系项平衡假设”。从形一义匹配关系看,平均分布、平衡用力符合语言象似性、经济性原则,是下指句小句主题与关联词互动的理想结果。

四、结 语

本文以自建语料库的统计结果与实例分析,初步证实了我们在前文中所提的汉语零形下指的“联系项平衡假设”。需要补充说明的是,鉴于以往研究主要以包含两个小句的简单下指句($\emptyset + P/N$)为对象,后续研究需要继续考察包含多个零形式以及下指连带回指的复杂下指句(如上文(7c)(8)等),以便进一步验证以下指句为据的“联系项平衡假设”、深化主题—关联词互动研究。此外,后续还需要探讨例(9)类句子为何缺省明示关联、需要比较该类句子与带有关联的下指句有何区别。

(9)a. \emptyset 没了女仆,她得自己去下厨房做饭。

b. \emptyset 看着院内的空棚,被水月灯照得发青,和撤去围裙的桌子,老头子觉得空寂无聊……

总之,作为特殊的复句/合句/“流水句”,下指句研究的深化、细化似乎还有很长的路要走。

注释:

①下标表示关联词/小句相对位置,照学界常规,我们的前期研究

以含两个小句的简单复句为典型语料。关联词不限于连词,下文又以“联词”简称。

②3/4,4/4 分别表示:在合计四个小句中的全句/大句中,分别位列第三、第四句。

③“轻转折”的对立体是“重转折”,“轻”“重”分别指转折语义的“弱”与“强”。

参考文献:

- [1] DIK S C. The Theory of Functional Grammar [M]. Berlin & New York: Mouton de Gruyter, 1997: 398.
- [2] 刘丹青. 语序类型学与介词理论[M]. 北京:商务印书馆, 2003:69.
- [3] 储泽祥,陶伏平. 汉语因果复句的关联标记模式与“联系项居中原则”[J]. 中国语文,2008(5):410-422.
- [4] 张建. 汉语复句关联标记模式的组合经济性[J]. 汉语学报, 2012(4):88-94.
- [5] 余泽超. 英汉语篇下指认知功能研究[M]. 上海:复旦大学出版社,2010:13.
- [6] 邢福义. 汉语复句研究[M]. 北京:商务印书馆,2001.
- [7] GIVON T. English Grammar: A Functional-based Introduction, Vol. 2 [M]. Philadelphia: John Benjamins, 1993: 285.
- [8] 周刚. 连词与相关问题[M]. 合肥:安徽教育出版社,2002:18.
- [9] 董秀英. 汉语后置假设标记与“联系项居中原则”[J]. 华文教学与研究,2012(3):60-65.

Corpus Verification of “the Balance Hypothesis for Relators” for Chinese Cataphora

YU Zechao

(College of Foreign Languages, Zhejiang Normal University, Jinhua 321004, China)

Abstract: Taking the example of paired conjunctions in Chinese cataphora, Yu (2008) proposed a new principle for the interaction between clause topics and conjunctions in Chinese cataphora, namely “the Balance Hypothesis for Relators”. This paper continues to verify the above principle with statistical analysis of a self-built parallel corpus.

Key words: zero cataphora; paired conjunctions; “the Balance Hypothesis for Relators”; corpus verification

(责任编辑 廖向东)