

文章编号: 1003-0077(2007)05-0131-05

否定词跨标点句管辖的判断

张瑞朋^{1,2}, 宋柔¹

(1. 北京语言大学 语言信息处理研究所, 北京 100083; 2. 中山大学 国际交流学院, 广东 广州 510275)

摘要: 现代汉语中基本否定词“不”以及扩充词“从不”、“很不”、“不能”、“不会”等的否定辖域受到学术界重视, 但前人研究一般局限于句内, 且主要局限于基本否定词, 其实否定辖域也涉及多个标点句, 否定词的管辖判断也涉及到扩充的否定词。跨标否定词跨标点句管辖的判断和否定词共享问题是整个跨标点句语法共享问题的一个重要组成部分。本文从形式上找到了一些否定词跨标点句的共享规律, 即着重从形式角度讨论了否定词跨标点句的辖域问题, 对现代汉语长句语法分析有重要作用, 并对汉外机器翻译有实用价值。

关键词: 计算机应用; 中文信息处理; 否定词; 标点句; 管辖; 共享

中图分类号: TP391

文献标识码: A

Negation Scope of Chinese Compound Sentences

ZHANG Rui-peng^{1,2}, SONG Rou¹

(1. Language Information Processing Center, Beijing Language University, Beijing 100083, China; 2. School for Overseas Educational Exchange, Sun Yat-Sen University, Guangzhou, Guangdong 510275, China)

Abstract: The scope of negative words, including the simple negator “不” and the compound negators “从不”, “很不”, “不能”, “不会”, are paid more attention in linguistics. However, many researches are focused on the negation scope in single sentence. In this paper, we find out some formal rules about the negation scope in compound sentences. It will be worthy to automatic syntax analysis of long sentences and machine translation.

Key words: computer application; Chinese information processing; negation scope; negator; compound sentence parsing; machine translation

1 引言

本文所说的标点句,指的是现代汉语书面语篇章中近邻的两个标点之间的词串。这里所说的标点,包括逗号、句号、分号、叹号、问号、冒号。其他标点符号这里暂不涉及。汉语的句法分析研究基本上以单句为对象,但在真实语料中,汉语单句边界的自动确定是很困难的。计算机处理汉语的前提是汉语的形式化,因此标点句自然而然就成了计算机处理汉语句子的基本单位。

标点句的边界是清楚的,但很多标点句的句法成分不完整,需要到上下文语境中寻找。

如果一个标点句 a 中的成分 A 和 C 有某种句

法关系, A 又与另一个标点句 b 中的成分 B 有句法关系,那么 A 称为共享成分, AB 是跨标点句的句法结构,标点句 a 称作原配句,标点句 b 称作续配句^[1,2]。

例: 1) 我租了一个房间, a

打算在香港住几个月。 b

b 句是一个述宾结构,描述的主体仍然是 a 句中的“我”,所以说 b 句的述宾结构共享 a 句的主语“我”,“我”和 b 句的述宾结构是主谓关系,所以 a 和 b 有跨标点句的主谓关系,其中 a 是原配句, b 是续配句,共享成分是“我”,原配成分是“租了一个房间”,续配成分是“打算在香港住几个月”。把 b 句换行排在共享成分“我”的后面,这种表示方法叫换行缩进表示法,它能清晰明了的表示出两个标点句之

收稿日期: 2007-04-30 定稿日期: 2007-06-29

基金项目: 国家自然科学基金资助项目(60572159)

作者简介: 张瑞朋(1979 →),女,博士生,研究方向为自然语言处理;宋柔(1946 →),男,教授,研究方向为自然语言处理。

间的句首成分共享情况^[1,2]。

2 为什么要研究否定词的共享问题

以往的研究多局限于单句内部的否定辖域^[3-7],很少涉及到跨标点句的否定辖域,而且主要针对基本否定词,很少论及基本否定词扩充词的辖域。跨标点句的否定辖域是整个跨标点句句法共享问题的一个组成部分。

跨标点句的主谓关系现象是语言学中关注比较多的,除了主谓关系,还有动宾关系、状中关系、定中关系、介宾、动补、连动等关系。下面例子中采用换行缩进表示法,要强调的共享成分用下划线标出,要强调的续配成分用波浪线标出。

例:动宾关系:

- 2) 信封里面只是一张信笺,
 一个红纸袋。

状中关系:

- 3) 他还是丕志高气满,
 改变本色。

定中关系:

- 4) 这吻的分量很轻,
 范围很小。

介宾关系:

- 5) 中国人民为反对霸权主义和推进世界多极化,
 推动建立公正合理的国际政治经济新秩序,
 促进世界和平与发展的崇高事业
 而不懈努力。

动补关系:

- 6) 现在学文科的人都穷得没有脸见人,
 不敢认同学了。

连动关系:

- 7) 苏小姐便打电话问唐小姐病好了没有,
 请她就来。

其他几种关系的讨论有另文详述,这里只讨论状中关系。两个标点句是状中关系时,原配句中共享成分可能是介词短语、副词、时间词等,这里重点讨论共享成分是否定词的情况。

在现代汉语中,一般的,基本否定词“不”不能单独被后一标点句共享,即:后一标点句若没有否定词,经常不能表达否定意。但否定词被后句共享的现象确实存在,在语料中虽然不多,但它是一种很有意思的语言现象。

跨标点句的句法分析问题尚无系统性方法,使得汉语长句分析和长句生成效果很差,并已经成为汉外机器翻译和汉语理解等深层次汉语处理应用系统的瓶颈。否定词的共享如果没有得到清楚地研究也会影响整体共享研究的效率,并且影响汉外机器翻译的效率。

本文调查语料主要是钱钟书《围城》,并辅助调查北京语言大学语言信息处理研究所《大陆小说文本库》和《名家小说文本库》。本文主要讨论否定词“不”及其他的扩充词的跨标点句共享情况。

3 基本否定词“不”的共享与不被共享

3.1 基本否定词“不”不被共享的条件

《围城》中前句是否定标点句,后句不完整的句子共2058句,剔除掉“不”的固定搭配、词语、习语,如:不好意思、惠而不费、经不起等,“不”在连动结构中,如:不小心打翻杯子,“不”在主语短语位置上,如:不做笔记只表示你们学问好,“不”构成并列结构,如:不是……而是,除此之外,“不”在原配句中作谓语动词的状语,续配句缺失成分的句子共668句,否定词共享情况如表1所示。

从表1可以看出:后一标点句没有否定词表达肯定意思的句子占绝大多数,即否定词管辖范围一般不能跨越标点句。

表1 原配句中的“不”被共享情况

相关总句数	续配句不共享“不”,表达肯定的句数及比例		续配句共享“不”的句数	续配句重复“不”,表达否定的句数
668	655	98.05%	4	9

例:续配句没有否定词,表达肯定意思:

- 8) 辛楣不爱米线,
 要一客三鲜糊涂面。

例:续配句表达否定意思,重复否定词:

- 9) 像我们这种小户人家,
 丕打牌,

不请客。

- 10) 那时候她不请医生打针，
不吃止痛药片。

例：后一标点句共享原配句否定词，表达否定意思：

- 11) 双方还没有彼此看破，
彼此厌恶。
12) 他没多动一条神经，
多用一丝声气。
13) 他还是坐二等车，
丕志高气满，
改变本色。

前句是否定句，后句是肯定句、不共享前句否定词，这可以从形式上找到一些不共享的条件。后句句首若有副词（除了程度副词和范围副词）、介词短语、关联词、名词性短语、能愿动词等，都不能共享前句否定词。在《围城》中原配句谓语前有否定词“不”，续配句没有否定词，表达肯定意思的有 655 句，符合这些条件的有 601 句，占 91.76%。

例：后句首有关联词或关联副词：

- 14) 这吵架没变严重，
因为不能到孙家去吵。

例：后句首有副词：

- 15) 柔嘉也没擦胭脂，
只敷了粉。

例：后句以名词性短语开头：

- 16) 那天刘小姐没吃早饭和午饭，
一个人在屋后的河边走来走去。

若续配句句首没有以上形式条件，还要看否定词后的语言成分和续配句结构是否相同，对仗是否整齐，语义是否相近或相反。若结构不同，语义没有相近或相反关系，也往往不共享否定词。

例：17) 顾尔谦不懂用意，
问：“什么？”

“不”后成分“懂用意”是动宾结构，续配句“问”是单个动词，二者结构不同，所以不共享“不”。

例：18) 辛楣不爱米线，
要一客三鲜糊涂面。

“不”后成分和续配句虽然都是动宾结构，但对仗太不整齐，也不能共享否定词。

3.2 基本否定词“不”被共享的条件

3.2.1 否定词后和续配句都是较长的形容词：四字短语或成语，则续配句可以共享否定词。

四字短语是一个严密的整体，否定词否定整体。如果形容词太短，就不行。

例：19) 她丕洋洋得意，
骄傲自满。

但不能说：

她丕漂亮，
可爱。

只能说：

她丕漂亮，
丕可爱。

或

她丕漂亮、可爱。

3.2.2 否定词后如果是动词 + 名词的动宾短语，则要分情况。

(1) 如果否定词后动词和续配句动宾短语的动词相同，续配句一般不能共享否定词。

例：20) 你们丕不吃肉，
吃长斋，
将来西天住不下。

a 句“不”后的动词和 b 句动词相同，都是“吃”，a、b 两句“不吃……，吃……”构成并列的对比结构，所以 b 句不共享 a 句否定词。

再比如：21) 那个姑娘很注重衣着，
丕愁吃，
愁穿。

同样的 d 句和 e 句中“不愁……，愁……”构成并列结构。所以，如果否定词后的动宾短语的动词和续配句的动词相同，“不 v……，v……”构成并列结构，因此续配句不共享否定词。续配句要表达否定意思，经常要重复否定词或用顿号，比如：不愁吃，不愁穿。或者不愁吃、愁穿。

(2) 如果否定词后动词和续配句动宾短语的动词不同，并且动宾短语字数较多（字数 4），结构相同，对仗整齐，则续配句可以共享否定词。

如上面例 12)、13)。如果字数较少，尽管动词不同，一般也不能共享否定词，很短的词一般不单说。如：上面例 9)，一般不能说：

我们这种人家丕打牌，
请客。

来表达和例 9) 相同的意思，因为“请客”字数太少，不倾向于单用。如果一定要省去否定词，则最好放到一个标点句中：他丕打牌请客。或他丕打牌、请客。

再比如，可以说：

22) 她不恭维领导,
欺压百姓。

续配句可以共享原配句“不”,但不能说:

她不恭维领导,
恭维同事。

上例的意思更倾向于:

23) 她不恭维领导,
恭维同事。

句子意义的理解往往是在前后对比中获得的。在“不 V N1, V N2”中,“不 V”与“V”可以构成对比,于是“不”倾向于不管“V N2”;在“不 V1 N1, V2 N2”中,“不 V1”与“V2”不能构成对比,“V1 N1”与“V2 N2”却能构成对比,所以“不”能管辖“V2 N2”。

4 基本否定词“不”前部扩充

4.1 否定词:并不、从不、很不、大不

从语料调查情况来看,上述扩充词和基本否定词“不”相同:总的倾向于不被后句共享,但在前后句谓语对仗整齐,结构相似时,也可以被共享。和基本否定词“不”不同的是:当前后句动宾短语的动词相同时,并不、从不、很不、大不也稍倾向于被后句共享。

24) 他从来都不浪费自来水,
浪费天然气。

25) 我刚才并没打开空调,
打开窗户。

上述两例如果去掉否定词前的副词“都”、“并”,续配句更倾向于不共享否定词。

4.2 肯定词:莫不、无不

“莫不”和“无不”双重否定表达肯定,相当于“都”,虽然结构形式上和“从不”、“很不”等类似,但不是否定词。《大陆小说文本库》语料中,原配句中“莫不”作谓语动词的状语,续配句缺失成分的句子共 24 句,续配句共享“莫不”的情况如下表:

表 2 原配句中的“莫不”被共享情况

相关总句数	续配句共享“莫不”	续配句不共享“莫不”	共享“莫不”的比例
24	17	7	70.83%

从表 2 可以看出,和前述基本否定词和扩充否

定词的共享情况不同,“莫不”被续配句共享的例子占绝大多数。这就是说,“莫不”、“无不”不但意义和否定词不同,共享形式也不同。由此也可以看出否定词的共享和肯定词不同,有其独特的特点。续配句不共享“莫不”的原因是:它们虽然是肯定词,但也都是副词,所以和“不”、“都”一样,后面不能接关联词、名词、副词(除了程度副词和范围副词),所以后句句首若是关联词、名词、副词,不能共享“莫不”、“无不”,如例 28)。

例:续配句共享“莫不”、“无不”

26) 鸿渐这学期莫不成绩优良,
名列前茅。

27) 通州城里的居民和四乡的百姓无不欢欣鼓舞,
拍手称快。

例:续配句不共享“莫不”

28) 而古今中外所谓好文章也莫不充满真挚动人其热如火的情感,

故能深深叩动读者
的心弦而起共鸣。

5 基本否定词“不”后部的扩充

“不”与能愿动词结合,组成:不能、不会等;与某些副词结合,组成:不大、不多等;与某些语素结合,组成:不必、不便、不宜等,构成一个新的否定词。

5.1 新的否定词被共享

若原配句新的否定词后的成分和续配句结构相似,则新的否定词被共享的可能性增大。即:

不能|不会|不要... V1 N1,
V2 N2。

不+能愿动词可以管辖后句 V2 N2。同样地,和基本否定词不同,V1 和 V2 可以不同可以相同。

例:29) 你不能浪费青春,
浪费生命。

30) 他不会出乱子,
闹笑话。

例 29) 新的否定词对后面整体成分否定,加之“浪费生命”和“浪费青春”结构相同,所以续配句“浪费生命”倾向于共享“不能”。如果去掉“能”,因为否定词后的动宾结构和后句动宾结构的动词相同,则续配句不能共享基本否定词“不”。例 30) 同理。

5.2 新的否定词不被共享

“不”与上述语言要素结合紧密,所以如果分别在两个标点句中,不能连用共享。不论原配句“不”后有没有能愿动词,若续配句以这些能愿动词或副词开头,不共享否定词“不”。即:

不[能|会|要...]V1 N1

能|会|要...V2 N2

和基本否定词不同的是,这里的 V1 和 V2 可以

不同可以相同。

例: 31) 她不东想西想, a

能很快睡得象死了似的。 b

a 句的“不”不能和 b 句的“能”在两个标点句中连用共享为“不能”,“不……能……”也形成并列结构,“不”不管辖后句能愿动词。在《围城》中原配句谓语句前有“不能、不必”,续配句缺失成分,共享“不能、不必”的情况如表 3 所示。

表 3 原配句中的“不能”、“不必”被共享情况

相关总句数	续配句不共享否定词,表达肯定的句数	续配句共享否定词的句数	不共享否定词的比例
45	42	3	87.5 %

续配句不共享“不能、不必”的条件除了上面所述句首是能愿动词,还有:续配句句首有关联词、副词(除了程度副词和范围副词)、形容词短语、感官动词、动宾短语的动词带“了”。

例: 32) 四十开外的人也不能一毛不拔,

于是剃去两缕。

33) 他不能喝酒,

吃了些菜。

“不能”表示一种常理,后面不可以用“了”表示的过去态。

6 总结

现代汉语中基本否定词“不”以及扩充词“从不”、“不能”等的共享问题是整个跨标点句句法共享问题的一个组成部分,它们常不被后句单独共享,扩充词“莫不”、“无不”肯定词的共享和否定词不同。合理运用它们,并且找到一定的形式规律,会使语言

运用变得经济有效,而且能促进跨标点句句法共享的研究,提高机器翻译结果的正确性和流畅性。

参考文献:

- [1] 宋柔. 汉语小句前部省略现象初析[J]. 中文信息学报, 1992, 6(3): 62-68.
- [2] 宋柔. 现代汉语书面语中跨小句的句法关系[D]. 香港城市大学: 香港城市大学语言资讯中心, 1999. 12.
- [3] 袁毓林. 并列结构的否定表达[J]. 语言文字应用, 1999, 3: 42-46.
- [4] 陈平. 现代语言学研究[M]. 重庆: 重庆出版社, 1991.
- [5] 吕叔湘. 疑问、否定、肯定[J]. 中国语文, 1985, 4: 245-248.
- [6] 钱敏汝. 否定载体“不”的语义-语法考察[J]. 中国语文, 1990, 1: 30-36.
- [7] 沈开木. “不”字的否定范围和否定中心的探索[J]. 中国语文, 1984, 6: 404-412.

[]表示括号中的内容可有可无, | 表示“或者”。