

# 基于广义话题结构语料库的报告语体与小说语体对比研究<sup>1</sup>

尚英, 宋柔

(北京语言大学 语言信息处理研究所, 北京 100083)

**摘要:** 传统语体学对于语体的描写多集中于词汇、句式、修辞等方面。近年来学者们开始在语法研究中更加重视语体因素, 但是目前的研究多为微观的分析, 没有宏观的理论体系支撑, 难以探索到语体深层次的问题。广义话题理论根据汉语篇章的特点, 以边界明确的标点句为基础, 提出了广义话题和话题结构的概念。本文从广义话题的角度对比了工作报告语体和小说语体的差异, 涉及到命名实体话题、状性话题、谓性话题、逻辑话题和关系话题等。并对这种差异的原因做了合理的解释。虽然工作报告与小说在语体上差异明显, 但没有人从话题-说明的角度进行过比较, 更从未有大规模语料库上的统计分析。本文的工作丰富了统计语体学的理论, 并且为计算机自动分析话题结构、自动评判作文水平、文本按语体分类等应用打下了扎实的基础。

**关键词:** 标点句; 广义话题; 话题结构; 语体;

## A Comparative Study on Report Style and Novel Style Based on Generalized Topic Structure Corpus

**Abstract:** Traditional Stylistics mainly concentrates on vocabulary, sentence pattern and rhetoric to describe the style. In recent years, scholars began to attach more importance to the style in grammar research. However, most of the current studies tend to be microanalysis and have no macro-theoretical support, thus it's hard for them to discover deeper problems of the style. According to the features of Chinese discourse and based on punctuation clause with definite boundaries, Generalized Topic Theory proposes the concepts of general topic and topic structure. The thesis makes a comparative study on the stylistic differences of work reports and novels, involving topic of named entity, adverbial topic, predicate topic, logical topic and relational topic, etc. The thesis also makes a reasonable explanation for these differences. Although work reports and novels differ greatly in style, no one has ever made a comparative study from the perspective of topic-comment, and there is no extensive statistical analysis based on a corpus up to now. The work of this thesis enriches the theory of statistical stylistics, and lays a solid foundation for the application of computer auto-analysis of topic structure, auto-judgment of composition level, and textual classification by style, etc.

**Key words:** Punctuation clause; Generalized topic; Discourse structure; Style

### 1 引言

传统的汉语语体研究主要在修辞学、风格学范围内进行, 主要成果集中在语体的分类以及各种语体风格特点的描写, 后者又多集中于词汇、句式、修辞格的运用等方面。近年来随着功能主义语法研究的兴起, 学者们开始在语法研究中更加重视语体因素, 尤其以陶红印(2007,2010)、张伯江(2007)等学者为代表, 他们强调语体意识在语法研究中的重要性, 并就某些语法现象进行语体调查, 从语体方面给出合理的解释。这方面的研究对汉语语法和语体研究都非常有价值, 但是目前的研究多是对某一语法现象的微观考察, 并没有宏观的理论体系支撑, 难以探索到语体深层次、规律性的问题, 而且从语用层面的话题角度进行语体对比研究的还很少见, 基于语料库统计的语体对比研究更是没有。

广义话题理论根据汉语语篇的特点, 以边界明确的标点句为基础, 提出了广义话题和话题结构的概念, 在数十万字、多种语体语料标注的基础上, 归纳了汉语的话题结构的流水模型和规律, 为汉语的篇章分析提供了理论基础和形式化手段。同时为我们的汉语研究提供了新的视角。

不同语体在广义话题理论的视角下, 呈现出不同的特点, 对于这些特点的研究一方面可以探索广义话题理论在不同语体中的覆盖情况, 另一方面可以挖掘更多的语体特点, 丰富现有的语体特征空间, 同

<sup>1</sup> 本文得到国家自然科学基金(61171129)的资助。

时为自然语言的文本按照语体自动分类、计算机自动分析话题结构、自动评判作文水平等应用提供参考。

本文选取了工作报告与小说两种语体进行比较。虽然这两种语体差别明显，但是目前尚未见到有人从话题-说明的角度对其进行描述和对比，更未见有大规模语料库上的统计分析。

本文首先简要介绍广义话题理论的相关概念，然后基于小说语体和报告语体的标注语料，在广义话题理论下对这两种语体在话题结构方面的差异进行比较，最后在此基础上进行原因分析和总结。

## 2 广义话题理论相关概念介绍

标点句就是逗号、分号、句号、叹号、问号、直接引语的引号以及这种引号前的冒号所分隔出的词语串。广义话题结构以标点句为基本单位。

如果标点句中的一个成分（包括整个标点句）被另一些标点句谈论，则称前者为话题，后者为它的说明。话题和它的所有说明组成一个话题结构。话题所在标点句中话题后面的部分如果也是谈论它的，那么也是它的一个说明。

这里强调话题至少要被所在标点句之外的另一些标点句谈论，因此构成话题结构的标点句一定多于1句。就是说这里的话题是一个篇章层面的概念，一定要在标点句序列的语境中才能显现。一个标点句不能成为话题的支持环境，其中的成分只是潜在的话题。潜在的话题被它所在的标点句中后面的成分说明，只是得到潜在的话题结构。只有被其它标点句所说明，潜在话题才能实现为话题，潜在的话题结构才能实现为实在的话题结构。在下文中，为了行文简炼，在不致引起混淆的情况下，把潜在话题、潜在话题结构也说成话题、话题结构。

话题是被谈论的对象。从语义上看，除了通常讲的实体类和时间、处所类话题外，还扩充进来了状性、谓性、推理前提等类型的话题，因此称为广义话题。广义话题与其说明所组成的结构称为广义话题结构。

如：

```
c1 {本宪法 [ {以法律的形式 [确认了中国各族人民奋斗的成果，  
c2           规定了国家的根本制度和根本任务， ]}  
c3     是国家的根本法，  
c4     具有最高的法律效力。 ]}
```

这是一个由4个标点句组成的话题结构。方括号括起来的是若干说明，方括号左面是话题，花括号括起来的是话题结构。其中c1后半部分“以法律形式确认了中国各民族人民奋斗的成果”和c2组成了内层的以介宾结构“以法律的形式”为话题的状性话题结构{以法律的形式[确认了中国各族人民奋斗的成果，规定了国家的根本制度和根本任务，]}；内层话题结构连同c3、c4一起都是“本宪法”的说明，组成了外层话题结构。

为了便于直观认知，把标点句分行排放，用于说明的标点句，其左端缩进后顶在它说明的话题的最后一个字后面，这种排列方法我们称为换行缩进图式。

关于广义话题结构的详细说明见[1]。

## 3 报告语体与小说语体广义话题视角下的对比

从语体的角度来看，小说和工作报告分别属于文艺语体和公文语体。两种语体在词汇、句式、修辞方式等风格方面的差异性很大，前人也做了较为详细的论述，我们拟从广义话题视角下对两种语体的话题及话题结构进行详细对比，以期从深层挖掘两种语体的差异，为语体分类提供更为细致的特征描述。

我们所做调查研究基于6079个标点句的小说体语料，5893个标点句的报告体语料。小说语料包括《鹿鼎记》（第一回）、《邢老汉和狗》、《围城》（第五章前2150句）。报告体语料主要来自政府工作报告，包括1983年、1985年、1996年、2005年政府工作报告以及《胡锦涛在庆祝中国共产党成立90周年大会上的讲话》。这些语料都经人工标注了广义话题结构。这两部分语料在下文中分别称作《小说》和《报告》。

下面我们基于这两种语体的标注语料，从命名实体话题、状性话题、谓性话题、逻辑话题、关系话

题几个方面分别对两种语体进行比较。

### 3.1 命名实体话题

命名实体话题包括时间话题、处所话题和实体话题三大类。

#### 3.1.1 时间话题

我们说的时间话题不是词类的概念，而是指能够起到时间表达作用的语言单位。包括时间词、与时间表达有关的方位短语、介词短语等。

我们将时间话题分为绝对时间话题和相对时间话题两种。绝对时间话题以确指的时间名称词语表示，如：“1985年”，“‘九五’期间”，“从1978年到1982年”等。相对时间话题有两种：一种是以说话时间为参照点的时间话题，如：“今天”，“下个月”，“今后五年”；另一种则是以某一事件的发生为参照点表达的时间话题，如：“中国同日本在1978年签订和平友好条约以来”，“抗战时期”等。

《报告》中有140例时间话题，共涉及396个标点句。

《小说》中有157例时间话题，共涉及357个标点句。

首先，工作报告中绝对时间话题出现频次明显高于小说。

时间话题类型	《报告》中频次	《小说》中频次
绝对时间话题	41	7
相对时间话题	99	150

《小说》中出现的绝对时间话题只有7例，其中《鹿鼎记》5例，《邢老汉》与《围城》各1例。

其次，两种语体中同是相对时间话题，语义内容也有所不同。

相对时间话题包括以说话时间为参照点和以事件发生为参照点两种。

《报告》中的相对时间话题最常见的如：“过去的五年”，“今年”，“今后五年”，“去年”，等等。从形式上看，它们都是相对时间话题，但因为工作报告发布的时间固定、明确，从时间话题真正表达的语义来说，相对时间话题是确指的某个时间位置，我们完全可以将相对时间话题转换为绝对时间话题。如1996年的政府工作报告中的“今后五年”指的就是1996-2000年。可以说工作报告中有一部分相对时间话题形式表达了绝对时间话题的意义，这类时间话题有73例，占相对时间话题的73.74%。

《报告》中也有少数相对时间话题具有模糊性：如“当前、目前、近年来、现在”，有26例，占相对时间话题的26.26%。

小说内部的时间序列沿着自己的逻辑顺序向前发展，环环相扣，彼此呼应，脉络清楚。尽管如此小说中相对时间话题是相对于说话时间或者某些事件的模糊时间，不能转换成绝对时间。模糊的原因在于，不论参照点是说话时间还是事件发生时间，都是模糊的，不清楚的。比如《邢老汉》中，邢老汉与要饭女人见面的一段话，

一天中午，邢老汉正准备做饭，……一会儿，要饭的女人看出了这个老汉做饭时笨手笨脚，就小声地说：“……。”

以“一天中午”邢老汉与要饭女人见面这一事件为参照点，“一会儿”“晚上”等表达的时间序列是清楚的，但是参照点本身“一天中午”是泛指的、模糊的，导致了以此为参照点的相对时间的模糊性。

小说时间话题的模糊性与小说本身并不追求时间刻度的精确性有关，并且小说本身是虚拟的人物、情节，时间的精确与否并不影响人物的塑造、情节的发展。时间话题于小说的意义在于提示情节自然且合逻辑的过渡，同时为篇章的衔接与连贯提供必要的手段。

### 再次，两种语体中时间话题的度量单位大小不同。

不管是相对时间话题还是绝对时间话题，工作报告中多数都是“年”为单位计数的时间话题；有时候出现“月、季度”等时间词。以“天”为单位的时间话题，《报告》中一共出现了6例，其中5例是确指讲话时间的“今天”，1例是对于特殊日子的说明，如下。

1997年7月1日和1999年12月20日，

我国将相继对香港和澳门恢复行使主权。

《报告》中以“月”、“季度”为单位的时间话题2例，明确以“年”为单位的时间话题87例。另有一类时间话题，如“八五期间”、“在本世纪上半叶”、“当前”、“现在”等实质语义以“年”为单位，这类时间话题有45个。以上两种《报告》中全部时间话题的95.71%。

小说中时间话题多是某天或者一天里的某个时间，如：“有一天，第二天，三天以后、从早晨起，吃晚饭时，晚饭后”等等。在《小说》中这类时间话题有130例，占《小说》中全部时间话题的82.80%。

《小说》中也有以“年”为单位或更大时间度量单位（如：朝代）的时间话题，以《鹿鼎记》最多，共18例；《邢老汉和狗》及《围城》共2例。《鹿鼎记》中我们选取的第一回，文中有很多篇幅涉及到作者对于社会背景的介绍以及小说人物语言中对于朝代更迭前后的社会情境介绍，以“年”为单位或更大时间度量单位的时间话题较多。

### 最后，两种语体中时间话题对于整个话题结构的意义也不相同。

小说中的时间话题是为人物行为动作或者事件发展变化提供一种环境信息，如：

二十二日下午近五点，

兄弟俩出门，

报告中的时间话题有时已经成为一种真正的实体信息。工作报告本身是对过去一段时间的总结，对未来一段时间的规划部署，所以时间是其陈述的对象或者是陈述对象的核心属性之一。比如：

刚刚过去的1984年，

是我国各族人民在实现八十年代以至九十年代的三大任务方面取得伟大成就的一年，  
是全国政治、经济、文化等各条战线的形势越来越好的一年。

此例中的时间话题便是话题结构的陈述对象，可以认为是狭义的话题。这类时间话题《报告》中有10例。

报告中更大量的时间话题，不是说明的核心对象，但它是描述社会发展状况或者复杂关系时必要的核心要素之一，这种时间话题一旦消失，话题结构的意义将无法正常表达。如：

1984年，

我国工农业总产值突破10,000亿元，

在1979到1983年平均每年递增7.9%的基础上，

比上年增长14.2%

此例中，“我国工农业生产总值”是说明的核心对象，但是离开“1984年”这个属性，所表达的语义是不清楚、不完整的。因此“1984年”这个时间话题成分所表达的“我国工农业生产总值”的必然属性，不可缺失。

统计数据显示，《报告》中这类时间话题有108例，占77.14%。

### 3.1.2 处所话题

	《报告》频次	《小说》中频次
处所话题	24	52

我们所说处所话题仅指语义上表示现实空间位置的话题。报告中的处所话题一般并不是指事情发生在某个具体的地点，而是泛指。如：“在城镇”，不是指在哪一个城镇，而是全国范围内所有的城镇，

“城镇”代表一个类属的意义，与之相对的是“在农村”。而小说中的处所话题多是行为处所，其作用在于为动作行为提供一种环境信息。

江南近海滨的一条大路上，

一队清兵手执刀枪，  
押着七辆囚车，  
冲风冒寒，

《报告》中只有两例处所话题指出的是行为发生的具体处所：“在黄河上游、长江中上游干支流和赤水河流域”，“在山西、内蒙古、安徽两淮、贵州六盘水等地的煤矿区”。

此外工作报告与小说中处所的度量范围不同。报告中大多数处所话题指称的度量范围要大于小说。

### 3.1.3 指人的名词性实体话题

指人的名词性话题在两种语体中的差异较大。

工作报告中所涉及到的指人的名词性话题都是集体称谓或者具有某种共同特征的群类，如“全国各族人民、人民解放军、中国政府和人民、国有企事业单位的领导人员、城乡人民”。这类话题《报告》中共有 21 例。

小说对于人物形象的塑造是其灵魂所在，所描写的人物都是独立的个体。小说中指人的名词性话题多是专有名词。《报告》中人名话题未出现。

### 3.1.4 指人的人称代词话题

报告中做话题被共享的人称代词基本都是“我们”，代表发布报告的组织机构，《报告》中以“我们”做话题的话题结构有 55 个，占《报告》中全部人称代词话题的 63.95%。只有下面 1 例以人称代词“他们”为人称代词话题。《报告》中未出现第二人称的代词话题。

他们全心全意为人民服务，

表现出了“先天下之忧而忧，后天下之乐而乐”的高尚思想和风格，

报告中还有一类话题与人称代词有关，就是“人称代词+名词”做话题。如《报告》中“我国” 26 例，“我们国家” 4 例。

小说中的人称代词话题更普遍。其中第三人称代词做话题的最多。

人称代词话题	《小说》中频次
我（们）	51
你（们）	27
他（们）/她（们）	101
咱（们）	8
合计	187

### 3.1.5 指物名词性话题

指物的名词性话题在两种语体中都存在，其最大的差异在于工作报告中指物的名词性话题多指称抽象事物，而小说中的指物名词性话题则多指称具体事物。本文在此不再详细论述。

## 3.2 状性话题

我们在此将从介宾结构话题和副词性话题两方面对比两种语体的状性话题。

### 3.2.1 介宾结构话题

	《报告》中频次	《小说》中频次
介宾结构话题	288	46

我们这里讲的介宾结构话题，不包括介宾结构形式指时间和处所的话题。

首先，从上表看出，《报告》中的介宾结构话题出现频次明显高于《小说》，并且《报告》中常常出现一个标点句前并列存在两个及以上的介宾结构话题。如：

随着生产建设的持续增长，  
通过各种渠道，

从1979年到1984年安排了4,200多万城镇劳动力就业，

其次，两种语体的介宾结构话题中介词有所不同。“在”字介词结构话题在两种语体中都是出现频率较高，工作报告中“在”通常形成框架式介词结构，如：在……下，在……中，在……方面，在……领域等；并与表达抽象意义的词语搭配，如：“在以江泽民同志为核心的党中央领导下”、“在我国中长期发展中”、“在水利、能源、交通、通信和重要原材料工业方面”等。“在”字介宾结构话题表达的语义是社会行为的范围、领域、条件。

报告中的介宾结构话题常用的介词呈现出书面化强的特点，多用二字介词。如：“经过、通过、根据、按照”等都是报告介宾结构话题中用的较多介词。

《小说》中出现频率最高的是“将”字介宾结构话题，8例。《小说》中介宾结构话题常用介词多为单字介词。如：

你对结婚和做事，

一切比我有信念。

两种语体中高频的介宾结构话题所用介词如下表：

《报告》		《小说》	
介词	频次	介词	频次
在	106	将	8
对	32	在	7
经过	20	对	4
对于	17	于	3
根据	14	因	2
按照	14	凭	2
通过	14	以	2

### 3.2.2 副词性话题

	《报告》中频次	《小说》中频次
副词性话题	57	53

《报告》中的副词性话题，充分体现了报告语体的正式性。所用的副词性话题除高频副词“将、都、就、也”外，多用二字形容词状语、副词状语和“四字格+地”。常见形容词、副词状语有“努力、坚决、仍然”等，四字格如“始终如一、坚持不懈、始终不渝、千方百计”等。而《小说》中副词常用的有“都、就、便、也、竟、却”等。

### 3.3 谓性话题

#### 3.3.1 话题-说明是述宾关系

两种语体中都存在大量的标点句以前一个标点句中述语做话题，但是做谓性话语的谓词差异性较大。

小说中充当话题的述语多是“说”、“道”，如：

魏老汉说：

“庄户人的狗要有这个本事，  
咱就不种庄稼了，  
领着狗四处耍把戏去。”

而政府工作中充当话题的述语最高频的动词为“是”，其他有：“证明、看到、告诉我们、继续”等，

《报告》一共有 146 例话题结构以述语充当话题。“说、道”未出现。

如：

我们的奋斗目标是，

让人民群众喝上干净的水、呼吸清新的空气，  
有更好的工作和生活环境。

要继续加强和改善宏观经济调控。

保持合理的固定资产投资规模和在建规模，  
加大投资结构调整力度，  
提高投资效益。

两种语体中述宾关系的话题-说明结构中的高频述语使用情况如下：

《报告》		《小说》	
述语	频次	述语	频次
是	55	道	269
证明	8	说	69
看到	7	想	20
告诉	7	是	17
认为	6	知道	5
其他述语	63	其他述语	52
总计	146	总计	432

谓词性话题以工作报告中更为常见，以下几类谓词性话题是工作报告的特有话题。

### 3.3.2 话题是谓词短语表致使义

报告中有一些标点句以动词“使”“致使”“以”“以利于”等开头，以前面谓性标点句为话题，表示前面行为的结果或者目的。《报告》中致使类话题结构有 29 个。如：

在国家机关和事业单位，

实行以职务工资为主的工资制度，

使职工工资同本人所肩负的职务、责任和劳绩密切联系起来。

《小说》中仅见 1 例。

### 3.3.3 助动词话题

以“要/必须/应当”等助动词为话题的话题结构是工作报告特有的类型。从句类的角度看，这种话题句都属于祈使句，祈使句是要求对方做或禁止做某事的句子，一般分为表指令、表号召、表请求、劝阻等几类。在工作报告中，常使用的是表号召和表指令祈使句，表号召多用“要”，表指令多用“必须”。

“要”类话题结构出现的几率最高，一般“要”类话题的说明部分有多个标点句，《报告》中“要”类话题结构有 270 个，共涉及 1074 个标点句。最多的“要”类话题结构包含 20 个标点句。下面是包含 19 个标点句的例子：

要继续稳定和完善多种形式的家庭联产承包责任制，

贯彻执行“决不放松粮食生产，积极开展多种经营”的方针，

同时逐步增加对农业的投资，

积极改善农业的生产技术条件。

进一步加强农田水利建设，

重点改善商品粮棉地区的排灌条件。

加快发展化肥工业，  
逐步提高复合肥料以及磷肥和钾肥在化肥施用量中的比重。  
积极发展饲料工业，  
大幅度提高配合饲料的产量，  
努力加强和改善草原建设，  
以∞促进畜牧业的发展。  
努力加强农业科学技术的研究和运用，  
抓紧优良品种的培育和推广，  
因地制宜地采取改良土壤、改革耕作栽培制度、防治病虫害等综合技术措施，  
提高单位面积产量。  
加快造林育林的步伐，  
合理利用现有森林资源，  
继续采取有力措施坚决制止乱砍滥伐森林的现象。

《报告》中“必须”话题结构 67 个，共涉及 200 个标点句。

必须从各自的实际情况出发，  
讲究经济效益，  
服从全国的综合平衡。

《小说》中一共出现以助动词为话题的话题结构 10 个，涉及 24 个标点句。

### 3.3.4 以动词为中心的偏正结构做话题

工作报告中有一类话题，从形式上来看是“np (vp) +的+vp”，与单纯的 Vp 相比指称性凸显，但动词原来所有的动作意义没有发生转指，仍然是自指。并且“的”字有时可以省略而不影响话题结构及意义。因此我们把这类话题也看做谓性话题。这种话题结构从语义上来讲，是对机构行为、举措的意义或结果进行说明。《报告》中此类话题结构有 46 个。如：

改革的推进，  
为国民经济和社会发展提供了强大动力，  
也为本世纪末初步建立社会主义市场经济体制奠定了基础。

对外开放的扩大，  
促进了国内经济发展，  
推动了经济体制改革，  
也增进了同各国政府和人民之间的交往。

《小说》中此类话题结构仅有 1 例，如下：

这一刹那的接近，  
反见得睽隔的渺茫。

## 3.4 逻辑性话题

对于这种广义话题，我们有明确的限制，就是这种推理关系必须是一种泛指规则，从形式上看两个标点句都没有实体主语，但有关联词，如“只有……才能……”、“要……只要……”等。

如：  
只有继续深化改革，  
才能解决妨碍经济社会发展的深层次矛盾。

这类话题结构在《报告》中有 14 例，而《小说》中我们只发现了 1 例，如下：

要上头满意，  
就得让大伙儿吃点亏。



### 3.5 关系话题结构

工作报告中对于复杂关系的描述，从表面语义上看是实体、时间、处所类话题结构，实质上是一种关系结构，这也是政府报告不同于小说的一个方面。如：

1982年底，

城乡储蓄余额达675亿元，

比1978年增长2.2倍。

上面的话题结构说明的是“1982年城乡储蓄余额”与“1978年城乡储蓄余额”的关系，前者比后者增长2.2倍。这种话题结构一般出现在报告中工作总结的部分。这一例表达的关系比较清晰简单，而有的关系话题结构很复杂，有的不能简单地用换行缩进图式表示各要素之间的关系，需要对前面标点句的部分成分进行提取、调序才能将关系表达清楚。工作报告中我们发现33例关系话题结构，其中有11例话题结构就属于这种情况。不能清楚表达的原因，有的是关系方要素齐全却不成句，有的是虽然能成句，但是关系方要素被省略，语义表达不完整。前一种情形如：

1984年银行的各项贷款总额比1983年增加28.9%，

其中12月一个月的增长额就占全年增加额的48.4%。

“1984年银行的各项贷款总额其中12月一个月的增长额就占全年增加额的48.4%。”不成句。

后一种情形如：

2004年，

国内生产总值达到13.65万亿元，

比上年增长9.5%；

财政收入2.63万亿元，

增长21.4%；

第四个标点句是成句的。但关系要素“比去年”被省略了。

目前我们看到最复杂的关系话题句为：

1982年同1978年相比，

aaaaaaaaa bbbbbbbbbb 社会商品零售总额由1559亿元增加到2570亿元，

增长64.8%，

cccccccccccccc 平均每年增加253亿元{{cb到a}}，

ddddddd 比1978年前的26年平均每年增加49亿元{{cb到a的d量}}，

高出4倍多。

这个话题结构最后一个标点句要表达的意思为：“社会商品零售总额1978年到1982年的平均每年增加量比1978年前的26年平均每年增加49亿元高出4倍多。”这一复杂的关系语义的表达，已经达到了汉语表达能力的极限，更适合于用公式、表格等形式化手段来表达。广义话题结构的现有模型对于这类结构的描述也有局限性，需要进行扩充。

## 4 两种语体广义话题差异的原因分析

我们认为两种语体在广义话题上的差异，根本原因在于两种语体语用意义本质不同。

小说是文艺语体的一种，从表达方式上讲，小说是叙事的，通过对环境的描写以及个体人物行为、命运的叙事性描写反映社会。我们所说的工作报告，一般指机构工作报告，属公文语体，就我们选取的语料政府工作报告来讲，内容主要包含过去一年的政府工作情况总结与当年的工作部署。政府报告的本质是对政府社会行为、政府指令的表述以及对各种复杂关系的客观说明。

由于小说核心在于塑造人物、以个体为对象进行叙事，因此时间、处所话题也必然为这一核心服务。其功能在于为推进小说情节发展，提示人物、事件变化，为个体行为动作提供附属的环境信息。从时间刻度上看，小说对于个人行为的描述更注重细节，个人行为的时间跨度小，时间话题的度量单位小。并且，小说中的时间转换依照人物、事件的发展变化往前推进，这就导致了相对时间话题多见。小说中行

为主体多为小说中拟塑造的人物形象，因此专有名词话题、人称代词话题多见；小说中人物语言都直接出自所塑造的人物之口，因此小说中述语话题以“说/道”为最为普遍。小说中个人的行为动作主体、客体、及关系语义相对简单，本身对于介宾结构的需求并不强烈，又因为句子较短，介宾结构与中心述语中间不用停顿而是组合成完整的标点句，难以形成介宾结构话题。

工作报告是对机构的社会行为、指令的表述以及对机构的社会行为、社会范围内的重要事件以及各种关系的说明。社会行为的时间跨度比较大，一般不会精细到某一天，某一天某一点的时间，因此从时间话题的刻度上讲，其度量范围多以“年”为单位，明显比小说要大。时间之于工作报告往往是描述对象之一或者是关系方的核心要素之一，时间话题一旦消失，关系方将无法进行关系描述，多用绝对时间话题也是工作报告内容客观而非虚拟的必然要求。

工作报告不管是过去的工作还是将来的工作所涉及到的主体都不是个人行为、个人事件，其主体都是国家、各级政府机构以及各种集体。因此人称代词话题多用“我们”，指人的话题多是集体称谓。由于社会行为、社会指令以及各种关系，涉及到社会行为的主体、条件、范围以及关系的关系方以及关系方的多个必要要素等等，简单的主谓宾结构难以容纳下那么多内容，为了使报告内容全面客观，突出主体、对象、范围、手段、依据等信息，会选择使用介词结构来引入。如果所涉及的要素不止一个，可以同时使用多个介宾结构引入多类要素。这样往往就会形成谓语中心前有多个介宾结构的情况。

工作报告的工作部署是其核心之一，指令性是这一部分本质内容，因此常常使用的是表号召和表指令的祈使句，表号召多用“要”，表指令多用“必须”。又因政府工作的复杂性，某一项工作涉及面广，一旦发出指令也涉及多方面，因此“要”类话题结构往往涉及多个标点句。

工作报告站在一定的高度上，在总结规律的基础上号召机构、团体成员按照正确的规律工作。规律的总结与提炼是工作报告高度与深度的体现，因此逻辑性话题结构在报告中较为普遍。而小说以个体为中心，对个体的行为表述上升到泛指的规则一般来说没有必要。

## 5 结语

广义话题理论根据汉语篇章的特点，以边界明确的标点句为基础，提出了广义话题和话题结构的概念。本文从广义话题的角度对一定规模的报告语体文本和小说语体文本进行了标注和调查研究，发现了两种语体在广义话题上的差异，给出了差异的实例和数据，并对这种差异的原因做了合理的解释。

需要说明的是，机构有大有小，我们选取语料为政府工作报告，其他各级地方政府、各类组织机构的工作报告在广义话题上会与政府工作报告呈现出一定的差异，比如：时间话题的度量单位、处所话题的所指范围等会有大小的差异。但是由于工作报告语体本质的一致性，它们在广义话题类型、话题结构类型等方面仍具有强相似性。

本文首次从话题-说明的角度对工作报告和小说语体进行比较，一方面展示了广义话题理论在不同语体中的覆盖情况，丰富了汉语语体对比以及统计语体学的理论，另一方面为计算机自动分析话题结构、自动评判作文水平、文本按语体分类等应用打下了扎实的基础。

### 参考文献

- [1] 宋柔. 汉语篇章广义话题结构的流水模型 [J]. 中国语文 (待发表).
- [2] 陶红印. 试论语体分类的语法学意义 [J]. 当代语言学, 1999, (3): 15-24.
- [3] 陶红印. 操作语体中动词论元结构的实现及语用原则 [J]. 中国语文, 2007, (1): 3-13.
- [4] 陶红印, 刘娅琼. 从语体关系到语法差异(上)——以自然会话与影视对白中的把字句、被动结构、光杆动词句、否定反问句为例 [J]. 修辞学习, 2010a, (1): 37-44.
- [5] 陶红印, 刘娅琼. 从语体关系到语法差异(下)——以自然会话与影视对白中的把字句、被动结构、光杆动词句、否定反问句为例 [J]. 修辞学习, 2010b, (1): 22-27.
- [6] 王德春. 语体学 [M]. 南宁:广西教育出版社, 1987.
- [7] 张伯江. 语体差异和语法规律 [J]. 修辞学习, 2007, (2).
- [8] 赵元任著. 汉语口语语法(吕叔湘译) [M]. 北京:商务印书馆, 1979.
- [9] 朱军. 汉语语体语法研究综述 [J]. 汉语学习, 2012, (5): 72-79.