

职务名称的构成形式及分布分析*

张金玲, 何伟

(中国传媒大学文学院, 中国传媒大学国家语言资源监测与研究中心有声媒体语言分中心, 北京 100024)

摘要: 本文对职务名称的定义、构成形式、音节特征以及使用频次等进行了研究与分析, 提出了职务名称的一般构成形式, 详细分析各构成要素, 并总结出各种类型职务名称的分布规律, 为实现职务名称的自动识别奠定了基础。本文选取 2010 年全年的新闻联播节目文稿作为语料, 借助中国传媒大学自动分词标注系统提取了语料中全部的职务名称, 并对得出的职务名称, 进一步统计与分析。语料中 444 个职务名称, 为本文的统计分析工作提供了数据支撑。

关键词: 职务名称; 类词缀; 新闻联播

中图分类号: H08

Constitution forms and distributions analysis of position titles

ZHANG Jinling, HE Wei

(National Language Resources Monitoring and Research Center, Broadcast Media Language Branch, College of Liberal Arts of Communication University of China, Beijing 100024, China)

Abstract: The paper analyses position titles from the definition, form, syllable characteristics, structure characteristics and frequency. In the end, we put forward the general constitution form, analyses each element of the form and summarizes the distributions of all kinds of types, which lay the foundation for realizing the automatic recognition of the position titles. The paper uses all the text corpora of CCTV News in 2010, extracts all the positions titles with Automatic word segmentation tagging system of the Communication university of China, then analyses further. 444 positions titles from corpora provides data support for the statistics and analysis.

Key words: position titles; the class affix; CCTV News

职务名称标示着人物的身份信息、社会地位和工作职责, 在社会领域中具有重要的作用。总结和分析职务名称的构成形式和分布规律, 有利于实现职务名称的自动识别, 从而推动信息抽取和信息检索等研究的发展。

目前, 专门针对职务名称的形式特点的研究并不是很多。其中, 有些文章主要从语言与文化的角度研究职务名称。但命名实体的研究与之具有极大的相关性, 我们可以借鉴命名实体的研究方法和角度, 如命名实体识别中基于规则的方

* 收稿日期:

作者简介: 张金玲 (1987.01—), 女 (汉), 山东淄博, 中国传媒大学文学院应用语言学专业硕士研究生, 主要研究领域: 计算语言学, 中文信息处理, E-mail: zhangjinling20062006@126.com

通讯作者: 何伟, 中国传媒大学文学院应用语言学专业硕士导师, 中国传媒大学国家语言资源监测与研究中心有声媒体语言分中心副研究员, E-mail: hewei@cuc.edu.cn

法。本文中，我们专注于带有“类词缀”的职务名称的研究。现代汉语中关于“类词缀”现象，语言学界的很多研究和讨论，主要研究的内容集中于类词缀的性质、与词缀、词根相比的特点以及新词语中出现的类词缀。此外，类词缀也应用于中文信息处理中，北京大学计算语言学研究所的《现代汉语语料库加工—词语切分与词性标注规范与手册》中，作为具有类化作用的后接成分为分词与词性标注服务。

1 职务名称的构成形式分析

1.1 职务名称的一般构成形式

职务名称在构成形式上存在共同点，具有一定的规律性。利用派生词的构词方法，词根指词语中表达概念意义的部分，词缀是词语中只起到虚化作用或语法作用的部分。还有一些类似于前缀或后缀的成份，叫做“类前缀”或“类后缀”：意思是意义已经开始虚化，但是，还没有达到真正的词缀那样的程度。¹类词缀表达的意义处于词根和词缀中间，既没有实词意义那么实，也无虚词的意义那么虚。

对于具有相同“类词缀”的职务名称，一般的构成形式为：

$P=(A)+(B)+C+(D)$

其中，P：职务名称，A：限定语；B：类前缀；C：词根；D：类后缀；（）表示可省略。具体类型又可以分为：

$P=B+C$,类词缀+词根，如，总编辑 副主任

$P=C+D$,词根+类后缀，如，董事长 设计师 评论员 教育家 发言人 研究者

$P=B+C+D$,类前缀+词根+类后缀，如，总队长 总会计师 副委员长 副部长

$P=A+C+D$,限定语+词根+类后缀，如，高级工程师 首席技术官

另外，还包含少数其他的构成形式。

如，a) $A+B+C$,限定语+类前缀+词根

执行总经理 执行总指挥 救灾总指挥

b) $B+A+C+D$,类前缀+限定语+词根+类后缀

副主任委员

1.2 职务名称的构成要素

1) 类词缀

根据与词根的相对位置，类词缀包含类前缀和类后缀。职务名称中的类前缀主要有“总”“副”；类后缀主要有“长”“师”“员”“人”“官”“家”“者”“手”，分别构成“总××”“副××”“××长”“××师”“××员”“××人”“××官”“××家”“××者”“××手”。

2) 词根

a) 词根表示职位名，职位名反映任职者的权力²。多与类前缀“总”“副”及类后缀“长”连用。

司令 经理 理事 参谋 董事 委员 秘书 首相 主席 总统 主任 领队 国务卿 总裁 总理 护士 领事

b) 词根表示职类名，职类名反映任职者工作内容²。多与类后缀“员”“师”“人”“者”“家”连用。

设计 会计 摄影 魔术 麻醉 理发 美容 药剂 建筑 经济 工程 农艺 摄像 检察 外交 指挥 联络 参议 乘务 运动 飞行 研究 宇航 公务 驾驶 评论 指导

指战 观察 技术 辅导 消防 航天 管理 普查 讲解 调度 监督 调研 服务 教练
调解 巡视 预报 众议 交易 潜水 裁判 宣讲 护林 教导 安检 联络 售票 宣传
监察 调查 放映 饲养 创作 炊事 管水 投递 协调 协管 测评 化验 试飞 信贷
巡线 引导 安保 安全 播音 操作 乘务 代理 检测 检察 检查 检验 交通 解说
联合会 领航 农技 陪审 气象 示范 特派 推销 维护 销售 巡逻 业务 营业 邮递
革命 教育 活动 歌唱 作曲 探险 发明 冒险 思想 登山 航海 军事 漫画 摄影
书画 主持 领导 负责 发言 代言 制片 经纪 志愿 工作 建设 消费 投资 劳动
组织 经营 援建 管理 领导 发明 研究 表演 创作 创办 创立 监督 游泳 播种
策划 创建 驾驶 监管 建言 解放 拍摄 培育 设立 宣传 制定 制作

c) 词根部分表示“学科领域”，多与类后缀“家”连用。

科学 数学 藏学 社会学 地质学 法学 火山学 美学 气象学 生物学 哲学 地震学
汉学 核科学 建筑学 历史学 经济学 考古学 天文学 人类学 生理学 文学
物理学 政治 音乐 艺术 美术 书法 金融 剧作 舞蹈

d) 词根部分表示“组织机构”，多与类后缀“家”“长”连用。

院 会 团 行 馆 州 署 站 台 园 银行 企业

e) 词根部分表示“领域名称”：多与表示职务品相的类后缀“师”连用，倾向于指在某一领域具备一技之长。

律 医 法 技 厨 军 画 教 艺 商

f) 词根部分表示“事物名称”，常与类后缀“师”“长”连用。

船 机 车 机 旗 枪 鼓 棋 红旗 机械 列车 飞机

g) 词根具有一定的顺序性和等级性，常与带有等级观念色彩的类后缀“长”连用，可以表现等级的高低，多为处于分类语义场的职务名称。

省—市—县（区）—镇—乡—村

队—总队—大队—中队

部—厅—司—局—处—所

军—师—旅—团—营—连—排

校—班—组

h) 其他词根

理 统 裁 监 督 编 工 诗 蛙 佣 哲 警 士 主 作 专 管 记 使 读 学 歌
御 骑 乐 推 助 射 议 导 技 厨 幼 委 党 官 队 演 会 球 社 盟 店 科

3) 限定语

职务名称中的限定语，既有表示级别的“高级”“首席”，也有限定具体的领域范围或内容的词语。

高级 首席 助理 执行 外交 国防 能源 退伍 产业 复员 第一 学部 抢险
减灾 政法 残奥 救护 考察 考古 政治 主创 防暴 救援 演职 国务 护卫 突击
共产 候补 少先 司乘 救护 共青 常务 神职 粮食 田径 救灾 大学生

2 职务名称的音节及结构分析

本文研究的语料来源于中国传媒大学有声媒体语言研究中心，选取的是2010年新闻联播365个文本。首先对365个文本语料进行切分标注处理（职务名称被标注为n，名词词性），然后利用TCRS（中国传媒大学熟语料库检索系统）统计词频和出现文本数，最后通过Microsoft Excel软件检索出带有“类词缀”的职务名称。

所选语料中的职务名称在音节分布上具有自己的特点。总体上看，以二音节、

三音节、四音节居多，对含有不同“类词缀”的词语而言，音节分布情况又各有不同。

“总××”“副××”中，三音节词语占多数，“副××”三音节比例以 74.3% 占绝对优势；“总××”的三音节比例也高于双音节比例。四音节、五音节所占的比例相对较少。

“××长”“××手”中双音节占优势，占 70%以上；“××员”“××家”“××者”中三音节占优势，“××者”中的三音节占 86.9%；“××人”“××师”中三音节比例高于双音节；“××官”双音节比例略高于三音节比例；“××家”“××人”“××师”“××员”“××官”中，四音节所占的比例相对较多；“××官”“××师”中，五音节占有一定的比例。

表 1、职务名称的音节分布

“总××”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	22	31.8%	36.4%	18.2%	22.7%	
“副××”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	35	0	74.3%	17.1%	8.6%	
“××长”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	65	70.8%	20%	6.3%	1.5%	
“××师”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	29	37.9%	44.8%	10.3%	6.9%	
“××员”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	六音节比例
	137	14.6%	60.6%	20.4%	3.6%	0.73%
“××人”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	24	37.5%	41.7%	16.7%	4.2%	
“××官”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	
	18	38.9%	33.3%	16.7%	11.1%	
“××家”	词种数	双音节比例	三音节比例	四音节比例	五音节比例	六音节比例
	57	7.0%	61.4%	26.3%	1.8%	1.8%
“××者”	词种数	双音节比例	三音节比例	四音节比例		
	38	13.2%	86.9%	0		
“××手”	词种数	双音节比例	三音节比例	四音节比例		
	19	73.7%	26.3%	0		

以下是对各种构成形式的职务名称在音节和结构上的分布特点分析。

2.1 对于“类前缀+词根”的构成形式

此类以“总××”“副××”为主的构成形式，三音节词语所占比例最高，其次为双音节词语，四音节占有一定的比例。

表 2、“类前缀+词根”形式的职务名称音节分布

构成形式	双音节比例	三音节比例	四音节比例	五音节比例
------	-------	-------	-------	-------

类前缀+词根	29.2%	62.5%	8.3%	0
--------	-------	-------	------	---

结构上, 该类构成形式中, 双音节为 1+1 式, 词根为单音节语素, 如, 总理 总统 总裁 总监 总督 总编; 三音节多为 1+2 式, 如, 总经理 总书记 总指挥 总领事 副主席 副经理。其中词根部分名词的比例略高, 同时也包含词根为动词和动、名词兼类的形式。四音节词语, 在结构上有 1+1+2 式, 如: 副总经理; 也有 1+2+1 式, 如: 总参谋长 总会计师 总经济师 副司令员 总工程师 副国务卿

2.2 对于“词根+类后缀”的构成形式

该种以类后缀“师”“员”“官”“人”“家”“者”“手”为主的构成形式中, 以双音节、三音节词语为主, 其中, 三音节以 65.2% 占优势。

表 3、“类前缀+词根”形式的职务名称音节分布

构成形式	双音节比例	三音节比例	四音节比例	五音节比例
词根+类后缀	34.8%	65.2%	0	0

其中, 双音节多为“××长”“××员”“××手”, 结构为 1+1 式, 如, 院长 会长 团长 律师 医师 法师 技师 厨师 军师 画师 船长 机长; 三音节的“××师”“××员”“××官”“××人”“××家”“××者”“××手”多为 2+1 式, 如, 设计师 司令官 经理人 司令员 组织者 经营者 科学家 数学家; “××长”既包含 1+2 式, 如大队长, 支队长, 总队长; 也含有 2+1 式, 如, 理事长 参谋长 护士长 董事长 委员长 秘书长。其中, “××长”“××师”“××家”“××手”中的词根部分名词的比例略高, “××官”“××者”“××人”“××员”中的词根多为动词。

2.3 类前缀+词根+类后缀

此结构形式中的职务名称类前缀为“总”“副”, 其中在同一词中, 类前缀“副”多与类后缀“长”共现。该类词语中多为三音节、四音节形式, 三音节占 65.5%, 五音节占有一定的比例。

表 4、“类前缀+词根+类后缀”形式的职务名称音节分布

构成形式	双音节比例	三音节比例	四音节比例	五音节比例
类前缀+词根+类后缀	0	65.5%	27.6%	6.9%

其中, 三音节词语多为 1+2 式, 如, 总队长 副省长 副县长 副队长 副厅长 副所长 副台长; 四音节词语的结构主要为 1+2+1 式, 如, 总参谋长 总会计师 总经济师 副司令员 总工程师 副国务卿

2.4 限定语+词根+类后缀

此类构成形式中, 类后缀主要为“长”“人”“员”“家”; 其中多为四音节、五音节词语, 四音节以 72.2% 占优势, 六音节也占有一定的比例。

表 5、“限定语+词根+类后缀”形式的职务名称音节分布

构成形式	双音节比例	三音节比例	四音节比例	五音节比例	六音节比例
限定语+词根+类后缀	0	0	72.2%	25.9%	1.9%

其中, 四音节词语的结构形式可分为:

- 1) 2+2 式, 如, 外交部长 国防部长 能源部长 退伍军人 产业工人 复员军人

2) 3+1 式, 如, 共产党员 少先队员 共青团员 突击队员 考古学家 天文学家 社会学家 地质学家 火山学家 气象学家 生物学家 地震学家 核科学家

3) 2+2+1 式 如, 田径运动员

五音节的结构形式包括:

1) 2+2+1 式, 如, 高级指挥官 高级农艺师 高级工程师 高级研究员 首席技术官 首席潜航员 首席试飞员

2) 3+1+1 式, 如, 大学生村官

2.5 对于其他构成形式的职务名称

音节数较多, 结构上有 1) 2+1+2 式, 如, 执行总经理 执行总指挥 救灾总指挥 2) 1+2+1+1 式, 如, 副主任委员

3 职务名称的使用频次分析

职务名称在新闻联播中的出现频率有两个影响因素, 一个是职务名称自身, 另一个是新闻联播这一电视新闻语体。新闻联播属于电视新闻语体, 电视新闻语体是话语主体在电视传播语境中, 选用适切的话语材料及表达方式传播新闻信息而形成的相对类型化的话语运用体式, 是全民语言在电视新闻传播过程中的功能变体³。因此, 语料中的职务名称在词种数、构成形式分布以及所使用的领域等方面有一定特点。

3.1 职务名称词种数的分布

语料中的职务名称具有不同“类词缀”的词语, 词种数不同, 所占比例不同, 其中“××员”的词种数最多, 占 30.9%, 其次为“××长”和“××家”, 分别占 14.6%和 12.8%。“××手”和“××官”的词种数相对较少。词种数的不同与类词缀本身的“构词频率”和“职务名称”范围的限定以及“新闻联播”的语料有关。

表 6、含不同“类词缀”职务名称的词种数分布

词语	词种数	所占比例
××员	137	30.9%
××长	65	14.6%
××家	57	12.8%
××者	38	8.6%
副××	35	7.9%
××师	29	6.5%
××人	24	5.4%
总××	22	5%
××手	19	4.3%
××官	18	4.1%

3.2 职务名称的构成形式分布

根据我们总结出的职务名称的一般构成形式, (限定语)+(类前缀)+词根+(类后缀); 其中, () 表示可省略, 在语料中, 各类构成形式的职务名称所占比例不同。在构成形式几种不同的类型中, “词根+类后缀”所占比例最高, 通用性最强, 占 75%, 其次为“限定语+词根+类后缀”, 占 12.1%。

表 7、各类构成形式的职务名称分布

构成形式	词种数	所占比例
------	-----	------

类前缀+词根	24	5.4%
词根+类后缀	333	75%
类前缀+词根+类后缀	29	6.5%
限定语+词根+类后缀	54	12.1%
其他	4	0.9%

3.3 职务名称的使用频次分布

在统计的语料中，频次较高的职务名称大都属于政治领域内，表8是频次最高的十个职务名称。新闻联播在电视传播中，最主要的功能是传播新闻信息，具有极强的政治性。政治领域内的职务名称使用频次高即在情理之中。作为新闻联播的重点信息内容，国家大事的参与者，国家各级领导人出现的频率高就是一个特色。相对应地，表示国家领导人的职务名称出现频次也较高。

表8、职务名称的分布使用频次分布

词语	频次	文本数	频次*文本
总理	1517	323	489991
总统	1163	305	354715
委员	639	220	140580
副主席	591	233	137703
领导人	579	229	132591
总书记	638	207	132066
部长	527	224	118048
专家	423	210	88830
党员	512	169	86528
委员长	476	170	80920

参考文献

- [1]邵敬敏，现代汉语通论[M]，上海：上海教育出版社，2001：120。
 [2]殷志平，企业职务命名：语言与文化的双重映现[N]，中国社会科学报 2011年4月12日。
 [3]李水仙，电视新闻语体研究 [M]，北京：中国传媒大学出版社，2007：第7页。

作者简介：

张金玲（1987.01—），女（汉），山东淄博，中国传媒大学文学院应用语言学专业硕士研究生，主要研究领域：计算语言学，中文信息处理；

2010年7月，本科毕业于曲阜师范大学中文系对外汉语专业；

E-mail: zhangjinling20062006@126.com

电话：15201228561

通讯作者：何伟，中国传媒大学文学院应用语言学专业硕士导师，中国传媒大学国家语言资源监测与研究中心有声媒体语言分中心副研究员，E-mail: hewei@cuc.edu.cn