

句法成分的词类构成统计考察

郭 锐

北京大学中文系

摘 要：汉语语法的特点之一是词类与句法成分不存在一一对应的关系，本文考察实际语料中各句法成分的词类构成，结果显示，句法成分与词类虽无简单的对应关系，但优势对应十分明显，只有定语是例外。

关键词：句法成分，词类，优势对应

A statistical investigation on the makeup of syntactic components in relation to word classes

Guo Rui

Department of Chinese, Peking University

Abstract The word classes do not simply correspond to syntactic components in Chinese. It is a characteristic feature of Chinese grammar. This paper investigates the makeup of syntactic components in relation to word classes. The conclusion is that there exist some preponderant correspondences between syntactic components and word classes with the exception of attributive, though there are no one-to-one relations between syntactic components and word classes.

Keywords syntactic components, word classes, preponderant correspondences

1 引言

朱德熙（1985）认为汉语的特点之一是词类与句法成分之间不存在一一对应的关系，我们希望通过统计方法，考察实际语料中各句法成分与词类的对应关系，即统计各句法成分的词类构成以及各词类在实际语料中的语法功能，看看汉语中词类与句法成分对应的复杂关系到底表现得怎样。由于篇幅限制，本文只统计前者。

我们抽样统计了2万字语料（口语和书面语各1万字），包括：1、口语：电视剧《编辑部的故事·飞来的星星》对话记录1.1万字，6672词；2、书面语：《人民日报》1995年12月16日一版全部及二版部分1.1万字，5635词，《儿童的心理世界》第四章第一节部分0.2万字，1042词，总计1.3万字，6677词。

2 分析方法

1、本文分 20 个词类，词类名称及代号如下：谓词：动词 (v)，形容词 (a)，状态词 (z)；体词：名词 (n)，时间词 (t)，处所词 (s)，方位词 (f)，量词 (l)；饰词：区别词 (b)，副词 (d)，数词 (m)，数量词 (q)，指示词 (j)，拟声词 (o)；虚词：介词 (p)，连词 (c)，语气词 (y)，助词 (u)；其他：代词 (r)，叹词 (e)。

2、离合词无论是合是离，一律作为一个切分单位，中间隔有其他成分时，看作不连续词，如：“洗/v- 了/u 一/m 个/l 澡/-v”、“离/v- 不/d 开/-v”，其中“-v-”和“-v”分别表示一个不连续动词的前后部分，两部分合成一个“v”。

3、成语标注成语标记及语法性质标记，如“中流砥柱 i/n”、“实事求是 i/v”、“全心全意 i/d”。代词除标注代词标记外，再加注语法性质标记，如“我 r/n”表示名词性代词，“这样 r/v”表示谓词性代词，“这么 r/d”表示副词性代词，“几 r/m”表示数词性代词。

4、句法成分功能分为：主语 (S)、宾语 (O)、谓语 (P)、定语 (A)、状语 (D)、补语 (C)、环境语 (H)、独立成分 (I)、连谓结构非核心成分 (W)、递系结构非核心成分 (X) 十种，此外，方位结构的定语 (F) 单独统计。

环境语指放在动词前的不作为动词论元的时间、处所成分，如 (“|” 前为词类标记，后为功能标记)，如：昨天/t|H 他 r/n|S 没/d|D 来/v|P。

独立成分指独立语和非主语句的核心成分，包括呼唤、应答语成分，对话中单独成句的体词性成分，单独成句的无主语的谓词性成分（不区分是否为省略）。

连谓结构的前项一般作为非核心成分，后项一般是核心成分，但“来、去”为后项时，作为非核心成分，如：我们 r/n|S 吃/v|W 了/u 饭/n 看/v|P 电影/n|O；我们 r/n|S 看/v|P 电影/n|O 去/v|W。

递系结构的后项处理为非核心成分，如：小王/n|S 要求/v|P 你 r/n|O 离开/v|X。

5、短语的整体功能算作该短语核心成分 (head) 的功能。核心成分包括：偏正结构的中心语、述补结构的述语、述宾结构的述语、主谓结构的谓语、连谓结构结构的后项 (“来、去” 做后项除外)、递系结构的前项，介词结构以介词为核心、“的/地” 字结构以 “的/地” 为核心。核心成分与其直接成分之间产生的语法关系规定的句法成分不再计入核心成分的功能，如 “红苹果” 做主语时只计 “苹果” 做主语，做中心语不再计入 “苹果” 的功能，但主谓结构的谓语部分仍算作做谓语。如：中文系/n|d3 的/u|A 同学/n|S 也/d|D 知道/v|P 他 r/n|S 来/v|P|O 自/p|C 我们 r/n|A 学校/n|O。（“d3” 表示与 “的 3” 构成直接成分）

6、联合结构、同位结构整体的功能计入其中的各项成分，如：我 r/n|S 和/c 他 r/n|S 去/v|P；北大 j/n|A 校长/n|S 陈·佳洱/n|S 出席/v|P 会议/n|O。

3 句法成分的词类构成

表 1 是统计数据（见后页），图 1—图 6 是各句法成分词类构成的柱形图（见后页）。

表 1-1 句法成分的词类构成 (1)

	名 词		动 词		形 容 词		状 态 词		区 别 词		副 词	时 间 词	处 所 词	方 位 词	数 词	其 他	
	次	比	次	比	次	比	次	比	次	比							
主 语	口语	184	24%	36	4.7%	2	0.3%	0	0.0%	1	0.1%	0	0%	0	0%	0	0%
	书面	450	79%	52	9.2%	1	0.2%	0	0.0%	1	0.2%	0	0%	1	0.2%	3	0.5%
	总计	634	47%	88	6.6%	3	0.2%	0	0.0%	2	0.1%	0	0%	1	0.1%	3	0.2%
宾 语	口语	428	50%	88	10%	23	2.7%	0	0%	5	1%	0	0%	1	0%	5	1%
	书面	746	65%	221	19%	23	2.0%	1	0.1%	5	0.4%	0	0%	13	1.1%	55	4.8%
	总计	1174	58%	309	15%	46	2.3%	1	0.0%	10	0.5%	0	0%	14	0.7%	77	3.8%
谓 语	口语	10	1.1%	811	89%	83	9.1%	0	0%	0	0%	0	0%	0	0%	0	0%
	书面	0	0%	525	88%	60	10%	1	0.2%	0	0%	0	0%	0	0%	0	0%
	总计	10	0.7%	1336	89%	143	9.5%	1	0.1%	0	0%	0	0%	0	0%	0	0%
补 语	口语	0	0%	97	68%	37	26%	0	0%	0	0%	0	0%	0	0%	0	0%
	书面	0	0%	66	57%	32	28%	0	0%	0	0%	0	0%	0	0%	0	0%
	总计	0	0%	163	63%	69	27%	0	0%	0	0%	0	0%	0	0%	0	0%
定 语	口语	63	8.5%	11	1.5%	31	4.2%	0	0%	12	1.6%	0	0%	3	0.4%	5	0.7%
	书面	801	35%	246	11%	106	4.7%	0	0%	115	5.1%	0	0%	22	1.0%	6	0.3%
	总计	864	29%	257	8.6%	137	4.6%	0	0%	127	4.2%	0	0%	25	0.8%	11	0.4%
~ 方 位	口语	34	83%	1	2.4%	6	15%	0	0%	0	0%	0	0%	0	0%	0	0%
	书面	58	74%	13	17%	2	2.6%	0	0%	0	0%	0	0%	1	1.3%	0	0%
	总计	92	77%	14	12%	8	6.7%	0	0%	0	0%	0	0%	1	0.8%	0	0%
状 语	口语	6	0.5%	0	0%	23	2.0%	1	0.1%	0	0%	837	74%	0	0%	0	0%
	书面	1	0.2%	0	0%	86	14%	0	0%	0	0%	290	46%	0	0%	1	0.2%
	总计	7	0.4%	0	0%	109	6.2%	1	0.1%	0	0%	1127	64%	0	0%	1	0.1%
环 境	口语	7	7.8%	0	0%	0	0%	0	0%	0	0%	37	41%	6	6.7%	19	21%
	书面	11	14%	1	1.3%	0	0%	0	0%	0	0%	46	58%	0	0%	10	13%
	总计	18	11%	1	0.6%	0	0%	0	0%	0	0%	83	49%	6	3.6%	29	17%
独 立	口语	13	2.0%	501	76%	71	11%	0	0%	0	0%	0	0%	1	0.2%	2	0.3%
	书面	21	5.8%	316	87%	7	1.9%	0	0%	0	0%	0	0%	2	0.6%	0	0%
	总计	34	3.3%	817	80%	78	8%	0	0%	0	0%	0	0%	2	0.2%	1	0.1%
其 他	口语	100	26%	183	48%	26	7%	4	1.1%	9	2.4%	0	0%	1	0.3%	3	0.8%
	书面	166	31%	236	44%	54	10%	1	0.2%	8	1.5%	0	0%	8	1.5%	11	2.1%
	总计	266	29%	419	46%	80	9%	5	0.5%	17	1.9%	0	0%	9	1.0%	4	0.4%

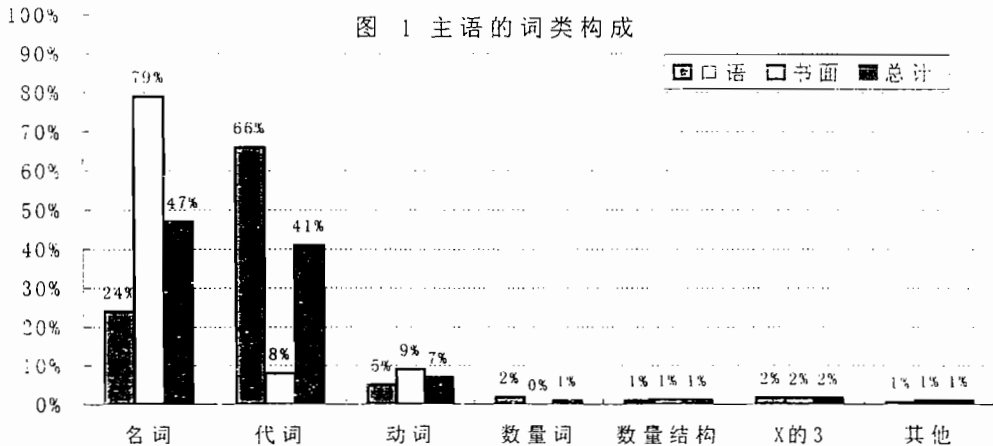
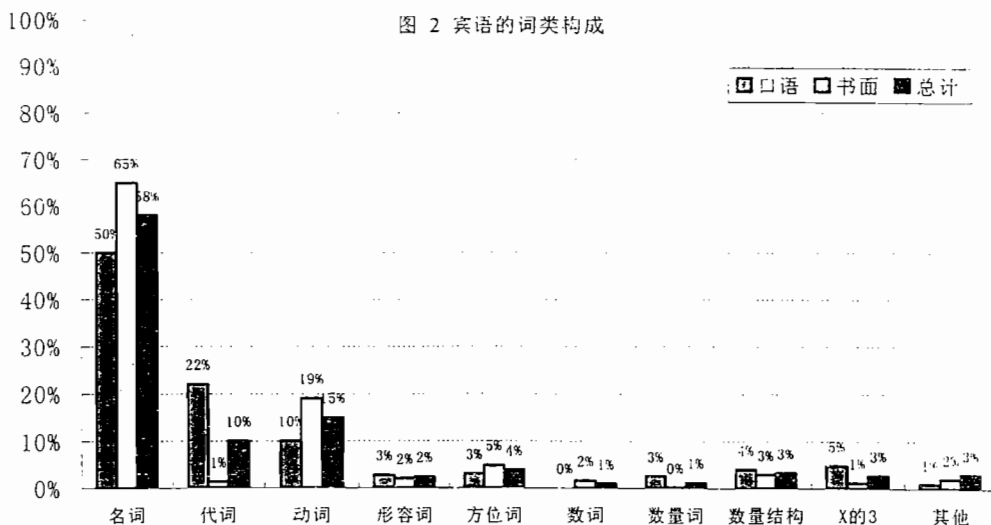
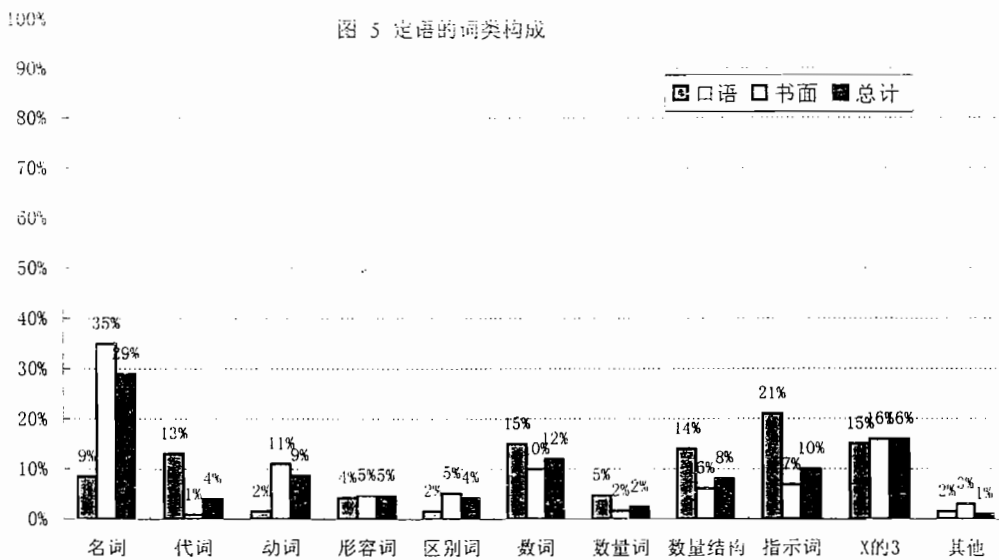
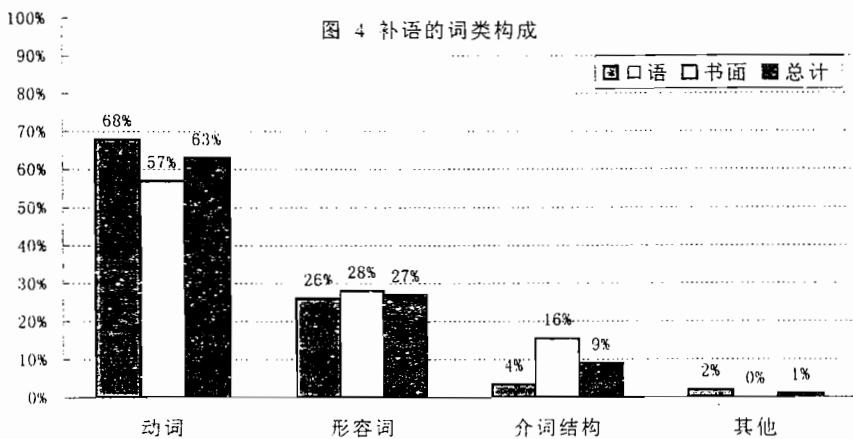
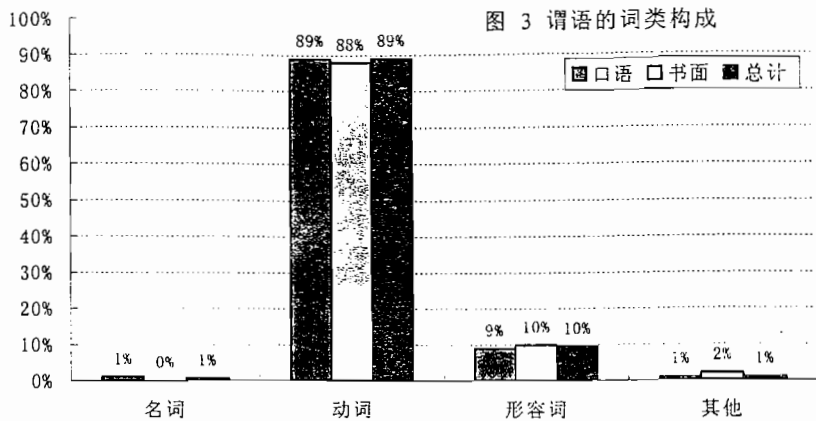
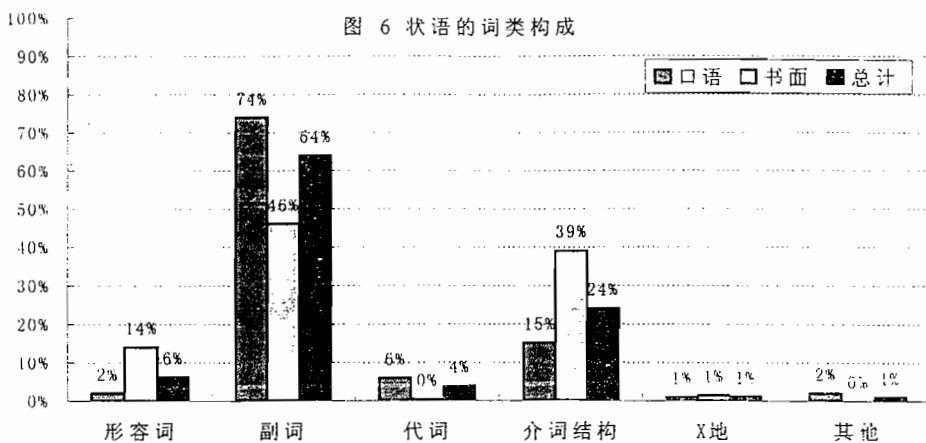


表 1-2 句法成分的词类构成 (2)

		数		量词		量		词		指		示词		拟声词		代		词		X		地		X		的2		X		的3		介		词组		总		计	
		次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次	比	次数	比例				
主	口语	14	1.8%	8	1.0%	0	0%	0	0%	0	0%	0	0%	511	66%	0	0%	0	0%	0	0%	0	0%	13	1.7%	0	0%	9	1.6%	0	0%	0	0%	769	15%				
	书面	0	0%	7	1.2%	0	0%	0	0%	0	0%	0	0%	43	7.6%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	567	9%						
	总计	14	1.0%	15	1.1%	0	0%	0	0%	0	0%	0	0%	554	41%	0	0%	0	0%	0	0%	0	0%	22	1.6%	0	0%	9	0%	0	0%	1336	11%						
宾	口语	22	2.6%	24	4.0%	0	0%	0	0%	0	0%	138	22%	0	0%	1	0.1%	41	4.8%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	860	15%						
	书面	2	0.2%	33	2.9%	0	0%	0	0%	0	0%	15	1.3%	0	0%	0	0%	0	0%	0	0%	0	0%	14	1.2%	0	0%	0	0%	0	0%	1152	18%						
	总计	24	1.2%	67	5.3%	0	0%	0	0%	0	0%	203	10%	0	0%	1	0.0%	55	2.7%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	2012	17%						
谓	口语	1	0.1%	1	0.1%	0	0%	0	0%	0	0%	5	0.5%	0	0%	0	0%	0	0%	0	0%	0	0%	2	0.2%	0	0%	0	0%	0	0%	914	16%						
	书面	0	0%	6	1.0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	594	9%						
	总计	1	0.1%	7	0.5%	0	0%	0	0%	0	0%	5	0.3%	0	0%	0	0%	0	0%	0	0%	0	0%	2	0.1%	0	0%	0	0%	0	0%	1508	12%						
补	口语	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	3	2.1%	0	0%	0	0%	5	3.5%	0	0%	0	0%	0	0%	142	2%						
	书面	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	18	16%	0	0%	0	0%	116	2%						
	总计	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	18	16%	0	0%	0	0%	258	2%				
定	口语	35	4.7%	104	14%	154	21%	1	0.1%	98	13%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	15	15%	0	0%	0	0%	0	0%	740	13%						
	书面	36	1.6%	139	6.1%	153	6.8%	0	0%	21	0.9%	0	0%	21	0.9%	0	0%	0	0%	0	0%	0	0%	##	16%	1	0.0%	0	0%	1	0.0%	2265	35%						
	总计	71	2.4%	243	8.1%	307	10%	1	0.0%	119	4.0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	##	16%	1	0.0%	0	0%	1	0.0%	3005	25%						
一	口语	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	41	1%						
	书面	0	0%	1	1.3%	0	0%	0	0%	1	1.3%	0	0%	1	1.3%	0	0%	0	0%	0	0%	0	0%	1	1.3%	0	0%	0	0%	0	0%	78	1%						
	总计	0	0%	1	0.8%	0	0%	0	0%	1	0.8%	0	0%	1	0.8%	0	0%	0	0%	0	0%	0	0%	1	0.8%	0	0%	0	0%	0	0%	119	1%						
状	口语	6	0.5%	4	0.4%	0	0%	0	0%	67	5.9%	10	0.9%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	171	15%	0	0%	0	0%	1133	20%						
	书面	0	0%	0	0%	0	0%	0	0%	2	0.2%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	245	39%	0	0%	0	0%	635	10%						
	总计	6	0.3%	4	0.2%	0	0%	0	0%	69	3.9%	10	1.1%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	416	24%	0	0%	0	0%	1768	15%						
环	口语	0	0%	7	7.8%	0	0%	0	0%	14	16%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	90	2%						
	书面	0	0%	6	7.6%	0	0%	0	0%	5	5.3%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	79	1%						
	总计	0	0%	13	7.7%	0	0%	0	0%	19	11%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	169	1%						
独	口语	1	0.2%	5	0.8%	0	0%	12	1.8%	38	5.8%	0	0%	5	0.8%	7	1.1%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	656	11%						
	书面	0	0%	2	0.6%	0	0%	0	0%	0	0%	0	0%	1	0.3%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	565	6%						
	总计	1	0.1%	7	0.7%	0	0%	12	1.2%	38	3.7%	0	0%	6	0.6%	7	0.7%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	1019	8%						
其	口语	1	0.3%	2	0.5%	0	0%	0	0%	25	6.1%	0	0%	0	0%	1	0.3%	2	0.5%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	378	7%						
	书面	0	0%	7	1.3%	0	0%	0	0%	13	2.4%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	17	3.2%	0	0%	0	0%	556	8%						
	总计	1	0.1%	9	1.0%	0	0%	0	0%	38	3.9%	0	0%	0	0%	1	0.1%	2	0.1%	0	0%	0	0%	0	0%	17	2.1%	0	0%	0	0%	934	8%						







4 结语

统计显示，口语中主语的总数（769，13%）比书面语（567，9%）还高，这表明主语省略在书面语中比口语中还多，这与一般认为口语中主语更易省略主语的说法不同。另外，有些句法成分的词类构成在口语和书面语中存在较大差异，最突出的是主语的名词和代词构成，口语中名词构成的主语只占 24%，但在书面语中占 79%；与此形成对照的是口语中代词构成的主语占 66%，而书面语中只有 8%。此外，状语的副词构成在口语和书面语中也有较大差异。

可以看到，谓语、补语和状语与词类的对应比较严格，而主语、宾语、定语与词类的对应较分散，特别是定语，比重最高的只有 29%（名词）。不过，尽管汉语词类与句法成分不一一对应，但优势对应关系仍然明显：谓语 99%由谓词（动词、形容词）构成，补语的 90%由谓词构成，状语 93%由副词性成分（副词、介词结构、副词性代词、X地结构）构成，主语 88%由体词性成分（名词、体词性代词）构成，宾语 68%由体词性成分构成。只有定语的构成无明显优势对应：42%由饰词性成分（区别词、数词、数量词、数量结构、指示词、“的”字结构）构成，33%由体词（名词、体词性代词）构成，14%由谓词（动词、形容词）构成。

参考文献

- [1]. 郭 锐.1999.语文词典的词性标注问题,《中国语文》第 2 期
- [2]. 胡明扬.1995.现代汉语词类问题考察,《中国语文》第 5 期
- [3]. 马 彪.1994.运用统计法进行词类划界的一个尝试,《中国语文》第 5 期
- [4]. 莫彭龄、单青.1985.三大类实词句法功能的统计分析,《南京师大学报》第 2 期
- [5]. 朱德熙.1985.《语法答问》.北京:商务印书馆