

基于传媒语音语料库的不同语言样式统计分析

邹煜, 侯敏, 陈玉东, 付莉

(中国传媒大学应用语言学系, 北京 100024)

摘要: 在广播电视有声语言传播中, 根据不同的节目类型以及不同的传播形式, 其有声语言表达有着不同的样式, 例如新闻节目中的播报、文学类节目中的朗读、各类专题节目中的讲解(或解说)和主持人节目中谈话等等。本文正是在传媒语音语料库的基础之上, 对播报、朗读和谈话三种不同语言样式的音节时长特性、音高音域的变化特性进行了初步的统计分析。通过分析认为, 在朗读、播报和谈话三种语言样式中, 音节的平均时长依次递减, 证明了三种语言样式在广播电视有声语言传播创作中的不同特性; 其次, 在音高及音高音域的变化中, 朗读、播报比谈话的音高音域的变化幅度要大得多。此外, 对于不同语言样式的统计分析, 其结果不仅有利于广播电视有声语言的教学, 还有利于语音识别和语音合成技术的发展。

关键词: 传媒语音语料库; 语言样式; 播报; 朗读; 谈话

Analysis of Various Speaking Styles Based on Broadcast Speech Corpus

ZOU Yu, HOU Min, CHEN Yudong, FU Li

(Dept. of Applied Linguistics, Communication University of China, Beijing, 100024)

Abstract: There are different speaking styles among different TV or radio programs. This paper is a preliminary statistical analysis for news announcing, oral interpretation and talking styles based on broadcast speech corpus, which the segmentation and annotation of the broadcast speech have been done. The purpose is to give some results from two aspects. Firstly, we conclude that the mean duration of tones of oral interpretation style is longer than that of news announcing style; that of talking style is the shortest among three styles. Secondly, we have calculated the F0 of each tone, the F0 and pitch range of news announcing style is high and has big fluctuation; that of talking style is high and has small fluctuation. However, the F0 of tone 3 of news announcing style is lower than that of oral interpretation and talking styles.

Keywords: Broadcast speech corpus; Speaking styles; News announcing; Oral interpretation; Talking

1 引言

在广播电视有声语言传播中, 根据不同的节目类型以及不同的传播形式, 其有声语言表达有着不同的样式, 例如新闻节目中的播报、文学类节目中的朗读、各类专题节目中的讲解(或解说)和主持人节目中谈话等等^[1]。这些不同的语言样式在我们听感上的差异是非常明显的, 那么它们各自在语音上的特性究竟体现在哪儿呢? 本文

基金资助: 本研究得到国家语委十五规划项目(项目号: YB105-61A)和中国传媒大学 211 工程项目(项目号: BBU211-15)的支持, 谨此致谢。

作者简介: 邹煜(1971-), 男, 四川德阳, 中国传媒大学计算语言学博士生 E-mail: zouiy@cuc.edu.cn

在传媒语音语料库的基础上，拟对其中的播报、朗读和谈话三种不同语言样式进行初步的统计分析，力求找到一些证据。此外，这些证据对于广播电视有声语言的教学、语音识别与合成技术也有很大的促进作用，例如，如何根据用户的需求合成用户喜好的不同语言风格(speaking styles)的语音？

2 语料库

为了更深入地研究广播电视语言，自 2002 年以来，我们开始建立传媒语言语料库。2005 年 9 日，教育部、国家广播电影电视总局和中国传媒大学三家共建的国家语言资源监测与研究有声媒体语言分中心挂牌成立。该中心的主要任务就是对有声语言资源进行收集、整理、加工，建设现代汉语传媒有声语言普通话样本库；利用该样本库提取传媒有声语言应用的各种参数和统计数据，为语言研究、语言规范研究、语言标准以及语言政策的制定提供依据；对传媒有声语言资源的语言文字应用情况进行监测和客观的描述；对不同语言现象的监测情况以不同的名义和方式进行发布；向国家有关部门提供咨询，为制定标准和规范提供依据和参考（详情可浏览 <http://ling.cuc.edu.cn/>）。

传媒语言语料库是一个面向播音主持艺术领域、以广播电视新闻语料为主要内容的综合性语料库系统，它包括两个密切相关的子语料库，即文本语料库和语音语料库。其中，现已建成的语音语料库的熟语料约 11 个小时左右。在这个语料库中，语料的选取在综合考虑到广播电视现有节目的基础上注重语料的代表性、典型性和全面性，所选节目都是广播电视中的主流节目和有代表性的节目；语料库中的所有声音文件都进行了音段切分和韵律标注，而且声音与文本并存，原始数据、转写数据、标注数据同时存在，便于多层面、多方位的综合开发和利用^{[2][11]}。

2.1 语料分布

如前所述，传媒语音语料库的语料要有代表性，并力求反映广播电视有声语言的实际使用情况，其目的主要是为研究传媒有声语言服务，尤其是传媒有声语言的韵律研究。因此，在语料选取上不仅要考虑不同节目类型以及不同语言样式（语体），而且还应考虑不同的年代，以便于对传媒语音发展作历时研究。表 1 所示为传媒语音语料库的语料整体分布情况：

表 1 不同类型语料的时间分配情况（单位：小时）

不同年代 \ 不同样式(语体)	播报式	朗读式	讲解式	谈话式
播音老前辈	0.5	0.5		
继往开来的时代	2	1	0.5	
新时代	2	1	1	1.5

注：1. 播音老前辈指 20 世纪五六十年代的录音文本；继往开来的时代指 20 世纪七八十年代的录音文本；新时代则是指 20 世纪九十年代以后的录音文本。

2. 上表中的不同语言样式（语体）的分类对应于传媒语言语料库的文本语料库，该分类与“中国播音学”中语言表达样式的分类是一致的^{[1][3]}。

2.2 语料标注

传媒语音语料库的语料标注由三人完成，分为浅层标注（2~3 层）和深层标注（5 层以上），具体标注内容如下：

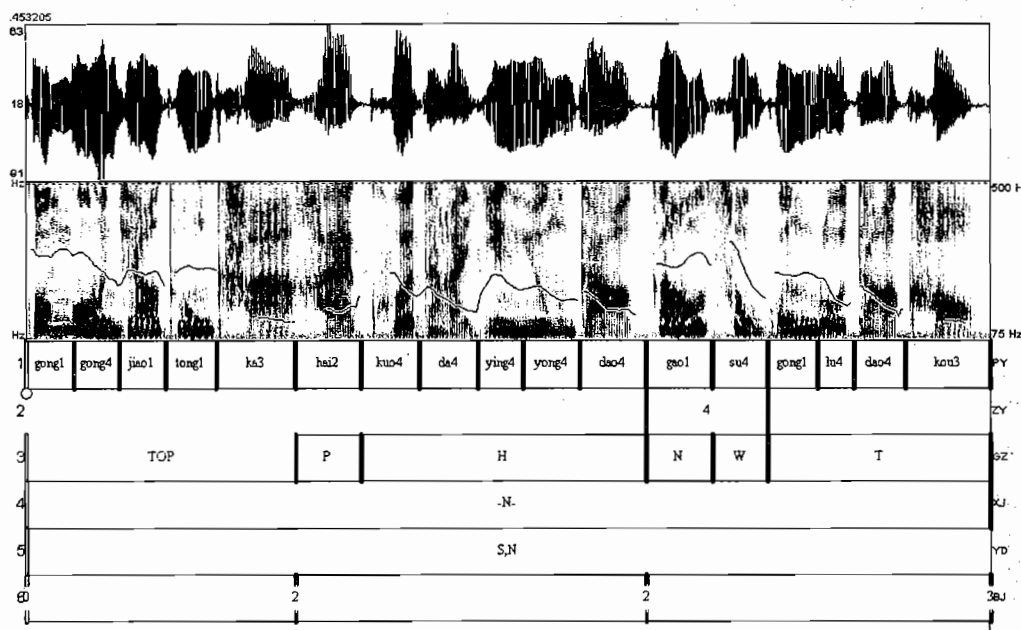


图 1 深层标注语料示例（共有六层，从上到下分别是拼音、重音、语调构造、小句焦点、语段构造、韵律边界）

3 数据分析

本文的研究从传媒语音语料库中挑选出了 4 个多小时的语料（包括 14 个男声、12 个女声）进行统计分析。所选语料的基本情况如表 2 所示：

表 2 所选标注语料的具体分布情况

语言形式		代表性广播电视节目	时间分布
独白	播报	整点新闻、新闻联播、全国新闻联播等	1:10:41
	讲解	音乐故事、世界周刊、中国周刊等	0:43:45
	朗读	主要是文学作品	1:13:35
	谈话	音乐故事、世界周刊、中国周刊、第一时间等	0:51:53
	综合	新闻纵横等	0:17:14
总计			4:17:08

由于所选语料中的讲解类主要是专题类节目的解说，也有新闻类或主持人节目中的专题解说，例如《世界周刊》中的“故事”、《中国周刊》中的“本周人物”、《音乐故事》中除主持人开场、衔接和结尾以外的部分等，这些几乎都有现场同期声或背景音乐，要过滤掉这些有一定难度。另外，即使过滤掉这些现场同期声或背景音乐，提取出来的语音数据（尤其是音高数据）也会有很大的失真，对这类语音数据进行比较分析意义不大。因此，本文在进行具体统计分析时没有采用这一类语言样式中的数据，只分析了播报、朗读和谈话三类广播电视语言样式的相关数据。

3.1 不同语言表达样式音节时长分析

本文首先分析了三类不同语言样式中的音节时长，具体统计数据如下：

表3 声调在不同语言样式中的时长均值(单位:秒)

	1	2	3	4	0
播报	0.189	0.199	0.192	0.180	0.130
朗读	0.338	0.337	0.324	0.335	0.277
谈话	0.167	0.173	0.163	0.163	0.124

表4 语料中出现频率最高的10个音节在不同语言样式中的时长均值(单位:秒)

	de0	shi4	shi2	zai4	le0	tal	bu4	guo2	yi4	zhong1
播报	0.112	0.190	0.166	0.174	0.133	0.185	0.139	0.187	0.142	0.179
朗读	0.149	0.290	0.319	0.246	0.169	0.293	0.174	0.223	0.228	0.297
谈话	0.105	0.170	0.166	0.159	0.112	0.169	0.141	0.134	0.150	0.200

从表3和表4的结果来看,朗读类语言样式相对来说比较夸张,语速的快慢变化比较明显,发音饱满,吐字讲究雕琢,因此,表现在音节时长上,朗读就比播报和谈话体要长得得多。而播报与谈话比较,播报又相对工整,音节时长要比谈话长,谈话体的音节时长最短。而且,不同语言样式中音节的时长差异显著(Sig.<.05)。

3.2 不同语言表达样式音节基频分析

音高是语音感知中除音长外的又一敏感因素,声调的变化也主要体现在音高的变化上。对不同语言样式中音高的统计分析,可以让我们更好的认识不同语言样式的特性。本文用时间归一化(time normalization)方式计算每个音节的音高,即不论每个音节的长短,统一取10个点的音高值;然后将赫兹值(Hz)转化为半音值(semitone),参考频率为100Hz;最后再从所取音高值中,阴平取最大值、阳平取最小值和最大值、上声取最小值、去声取最大值和最小值。相关统计数据见表5和表6:

表5 四声在播报、朗读和谈话三种语言样式中的基频值描述统计量(单位:st)

		样本数	均值	标准差	标准误	均值95%置信区间	
						下限	上限
阴平	F0_max 播报	419	9.0614	4.21747	.20604	8.6564	9.4664
	朗读	455	8.5889	7.07356	.33161	7.9372	9.2406
	谈话	438	13.0966	4.73253	.22613	12.6522	13.5410
阳平	F0_min 播报	402	2.2973	4.25707	.21232	1.8799	2.7147
	朗读	678	1.1017	5.13432	.19718	.7145	1.4888
	谈话	329	7.3282	4.84014	.26685	6.8032	7.8531
	F0_max 播报	402	6.9560	4.48367	.22363	6.5164	7.3956
	朗读	678	8.2555	7.26314	.27894	7.7078	8.8032
	谈话	329	11.0841	4.98558	.27486	10.5434	11.6248
上声	F0_min 播报	381	-1.7810	5.06812	.25965	-2.2915	-1.2705
	朗读	473	.0251	4.83923	.22251	-.4121	.4623
	谈话	278	5.3948	5.08287	.30485	4.7947	5.9950
去声	F0_max 播报	670	9.0599	4.50105	.17389	8.7185	9.4013
	朗读	910	9.2981	7.08333	.23481	8.8373	9.7589
	谈话	692	12.9382	5.09606	.19372	12.5579	13.3186
	F0_min 播报	670	.9595	5.43811	.21009	.5470	1.3720
	朗读	910	1.3753	5.65977	.18762	1.0071	1.7436
	谈话	692	8.3773	5.20357	.19781	7.9890	8.7657
轻声	F0_min 播报	57	-1.9177	4.51884	.59854	-3.1167	-.7187

	朗读	351	1.0604	5.62220	.30009	.4702	1.6506
	谈话	212	6.6233	5.29968	.36398	5.9058	7.3408

表6 播报、朗读和谈话三种语言样式不同声调的音高单因素方差分析

			Sum of Squares	df	均方	F	Sig.
阴平	最大值	组间	5396.525	2	2698.263	88.437	.000
		组内	39938.429	1309	30.511		
		总计	45334.955	1311			
阳平	最小值	组间	8789.912	2	4394.956	188.406	.000
		组内	32797.770	1406	23.327		
		总计	41587.682	1408			
	最大值	组间	3192.878	2	1596.439	43.225	.000
		组内	51928.101	1406	36.933		
		总计	55120.979	1408			
上声	最小值	组间	8686.720	2	4343.360	175.315	.000
		组内	27970.455	1129	24.775		
		总计	36657.175	1131			
去声	最大值	组间	6757.275	2	3378.638	99.423	.000
		组内	77106.477	2269	33.983		
		总计	83863.753	2271			
	最小值	组间	24863.865	2	12431.933	417.201	.000
		组内	67612.608	2269	29.798		
		总计	92476.474	2271			
轻声	最小值	组间	5422.133	2	2711.066	92.248	.000
		组内	18132.976	617	29.389		
		总计	23555.108	619			

通过对表 5 的数据分析可以证明，文学作品的朗读和新闻播报中音高音域的变化要比谈话明显得多，也丰富得多。也正因为这样，前者在听感上就更富有节奏感，也更清晰，正所谓“字字如珠玑”，而后者就更接近我们生活中的语言，在凸显主要信息的基础上，其余的冗余信息有时候就一带而过了。由表 6 可以得知，阴平、阳平、上声、去声和轻声的音高在播报、朗读和谈话中有着显著的差异（Sig=.000<.05）。

4 结论

通过对广播电视语言传播中的播报、朗读、谈话等三种不同语言样式的统计分析，我们初步得出如下一些结论：

首先，在朗读、播报和谈话三种语言样式中，音节的平均时长依次递减，证明了三种语言样式在广播电视有声语言传播创作中的不同特性。例如，播音创作中，在处理重音时，要求重音词语吐送时音调明显上升，即把音节相对拉长（快慢法）；或者加快速度，缩短音长（强弱法）等^[9]，而这种处理要求也主要是针对新闻播报和文学作品的朗读。

其次，在音高及音高音域的变化中，朗读、播报比谈话的音高音域的变化幅度要大得多。同样，播音创作中在处理重音时，要求把重音音节声调的调域拉宽（即重音表达中的高低法）^[9]，而在新闻播音中，播报是其“主流形态”，而“字正腔圆，呼吸无声，感而不入，语尾不坠，语势平稳，节奏明快，新鲜感强，基调各异，分寸恰当，语流畅达”则是其一般范式^[10]。

再有,对于不同语言样式的统计分析,其结果不仅有利于广播电视有声语言的教学,还有利于语音识别和语音合成技术的发展,例如,如何根据用户的需求合成用户喜欢的不同语言风格(speaking styles)的语音?有人^[12]对日语不同话语风格语音的频谱倾斜(spectral tilt)特性、共振峰频率特性进行了语音声学分析,为能合成不同语言风格的语音奠定了一定的基础。这也将是本文进一步的研究工作。

参考文献

- [1] 张颂. 话语样式简论[A], 语言传播文论(续) [C], 北京:北京广播学院出版社, 2002: 215~228.
- [2] 邹煜,胡凤国. 传媒语音语料库的语料选取与标注[A], 声学技术: 第八届全国人机语音通讯学术会议(NCMMSC-2005)[C], 北京, 2005-10: 490~494.
- [3] 姚喜双. 播音学概论[M]. 北京:北京广播学院出版社, 1998.
- [4] 林茂灿. 普通话孤立句的韵律结构和 F0 下倾[A], 新世纪的现代语音学(第五届全国现代语音学学术会议论文集) [C]. 北京:清华大学出版社,2001: 180~184.
- [5] 林茂灿. 普通话语句的韵律结构和基频(F0)高低线构建[J],当代语言学,2002,(4): 254~265.
- [6] 陈玉东. 传媒有声语言语段的构造和调节, 博士论文. 北京大学, 2004.
- [7] 沈炯. 汉语语调构造和语调类型[J],方言,1994,(4): 221~228
- [8] 沈炯. 关于韵律和语调的一些看法[A],第三届全国语音学研讨会论文集[C],1996: 115~119.
- [9] 李晓华. 广播电视语言传播发声艺术概要[M]. 北京:北京广播学院出版社, 1999.
- [10] 张颂. 试论新闻播音的创新空间[A], 语言传播文论(续) [C], 北京:北京广播学院出版社, 2002: 152~158.
- [11] Zou Yu, Hou Min, Chen Yudong, Hu Fengguo and Fu Li. *Broadcast Audio and Video Bimodal Corpus Exploitation and Application*. In Proceeding of fifth SIGHAN Workshop on Chinese Language Processing, Sydney, Australia, July 22-23, 2006.
- [12] Masanobu Abe. Speaking Styles: Statistical Analysis and Synthesis by a Text-to-Speech System. In Jan P.H. van Santen et al (eds). *Progress in Speech Synthesis*. Springer-Verlag New York, Inc. 1997:495-512.