

自然话语音高研究

王茂林

中国社会科学院语言研究所
wangmaolin26@hotmail.com

摘要

本研究使用 973 电话语料,对 79 个说话人的基频低点和音域进行了考察,对 1109 个语调短语的音高范围和音高位置作了统计分析。我们发现虽然男女讲话人的基频低点相差很大,但是音域却很接近。语调短语可以根据音高范围和音高位置分出不同的类,这与所传达的信息量有关,传信传疑度越高,音高范围就越大,位置也越高。

1. 引言

近年来,韵律研究日益受到语言学界和言语工程界的重视。深入全面的理解自然语言的韵律特征无论对语音学研究,还是对提高语音合成的自然度和识别语音的准确性来说,都是至关重要的。提高自然度、使合成的声音能够表现出话语的抑扬顿挫变化,是当前合成中需要解决的重点问题。如果能使合成系统构造出人进行自然会话时的音高模式,势必会提高合成的自然度。这一工作的第一步,就是要对人会话的机制及自然话语的音高模式有一个清楚的认识。本文采用电话语料为研究材料,讨论自然话语的音高问题。

2. 语料

我们使用的语料是 973 电话语料库中的材料,该语料已由我们实验室标出了音节、声韵、功能等。并用 praat 软件提取了基频,该软件提取的正确率在 80% 以上,对于不正确的我们参照窄带语图进行了人工修改。语料中有不少讲话人讲话带有方言色彩,这样的语料我们没有使用。仅使用属于标准普通话的材料。

由于我们研究的是自然话语,所以其中有很多重复、迟疑、说半句话的现象(陈建民 1984)。另外说话人还经常使用填充词(filler),填充词指的是对话中出现的没有实际意义的词语(Watanabe & Ishi 2000),如“嗯”、“这个”、“那个”等,用于表示停顿、思考,具有占有说话权的作用。研究说话人在什么情况下会出现重复、迟疑等现象,也是非常有趣的,但是本文是要探讨一下在正常情况下,讲话比较连贯时话语的音高表现,所以在选择语料时,我们没有选择重复和迟疑较多的语料。经过挑选,我们使用了 69 段对话,其中涉及讲话人 79 人,男 37 人,42 人。他们都是年龄在

20-40 之间的成人。

3. 结果与讨论

3.1 讲话人基频低点

首先我们对男女讲话人的基频低点分别进行了考察,结果如下:

讲话人	男	女
平均低点基频 (Hz)	85.9	146.3
标准差	18.82	33.4
讲话人数	37	42
最大值	133.2	196.52
最小值	55.11	80.29

表 1 男女讲话人基频低点

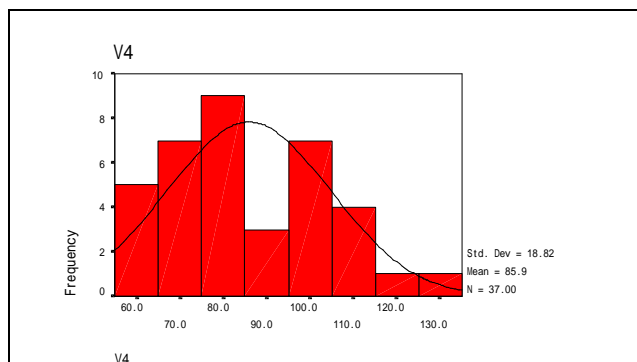


图 1 男讲话人基频低点分布图

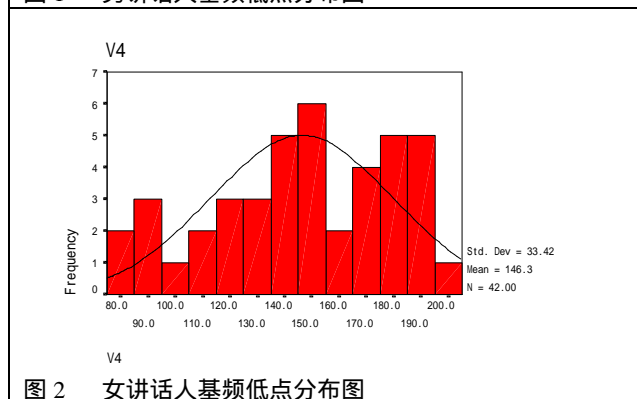


图 2 女讲话人基频低点分布图

从图 1 和图 2 可以看出,男说话人的基频低点主要分布在 75-100 赫兹之间,女说话人的基频低点主要分布在

130-190 赫兹之间。

3.2 讲话人音域

我们又对不同讲话人在一次谈话中的音域进行了考察，这里用半音来表示。具体计算方法是：找出讲话人在一次谈话中 F0 的最大值和最小值，然后用下面的公式计算：

$$\text{range}_{\text{spe}} = 12 * (\lg (F0_{\text{max}} / F0_{\text{min}}) / \lg 2)$$

$\text{range}_{\text{spe}}$ 表示讲话人在一次谈话中的音域， $F0_{\text{max}}$ 和 $F0_{\text{min}}$ 分别表示这个讲话人 F0 的最大值和最小值。结果如下：

讲话人	全部	男	女
平均音域 (半音)	16.3	16.4	16.2
标准差	4.41	4.75	5.07
讲话人数	79	37	42
音域最大值 (半音)	27.89	27.08	27.89
音域最小值 (半音)	8.67	8.74	8.67

表 2 讲话人的音域

973 语料是客人预订宾馆和车票的电话录音，讲话人目的在于办事，所以讲话时态度比较平和，整个音域适中。如果是在大礼堂讲话，或是在激烈的争论问题，音域肯定要比这里大。同时我们可以看出，男女的差异也不大，音域主要分布在 13-19 个半音之间。

3.3 语调短语音高范围和音高位置

近几年关于韵律结构研究的成果很多 (林茂灿 2000、曹剑芬 2001、李爱军 2001)，我们参考这些成果，并结合自然话语的特点，对语料的韵律进行了标注。我们分出韵律词、中间短语、语调短语和韵律组 4 级单位。韵律词包括音步、附加组和复杂韵律词。如果一个语句中含有多个韵律词，它们之间的间断不会是完全一致的，我们把间断大的地方定为中间断语的边界。人在讲话时先在大脑中有一个想法，再利用自己的语言知识，将话语脱口而出。一个思想单元就是一个语调短语 (Chafe 1980)。标注的结果，得到语调短语 1109 个。

在上面的讨论中我们发现男女的音域差异不大，所以在计算语调短语音高范围和位置时我们对男女就不分别进行了，具体计算方法是：

$$\text{range}_{\text{ip}} = \text{st}_{\text{max}} - \text{st}_{\text{min}}$$

range_{ip} 表示语调短语音高范围， st_{max} 和 st_{min} 分别表示这个语调短语中高出讲话人音域低点的最大值和最小值。

平均音高范围 (半音)	10.2
标准差	3.56
语调短语个数	1109

最大值 (半音)	26
最小值 (半音)	1.38

表 3 语调短语音高范围

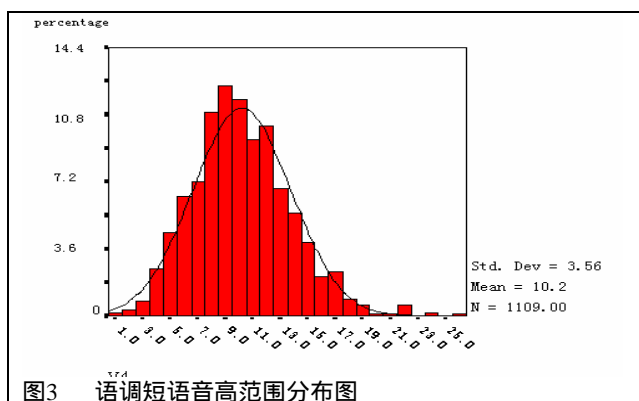


图 3 语调短语音高范围分布图

因为不同说话人的音域有很大不同，所以在考察语调短语的音高范围和音高位置时，我们对不同说话人进行了规整处理。具体办法是：

$$\text{range}_{\text{nom}} = \text{range}_{\text{ip}} / \text{range}_{\text{spe}}$$

$$\text{reg}_{\text{inom}} = \text{st}_{\text{max}} / \text{range}_{\text{spe}}$$

$\text{range}_{\text{nom}}$ 和 reg_{inom} 分别表示规整的语调短语音高范围和音高位置，结果如下：

类别	音高范围	音高位置
平均值 (规整)	0.59	0.8
标准差	0.19	0.13
语调短语个数	1109	1109
最大值 (规整)	1	1
最小值 (规整)	0.09	0.34

表 4 规整的语调短语音高范围和音高位置

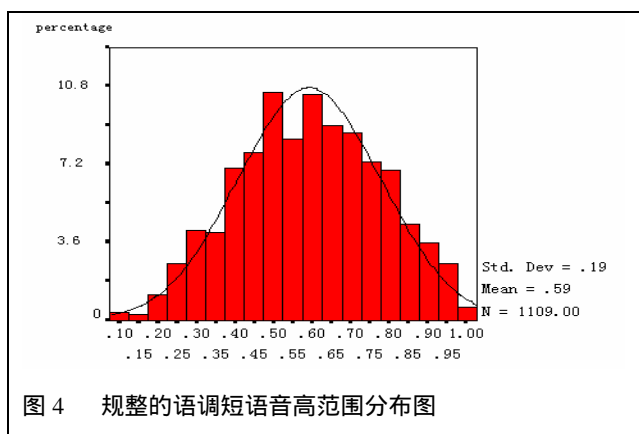


图 4 规整的语调短语音高范围分布图

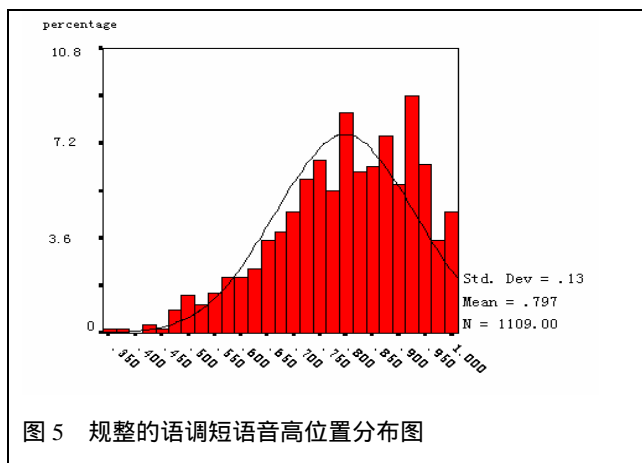


图 5 规整的语调短语音高位置分布图

0.4-0.8 之间,我们据此将其分为三类,0.4 以下的标为“窄”(N),0.4-0.8 之间的标为“中”(M),0.8 以上的标为“宽”(W)。经过统计,范围窄的有 179 个语调短语,范围宽的 177 个,范围居中的最多,有 753 个。这里从图 4 中看到数据在 0.4 和 0.8 之间比较集中,并参考沈炯(1994)的研究,我们就分成了三类。对于范围窄的,我们根据规整的音高位置又分为三类,位置在 0.62 以下的标为“低窄”(LN),位置在 0.62-0.85 之间的标为“中窄”(MN),位置在 0.85 以上的标为“高窄”(HN)。对于范围居中的,因为它本身范围已经比较宽(在 0.6 左右),没有必要分为三类了,我们就根据规整的音高位置分两类,位置在 0.8 以下的标为“低中”(LM),位置在 0.8 以上的标为“高中”(HM)。对于范围宽的,我们就一律标为“宽”(W),见下表:

从图 4 可以看出,语调短语的音高范围主要集中在

类别	窄(N)			中(M)		宽(W)
数量	179			753		177
类别	低窄(LN)	中窄(MN)	高窄(HN)	低中(LM)	高中(HM)	
数量	58	96	15	334	419	

表 5 语调短语音高范围和位置分类

从图 3 可以看出,语调短语的音高范围主要分布在 7-14 之间。语调短语音高范围有的很小,只有不到两个半音,有的却很大,可以达到二十几个半音,我们认为造成这种差异的一个重要原因就是话段的功能。言语有传信和传疑两种功能(吕叔湘 1947),不仅如此,话段传信和传疑又有程度的差别。如果一句话所传达的信息量比较大,我们就说这句话的传信度高,否则就低;如果一句话所表达的疑问语气比较强,我们就说这句话的传疑度高,否则就低。着也反应了语音和语义上的象似(iconicity)关系,看下面的例子:

.....
b: 我是曲洪
老到你们那儿去
据说是九折
a: 是吗?(NL)
b: 好吗?

.....
这里说话人 a 说了一句“是吗”,并没有疑问,也不需要对方回答,只是在谈话中起话语交际作用,所以它的音高范围很小,位置也很低,再如:

.....
a: 您要是住双间
不在一层可以吗?
b: 不在一层可以

a: 可以哈?(NM)
那行(NM)

.....
这里的“可以哈”和“那行”并不传达多少实际信息,也是主要起交际作用,所以范围很小,但是比上面的“是吗”又实在一些,故不是很低,再如:

.....
a: 一间房是一百三十块钱一个晚上
b: 双人间?(NH)
a: 呵
双人间

.....
这里说话人是要确认一下,疑问语气比较重,所以位置高。“双”和“间”都是阴平调,中间的“人”受协同发音的影响,也变阴平,所以范围很窄,这种情况不多(见表 5),我们发现音高范围窄的语调短语的音节数目都不多,也就是说,短句倾向于音高范围小。再看下面的例子:

.....
a: 双人间有二百、二百四、三百、四百
这四种
b: 差别在哪里?
a: 主要是房间面积(MH)
b: 呵
主要房间面积啊(ML)
行

到时候再说吧

.....

音高范围适中是出现率最高的情况(见表5),讲话人 a 说“主要是房间面积”是回答 b 的问话,传达的信息量大,所以位置高。讲话人 b 说“主要房间面积啊”,表示自己明白了,传达的信息量小,所以位置低。再看下面的例子:

.....

b:应该给我八折

我们去过多少趟了

a:我们院里来的都是九折 (W)

.....

这里说话人 a 在说“我们院里来的都是九折”时有强调的意味,所以音高范围很宽。

4. 结论

我们发现人在正常谈话时音域主要分布在 13-19 个半音之间,而且男女差异不大。语调短语的音高范围主要分布在说话人音域的 0.4-0.8 之间,位置主要分布在 0.62 以上。如果一句话所传达的信息量比较大,并且其中有被强调的成分,则它的音高范围就会比较大,如果一句话所传达的信息

量比较小,则它的音高范围就会比较小,位置也会比较低。

参考文献

- 曹剑芬(2001)汉语韵律切分的语音学和语言学线索 《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
- 陈建民(1984),《汉语口语》,北京出版社。
- 李爱军(2001)普通话对话中韵律特征的声学表现,《第六届全国人机语音通讯学术会议论文集》
- 林茂灿(2000),《普通话与句中中断和语句韵律结构》,《当代语言学》2000年,第2期。
- 吕叔湘(1947),《中国语法要略》,商务印书馆(1982年版)。
- 沈炯(1994)汉语语调构造和语调类型,《方言》1994年第四期
- Chafe, Wallace L.(1980), The Development of Consciousness, in *The Pear Stories Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Ablex Publishing Corporation, Norwood, New Jersey.
- Watanabe, Michiko and Carlos Toshinori Ishi (2000) *The Distribution of Fillers in Lectures in the Japanese Language*. Proceedings in ICSLP 2000, Beijing.

A STUDY OF PITCH OF SPONTANEOUS SPEECH

Wang Maolin

Institute of Linguistics, CASS
wangmaolin26@hotmail.com

ABSTRACT

In this paper the bottom F0 values and pitch ranges of 79 speakers and the pitch range and register of 1109 intonation phrases in spontaneous speech corpus is analyzed. We found that there is no difference between the pitch ranges of male and female speakers

although the difference between the bottom F0 values of male and female speakers is great. We can make classification of intonation phrases by their pitch ranges and registers, which is related to the information the clause carried.