

## 汉语语调降阶的实验研究\*

黄贤军<sup>1,2</sup> 杨玉芳<sup>1</sup> 吕士楠<sup>1</sup>

(1 中国科学院心理研究所脑与认知科学国家重点实验室 北京 100101)

(2 中国科学院研究生院 北京 100039)

(3 中国科学院声学研究所 北京 100080)

2006 年 3 月 15 日收到

2006 年 7 月 17 日定稿

**摘要** 采用设计特定声调组合的实验室语句的方法,研究了汉语普通话语调降阶的规律。结果发现,在普通话语调中,同样存在自动降阶和非自动降阶两种音高现象。在有低音介入的非自动降阶中,低音对后音节音域的影响表现在使音域上限降低,音域的下限保持不变;因此汉语中的降阶是语调高音特征的表现;低音特征声调会使低音后的高音发生正向降低和低音前高音的逆向提高,并且这两种作用是可以相互包容的。在自动降阶中,高音线的下降趋势是线性的,它受重音位置及重读程度的影响,在发音人中有较大差异。与已有的其他语言降阶研究的结果相比,汉语的降阶率不是固定的,且降阶的作用范围并不只局限于同低音相邻的音节。

PACS 数: 43.70

### Experimental studies on downstep in Chinese intonation

HUANG Xianjun<sup>1,2</sup> YANG Yufang<sup>1</sup> LÜ Shinan<sup>3</sup>

(1 State Key Laboratory of Brain and Cognitive Science, The Institute of Psychology,  
Chinese Academy of Sciences Beijing 100101)

(2 Graduate School, Chinese Academy of Sciences Beijing 100039)

(3 Institute of Acoustics, Chinese Academy of Sciences Beijing 100080)

Received Mar. 15, 2006

Revised Jul. 17, 2006

**Abstract** Downstep in pitch contour of Mandarin Chinese is examined using subtly designed sentences, by controlling tone combinations. The results show both automatic and non-automatic downstep exist in Putonghua. In non-automatic downstep, the low tone will compress downwards the pitch range of following syllables, and the main influence of downstep is on topline. Low tone will not only lower the topline behind it, but also raise the high tone before it, both effect can be compatible with each other in Chinese. In automatic downstep, the topline of intonational phrase pitch contour is presented as a linear downtrend. But there are different results among speakers because of the effect of stress. In addition to these, by comparing with the downstep in other tone or non-tone languages, in Chinese, the ratio of downstep is not constant, and the domain of downstep is not limited within the adjacent tones.

### 引言

不论声调语言或非声调语言,陈述语句的音高曲线中都存在明显的下降趋势(downtrend)。Shih<sup>[1]</sup>指出影响音高曲线的因素有很多,包括下倾、降阶、末尾下降、重音和声调、音段、语调类型、话语结构等。由于上述因素同时对基频曲线的下降趋势产生影响,因此它们对下降趋势会存在叠加效应。其中下倾

和降阶最受关注,争议也最多。降阶现象最初在非洲声调语言中发现的,相应的研究结果也比较多<sup>[2,3]</sup>;由于非声调语言中高音和低音不与语法上的词素对应,尽管证据比非洲声调语言少,但仍有一些研究者指出英语和日语中也存在降阶现象<sup>[4,5]</sup>。

降阶包含自动降阶(automatic downstep)和非自动降阶(non-automatic downstep)两类<sup>[6]</sup>。自动降阶则指在没有低音介入的情况下,高音的自动降低现

\* 国家自然科学基金资助项目(60175013, 30370481)

象。非自动降阶也称为 downdrift, 指在低音介入的情况下, 构成通常语调标注系统中的  $H^1LH^2$  序列, 后面的高音 ( $H^2$ ) 明显低于前面的高音 ( $H^1$ ) 的现象。通常所说的降阶指的是非自动降阶。降阶的作用除了表现为  $H^1LH^2$  序列中  $H^2$  的降低外, 还会使  $L$  前的  $H^1$  提高<sup>[3]</sup>(H-raising)。

由于降阶现象只在局部出现, 后面的高音只受其前高音的限制, 因此不少研究者使用固定的降阶率 (ratio of downstep) 来描述音高重音的递降现象<sup>[4,7]</sup>。他们提出, 在一个语调单元中, 高音成分是有规律可循的, 后面高音目标的峰值可以通过其前的峰值乘以降阶率预测得到。Prieto 等<sup>[8]</sup>还发现降阶率不受短语长度和两个高音之间时间长短的影响。

对汉语降阶现象的研究, 最值得关注的是 Yi Xu<sup>[9,10]</sup> 和 Shih<sup>[1,11]</sup> 的研究。Xu 发现汉语声调之间存在正向 (carry-over effect) 和逆向 (anticipatory effect) 两种效应, 即前一个声调的终值高低会影响后一声调的起始音高, 后一声调的起始音高的高低会影响前一声调的终值; 并且正向效应大于逆向效应。Shih<sup>[11]</sup>对汉语降阶的研究发现, 当高音前为上声 (L) 时, 降阶的作用最大; 当高音前为阳平声 (R) 和去声 (F) 时, 降阶的作用比较弱; 当高音前为阴平声 (H) 时, 不发生降阶。

综上所述, 以往的研究尤其是在非声调语言的降阶研究中, 大都关注语调曲线中的高音部分, 同时考察高音和低音两部分的研究比较少。以前的汉语研究对降阶规律描述的不够明确。本研究在实验方法上, 通过控制相关条件, 设计了不同声调组合的主动宾 (SVO) 句式的实验室语句; 研究降阶这种语调现象时, 对汉语普通话基频曲线中的高音和低音两部分同时作了系统考察, 总结汉语普通话语调降阶的规律。

## 1 方法

### 1.1 实验材料

实验语句为 6 音节的 SVO 语句, 每个语法成分由双音节词构成。句子声调组合采用以具有高音特征的阴平声音节为主体的如下设计: 1AB111、11AB11、1A1B11, 其中 1 表示阴平声音节。A 和 B 表示声调不同的替换音节, A 设计成阴平声或上声, B 设计成阴平声、上声或去声。由于前面的声调对后面声调的影响大于后者对前者的影响, 并且上声产生的降阶作用大于阳平声和去声, 因此着重考察 A 位置上的高调和低调对 B 位置音域中高音成

分和低音成分的影响。每种组合条件下各编制 10 个实验句, 共 120 句 (例句见附录)。

### 1.2 发音人与朗读方式

两位男发音人朗读了所有的句子。一个为播音系的本科生 (M1), 另一个为研究生 (M2), 都说标准普通话。要求发音人不带特殊强调重音、用中等语速、以接近新闻广播风格的方式, 把句子自然、流畅地朗读出来。

### 1.3 声学分析

对采集到的句子使用 MultiSpeech 语音分析软件进行声学分析, 将去声起点定义为高音点, 去声末点和上声最低点定义为低音点。并对提取的音高进行半音转换。其计算公式为:

$$St = 12 \log_2(F_0/F_{0ref}), \quad (1)$$

其中,  $F_0$  为基频测量值;  $F_{0ref}$  为基频参考值, M1 取 70 Hz, M2 取 90 Hz。

## 2 结果与讨论

首先, 对全部用阴平音节组成的句子的音高曲线进行考察, 以阴平声音高的平均值为其特征值, 我们把阴平声的特征值曲线称为基线。由于基线中不包含任何低音成分, 因此按照 Bruce<sup>[6]</sup>的观点, 基线的音高下降是自动降阶的结果。

### 2.1 基线自动降阶的规律

在基线中, 发音人 M1 句首和句尾音高差值为 1.54 个半音, 发音人 M2 句首和句尾音高差值为 4.70 个半音。统计结果表明基线始末音高具有显著性差异 (发音人 M1:  $t_{(9)} = 6.01^{***}$ ; 发音人 M2:  $t_{(9)} = 35.41^{***}$ ), 因此两位发音人的基线具有明显的自动降阶趋势, 但不完全相同。两位发音人朗读 6 个阴平声组合的句子后, 10 个句子的平均基线如图 1 所示。

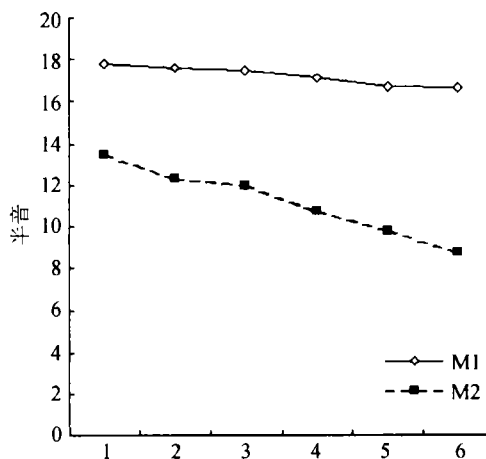


图 1 两位发音人 10 个句子的基线平均值

分别对基线的下降趋势进行线性拟合和二次曲线拟合, 结果如表 1 所示。

表 1 基线线性拟合与曲线拟合结果

发音人	线性拟合			二次曲线拟合			
	B0	B1	P	B0	B1	B2	P
M1	17.69	-0.31	0.001	17.67	-0.29	0	0.000
M2	14.44	-0.93	0.000	14.13	-0.70	-0.03	0.001

由表 1 可以看出, 线性拟合与二次曲线拟合的结果差别不大, 主要表现在: 使用线性拟合后的结果显著性水平都达到 0.001, 二次曲线拟合系数 B2 的值很小。因此, 可以认为基线的下降趋势是线性的。这与 Shih<sup>[1]</sup> 和 Xu<sup>[9,10]</sup> 的研究结果不同。Shih 发现: 即使在无焦点的句子中, 当声调全部为阴平声时, 音高下倾的速度是先快后慢的。Shih 使用句首和句尾固定的实验材料, 通过逐步增加和变换句子中间部分的内容来改变句子长度; 由于不断变换句子中间部分的内容, 在句子的中间部分引人的新信息, 容易造成发音人对变换的部分重读, 使得基频曲线的下降速度先快后慢。Xu 给出的 5 音节无焦点的阴平声组合句的基频曲线始末差值仅为 1 Hz, 即没有明显的下降趋势, 有可能是宽焦点句中句末的宽焦重音造

成的。从以上实验结果的不同, 可以看到重音对高音线的影响, 而且重音的重读程度有较大的自由度, 使得自动降阶现象在发音人之间有较大的差异。

## 2.2 低音的非自动降阶作用对其后音节音域的影响

当语句中介入低音成分后, 低音成分会对音高曲线产生非自动降阶作用, 因此在考察基线自动降阶的基础上, 接下来进一步考察非自动降阶的作用对语调曲线的影响。在 1AB111、11AB11、1A1B11 三类组合中, 以 B 位置去声起点为音域上限, 以上声最低点为音域下限; 当 A 位置分别为阴平声和上声时, 考察 A 位置的低音(上声)对其后 B 位置音域上下限的影响。三类组合的统计结果如图 2—图 7 所示。

由图 2—图 7 可以得出: 非自动降阶现象是普遍存在的, 但显著程度在发音人间有差异, 在我们的实验中 M1 要比 M2 显著。降阶程度也和低音音节的距离有关系, 11AB11 要比 1A1B11 显著。所以, 除发音人 M2 在 1A1B11 组合中(图 7), A 位置的低音对 B 位置音域没有显著影响外; 其他组合条件下, A 位置的低音使 B 位置的音域发生明显的变化。由图可见音域的上限降低了, 而下限变化很小。进一步

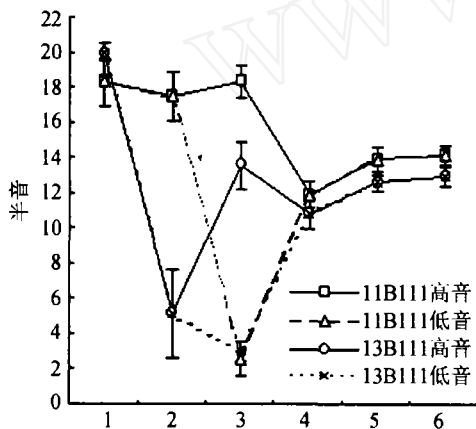


图 2 发音人 M1 在 1AB111 组合中的结果

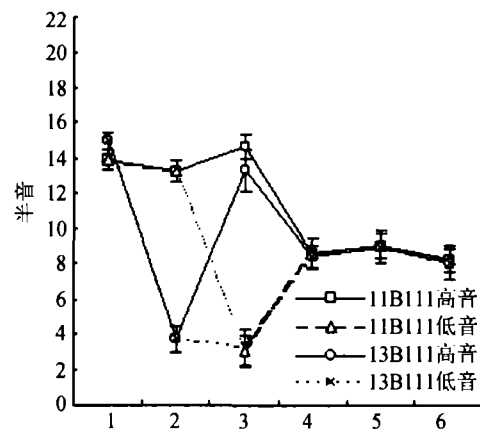


图 3 发音人 M2 在 1AB111 组合中的结果

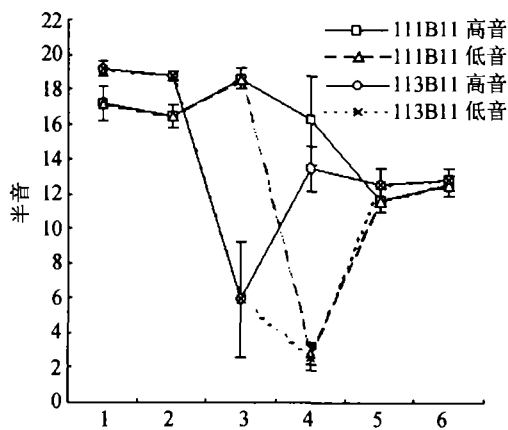


图 4 发音人 M1 在 11AB11 组合中的结果

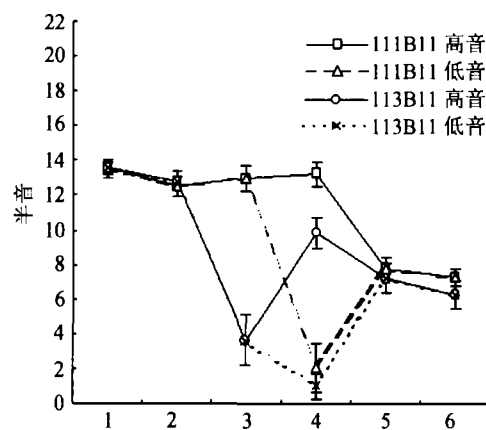


图 5 发音人 M2 在 11AB11 组合中的结果

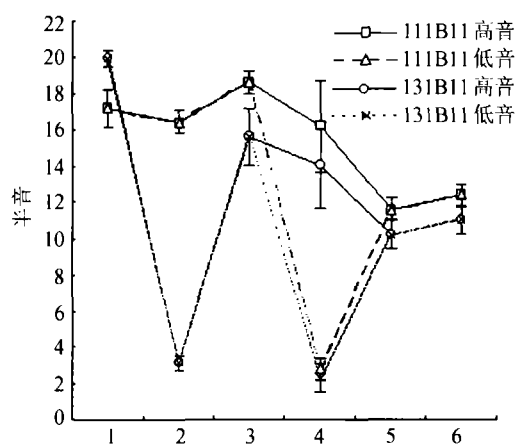


图 6 发音人 M1 在 1A1B11 组合中的结果

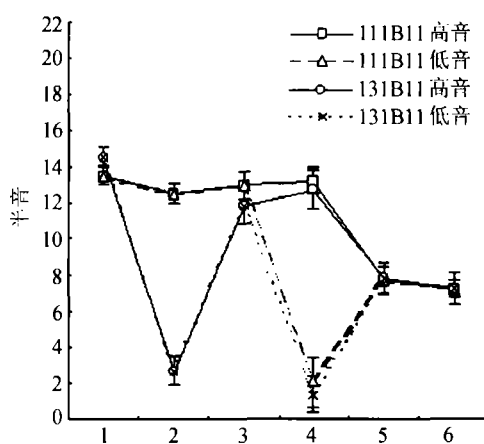


图 7 发音人 M2 在 1A1B11 组合中的结果

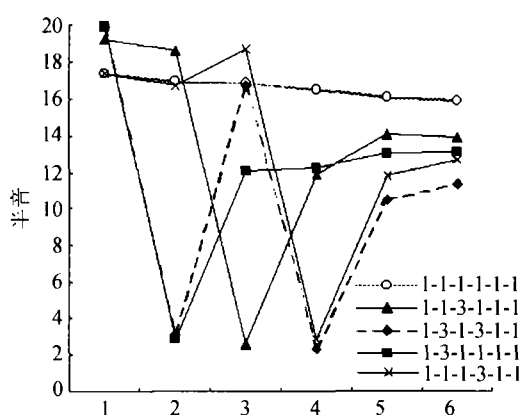


图 8 发音人 M1 五种组合结果

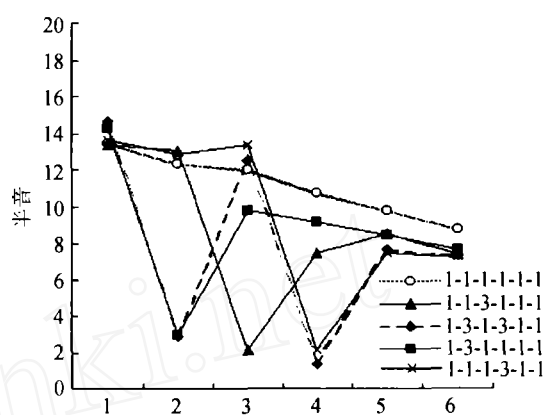


图 9 发音人 M2 五种组合结果

地, 当 A 位置分别为阴平和上声时, B 位置的去声末点的音高值变化也很小。因此, 汉语中的降阶是语调中高音特征的表现。两位发音人音域上限降低的程度不同, 发音人 M1 对三种组合的音域降低范围都比较大, 可以一直延续到句末; 发音人 M2 除了在 11AB11 组合中音域降低范围较大外, 其他组合中只对低音后相邻的高音有显著的降低作用。

### 2.3 低音的正向降低与逆向提高作用及其作用范围

由上可知, 汉语的降阶表现为语调曲线中高音成分的变化。并且在自然语句中, 全部为阴平声组合的句子很少, 大部分都是由包含高音成分和低音成分的声调共同组成的。因此以典型的高音(H-阴平)和典型的低音(L-上声)组合得到的句子为研究对象, 系统地考察非自动降阶的规律。我们称阴平的音高均值和上声最低点的音高值的连线为特征值曲线。当低音出现在第 2、3、4 音节时, 计算每种条件下 10 个句子的特征值曲线, 两位发音人的平均结果分别如图 8 和图 9 所示。

由图 8 和图 9 可见, 在有低音(L)介入的四种组合 131111、113111、111311、131311 中, 音高曲线的高音成分与基线相比下降程度更大, 发音人

M1 平均下降 6.33 个半音, 发音人 M2 平均下降 6.58 个半音; 而且高音的降低是从低音后的邻接高音开始, 作用域一直扩展到句末。说明低音(L)对语调曲线中的高音(H)存在显著的正向降低作用。在汉语中表现为从低音后的第一个高音开始, 建立了新的、降低了的音高水平。这与 Mambila 语言的研究结果不同<sup>[6]</sup>, 在 Mambila 语言中, 低音只对其后的第一个高音存在正向降低作用, 使得此高音明显降低, 但此后的高音则又重新升至语句固有的音高水平上。这说明汉语和 Mambila 语在低音正向降低的作用域上明显不同。Xu 研究了汉语中不同声调及不同的焦点位置对音高曲线的影响, 在他的女性发音人的结果中, 可以看到降阶的正向降低作用范围较大, 但他认为在较大范围上的降阶作用比与低音相邻的降阶作用更复杂; Shih 研究发现了汉语中具有低音特征的声调不同, 降阶力度也不同; 但对汉语中降阶的作用域上都没有得出明确的结论。

由图 8 和图 9 同时可以看到低音对其前的高音具有显著的逆向提高作用, 分别计算 131111、113111、111311 组合中低音前音节的音高和基线对应位置上音高的差值如表 2 所示。

表 2 三种声调组合中低音前的音节音高和基线的差值 (单位: 半音)

发音人	组合 131111		组合 113111				组合 111311					
	第一个音节		第一个音节	第二个音节		第一个音节	第二个音节		第三个音节			
	平均值	标准差	平均值	标准差	平均值	标准差	平均值	标准差	平均值	标准差		
M1	2.47	0.57	1.77	0.74	1.60	0.61	0.09	1.10	0.31	0.98	1.88	0.72
M2	0.82	0.49	0.82	0.50	1.40	0.46	0.17	0.51	0.63	0.51	1.39	0.94

由表 2 所列的平均值均为正值, 说明从统计结果看, 低音前的高音不同程度地被提高了, 证明了汉语的降阶也同有逆向提高作用。在组合 131111 中, M1 提高了 2.47 个半音, M2 提高 0.82 个半音, 发音人的差异比较大。在组合 113111 中, 上声前的两个阴平声都提高了。发音人 M1 分别提高了 1.77 和 1.60 个半音; 发音人 M2 分别提高了 0.82 和 1.40 个半音。在组合 111311 中, 第三个音节位置上的高音 (即低音前的邻接高音) 被提高了, 上声前其他的两个音节提高不明显。说明低音对其前邻接高音的逆向提高作用最大; 逆向提高作用的作用范围会大于一个音节, 但这种提高作用会随着前面的高音与低音距离的变大而变弱, 而且韵律词边界会减弱逆向提高的作用, 这和正向降阶作用域的表现不同。

有趣的是在 131311 组合中, 第三个音节位置的高音, 偏离基线很小。这是由于第二个音节位置上的低音的正向降低作用, 使它的音高降低; 又由于第四个音节位置上低音的逆向提高作用, 使它音高升高。因此, 综合的结果, 它的音高显著高于只存在低音正向降低作用的组合 131111 中相同位置的高音 (发音人 M1:  $t_{(18)} = 9.13^{***}$ ; 发音人 M2:  $t_{(18)} = 7.21^{***}$ )。这一现象说明汉语语调曲线中的某一位置的高音可以同时受正向降低和逆向提高两种作用的制约, 正向和逆向作用是可以互相包容的, 而不是互相排斥的。

#### 2.4 不同位置低音对高音的降阶率

在非声调语言中, 降阶率几乎是固定的<sup>[4,7,8]</sup>。对上述阴平声和上声组合的句子, 参照 Liberman 与 Pierrehumbert<sup>[4]</sup> 对降阶率的定义, 将降阶率定义为上声前后, 阴平声的半音比值的倒数。降阶率的值越小, 表示降阶程度越大。两位发音人不同声调组合语句中的降阶率如表 3 所示。

从两位发音人的结果可以看出, 在句中只有一个低音时, 降阶率的数值比较接近。但在一个句子有两个低音的组合 131311 中, 前后的降阶率有明显的差别, M1 分别为 0.84 和 0.64, M2 分别为 0.85 和 0.61。这与 Liberman 等<sup>[4]</sup>、Prieto<sup>[8]</sup> 等研究结果

不同, 他们分别在英语和西班牙语中发现在一个韵律单元中, 高音和低音交替出现时, 降阶率几乎是固定的。这说明具有相对稳定的降阶率只是非声调语言的特点, 汉语在降阶特征上与印欧语系不同。

表 3 两位发音人降阶率统计结果

组合类型	发音人 M1		发音人 M2	
	平均值	标准差	平均值	标准差
131111	0.60	0.04	0.68	0.06
113111	0.64	0.04	0.59	0.05
111311	0.63	0.04	0.55	0.03
131311 <sup>1</sup>	0.84	0.06	0.85	0.04
131311 <sup>2</sup>	0.64	0.06	0.61	0.06

#### 2.5 讨论

本研究在综合考察降阶对汉语语调曲线中的高音和低音成分的影响后, 发现降阶作用主要使低音后的音域上限明显降低, 对音域下限成分无显著性影响。仲晓波等<sup>[12]</sup>、王蓓<sup>[13]</sup> 等通过统计大规模语料库中不同等级韵律边界前后音节的高音、低音差值, 发现边界前后低音点的重置是系统变化的; 高音点由于受重音的影响而比较复杂, 不规则。从本文的研究结果可以看出, 在语句中, 高音点除了受重音的影响外, 还受低音产生的降阶的作用。

Xu<sup>[14]</sup> 认为普通话语句的基频曲线受主观因素和客观因素共同决定。其中主观因素受发音人的主动控制, 由语调、情感表达等高层次的语言功能决定; 包括新话题的引入以及说话人表达焦点重音, 主观因素通过音高目标和调域来表现。客观因素不受发音人的控制, 无意产生的, 如说话人的总音域、音高升降的最快速度和音高转向的最快速度等。Ohala 与 Ewan<sup>[15]</sup> 等发现音高下降的速度要快于音高上升的速度, 产生一个升调要比产生一个降调困难。因此使得低音后的高音不能达到原有的音高目标值。所以, 降阶作用尤其是非自动降阶作用, 是由于发声器官的生理局限而产生, 进一步影响到基频曲线音高走势的。但这种局部发音事件只影响基频曲线的高音成分。

<sup>1</sup> 131311 组合中第一个低音的降阶率

<sup>2</sup> 131311 组合中第二个低音的降阶率

### 3 总结

本研究通过特别设计实验室语句, 综合考察了汉语普通话语调的降阶现象。观察发现, 汉语中同样有自动降阶和非自动降阶两种音高表现, 它们都使音高曲线产生降低。自动降阶作用的表现会受重音的影响, 在发音人中有较大的差异。低音产生的非自动降阶, 一方面表现为正向降低作用, 使得低音后的高音音高明显降低, 进而影响到低音后侧语调曲线中的其他高音成分; 另一方面还表现为逆向提高作用, 使得低音前的高音音高明显提高。综合考察低音介入对其后汉语语调曲线中的高音和低音成分的影响后, 发现降阶作用主要使低音后的音域上限明显降低, 对音域下限成分无显著性影响, 证明降阶是语音中高音特征的表现。除了考察降阶的正向降低和逆向提高作用这一语言的共性外, 还进一步考察了正向降低和逆向提高的作用范围。发现低音的正向降低作用是从低音后的高音开始, 高音线降至一个新的音高水平, 可以一直扩展到句末。这与 Mambila 语的研究中, 正向降低作用只局限于低音后的第一个高音的结果不同。逆向提高的作用范围也超过低音前的相邻音节, 但随着低音前的高音与低音之间距离的变大而变弱。另外, 通过对非自动降阶中降阶率的测量, 发现汉语语调不存在固定的降阶率, 这与非声调语言也不同。

本实验所揭示的汉语语调曲线降阶的规律, 为预测合成语句的基频曲线、提高合成语句的自然度等提供了实验依据。

### 参 考 文 献

- 1 Chilin Shih. Declination in Mandarin. Intonation: theory, models and applications. Proceedings of an ESCA workshop. Athens, Greece, 1997: 293—296
- 2 Stewart J M. Downstep and Floating Low Tones in Adioukrou. *Journal of African Languages and Linguistics*, 1983(5): 57—78
- 3 Laniran O, Clements G. Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics*, 2003; **31**: 203—250
- 4 Liberman M, Pierrehumbert J. Intonational invariance under changes in pitch and length. In Mark Aronoff and Richard T. Oehrle (eds.), *Language Sound Structure*, Cambridge, MA: MIT Press. 1984: 157—234
- 5 Beckman M, Pierrehumbert J. Intonational structure in Japanese and English. *Phonology Yearbook*, 1986(3): 255—310
- 6 Connell B. Downdrift, downstep and declination. In *Typology of African Prosodic Systems*, U. Gut and D. Gibbon, Eds. 2002: 3—12
- 7 Bruce G. Developing the Swedish intonation model. Lund: Department of Linguistics and Phonetics, Lund University. 1982; **22**: 51—116
- 8 Prieto P, Shih C, Nibert H. Pitch downtrend in Spanish. *Journal of Phonetics*, 1996; **24**: 445—473
- 9 Yi X. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of phonetics*, 1999; **27**: 55—105
- 10 Yi X. Contextual tonal variations in Mandarin. *Journal of Phonetics*, 1997; **25**: 61—83
- 11 Shih C. Tone and intonation in Mandarin, working Papers, Cornell Phonetics laboratory, 1988: 83—109
- 12 仲晓波, 郑波, 杨玉芳. 关于普通话韵律短语重音的实验研究. *声学学报*. 2002; **27**(2): 141—148
- 13 王蓓, 杨玉芳, 吕士楠. 汉语韵律层级结构边界的声学分析. *声学学报*. 2004; **29**(1): 29—36
- 14 Yi Xu. From Simple Pitch Targets to Complex F0 Contours. 现代语音学论文集. 1999: 93—98
- 15 Ohala J, Ewan W G. Speed of pitch change. *Journal of the Acoustical Society of America*, 1973; **53**: 345—346

### 附录

#### 实验设计句子示例

- 111111: 公司装修车间。  
 113111: 专家主编书刊。  
 113311: 张辉采访嘉宾。  
 134111: 工厂印刷书刊。  
 113411: 官兵保卫边疆。  
 111311: 公司生产沙发。  
 111411: 师生歌颂家乡。  
 131111: 商场批发香烟。  
 131311: 张伟清点资金。  
 131411: 交警发布通知。  
 114111: 专家定居乡村。  
 133111: 工厂补发津贴。