

汉语自然话语的音高下倾

王茂林, 訾广玲

暨南大学华文学院, 广东, 广州, 510610

摘要: 本研究使用电话对话语料, 在统计的基础上对语句的音高下倾进行了考察。发现绝大多数语句的音高都是逐渐下降的, 音高曲线前高后低的走势有其生理上的原因, 并且具有标界功能。少数语句音高不下降, 这与词语的载义重度、焦点及音节本调有关。本文又对陈述句和疑问句的音高进行了考察, 发现与陈述句相比, 疑问句的整体音域较大, 句末无疑问语气词的是非问句末尾两音节间的降幅较小。

关键词: 自然话语; 音高; 下倾

中图分类号: TP391

文献标识码: A

Pitch Declination of Chinese Spontaneous Speech

WANG Maolin, ZI Guangling

College of Chinese Language and Culture, Jinan University, Guangdong, Guangzhou, 510610

Abstract: In this paper, based on a telephone conversation corpus, the pitch declination of sentence is done. It is found that pitch declination occurs in most cases, which is due to physiological reason, and there is demarcative function as well. In some cases declination does not occur, and this is related to semantic strength, focus and tone. The pitch representation of statement and question is analyzed, and it is found that compared to statement, the pitch range is great for question. The pitch drop is the least between the final two syllables for yes-no question without final particle.

Key words: spontaneous speech; pitch; declination

1 引言

音高下倾 (pitch declination) 是指语句的音高曲线逐渐下降的现象。林焱、王理嘉 (1992: 187) 指出, 电台播音员播出“北京人民广播电台”的台名时, 是一般的陈述语气, 语调是下降的。开头上声“北”的调值是 21, 结尾阳平“台”的调值应该是 35。可实际测量的音高结果显示, 末尾“台”字调值 35 的 5 比开头“北”字调值 21 的 1 还低得多。音高从起始到末尾是逐渐下降的, 这就是音高的下倾现象。

下倾是语句音高的普遍表现 (Cruttenden, 1997), 王安红等 (2004) 基于言语数据库, 研究了汉语的音高下倾现象, 发现低音线的下倾可以出现在音步、语调短语和语句等不同的韵律单元中, 最明显的是在语调短语中。黄贤军等 (2009) 使用设计语料, 考察了语句的音高下倾趋势, 发现低音线清晰地呈现出以韵律短语为单元的下倾现象, 下倾的斜率与韵律短语的长度成反比。声调组合不同, 会导致低音线下倾的斜率不同。韵律短语音高下倾程度受其在句中所在位置的影响, 句首韵律短语的下倾程度大于句末韵律短语的下倾程度。

目前在汉语语音合成方面, 合成语音的质量已经明显提高, 然而在自然度方面仍需进一步改进。语音的自然度主要表现为词语的轻重、节奏及音高的变化等方面, 也就是话语的韵律。近年来, 韵律研究日益受到语言学界和言语工程学界的重视, 深入了解话语的韵律特征, 对于提高语音合成的自然度具有重要的意义, 而音高下倾就是话语韵律的重要特征。本文拟使用多人较大规模的语料, 对汉语自然话语的音高下倾现象进行量化的分析。

2 语料说明

本研究使用的语料是中科院自动化所录制的 973 电话语料库中的材料（宗成庆等，1999），这是在真实场景下关于客人预订宾馆房间及火车票的电话录音，属于自然口语。该语料由中国社科院语言所语音室进行了音节、声韵及句子语气等的标注，又用 Praat (www.praat.org) 提取了基频。该软件提取基频的正确率在 80% 以上，对于提取不正确的部分参照窄带语图进行了人工校正。语料库中有不少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。挑选语料时，也没有选择包含重复和迟疑的材料。经过选择，我们使用了 69 段对话，其中涉及说话人 79 人，男 37 人，女 42 人。

3 结果与讨论

3.1 韵律结构

近年来关于汉语韵律结构研究的成果较多。研究发现，一句话中词与词之间在韵律上的疏密关系各不相同，因此可以划分出不同层级的间断 (break)，进而构成语句的韵律结构（曹剑芬，2002；林茂灿，2002；李爱军，2002）。依据这些研究，并结合自然话语的特点，我们对语料进行了韵律标注。在句子以下标出韵律词、韵律短语两级韵律单位。各韵律单位的界定是：组成韵律词的几个音节听起来是紧密连在一起的；韵律短语后面通常有音节延长或小的停顿。语料标注的结果，得到语句 1109 个。

3.2 下倾斜率

本文对语句音高下倾斜率的考察，是基于韵律结构及音高高低音线进行的。下面先看一句话的音高曲线（见图 1）：

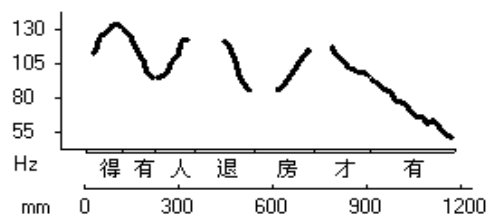


图 1 语句音高曲线示例

图 1 所示音高曲线的总趋势基本上是逐渐下降的，但音高的局部走势有升有降，这是由音节的本调引起的。汉语是声调语言，汉语语句的音高曲线同时负载了声调和语调等多方面的信息。声调影响音高曲线的局部走势，语调对声调的音阶起整体调节作用，二者的音阶相互叠加（曹剑芬，2002）。这里“音阶”是指音高的位置，例如图 1 中“得”在“有”前声调变为阳平，变调后的“得”与“人”和“房”声调相同，但是“得”的音高位置比“人”高，“人”又比“房”高，也就是“得”的音阶高，“人”的音阶次之，“房”的音阶更低。正因为汉语有声调，所以在讨论汉语语调问题时，研究者常使用音高的高音线和低音线来反映句子音高的整体走势（沈炯，1994；林茂灿，2002）。本文讨论汉语自然话语的音高下倾，主要就是要考察高低音线的走势问题，语句的高低音线是将其所包含韵律词的高音点和低音点分别连接起来而成的线（林茂灿，2002）。依据听感进行判断，图 1 所示的语句包含三个韵律词：“得有人”、“退房”和“才有”。音高的高音线就是把“得有人”的高音点和“才有”的高音点连接而成的线，低音线就是把它们的低音点连接而成的线，见图 2：

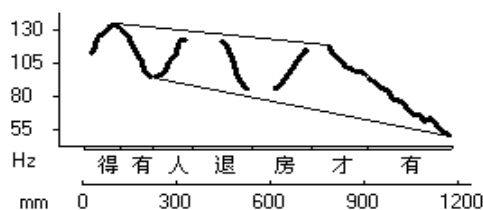


图 2 语句音高的高低音线示例

如果一句话包含两个或两个以上的韵律词，就对其高音线和低音线的斜率进行了计算，计算的公式如下：

$$Sl_t = \frac{T_f - T_e}{D_t} \quad (1) \quad Sl_b = \frac{B_f - B_e}{D_b} \quad (2)$$

Sl_t 和 Sl_b 分别表示语句高音线和低音线的斜率， T_f 和 T_e 分别表示起首和末尾韵律词的高音点， B_f 和 B_e 分别表示起首和末尾韵律词的低音点， D_t 表示从起首韵律词的高音点到末尾韵律词高音点的距离， D_b 表示从起首韵律词低音点到末尾韵律词低音点的距离。因为语料中包含男女不同性别的说话人，说话人之间的音域不同，实际基频值就会相差很大。例如从图 1 和图 3 所示的三句话的基频值看，有的高音点接近 400 赫兹，有的高音点只有 100 多赫兹，所以需要进行归整处理。吴宗济先生（2002）指出，人们在听觉上辨别语音的高低升降，是和听辨旋律的音阶相同的，所以在计算斜率时我们没有使用实际基频值，而是将基频值转换为半音值计算的，半音的计算公式如下：

$$St = 12 \times \log_2 \left(\frac{F_0}{F_{0ref}} \right) \quad (3)$$

F_0 是某点的基频值， F_{0ref} 是参考频率值，这里用 27.5Hz， St 即为这一点的半音值。上述三个公式都是在已有研究的基础上，稍作改动得到的。另外，普通话有 4 个声调，其中阴平是高平调，上声是低降调，前者没有“低”特征，后者没有“高”特征。如果一句话的起首韵律词或末尾韵律词中的几个音节都是阴平调，或者起首韵律词或末尾韵律词只有一个音节，而且这个音节是上声调，对这样的语句就没有统计。下倾斜率的统计结果见表 1：

表 1 语句的下倾斜率

	平均斜率（半音/秒）	样本数 ¹
高音线	2.49	731
低音线	3.39	731

表 1 显示，从整体平均值看，语句的高音线和低音线的斜率都是正值，表示它们都呈下降走势，音高的这种前高后低表现跟语音产生的生理机制有关。发音的原动力是声门下的肺气压，声门下压力的大小能影响基频的高低。随着一句话发音的进行，声门下压力会越来越小，音高也会逐渐下降（Lieberman, et al, 1969；曹剑芬，2002）。另外从功能的角度看，前高后低的音高模式具有“标界”（demarcative）功能（Fletcher, et. al, 2002），可以表示一句话结束了，有表示语句“完结”（finality）的功能。相反，如果是话说到一半的时候出现迟疑，音高就不会有明显的下降，而且说话人还会使用“嗯”、“这个”等填充词（filler），表示自己还没有说完。所以，语音表现是服务于表达需要的。从表 1 还可以发现，低音线的平均斜率比高音线大，说明末尾韵律词的平均音域比较大。音域是高音点与低音点音高之差，在图 2 所示的语句中，“才有”的音域比“得有人”的大。末尾韵律词的音域增大，表明语句的音高有降尾（final lowering）存在，降尾同样具有标界作用。观察图 2 还可发现，不但语句的整体音高是下降的，韵律词内部音节的音高位置也是下降的，图 2 中的“得有人”是一个韵律词，“得”和“人”声调相同，但是前者的位置比后者高，这是由声带紧张度的变化决定的。声带紧张度的改变可以引起音高的变化（Poletto, et. al, 2004），人在说一个韵律词的时候，通常首音节发音时声带较紧，后面音节发音时声带会相对松弛一些，所以韵律词内部音节的音高也是下降的。

表 1 给出的是语料库中 731 个语句高低音线的整体平均斜率，事实上，高低音线并非全部都是下降的，也有少部分是上升的。将不同走势的高低音线分别进行统计，便得到表 2

¹我们选择的语料共有 1109 个句子，但是有的句子只包含一个韵律词，有的起首或末尾韵律词的音节都是阴平调，或者起首或末尾韵律词只有一个音节，且这个音节是上声调，这样的句子就没有统计下倾斜率，符合统计条件的有 730 个。

所示的结果:

表 2 不同走势高低音线的平均斜率

	高音线下降	高音线上升	低音线下降	低音线上升
平均斜率(半音/秒)	3.42	-1.51	4.25	-1.14
样本数	593	138	615	116
占总体比例(%)	81.1	18.9	84.1	15.9

从表 2 可以看到,不管是高音线还是低音线,下降的都占百分之八十以上,同时也有百分之十几是上升的。高低音线上升的例子如图 3 所示:

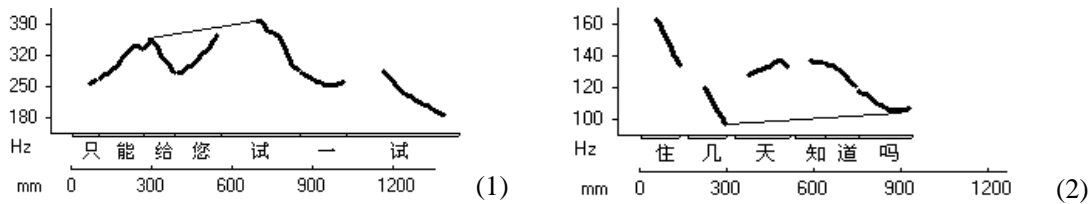


图 3 高音线和低音线上升举例

图 3-1 的高音线是上升的,图 3-2 的低音线是上升的。表 2 显示,大多数语句的高低音线都是下降的,音高前高后低的模式有标界功能。那么这种上升的走势是什么原因造成的呢?观察发现,高低音线上升与三个因素有关,一是词语负载信息量的大小,一是焦点的位置,再就是音节的声调。通常焦点总是出现在句子的末尾(Chafe, 1994; 张伯江、方梅, 1996; 陈昌来, 2000),句末焦点位置上的词语一般都会延长,但是往往音阶却相对较低。然而如果一句话开头词语所负载的信息量很小而末尾词语负载的信息量很大,二者负载的信息量相差悬殊,则末尾词语的语音形式会变得非常突显(prominent),音域增大很多,高音点会超过前面的词语,结果使得语句的高音线出现上升走势(如图 3-1 所示)。焦点通常都在句末,但是有时也会前移至句首,图 3-2 所示的“住几天知道吗”。这句话就是先把重要的信息说出来,再说次要的信息,其焦点在“住几天”上。在“住几天”中,处在焦点核心位置上的音节声调是上声,上声成为焦点,音高的低点会变得更低(林茂灿, 2002),结果使得语句的低音线出现上升走势。

在汉语中音高具有多重功能,首先声调就取决于音高,汉语声调具有区别意义的作用;其次句末音节音高的变化是决定陈述和疑问语气的重要因素(林茂灿, 2006);再就是韵律单元的音高逐渐下降具有标界功能;除此之外,音域扩大还是表征焦点的重要手段(Xu, 1999; 王蓓等, 2002),而且人的喜怒哀乐等不同的情感在音高方面也有不同的表现。既然音高具有多重功能,它的具体表现就是这几种功能的综合体现,并且这几方面的力量还表现出相互竞争、此消彼长的态势。句子的信息编排往往是遵循从旧到新的原则,越靠近句末的信息内容就越新,焦点常出现在句末。然而句末焦点通常可以通过音节延长来表示,所以大多数小句的末尾音阶都相对较低,这种模式具有标界功能。从信息传递的角度而言,不同词语所负载的信息量大小各不相同,这里称之为载义重度。载义重度大的词语,语音形式会比较突显,这种突显主要表现在音高和时长两个方面,而且载义重度越大,语音形式就越突显,这反映了语音和语义的“象似性”(iconicity)(沈家煊, 2000)。如果开头词语的载义重度很小而末尾词语的载义重度很大,并且处在焦点核心位置的音节又不是上声,就可能出现高音线上升的情况。焦点通常是在句末,但有时也会前移至句首,如果焦点在句首,而且处在焦点核心位置的音节是上声或去声,则有可能低音线会上升。

3.3 陈述句和疑问句的音高表现

本研究使用的语料具有语气标注,主要标出了陈述、疑问、祈使、感叹等语气。由于祈使句和感叹句的样本较少,没有统计意义,本节将主要考察陈述句和疑问句的音高表现。疑问句有不同的类型,语料中是非问和特指问句出现较多。在是非问句中,有的句末有疑问语气词,如“吗”、“吧”等,有的句末没有疑问语气词。不同类型句子高低音线的斜率如表 3

所示:

表 3 不同类型句子的高低音线平均斜率

	陈述	疑问		
		特指问	是非问	
			有语气词	无语气词
样本数 ²	329	82	76	50
高音线平均斜率 (半音/秒)	2.46	2.61	2.56	2.48
低音线平均斜率 (半音/秒)	3.38	3.42	3.37	3.52

表 3 显示,从平均斜率看,陈述句及疑问句的高音线和低音线都是下降的,疑问句的音高并没有上升。有一种观点认为,陈述句和特指问句的语调是下降的;是非问句,特别是没有句末疑问语气词的是非问句的语调是上升的。但是这里基于自然话语语料,以句首和句末韵律词为基准进行考察,发现句末没有疑问语气词的是非问句的高低音线也是下降的,而且斜率并不比陈述句的斜率小。林茂灿(2006)基于朗读语料,发现疑问语气主要表现在末尾音节上,这里我们基于自然话语语料,再来考察一下句末韵律词各音节的音高表现。句末韵律词有双音节的,也有单音节、三音节和四音节的³。但是受汉语词语双音化的影响,单音节、三音节和四音节的韵律词都比较少,所以这里只讨论双音节的情况。另外,普通话阴平没有“低”特征,上声没有“高”特征,故此这里没有统计含有这两种声调的韵律词;如果陈述句、特指问句和无语气词的是非问句末尾音节是轻声,也没有统计。我们对陈述句、特指问句、有语气词的是非问句和无语气词的是非问句末尾两音节韵律词的音高表现分别进行了考察。具体做法是,用前音节的高音点减去后音节的高音点,得到高音点降幅,用前音节的低音点减去后音节的低音点,得到低音点降幅。结果见表 4:

表 4 末尾两音节韵律词高低音点降幅

	陈述	疑问		
		特指问	是非问	
			有语气词	无语气词
高音点平均降幅 (半音)	1.21	2.12	2.24	0.71
低音点平均降幅 (半音)	1.97	2.37	2.91	1.31
样本数	66	19	15	10

从表 4 可以看出,无论是高音点还是低音点,平均降幅都是句末没有疑问语气词的是非问句最小,这正是它与陈述句的差别所在。句末有疑问语气词的是非问句末尾音节都是轻声,它的高音点和低音点降幅都是最大的,特指问句末尾音节的降幅也很大。

上文谈到,以韵律词为基准进行考察,句末没有疑问语气词的是非问句下倾的斜率并不比陈述句小(见表 3)。考察句末双音节韵律词的音高表现可以看出,无疑问语气词的是非问句高音点和低音点的降幅都是最小的(见表 4)。那么陈述句和疑问句的整体音域和音阶是否有差别呢?本文对此也进行了考察,小句音域的计算公式如下:

$$Ra_i = St_{\max} - St_{\min} \quad (4)$$

Ra_i 表示小句的音域, St_{\max} 和 St_{\min} 分别表示该小句的高音点和低音点,单位是半音。至于小句的音阶,这里用它的低音点表示,因为人在说话时通常低音点是比较稳定的。另外,本研究使用的是数十个说话人的语料,每个说话人的自身音域各不相同,这就会影响小句的音域和音阶。说话音域宽者的小句音域也会比较宽,说话音域窄者的小句音域也会比较窄。为了便于比较,这里又对小句的音域和音阶进行了归整处理,归整公式为:

²这里的样本总数只有 537 个,祈使句、感叹句、选择问和正反问句没有统计。

³因为语气词不能独立成为韵律词,它必然要附着在前面的词语上。如果前面的词语恰好是三音节的,就会出现四音节韵律词。

$$Ra_n = \frac{Ra_i}{Ra_s} \quad (5) \qquad Rg_n = \frac{Rg_i}{Ra_s} \quad (6)$$

Ra_n 和 Rg_n 分别表示小句音域和音阶的归整值, Ra_i 和 Rg_i 分别表示小句的实际音域和音阶, Ra_s 表示说话人的音域。说话人的音域是用说话人在整个谈话中的音高最大值减去音高最小值得到的。因为 Ra_s 是说话人谈话中表现出的整体音域, Ra_i 和 Rg_i 只是其中某个小句的音域和音阶, Ra_i 和 Rg_i 必然都小于 Ra_s , 所以 Ra_n 和 Rg_n 的值都在 0 到 1 之间。陈述句、特指问句和是非问句的音域和音阶归整值如表 5 所示:

表 5 不同类型句子的音域和音阶归整值

	陈述	疑问	
		特指问	是非问
音域平均归整值	0.57	0.58	0.59
音阶平均归整值	0.21	0.23	0.21
样本数	329	82	126

表 5 显示, 与陈述句相比, 疑问句的音域比较大, 特指问句的音阶比较高。结合表 5 我们再看表 4, 表 4 显示, 特指问和句末有疑问语气词的是非问句末尾两音节音高降幅较大, 这与它们的整句音域较大有关, 正是因为疑问句的整句音域比较大, 才使末尾两音节音高有了降幅较大的空间。

在英语中, 是非问句末尾语调通常会上升, 但是这里的考察发现, 汉语疑问句末尾音高并不上升, 只是句末没有疑问语气词的是非问句末尾音高下降的幅度比较小。虽然无疑问语气词的是非问句末尾音高没有上升, 但与陈述句相比, 其末尾音节音高的相对位置比较高。可以说, 相对于陈述句而言, 它是“上升”了。对于一个没有词汇和句法形式标记的疑问句, 人们正是依据这种相对的“上升”来识别它的。我们还发现, 无论是非问句还是特指问句, 其音域都比陈述句大, 这也是分辨疑问句和陈述句的重要语音特征。

与朗读话语相同, 本研究发多数语句的音高都表现出下倾走势。不过这里使用的语料是自然话语, 语句都比较短小, 而且语句内部词与词都比较连贯, 多数语句内部并无韵律短语边界。所以, 语句内部韵律单元之间的音高重置现象便比较少。此外, 朗读话语的音高下倾趋势是, 句首韵律短语的下倾程度大于句末韵律短语的下倾程度。本研究使用自然话语语料, 没有发现这种现象, 这可以认为是自然话语和朗读话语的不同之处。

4 结 语

下倾的产生有其生理上的原因, 而且在功能上又具有标界的作用, 所以这一现象具有普遍性。汉语也不例外, 汉语的语句也存在下倾现象。本文发现在汉语自然话语中, 大多数语句的高音线和低音线都是下降的, 而且最后还有降尾出现。但是也有一小部分语句的高低音线并不下降, 这主要与词语的载义重度、焦点的位置及声调有关。另外, 无论是陈述句还是疑问句, 大多数小句的高低音线也都是下降的。句末没有疑问语气词的是非问句末尾音节音高并不上升, 只是音高位置相对较高。对小句整体音域和音阶的统计发现, 无论是非问句还是特指问句, 其音域都比陈述句大。句末没有疑问语气词的是非问句末尾音节的音高相对较高, 疑问句的整体音域较大, 这是疑问句的两个重要语音特征。

参考文献

[1]曹剑芬. 汉语声调与语调的关系[J]. 中国语文, 2002, (3).
 [2]陈昌来. 现代汉语句子[M]. 武汉: 华中师范大学出版社, 2000.
 [3]黄贤军, 高路, 杨玉芳, 吕士楠. 汉语语调音高下倾的实验研究[J]. 声学学报, 2009, (2).
 [4]李爱军. 普通话对话中韵律特征的声学表现[J]. 中国语文, 2002, (6).
 [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).
 [6]林茂灿. 疑问和陈述语气与边界调[J]. 中国语文, 2006, (4).
 [7]林 焘, 王理嘉. 语音学教程[M]. 北京: 北京大学出版社, 1992.

- [8]沈家煊. 认知语法的概括性[J]. 外语教学与研究, 2000, (1).
- [9]沈 炯. 汉语语调构造和语调类型[J]. 方言, 1994, (3).
- [10]王安红, 陈 明, 吕士楠. 基于言语数据库的汉语音高下倾现象研究[J]. 声学学报, 2004, (4).
- [11]王 蓓, 吕士楠, 杨玉芳. 汉语语句中重读音节音高变化模式研究[J]. 声学学报, 2002, (3).
- [12]吴宗济. 中国音韵学和语音学在汉语言语合成中的应用[J]. 语言教学与研究, 2002, (1).
- [13]张伯江, 方 梅. 汉语功能语法研究[M]. 北京: 北京大学出版社, 1996.
- [14]宗成庆, 吴华, 黄泰翼, 徐波. 限定领域汉语口语对话语料分析[A]. 计算语言学文集[C], 清华大学出版社, 1999.
- [15]Chafe, W. L. *Discourse, Consciousness, and Time*[M]. Chicago: The University of Chicago Press, 1994.
- [16]Cruttenden, A. *Intonation*[M], Cambridge: Cambridge University Press, 1997.
- [17]Fletcher, J., Evans, N. and Round, E.. Left-edge tonal events in Kayardild (Australian) — a typological perspective[A], in B. Bel and I. Marlien (eds.) *Proceedings of Speech Prosody 2002*[C], Aix-en-Provence, France, 2002.
- [18]Lieberman, P., Knudson, R. and Mead, J.. Determination of the rate of change of fundamental frequency with respect to subglottal air pressure during sustained phonation[J]. *Journal of the Acoustical Society of America*, 1969, (45).
- [19]Poletto, C. J., Verdun, L. P., Strominger, R. and Ludlow, C. L. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures[J], *J Appl Physiol*, 2004, 97(3): 858-866.
- [20]Xu, Y.. Effects of tone and focus on the formation and alignment of F0 contours[J]. *Journal of Phonetics*, 1999 (27): 55-105.