

# 单念条件下作用于普通话两字组之上的音高降阶效应

熊子瑜

中国社会科学院语言研究所

xiongziyu@163.com

## 摘要\*

为了细致地分析普通话两字组的音高模式,本文基于大规模语音语料库,考察了单念条件下作用于后字为零声母或浊声母的两字组之上的音高降阶效应。统计数据表明,单念条件下,当普通话两字组的声调音高特征构成“..H..L..H..”的音调序列时,处于该序列中L之后的H调会发生显著的调阶降低现象,这种现象被称为“降阶效应”,降阶效应可以用来简便而精确地解释普通话两字组音高模式中出现的一些有规则的异常现象。统计数据还表明,降阶效应应在不同的语音环境里都具有相对稳定的音高表现。基于研究结果,本文认为,在普通话语流中,每个音调位置上的音高值受声调音高特征(如“H”和“L”)和作用于声调音高特征之上的各类音高调节效应(如降阶效应)共同决定着。

关键词: 音高模式 音高特征 降阶效应

## 1. 引言

在AM理论(Autosegmental-Metrical theory)看来,任何语言的音高曲线(Pitch contour)都是由局部音高事件(Local pitch event)和音高过渡(Pitch transition)按照一定的规则线性组合而成[2]。英语里最为重要的局部音高事件是音高重音(Pitch accent),普通话里最为重要的局部音高事件是字调(syllabic tone)。在AM理论里,局部音高事件用音高特征H和L及其组合形式来描写,如Pierrehumbert [4]将英语的音高重音分析为七种类型,分别标示为:H\*, L\*, L+H\*, L\*+H, H+L\*, H\*+L和H\*+H。语流中各音高特征的语音实现不仅受控于其自身的音高范畴,如高调和低调,还会受到一些音高调节效应(Modification effects)的影响。D. Robert Ladd [1]研究表明,在调峰(intonational peaks)之上可能出现三类音高调节效应:[±滞后性(delayed)]、[±降阶性(downstepped)]和[±升阶性(raised)]。因此研究一个语言的音高变化模式,不仅需要研究这个语言中存在的局部音高事件,还需要研究这个语言中存在的音高调节效应。

本文在此仅研究降阶现象。从形式上讲,降阶是指短语中一个H调的调值与其前一H调的调值之间存在一个显著的落差。从功能上讲,产生降阶的因素可能有很多,如有研究表明,在非洲的某些声调语言里,“..H..L..H..”音调序列中的后一H调的调值会明显的低于其前一H调的调值[7],这种存在于“..H..L..H..”音调序列中的降阶现象不仅出现于非洲的声调语言中,也同样

出现在英语和日语等非声调语言里[4][5],因此被认为是一种很普遍的音高变化现象。

普通话是声调语言,有阴平、阳平、上声和去声四个声调,本文用音高特征“H”和“L”以及它们的组合形式“HL”、“LH”来描写普通话的声调,阴平调标示为“H-”,阳平调标示为“LH”,上声调标示为“L-”,去声调标示为“HL”,“-”表示对前一音高特征的扩展(tone spread)。除了“上上”连读变调外,普通话两字组可以组合成15种不同的音高序列,其中有5组包含“..H..L..H..”的音调序列,其余的都不包含“..H..L..H..”的音调序列,因此很方便用来研究“..H..L..H..”音调序列中的降阶现象。另外,由于普通话是声调语言,可以确定音高曲线的声调音高特征,因此便于将声调音高特征的音高表现与降阶效应的音高表现分离开来,从而考察降阶效应的音高表现。

## 2. 语音现象

本文分析所采用的数据取自中国社会科学院语言所建立的两音节语料库,该语料库由一位说标准普通话的成年男性发音人朗读。本研究仅从该语料库中选取了1139个后音节为零声母或浊声母的两字组进行数据分析,其中各种声调组合的样本数目如下表(一)所示:

表(一): 16类声调组合的样本数目

声调组合	数目	声调组合	数目	声调组合	数目	声调组合	数目
阴阴	69	阴阳	114	阴上	55	阴去	107
阳阴	56	阳阳	55	阳上	49	阳去	103
上阴	34	上阳	39	上上	31	上去	72
去阴	70	去阳	73	去上	77	去去	135

数据分析过程如下: 1.用Praat<sup>1</sup>语音分析软件所提供的标注工具对每个两字组进行音节以及声韵调的标注; 2.用Praat语音分析软件所提供的音高分析工具对每个两字组进行音高分析,并对所提取出的音高数据进行了人工纠正以保证音高数据的准确性; 3.根据音节标注数据和修改后的基频数据,对每个音节的基频段按其时长比例等分并提取出10个基频值; 4.按照声调组合类型的不同进行分类,用SPSS<sup>2</sup>统计软件计算出各个点的音高平均值来。

下面图1.1至图1.8中的音高曲线就是根据最后的音高平均值做出的,每条音高曲线都是根据几十个甚至更

\* 本研究部分受国家自然科学基金项目资助(项目编号:60475043),部分受中国社会科学院青年科研启动基金项目资助。

<sup>1</sup> <http://www.praat.org>

<sup>2</sup> <http://www.spss.com>

多的样本数据综合平均后得到的结果，因此能够在一定程度上消除随机误差对系统的影响。

图（一）：发音人X所读的两字组音高均值曲线<sup>3</sup>

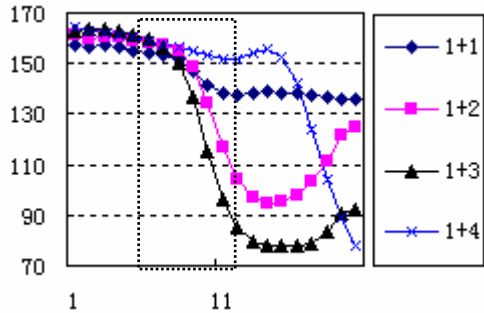


图1.1 “阴平+四种声调”组合

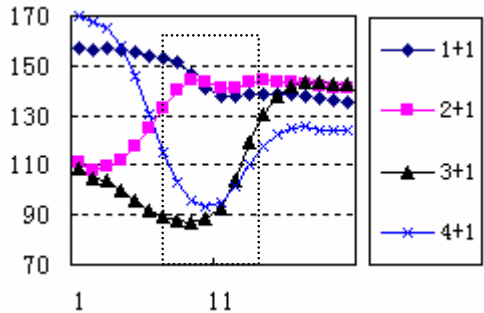


图1.2 “四种声调+阴平”组合

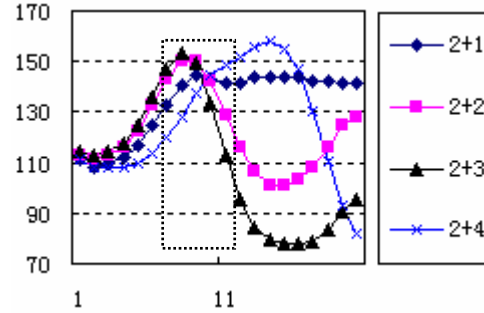


图1.3 “阳平+四种声调”组合

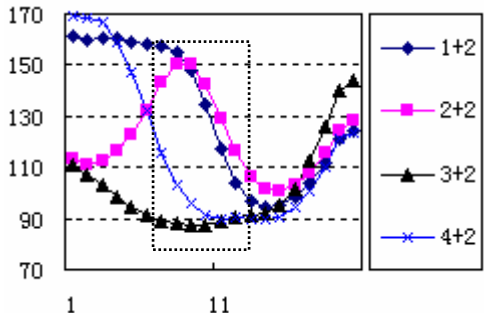


图1.4 “四种声调+阳平”组合

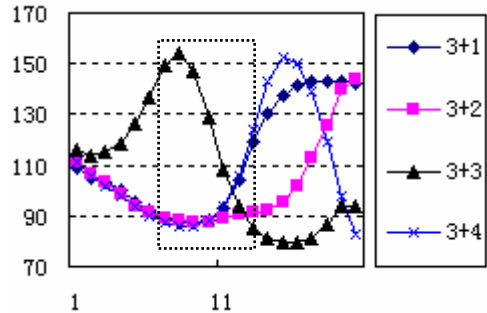


图1.5 “上声+四种声调”组合

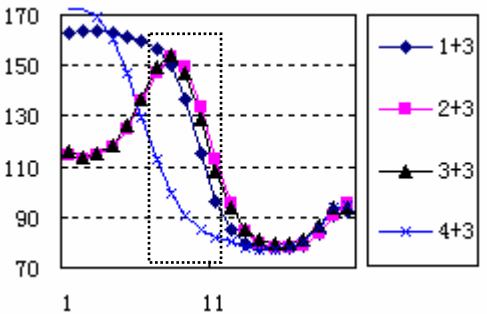


图1.6 “四种声调+上声”组合

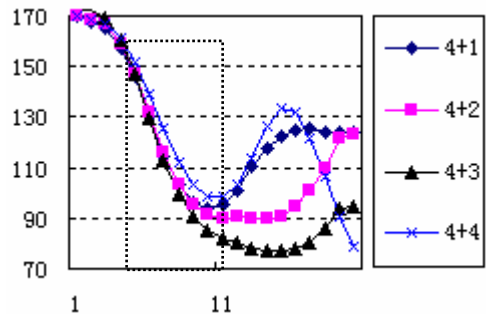


图1.7 “去声+四种声调”组合

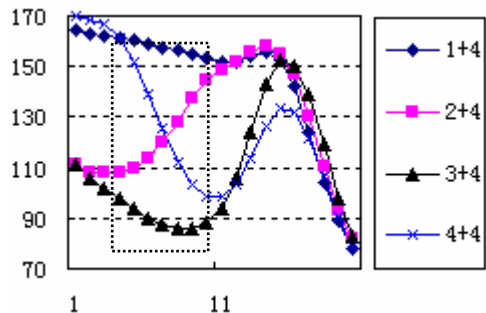


图1.8 “四种声调+去声”组合

从上面的这些数据可以清楚地看出：（1）不管后接哪种类型的字调，都不会影响前字调首的音高值，这可以从图1.1，图1.3，图1.5和图1.7中第一个音节的起始音高数据看出；（2）不管前接哪种类型的字调，通常都不会影响后字调尾的音高值，这可以从图1.2，图1.4，图1.6和图1.8中第二个音节的末尾音高数据看出（当第二个音节为阴平和阳平时，其尾调的音高实现情况显著不同，后面对此进行了解释）；（3）前字调尾和后字调首之间存在一定长度的音高过渡现象，大致相当于以上各图中虚框圈住的部分，音高过渡段的时长可以根据一定的因素加以精确预测[8]。从数据可以看出，尽管音高过渡段在不同的声调组合中表现出很大的差异，但如果撇开音高

<sup>3</sup> 图中的“1”、“2”、“3”和“4”分别表示“阴平”、“阳平”、“上声”和“去声”，因此“1+1”表示“阴平+阴平”组合，其他的以此类推。

过渡段, 不难发现: a) 不管后接哪种类型的字调, 都不会显著影响前字调尾的音高值, 这可以从图1.1, 图1.3, 图1.5和图1.7中第一个音节的末尾音高数据看出; b) 不管前接哪种类型的字调, 都不会显著影响后字调首的音高值, 这可以从图1.2, 图1.4, 图1.6和图1.8中第二个音节的起始音高数据看出(当后字为阴平和去声时, 其首调的音高实现情况显著不同, 后面对此进行了解释)。总之, 观察数据表明, 两字组单念时, 前后字调之间的影响实际很小。另外, 从这些数据可以看出, 发音人能够较为精确地控制音高目标值的语音实现, 否则就不会出现这些稳定的音高表现。

除了上面所说的这些较为一致的音高表现外, 还可以看出一些不太一致的音高现象, 归结为以下几点: 现象(一), 根据图1.2可以看出, 阴平字处于去声之后时, 其调值整体偏低; 阴平字处于其他三类声调之后时, 其调值整体偏高。现象(二), 根据图1.8可以看出, 去声字处于另一个去声之后时, 其调首的调值偏低; 去声字处于其他三类声调之后时, 其调首的调值偏高。现象(三), 根据图1.4可以看出, 阳平字处于上声之后时, 其调尾的调值偏高; 阳平字处于其他三类声调之后时, 其调尾的调值偏低。现象(四), 比较图1.7和图1.5、图1.3、图1.1中的末音节之四声音高聚合数据可以看出, 在去声之后的四声聚合音域明显压窄, 而这种压窄主要是由于后阴平字和后去声字的高调被降低所造成。

上面各图中的音高曲线是根据几十个甚至更多的样本进行平均后做出的, 因此可以肯定这些不一致现象是有规则的, 应该有其背后的成因。那么如何解释这么四种有规则的不一致现象呢? 本文接下来将考察这些有规则的不一致现象, 以分析其背后的成因。

### 3. 降阶效应

如前文所述, 已有的研究表明, 非洲声调语言的“..H.L.H..”音调序列中, 后一个高调的调值会明显的低于前一个高调的调值, 这类现象被学界称为“降阶(downstep)”。降阶现象不仅出现于非洲的声调语言中, 还同样出现在日语和英语等非声调语言里, 因此被认为是一种很普遍的音高变化现象。这种由于“..H.L.H..”音调序列所引发的“降阶现象”包含两点基本含义: (1) 只有当高调特征(H)与低调特征(L)一起构成“..H.L..”序列时, 才可能引发后面的“..H..”产生降阶现象, 仅靠一个“H”特征或一个“L”特征不会引发后面的降阶现象。(2) 降阶效应只作用于“..H.L..”序列之后的高调, 而不会对其后的低调特征“L”产生影响。

普通话是声调语言, 是否也存在类似的降阶现象呢? 下面将普通话四声组合的16种调型按其底层的字调音高特征描述如下表(二)所示:

表(二) 普通话四声组合及其底层调型的描写

声调组合	组合调型	声调组合	组合调型
阴阴	H- --	阴阳	<u>H- LH</u>
阴上	H- L-	阴去	H- -L
阳阴	LH --	阳阳	<u>LH LH</u>
阳上	LH L-	阳去	LH -L

上阴	L- H-	上阳	L- -H
上上	LH L-	上去	L- HL
去阴	<u>HL H-</u>	去阳	<u>HL -H</u>
去上	HL --	去去	<u>HL HL</u>

可以看出, 普通话四声组合的16种调型中具有“..H.L.H..”音调序列的共有5种, 在上表中分别用方框标明。如果普通话中也存在“..H.L.H..”音调序列的降阶现象, 那么就一定会也只能会发生在在这五种组合之上。观察图1.1至1.8中的数据, 会发现在这五种声调组合都无一例外的发生了降阶现象, 导致“..H.L.H..”音调序列中后一个高调的调值明显降低。

普通话里存在“..H.L.H..”音调序列的降阶现象, 这可以用来解释前文所描写的四种有规则的不一致现象: 发生在“去声+阴平”组合上的降阶效应解释了现象(一); 发生在“去声+去声”组合上的降阶效应解释了现象(二); 发生在“阴平+阳平”、“阳平+阳平”和“去声+阳平”组合上的降阶效应解释了现象(三); 发生在“去声+阴平”、“去声+去声”和“去声+阳平”组合上的降阶效应, 导致去声之后的所有高调调值下降, 进而产生了去声之后的四声聚合音域明显压窄现象, 这就解释了现象(四)。再观察图1.1至1.8中的数据, 会发现除了这五种声调组合外, 其他组合都没有发生降阶效应。总之, 用一条“降阶效应”就可以简便而有效地解释前文所述的四种有规则的不一致现象。降阶效应还被有些学者看作是导致语句音高下倾的一种重要因素。

### 4. 降阶效应的分解

基于上文的论述, 可以看到: 在降阶效应的作用下, H调的调值会发生显著降低。本文接下来将受到降阶效应作用的H调标记为!H, 其中“!”表示降阶效应, 比较“!H”与“H”在音高表现上的差异, 从而得出“!”的音高表现。

根据图1.1至1.8的数据, 可以得到三组与降阶效应相关的数据: (1) 当一个去声处于另一个去声之后时, 发生降阶现象, 其“!H”的调值为134赫兹; 当一个去声处于一个阴平或阳平或上声之后时, 不发生降阶现象, 其“H”的调值为152赫兹。二者相比, “!H”的音高值是“H”的88%左右。(2) 当一个阳平处于一个阴平或阳平或去声之后时, 发生降阶现象, 其“!H”的调值为123赫兹; 当一个阳平处于一个上声之后时, 不发生降阶现象, 其“H”的调值为140赫兹。二者相比, “!H”的音高值是“H”的88%左右。(3) 当一个阴平处于一个去声之后时, 发生降阶现象, 其“!H”的调值为122赫兹; 当一个阴平处于另一个阴平或阳平或上声之后时, 不发生降阶现象, 其“H”的调值为138赫兹。二者相比, “!H”的音高值是“H”的88%左右。这三组数据都表明, 降阶效应的作用力是很稳定的, “!H”的音高值是“H”的88%左右, 也就是说, 其作用力大约是对H调的音高值降低2个半音(S<sub>1</sub>)。有关其他语言中降阶现象的研究也有类似的结果, 即根据前一峰值, 用一个恒定的衰减系数就可以精确预测出被降阶后的峰值, 如英语[3]和荷兰语[6]。

## 5. 结论和讨论

普通话两音节四声组合的十六种调型中共有五种调型满足“..H..L..H..”的音调序列特征，分别是“阴平+阳平(H-LH)”“阳平+阳平(LH LH)”“去声+阴平(HL H-)”“去声+阳平(HL -H)”和“去声+去声(HL HL)”。单念时，这五种后字为零声母或浊声母的两字组中都无一例外地发生了降阶现象，而且降幅一致：“!H”的调值与“H”的相比约降低了2个半音。这一结果表明，降阶效应可以用来简便而精确地解释普通话两字组音高模式中的一些有规则的异常现象。

普通话里存在的降阶现象表明，尽管普通话是声调语言，有阴平、阳平、上声和去声四个调类，但分析普通话的音调系统时应将声调描写为由声调音高特征(H和L)组成的线性序列。首先，上文分析表明，单念条件下普通话中的“去声+去声(HL HL)”组合会出现降阶现象，使得后一去声“HL”中的“H”发生调值降低，而“L”的调值却不受影响，这说明降阶现象并非对后一去声整个儿产生影响，而只是对其中的一个“H”音高特征产生影响，要解释降阶现象的这种作用方式则需要将普通话的声调分解为音高特征。其次，上文分析还表明，单念条件下普通话两字组中的“阴平+阳平(H-LH)”、“阳平+阳平(LH LH)”两类声调组合也会出现降阶现象，这说明跨音节边界的“HL”序列也能激发其后的“H”音高特征发生降阶现象，要解释降阶现象的这种作用条件也需要将普通话的声调描写为由音高特征组成的线性序列。

基于以上考虑，本文把普通话的声调描写为两个作用于同一个音节之上的声调音高特征组成的线性序列。声调音高特征一方面是自主的，独立于音段成分，但另一方面又按照一定的规则与音段成分发生关联(association)，声调音高特征与音段成分发生关联的位置，本文称之为“音调位置”，用“P”表示。一个音调位置不仅可以承载声调音高特征，还可以同时承载多种音高调节效应，如“!H”，降阶只是诸多音高调节效应中的一种，这些音高调节效应和声调音高特征共同决定着该音调位置P的音高值，可用公式表示为：

$P=Ps+[\Sigma Pm]+k$ ，其中P表示某个音调位置的音高值，Ps表示声调音高特征的音高值，Pm表示作用于该声调音高特征之上的某种音高调节效应的音高表现，“+”在此不表示代数和关系，只表示音高调节效应和声调音高特征共同作用于一个音调位置， $\Sigma$ 也不表示累加，只表示作用于一个音调位置上的音高调节效应可能不止一种，方括号表示可选，也就是说一个音调位置上可以不负载任何音高调节效应，k表示随机误差。而至于普通话中除了存在降阶效应之外，还存在哪些类型的音高调节效应以及这些音高调节效应的作用条件和方式等等问题还有待深入的研究。

## 参考文献

- [1]Ladd, D. Robert. Phonological features of intonational peaks [J]. *Language* 1983. 59. 721-59.
- [2]Ladd, D. Robert. *Intonational Phonology*[M]. New York: Cambridge University Press. 1996.
- [3]Lieberman, M. and J. Pierrehumbert. Intonational Invariance under Changes in Pitch Range and Length [C]. in M. Aronoff and R. Oehrle, eds, *Language Sound Structure*, MIT Press, Cambridge MA. 1984. 157-233.
- [4]Pierrehumbert, J., *The Phonology and phonetics of English intonation* [D], PhD thesis, MIT. 1980.
- [5]Pierrehumbert, J. and M. Beckman. *Japanese Tone Structure* [M]. Cambridge: MIT Press. 1988.
- [6]R. van den Berg, C. Gussenhoven and T. Rietveld. Downstep in Dutch: Implications for a model [C]. In G. J. Docherty and D.R. Ladd, eds, *Papers in Laboratory Phonology II: Segment, Gestures, Tone*. Cambridge: Cambridge University Press. 1992.
- [7]Stewart, J.M. Downstep and Floating Low Tones in Adioukrou [J]. *Journal of African Languages and Linguistics*. 1983. 5. 57-78.
- [8]Xiong, Ziyu. The duration of pitch transition phase and its relative factors [C]. in *Proceedings of the 8th International Conference on Spoken Language Processing*, Jeju, South Korea. 2004.

## Downstep Effect on Disyllabic words of Citation Form in Standard Chinese

XIONG Ziyu

Institute of linguistics, CASS.

xiongzuyu@163.com

**ABSTRACT:** The purpose of this study is to specifically analyze the pitch contours of disyllabic words in Standard Chinese. Based on a huge speech corpus, the paper investigates the downstep effect on the second syllable with a zero initial or a voiced initial in disyllabic word of citation form. Statistical results show that when the tonal pitch features of a disyllabic word constitute a “..H..L..H..” tonal sequence, the H tone behind the L tone will be lowered

contrasted to the preceding H tone obviously, and such phenomenon is called downstep. It is also shown in this paper that downstep effect happened on disyllabic words in Standard Chinese has relatively stable pitch manifestation in different kinds of tonal combinations. The results found in this paper can be applied to generate pitch contours properly to improve the naturalness of the speech synthesis.