心 理 学 报
第3期 ACTA PSYCHOLOGICA SINICA 1988年

塞辅音和声调知觉问题[1)]•

杨玉芳 金凌娟

中国科学院心理研究所

摘 要

本文报告的两个实验以合成的汉语普通话CV音节作刺激，研究了塞辅音与声调之间在知觉上的相互作用问题。主要结果是：(1)塞辅音的发音方式影响声调的知觉，不送气音使听者在辨别声调时倾向于基频曲线起点高的声调反应；(2)音节的声调也影响对塞音发音方式的判断，在一、四声音节里，听者倾向于将塞辅音听成不送气音，在二、三声音节里，听成送气音。

引 言

语音学和言语知觉研究中的一个明显事实是，言语的基频不仅取决于高层次上的语言学信息，如句调、声调等，还受到组成言语声的音段，如元音、辅音等的影响。由音段引起的基频变化一般幅度较小，文献中称之为微观韵律语音学特征(Microprosody)或局部扰动(Localperturbation) (Hombert, 1978[1])。通常认为，这种变化是由发音器官的生理限制引起的，不能由发音人控制；在言语知觉中，它又成为识别元音或辅音的声学线索。长期以来，人们已对此作了一定的研究。但是，要彻底了解自然语言的韵律规则，并把它成功地运用到语音合成和识别中去，还必须从多方面对它加以细致的研究。

在这个领域里，很大部分工作集中于研究塞辅音与元音的基频之间的关系。实验结果证明，一方面，塞辅音的清浊影响后面的元音的基频，特别是它的开始部分。清音后面的元音基频高，浊音后面的低。但清塞音的送气程度怎样影响元音的基频，研究结果很不一致。另一方面，元音基频又会影响到前面的辅音。目前尚未见到有关声学分析结果。但已有不少知觉方面的实验结果证明，元音基频起始点和走向可以影响前面塞辅音发音方式的识别，特别是当决定发音方式的声学线索VOT值比较模糊时(Haggard, 1970)[2]。

在汉语普通话和其他一些方言里，塞辅音都是清音，有送气与不送气之分，可以用来研究送气程度与元音基频扰动的关系。特别是，汉语是声调语言，使得问题变得更复

1) 本文于1988年1月7日收到。
• 本实验使用的刺激由中国社会科学院语言研究所杨顺安同志协助合成，在此表示感谢。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

杂，更有意义。Ling King等(1986)[3]分析了一名发音人用上海方言发的CV(辅音+元音)音节中，塞辅音发音方式、元音类型与声调之间的相互作用。分析结果表明，在不送气塞音之后的元音基频高，送气音之后低，这一差别在浊音开始后15—30ms左右消失；塞音的发音位置不显著影响基频模式；元音的高度对基频有显著影响，这种影响又与塞音的发音方式以及音节的声调有关。Zee(1980)[4]对广东话作的研究结果与Ling King的相反，在送气音之后元音基频起点高，在不送气音之后低。

本文报告的实验使用汉语普通话的CV音节，从知觉角度初步研究塞辅音与声调之间的相互影响，即塞音送气与否怎样影响音节声调的知觉，音节的声调模式又如何影响塞音发音方式的辨别。实验结果可能为清塞音送气与否与声调的关系提供一个例证，也可能为语音合成中处理VOT与声调的关系提供参考。实验使用音位或声调的范畴边界和平均反应百分数的变化作指标。

实验一、塞辅音与声调知觉

这个实验研究清塞音的发音方式，即送气与不送气，对声调模式知觉的影响，特别是声调范畴边界随辅音变化的方向。

1、方法

(1)、刺激：

考虑到汉语普通话中塞音、元音和四声之间可能的搭配，CV音节中的辅音选了p，b和t，d两对塞音，元音为a(本文中的音标是汉语拼音符号)。四个声调连续体是：pā—pá，bā—bá，tǎ—tà和dǎ—dà。

实验中使用的所有音节为合成语音，用在GW0520CH微型计算机上运行的共振峰合成软件系统[5]合成，D/A转换器为12bit，采样频率10KHZ，低通滤波器截止频率3.9KHZ.

使用发音参数合成，每个CV音节由爆破、音渡(Transition)和稳态元音三部分组成。不送气音音节长260ms，其中爆破部分10ms，音渡50ms，元音200ms，音渡部分由浊声源激励，爆破与浊声源开始之间的时间间隔，即VOT，为0ms，送气音音节长300ms，爆破部分10ms，音渡60ms，元音230ms。音渡部分前40ms用清声源激励，后20ms用浊声源，VOT为40ms。所有音节浊音部分时长相等，以保证对同一种声调模式有相同的基频变化率。合成语音为男声，塞音的音渡频率F_{T1}—F_{T3}和元音a的共振峰频率F_1—F_5 见表1

表1　辅音音渡频率和元音共振峰频率　　单位：HZ

p，b	t，d	a
$F_{T1}=400$	$F_{T1}=300$	$F_1=850$
$F_{T2}=700$	$F_{T2}=1800$	$F_2=1250$
$F_{T3}=2000$	$F_{T3}=2800$	$F_3=2600$
		$F_4=3500$
		$F_5=4500$

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

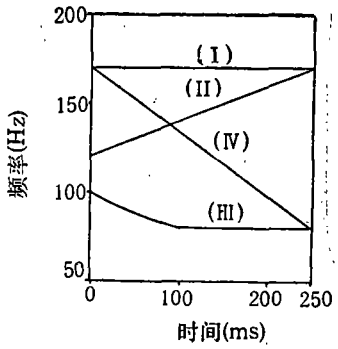

图1 四声声调模式

四声的声调模型如图1所示。基频变化的范围为80—170HZ，为典型的男声声调。为简化分析，没有考虑弯头降尾。将两对声调，一声、二声和三声、四声所复盖的频率范围等分为10份，构成声调连续体中刺激的基频模式。每个声调连续体中11个刺激，一、二声连续体中，相邻两个刺激的基频起点相差5HZ；三、四声连续体中，相隔7HZ。就是说，一、二声和三、四声连续体中步长分别为5HZ和7HZ.

(2)程序：

刺激是通过GF666录音机提供的。录音机放在听者前方一米左右，音量调到听者感觉适度。实验在隔声室中进行，室内噪声不高于40dB。共两组刺激。第一组是将两个一、二声连续体中的刺激混合后以随机方式提供，每个刺激重复10次。第二组是将三、四声连续体中的刺激混合后随机呈现，每个重复10次，两个刺激之间的时间间隔是5秒左右。

听者的任务是分辨所听到的刺激属哪个声调，并在答案纸上作出相应的记号。有15名听者参加了实验，年龄在20—30岁，没有听力问题和耳疾，讲普通话。

2、结果和分析

分析中，首先计算了所有听者对每个刺激所作的一声和三声反应的平均反应百分数。结果由图2中的反应曲线表示，音节中的辅音作曲线的参数。看来，塞音的发音方式对于音节的声调知觉是有一定影响的，两条一声反应曲线和三声反应曲线有偏移。

为说明塞音的发音方式对声调知觉的影响是否显著，分析中又进一步计算了一声和三声反应的平均百分数R_M、标准差S_D，并对在送气音和不送气音条件下得到的结果作了t考验分析。结果见表2。

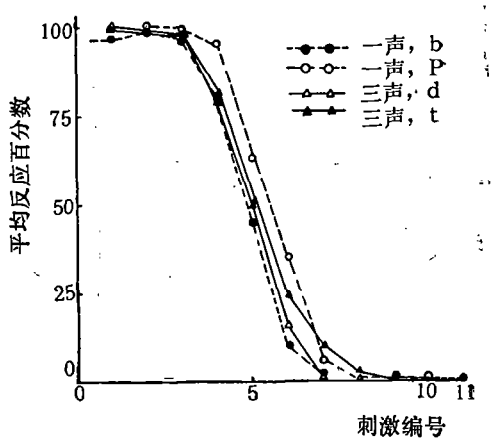

图2 一声和三声反应曲线

分析结果是，对pā—pá连续体中的刺激作的一声反应($R_M=45.8\%$)比对bā—bá连续体作的($R_M=48.0\%$)少，平均反应差R_D为2.4%。t考验分析结果证明，一声反应百分数差异显著($P<0.001$)。这说明，元音前的送气音使一、二声之间的边界向一声方向移动。

对tǎ—tà连续体作的三声反应($R_M=41.3\%$)多于dǎ—dà连续体($R_M=40.0\%$)，反应差为1.3%。t考验分析结果是，两个三、四声连续体的三声反应百分数差异显著($p<0.05$)。如同一、二声连续体中的p的作用一样，t的作用使三、四声的边界向基频起点

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表2　平均反应百分数和统计分析　　(%)

目标声调	一声		三声	
塞辅音	p	b	t	d
平均反应% R_M	45.6	48.0	41.3	40.0
标准差 S_D	4.4	4.3	6.8	6.4
平均反应差 R_D	2.4		1.3	
t 考验结果	$t(13)=6.612$, $p<0.001$		$t(13)=2.012$, $p<0.05$	

高的四声方向移动。

有两点值得一提。第一，结果中的平均反应标准差S_D比平均反应差R_D大。反映出这样一个事实，听者对一、二声和三、四声作的辨别差别较大，但塞音发音方式对声调辨别的影响却是一致的。第二，p对一、二声连续体的辨别的影响似乎大于t对三、四声连续体的影响，从平均反应差R_D和t考验结果中可以看到这一点。

3、讨论

本实验要观察的是塞辅音的发音方式，即送气与不送气，对声调知觉的影响。两对声调连续体的结果都说明，在送气音之后作的一声和四声反应比在不送气音之后少。一、四声的共同点在于基频曲线的起点高。这样，送气音的作用似乎是相当于降低了基频曲线的起点。利用图2中的反应曲线，可以粗略地估计一下发音方式影响的幅度。根据一声和三声反应曲线上面对百分之五十点之间的步数间隔和各自的步长，由发音方式引起的知觉上的基频起点变化约为2—4HZ。根据文献中提供的数据，声学测量中观察到由发音方式引起的元音基频变化约为10—20HZ。可见，发音方式引起的知觉上的变化要小得多。

从反应曲线还可以看出，两对反应曲线的差异集中于声调范畴边界附近，对连续体两端附近的刺激的反应，几乎不受发音方式的影响。因此，与音节的基频模式相比，塞辅音对于声调的辨别的作用是次要的。只有当声调知觉的主要线索——基频曲线不够清楚时，塞音的发音方式才会产生影响。

实验二、声调与发音方式的辨别

这个实验的目的是检验音节的声调怎样影响对塞音发音方式的辨别。决定塞音发音方式的主要线索是VOT。所以，具体地说，这里是检验音节的声调与送气和不送气音沿VOT连续体的音位边界的关系。

1、方法

通过合成产生四个VOT连续体：bā—pā，bá—pá，dǎ—tǎ和dà—tà。语音合成条件与实验一相同，辅音的音渡频率F_{T1}—F_{T3}和元音的共振峰频率F_1—F_5等参数同表1所列一样。四声声调模式如图1所示。处于连续体两端位置的音节的VOT及其他时间参数

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

见表3。ba—pa连续体中有9个刺激，VOT变化范围0—40ms，步长5ms。da—ta连续体中10个刺激，VOT为0—45ms。

表3 音节的时间参数 单位：ms

音节	爆破	音渡		元音	总时长	VOT
		清声源	浊声源			
ba	10	0	60	190	260	0
pa	10	40	20	230	300	40
da	10	0	65	185	260	0
ta	10	45	20	230	305	45

实验中，分别将两个ba—pa连续体和da—ta连续体中的刺激混合后以随机顺序呈现给听者。刺激的呈现条件以及听者的反应方式如实验一所述。听者是参加过实验一的同一批人，他们的任务是判断每个CV音节中的辅音是送气音还是不送气音。

2、结果和分析

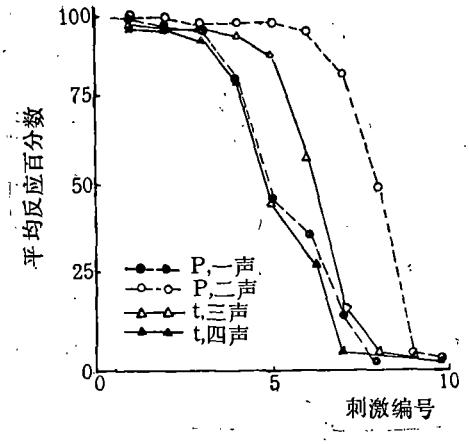

图3 塞音识别反应曲线

所有听者辨别塞音发音方式的平均反应曲线如图3所示。可以看出，声调模式对于区分送气与不送气音的影响很大。

计算出的送气音反应的平均百分数R_M，标准差S_D，平均反应差R_D以及t考验结果见表4。

统计分析结果表明，当音节声调为一声时，将辅音听成p的可能性比在二声时要小得多，平均反应百分数差R_D为29%。t考验结果证明，声调的影响显著($t(12)=2.740$，$P<0.01$)。当声调为四声时，听成t的可能性比在三声时要小得多，平均反应百分数差R_D为11%，t考验证明声调对塞音辨别作用显著($t(12)=1.784$，$p<0.05$)。这两组结果都说明，当声调模式的起点比较高时，塞音更容易被听成为不送气音。如同实验一的结果一样，三、四声引起的发音方式辨别上的差异小于一、二声引起的

表4 平均反应百分数和统计分析

目标音位	p		t	
声调模式	一声	二声	三声	四声
平均反应% R_M	51.0	80.0	55.0	44.0
标准差 S_D	11.9	7.1	8.3	8.6
平均反应%差 R_D	29.0		11.0	
t 考验结果	$t(12)=2.740$，$p<0.01$		$t(12)=1.784$，$p<0.05$	

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

变化。

3、讨论

这个实验检验音节的声调模式对塞音发音方式辨别的影响。实验结果表明，当声调的起点比较高时，前面的塞音容易听成为不送气音。似乎是，高的基频的作用相当于缩短了VOT。从平均反应百分数差R_o看，基频曲线的起点的影响是相当大的。从反应曲线粗略地估计一下，一、二声引起的b、p音位边界差约为15ms，三、四声引起的d、t音位边界差约为7 ms。

与四声声调复盖的频率范围和送气音与不送气音的VOT变化范围相比，塞音发音方式引起的声调知觉上的变化比声调引起的发音方式辨别的差异要小得多。如果这些变化可以用某种替换关系描述，则VOT与声调和声调与VOT之间的替换关系是不同的。为什么在知觉上有这样的相互作用的不对称，是一个有意思的问题。

讨　论

这两个实验以合成的汉语普通话CV音节作刺激，研究了塞辅音发音方式与声调之间在知觉上的相互影响问题。得到的主要结果是：(1)塞辅音的发音方式，送气与不送气，影响到音节声调的知觉，不送气音b和d使听者在辨别声调时倾向于基频曲线起点高的声调反应(一声和四声)；(2)音节的声调也影响着对塞音发音方式的判断，当VOT线索比较模糊时，听者倾向于将一、四声前的塞音听成不送气音，将二、三声前的听成送气音。

引言中提到，关于送气和不送气音对元音基频的影响，文献中报告的研究结果很不一致。根据本实验得到的知觉上的研究结果和发音活动与知觉之间可能的相互关系，可以推测，汉语普通话的CV音节中，对同一种声调模式，送气音使基频曲线起点降低，不送气音使起点升高。声调不同时，送气音和不送气音的VOT值也可能有所变化。但应当看到，发音与知觉之间的关系是相当复杂的，有必要对音段与声调之间的关系作系统的测量和分析。

参考文献

[1] J. M. Hombert (1978), Consonant types, vowel quality, and tone, in Tone, A Linguistic Servey, Ed by Fromkin, pp. 77—111.

[2] M. Haggard, S. Ambler, and Mo Callow (1970), Pitch as a voicing cue, Journal of the Acoustical Society of America, vol. 47, pp. 613—617.

[3] Ling King, H. Ramming, L. Schiefer and H. G. Tillmann (1987), Initial FO contours in Shanhai CV-syllables; An interactive function of tone, vowel hight, and place and manner of stop articulation, in Proceedings of 11th International Congress on Phonetic Sciences, vol. 1, pp. 154—156.

[4] E. Zee (1980), The effect of aspiration on the FO of the following vowel in Cantonnese, UCLA working papers in phonetics 49, pp. 90—97.

[5] Yang Shun-an, Xu Yi (1987), A Software system for synthesizing Chinese, in Proceedings of 2nd International Conference on Information Processing of Chinese Language, pp. 405—414.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

STOP CONSONANTS AND TONE PERCEPTION

Yang Yu-fang Jin Ling-juan

Institute of Psychology, Chinese Academy of Sciences, Beijing

Abstract

Two experiments reported here were conducted with Standard Chinese CV-syllables as stimuli to investigate the interinfluence between tone and stop in consonant perception. It is shown that the manner of stop consonant articulation, aspirated or not aspirated, has a significant effect on tone perception, that is, unaspirated stops make listeners more responsive to tones with higher onset F_0 foo and that the tone of a syllble strongly affects the identification of stop manners. When the cue of VOT is ambiguous for judging higher onset F_0 of tone one and tone four, it leads to more responses of unaspirated stops.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net