

正常成人低通滤波汉语清塞音的识别研究<sup>△</sup>张梦超<sup>1,2</sup> 刘巧云<sup>1,2</sup> 卢海丹<sup>1</sup> 黄昭鸣<sup>1,2</sup> 杜晓新<sup>1,2</sup> 孙喜斌<sup>2,3</sup>

**【摘要】** 目的 获得健听成人识别低通滤波汉语清塞音所需的最低截止频率及识别的影响因素。方法 以汉语清塞音与韵母/a/组成的/ba/、/pa/、/da/、/ta/、/ga/、/ka/单音节为言语测试材料,步长为 200 Hz,通过上升法测试 33 名健听成人不同截止频率时的识别情况,获得正确识别的最低截止频率,并对结果进行方差分析。结果 ①33 名健听成人识别/ba/、/da/、/ga/、/ta/、/ka/所需的截止频率在 3 000 Hz 以下,分别为 2 239.39、2 936.36、2 084.55、2 372.73、1 693.94 Hz,识别/pa/需 7 000 Hz 左右的截止频率,为 6 784.85 Hz。②识别清塞音所需的低通截止频率受到送气方式和发音部位的影响,但不受受试者性别影响。结论 识别低通滤波汉语清塞音/pa/所需的最低截止频率均值高于 6 000 Hz,其余清塞音的最低截止频率均值均在 3 000 Hz 以下;识别清塞音的截止频率由低到高依次为/kɑ/、/gɑ/、/bɑ/、/tɑ/、/dɑ/、/pɑ/;发声的送气方式和发音部位影响低通截止频率。

**【关键词】** 低通滤波; 清塞音; 言语识别

DOI:10.3969/j.issn.1006-7299.2012.02.003

网络出版时间:2012-3-7 19:23

网络出版地址: <http://www.cnki.net/kcms/detail/42.1391.R.20120307.1923.020.html>

**【中图分类号】** H018.4 **【文献标识码】** A **【文章编号】** 1006-7299(2012)02-0105-03

Recognition of Low-pass Filtered Voiceless Stop Consonants in Chinese  
Mandarin in Normal-hearing Adults

Zhang Mengchao\*, Liu Qiaoyun, Lu Haidan, Huang Zhaoming, Du Xiaoxin, Sun Xibin

[\* Key Laboratory of Speech and Hearing Sciences(East China Normal University), Ministry of Education, Shanghai, 200062, China)

**【Abstract】** Objective We aim to get the lowest cut-off frequencies for recognizing low-pass filtered voiceless stop consonants in normal-hearing adults and the factors affecting recognition. Methods Test materials consisted of six consonant-vowel syllables as /pa/, /ta/, /ka/, /pa/, /ta/, /ka/. By ascending the cut-off frequency, we got the proportion of correct answers on each frequency and the lowest cut-off frequency for the best performance as well. Repeated measure was used to analyze the significance among the levels of factors. Results ①3 000 Hz or lower was required for /pa/, /ta/, /ka/, /ta/, /ka/ recognition and higher than 7 000 Hz for /pa/ in 33 normal-hearing adults. ② Frequencies necessary for low-pass filtered voiceless stop consonants recognition was affected by methods of aspiration and place of articulation, but not affected by gender. Conclusion We get the lowest cut-off frequencies for recognizing /pa/, /ta/, /ka/, /pa/, /ta/, /ka/ and provide reference for low-pass filtered monosyllable test.

**【Key words】** Low-pass filter; Voiceless stop consonant; Speech recognition

言语信号经过滤波后,外在冗余度下降,随着截止频率降低,个体需要动用更多听处理能力提取其声学特征。识别时所需的声学特征越少,说明个体的听处理能力越强。早在 1954 年, Bocca 等首先使用低冗余度的滤波言语作为听处理障碍患者的测试

材料<sup>[1]</sup>, 并发现病变对侧耳的滤波言语识别得分降低。随后几十年间,滤波材料开始被应用于中枢病变的诊断。上世纪 80 年代,程锦元等编制了一部分低通滤波言语测听词表,并在中低频区域划分了 7 档截止频率,有 350、500、750、1 100、1 600、2 500 和 3 500 Hz<sup>[2]</sup>。2006 年苏金霏等的研究发现,500~8 000 Hz 是健听成人言语识别的重要频带,单音节词是外在冗余度最少的语音单位,比有义词更难听懂,言语测听最好使用单音节词<sup>[3]</sup>。为探讨识别低通滤波汉语到底需要多大的通带范围及不同声韵母会受到何种因素的影响,本研究选取汉语声母中

<sup>△</sup>上海市哲社规划青年基金资助项目(2009EJY002)、言语听觉科学教育部重点实验室开放课题项目资助(KFZR06201014)

1 华东师范大学言语听觉科学教育部重点实验室(上海 200062);  
2 华东师范大学言语听觉康复科学系; 3 中国聋儿康复研究中心  
作者简介:张梦超,女,上海人,在读硕士,主要从事言语听觉康复研究。  
通讯作者:张梦超 (Email:lafela0716@126.com)

的塞音,将塞音开头的单音节词作为言语识别材料,通过改变通带范围的方式,获得健听成人识别低通滤波清塞音所需的最小通带及影响因素,为进一步完善听处理相关的言语测听材料作前期探索,报告如下。

## 1 资料与方法

**1.1 研究对象** 本实验对象为华东师范大学 33 名在读大学生,其中男 17 人,女 16 人,年龄 19~24 岁,平均  $21.5 \pm 1.2$  岁,双耳各频率听阈均在 25 dB HL 以下。

## 1.2 研究方法

**1.2.1 测试材料** 本实验选用汉语清塞音与韵母 /a/ 组合的单音节(表 1)为言语测试材料,声调均为一声,实验材料由一名 25 岁的成年女性用标准普通话朗读,平均言语基频为 236.67 Hz。在隔声室内用 Sony Soundforge 7.0 软件单声道录制声音材料,采样频率为 44 100 Hz,录制后对材料进行音量归一化。采用言语听觉科学教育部重点实验室开发的 Lab-View 平台下的低通滤波语音识别软件,由一根音频线连接至 GSI 61 听力计,通过听力计控制输出音量在 65 dB SPL,通过 TDH 50 压耳式耳机播放声音。

表 1 实验测试音节

送气方式	发音部位		
	双唇	舌尖	舌根
不送气	/ba/	/da/	/ga/
送气	/pa/	/ta/	/ka/

**1.2.2 测试程序** 首先选择 5 名在读大学生进行预实验,采用 100 阶巴特沃斯低通滤波器,设置起始截止频率为 300 Hz,受试者在听某一音节时,若报告错误,则以 100 Hz 的步长上升,直至受试者在连

表 2 不同性别受试者低通滤波汉语清塞音识别的最低截止频率(Hz,  $\bar{x} \pm s$ )

性别	/ba/	/da/	/ga/	/pa/	/ta/	/ka/
男	2 311.76 $\pm$ 904.08	3 017.65 $\pm$ 748.53	2 081.76 $\pm$ 477.39	7 241.18 $\pm$ 1 403.15	2 405.88 $\pm$ 490.50	1 723.53 $\pm$ 272.79
女	2 162.50 $\pm$ 869.39	2 850.00 $\pm$ 721.11	2 087.50 $\pm$ 904.71	6 300.00 $\pm$ 1 025.02	2 337.50 $\pm$ 255.28	1 662.50 $\pm$ 233.45
总体	2 239.39 $\pm$ 876.76	2 936.36 $\pm$ 728.79	2 084.55 $\pm$ 705.43	6 784.85 $\pm$ 1 305.79	2 372.73 $\pm$ 389.93	1 693.94 $\pm$ 252.41

对获得的截止频率进行一元方差分析,发现送气方式和发音部位的主效应均极显著,说明送气方式、发音部位影响滤波塞音的识别;受试者性别的主效应不显著。送气方式、发音部位的交互作用极显著;其余的交互效应均不显著(表 3)。对发音部位进行多重比较后,发现双唇音与舌尖音( $P < 0.01$ )、双唇音与舌根音( $P < 0.01$ )、舌尖音与舌根音( $P < 0.01$ )之间差异均有显著统计学意义。

在同一发音部位上,双唇送气音/pa/与不送气音/ba/、舌尖送气音/ta/与不送气音/da/、舌根送气音/ka/与不送气音/ga/低通滤波的最低截止频率差异均有显著统计学意义。其中,双唇送气塞音/pa/

续增长的十个步长中均正确识别。要求受试者口头报告所听到的音节,报告完全正确记“1”,报告错误记为“0”,并由测试者在软件中输入受试者实报的音节作为错误走向。通过预实验,研究者在正式实验中作了以下调整:①预实验中当截止频率为 500 Hz 以下时,受试者均表示无法识别,因此在正式实验中提高了截止频率;②预实验中步长 200 Hz 的结果与 100 Hz 的情况基本一致,因此在正式实验中调整了步长,缩短了实验时间;③排除了实验操作中可能存在的硬件、软件问题。

正式实验采用 100 阶巴特沃斯低通滤波器,设置起始截止频率为 500 Hz,步长为 200 Hz,受试者连续正确报告 5 次时停止放音,并记录 5 次连续正确报告中的最低截止频率,作为该受试者识别该音节的截止频率,其余流程与预实验基本一致。整个实验在隔声室内完成,本底噪声低于 45 dB SPL。

**1.3 统计学方法** 本实验用 Microsoft Office Excel 2007 进行数据记录,用 SPSS 16.0 作数据分析处理。采用  $2 \times 3 \times 3$  的三因素混合实验设计,受试者间变量为性别,有男、女两个水平;受试者内变量为送气方式和发音部位,其中送气方式有送气和不送气两个水平,发音部位有双唇、舌尖、舌根三个水平。因变量为低通滤波塞音正确识别的最低截止频率。

## 2 结果

清塞音音节的平均截止频率见表 2。可见,识别/pa/所需的最低截止频率均值高于 6 000 Hz,其余清塞音的最低截止频率均值均在 3 000 Hz 以下。识别清塞音的截止频率由低到高依次为/ka/、/ga/、/ba/、/ta/、/da/、/pa/。

比双唇不送气塞音/ba/的均值高,但舌尖音、舌根音的情况下,都是不送气塞音比送气塞音的均值高(表 4)。

表 3 送气方式、发音部位和性别对低通滤波汉语清塞音识别的影响

影响因素	F	P
送气方式	30.429	0.000
发音部位	190.1	0.000
性别	3.378	0.076
送气方式 $\times$ 性别	1.813	0.188
发音部位 $\times$ 性别	2.008	0.143
送气方式 $\times$ 发音部位	246.827	0.000
送气方式 $\times$ 发音部位 $\times$ 性别	1.655	0.199

表 4 同一发音部位不同送气方式的影响

发音部位	差值(Hz)*	P
双唇	-4 533.46	0.000
舌尖	562.13	0.000
舌根	391.62	0.004

注: \*差值=不送气塞音最低截止频率均值-送气塞音均值

在不送气塞音水平上,双唇音/ba/与舌尖音/da/、舌尖音/da/与舌根音/ga/低通滤波的最低截止频率的差异有极显著统计学意义,双唇音/ba/与舌根音/ga/的差异无统计学意义。在送气塞音水平上,不同发音部位间两两差异均有显著统计学意义(表 5)。

表 5 同一送气方式不同发音部位的影响

送气方式	发音部位	差值(Hz)	P	
不送气	双唇	舌尖	-696.69	0.004
		舌根	152.50	0.807
	舌尖	舌根	849.19	0.000
送气	双唇	舌尖	4 398.90	0.000
		舌根	5 077.57	0.000
	舌尖	舌根	678.68	0.000

### 3 讨论

本研究发现,影响清塞音识别的低通截止频率的第一个主要因素是发音部位:当送气方式相同时,除了/ba/与/ga/之间差异没有统计学意义外,其余不同部位之间差异均有显著统计学意义,说明发音部位影响了大部分清塞音识别的低通截止频率。第二个主要影响因素是送气方式:当两个塞音的发音部位相同时,识别送气音与不送气音的截止频率差异均存在显著统计学意义。此外,发音部位和送气方式之间还存在交互效应。舌相关的塞音(舌尖音、舌根音)中,不送气音的截止频率高于送气音,唇相关的塞音(双唇音)中,送气音的截止频率高于不送气音。

从文中结果看,识别/ka/、/ga/、/ba/、/ta/、/da/的最低截止频率均处于低中频区域,说明识别/ka/、/ga/、/ba/、/ta/、/da/发音部位和送气方式的声学信息基本集中在低中频区域。以/ka/为例,其截止频率均值约为 1 700 Hz,说明截止频率 1 700 Hz 的/ka/包含了舌根送气音的声学线索,1 700 Hz 以上的声学信息就可作为冗余信息。

/pa/的识别情况比较特殊,与其他清塞音比较其截止频率在高频范围。从声学分析的角度来说,识别塞音发音部位可依靠辅音共振峰判断<sup>[4]</sup>,双唇音/b/、/p/的第二共振峰在 1 100~1 500 Hz 左右的低频部分<sup>[4,5]</sup>。结合本研究结果,发现/pa/在 1 500 Hz 时识别率曾达到 50%,而在 2 300~5 700

Hz 的频段之间,/pa/大量误听为/ta/,错误率超过 70%,直至截止频率达 7 000 Hz 以上时,/pa/的识别率才达到 70%以上。吴宗济发现/pa/中的/p/除了在中低频区有一共振峰<sup>[5]</sup>,还有一个 7 kHz 左右的辅音共振峰,本实验也从另一方面证实了这点。从听觉感知的角度来说,曾有研究提示低通截止频率为 3 500 Hz 时言语基本能被识别<sup>[2]</sup>,但该类实验采用了阶数较低的低通滤波器,过渡带中遗留大量声学信息助于识别。本实验采用高阶滤波器,排除了过渡带的影响。结合两者可知,声学分析与听觉感知之间存在着一定差异,/pa/中/p/的发音部位体现在其高频的辅音共振峰。

男、女受试者识别低通滤波的最低截止频率没有差异,说明性别不影响低通滤波清塞音识别,在今后的实验中可以不考虑男女受试者比例。

综上所述,健听成人识别/ba/、/da/、/ga/、/ta/、/ka/所需的截止频率在 3 000 Hz 以下,识别/pa/需 7 000 Hz 以上。识别清塞音所需的低通截止频率受送气方式和发音部位的影响,但不受性别影响。同一部位不同送气方式时,舌塞音中不送气音的截止频率高于送气音,唇塞音中送气音的截止频率高于不送气音。采用高阶的低通滤波器可以降低过渡带的冗余信息,更符合中枢听觉处理评估的低冗余信息测试材料。本实验获得识别/ba/、/da/、/ga/、/ta/、/ka/的最小截止频率,能为选择低通滤波单音节词的通带范围提供参考。此外,听觉感知的机理与声学分析之间的差异也值得进一步探索,在后续研究中将对不同韵母类型的情况及其他汉语声母进行研究,并加入单、双耳的比较,为进一步完善滤波言语测听提供依据。

### 4 参考文献

- 1 Bocca E, Calearo C, Cassinari V. A new method for testing hearing in temporal lobe tumors: preliminary report[J]. Acta Otolaryngo, 1954, 44: 9.
- 2 程锦元, 胡逢昶, 庞慧娟, 等. 滤波汉语清晰度的初步介绍[J]. 中华耳鼻咽喉科杂志, 1983, 18: 138.
- 3 苏金霏, 王宁宇, 张秋航, 等. 测试材料的滤波通带对言语识别率的影响[J]. 中国实验诊断学, 2006, 10: 750.
- 4 曾进兴. 语言病理学基础(第二卷)[M]. 台北: 心理出版社, 1996. 48~51.
- 5 吴宗济, 林茂灿. 实验语音学概要[M]. 北京: 高等教育出版社, 1989. 127~129.

(2011-08-02 收稿)

(本文编辑 李翠娥)