

# 探析元音与声强对嗓音声学共振峰检测结果的影响\*

谢冰<sup>1</sup> 李哲<sup>2</sup> 王红星<sup>2</sup> 匡栩源<sup>2</sup> 倪伟<sup>2</sup> 钟润琪<sup>2</sup> 李艳<sup>2</sup>

**[摘要]** 目的:本研究旨在探析元音与声强对共振峰的影响,为声学检测中声音样本和发声方式的选择提供参考。方法:招募年龄为 19~24 岁健康受试者 38 例,男、女各 19 例。在不同元音(/a/、/æ/、/i/和/u/)以及不同声强(最低音、舒适音、真声最高音和假声最高音)下的共振峰进行差异性分析,并对差异显著的组间进行两两比较。结果:①在第一共振峰中元音/a/、/æ/大于/i/和/u/,第二共振峰中/i/最大。第一共振峰的最小值是/i/的最低音,最大值是/a/的最高音。②第一共振峰中胸音区随着声强增加而增大,第二共振峰进入假声最高音明显下降。结论:不同元音与声强共振峰分布各异,即元音和声强对共振峰产生不同程度的影响。根据第一共振峰的极值初步明确其最大正常范围,有助于完善声学检测。

**[关键词]** 元音;声强;嗓音声学分析;共振峰检测

**DOI:**10.13201/j.issn.2096-7993.2024.12.011

**[中图分类号]** R767.92 **[文献标志码]** A

## The influence of vowel and sound intensity on the results of voice acoustic formant detection was analyzed

XIE Bing<sup>1</sup> LI Zhe<sup>2</sup> WANG Hongxing<sup>2</sup> KUANG Xuyuan<sup>2</sup> NI Wei<sup>2</sup>  
ZHONG Runqi<sup>2</sup> LI Yan<sup>2</sup>

(<sup>1</sup>Department of Rehabilitation Medicine, The First Hospital of Changsha, [The Affiliated Changsha Hospital of Xiangya School of Medicine, Central South University], Changsha, 410005, China; <sup>2</sup>Department of Rehabilitation Medicine, Xiangya Hospital, Central South University)

Corresponding author: LI Yan, E-mail: kangfuyiban62@163.com

**Abstract Objective:** This study aims to explore the influence of vowels and sound intensity on formant, so as to provide reference for the selection of sound samples and vocal methods in acoustic detection. **Methods:** Thirty-eight healthy subjects, 19 male and 19 female, aged 19-24 years old were recruited. The formants of different vowels(/a/, /æ/, /i/ and /u/) and different sound intensities(lowest sound, comfort sound, highest true sound and highest falsetto sound) were analyzed, and pairings were compared between groups with significant differences. **Results:** ①The vowels /a/ and /æ/ in the first formant were larger than /i/ and /u/, and /i/ was the largest in the second formant. The minimum value of the first formant is the lowest sound of /i/ and the maximum is the highest sound of /a/. ②In the first formant, the chest sound area increases with the increase of sound intensity, while the second formant enters the highest falsetto and decreases significantly. **Conclusion:** Different vowels and sound intensity have different distribution of formant, that is, vowel and sound intensity have different degree of influence on formant. According to the extreme value of the first formant, the maximum normal range is determined initially, which is helpful to improve the acoustic detection.

**Key words** vowel; sound intensity; voice acoustic analysis; formant detection

嗓音障碍是多数咽喉部疾病的常见主诉和早期表现。调查显示 3%~9% 人群患有不同程度的嗓音障碍<sup>[1]</sup>。嗓音评估能判定嗓音障碍的类型和程度,也是制定及动态调整治疗方案的前提和基

础。嗓音声学检测具有简便无创、能提供定量数据等优点,是临床评估嗓音障碍的主要方法之一。声学检测所提供的基频、共振峰、微扰以及谐波比等数据多方面反映声音的特性。共振峰作为腔体共鸣参数,能反映发声时舌和腔体的位置、大小和形状,较大程度地影响言语音色和清晰度,是目前声学研究的热点。有研究证实,共振峰检测在鼾症、唇腭裂以及帕金森等疾病诊疗中体现出较高的准确性和灵敏度<sup>[2-3]</sup>。

\*基金项目:湖南省自然科学基金面上项目(No: 2023JJ30945)

<sup>1</sup>中南大学湘雅医学院附属长沙医院(长沙市第一医院)康复医学科(长沙,410005)

<sup>2</sup>中南大学湘雅医院康复医学科

通信作者:李艳, E-mail: kangfuyiban62@163.com

引用本文:谢冰,李哲,王红星,等.探析元音与声强对嗓音声学共振峰检测结果的影响[J].临床耳鼻咽喉头颈外科杂志,2024,38(12):1149-1153. DOI:10.13201/j.issn.2096-7993.2024.12.011.

既往关于共振峰检测的研究多聚焦于病理嗓音与正常嗓音声学上的差异性比较。由于不同的语音具有各自的共振峰,因而不同的检测语料会导致评估结果出现偏差<sup>[4]</sup>。声学检测中的语料包括元音和连续语音。其中元音具有稳定、持续、易获取准确数据等特性,是多数声学研究的基本选项<sup>[5]</sup>。目前临床声学检测在具体元音类型和声强选择以及语料对共振峰影响方面尚未形成共识。基于此,本研究从元音和声强 2 个维度,探索对共振峰检测结果产生影响的语料因素,进而为临床选择适用于嗓音共振峰检测的语料提供一定参考。

## 1 资料与方法

### 1.1 临床资料

受试者来源于 2022 年 9 月—2022 年 12 月在中南大学湘雅医院康复科实习的健康成人。对其进行嗓音声学检测,筛选出 38 例受试者(男、女各 19 例)所发出的 608 个声音样本纳入研究。受试者年龄 19~24 岁,该年龄段变声期结束,受到解剖结构改变或激素水平变化等干扰因素较少,声音较稳定。本研究获得中南大学湘雅医院临床伦理委员会批准(No:202206135)。

### 1.2 纳入及排除标准

纳入标准:①以汉语为母语环境;②听力、语言和认知功能正常;③年龄 19~24 岁。排除标准:①嗓音障碍指数数量表>0 或自述有嗓音不适;②既往有口腔或喉结构病变或头颈部手术;③受过专业声乐或器乐演奏训练;④测试前 24 h 内有吸烟或饮酒;⑤近 1 周患有上呼吸道疾病;⑥有哮喘,有反流或其他影响声音的慢性疾病;⑦近期服用激素可能对嗓音产生影响的药物。

### 1.3 研究方法

**1.3.1 嗓音声学检测语料** 本研究采用 4×4 双因素(元音与声强)重复测量的方法同时探索元音和声强对声学检测的影响。选择汉语中稳定性强、最具代表性的元音/a/、/i/和/u/,再纳入元音/æ/作为相应对照。声强则是对最低音、舒适音、真声最高音以及假声最高音进行全音域的研究。

**1.3.2 声音的录制与数据采集** 检查在室内温度 27℃左右的安静房间内(<40 dB)进行。采用上海泰亿格公司的 Dr. Speech Since for Windows 的 Voice Assessment 软件进行语音录制及声学分析。受试者坐位,麦克风采用 45°离轴定位距离嘴唇 10~15 cm。引导受试者在气息平稳状态下发声出所需的 16 个声音 3~5 s。从各声音样本中提取平稳段(≥1 s)采样频率为 4 400 Hz,经前置放大器将嗓音录入计算机,进行频谱分析后,经快速傅里叶转换计算出共振峰的具体数值。测试过程中如出现错误则重新录制,直至最终完成。最终对 38 例受试者所发的 608 个声音样本中提取第一共

振峰 Formant 1(F1)和第二共振峰 Formant 2(F2)数据进行分析,具体流程见图 1。

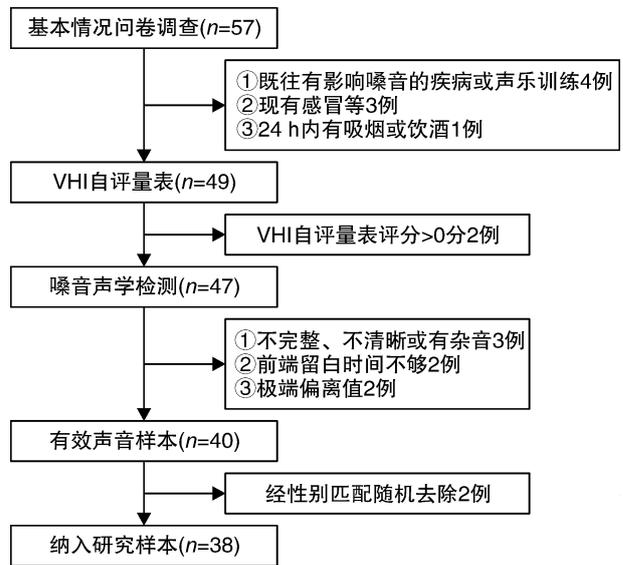


图 1 声音样本获取流程

### 1.4 统计学方法

采用 SPSS 26.0 统计学软件对纳入研究的受试者的年龄、身高和体重等一般情况做统计描述,根据数据分布情况分别用  $\bar{X} \pm S$  和  $M(P_{25}, P_{75})$  描述。为探索元音和声强对声学的影响,分别以相同元音不同声强、相同声强不同元音分为 2 组进行比较。首先对嗓音声学参数结果分别进行正态性检验与方差齐性检验,若检验通过使用 ANOVA 方差分析,若是非正态分布多个相关样本,则使用非参数 Friedman 检验进行差异性比较,然后对差异有统计学意义的组间进行两两比较,并使用绘图软件 GraphPad Prism 9 绘制共振峰在不同元音、强度中的分布特点和变化趋势, $P < 0.05$  表示差异有统计学意义。

## 2 结果

第一共振峰相同元音不同声强和相同声强不同元音的差异性比较,分别见表 1、2,元音的类型和声强均对 F1 有显著影响( $P < 0.05$ )。

### 2.1 元音与声强对 F1 的影响

**2.1.1 元音对 F1 的影响** 元音在 4 个声强中均对 F1 中有显著影响( $P < 0.05$ ),见表 1。并且整体趋势上在胸音发声区高元音/i/、/u/要低于低元音/a/、/æ/,见图 2,通过两两比较,差异有统计学意义( $P < 0.05$ )。

**2.1.2 相同元音、不同声强 F1 比较** 声强在 4 个元音中对 F1 均有显著差异( $P < 0.05$ ),/i/和/u/随着强度增加而 F1 值增加,而/a/、/æ/则在假声最高音出现回落,见表 2、图 2。通过两两比较,差异有统计学意义( $P < 0.05$ )。

表 1 不同元音在各声强 F1 分布及差异性分析

声强	元音	$M(P_{25} \sim P_{75})$	$\chi^2$	$P$
最低音	/a/	543(249~771)	11.133	<0.05
	/æ/	620(296~759)		
	/i/	295(276~340)		
	/u/	363(330~413)		
舒适音	/a/	749(604~996)	51.301	<0.05
	/æ/	625(469~877)		
	/i/	307(277~350)		
	/u/	378(346~408)		
真声最高音	/a/	782(698~922)	70.263	<0.05
	/æ/	706(584~855)		
	/i/	365(327~417)		
	/u/	401(359~455)		
假声最高音	/a/	596(491~674)	19.202	<0.05
	/æ/	584(501~647)		
	/i/	541(454~643)		
	/u/	497(474~563)		

最低音:/i/ < /æ/; 舒适音、真声最高音:/i/ < /a/、/æ/、/u/ < /a/、/æ/; 真声最高音:/i/、/u/ < /a/、/æ/; 假声最高音:/u/ < /a/、/æ/

表 2 不同声强在各元音 F1 分布及差异性分析

元音	声强	$M(P_{25} \sim P_{75})$	$\chi^2$	$P$
/a/	最低音	543(249~771)	19.168	<0.05
	舒适音	749(604~996)		
	真声最高音	782(698~922)		
	假声最高音	596(491~674)		
/æ/	最低音	620(296~759)	16.263	<0.05
	舒适音	625(469~877)		
	真声最高音	706(584~855)		
	假声最高音	584(501~647)		
/i/	最低音	295(276~340)	70.249	<0.05
	舒适音	307(277~350)		
	真声最高音	365(327~417)		
	假声最高音	541(454~643)		
/u/	最低音	363(330~413)	46.050	<0.05
	舒适音	378(346~408)		
	真声最高音	401(359~454)		
	假声最高音	497(474~563)		

/a/最低音 < 舒适音、真声最高音,假声最高音 < 真声最高音;/æ/最低音 < 真声最高音,假声最高音 < 真声最高音;/i/最低音 < 真、假声最高音,舒适音、真声最高音 < 假声最高音;/u/最低音、舒适音、真声最高音 < 假声最高音。

2.2 元音与声强对 F2 的影响

2.2.1 元音对 F2 的影响 元音在 4 个声强中均对 F1 有显著影响 ( $P < 0.05$ ),且/i/在各个强度中均显著高于其他元音,除最低音差异不显著外,/u/都是最小的,/a/、/æ/则位于二者之间,见表 3、图 3。通过两两比较,差异有统计学意义 ( $P < 0.05$ )。

2.2.2 相同元音、不同声强 F2 比较 声强对/a/、/æ/和/i/的 F2 值影响显著 ( $P < 0.05$ ),并且这 3 个元音随着声强的增加,但同时又都在假声最高音中都出现了回落,见表 4、图 3。通过两两比较,差异有统计学意义 ( $P < 0.05$ )。

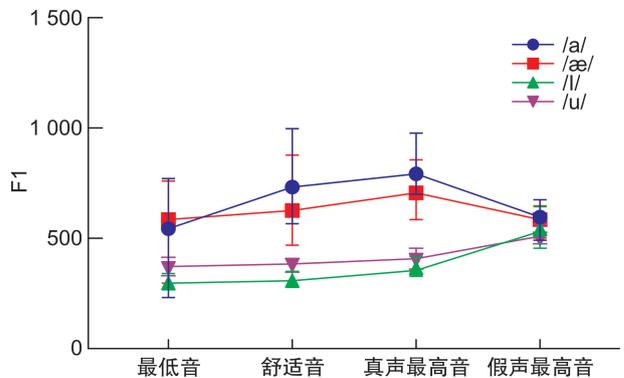


图 2 元音在不同声强第一共振峰随声强变化趋势

表 3 不同元音在各声强 F2 分布及差异性分析

声强	元音	$M(P_{25} \sim P_{75})$	$\chi^2$	$P$
最低音	/a/	1 029(935~1 312)	27.884	<0.05
	/æ/	1 613(1 320~1 955)		
	/i/	2 229(1 933~2 596)		
	/u/	946(835~1 610)		
舒适音	/a/	1 567(1 121~2 095)	32.305	<0.05
	/æ/	1 752(1 446~1 970)		
	/i/	2 132(1 924~2 799)		
	/u/	883(773~1 466)		
真声最高音	/a/	1 263(1 129~1 396)	51.000	<0.05
	/æ/	1 645(1 409~1 892)		
	/i/	2 207(1 997~2 676)		
	/u/	861(798~1 060)		
假声最高音	/a/	1 140(983~1 330)	33.697	<0.05
	/æ/	1 139(1 024~1 284)		
	/i/	1 557(1 129~2 109)		
	/u/	985(893~1 115)		

最低音:/a/、/æ/、/u/ < /i/; 舒适音:/u/ < /a/、/i/、/æ/ < /i/; 真声最高音:/u/ < /a/、/æ/ < /i/; 假声最高音:/u/ < /a/、/æ/、/i/、/æ/ < /i/。

2.3 F1-F2 散点图分布

以第一共振峰为横轴,第二共振峰为纵轴将所有数据纳入得到第一共振峰和第二共振峰散点图。把图的高度分成第一共振峰和第二共振峰对应的分区,根据第一共振峰和第二共振峰频率的值,绘制包含全部受试者产生跨强度的各个元音的空间示图,每个数据点代表不同音调中每个元音的第一共振峰和第二共振峰值,普通话的元音空间类似于三角形。见图 4。

表 4 不同声强在各元音 F2 分布及差异性分析

元音	声强	M(P <sub>25</sub> ~P <sub>75</sub> )	χ <sup>2</sup>	P
/a/	最低音	1 029(935~1 312)	18.325	<0.05
	舒适音	1 567(1 121~2 095)		
	真声最高音	1 262(1 129~1 396)		
	假声最高音	1 140(983~1 330)		
/æ/	最低音	1 613(1 320~1 955)	16.583	<0.05
	舒适音	1 752(1 446~1 970)		
	真声最高音	1 645(1 409~1 891)		
	假声最高音	1 138(1 024~1 284)		
/i/	最低音	2 229(1 933~2 596)	11.211	<0.05
	舒适音	2 132(1 924~2 799)		
	真声最高音	2 207(1 997~2 676)		
	假声最高音	1 557(1 129~2 109)		
/u/	最低音	947(835~1 610)	/	>0.05
	舒适音	883(773~1 466)		
	真声最高音	861(798~1 060)		
	假声最高音	985(893~1 115)		

/a/:最低音<舒适音,真声最高音<舒适音;/æ/:假声最高音<舒适音、真声最高音;/i/:假声最高音<真声最高音。

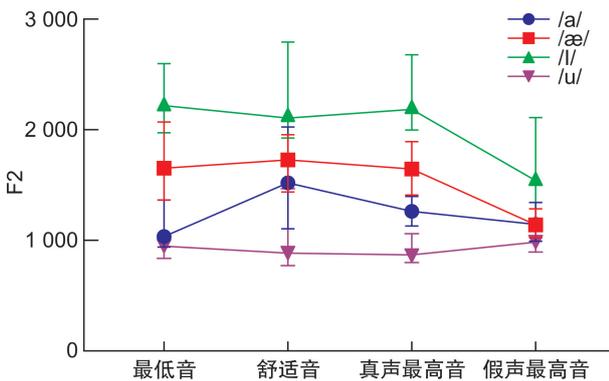


图 3 元音在不同声强第二共振峰随声强变化趋势

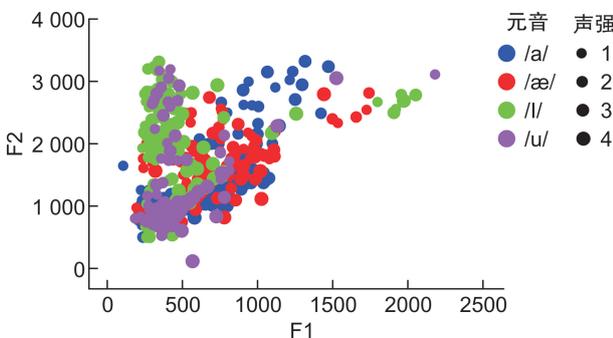


图 4 第一共振峰和第二共振峰散点图

### 3 讨论

声带振动产生的信号通过喉腔、咽腔传递至口腔和鼻腔。由于声道形状和大小因人而异,在此过程中某些声音被放大而形成共振,声道共振产生共

振峰。共振峰能实时定量反映发声过程中共鸣的相关信息,广泛运用于口鼻和咽腔的生理和病理评估<sup>[6]</sup>。现有研究多侧重于病理与正常的共振峰对比,对于具体的评价方法和异常区间值的界定目前尚未确定。本研究采用随机区组单因素分析,创新从元音和声强两个维度分析对共振峰的影响,并结合常用端点元音和全音域的共振峰检测,较全面细致体现不同元音和声强共振峰的分布及差异性比较,以此探寻适用于共振峰检验的语料。

#### 3.1 元音对共振峰的影响

研究表明对共振峰影响最大的是唇形、前下颌骨和舌体的调整,其次是软腭的变化。相比之下,后下颌运动和舌尖被认为影响较小<sup>[7]</sup>。本研究基于汉语语系,选择元音中占比最大、最具代表性和稳定性强的/a/、/i/和/u/,这3个元音属于声学空间中的框架元音,涵盖了舌前后、下颌前上下等构音器官最大运动范围<sup>[8]</sup>。此外纳入元音/æ/作为参考比较,/æ/为喉低位音,声带和声道交互作用最小。结果显示:在全音域中4个元音的F1和F2都有显著差异,说明元音的类型对F1和F2均有显著影响。

研究认为F1反映舌头的垂直运动,而F2反映舌头的前后运动。元音/i/发声时随着舌尖前伸直,舌根上抬,扩大了F1共振的咽部空间,F1值随之减少<sup>[9]</sup>。另一方面,鼻窦部的上颌窦和蝶窦有吸收声音的作用。研究表明舌位较高、较前的元音鼻流量较大,健康成人发元音/i/时的鼻流量最高<sup>[10]</sup>。因此,F1在胸声发声区都是/i/最低。本研究在假声最高音中F1值出现了上升。这可能与发假声最高音时,咽部肌肉持续收缩以满足更高的声带振动频率,导致咽腔空间缩小有关。而其他元音发声时舌位相对较低,咽部空间相对较小,因而咽部收缩所引起的咽腔体积变化不如/i/显著。

为了达到高元音/i/、/u/所需的舌位最大高度,颊舌肌后纤维收缩从而为前部的舌体提供支撑。这种收缩伴随结果是变宽的咽部与上抬的舌根一起构建了宽大的咽腔<sup>[11]</sup>。因而在F1中,高元音/i/、/u/小于低元音/a/、/æ/。本研究在胸声发声区的F1中都表现出/i/、/u/小于/a/、/æ/的趋势。有学者研究腭咽闭合不全和软腭或硬腭分裂对元音/a/、/i/和/u/的影响,发现对/a/产生的影响最小<sup>[12]</sup>。这也从侧面说明高元音/i/、/u/的鼻腔反共振较/a/明显,故其共振峰值也小于/a/。因此在对唇腭裂患者评估时,低元音/a/和/æ/不适于作为嗓音声学的检测语料。鉴于多数病理性嗓音最早出现在假声最高音,在检测假声最高音时,应考虑到不同元音头声发声区的差别。/i/、/u/在与胸声区比较方便程度和音域下降敏感度方面优于/a/、/æ/。

### 3.2 声强对共振峰的影响

声强是单位时间通过垂直于声音传播方向单位面积的声能量,与声速、声波频率以及振幅有关。日常可通过压缩声带与加强声门下压力获取更有力量声强。临床根据嗓音强度可分为最低音、舒适音、真声最高音和假声最高音。国外研究发现声门下气压和声音强度之间存在线性关系,并随着声门下压的增强,F1 有增加的趋势<sup>[13]</sup>。本研究数据显示声强对各元音的 F1 有显著影响,并且在胸声发声区,随着强度增加,F1 表现出与之相应的增大。由于较高的声强通常意味着下颌开口增大,同时伴随着舌位的下降从而导致 F1 增加,因而声强最小的最低音 F1 值最小。在临床应用共振峰评估声道过程中应考虑声强对 F1 的影响。呼吸弱、音量低的受试者可能产生更低的 F1 值。

有研究证实 F2 数值与舌尖-齿列间的空间相关,空间越小 F2 值越大<sup>[14]</sup>。这与本研究中前元音/i/最大而后元音/u/最小的结果相一致。此外,本研究还发现声强在/a/、/æ/和/i/都对 F2 值有显著影响,且都在假声最高音表现出回落。可能原因是发假声最高音时,咽喉部肌肉及舌肌收缩力度较低音更强大,但齿列未相应发生位移,因而舌尖与齿列间的空间扩大从而导致 F2 值下降。

相关研究发现第一共振峰值越大,下颌越低;第二共振峰值越大,舌肌收缩越小,即舌的位置相对在口腔的前面<sup>[15]</sup>。根据第一共振峰和第二共振峰频率的值绘制的图形空间类似于三角形,这与英语在元音空间图中呈现出四边形有所差别。在英语中/i/是高前元音,/æ/是低前元音,/a 是低背元音,/u/是高背元音,而普通话中元音/æ/几乎交织/a/与/i/中间,可能与普通话中并无/æ/元音,因而在模仿发声过程中易向/a/与/i/转化有关。图中元音/i/第一共振峰最低,第二共振峰最大。这是由于发/i/时舌尖前伸直,舌根上抬,口腔前部变小而口咽部比其他元音更长。

综上所述,本研究初步探析了元音和声强都会对共振峰产生不同程度影响。为临床声学检测中元音类型和发声方式的选择提供部分依据。结果显示/i/的最低音和/a/的最高音分别为 F1 的最小值和最大值,因而明确 F1 的最大正常范围,有助于根据 F1 值检测出疾病状态下的舌肌运动受限程度,利于疾病早期诊断。本研究样本量相对较少,此外病理性嗓音中不同元音和声强的特征性表现有待进一步研究。

**利益冲突** 所有作者均声明不存在利益冲突

### 参考文献

- [1] Selvakumari SN, V. Radha. Voice Pathology Identification: A Survey on Voice Disorder[J]. IJEM, 2017, 7(2): 39-49.
- [2] 张舒,徐洁洁. 鼻、咽发音共鸣器官疾病与嗓音障碍[J]. 听力学及言语疾病杂志, 2014, 22(3): 332-335.
- [3] 唐静,陈林丽,杨荆生,等. 元音发音障碍作为帕金森病早期标志的评估[J]. 中国神经免疫学和神经病学杂志, 2023, 30(2): 110-114.
- [4] Belean B. Comparison of formant detection methods used in speech processing applications[J]. AIP Conf Proceedings, 2013, 1565(1): 85.
- [5] 赵学彬. 声学、生理框架下的汉语基础元音发展研究[D]. 北京:首都师范大学, 2009.
- [6] 韩仲明. 嗓音声学检测分析[J]. 中国耳鼻咽喉头颈外科, 2006, 13(5): 351-353.
- [7] Graf S, Schwiebacher J, Richter L, et al. Adjustment of Vocal Tract Shape via Biofeedback: Influence on Vowels[J]. J Voice, 2020, 34(3): 335-345.
- [8] Yang J, Fox A R. Acoustic development of vowel production in native Mandarin-speaking children [J]. Journal of the International Phonetic Association, 2017, 49(1): 33-51.
- [9] Dromey C, Jang G, Hollis K. Assessing correlations between lingual movements and formants[J]. Speech Communication, 2013, 55(2): 315-328.
- [10] Ha-kyung K, 段弘艳, 惠芬芬, 等. 正常成人发不同元音时口腔压力和鼻流量研究[J]. 听力学及言语疾病杂志, 2016, 24(2): 126-129.
- [11] LaCross A, Takano S, Kidder JI, et al. Single motor unit activity in the genioglossus muscle during vowel articulation[J]. J Acoust Soc Am, 2012, 132(3): 2087.
- [12] Vampola, Tomas, Horacek, et al. Modeling the Influence of Piriform Sinuses and Valleculae on the Vocal Tract Resonances and Antiresonances [J]. Acta Acustica, 2015, 101(3): 594-602.
- [13] Plant RL. Aerodynamics of the human larynx during vocal fold vibration[J]. Laryngoscope, 2005, 115(12): 2087-2100.
- [14] 万萍, 黄昭鸣, 郑钦. 中国人嗓音共鸣功能的声学基础研究[J]. 临床耳鼻咽喉头颈外科杂志, 2010, 24(6): 250-252.
- [15] 王臻旒, 陈旸, 吴民华, 等. 成人普通话元音的声学分析(英文)[J]. 康复学报, 2020, 30(3): 183-191.

(收稿日期: 2024-02-21)