语音实验分析的两个基本问题[1]

王 迈[2]

摘要

归一化处理是语音基频分析的基础，原始值经过频率的平移、频域的压扩而进入标准空间，其中频率的对数处理和频域上下限的确定方法是归一化效果的决定因素。汉语音节的声母无论是否带音，都不承载声调信息；辅音韵尾则是重要的声调承载段。共振峰的分布同元音的舌位有密切关系，其中第一共振峰（F1）的频率同舌位的高低成反向关系，第二共振峰（F2）的频率同舌位的前后成正向关系。汉语的一对舌尖元音的第一共振峰（F1）基本一致，平舌元音[ɿ]的F2与F3相距靠近，卷舌元音[ʅ] 的F2与F3相距较远。

关键词： 语音实验分析；基频归一化；声调；共振峰；舌位

---

[1] 2012年度中国教育部人文社科规划项目“以自主学习为特征的外国学生汉语语音训练研究”，编号：12YJAZH008。2012年度上海市教育委科研创新重点项目“基于语料库的外国学生汉语语音训练研究”，编号：12ZS063。

[2] Ph.D.

Two Basic Questions on Phonetic Experimental Analysis

Wang Mai[3]

Abstract

Normalization processing is the basis for phonetic fundamental frequency analysis. The original value enters the standard space through the translation of frequency and the compression and expansion of frequency range, during which the logarithmic transformation of frequency and the determination methods of frequency range are the decisive factors of the effect of normalization. Chinese initials, either voiced or voiceless, carry no tone information, while consonant endings act as the major tone carriers. Formant distribution bears a close relation to the tongue position of vowels; therein the first formant (F1) frequency is negatively related to the height of tongue position and the second formant (F2) frequency is positively related to the front-and-back of tongue position. In Chinese, the first formants (F1) the two tongue-tip vowels are consistent, while F2 and F3 of the flat-tongue vowels [ɿ] stay adjacent to each other, but F2 and F3 of the rolling-tongue vowels [ʅ] are distant.

Keywords : phonetic experimental analysis; fundamental frequency normalization; tone; formant; tongue position

[3] Ph.D., Lecturer, Shanghai International Studies University, P.R.China
E-mail: wangmai@126.com

ประเด็นปัญหาพื้นฐานสองประการเกี่ยวกับการวิเคราะห์การออกเสียงเชิงประสบการณ์

Wang Mai[4]

บทคัดย่อ

กระบวนการปรับค่ามาตรฐานเป็นหลักการวิเคราะห์ความถี่การออกเสียงในระดับพื้นฐาน โดยค่าเดิมจะมีการป้อนช่องว่างปกติเพิ่มเข้าไปในระหว่างที่มีการแปลงความถี่ การบีบอัดและการขยายช่วงความถี่ ซึ่งในช่วงนี้เองที่การแปรผันทางลอการิทึมของความถี่และช่วงค่าความถี่ในการคำนวณจะเป็นปัจจัยตัดสินผลกระทบของการปรับค่ามาตรฐาน การออกเสียงอักษรตัวแรกในภาษาจีนไม่ว่าจะเป็นเสียงก้องหรือเสียงไม่ก้องจะไม่มีข้อมูลเสียงวรรณยุกต์ ในขณะที่เสียงพยัญชนะสะกดจะทำหน้าที่เสมือนเป็นตัวเก็บเสียงวรรณยุกต์หลักเอาไว้ การกระจายกลุ่มความถี่สั่นพ้องเป็นตัวเชื่อมความสัมพันธ์อันใกล้ชิดกับตำแหน่งลิ้นของเสียงสระ ดังนั้นกลุ่มความถี่สั่นพ้องกลุ่มแรก(F1) จึงเกี่ยวโยงในเชิงลบกับตำแหน่งความสูงของลิ้น และกลุ่มความถี่สั่นพ้องช่วงที่สอง(F2) จะเกี่ยวโยงในเชิงบวกกับตำแหน่งลิ้นส่วนหน้า-หลัง ในกรณีของภาษาจีนนั้นกลุ่มความถี่สั่นพ้องที่หนึ่ง (F1) สระปลายลิ้นทั้งสองด้านนั้นสอดคล้องกันในขณะกลุ่มความถี่สั่นพ้องที่สอง(F2) และสาม (F3) ของสระกลางลิ้น [ɿ] ยังคงติดประชิดกันและกัน แต่ F2 และ F3 ของเสียงสระที่เกิดจากการม้วนลิ้น [˞] นั้นห่างกัน

คำสำคัญ : การวิเคราะห์การออกเสียงเชิงประสบการณ์;การปรับมาตรฐานความถี่พื้นฐาน;เสียงวรรณยุกต์; กลุ่มความถี่สั่นพ้อง; ตำแหน่งลิ้น

---

[4] Ph.D., Lecturer, Shanghai International Studies University, P.R.China
E-mail: wangmai@126.com

基频归一化方法

汉语声调调值的主要决定因素是音高，广为流传的五度制标音法就是依据语流音高变化而创制的，然而这里的音高指的是相对音高，只有相对音高才对调值的判断有意义。我们都有这样的经验，男女老少，其话语的绝对音高各不相同，即使是同一个人的两次讲话，其绝对音高也不相同，但这并没有影响我们对声调的感知和判断，这是由于，语流内部各成分的相对音高是基本保持不变的，而我们感知声调正是基于稳定的相对音高，这就如同一支曲子可以用不同的调来演奏，但旋律并未改变。以汉语为例，当语流进入人耳，大脑会自动进行如下处理：①获取说话人嗓音的频率范围，估算出音高的标准值，作为参照；②把语流切成音节，再把音节切成若干个带音的片段，把每个片段同标准值进行比较，获得其偏离标准值的距离；③勾勒出每一个音节内部的相对音高变化，并与汉语声调标准调值进行比较，判断其调位归属。这是一个基本的基频归一化处理过程。

同样，我们对语音进行实验分析，也要模拟上述过程，对基频作归一化处理。归一化的目标是消除人际差异对声调判断的干扰，从变化的外在数据中提取恒定的参数。朱晓农（2004）介绍了六种基频归一化方法，其基本公式是 $x' = \frac{x_i - x_1}{x_2}$，其中$x'$是经归一化处理得到的标准值，$x_i$是实验测得的原始基频值，$x_1$和$x_2$是两个参照值，前者用来作频率的线性平移，后者用来做频域的压缩或扩大。下面对这六种归一化方法作简要介绍：

z分数转换法（z-score）。公式为 $z_i = \frac{x_i - m}{s}$，其中 z_i 为归一化值，x_i是测得的原始值，m 是样本的均值，s 是标准差，

$m = \frac{1}{n}\sum_{i=1}^{n} x_i$，$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i)^2}$。这种方法的实质是把基频数据表达为一个比值，即与样本均值的距离是标准差的多少倍。z-score没有改变原始数据的大小和分布，只是对原始数

据进行了线性变换，获得的一组z值就有了特别的性质——均值为0，标准差为1，这样就完成了对原始数据的归一化处理。

频域分数转换法（FOR）。公式为 $v_i = \frac{x_i - x_L}{x_H - x_L}$，其中$v_i$为归一化值，$x_i$是测得的原始值，$x_H$和$x_L$分别是样本频域的上限和下限值。这种方法的实质是把基频数据表达为距样本下限的距离与频域宽度的比值，它比z-score计算方法更简单，但易受极端值的影响，且没有考虑数据的分布情况。

频域比例转换法（POR）。公式为 $g_i = \frac{x_i - (m - cs)}{(m + cs) - (m - cs)}$，其中$g_i$为归一化值，$x_i$是测得的原始值，m为均值，s是标准差，c是常数。POR的思路与FOR相似，不同之处在于FOR只取两个最值作为频域的上下限，而POR则参考所有样本来确定频域的上下限。有时候最值受非常规因素影响，不能代表真实的上下限（例如气息冲击话筒出现的破音等），POR无疑可以通过多样本来冲淡这种非常规性，因此它比FOR更稳定可靠。

对数半音差比转换法（LD）。公式为 $d_i = \frac{\lg x_i - \lg x_L}{\lg x_H - \lg x_L}$，其中 d_i 为归一化值，x_i 是测得的原始值，x_H 和 x_L 分别是样本频域的上限和下限值。这种方法的实质是把基频数据表达为与两个半音差距的比值。与FOR不同，LD把每一个取值做了对数处理，它基于这样的事实：虽然音阶的升降是一种线性关系，但与之相应的频率升降却是非线性的，而是一种指数关系，声音每升高一个八度，频率为原来的两倍（如低音、中音、高音C的频率分别为261.6Hz、523.3Hz、1046.5Hz,）。因此，通过对频率进行对数处理，就可以把指数关系降阶为线性关系，使之与五度制标音法各阶的线性关系相一致。相邻半音的频率比值进行对数化处理后，可以得到一个常数0.083，这是对参照值所作的归一化处理。类似地，频域的上下限 x_H 和 x_L 的确定也可以使用 FOR 和 POR 两种方式，我们经常使用的T值法（石峰，1986）为前一种方式，直接把频率最值作为频域的上下限；而D值法（沈炯，1985）为后一

种方式，频域上下限由多个样本共同确定。

对数 z-score 转换法（LZ）。公式为 $z_i' = \frac{y_i - m_y}{s_y}$，其中

$y_i = \lg x_i$，$m_y = \frac{1}{n}\sum_{i=1}^{n}\lg x_i$，$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\lg x_i - \frac{1}{n}\sum_{i=1}^{n}\lg x_i)^2}$，

c 为常数。与前述 z-score 法相比，它在计算均值 m_y 和标准差 s_y 时，都首先对原始数据做了对数化处理，把频率的指数关系降阶为线性关系。

对数频域比例转换法（LPOR）。公式为

$g' = \frac{y_i - (m_y - cs_y)}{(m_y + cs_y) - (m_y - cs_y)}$，其中 $y_i = \lg x_i$，

$m_y = \frac{1}{n}\sum_{i=1}^{n}\lg x_i$，$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\lg x_i - \frac{1}{n}\sum_{i=1}^{n}\lg x_i)^2}$

c为常数。它保留了前述POR法的优点，用全部样本计算频域的上下限，具有稳定性；同时对原始数据做了对数化处理，使音高关系具有了线性特征。

上述六种归一化方法各具特色，前三组是基本的，后三组在前面的基础上增加了对数降阶处理。优秀的归一化方法可以将不同人相同内容的语流转换成基本重合的基频曲线，这正符合人们从不同声源听辨并获取相同信息的过程。朱晓农（2004）采用标准指数（NI）、离散系数（DC）和人际均方差（MBSV）对归一化效果作出评价，通过对沪方言阴平调11个人的发音所做的归一化处理，得到两组急降型和缓降型基本重合的归一化曲线，从而认为LZ法是目前最佳的基频归一化方法。

音节时长归一化及带调段的选择

声调是复杂的语音现象，决定声调的主要物理参数是音高，但音长和音强也是重要的伴随特征。例如，汉语方言中，入声比舒声要短促得多；普通话中，上声较其他三声要长，轻声则极为短促，且音强也小。人际差异则表现为有的人说话慢吞吞，拉长调，有的人则急促有力。因此，我们在进行声调分析时，有必要对音长和音强特征也进行归一化处理。在提取调值的样本数据时，通常按照相对时间坐标，取5%、10%、20%……90%、100%处的数值，这就把样本间的随机因素尽可能地剔除了。当然，声调的绝对时长需另行记录，它也是判断调类归属需要参考的重要参数。

另一个需要讨论的是相对采样时点的选择问题。我们需要考虑两方面的因素：①声调的附着性。声调是超音段音位，它附着在占据时轴空间的音段音位之上，但并非所有的音段音位都附着有声调。汉语的音节结构是CVC，其中前后辅音可以不出现。前辅音有两种情况：为清音时，自然是不带声调的；为浊音时，声带振动，可以带声调，但平悦玲等（2001）认为浊声母的带音现象纯粹是这些音段本身的发音特征，不具有调位意义，应排除在样本之外。后辅音只有 [n] 和 [ŋ]，它们都是带音的，与前辅音不同，它们是重要的声调承载段。元音部分比较复杂，可以是单元音、双元音，也可以带半元音介韵，但无论是那种情况，它都是声调的主要承载段。②发音器官的生理干扰。声调曲线往往带有弯头和降尾，这是受发音器官生理限制形成的：发声时，声带由松弛状态进入振动状态，有一个渐变的过程，高调起始时弯头为上升状，低调起始为下降状；相反，语流终止时，声带由紧张状态进入松弛状态，高调终止时弯尾为下降状，低调终止时为上升状。弯头降尾不具有声调意义，应排除在样本之外。

归一化效果测试

我们用 praat 对两个样本分别进行基频数据提取，并使用T值法（石锋，1986）对其进行归一化处理，计算两个样本的五度制

调值，并对比调值以验证T值法的归一化效果。样本为成语“千奇百怪”，它正好包含普通话的四个声调；样本分别由男女声朗读，男声的普通话水平为二级甲等，女声未测试，听感上基本标准。

首先我们获取样本的基频，得到如图1的基频曲线图：

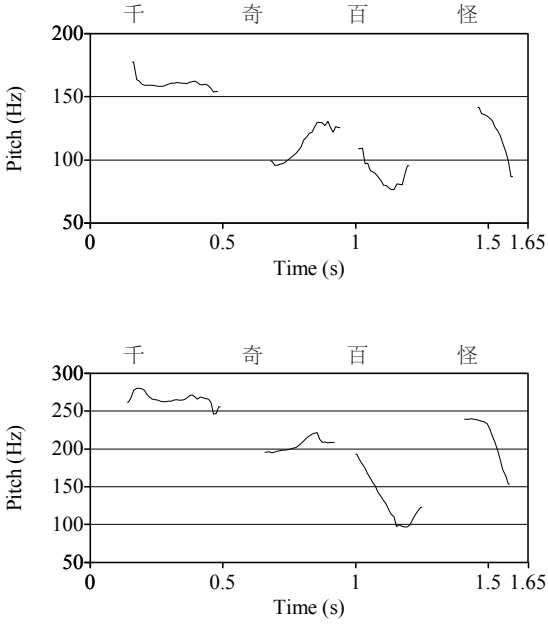


图1 基频曲线（上图为男声，下图为女声）

然后获取样本相对时点的基频值，为方便批量操作，可以使用下面的脚本。运行脚本时，在弹出窗中输入适当的采样点间隔，即可获得所需数量的基频数据。

```
# Getting the interval from the user:
form Pitch listing by specific interval value...
  comment Supply the interval for listing:
  positive Interval(sec) 0.03
endform
```

```
# Printing the table header:
echo Pitch listing by the interval of 'interval' sec:
printline
printline Time(sec)'tab$'F0(Hz)
printline ------------------------

# Initializing loop variables:
begin_from=Get begin of selection
end_by=Get end of selection
time=begin_from

# Listing by specific interval:
while time<=end_by
  Move cursor to... 'time'
  f_0=Get pitch
  printline 'time:8''tab$''f_0:2'
  time=time+interval
endwhile
printline ------------------------
```

获取的数据记录如下，每组10个数据分别表示相对时长5%、15%……95%处的基频值：

样本	“千”		“奇”		“百”		“怪”	
朗读者	男	女	男	女	男	女	男	女
基频数据（Hz）	161.37	264.36	95.81	195.24	104.45	193.09	141.50	239.38
	159.14	280.24	97.71	197.67	95.65	180.56	137.63	239.55
	158.72	270.91	101.87	198.85	90.14	164.29	135.72	238.41
	158.69	264.68	107.05	200.98	86.42	150.38	133.80	236.78
	160.66	262.79	116.59	205.56	79.98	134.49	131.43	233.50
	160.97	264.98	122.20	214.11	78.08	119.45	126.16	221.33
	160.82	265.62	129.74	220.61	76.35	104.66	121.93	206.14
	162.01	269.87	127.88	211.93	80.73	97.91	114.89	187.16
	159.69	267.83	122.72	208.49	85.09	98.00	106.71	167.67
	155.69	254.66	125.52	208.43	95.16	111.83	96.45	153.75

表 1 基频采样值

我们对采样值进行归一化处理，公式为 $T=\frac{\lg x-\lg b}{\lg a-\lg b}\times 5$，其中 x 为测量值，a、b 为声调频域的上下限。上下限通过查找最值得到，此处，男声频域为[76.35，162.01]Hz，女声频域为[97.91，280.24]Hz。经过计算，我们得到每个采样数据的T值：

样本	“千”		“奇”		“百”		“怪”	
朗读者	男	女	男	女	男	女	男	女
T值	4.97	4.72	1.51	3.28	2.08	3.23	4.10	4.25
	4.88	5.00	1.64	3.34	1.50	2.91	3.92	4.25
	4.86	4.84	1.92	3.37	1.10	2.46	3.82	4.23
	4.86	4.73	2.25	3.42	0.82	2.04	3.73	4.20
	4.94	4.69	2.81	3.53	0.31	1.51	3.61	4.13
	4.96	4.73	3.13	3.72	0.15	0.95	3.34	3.88
	4.95	4.75	3.52	3.86	0.00	0.32	3.11	3.54
	5.00	4.82	3.43	3.67	0.37	0.00	2.72	3.08
	4.90	4.78	3.15	3.59	0.72	0.00	2.22	2.56
	4.74	4.54	3.30	3.59	1.46	0.63	1.55	2.15

表 2 计算所得 T 值

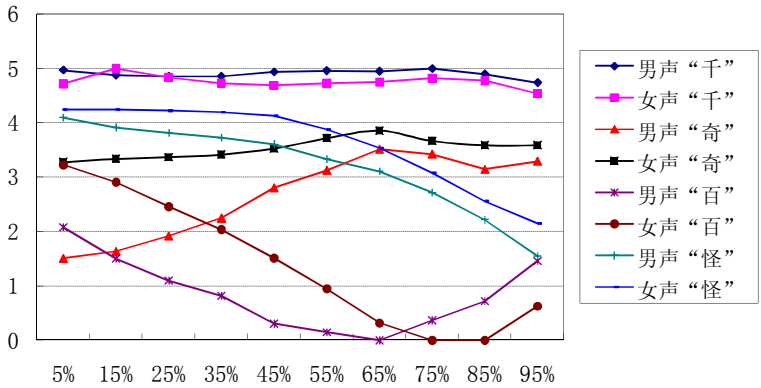


图 2 T 值折线图

图2 是根据T值制作的声调折线图，我们作如下分析：①平声“千”的男女声折线基本重合，调值为高平55。②去声“怪”的两条折线的重合度也较高，调值可定为男声52，女声53。③上声“百”的两条折线，重合度不高，主要是时间差造成的，女声的低点较男声晚了约1/6个音节时长，但折线的总体走向是一致的，调值可定为男声312，女声411，约为半上的样子。④差异最大的是阳平“奇”的两条折线，男声调值为24，女声调值则接近44。仔细观察，我们发现，女声“奇”的声调确实有些不同，其起始段受到前方阴平调的影响而显著抬升；其结束段又受到后续上声调的影响而略有下降，形成了独特的升降调型。有趣的是，男声的“奇”也出现了相似的情况，虽然其调型总体上是上升趋势，但其基频的峰值并非我们预测的那样出现在音节末尾，而是在中间65%处，这种现象不是偶然的，应该予以重视。上述音变现象可能是在语流中受相邻声调的影响而产生的条件变体，它不会给人造成听感上的别扭，是完全可以接受的，但若将其单独切分出来，作为单字调来审听，我们就会立即发现其与标准阳平调的显著差异，听感上也难以接受。可见同一声调的调值在单字状态和语流中存在较大差异，心理接受度也存在较大差异。基频分析以及相应的归一化处理忠实地记录了汉语声调的客观面貌，使我们拥有了发现靠主观认知很难发现的某些现象和规律的能力。

元音共振峰分布与舌位的关系

语音中除了基音，还有与其呈倍数关系的若干陪音，语音能量在各陪音上的分配不是平均的，而是有规律地集中在部分陪音上，形成一个个能量区。这些由若干相邻高能量陪音聚集而成的陪音簇，在语图中表现为一道道深色的横杠，称为共振峰（formant）。共振峰处于基频之上，自低频区向高频区依次排列，分别记做F1、F2……Fn，基频处于最下方，记做 F0。共振峰同声音的音色有密切关系，其数量的多寡、排列的疏密，是决定音色差异的主要因素。钢琴与小提琴的音色差异，语言中不同元音间的音色差异，都可以分析为共振峰的不同排列组合。图5是语流“上海外国语大学”各带音段的共振峰分布状况，有些带音段的

共振峰分布有一个变化过程，正与双元音音节内音色变化的过程相对应。简而言之，基频决定音高，共振峰决定音色，它们都是语音分析中的重要参量。

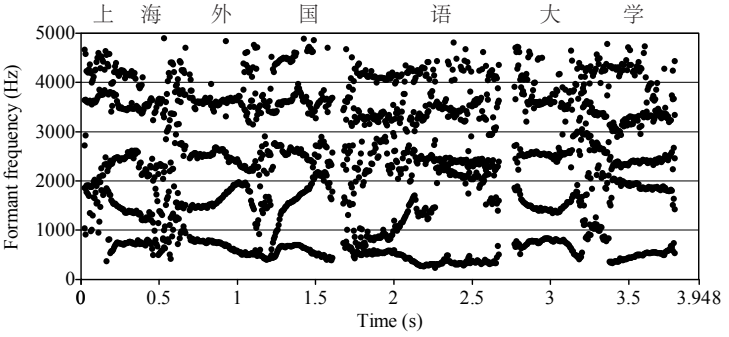


图 3 共振峰分析

音色的样态极为丰富且具灵活性，是元音等音段音位最主要的辨义手段，也是语言承载意义的最重要的形式。从生理发声学分析，一个元音可以由舌位的高低、前后及唇的圆展来确定；实验语音学则告诉我们音色主要由共振峰数量及分布状况决定。语音的生理属性和物理属性不是彼此独立的，它们是从不同的角度对语音做出的描述，因此，两者间应该具有某种形式的关联，或者可以通过某种函数关系相互转换。

研究证实，元音的舌位差异主要表现为第一共振峰（F1）和第二共振峰（F2）的频率变化。舌位的高低主要与F1相关：舌位越高，F1的频率就越低；舌位越低，F1的频率就越高。舌位的前后主要与F2相关：舌位越前，F2的频率越高；舌位越后，F2的频率越低。唇的圆展也与F2相关，展唇元音F2偏高，圆唇元音偏低。

我们以第一共振峰频率为Y轴、第二共振峰频率为X轴，建立声学元音图，标示16个定位元音在坐标系中的位置。为了同标准元音图的方向一致，便于观察，X、Y轴都取反向；同时取以2为底的对数刻度，这样可以把频率的指数级关系降阶为线性关系。共振峰数据可以通过下面的脚本获取，在弹出窗中设定适当的取样间隔，就可以获得相应时点和密度的数据。缺省状态下，它只获取F1~F4四个共振峰数据，这对音色研究已经足够了，更高的共振峰能量衰减明显，同音色的关系也不明确。

```
# Getting the interval from the user:
form Formant listing by specific interval value...
  comment Supply the interval for listing:
  positive Interval(sec) 0.01
endform

# Printing the table header:
echo Formant listing by the interval of 'interval'
sec:
printline
Time(sec)'tab$'F1(Hz)'tab$'F2(Hz)'tab$'F3(Hz)'tab$'F4
(Hz)
printline ------------------------

# Initializing loop variables:
begin_from=Get begin of selection
end_by=Get end of selection
time=begin_from

# Listing by specific interval:
while time<=end_by
  Move cursor to... 'time'
  f_1=Get first formant
  f_2=Get second formant
  f_3=Get third formant
  f_4=Get fourth formant
  printline
'time:8''tab$''f_1:2''tab$''f_2:2''tab$''f_3:2''tab$'
'f_4:2'
  time=time+interval
endwhile
printline ------------------------
```

表 6 是 Catford, J.C.（1988）测得的十六个标准元音的第一共振峰和第二共振峰的频率值。图7是根据表6绘制的声学元音图。图8是作为参照的标准元音舌位图。

元音 (IPA)	共振峰 f1	共振峰 f2	元音 (IPA)	共振峰 f1	共振峰 f2
i	240 Hz	2400 Hz	ɑ	750 Hz	940 Hz
y	235 Hz	2100 Hz	ɒ	700 Hz	760 Hz
e	390 Hz	2300 Hz	ʌ	600 Hz	1170 Hz
ø	370 Hz	1900 Hz	ɔ	500 Hz	700 Hz
ɛ	610 Hz	1900 Hz	ɤ	460 Hz	1310 Hz
œ	585 Hz	1710 Hz	o	360 Hz	640 Hz
a	850 Hz	1610 Hz	ɯ	300 Hz	1390 Hz
æ	820 Hz	1530 Hz	u	250 Hz	595 Hz

表 3 十六个定位元音的共振峰数据

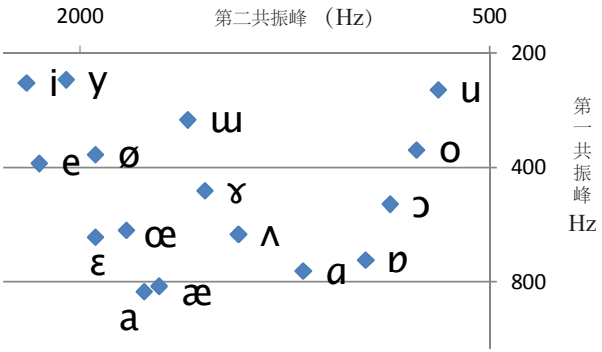


图 4 声学元音图

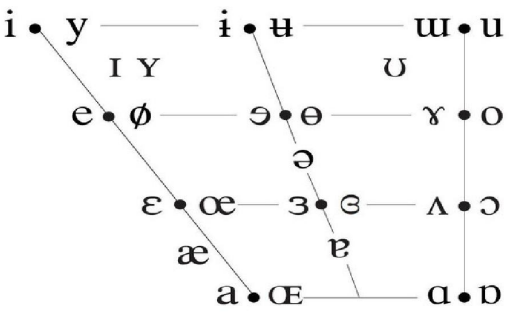


图 5 标准元音舌位图

图例显示，根据共振峰频率绘制的元音图，各元音的位置大致和标准舌位图的元音位置相当。这表明，生理语音学归纳的元音舌位图得到了现代声学实验的支持，两者可以相互印证。

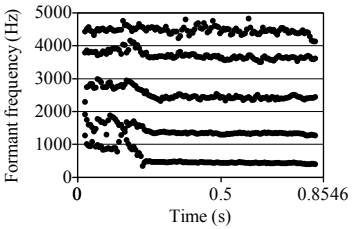
但我们也发现了一些差异：圆唇元音在声学元音图中较相同舌位的展唇元音要靠后；在生理舌位图中却共用一个位点。我们可以这样解释：由于圆唇元音有一个嘴唇聚拢撮起的动作（俗称撅嘴），这无疑增加了口腔（共鸣腔）的空间，与舌位后移增大共鸣腔有相似的效果，共振峰数据忠实记录了这种联动效应，唇的展圆与舌位的前后在声学元音图中就有了相似的特性。与声学分析相对，生理元音图表述的是一种理想状态，发音心理上，唇的圆展是一个独立的参数，不与舌位、共鸣腔大小等参数产生联动，即使由此引发了其他参数的变化，也只是额外的副作用，不在考虑范围之内。关于心理音值和物理音值的对立，语言中可以找到很多例子，有一种理论认为，汉语声调的全上和半上在心理音值上都是低平调，在具体的语流中，受到前后音节的影响以及发音器官自身的生理限制，而形成了214或211的声学调值，等等。上述现象与此类同，也可以归为心理音值同物理音值的差异现象。

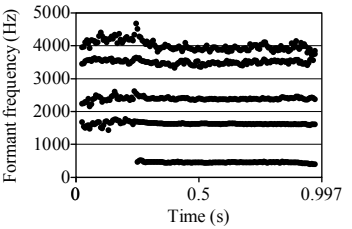
进一步，我们还发现，展唇元音同圆唇元音的差异，在后高元音上表现得尤为显著，[ɯ]的F2频率几乎是[u]的近2.3倍，甚至比央元音的F2还要高而接近前元音。 这是什么原因呢，这也可以

从发音器官的生理特性得到解释。我们知道，央元音的舌位处于中间位置，是最自然放松的状态；前元音的舌位靠前，这符合舌头舔舐食物的自然生理机能，因此也较轻松；后元音却不同，舌位后移不符合一般生理机能，且会挤压气道，因此肌肉的紧张度最高。相应地，唇在放松状态下是平展的，因此发展唇元音时肌肉松弛；而发圆唇元音则必须使肌肉紧张，唇部隆起。受颌面部神经控制的影响，唇与舌这一前一后的肌肉状态往往产生联动效应：发圆唇后元音[u]时，唇舌两处的肌肉都处于紧张状态，共鸣腔拉长，舌位移后；发展唇元音[ɯ]时，唇部肌肉放松，舌部肌肉也联动地放松，使得舌位略微前移。舌位与共鸣腔大小此消彼长，两者的比值也就产生了显著的变化，这与舌位的前后移动有相似的声学特征，在声学元音图上得以清晰地展示。需要强调的是，这仍然是生理层面的联动变化，在心理层面上，唇和舌的状态依然是彼此独立的理想状态，因此在生理舌位图上没有体现。

平卷舌元音的共振峰分布差异

汉语元音中，存在一对舌尖元音[ɿ]和[ʅ]，与舌面元音不同，它们的区别性特征不是舌位的高低、前后或唇的圆展，而是平卷舌的对立，它们的共振峰分布有何差异？图9是“思”和“诗”的共振峰轮廓图。前约1/4部为辅音部分 [s] 和 [ʂ]，我们从较稳定的后3/4，即 [ɿ] 和 [ʅ] 的部分提取两者的共振峰数据，如下表所示：





	元音部分	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
思	均值	442.54	1323.50	2519.66	3674.20
诗	均值	449.95	1625.27	2369.18	3479.26

图6 “思”和“诗”的共振峰分布

两者的F1频率基本一致；F2频率“思”比“诗”低些；F3频率“思”又比“诗”高些，因此可以大致得出这样的结论：汉语的两个舌尖元音 [ɿ] 和 [ʅ] 的F1频率基本相同，F2和F3的分布略有差异，平舌元音F2与F3相距较近，卷舌元音F2与F3相距较远。

References

朱晓农. (2004). 基频归一化——如何处理声调的随机差异？. **语言科学**. 2004 (4).

平悦玲等. (2001). **吴语声调实验研究**[M]. 上海: 复旦大学出版社. P5-16.

石锋. (1986). 天津方言双字组声调分析. **语言研究**. 1986 (1).

石锋. (2003). 语音格局的分析方法. **第六届全国现代语音学学术会议论文集（上）**.

沈炯. (1985). 北京话声调的音域和语调. 林焘, 王理嘉. **北京语音实验录**. 北京: 商务印书馆.

岑麒祥. (1982). **国际音标**[M]湖北: 湖北人民出版社.

马大猷. (2001). **现代声学理论基础**[M]. 北京: 科学出版社.

Catford, J.C. (1988). **A Practical Introduction to Phonetics**, Oxford University Press. pp.161.

2012年度中国教育部人文社科规划项目“以自主学习为特征的外国学生汉语语音训练研究”，编号：12YJAZH008.

2012年度上海市教育委科研创新重点项目“基于语料库的外国学生汉语语音训练研究”，编号：12ZS063.