

第二章 · 声波

不同颜色是光谱中波长不同的光波，
不同音色是声谱中波长不同的声波。

物不得其平则鸣。（韩愈）

1. 波的性质和分析

5. 语音信号分析

前文介绍了语音有音高、音长、响度、音色四大要素及其波形特征。本节内介绍如何从语图上确定相应的基频、时长、音强、共振峰。‘语图’是个专门用语，指三维声谱图，横轴表示时间，纵轴表示频率，颜色深浅表示能量高低。制作语图过去是用模拟式的语图仪，昂贵而笨重。现在已经全面换用方便、灵巧、高效的电脑软件。下面介绍一种目前非常流行的语音分析软件 Praat。这个软件由荷兰阿姆斯特丹大学的 Paul Boersma 和 David Weenink 编写，功能强大而使用方便，可以从网上免费下载：http://www.fon.hum.uva.nl/praat/download_win.html。

5.1. 文件操作

启动 Praat，会出现两个界面（见图 18）。左面是操作语音分析用的 Praat objects（对象窗口），这是 Praat 的主界面。右面是编辑和制作语音视图用的 Praat picture（图像窗口）。

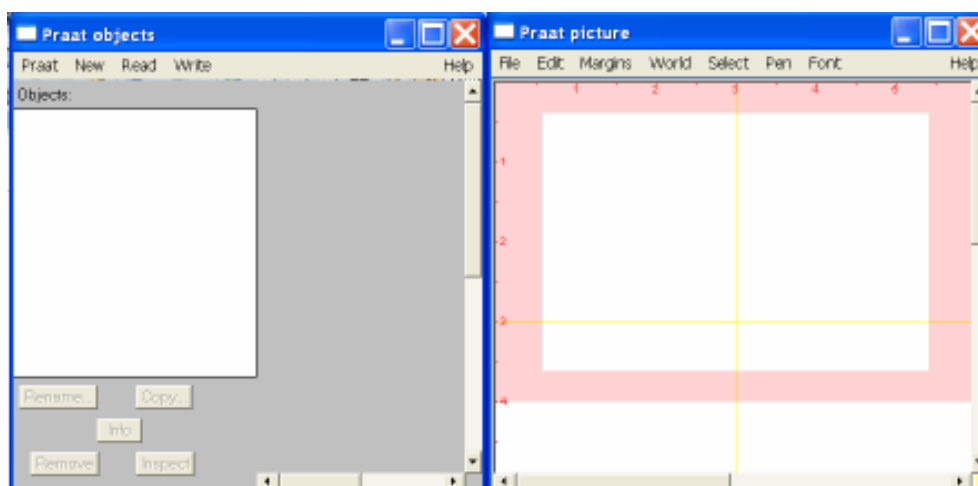


图 18. Praat 软件的 [左] 对象窗口, [右] 图像窗口.

输入待分析的语音文件有三种办法, 也即 Praat 软件提供三种输入语音文件的功能。

第一是利用 Praat 来制作一个声音。

第二是打开已有的声音文件, 可以从主界面即对象窗口的主菜单中选择 Read (读取), 然后点击 Read from file... (文件读取), 就能搜寻并打开已有的音档。

第三是新录音, 可以在菜单中选择 New (新文件), 见下图 19 左。点击 Record mono Sound... (单声道录音), 就有一个 SoundRecorder (录音机) 的窗口弹出 (见图 19 右)。录音采样率 (Sampling frequency) 预设设在 22,050 Hz。这对于一般的语音分析来说是很合适的。设定过高, 占内存太多; 设定过低, 高频段的信号就测不到。

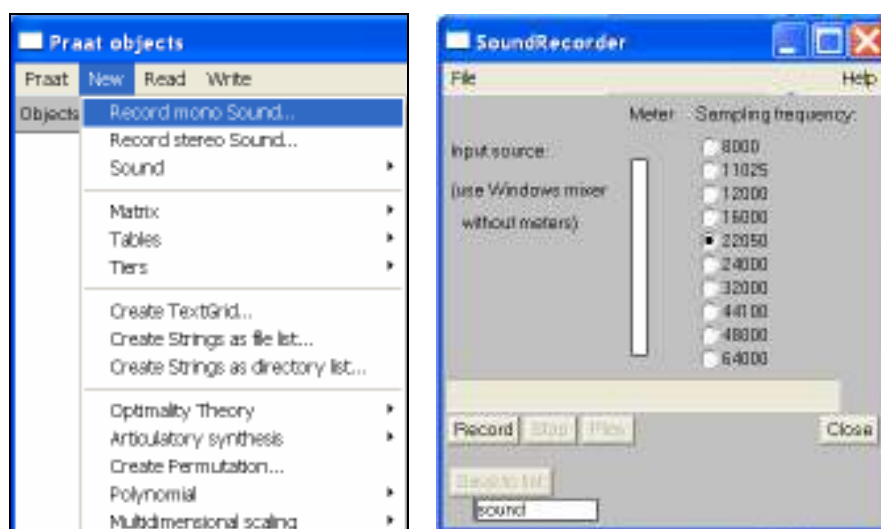


图 19. [左] 在对象窗口的菜单中选取 New → Record mono Sound..., [右] 录音机窗口

在 SoundRecorder 窗口内点击 Record, 即开始录音。录音时注意音量 (响度) 控制, 音量由 Meter (音量计) 中的三种颜色显示。绿色表示正好, 黄色警告, 红色表示音量过大 (见图 20

左)，此时录下的语音的声波会出现‘削波’现象。录音结束，按 Stop（结束），然后点击 Save to list（存取至对象单）。此时在 Praat objects 窗口内就出现一个预设名为 sound 的文件（见图 20 右）。如果想自取文件名，可以在录音机窗口左下方的方框中（图 20 左），删去 sound，键入自取的文件名。

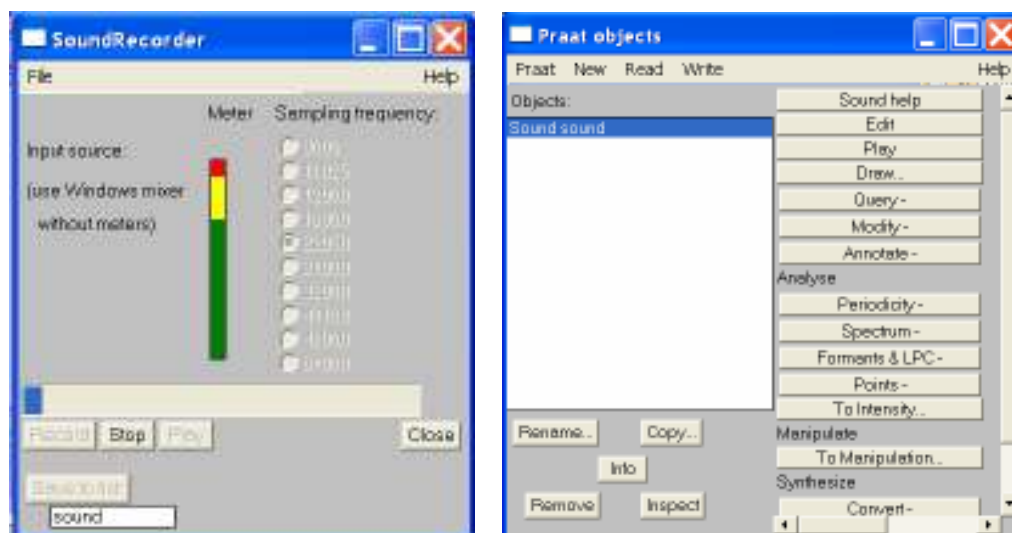


图 20. 录音和存取.

在主界面的菜单中拉下 Write（存写），点击 Write to WAV file...（存储为 WAV 文件），储存这个声音文件（图 21 左）。

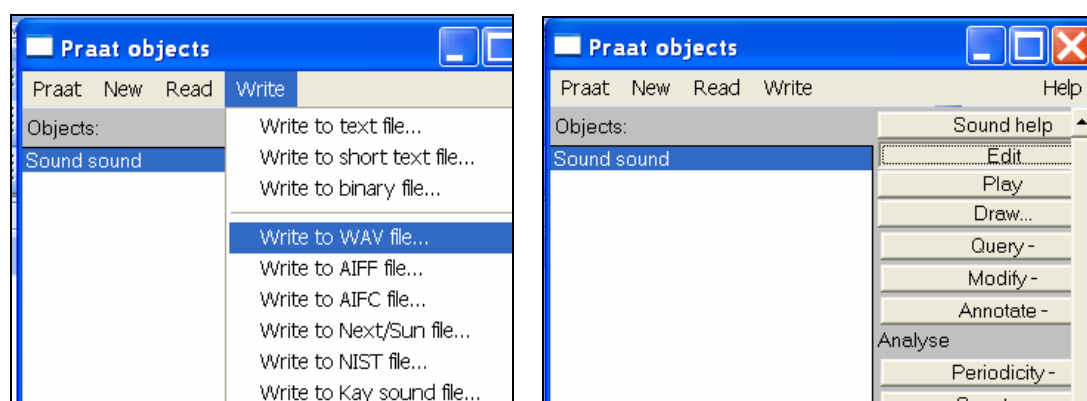


图 21. [左] 储存文件, [右] 编辑文件.

5.2. 语音分析

使用 Praat 软件，可以进行多种语音分析工作，如共振峰分析、基频分析、音强分析、时长分析、频谱分析、波形编辑等等。这些都可以通过一个 SoundEditor（语音编辑器）的窗口来进行。点亮对象窗口中的目标声音文件，然后点击右面的 Edit（编辑）键（图 21 右），就有一个显示波形的窗口出现（图 32），可以播放、测量、编辑声波信号等。

图 22 上半部是声压波形图，简称波形图，纵轴为振幅，单位是任意的，或者说相对的，已

被规整到 ± 1 范围之间。下半是宽带语图。如果宽带图没有在窗口中显示出来，则可在菜单中选取 Spectrum（频谱），点击 Show Spectrogram（显示声谱）。宽带图的纵轴是频率，单位是赫兹 Hz。横轴是时间，单位是秒 s。并通过灰度的深浅表示能量的强弱。点击下面的时间杠，可以播放相应段的声音。

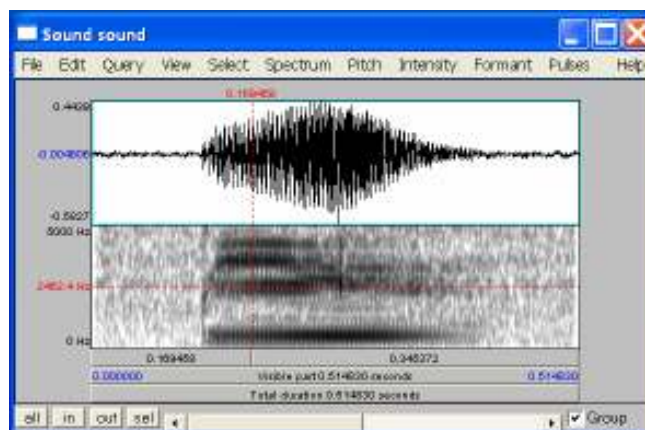


图 22. 声波图和宽带频谱图.

语图的左下角有四个小按钮：all, in, out, sel，其功能分别为显示整个音档、放大、缩小、显示所选点亮的语音段。图 23 左图中元音有一段 134 毫秒的点亮部分。点击左下角 sel（显示选择部分），就把这 134 毫秒放大至右图的形式。下面宽带图中一条一条的竖条表示喉门脉冲，即由声门开闭引起的快速声压变化。图 4 上面声波图中两条小竖线指向两个相邻的正值波峰，表示一个与喉门脉冲相对应的声波周期。

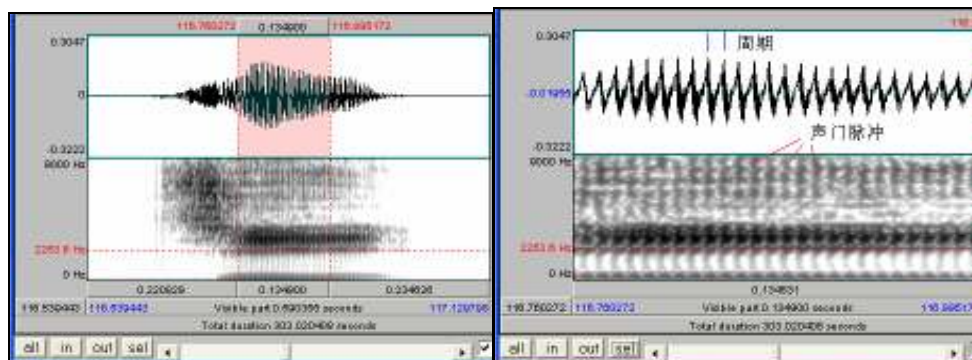


图 23. 【左】 $jī$ 的声波图和宽带图. 【右】放大图.

5.2.1. 宽带图和共振峰

共振峰指的是在声道的共振频率附近，一些谐波的能量因声道共振作用而得到加强提升，共同构成的一个强频区。图 24 宽带图中的那些黑色横条状是共振峰的走向。在 5,000 赫兹范围内，共有四条黑带，表示第一到第四个共振峰（即 F1, F2, F3, F4）。那条红色的横虚线经过第二共振峰的中心，表示 F2 的频率目测大致为 2,462 Hz。在以前模拟式语图时代，获得共振峰数据的方法，就是用手工在黑带中心画条线，然后估算其共振频率。现在只需在菜单中选择 Formant（共振峰），点击 Show formant（显示共振峰），就能得到如下图 24。图中有四条红色点状线，显示

了头四个共振峰。其中 F2 在游标竖线处的频率，可以从左面的纵轴上读出：2,475 赫兹。有关共振峰的测量，后文‘元音’章再详谈。

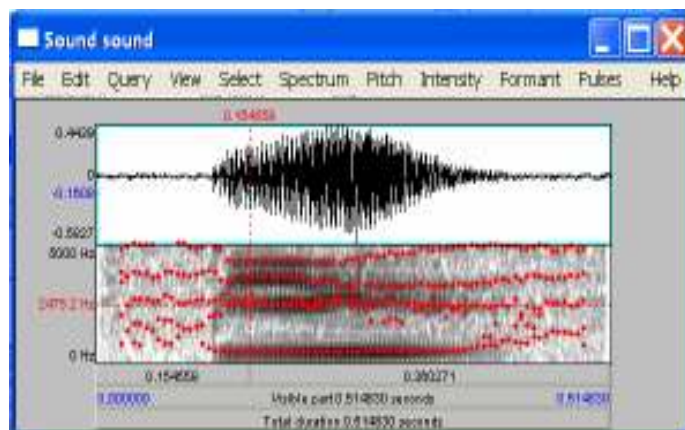


图 24. 共振峰走向显示

5.2.2. 窄带图、谐波、基频

上文所述的宽带图是 Praat 软件的预设图。如果想要窄带图，可以选择菜单中的 Spectrum（频谱），点击 Spectrogram settings...（声谱设置）（见图 25 左），就会弹出一个窗口（图 25 右）。其预设的宽带图的频率可视范围（View range）为 0 到 5,000 赫兹，窗口长度（Window length）0.005 秒。如果把窗口长度改为 0.04 秒，可视范围改为 0~3,000 赫兹，就能得到如图 25 那样的窄带图。

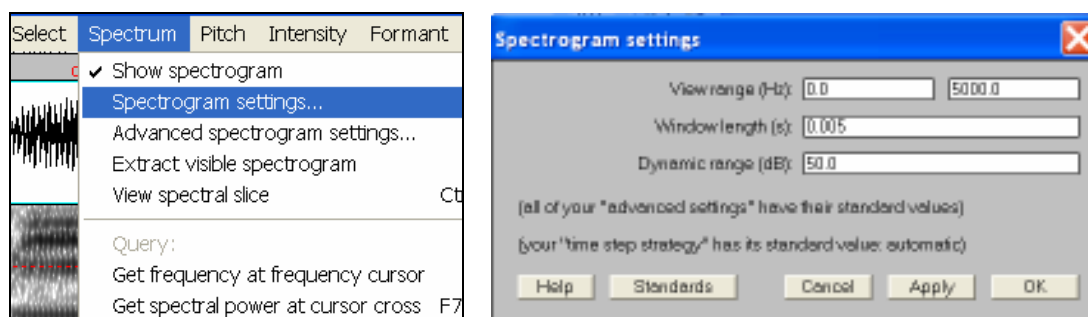


图 25. 频谱图设置

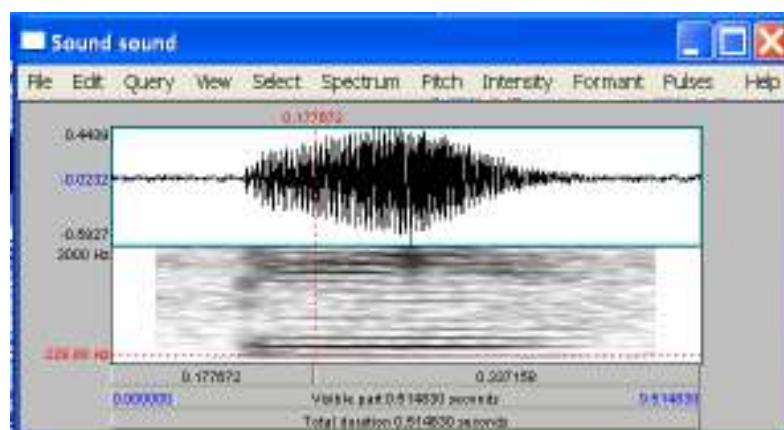


图 26. 窄带图

窄带图中有很多根横向细条，表示基频和谐波。最下面的一条是基频，用 F_0 或 F_0 表示。现在的习惯是把基频也叫做第一谐波 H_1 。图 26 中有条红色横虚线经过第一根细条的中心，从左面的纵轴可以读得基频为 229 赫兹。基频上面的细条，即从下往上数第二根细条是第二谐波 H_2 ，再往上是 H_3 , H_4 , 以此类推。在以前模拟式语图仪时代，这些谐波被用来测量估算基频。不过，现在已经能直接从软件上获得数据。在主界面菜单上选择 Pitch（音高），点击 Show pitch（显示音高），就出现图 27。图中的蓝线就是基频曲线，竖的游标线通过处的基频可以从右面的纵轴上读出：234 赫兹，与图 26 中测得的 F_0 229 Hz 相比，误差 2% 左右，不算太大。如果测量图 26 中的 H_{10} 或 H_{11} ，误差会有实质性的减小。有关基频测量作图我们到后面‘声调’章详谈。

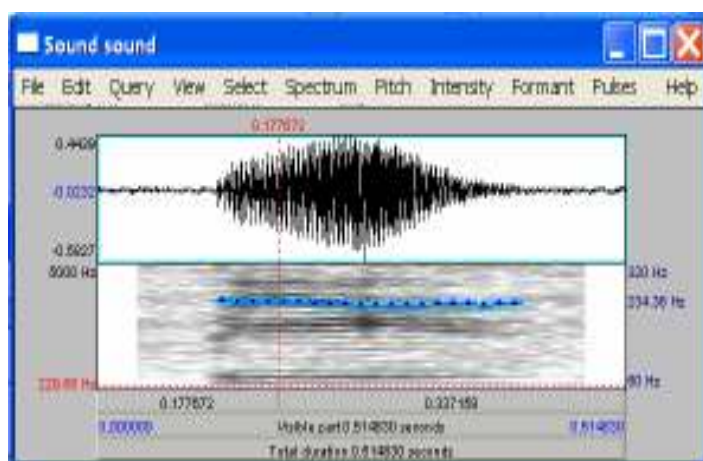


图 27. 基频曲线显示

5.2.3. 音强曲线

不同的音强表现在传统的模拟式语图上即为颜色深浅不同，而在现在的数字语图上就像基频曲线一样能显示出音强曲线。在菜单中选择 Intensity（音强），点击 Show intensity（显示音强），就得到图 28。图中黄色细线即为音强曲线。竖的游标线通过处的音强在右面的纵轴上可以读出：81 dB 分贝。这是这个音节的音强最大处，与上面波形图中的振幅最大处相吻合。注意：分贝是一个信号强度相比于噪声强度的相对的量度。

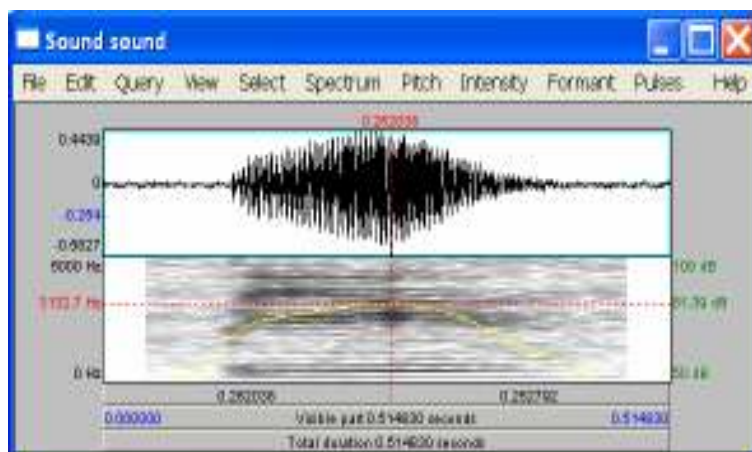


图 28. 音强曲线显示

5.2.4. 音强声谱图

我们知道，语图实际上是三维的，横轴时间，竖轴频率，音强是用颜色深浅显示的。音强能量的截面声谱图 FFT Spectral Slice 就是在某个时间点上的频率声谱图的切面，显示的是频率 Hz（横轴）和音强 dB（竖轴）的二维关系，即在某个频率段上的音强能量。制作截面图有几种办法，最简单的一种如下。先制作一个窄带图，即把 window length 设成 0.04s，见图 29 左小图。然后把游标放在元音起始后三五十毫秒处，大约是音节短就稍往前点，音节长就稍往后点。图中的红竖虚线在 36 毫秒处。然后从菜单上拉下 spectrum（强谱），选取 view spectral slice（看强谱截面图），见右小图。

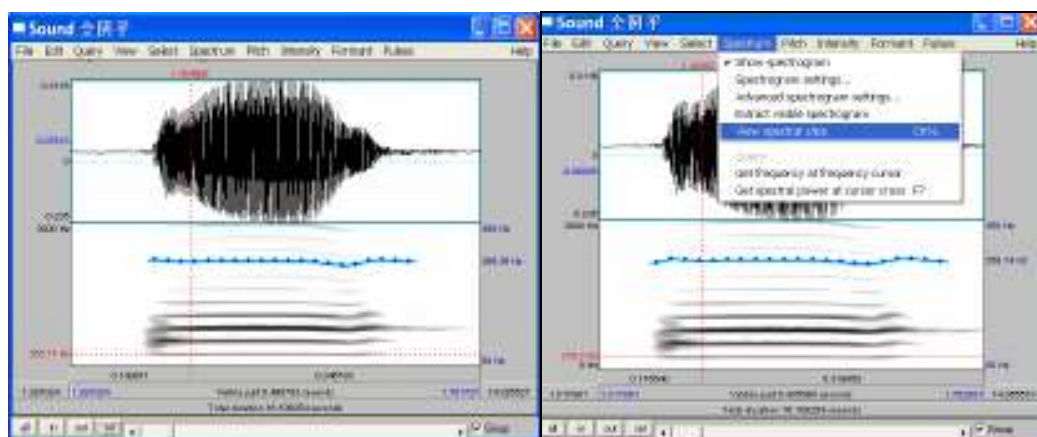


图 29. 制作强谱截面图过程，【左】吴江松陵女发音人 PMF ‘刀’ tō 的窄带图，
【右】选取 view spectral slice

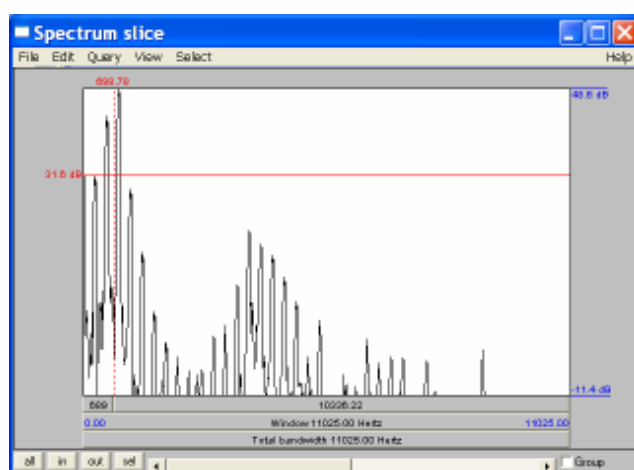


图 30. 强谱截面图

于是得到图 30 的强谱截面图。横轴为频率，单位赫兹 Hz；右面竖轴为音强，单位分贝 dB，标度从 -11.4 到 48.6 dB。图中一根根竖窄条即为谐波。第一条是第一谐波 H1，也就是基频 F0，音强可从左面的标尺上读出：31.6 dB。第二条是第二谐波 H2，第三条是 H3 等等。同时可以估算共振峰，例如在 H2 和 H3 之间靠近 H3 处，即竖虚线处估得 F1 为 689 Hz。