

声图及其在 喉科的临床应用

编著 郭志祥

上海医科大学出版社

声图及其在喉科的临床应用

郭志祥 编著

上海医科大学出版社

责任编辑 高敬泉
封面设计 吴 平

声图及其在喉科的临床应用

郭志祥 编著

上海医科大学出版社出版

上海市医学院路 138 号

邮政编码 200032

新华书店上海发行所经销

常熟市新华印刷厂印刷

开本 850×1168 1/32 印张 5 插页 1 字数 129 000

1991 年 4 月第 1 版 1991 年 4 月第 1 次印刷

印数 1—2,800

ISBN 7-5627-0075-3/R·66

定价 5.00 元



郭志祥 男 1932年生。1952年毕业于第一军医大学医疗专业。现任国际嗓音学会会员、中华耳鼻咽喉科学会委员、《中华耳鼻咽喉科杂志》编委、中国头颈外科学会常委、中国人民解放军耳鼻喉科中心主任、第四军医大学兼职教授、空军沈阳医院副院长。

他从事耳鼻喉科临床38年来，着重从事喉外科临床及喉病理及嗓音疾病的研究。曾先后获全国科技大会奖、军队科技进步2、3、4等奖10次。发表学术论文200余篇。出版《喉癌的手术》录像带，并参加编写《中国大百科全书(医学卷)》、《全军医疗技术操作常规》等。

前 言

嗓音和言语医学是研究人类发声及喉发声疾病的多学科交叉的科学,它涉及声学、电学、电子学、微电子技术、电子计算机技术、语言学、语言声学、语言病理学、喉科学等范畴,形成了对人声或嗓音进行生理、病理研究的新科学。近年来我国研究工作者和临床医师,应用声图仪(又称语图仪或声谱仪)、声门电阻抗仪、喉摄影术、喉录像等现代科学手段,研究喉的发声学、嗓音医学,并取得了一定的成就。

在耳鼻咽喉科领域中,听力学的检查已比较普及,在我国许多医院都有脑干电反应测听仪,可以客观检查听力。然而对嗓音的检查,大多数医院仍处于凭藉耳听觉进行主观判定阶段。

作者采用声图仪研究分析喉生理和病理嗓音的改变,仅是初步尝试,目的在于引起喉科同道重视嗓音的客观检查,以推动本学科的发展。鉴于作者是一名长期从事耳鼻咽喉科临床工作的医师,对多学科的知识甚感匮乏,本书编写过程中,虽反复推敲,几易其稿,但不当之处恐难避免,敬请读者批评指教。

作者

1990年4月

目 录

前言

第一章 绪论	1
第二章 声学概论	7
第一节 物理声学.....	7
第二节 生理声学.....	10
第三节 语言声学.....	18
第三章 噪音的声学分析	29
第一节 声学分析的信号记录.....	29
第二节 声学分析的样本.....	30
第三节 估价发声功能的声学参数.....	31
第四章 声图仪	43
第一节 声图仪的应用.....	43
第二节 频谱仪.....	48
第三节 频率分析仪.....	50
第四节 数字语图仪.....	57
第五章 Y 512 声图仪	69
第一节 性质及用途.....	69
第二节 仪器构造.....	69
第三节 技术指标.....	70
第四节 工作原理.....	72
第五节 仪器的电路系统.....	73
第六节 仪器使用方法.....	74
第六章 声图在喉科的应用	79
第一节 正常人的声图.....	79

第二节	艺术噪音的声图·····	83
第三节	病理噪音的声图·····	97
第四节	喉癌和喉手术后发音的声图·····	130

第一章 绪 论

声图(sonagram)是将人喉发出的声音信号记录下来,经过电处理转换而成的描记图形,藉此图分析声音和话声。经仪器描记的声音图谱所反映的是声音的物理特征,它是一种客观检查方法,但不能代替人用听觉辨别声音音色及其响度的生理作用,故声图检查应结合主观检查法使用。

声谱的研究始于本世纪40年代,美国贝尔实验室的老贝尔(A.G.Bell),希望给他那听觉困难的妻子安装一部特殊的“电话分机”——可见言语转换器(visible speech translator),它可将言语的谱型显示在阴极射线管上,借此她可看懂他人所讲的话。不久他的希望终于实现了,1946年该实验室发表了言语的声谱描记技术的报告;1947年实验室的Potter等发明了声谱仪,并著有《可见言语》(《Visible Speech》)一书,使实验语音学的研究进入一个新的时代。利用声图仪显示的谱图研究因声音嘶哑所造成的声学改变,以鉴别各种喉科疾病,已成为近年的一种新技术。它用声学仪器和方法抽取各种嗓音中共有的和特有的参数,如基频、强度、持续时间、各共振峰等的异常变化,再将这些参数与临床相对照,寻找其中的关系。由视觉和分析(波型和数据)代替听觉,作出客观的记录和判断。但对喉科疾病,目前只能判定有无疾病而不能作出疾病的准确诊断,即使患同一种疾病因其程度不同、个体差异等会呈现不同的参数和图型。这些都有待在发展中完善。

喉科中用于客观分析声音和语言的常用的仪器是声图仪(sonagraph)和声谱仪(sound spectrograph),其构成分3部分:①录音装置,将原始信号记录在磁带上;②外差式频率分析装置,将所记录的信号以45Hz(狭带)或300Hz(宽带)进行频谱分析;③显示装

置,将分析的结果显示在荧光屏上或记录在专用记录纸上。显示的图形有两种:一种是时间—频率—强度的三维图形(语图);另一种是在某一时间断面上的频率—强度的二维图形(声谱)。测定时用元音、字、句子均可,但为了测定准确,根据仪器的特性,可连续发较长时间的声音,一般每份声样需 3min。这样才能较正确地分析嗓子的功能。

70年代以来的研究表明,嗓音的声音分析可以提供有关喉或嗓音疾病病理改变的第一手资料。嗓音可因喉的功能性或器质性改变而发生变化,这是因声带振动方式的改变所引起。声学反射引起经声门的暂时性压力改变而使声带不规则的振动,致嗓音信号紊乱。1970年 Iwata 和 Leden 首先应用声图(sonagram)和声纹(voice print)客观地估价各种不同程度的声音嘶哑的声图。Nessel等的研究发现,声音嘶哑的主要声学改变有:①元音的主共振峰中混有噪音成分;②乐音声能消失;③基频无规律性。

在正常人中,男、女、儿童的声谱基本特征都是一样的。波形的周期性非常规律。女性频率较高,声谱上的波纹较密。耳语声时声门不靠拢,都是气流摩擦声,因此声能集中在高频区。嗓子疲劳时或有痰时,在有规律的波形中,常会突然出现非周期的低频波,但很快就可恢复正常。训练有素的职业用嗓者,常有其特定的声能集中区(歌唱共振峰),歌唱时的颤动(vibrato)则表现为有规律的波浪形起伏。这都是正常的变异。

Gruetzmacher 曾用声谱分析法作过 1 040 例声音普查,结果认为此种纯物理的声音检查,对各人的音色特点、呼吸方式与发声技巧是不易测知的,而这些正是临床上想要借助仪器以预测声病的内容。

声图分析的要点是:①非周期性的波型(表示声门开闭不规则);②声谱中高频区能量增加(表示漏气,闭合不佳);③共振峰遭到破坏(形态和能量)。根据谱图的特点,Yanagihara将声音嘶哑按照不同的程度分为 4 型,此分类法获得的声音嘶哑的客观程度

与主观程度非常一致,故曾被广泛应用,但其并不完善,主要是它只考虑到噪音成分而没有考虑其他方面,限制了临床应用,尚未脱离对声图的主观估计。

声图的实用价值是:①正常人图型的鉴定;②声乐工作者谱图分析以判定声部,早期声病的辨识等;③估计喉发声重建手术的效果;④估计声带小结、声带息肉等良性病变的程度,手术前后的对比;⑤估计喉癌各型对声门区侵犯的程度,为手术提供指征;⑥对声带麻痹病程的预估及治疗方式的选择提供参考;⑦检查药物对甲状腺功能减退或重症肌无力患者的疗效,或用注射氯化钾期隆前后的声谱变化来帮助诊断重症肌无力;⑧对功能性发声困难的病程作出客观判断。故有人认为声图用于发声检查与电测听仪用于听力检查,具有同样的作用。

人声基本频率(简称基频)的测定是声图基本的测定项目,此为声带振动的最低固有频率。不同频率声带振动的形式不同,一般低音或真声在 200 Hz 左右,高音或假声在 400 Hz 左右。男女基频不同,男子基频在 120Hz 左右,女子在 200Hz 左右。1959年Mysak 曾提出 80~92 岁男子的讲话基频较 65~79 岁组高。女子成年后变化不大。很多学者否认妇女有“老年嗓音”的存在。

声纹为声谱分析仪显示人声谱图的一种方式。此一术语最早出自美国,当时想通过比较嫌疑犯的语言声学的图型能象指纹那样达到破案的目的,但至今未能如愿。1970年Iwata等首次用声谱仪(KAY-6061A)联接轮廓显示单位(KAY-6070A),用 H-S 仪器复制控制强度值。为了提高分析声图的正确性,对受试者每人录制两种不同的声谱图型,第一种是 0~8 000Hz 内所有频率范围的声纹,第二种是 0~2 000Hz 内的声纹,详细地显示了第一共振峰中的声能分布。图中的深和浅代表声能的高与低(0~6dB),越浅,则强度越低。为了进行比较,用 H-S 仪器获取同样的声音信号,按 Yanagihara 分类法分析语图,研究结果显示正常人的轮廓线排列非常稳定而有规则,最大声能集中在第一共振峰区。病变时,

I型轻度声音嘶哑者, 声谱图显示第一共振峰区里有少量噪音成分, 但基频及乐音成分正常有规律性; 声纹图在 4 000Hz 以上的共振峰中声能分布轻度不规则, 然而在第一共振峰中声能的分布很有规律。II型中度声音嘶哑的声纹图可见声能不正常分布, 主要集中在 2 000~5 000Hz 之间的中频范围, 第一共振峰上缘呈锯齿形, 峰带增宽, 且有不规则的中断。III型重度声音嘶哑的声纹图示全部共振峰范围内声能分布不均匀, 尤其是 6 000Hz 以上的较高频范围, 而较低频范围的共振峰有这样的特点: 第一共振峰的上下边缘明显不规则, 第二共振峰中声能呈岛状分布, 第一共振峰加宽有几段与基频混合。IV型严重声音嘶哑的声纹图示: 第一共振峰带异常增宽, 多处与下面的基频及上面的第二共振峰相连, 整个声能杂乱呈岛状分散在频谱范围内, 共振峰难以分清; 而另一 IV 型的声带麻痹患者, 他的声纹图都呈现相反的图示, 出现声能分布较规则, 各共振峰清楚, 只是声能的对比度降低, 第一共振峰中的能量较弱, 它的扩大显著减少, 尤其是上缘。以上的声纹图特点有别于声谱图, 但作为喉科疾病的诊断依据尚需做大量的研究工作。

人嗓音的周期不是恒定不变的, 它的微小变动就称之为振幅和音调的微扰因子(amplitude and pitch perturbation factor)。在喉部病变时微扰因子增高, 这是因为声学信号中那些微扰与声门区的功能改变一致, 而音调的微扰因子参数是用以测定声谱中相邻的基本周期间的变化(包括变化的类型、频率、数量和大小)。Lieberman(1961, 1963)连续研究微扰因子及嗓音病理, 将音调微扰参数的正常值定为 $>0.5\text{ms}$ 。他认为这是声门处气压短暂的变化(降低)所产生, 又可因声道形状的复原而恢复。喉部病变者音调干扰较大, 声带损害的程度越重, 音调干扰越大, 因为声带损害影响声门的关闭。

以后, Risberg 用瑞士语和英语, Saito, Sugimoto, Koshihawa 等用日语, 经接触麦克风拾取信号, 都取得了类似的资料。1965 年 Moore 用声学分析和高速电影对照证实了声音嘶哑越重周期间

频率不同者越多,每个周期所占的照片数相差也越大。1969年Koike提出用主振幅峰值的参数来分析喉肿物、声带麻痹和正常嗓音,结果3组近似,说明病变对主振幅峰值的影响不大。但在喉肿物组中却有40%可见到主振幅之间还有各种不同的短波周期,该现象在麻痹组中未见,而正常组中仅有10%,说明喉肿物可影响声学信号的振幅或称之为声学信号的微扰。

1975年Takahshi和Koike运用频率微扰商(FPQ)和振幅微扰商(APQ)表示嗓音病变时的频率和振幅波动情况。1977年Koike进一步提出在测得FPQ和APQ后再绘制成“临界椭圆”,用临界椭圆的计算方法来测定病理和正常嗓音。

频率颤动(jitter frequency)也是常用的声学参数,它亦可表示音调的微扰情况。实际上是用百分比来表示频率的变化。

1959年Miller首创间接法研究喉部的振动,以后许多学者不断改进,称之为负滤波技术(inverse filtering technique)。因为口声不能完全代表喉部的情况,还包括声道的因素。因此许多学者用放在胸骨柄上气管前的接触麦克风来拾取声音信号(residue signal),将有利于准确地测定声源。其优点是:不受喉部以上发声吐字结构的影响;波形简单;如有环境噪声,则信号/噪声比变化明显。

我国学者采用声图仪(或称语图仪)研究艺术嗓音和喉科疾病,近年来取得不少成绩。顾立德(1984)应用连续式语图仪研究专业歌唱及戏曲演员60名,证实听感特征与声图所见一致,并发现我国民族歌声的歌唱共振峰所处频带较高,某些假音中各谐波所载声能规律的强弱相间,以及半真假艺术嗓音的声图特征等。商泽民等(1984)应用声图仪和频谱仪对照,研究声乐工作者86人,对音色个性、共鸣运用、声部鉴别、民族与西洋唱法区别等进行客观分析,认为对指导艺术演唱和教学具有重要的实用意义。黄平(1984)应用BP6A型声频频谱记录仪分析艺术嗓音的组成成分,认为在声音的定性、定量方面,较声图仪更为具体而准确。包紫薇等

(1984)应用频谱研究各种歌唱嗓音的特征,通过频域和时域的观测,认为在音色听感和声学参数之间找到了一定的对应关系,并发现美声唱法和我国几种唱法的歌唱共振峰分布的差别很典型。吴卫彬等(1984)对100多名歌剧演员和戏曲演员进行录音,后来用频谱仪、窄带实时分析仪、积分声级计等进行分析,结果认为对演员歌声分部可得到满意的结果。鲍怀翘(1987)报告嗓音频谱及声门波形和病态嗓音的声学特征。周敏好、蔡钺候(1987)应用我国首创的声谱实时伪彩色编码显示仪(该仪器出现的是彩色声谱图,除时间和频率二维图象显示外,把声强级分为10级,用伪彩色编码显示,并有动态伪彩色分析、空场分析、编码打印及数据储存多种功能),抽取正常嗓音30例和各种喉病的声样进行分析,显示客观分析效果良好。杨式麟等(1986)应用频谱分析仪检查841名正常人和100名喉病患者,发现正常人谐音清晰,无噪声;儿童、女性、老年人基频有不同的改变;并将病态嗓音频谱的噪声按轻重分为5度。该氏(1987)应用2031型窄带频谱分析仪检测17~22岁男女评剧学员发*i*音基频(F_0)及第二共振峰的频率,与相应年龄的正常人(非歌唱者)作对比,发现学员男子组或女子组的真假声音域皆高于相应年龄组的非歌唱者。作者自1980年以来,采用语图仪对正常人、声乐演员、各种喉科疾病,以及喉癌手术前后的大量声样进行分析,认为:客观记录分析发声样本是当前喉科学的一项必不可少的手段,随着实验仪器的改进,不久的将来它将成为喉科检查的常规设备。

第二章 声学概论

第一节 物理声学

只有对物理声学的有关基本原理和知识有所了解,才能对声图作出正确的分析和判断。以下将介绍有关的基本知识。

一、声振动过程中的物理量

一切谐和振动过程都表现为两个互不依赖的量——振幅和频率。

1. 振幅 声振动的幅度。是指声压、位移、振动速度或相应于这些量的声强度的幅度。

声振动的幅度引起对于某一声响度的感觉。由此可知,响度是对已知振幅在听觉上的一个估计量。研究听觉的量时,常引用声压或其相应的声强的有效值而不用其幅值。

2. 频率 声音每秒振动的周期数,俗称音调。声振动的频率产生一定高度的声音的感觉称为音调。一个固定频率的声振动相当于一个单音或纯音。频率越低,人听觉感受到的音调也越低。低的音调可称为低音。声振动的频率增加,听觉将感受到音调提高。

3. 基频和谐波 话声、歌声、乐器声的声源所产生的声振动都是复杂的声振动过程。复杂的周期振动过程 $\vartheta(t)$ 可以用谐函数的级数形式代表:

$\vartheta(t) = \sum_{n=0}^{\infty} A_n \sin(n\omega t + \varphi_n)$, 式中 A_n 和 φ_n 为级

数中各项相应的振幅和相位, n 为任意整数。此级数可能是有限的,也可能是无限的,视函数 $\vartheta(t)$ 的形状而定。级数中各项的整体称为已知函数的谱。级数中的每一项称为谱线。表征级数中每一项

的谐振动称为谐波。第一个谐波的频率,即级数中 $n=1$ 的一项,称为基本频率或基频,即: $\omega_1 = \frac{2\pi}{T}$; 以次的谐波频率将是: $\omega_2 =$

$2\frac{2\pi}{T}$, $\omega_3 = 3\frac{2\pi}{T}$, 由此可见复杂振动的周期 T 越大,则两个

相临谱线间的频率间隔越小。谐波频率依数学的自然级数律增加。与此相应,谐波则以顺序的号数来表示,如第二次谐波的频率2倍于基频,第三次谐波的频率则3倍于基频,依此类推。

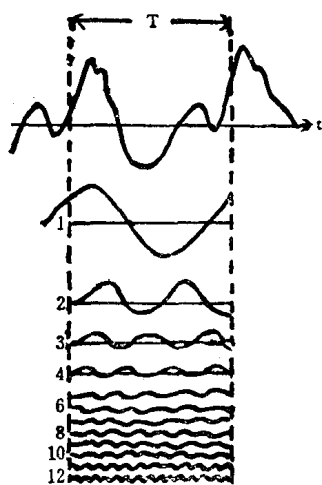


图1 声的谐波

图1以曲线表示一个复杂周期性振动过程的谐波成分,直到第十二次谐波。有趣的是,一个训练有素的习惯性音乐听觉的耳,有能力分辨在一个复杂声音中的谐波成分(纯音)。

二、声图分析的有关概念

1. 噪声(noise) 有时也称无调声,它是很多紊乱、断续、不协调的基音和它们的谐音一起形成的声音。噪声可能只包含一个纯音,也可能包含两个或更多的

分音。如果它包含了许许多多均匀地分布在某一频带范围内的分音,那么这些音可能非常靠近,以致实际上不能加以分解,这样的频带称之为“连续频谱噪声”;如果频率上下限度已知,则应该加以注明,否则可用一般名词来说明,例如“高频”、“中频”、“低频”或“全声频”等。

噪声有时也指电讯号中任何不需要的成分,故又指不需要的声音,如电干扰,既使这些电流并不一定转化为相应的可听声音。在这种情况下用“电噪声”为妥。

噪声的响度级以“昉”为计算单位。

(1) 无规噪声(random noise): 瞬时值不能预先确定的声振荡。无规噪声的瞬时值对时间的分布只服从一定统计分布规律。

(2) 白噪声(white noise): 用固定频带宽度测量时, 频谱连续并且均匀的噪声。

(3) 粉红噪声 (pink noise): 用正比于频率的频带宽度测量时, 频谱连续并且均匀的噪声。

2. 共振峰(formant) Helmholtz (1862) 对语言声进行广泛研究后指出: 特殊元音是由一特殊的共振系列表征的, Hermann 称之为“共振峰”, 意指在复音频谱中一些谐波幅值较大的范围。语言共振峰是由声带振动发声再经口、唇、舌、齿等共振腔的控制呈现共振而产生的, 共振峰中心频率是共振峰频率。涉及语言声学范畴, 容后再述。

3. 频谱(frequency spectrum) 把时间函数的分量, 按幅值或相位表示为频率函数的分布图形, 即幅值片段图。在谱图上, 时间在横轴上显示, 频率在纵轴上显示, 强度则由图案痕迹的明暗度表示。根据声音性质的不同, 声谱图可能是线谱、连续谱或两者之和。线谱是由一些离散频率成分形成的谱, 连续谱是在一定频率范围内含有连续频率成分的谱。

4. 声强(sound intensity) 在某一点上, 一个与指定方向垂直的单位面积上, 在单位时间内通过的平均声能。单位为瓦每平方米 (W/m^2)。通常声强是指声音的强弱。声音的强弱与振幅的大小成正比, 也就是说振幅越大, 声音越强; 振幅越小, 声音越弱。为了计算方便, 用对数算法称为声强级。计算方法为: 以声强与基准声强(常用基准声强为 $1PW/m^2$) 之比的以 10 为底的对数乘以 10, 单位以分贝 (dB) 计。分贝是贝尔 (bel) 的十分之一, 贝尔是一种级的单位。在声强图上声强的大小以谱线灰度的深浅来表示, 灰度越深, 声强越大; 灰度越浅, 声强越小。

第二节 生理声学

噪音的产生是一个复杂的生理过程，它系由中枢神经的控制系统和末梢的发声、构语、共鸣器官等形成。

人类的发声器官系由动力、振动、共鸣、构语 4 部分组成。肺和有关的呼吸肌成为发声的动力器官，喉和声带因呼出的气流使声带振动而发声成为振动器官，共鸣器官是指喉腔、咽腔、口腔及鼻腔和副鼻窦等，统称为声道，构语器官由口腔、舌、唇、齿、软腭等组成，声音通过它们的各种协调动作而产生语言。语言和歌唱是发声器官各部分在中枢神经的控制调节下的一种复杂的运动过程。

一、噪音产生的中枢活动

噪音虽偶可在无意中或反射性产生，但它主要是人类社会中以交流语言的工具。噪音也用于艺术活动领域中，包括歌唱和戏剧演出等。可见人类的行为和发声是后天经过学习训练而发展的，是由大脑所控制的。

图 2 示噪音产生的中枢活动过程。讲话和歌唱时，高级中枢，包括脑皮层语言中枢，首先决定声音产生的顺序。高级中枢包括语言和艺术活动中枢，从高级中枢发出指令，传递至位于中枢前脑回的运动皮层，运动皮层再发出一系列指令达脑干和脊髓的运动核，轮番地传递指令达呼吸肌、喉肌和发声肌肉。椎体外系统，包括脑皮层的若干部分、小脑和基底神经节给呼吸、喉和发声肌肉以附加的微细调节。

肌肉活动的结果使发音器官运动，从而产生噪音，再传至收听者的耳。噪音同样也传至讲话者或歌唱者的耳内，并产生声音的反馈。反馈机制也经深或浅感受体起作用，它供给肌肉收缩和发音器官运动的信息。心理状态能影响整个发音过程，自主神经系