

厅堂音质设计中的评价问题*

王季卿

(同济大学声学研究室)

厅堂音质评价是设计的依据,而评价的参量应该具有明确的物理意义,并与主观反应有密切的联系。本文对此作了评述,并结合本人工作经验,对一些实际问题进行了探讨。

一、引言

厅堂音质的评价决定于听音者的判断,是主观的属性。但作为工程上的厅堂音质设计服务,就不仅要使这些主观的音质评价与客观的物理量取得某种联系,而且还应提出与工程设计有关的变化量,使厅堂音质设计能遵循一定的规律进行,才能满足使用要求。

混响时间 T_{60} 之所以成为厅堂音质评价重要指标,不仅在于它与主观听音中的干涩和丰满、清晰和浑浊等等密切相关,而且掌握了 T_{60} 与大厅体积 V 和室内吸声量 A 的变化规律($T_{60} \sim \frac{V}{A}$),即改变 V 和 A 可以达到调节 T_{60} 至合适程度的目的。而目前其它评价指标还不具备这样明确而密切的联系。

近年来,在电声器件和乐器方面也开展了大量的音质评价研究和评比活动,但厅堂音质评价还有其自身的特点:(1)同一个大厅,由于厅内位置不同,音质可以有较大的差别;(2)演奏者受大厅音质的影响,例如演奏是否舒畅,各演奏者之间的相互听闻条件影响到协奏能力等;(3)受到干扰声的限制,包括外来噪声和厅内回声、聚焦等;(4)互比条件较为困难,各厅之间和同一厅的不同位置之间的互比均存在一定的困难,因为依

靠记忆来判断很容易失误。

二、评价参量的演变和进展

混响时间对厅堂音质的重要性已无需赘述,而且也早已认识到它决不是厅堂音质评价的唯一参量,所以这方面的进展始终受到音质设计者的关注。至于作为反面作用的参量如回声和噪声等,均是干扰因素,本文不拟多谈。

1. 传统的混响时间 T_{60} 定义是声音达到稳态后停止发声,其平均声能密度自原始值衰变60分贝所需的时间。但是实际厅堂测量中,往往只计量衰变过程的-5分贝到-35分贝范围,然后推算出 T_{60} 。事实上在连续变化的语言和音乐节目中,觉察到的衰变过程范围还要小得多,除非在片断的结尾。再则如果衰变过程随时间不按线性下降,即实际上常遇的衰变曲线往往是弯曲的,那么究竟取哪一段衰变斜率最能代表主观感觉上的混响感呢?于是有以(1)“等效混响时间” T_{160} 来表示混响曲线中前160毫秒的斜率,(2)根据衰变过程中开始下降15分贝的“初始混响时间”乘上4,或是(3)按开始下降10分贝的

* 本文系在“厅堂音质学术讨论会”(1983年9月7—10日,北京)上的报告。

“早期衰变时间”(EDT)乘上6等等指标的提出。目前趋向于最后的一种指标,认为它与主观印象最为符合。无论 T_{60} 或是EDT,都不考虑随声源和听者在室内的位置而变化,因此要评价音质随位置不同而可能出现的变化时,就显得不足了。

2. 五十年代起受人瞩目的室内音质评价第二参量研究中^[1],室内声漫射(扩散)的重要性,原则上是认识了,但涉及具体设计,迄今还处于半定量阶段,往往只得凭经验从事。

3. 差不多同时,早期反射声的重要作用亦很受重视。因为直达声到达听者之后几十毫秒以内的反射声,对音质很有帮助,它可以增加响度和清晰度。最早的定量研究结果是Haas(1951)提出的,后来Thiele(1953)根据延迟时间 τ 小于50毫秒的反射声为有用的Haas这一假设,首次提出了一个定名为“明晰度” D 的评价指标:

$$D = \frac{\int_0^{\tau} I(t)^2 dt}{\int_0^{\infty} I(t)^2 dt}$$

即包括 $\tau=50\text{ms}$ 早期反射声在内的直达声能占整个混响声能的比例,将是一个重要的评价参量。许多人认为楼下正厅前排(贵宾席)音质不佳的一个主要原因即在于这个 D 值偏低,即缺乏早期反射声。由于房间的几何形状对这一指标十分敏感,因此早期反射声可在厅堂的平面和剖面设计中用作图求出。

以后,由此而衍生出了许许多多作不同程度修改的评价指标^[2]。本文只着重讨论下列三个方面:

(1) 认为有用反射声到有害反射声之间没有“一刀切”的界限,而是渐变的连续过渡。因此在处理早期反射声的能量时还应乘上一个与延迟时间有关的计权函数。Beranek则具体地提出了它们之间的数值关系^{[3][4]}(见图1),如对交响乐可从20ms的最佳值(满分)算起,一直考虑到70ms(不起有用效果,零分),对歌剧则从24ms考虑到85ms。对语言有

人更把有用反射声的延迟界限推至95ms^[5]。

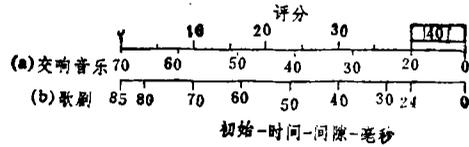


图1 听者接收到的第一个反射声与直达声之间的间隔(初次-时间-延迟间隔,毫秒)与音质评价(满分得40分)的关系

(2) 对交响乐来说,早期反射声的到达方向和“空间感”关系很大。由于人耳对侧向来的反射声最为灵敏,也正是空间感所最需要的。所以有人把它定义为侧向早期反射声能和非侧向早期反射声能之比作为衡量空间感的指标^[6]。它实际上起到使乐队声源有表观展览的效果,故又称为“立体感”。(过去我们也用过“环绕感”一词,表达这种评价的其它名称不下七、八种之多,在此不一一列举)。根据这一观点,于是有大面积侧向悬挂反射板之出现,并使厅堂设计面目为之一新。

(3) 为了补偿直射声沿听众席出现的低频衰减低谷,和高频声随距离有较大损耗等影响,在早期反射声中还应对频率有所考虑。这方面的实践经验已有报导^[7]。

上述诸方面内容都可用脉冲-时间响应的测量来进行。由于信号处理手段的计算机化,使数据资料的积累推进了一大步,这一评价参量当可更趋于成熟。

4. 音质评价是一个多元函数关系,各因素之间不是完全独立的参量,所以应该看到评价指标方面存在的复杂性。例如,为了创造“立体感”,便要获得良好的侧向早期反射声,这样必然使声场扩散程度下降,丰满的混响感会使清晰度下降等等。

当然简单一点的办法是,把各项评价指标看成独立的因素,然后规定各指标在总评价中所占比重来加以综合。Beranek^[4]曾提出过一套完整的评分设计办法,对音质有利

作用的8个评价指标看成独立的因素，并分别给出最高得分额(见表1)，如果出现不利因素包括回声、噪声、失真和厅内不均匀性等时，则要扣分若干，最多可扣除50分。然后从总得分来评定一个大厅的音质。90—100分者评为A+，优异；80—89分者评为A，很好到优异；70—79分者评为B+，好到还好；60—69分者评为B，还好到好；50—59分者评为C+，还好；50分以下者评为很差。

表1 8项音质评价指标的评分方法

	最高得分	
	音乐厅	歌剧院
亲切感	40	40
活跃度	15	15
温暖感	15	15
直达声的响度	10	10
混响声的响度	6	6
平衡和融洽	6	10
扩散	4	0
协奏能力	4	4
合计	100	100

这是一种很实用的方法，也便于设计者知道如何着手达到“高分”的各种途径和措施。不幸的是 Beranek 根据自己提出的方法来进行的纽约 Philharmonic 音乐厅音质设计，按他的评分方法当在A以上，却遭到意想不到的失败。这至少说明评分标准应该调整，8项指标以外是否还有重要的遗漏也值得考虑。

三、音质评价工作中的一些经验

五十年代后期开始，我国在语言清晰度方面开展了较多的基础研究，以适应厅堂和通讯、广播等多方面的需要。本文只想着重地谈谈欣赏方面的厅堂音质评价工作经验。

长期来国外在谈到音质评价时，往往较多地探讨表演古典交响乐的音乐厅(包括录音室)设计。一是由于音乐厅对音质有苛求，

二是国外交响乐演出较普遍。所以不能用这些指标来作为我们设计的标准。就以国外来说，也有人对于多功能大厅的特点作了调查分析，与常见的音乐厅指标相比(见表2)，两者之侧重点也根本不同。

表2 多功能大厅与音乐厅的音质评价比较

Beranek 提出的音质评分方法(1962年)	
初始时间间隔(早期反射声) (强调浮云式反射板的作用)	占总分40%
混响时间	占总分30%
响度	占总分16%
Long对多功能大厅之调查结果(1980年)	
喜爱声音的力度	占调查者47%
明晰清楚 (强调体积小些，有强的侧向反射声)	占调查者28%

我们在1964年前后曾组织了上海影剧场音质的调查*，着重考虑了下列几个方面的内容：

1. 响度 在观众厅噪声普遍较大的实际情况下，响度要求就非常突出。当然这项指标也涉及室内安静措施的重要性。

2. 清晰度 戏曲和电影都有语言可懂度问题。

3. 混响感 各种带有音乐感的节目都要求中高频活跃，低频温暖。

4. 其它 对音乐戏曲表演还要重视亲切感、音色的平衡等；对电影则自然度较为重要。

当时，曾试图找出评价的重点和它们的最佳值范围。为了使调查工作定量化，我们采用了属于非参数统计中的“等级序列”(Rank order)评分方法，用中位数(或称中值)来说明一种趋势。有关评分标尺见图2。全厅分成几个区分别进行评定，然后汇总，评价时，一律不用扩声系统。

同一大厅内各区的音质会有差异(见图

* 参加这一工作的还有上海民用建筑设计院及上海市文化局、电影局等单位，和上海地区的许多声学工作者。

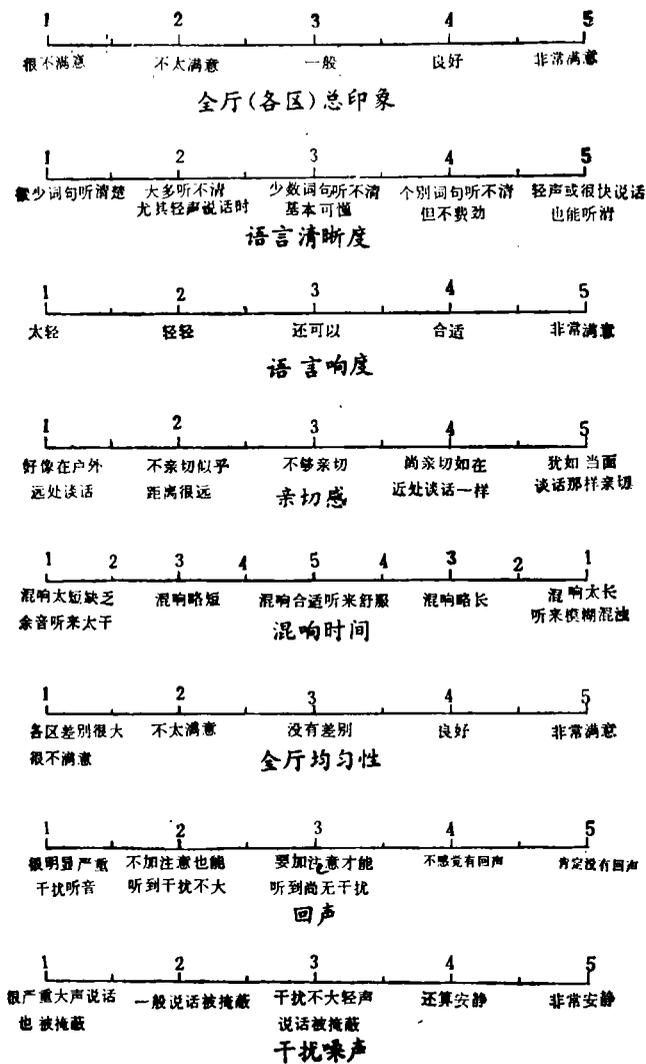


图2 各项音质评价指标的评分标尺(允许以0.5作为评分进位)

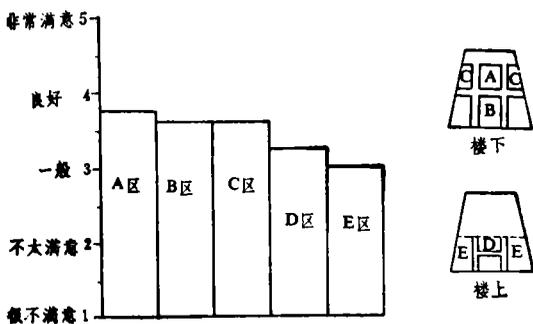
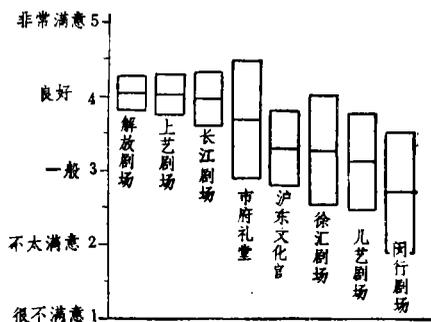


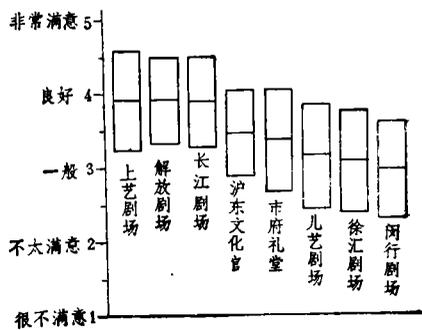
图3 一个厅内各区总印象的比较

3),即使同一区内不同评价人之间的差异也可能不小,这是可以预料的。但是由中位数的统计结果,可以看出各厅堂音质的优劣次序和其中存在的问题及严重程度。图4所示为8个厅堂演出话剧时的评价结果。其中(a)为全厅总印象评分,(b)为全厅各区总印象评分的汇总,两者基本相符。

对于评价人员,我们并未进行特殊训练,只是对评价标尺的使用向全体评价人员作了充分解释,并组织讨论以统一认识。人员来自三个方面:录音部门、电声制造系统和建声、土建设计部门共约40人,



(a)



(b)

图4 8个剧场演话剧时的评价结果(中位数和偏离度)
(a) 全厅总印象 (b) 全厅各区总印象的平均
1. 解放剧场4.03(3.90) 5. 沪东文化宫3.31(3.45)
2. 上艺剧场4.02(3.91) 6. 徐汇剧场3.28(3.07)
3. 长江剧场3.93(3.88) 7. 儿艺剧场3.12(3.16)
4. 市委礼堂3.70(3.37) 8. 闵行剧场2.71(2.95)
以上所列数据为中位数,括号内为图(b)的结果。

每次出席人数有一半以上。考虑到职业性的偏爱和听音经验，在统计结果时将前两方面人员列为一类。事实上这两类人员之间的评价差异小于本类人员之间的偏离度，而且在判断优劣次序方面也是基本一致的。图5所示为选自一个剧场的评价资料。

以上情况说明这一方法是可取的。至于这些评价如何与厅堂声学设计联系起来，有待深入研究。调查的初步结果也表明了一些设计上的问题。例如这8个大厅中，A区音质普遍认为是全厅各区的最好者；尽管通常认为“贵宾席”音质差，但在本次调查中则未发现。B区为挑台下的正厅后座，评价为最差，故今后设计应特别注意开口比和挑台下的反射等措施。

有关这次调查结果的详细资料见参考文献^[9]。

四、音质评价中的容忍度和模糊性

音质评价终究是一种主观属性，而且是多因素的综合判断，因此很难有一个划一的指标。但作为设计的依据，总得要有一个数量概念和最佳范围。如果这一范围很大，则降低了控制指标的意义。再说多因素中有主有次，但互有影响，因此不能强调一点，不及其余。音质评价之复杂性也在于此。下面来谈谈评价工作中的容忍度和模糊性。

1. 容忍度 以我们最常用的混响时间来说，它的最佳值一直认为是随用途和房间体积而异。但早在1954年Kuhl就根据大量调查提出，演出交响乐的录音室，当体积在2000~15000m³范围内的最佳混响时间几乎是不变的。例如对浪漫派交响乐均应在2.1秒左右。这和常用的最佳混响时间曲线有很大差异。甚至有人还认为室内扩散很好时，混响时间可容许有100%的差别！这里就提

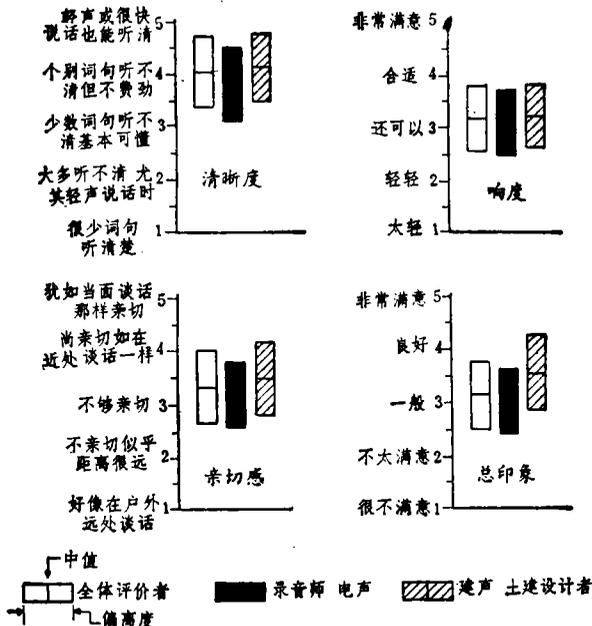


图5 两类评价人员对儿童艺术剧场的评价结果比较

出了人们对最佳混响时间的容忍度究竟有多大的问题。

从20多年前纽约林肯中心新建Philharmonic音乐厅所发生的事件来看，有些问题也是令人费解的。这样一座重要的厅堂，在正式开幕之前组织了为期一周的调试活动，包括声学测量和试听评价。参加者有著名指挥、音乐评论家、富有经验的声学工程师以及几个音乐团体的代表等组成的高水平听音评价队伍，在试音周内演出了30个不同类型的音乐作品。演奏者也都要求填表发表意见。室内布置的调整最后是这样定下来的，可以说经过专家“审查”通过的。在试音周中，专家们曾要求加强低音，并调整到被认为满意的效果。从测量结果来看也是符合设计要求的。实测混响时间 $T_{125} = 3.0$ 秒， $T_{500-1000} = 2.2$ 秒。但是开幕后不到半年，舆论哗然，批评尖锐。其中反应强烈的一个严重缺点是低音不足。这里不打算讨论造成这一后果的原因，而是要问，为什么这一明显缺点能“混过”一大批专家的评判？如果说容忍度较大，那又为什么公众对此反应如此强烈呢？

再如作者本人也有过这样的经历。漕河泾上海音乐学院礼堂是1954年建成的。该校从江湾迁入新址后,几乎有半年的时间里教师们一致反应声音太糊(混响太长),无法使用。但是后来1958年该校又迁入市中心区时,他们又一致为失去这座音质如此优美的演出场所而惋惜。作为听音评价训练有素的音乐家们,其判断竟然可以从一个极端变到另一个极端!

有人认为^[8]:如果一个音乐厅在风格上超前了一些,或是与大数人的习惯非常不同,则它将一定会有一个“过渡”时期,或者说“成长”时期。

由此可见,人们听音的容忍度有时很宽,而有时又很苛刻,其中必有缘故。

2. 模糊性 上面介绍的音质评价评分方法(见图2),在某种意义上说是带有任意性的,所用评语也缺乏严格的定义,各人的体会也就可能有较大的出入。所以,虽然用了等级评分制,但并不存在绝对的界限,而只有相对的意义。这是一种语言变量,描述了程度的差异,在此用数量来表达总是有很大局限性的。

另外,在评分时也包含了不确定性因素,因为1至2,2至3之间并无明确界限,是有连续性的,而且并非等距的。即使两个极端语变量,例如1分和5分分别代表“很不清楚”和“非常清楚”的评价,但它们之间的界限也不是一刀切的,其中存在着相当的过渡区。对于这类问题,自然不能用二值逻辑的计算程序来表达。这样一组评价结果是属于“模糊集合”,所以只能用近似推理——模糊逻辑来处理。有人认为这部逻辑很可能比经典的二

值逻辑更接近人类决策过程中所包含的逻辑。

再以听得清楚与否为例。造成清晰度高低的因素很多,例如响度不够、混响干扰、语音失真等等。因此用它来评价厅堂音质本身就具有模糊性。这就要求采用模糊数学方法去进行多因素的综合评价。另外还要看到模糊事件在各人评价中的界限虽不相同,但总有一定的分布,这里可用一致性函数来表达这种分布的规律。可以预料,这种新的处理方法反应了评价工作的综合性,使评价工作考虑得更全面一些。

参 考 文 献

- [1] 马大猷:“广播中的声学问题”,《电信科学》,1956年,2期,58~61页。
- [2] 王季卿:“厅堂音质发展概况”,《建筑声学与环境控制技术资料选编》1—5页,华东地区建筑设计标准化办公室,1981年10月。
- [3] Beranek, L. L. —Music, Acoustics and Architecture. John Wiley & Sons, Inc., 1962. 421—423 页。
- [4] Beranek, L. L.—音乐的音质评价名词,《声学译丛》总第10号;《建筑声学》(2),22~25页,上海科技翻译馆出版,1963年。
- [5] Lochner and Burger,《Acustica》, Vol.10(1960) 394; Vol.11 (1961), 195.
- [6] Barron,《J. Sound and Vibration》, Vol.15 (1974), 475.
- [7] 王季卿:“国外音乐厅声学设计中的一些新设想”,《声学学报》,1981年,2期,123—124页。
- [8] Lanier, R. S. —“费拉第尼克音乐厅发生了什么事?”,Arch. Forum, 1963年12月号。中文介绍见王季卿、戴复东:“大型厅堂中透空的悬吊平顶反射板的声学性能及其实用价值”,42—48页。同济大学科技情报站资料F66009,1966年2月。
- [9] 徐之江、刘逞鹤:“上海市剧场音质主观评价”——厅堂音质学术讨论会”资料(1983年9月7—10日,北京)。