

语言传输指数直接与间接测量方法实验对比

祝培生¹, 莫方朔², 康健³

(1. 大连理工大学 建筑与艺术学院, 大连 116023; 2. 同济大学 声学研究所, 上海 200092;

3. 谢菲尔德大学 建筑学院, 英国 谢菲尔德, S10 2TN)

摘要: 通过对 4 个不同房间内 12 个测点共 48 个听音条件的测量, 分析了直接测量方法和间接测量方法测量语言传输指数(STI)、公共广播系统语言传输指数(STIPA)时的差异以及不同频率响应的测量系统对这两种测量方法的影响. 结果表明, 只要采用标准化的测量系统、测试程序, 测量过程严格按照规定执行, 这两种方法测得的 STI 和 STIPA 都能准确反映出语言经过声传输通路后与语言可懂度相关特征的变化, 测量结果都可以用于语言可懂度的客观评价, 差别较小.

关键词: 语言可懂度; 语言传输指数; 直接方法; 间接方法
中图分类号: TU112.1 **文献标志码:** A

Experimental Comparison Between Direct and Indirect Measurement Methods for Speech Transmission Index

ZHU Peisheng¹, MO Fangshuo², KANG Jian³

(1. School of Architecture and Fine Art, Dalian University of Technology, Dalian 116023, China; 2. Institute of Acoustics, Tongji University, Shanghai 200092, China; 3. School of Architecture, University of Sheffield, Sheffield S10 2TN, UK)

Abstract: The speech transmission index (STI) and speech transmission index for public address system (STIPA) metrics at 12 receivers in 4 rooms, with a total of 48 acoustic conditions, were measured using the direct and indirect method and the measurement results of the two methods were compared, with a particular attention on the influence from the frequency response of the measurement system. The results show that, there is no substantial difference between STI and STIPA metrics; both direct and indirect measurement can be used as an objective rating method for speech intelligibility, as long as the standardised measurement program and measurement system are adopted.

Key words: speech intelligibility; speech transmission index (STI); direct method; indirect method

语言可懂度是音质评价的重要指标, 可以用来评价厅堂或扩声系统的声音传输质量. 对汉语语言可懂度的评价, 可分为主观评价和客观评价. 语言传输指数(speech transmission index, STI)方法因能够较好地反映混响时间、信噪比对语言可懂度的影响, 并具有一定的抗系统失真能力, 被 IEC (International Electrotechnical Commission) 标准^[1] 采纳并推荐, 并在建筑声学领域得到广泛应用. IEC 60268-16^[1] 自发布以来共经过了 4 次修订(1988; 1998; 2003; 2011), 2011 修订首次将间接测量方法纳入推荐, 这样推荐的 STI 测量方法一共有两种: 基于信号调制的直接测量方法(以下简称直接法)和基于房间脉冲响应的间接测量方法(以下简称间接法). 关于 STI 的计算模型及测量方法在相关文献^[2] 中已有详细描述.

作为 IEC 同时推荐的两种方法, 虽然理论上可以认为当直接法的测量信号包含调制频率的周期数达到无限多时, 就接近使用脉冲响应计算调制转移函数时的测量结果, 但两种方法是否可以认为完全等效以及影响测量结果精度的因素有哪些, 也是测试人员关心的问题. 由于 IEC 60268-16 推荐间接法的时间还比较短, 国内、外相关的实验对比研究还很难见到.

国内与汉语语言可懂度有关的现行规范有两本^[3-4], 分别对应于 IEC 标准^[1] 和美国标准^[5]. 考虑到这两本规范都有近 20 年没有修订的现状, 将 STI 直接法和间接法按照最新标准进行系统的对比研究具有较好的现实意义. 本文通过对 4 个房间内的 12 个测点共 48 个听音条件进行了测量, 主要研究以下内容: ① 分别用直接法、间接法测量汉语标准频谱的 STI、公共广播系统语言传输指数(speech transmission index for public address system,

STIPA) 并进行对比; ② 不同频率响应的测量系统对直接法和间接法 STI 测量结果的影响。

1 实验方法

1.1 实验设置

本文选取了 4 个不同的房间作为测试场所, 包括一间办公室、一间实验教室、一间多媒体报告厅和一间半消声室(内部放有 1 张办公桌和 4 把椅子)。其中办公室、实验教室和半消声室的平面均为矩形, 多媒体报告厅的平面为八边形, 表 1 给出了 4 个房间的基本情况。

表 1 4 个测试房间的基本情况

Tab.1 Characteristics of the four test rooms

房间类型	房间容积/ m^3	混响时间/s
办公室	108	0.59
实验教室	238	1.70
多媒体报告厅	1 674	0.90
半消声室	550	0.05

在办公室内共布置了 2 个测点, 测点和声源的布置情况如图 1 所示; 实验教室内共布置了 3 个测点, 测点和声源的布置情况如图 2 所示; 多媒体报告厅内共布置了 6 个测点, 测点和声源的布置情况如图 3 所示; 半消声室内布置了 1 个测点, 测点和声源的布置情况如图 4 所示。在这 4 个测试房间中, 接收点的高度都为 1.2 m, 声源的高度都为 1.5 m, 图 1—3 中 S_1 为信号声源, S_2 为干扰噪声声源, $R_1 \sim R_{12}$ 为接收点。

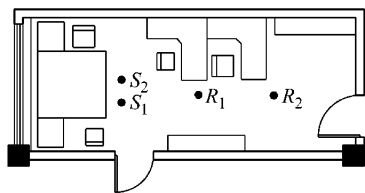


图 1 办公室声源与测点布置示意图

Fig.1 Source and receiver positions in the office

为获得从非常差到非常好的听音条件, 对于 4 个房间中的每个测点, 都设计了 4 种发声条件, 对应于 4 种不同的信噪比。测量时, 在发声位置处布置了两个声源同时发声(两声源相距 0.5 m), 用人工嘴播放测试信号(直接法使用 IEC 60268-16 规定信号, 间接法使用对数扫频信号^[1]), 球面声源播放经过汉语语言频谱调整的粉红噪声。语言信号播音条件按照 IEC 60268-16 要求, 在半消声室内距离人工嘴正前方 1 m 处声压级设定为 60 dB, 同时调整干扰噪声

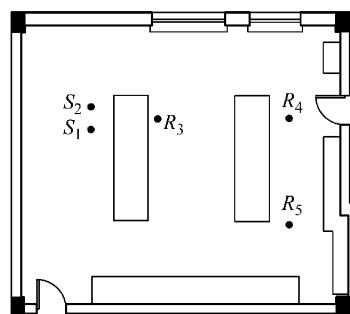


图 2 实验教室声源与测点布置示意图

Fig.2 Source and receiver positions in the laboratory

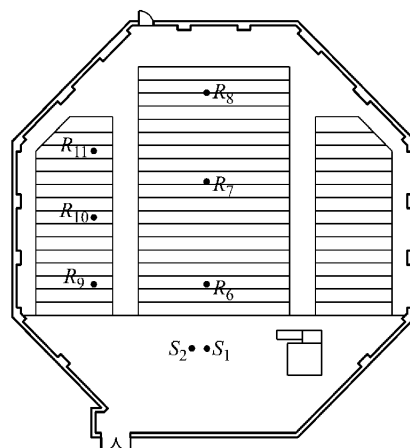


图 3 多媒体报告厅声源与测点布置示意图

Fig.3 Source and receiver positions in the multimedia lecture hall

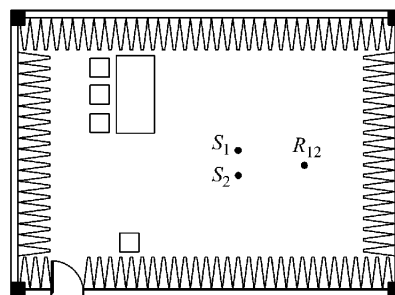


图 4 半消声室声源与测点布置示意图

Fig.4 Source and receiver positions in the semi-anechoic chamber

声源的声压级, 使得在距这两个声源 1 m 处分别对应于 4 个不同的信噪比: -5; 0; 10; 20 dB(这些信噪比并不代表现场测点处的实际信噪比)。然后将消声室内设定好的信号声源和干扰噪声声源摆放到测试房间的相应位置, 依次对各测点进行测量。测量时人工嘴使用的信号共有两种, 一种已经对测量系统进行了频率响应均衡, 使用该信号测量了 4 个房间内 $R_1 \sim R_{12}$ 测点的 STI、STIPA(分别采用直接法和间接法); 另一种没有进行均衡, 使用该信号测量了办公室和实验教室两个房间内 $R_1 \sim R_5$ 测点的 STI(分

别采用直接法和间接法). 球面声源播放的干扰噪声信号经过了频率响应均衡.

除了一些 STIPA 测试仪采用直接法以外, 当前常用的 STI 测量平台主要有 Dirac^[6]、Aurora^[7] 等, 都是采用的间接法. 直接法由于测量时间较长, 数据处理对硬件要求较高, 尤其是完整 STI(Full STI, 在 IEC 标准中建议一般只在科研时使用) 更加难以测量, 因此完整 STI 的直接法商业测量平台很难见到. 为能真实体现 IEC 60268-16 推荐的 STI 测量方法, 并且采用汉语的标准语言频谱, 编写了直接法的测量与计算程序, 间接法测量则采用了 Aurora 测量平台. 间接法没有编写汉语标准语言频谱的测试程序而采用了商用测量平台, 是因为语言频谱对间接法的影响主要体现在后期信噪比的修订中, 后面会对此做进一步的讨论.

1.2 提高测量结果的有效性

在测量过程中, 系统的选用与匹配、声源声压级的标定、测试信号的生成及后期处理都可能带来较大误差. 降低这些误差的关键是采用标准化的测量系统、测试程序, 测量过程也应严格按照规定执行.

实验所采用的测量系统包括信号声源 GRAS 44AA、干扰噪声声源 PYRITE、功率放大器 AMPHION(对应于 PYRITE)、声卡 B&K ZE-0948、传声器 B&K 4189(供电系统为 B&K 1704) 以及录音软件 Audition3.0. 测试前对整个系统进行了回路校验, 以保证其为线性时不变系统, 没有谐波产生.

声源声压级的标定较易产生误差, 进而影响信噪比大小和 STI 测量结果, 因此实验中两个声源都在半消声室内进行了标定, 测试信号也在信号生成时通过滤波对声压级进行了精确控制.

所用测试程序是否有效是测量结果可信的关键. 假定已经根据规范生成了调制深度为 0.8 的测试信号, 然后将该信号直接当做接收信号进行后期处理(回路校验), 得到的 98 个调制转移值并不是 0.8. 造成这种情况的原因很多, 比如所用粉红噪声是伪随机信号; 信号生成时需要滤波、调制; 接收后还要滤波、包络提取等. 这些都会造成一些信号损失. 为判断这个误差的大小, 采用了统计学中单一样本 T 检验的方法. 计算得到的相伴概率值为 0.819, 可认为该 98 个数的均值和指定值 0.8 相比, 没有显著性变化. 98 个 0.8 的调制转移值和 98 个经过回路后的调制转移值还被同时用来计算 STI, 两种情况下的 STI 相差仅为 0.001, 远小于 1 个最小可察觉差

(just noticeable difference, JND), 约为 0.03^[8]. 由此可见, 由测试程序带来的误差是非常小的.

2 实验结果

2.1 直接法、间接法测量结果对比

考虑到当前采用直接法测量的商用仪器主要是 STIPA 测试仪, 商用软件平台基本上采用间接测量方法, 因此选用以下三组数据进行对比分析: 直接法测量的 STI、STIPA; 直接法与间接法测量的 STI; 间接法测量的 STI、STIPA.

图 5 给出了直接法测得的 STI 与 STIPA 的差值(图中横坐标每个测点的 4 个信噪比条件为 -5, 0, 10, 20, 下同, 纵坐标为量纲一). 可以看到 4 个房间的数据没有明显的分布规律, 48 个测试条件的差值都非常小. 这些差值基本分布在 0 轴线附近, 最大差值为 0.032, 略大于 1 个 JND, 出现在多媒体报告厅的 R_6 测点 -5 dB 信噪比情况. 可见直接法测得的 STI 与 STIPA 是非常接近的. 测量时发现多媒体报告厅的前面和后面墙面之间具有较强颤动回声, 前后墙面之间的距离为 18.7 m, 颤动回声的频率约为 9.1 Hz, 这可能会对附近的调制频率带来影响, 测点 R_6 结果差别较大或许与此有关.

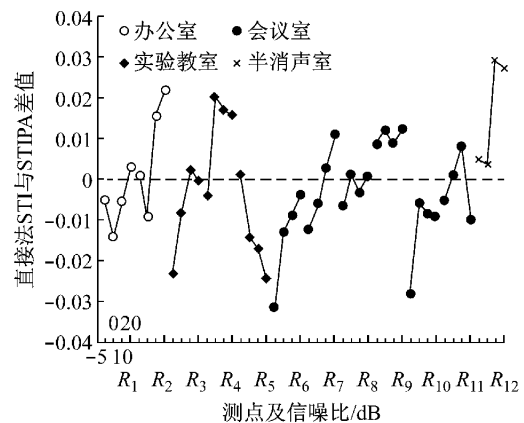


图 5 直接法测量的 STI 与 STIPA 差值

Fig. 5 The difference between STI and STIPA measured with direct method

图 6 给出了直接法与间接法测得的 STI 差值. 可以看到 48 个测试条件的差值比较小, 并且间接法整体上要比直接法略大, 最大差值为 0.038, 略大于一个 JND, 出现在多媒体报告厅的 R_6 测点 0 dB 信噪比情况. 由此可见采用直接法与间接法测量的 STI 并没有本质区别.

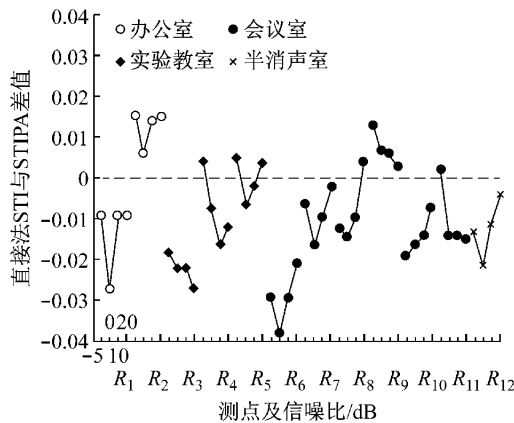


图 6 直接法与间接法测量的 STI 差值

Fig. 6 The STI difference between direct and indirect methods

图 7 给出了间接法测量的 STI 与 STIPA 的差值。一般来说,间接法测量 STIPA 并不比 STI 省事,因此很少有人使用间接法测量 STIPA。可以看到,48 个测试条件的差值比较小,平均值仅为 0.005 2,最大差值为 0.03,小于一个 JND,出现在半消声室的 R_{12} 测点 20 dB 信噪比情况。可见间接法测量的 STI 与 STIPA 并没有本质区别。另外从图 7 中也可以发现一个有趣现象,间接法测量的 STI 普遍比 STIPA 大,出现这种现象的原因还有待于进一步研究。

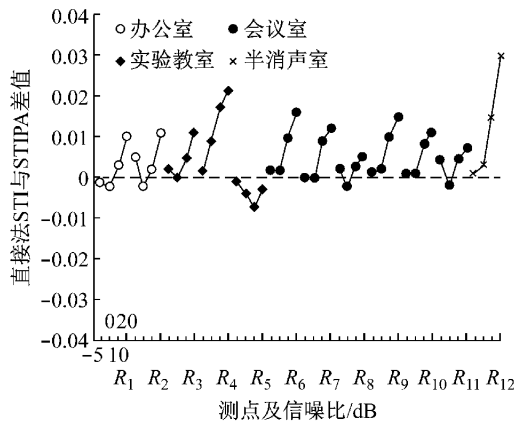


图 7 间接法测量的 STI 与 STIPA 差值

Fig. 7 The difference between STI and STIPA measured with indirect method

2.2 测量系统的频率响应对直接法与间接法影响

图 8 给出了直接法、间接法在系统频率响应修正与不修正两种情况下的 STI 差值,包括办公室、实验教室内 5 个测点共 20 个测试条件,测试信号的频率响应修正按照 1 000 Hz 声级相等设定的。从图中可以看到,直接法、间接法的差值曲线基本相同,并且呈现出一定的规律性:两条曲线的 STI 差值都

大于 0,表明修正后的测量值都大于修正前的测量值;对于每个测点,信噪比为 -5 dB 时,频率响应修正给两种方法带来的改善量都较小,但随着信噪比的增加,改善量明显加大,最大的改善量均出现在信噪比 10 dB 时;而信噪比 20 dB 时,改善量又有所降低;两种方法的改善量约在 0.050~0.279 之间,都超过了 1 个 JND,说明对于使用像人工嘴这样频率响应较差声源的测试系统来说,如果不进行频率响应修正,测量结果是无法使用的。

一般来说,频率响应修正只会带来信噪比的变化,对混响时间的影响较小,图 8 中的测量结果也说明了这一点:-5 dB 时,由于信噪比非常低而对语言可懂度基本没有贡献,频率响应是否修正对 5 个测点产生的影响都非常小;0 dB 时,影响明显增大;10 dB 时影响迅速增大并达到最大值;20 dB 时 STI 差距没有进一步扩大反而缩小,说明信噪比对语言可懂度的影响在超过一定值之后就会逐渐稳定下来。另外从图中也可以看到,在混响时间较长的实验教室内($R_3 \sim R_5$ 测点),频率响应修正带来的变化要小于混响时间较短的办公室内($R_1 \sim R_2$ 测点),说明混响时间也起到了一定的协同作用。由于混响时间和信噪比对语言可懂度的共同影响作用比较复杂,这些现象体所现出的内在规律还有待于进一步研究。

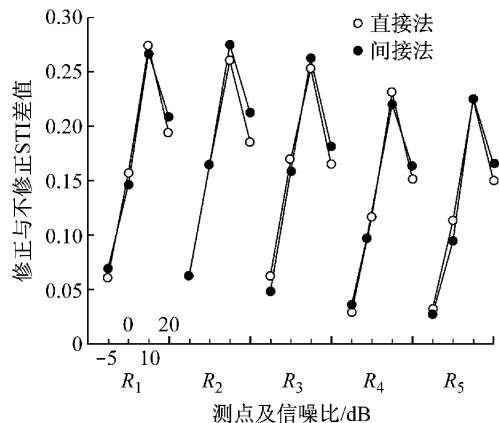


图 8 直接法与间接法频率响应修正与不修正的 STI 差值
Fig. 8 The STI difference between the corrected and the uncorrected frequency response conditions for the direct and indirect method

间接法测量一般分两部分完成,包括脉冲响应的测量和修正信噪比的测量,系统的频率响应在这两个环节中都有可能对测量结果产生影响。图 9 给出了脉冲响应进行和不进行频率响应修正对间接法测量结果的影响。从图中可以看到,脉冲响应是否修正对测试结果的影响非常小。这种情况也是正常的,

因为如果测得的脉冲响应各载波频带的信噪比非常高(例如超过 20 dB),此时频率响应修正带来的信噪比变化就可能不会影响到混响时间的计算结果,对 STI 的影响也就基本可以忽略了.同时这也说明,是后期信噪比的修正对间接法测量结果产生了决定性影响.修正信噪比的测量一般是通过人工嘴播放没有经过调制频率调制的 7 个载波频带合成信号,在测点位置处测量该信号的 7 个谱级;然后测量不播放该信号时的背景噪声谱级,经过计算可得到 7 个频带的信噪比.可见用来测量操作语言谱级的合成信号是否经过频率响应修正、频谱是否正确以及播放的声压级是否经过准确标定都会对结果产生较大影响,这也是为什么会强调间接法对测量人员的素质要求较高的原因.

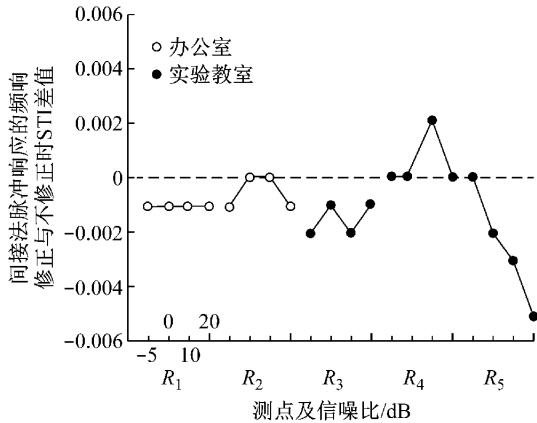


图 9 间接法脉冲响应的频响修正与不修正时 STI 的差值
Fig.9 The difference between the STI with and without frequency response correction for the measured impulse response of indirect method

2.3 对测量的多因素方差分析

实验中影响测量结果的因素较多,本文采用了统计学中的多因素方差分析方法进行分析.在多因素方差分析模型中,共有三个因素,其中因素 A 为测试方法,有直接法和间接法两个水平;因素 B 为混响时间,有 0.59;1.7;0.9;0.11 s 四个水平;因素 C 为信噪比,有 -5;0;10;20 dB 四个水平;整个实验次数为 96.一般影响测量结果的原因可分为两大类:由测试系统本身和测试过程中的偶然性带来的误差,称为随机误差,又称组内误差;由混响时间、信噪比、测试方法作为测试控制变量对因变量 STI 的影响,称为组间误差,总误差为组内误差和组间误差之和.于是构造了一个统计量 F :

$$F = \frac{S_{SA}/(k-1)}{S_{SE}/(n-k)}$$

式中, k 为水平数; n 为总的样本容量; S_{SA} 为组间离差平方和; S_{SE} 为组内离差平方和. S_{SA} 可计算如下:

$$S_{SA} = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$$

式中, n_i 为第 i 个水平的样本容量; \bar{X}_i 为第 i 个水平的样本平均值; \bar{X} 为所有样本的平均值. S_{SE} 可计算如下:

$$S_{SE} = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

式中, X_{ij} 为第 i 个水平的第 j 个样本; \bar{X}_i 为第 i 个水平的样本平均值.

表 2 给出了计算结果,表中“*”表示不同方法之间的交互影响, P 为相伴概率值.

表 2 多因素方差分析结果
Tab.2 Multi-way analysis of variance

影响因素	F	P
混响时间	205.68	0
信噪比	1 149.00	0
测试方法	1.17	0.28
混响时间 * 信噪比	6.72	0
混响时间 * 测试方法	0.12	0.98
信噪比 * 测试方法	0.09	0.97
混响时间 * 信噪比 * 测试方法	0.02	1.00
随机误差	1.00	

从表 2 中可知,混响时间的 F 值为 205.68,相伴概率 P 值为 0;信噪比的 F 值为 1 149, P 值为 0;测试方法的 F 值为 1.17, P 值为 0.28.这些结果说明测试方法对测量结果的影响非常小,大小与随机误差差不多,而混响时间和信噪比的影响则要大得多,尤其是信噪比的影响最大.这也很好地解释了为什么间接法的后期信噪比标定会对测量结果产生较大的影响.另外表 2 还给出了混响时间、信噪比、测试方法之间的交互影响数据.可以看到混响时间与信噪比之间的 F 值为 6.72, P 值为 0,说明二者的交互影响是显著的.混响时间与测试方法之间、信噪比与测试方法之间、混响时间与信噪比以及测试方法之间的交互影响不显著,相应的 F 值、 P 值见表 2.

图 10 给出了混响时间与信噪比的交互影响关系.可以看到,在混响时间较短的半消声室,随信噪比的增加 STI 增长较快;在混响时间较长的实验教室,随信噪比的增加 STI 增长较慢;办公室与实验教室的混响时间接近,随信噪比的增加二者 STI 的增长程度也接近(接近于平行).

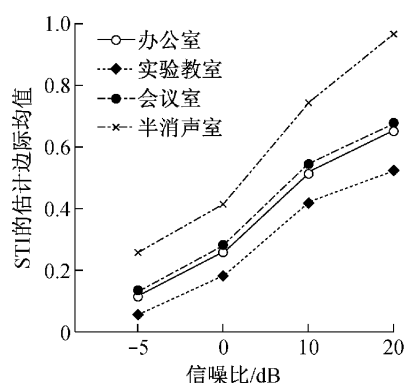


图 10 混响时间与信噪比交互影响关系图

Fig. 10 Interaction diagram of reverberation time and signal-to-noise ratios

3 结论

对于建筑声学来说,无论是直接法还是间接法,测得的 STI 和 STIPA 都能准确反映出语言经过声传输通路后与语言可懂度相关特征的变化,测量结果都可用于语言可懂度的客观评价,差别较小。但要获得可信的测量结果,应采用规格标准化的测量系统、测试程序,测量过程严格按照规定执行,否则测量的数据之间没有可比性,差别也会较大。

测试系统的频率响应对直接法和间接法都有较大影响,但影响的方式并不相同。直接法需要对测试信号做针对测量系统的频率响应预均衡,否则结果可能会有较大误差;间接法的对数扫频信号是否均衡对测量结果影响不大,前提是测得的脉冲响应应具有较大的信噪比;间接法后期修正信噪比的测量非常重要,用来测量操作语言声级的信号是否经过频率响应修正、频谱是否正确以及播放的声压级是否经过准确标定都可能会对测量结果产生较大影响。

对测量的多因素方差分析表明,测试方法对测量结果的影响较小,大小与随机误差差不多,而混响时间和信噪比的影响则要大多,尤其信噪比的影响最大;混响时间与信噪比之间存在着显著的交互影响。

随着 IEC 60268-16 第 4 次修订的发布,语言可

懂度客观评价指标在今后必将作为一个通用的声学参量而获得广泛应用,对于剧院、多功能厅、礼堂、体育馆、教室等领域的传统声学测量也会向该领域拓展。但随着测量领域迅速扩大,测试本身也会变得更富有挑战性及意义重大,本文研究结果对此具有较好的参考价值。

参考文献:

- [1] International Electrotechnical Commission. IEC 60268-16 Sound system equipment—part 16; objective rating of speech intelligibility by speech transmission index[S]. Geneva; IEC, 2011.
- [2] 祝培生,莫方朔,路晓东,等. 语言清晰度客观评价方法——对 IEC 60268-16(4.0 版,2011.6) 规范的解读[J]. 电声技术, 2012,36(5):40.
ZHU Peisheng, MO Fangshuo, LU Xiaodong, *et al.* Objective rating methods of speech intelligibility—the interpretation on the IEC 60268-16 (4th ed, 2011. June) standard [J]. *Audio Engineering*, 2012,36(5):40.
- [3] 中华人民共和国机械电子工业部. GB/T 14476—93 客观评价厅堂语言可懂度的“RASTI”法[S]. 北京:中国标准出版社, 1994.
Ministry of Mechanical and Electric Engineering of the People's Republic of China. GB/T 14476—93 “RASTI” method for the objective estimating of speech intelligibility in auditoria[S]. Beijing; China Standard Press, 1994.
- [4] 全国声学标准化委员会. GB/T 15485—1995 语言清晰度指数的计算方法[S]. 北京:中国标准出版社,1995.
National Technical Committee on Acoustics of Standardization Administration of China. GB/T 15485—1995 Acoustics methods for the calculation of the articulation index of speech[S]. Beijing; China Standard Press, 1995.
- [5] Standards Secretariat Acoustical Society of America. ANSI S3.5-1997 Methods for the calculation of the speech intelligibility index[S]. New York; American National Standards Institute, 1998.
- [6] Brüel & Kjær Sound & Vibration Measurement A/S. DIRAC room acoustics software, V5.0[Z]. Lyngby; Brüel & Kjær Sound & Vibration Measurement A/S, 2010.
- [7] Farina A. User manual of Aurora43[Z]. Parma; University of Parma A/S, 2012.
- [8] Bradley J S, Reich R, Norcross S G. A just noticeable difference in C50 for speech[J]. *Applied Acoustics*, 1999,58:99.