

关于声学比—混响时间—语言清晰度 关系的实验与理论计算*

饶 宇 安**

以前有关室内声场中语言清晰度问题的研究,都仅涉及混响时间“ T ”对语言清晰度“ S ”的变化关系,至于同一个“ T ”而不同声学比“ R ”的影响考虑很少。本文专为对各个不同 T 的条件下,“ R ”对“ S ”的影响进行定量研究。得到清晰度 S 与混响时间 T 乘声学比 R 的归一化实验曲线,并用声学比—混响系数 K_{R} 的形式表示为 $K_{R} = 1.0 - 0.22\lg(TR)$ 或 $S\% = 98.7 - 21.7\lg(TR)$ 。这个结果说明:混响场中的语言清晰度 S 与 (TR) 乘积的对数成反比。令人惊奇的是,这个实验所得的物理律的形式类似于 Weber-Fechner 定律,但要比它复杂得多。

Bolt 曾按等概模型计算混响条件下的语言清晰度,并和 Knudsen 的实验结果比较。我们在声学比实验的基础上按等概和一维分布计算,理论曲线的形状与实验符合。

双耳听觉具有指向能力,会使现场听音时的有效声学比减少从而提高清晰度,在这次实验中证明这个看法是对的。因此建议清晰度可表示为 $S = 98.7K_l \cdot K_a \cdot K_{R,R} \cdot K_{d(t)} \cdot K_e$ 。

自 1925 年以来,用语言清晰度来度量建筑中声学质量的办法一直是常用的^[1,2]。影响清晰度的物理量主要是混响、声级、频率与噪声。混响特性的测量通常假设扩散为理想,且混响时间 T 与频率 f 间的关系为低频稍长、500 赫以上为平直的,从而选用 500 赫上的混响时间表示清晰度与混响时间长短的关系^[3,4]。但在厅堂的现场中进行混响与清晰度关系的研究时,它不仅掺杂着客观条件中其他许多非线性畸变量的影响^[5,6],还与参加测听的人的主观状况有关,使现场测试的参数不易控制,并使客观环境对主观状况的影响很大。所以我们设计在实验室中进行这项工作。本文专为对各不同混响时间 T 时,不同声学比 R 对清晰度 S 的变化关系进行研究。

对清晰度队员的要求及测验材料的选编以前都已有过很多研究^[7],我们的测量是满足这些要求的。

一、测量原理和装置

实验装置如图 1 所示。图中发声室 1 是个特制播音亭,其混响时间低于 0.5 秒。发声人距电容传声器 10 厘米并可观察放大器上的电表以掌握声量。发声人所唸的语声信号分为两

* 本文 1978 年 11 月 7 日收到。

** 该工作是 1964 年于声学所完成,作者今于中科院生理所。

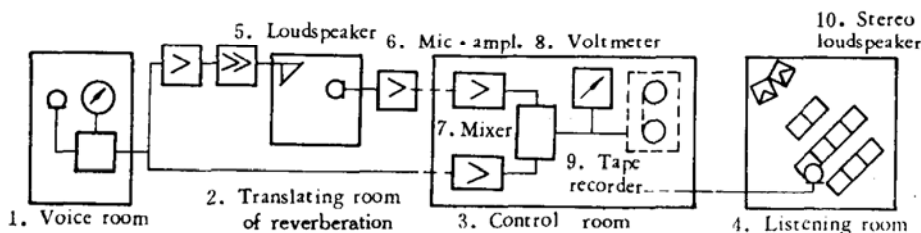


图1 实验装置原理图

- 1.发声室; 2.混响转播室; 3.控制室; 4.测听室; 5.转播扬声器;
6.电压放大器; 7.混合器; 8.测量混合器两路输入栅压的电压表;
9.磁带记录器; 10.立体声放声系统

The block diagram of experimental arrangement

路,一路经控制室成为测听室4中听到的直达语音。另一路经放大后送至混响转播室,转播室中的转播扬声器都放在房间的一角以有助于声场的均匀激发,并使声级保持在85分贝左右,避免扬声器过载失真。转播传声器都放在混响声场占优势的位置上(其室内声学比 R 的数值在 $T > 1.2$ 秒后都在10以上)。混响转播室多数是固定使用体积为104米³的混响室^[8],用吸声材料改变它的混响时间,吸声材料的布置尽量有利于扩散。由混响室转播来的语音信号和播音亭直接送来的信号在双路混合器7中合并后,送入测听室4中进行清晰度试验。在混合器7的栅极上接电压表8,分别测量两路的输入电压,令混响声一端来的电压值 e_r 与直达声一端来的电压 e_d 的平方比为 R_E 。混合器的输入与输出电压之间在3伏以内保持良好线性关系,为了保证两路信号的电压输入混合器时工作在线性范围内,我们对 e_r 、 e_d 的电压值(伏)与 R_E 的关系按下式来取:令 $e_r + e_d = 2.4$ 伏,由 $R_E = e_r^2/e_d^2$,可得:

$$e_d = 2.4 / (1 + \sqrt{R_E}) \text{ 及 } e_r = 2.4 - e_d \quad (1)$$

混合器在实际使用条件下,直达声与混响声两路的频率特性保持平直(图2中曲线2、3)。混合器的输出信号既然包含了不同比例的直达声与混响声能,我们即可以在通常测量混响曲线用的声级记录仪上验证具有不同声学比 R 时混响曲线的形状。在图3中给出了四种不同比例的 R_E 的混响曲线。这些曲线是在体积为425米³的混响室中,将声源放在某个墙角上、传声器与声源相距5米(房间常数 R_H 为50.3)时,在混合器输出端用记录器所测。在测得的曲线上,直达声垂直下降的分贝数与按公式 $10 \lg[(1 + R)/R]$ 计算所得相符^[9]。我们曾指出:

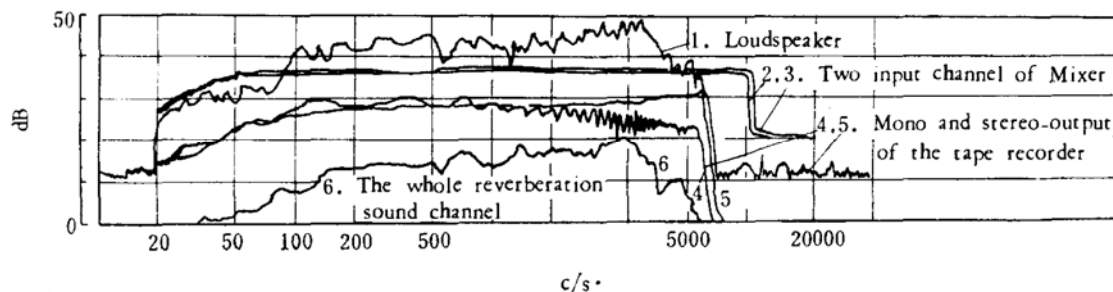


图2 频率特性

- 1——转播扬声器的频率特性; 2,3——混合器两路输入的频率特性;
4,5——录音机单路输出与双路输出的频率特性; 6——混响声一路全通道频率特性

The frequency characteristics

对于测量混响常用的 50 分贝电位计来说, 混响曲线上所能分辨的起伏精度为 2 分贝, 则由 $10 \lg[(1+R)/R]$ 式可知, 当声学比 R 大于 1.8 以后, 在混响曲线上即不再能分辨出 R 的变化. 由图 3 中的曲线 3 可见, 当 $e_r/e_d = 1.4:1.0$, 即 $R_E = 1.96$ 时, 混响曲线已成为一个标准直线, 目视分辨不出不同比例的直达声, 但对听觉和清晰度的影响却仍很大.

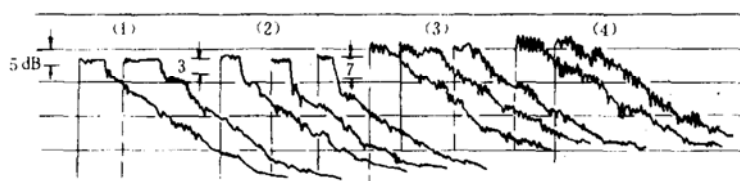


图 3 改变 R_E , 由混合器输出端接记录器所得混响曲线

(1) $e_r/e_d = 1:1$; (2) $e_r/e_d = 0.8:1.58$;

(3) $e_r/e_r = 1.4:1.0$ (4) e_d 断开, 全是 e_r 即 $R_E = R_H = 50.3$

The reverberation curves for varying R_E recorded at the output of mixer.

进行清晰度试验的测听室体积为 124 米³, 以两层窗及两道加厚的门来隔声, 所布置的混响时间为 0.82 秒, 频率特性平直, 室内各点声学条件接近相同. 队员的室内分布如图 4 所示, 房中央虚线所围部分算作 I 区, 其余周围算 II 区. 可认为测听室本身的音质特性是满意地合乎使用要求的.

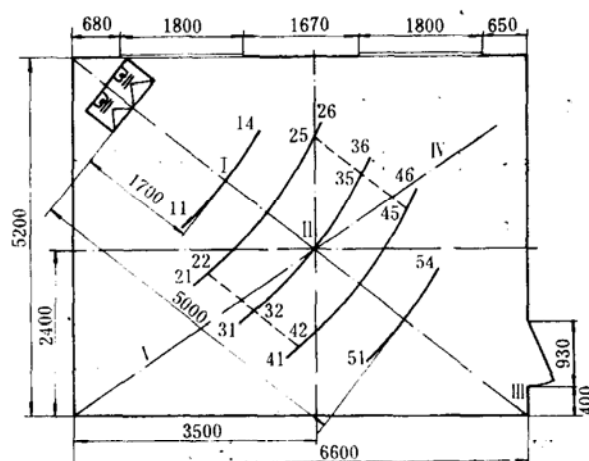


图 4 测听室平面图, 测听队分布排号示于图中, 11—54 为坐位编号,

2—4 排虚线内算 I 区, 其余算 II 区

The plan of listening room, in which the distribution of the listeners is described

二、实验方法与步骤

在测试语言清晰度与 R_E 变量关系的实验中, 测听队的基本队员放在中央 I 区. 他们经过多次不同条件的训练后, 在生理和心理状态上, 可以认为已是正态分布. 每个测听队员的清晰度测量结果随 $R-T$ 变化而变化的规律皆相同, 但绝对值不同. 我们解释为各人听觉本领存在系统差别, 这种差别在条件差时更显著. 这还可以从下列实验统计 (表 1) 得到证明: 当条

件差、平均得分较低时,其自身离差也大,反之则小。在清晰度平均值降低时,各测听队员之间分的离散性也增加,相对误差变大,按 z 分布计, $\Delta S\%$ 超过 2% 者占 42%。

表 1 四个听声人在不同条件下所得清晰度偏差

听声人	按人计算抽取31个不同实验条件所得平均清晰度 $\bar{x}_{ik}$	相应的相对误差 σ_{ik}	按人计算抽取48个条件所得平均清晰度 $\bar{x}'_{ik}$	相应的相对误差 σ'_{ik}	听声能力
A(女)	86.8%	3.13%	83.1%	1.33%	最佳
B(女)	81.9%	3.88%	81.9%	2.77%	较佳
C(女)	76.5%	4.82%	79.0%	3.21%	中等
D(男)	72.2%	5.23%	71.8%	3.59%	较差

任何清晰度测验值都与发声者的发声速率及口声特性有关^[10],为了接近实际,我们的实验中用男女发声人各一名,采用通常电台播音时的发声速率,平均音节速率约为每秒 5 至 6 个。

在整个 $R-T-S$ 实验中,由混响转播室的声学布置所得到的一组混响时间 T 与频率 f 的特性曲线示于图 5 中。

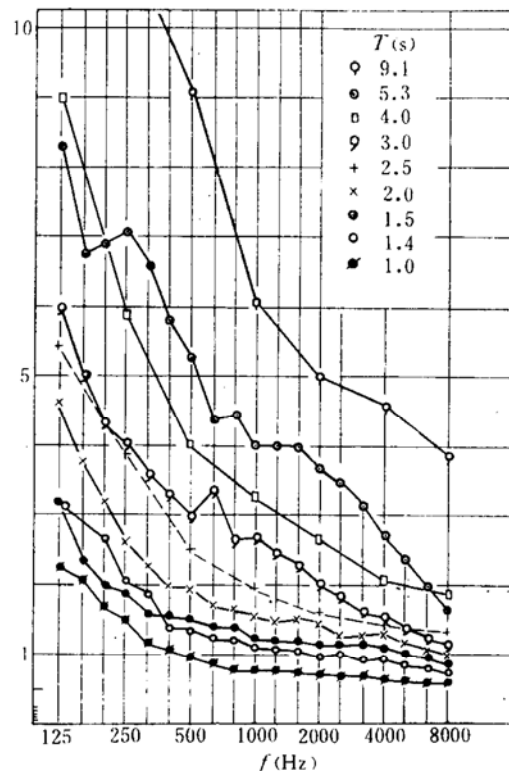


图 5 在混响声一路电压放大器输出端所测各次实验的混响时间 T -频率 f 的特性曲线(每 $1/3$ 倍频程测一点),混响曲线衰减均良好。

The T - f curves measured at the output of voltage amplifier of the reverberant sound in each experiment(at $1/3$ octave interval).

All of the reverberation curves are normal

三、实验结果与归纳

在观察 $R-T-S$ 三量关系之前, 先观察单独由混响室转播到测听室试验清晰度 S 与混响时间 T 、 $T-f$ 形状的关系. 从单独混响一路转播所得清晰度实验结果可见: 混响时间 T 与频率 f 间的曲线形状对清晰度有着显著影响. 在建筑声学中以前习惯于用频率为 500 赫的混响时间作为整个 $T-f$ 曲线的代表来估计清晰度, 但这种办法只适用于“正常”形状的 $T-f$, 即 500 赫以上基本平直, 500 赫以下低频部分曲线比 500 赫高出 20—30% 的情况. 当 $T-f$ 曲线较特殊时, 仅以 500 赫的混响时间来代表各长短不同的混响, 它们相应的清晰度也将有较大出入. 因此, 我们在 $R-T-S$ 三量实验中所设置的 $T-f$ 混响曲线以及混响声级的衰减曲线都要求是“正常”形状的. 在不同 T 的条件下改变声学比 R_E 所得清晰度 S 的曲线族示于图 6 中.

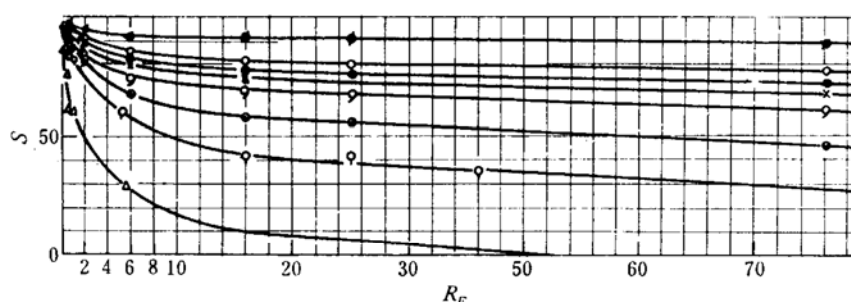


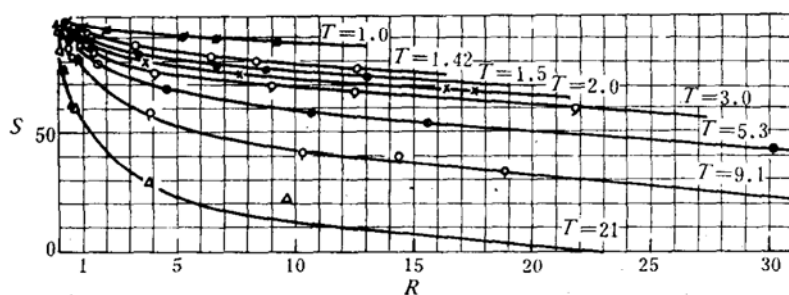
图 6 发声人 A(女)的 R_E 与 T, S 的实验曲线族,
图中横坐标 R_E 与纵坐标 S 都是线性标度

The experimental curve family of R_E vs T, S for speaker,
A(female), in which abscissa R_E and ordinate S are linear

在这个曲线族中, 是把除直接通路来的直达声以外所有其余(后到)的声能都算作混响声能来计算声学比得出的关系. 考虑到人耳的积分效应, 即直达声之后的近次反射声对直达声的接收是有利的, 而我们在所有对 R_E 的实验里, 混响转播室中扬声器到传声器间的距离均小于 7 米, 亦即经混响室中扬声器到传声器的直达声部分, 在混合器上与直接通路的信号合并时, 其时差不大于 50 毫秒, 因此我们至少可以把混响室中扬声器到传声器的直达声合并到直达声电路内, 都算作直达声. 由此概念, 声学比的解析关系可推导如下: 由前已知, R_E 是混合器上由两路输入电压值的平方比算得的声学比, 直达声电路的电压为 e_d , 混响声电路电压 e_r , 则是由混响转播室来的信号中本身的直达声声压 p_{dH} 与混响声压 p_{rH} 共同作用产生的, 所以混响声电路的电压 e_r 可写成 $e_r = e_{dH} + e_{rH}$, e_{dH} 和 e_{rH} 各为混响转播室中直达声声压与混响声压所相应的电压, 混响室内的声学比 R_H 即为 $R_H = (e_{rH}/e_{dH})^2$, 所以 $R_E = e_r^2/e_d^2 = (e_{rH} + e_{dH})^2/e_d^2 = (\sqrt{R_H} + 1)^2 e_{dH}^2/e_d^2$, 当 e_{dH} 和 e_d 间的时差小于 50 毫秒(即转播室中扬声器与传声器距离小于 17 米)时, 把 e_{dH} 能量和 e_d 合并作直达声, 听觉上起作用的总声学比 R 为:

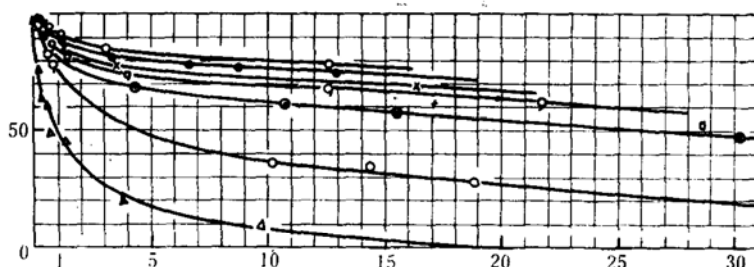
$$R = e_{rH}^2/(e_d^2 + e_{dH}^2) = R_H R_E / (R_H + 2\sqrt{R_H} + 1 + R_E) \quad (2)$$

当 e_{dH} 与 e_d 之间时差较大时, 则总声学比就等于电声学比 R_E . 由这个考虑了人耳积分效应的总声学比 R 与实验测定的清晰度 S , 所得到的 $R-T-S$ 曲线族(图 7) 可见, 音节清晰度和



(a)

女发声 $T = 1.0-5.3$, $r = 6m$
 $T = 9.1$, $r = 4.1$,
 $T = 21$, $r = 5$
 $T = 2.5$, $r = 4.4$
 $T = 4.0$, $r = 4.1$



(b)

$$\text{男发声 } R = \frac{R_E}{\frac{R_E}{R_H} + \left(\frac{2}{\sqrt{R_H}} + \frac{1}{R_H} + 1 \right)}$$

图7 总声学比 R 、混响时间 T 与清晰度 S 的实验曲线。用白噪声调 R_E , 但最下端的 $T=9.1$ 及 $T=21$ 秒曲线的 R_E 是用 1000 赫单频调的。

The experimental curve of the total acoustic ratio R , reverberation time T and articulations. R is adjusted by the white noise, but R_E of the lowest curve of $T=9.1$ s and $T=21$ s is adjusted by pure tone of 1000 Hz

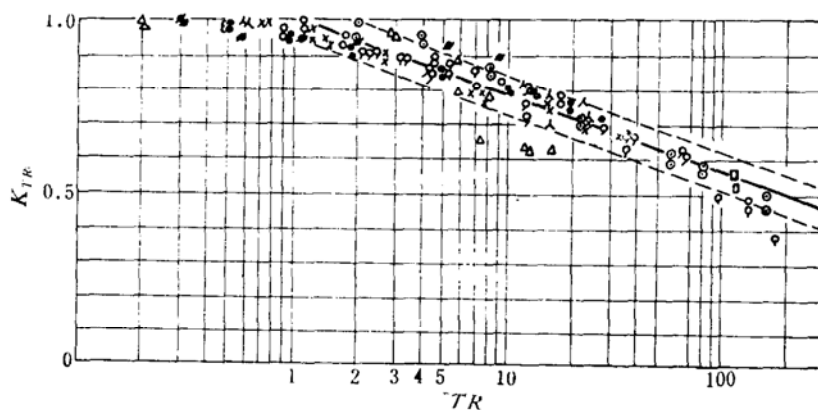


图8 以普通话清晰度比值 K_{TR} 作纵标, 与各次实验 (包括男女发声人) 所得 TR 作横坐标的曲线关系。虚线为按其最大误差表示的分布范围。

The dependence of the efficiency K_{TR} (S divided by the maximum articulation of the Chinese speech 98.7) on the TR product obtained from experiments (including male and female speakers). The dotted lines show the distribution range of maximum deviation

声学比、混响时间之间的关系呈对数变化,我们可找出 R - T - S 的归一化关系: 将男女发声的测听结果加以平均,并以汉语普通话清晰度的最大值 98.7%^[11] 作为各次清晰度 S 的分母所得到的比值系数 K_{rR} 作纵坐标变量,而以混响时间 T 与声学比 R 的乘积的对数作横坐标变量,得到图 8 所示直线,它可近似表达为:

$$K_{rR} = 1.0 - 0.22 \lg(TR) \quad (3)$$

由上式可见,当混响时间与声学比的乘积 TR 减少到趋近 1 时,这个声学比-混响系数 K_{rR} 均趋近于 1,清晰度将趋近于最大值. 上式只适用于 TR 大于 1. 准确度约为 5%.

四、结果讨论与验证

我们由实验得到的声学比-混响与清晰度间的系数关系式:

$$K_{rR} = 1.0 - 0.22 \lg(TR)$$

或直接用汉语清晰度表示式表示

$$S\% = 98.7 - 21.7 \lg(TR) \quad (4)$$

T 和 R 是混响内语言清晰度的两个独立参量,混响声场中的清晰度 S 与 (TR) 乘积的对数成反比变化. 这就解释了长期以来仅用混响时间作音质第一评价参数时,之所以在相同混响时间下会得到不同清晰度实验结果,是忽略了参数 R 之故. 例如 Steinberg 对英语的实验中^[12],当 $T = 2.2$ 秒时 2 米处的 $S = 95\%$ 到 6 米处为 85%; 在俄语的实验中当 2 米处的 $S = 83\%$ 时,在 6 米处为 50%^[13]. 用有指向性传声器所得 S 比用无指向性的为高.

但使我们更感到惊奇的是这个物理实验所得到的物理律公式(3)和(4)的形式,却和生理和心理学领域中的 Weber-Fechner 定律相似但要复杂得多,因为“混响”本身就是个包含了空间与时间的复杂过程.

我们还在一间 425 米³的混响室中进行过直接测听实验. 测听人距声源平均距离为 3.4 米 (R_H 较小),发声人仍在另外的播音亭中,这样测得的 T - S 变化规律(图 9 中实线),即与

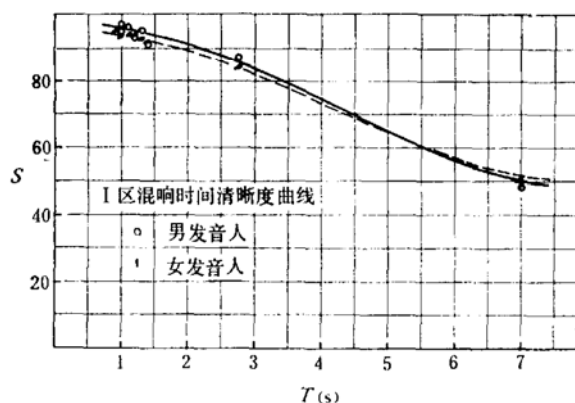


图 9 测听队直接在体积为 425 米³的大混响室中测得的 S - T 曲线. 虚线为 Knudsen 现场归纳的 S - T 曲线.

S - T curve obtained by directly listening in a large reverberation room. The dotted curve is the S - T curve of Knudsen's obtained in a real auditorium.

Knudsen 所得曲线相似,因为在逐渐减少吸声来增长混响时间的同时,其室内声学比 R_H 的增加也越来越慢。

由于人耳的双耳效应,使人在现场面对声源听声时具有指向效应(用双耳指向性系数 Q_p 代表),它使得现场听声时的有效声学比减小,从而能提高清晰度,但用混响转播法测清晰度时,无方向性的转播传声器不能象人在现场测听时那样,用双耳去甄别混响声与直达声,这就使有效声学比加大。为了进一步验证这个看法,我们在体积为 104 m^3 的混响室中布置了长短两种混响时间,使听音人先直接在混响室、然后经混响室转播至测听室听男女发声的清晰度,结果列于表 2,所得现场和转播测听的清晰度之间约差 30%,按前面公式(4)计算,所相当的 $Q_p = 13.7$ 。这证明了我们关于人耳听觉存在一个指向性效应 Q_p 的假设是成立的。总之,由

表 2 转播对清晰度的影响

$T = 2.5 \text{ (s)}$			$T = 4.0 \text{ (s)}$
现 场	男发声	92.7%	86.3%
	女发声	93.0%	81.8%
转 播	男发声	68.5%	56.2%
	女发声	61.2%	50.9%

物理定义所计算的声学比 R 与由人的双耳效应产生的有效声学比 R_e 之间的差 ΔR ,都与声场的瞬态方向扩散 $d(i)$ 的状态有关^[14]。在厅堂不同的区域上,不论用扩声系统与否,各点的清晰度都将处于图 7 所示 R - S - T 曲线族的范围中。若厅堂前排坐位听者还能看清发声人口部动作,则视觉信息还能帮助提高清晰度^[15]。

五、理论考虑与算式

由实验已知声学比 R 和混响时间 T 是影响清晰度的两个独立参量。Bolt 在 1949 年曾按语音振幅为等概分布的假设但没有考虑 R 的情况计算混响对清晰度的掩蔽^[16],并认为和 Knudsen 的实验结果相合。现在将声学比、混响时间合在一起来考虑清晰度的理论计算模型,并和实验结果进行对比。我们用了下列理论考虑:(1) 汉语普通话声级动态范围 ΔI 和英语相似^[17,18];(2) 在英语等语言的通讯问题中,利用清晰度面积表示法进行的清晰度计算^[19-21],在建筑声学中则以“混响语言谱”^[22-24]计算,这实际都是认为,在语言的平均频谱及动态范围所包围的曲面内,各点上语声出现的概率相同。但这个等概假设只有在满足最佳声级并只求近似计算时才能成立。当存在着非线性畸变及噪声或需要较精确计算时,则需按语声声级的概率分布密度函数 $\omega(h)$ 来计算^[25-27];(3) 在混响场中的清晰度与发声人发声速率有关。我们将普通话每个字音的音节看成是一个独立脉冲,一组音节即组成了一组高度为 ΔI 分贝、平均音节脉宽为 τ_1 和音节间隙为 τ_2 的语言脉冲系列,由混响引起前后脉冲间的掩蔽。当我们得出音节宽及音节间隙的平均值后,可以象 Bolt 那样把音节脉冲看成矩形,但并不假设信息量全部集中在矩形脉冲的前沿。各脉冲系列的衰减斜率 m 按混响定义即为 $m = 60/T \text{ (dB/s)}$,由于声学比 R 的存在,比 Bolt 等概分布模型的声级区间要多一项 $h_d = 10 \lg[(1 + R)/R]$,而 $h_{\max}$ 最多

只能等于 ΔI 分贝(图 10), 在 T 和 R 都是连续变数时, $h_{\max}$ 不会都恰等于某个条件下相应的

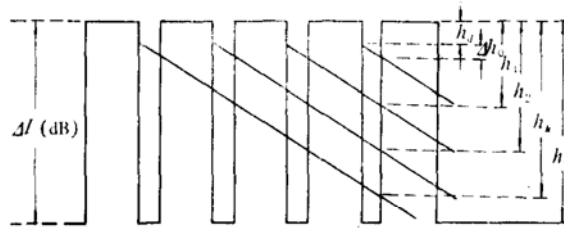


图 10 取时间-声级平面的语脉冲计算图

The calculating model of speech sound pulse taking the time-sound level plane.

Δh 的整数倍, 因此整个语脉冲所能分割的区间数 n 是, $n = \{\Delta I - [10 \lg(1/R + 1) + h_0]\} / (\Delta h = \text{整数 } k + \text{小数部分})$, 令 $\Delta h / \Delta I = c$, 取 $\Delta I = 30$ 分贝, 可导出语脉冲各声级区间未受掩蔽的概率 $A. I.$ 各为:

$$A. I. (\Delta h_d + \Delta h_0) = [\lg(1/R + 1)]/3 + 2\tau_2/T,$$

$$\sum_{i=1}^k A. I. (\Delta h_i) = \sum_{i=1}^k c(2-c)(2-3c)(2-5c) \cdots [2 - (2i+1)c]/2^i$$

语脉冲最下端非整数部分的清晰度指数为

$$A. I. (\Delta h_n) = N(2-3c)(2-5c) \cdots [2 - (2k-1)c][2 - (N+2kc)]/2^k,$$

其中

$$N = 1 - \left[\frac{1}{3} \lg \left(\frac{1}{1+R} \right) + \frac{2\tau_2}{T} + kc \right],$$

总的清晰度指数为

$$A. I. = A. I. (h_d + h_0) + \sum_{i=1}^k A. I. (\Delta h_i) + A. I. (\Delta h_n)$$

这个 T 及 R 的函数 $A. I. (T, R)$ 是个末项逐渐变小的有限级数, 在变数 R 及 T 从 $0.5 \rightarrow \infty$ 时满足均匀收敛级数的 Weierstrass 判则, 收敛区间为 $[0, 1]$.

语言的声级在它的动态范围中实际上并非等概出现, 非等概的影响对混响场中的语言比通讯广播时大. 下面为考虑了声级分布概率的混响掩蔽算法: 设 $w(\Delta h_i)$ 为语言声级振幅落在 $(h_i - \Delta h_i/2 \leq h \leq h_i + \Delta h_i/2)$ 区间的概率, 表示为

$$w(\Delta h_i) = \int_{h_i - \Delta h_i/2}^{h_i + \Delta h_i/2} \omega(h_i) dh_i,$$

$\omega(h_i)$ 是语声在 ΔI 中的分布密度函数. 清晰度算式则成为

$$A. I. (\Delta h_d + \Delta h_0) = \frac{10}{\Delta I} \left\{ \lg [(1/R) + 1] w(\Delta h_d) + \frac{60\tau_2}{\Delta I T} w(\Delta h_0) \right\}$$

$$\sum_{i=1}^k A. I. (\Delta h_i) = \sum \left[1 - \frac{1}{2} c w(\Delta h_1) \right] \left[1 - c w(\Delta h_1) - \frac{1}{2} c w(\Delta h_2) \right] \cdots \left[1 - c w(\Delta h_1) - c w(\Delta h_2) \cdots - c w(\Delta h_{i-1}) \right]$$

$$- \frac{1}{2} c w(\Delta h_i) \Big] c w(\Delta h_i)$$

语脉冲最末顶小数部分清晰度指数为

$$A. I. (\Delta h_{n'}) = \left[1 - \frac{1}{2} c w(\Delta h_1) \right] \cdots \left[1 - c w(\Delta h_1) - c w(\Delta h_2) - \cdots c w(\Delta h_{k-1}) - \frac{N}{2} w(\Delta h_{n'}) \right]$$

它的总和同样满足 Weierstrass 级数收敛判则。若有噪声干扰时,也按类似方法计算,取不同信噪比 S/N 则有不同语脉冲的区间数 n' 。因目前尚未得出汉语普通话概率分布密度函数的标准,但将已有结果^[27]与英语^[25,28]对比,除小信号因测量不准而稍有差别外,均趋近相同的指数律分布,故取英语积分分布曲线换算成密度分布函数 $\omega(h_i)$ (图 11)来计算。

六、参数测量及计算

计算所需语声脉宽 τ_1 和 τ_2 测自前面发音人的录音,以动片摄取双线示波器上语幅随时

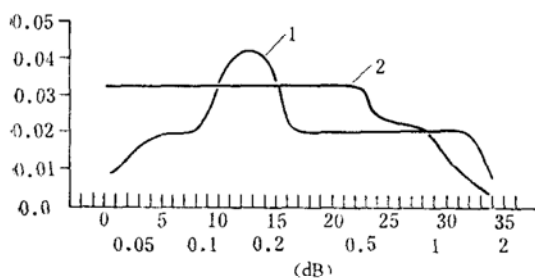


图 11 由英语累积分布曲线换算得来的声级分布概率密度函数:

1. 为峰值分布^[14], 2. 为瞬时值分布^[10]。

The probability density distribution function curve derived from cumulated distribution curve of English: 1. the peak distribution; 2. The instantaneous distribution.

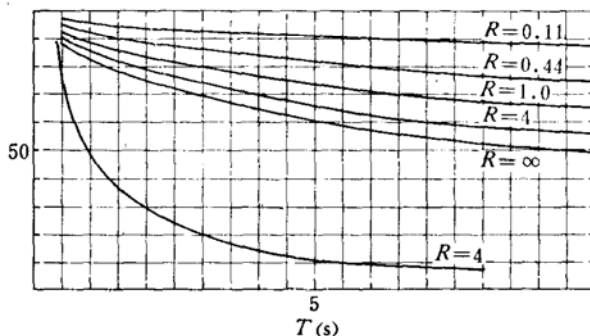


图 12 以 $\tau_1 = 172.5$ (ms)、 $\tau_2 = 16.4$ (ms) 代入等概分布算式得上部一组曲线,代入非等概分布算式得最下一条曲线。

The upper curves are calculated with $\tau_1 = 172.5$ ms and $\tau_2 = 16.4$ ms by means of the equal probability expression, the lower by the unequal probability expression.

间变化的过程。因目前未规定音节测法,我们所量 τ_1 值以实际语声出现时间来量度,底片上看出语言波时就起算(表 3)。由十几张字表的 600 个音节测得 $\bar{\tau}_1 = 172.5$ 毫秒, $\bar{\tau}_2 = 16.4$ 毫秒。将测得的 τ_1 及 τ_2 值代入前面按等概分布及按声级分布律的计算式,可得出不同 T 及 R 变数下相应的清晰度指数的曲线族,由已有的普通话实验结果可换算成不同 T 及 R 时的清晰度百分数曲线(图 12)。

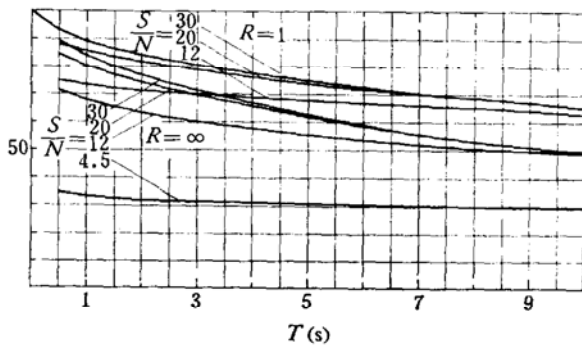
当混响场还有另外的环境噪声存在时,对 $\Delta I = 30$ 分贝,取信噪比 S/N 的不同比值,则不同语声脉冲的区间数 n (或 n'),例如当 $S/N = 20$ 分贝时

$$n(\text{或 } n') = \left\{ 0.67 - \frac{1}{3} \lg \left(\frac{1}{R} + 1 \right) + 2 \frac{\tau_2}{T} \right\} / c(\text{或 } c')$$

当 $S/N = 12$ 分贝时

表 3 语声脉冲宽度和间隙时间(毫秒)

τ_1	τ_2	τ_1	τ_2	τ_1
度 138.6	23.7	纸 120.7	46.5	折 171.3
又 267.5	45	发 192.5	7.5	镜 210
爱 149.4	77.4	偏 198	14.4	见 162
苦 117	0	追 140.4	41.4	等 86.4
色 154.8	43.2	笛 151.2	10.8	穿 189
雅 198	39	擦 188	0	会 157
固 150	48	得 187	0	叫 195
懂 174	66	舌 167	0	卧 148
如 170.3	$\frac{1}{2} \tau_2$	187.1	0	159.3 (ms)

图 13 ΔI 取为 30 分贝时, 分别对信噪比 S/N 为 30、20、12、4.5 分贝及 $R = 0.1$ 、 $R = \infty$ 两种情况计算的等概清晰度曲线

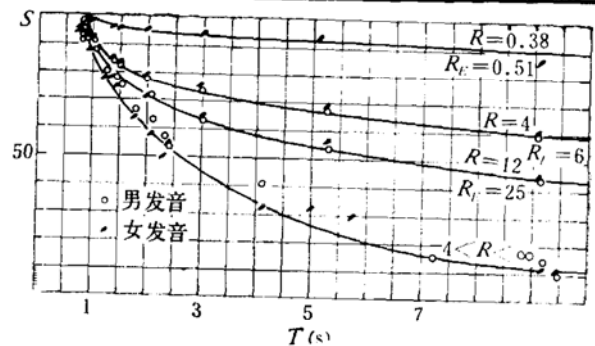
Calculated curves of articulation with $R = 1$ and $R = \infty$, $\Delta I = 30$ dB and signal-noise ratio = 30, 20, 12, 4.5 dB respectively for equal probability.

$$n(\text{或 } n') = \left\{ 0.40 - \lg\left(\frac{1}{R} + 1\right) + \frac{2\tau_2}{T} \right\} / c(\text{或 } c')$$

图 13 为在不同信噪比时按等概分布式计算的一些曲线

为将理论计算与实验对比, 将实验测得的 R - T - S 的一些曲线绘于图 14 中.

由这些计算结果与实验对比后可见: (1) 当直达声较多、 R 不是很大, 总声学比 R 为 4.0 及 0.11 时, 按等概分布计算结果与实验很接近, 仅原点有差别; (2) R 越大, 所算得的曲线下

图 14 不同 R 及 T 的实验结果, 最下一条曲线为 R 都是很大仅有混响一路时的实验曲线.

The experiment results of various R and T . The lowest curve is the experimental result in the large value R with only reverberation channel

降越陡,和实验所得一致;(3) 由噪声与混响并存掩蔽的计算结果可见: 环境噪声对清晰度的影响只有在短混响时才显著,在长混响时影响小得多,甚至看不出,这是因为长混响相当于一种与信号频谱相近的噪声掩蔽,因此它的实效就相当于使语言信噪比降得很低的强噪声。在声学比 R 趋于大数时,混响时间 T 的增长相当于迅速的降低信噪比。图 13 及图 14 最下端的曲线即为这类情况:

在 R - T - S 实验的基础上进行的理论计算曲线与实验规律比较符合,也说明以前 Bolt 计算与 Knudsen 实验间的符合带有偶然性。但是某些有关的听觉特性(如积分效应^[29]、音节过渡特性^[30-32]等)的研究还有待深入,以进一步完善这些计算模型。

七、实验结果与建议

比较各种语言条件的清晰度可知,汉语普通话四声的声调是很不容易听错的,但在长混响时也会听错声调,这可能是混响时间 T 较长时,前后声调叠加产生的谐波干涉,在听觉上造成“简并音”效果之故。前面实验中已比较了混响频率特性对清晰度的影响。因此对非‘正常’的混响频率特性,计算清晰度所取的混响时间应作如下计权: 由现有的汉语等清晰度指数分量的频带与通常测量混响时间的倍频程频带对比,可见中频 900 至 4000 赫内,在每 1/3 倍频程测一个混响时间 T 值的频段当中都包含了两至三个等清晰度指数的频带,因此对中频段的混响时间作用应加权重。由于实际的 T - f 曲线的中高频部分不会起伏很大,为应用的方便,可按下面的简单关系式算出平均混响时间 $\bar{T}$,再代入前面诸式计算。计权系数 b_i 列于表 4。

表 4 计权系数 b_i

频 道	频道边界 $f(\text{Hz})$	带 宽 $\Delta f(\text{Hz})$	中心频率 $f(\text{Hz})$	按 1/3 倍频 的中心频率	系 数 b_i
1	200—400	200	300	320	1
2	400—550	150	475	500	1
3	550—730	180	640	640	1
4	730—900	170	815	800	1
5	900—1020	120	960	1000	2
6	1020—1150	130	1085		
7	1150—1270	120	1210	1250	3
8	1270—1400	130	1335		
9	1400—1570	120	1460		
10	—1700	180	1610	1600	2
11	—1900	200	1800		
12	—2100	200	2000	2000	2
13	—2400	300	2250		
14	—2700	300	2550	2500	2
15	—3000	300	2850		
16	—3400	400	3200	3200	2
17	—4000	600	3700		
18	—4700	700	4350	4000	1
19	—5700	1000	5200	5000	1
20	—7000	1500	6350	6400	1

$$\bar{T} = \left\{ \sum_1^{13} b_i T(f_i) \right\} / 13$$

由此我们建议：在任何一个声源为自然声源或是加了扩声系统的实际厅堂的任一区域上的语言清晰度可表示为下面五个系数与清晰度最大值之间的乘积：

清晰度百分数 $S = 98.7 K_l K_n K_{rp} K_{d(t)} K_e$, K_l 及 K_n 各为响度及噪声的影响系数； K_{rp} 为加权平均的混响时间与作用在人耳上有效声学比所决定的系数，人耳双耳效应的指向性系数 Q_p 约为 13.7, 其他直达声能由现场具体情况决定； $K_{d(t)}$ 是由瞬态方向扩散度 $d(t)$ 决定的系数（它与回声效应、声源特性、大厅形状及内表面吸声分布、声级衰减中的起伏幅度有关，作简化估计时可取为 1）； K_e 为视觉信息所增加的清晰度系数（通常可取作 1）。

谨向马大猷教授及魏荣爵教授热情有力的帮助致以衷心感谢。

参 考 文 献

- [1] Knudsen, V. O., *Phy. Rev.*, **26**(1925), 287.
- [2] Cavanaugh, W. J., Farrel, W. R., et al., *J.A.S.A.*, **34**(1960), 478.
魏荣爵、包紫薇, 电信科学(1957), No. 9, 4. “汉语的混响性质”, 1957年东德建筑与房屋声学会议论文集.
- [3] Knudsen, V. O., *J.A.S.A.*, **1**(1929), 56.
- [4] Sabine 编, 马大猷译, “实用建筑声学”(1953年商务印书馆).
魏荣爵, 物理通报(1954) No. 12, 705.
- [5] French, N. R. & Steinberg, J. G., *J.A.S.A.*, **19**(1947), 90—119.
- [6] E. G. 里查孙等著, “声学技术概要(上册)”(科技出版社1961), 256—263.
- [7] Fletcher, H. & Steinberg, J. G., *J.A.S.A.*, **1**(1929), 292.
- [8] 饶宇安, “房屋形状对声场扩散影响的测量”, 中国建筑物理学会学术论文选集(1963年中国工业出版社).
- [9] Фурудуев, В. В., “Акустический основы вещение”, (1960).
- [10] Pollack, I., Pickeff, I. M. & Sumby, W. H., *J.A.S.A.*, **1**(1929), 121.
- [11] 张家骥、马大猷, 声学学报, **2**(1965), 80.
- [12] Steinberg, J. G., *J.A.S.A.*, **1**(1929), 121.
- [13] Рабинович, А., *Ж. Т. Ф.*, **4**(1934), 1134.
- [14] 饶宇安, “声场瞬态扩散与音质”第一届全国声学会议论文摘要, (1964, 4, 北京).
- [15] Sunby, W. H. & Irwin Pollack, *J. A. S. A.*, **26**(1954), 212—215.
- [16] Bolt, R. H., MacDonald, A. D., *J.A.S.A.*, **21**(1949), 577—580.
- [17] 魏荣爵、包紫薇, 物理学报, **16**(1960), 338—347.
- [18] 徐焕章, 声学学报, **2**(1965), 203.
- [19] French, N. R., Steinberg, J. C., *J.A.S.A.*, **19**(1947), 90—119.
- [20] Karl, Kryter, D., *J.A.S.A.*, **34**(1962), 1698—1702.
- [21] Weberter, J. C., *J. A. S. A.*, **36**(1964), 1662—1669.
- [22] Baranek, L. L., “声学”(高教出版社1959), 430—464.
- [23] Jassen, J. H., *Acustica*, **7**(1957), 305—309.
- [24] Lochner, J. P. A., & Burgur, J. F., *Acustica*, **10**(1960), 394.
- [25] Wibup. Daveport, JR, *J.A.S.A.*, **24**(1952), 390—399.
- [26] Римский-кореаков, А. В., *Акус. Жур.*, (1960), Выч 3.
- [27] 汪润生等, “汉语信号及其变化率的一维概率分布密度的测量”第一次全国信息论专业学术会议论文选集, (1964), 132—140.
- [28] Dunn, H. N. & White, S. D., *J.A.S.A.*, **11**(1940), 278—288.
- [29] Oettinger, R., *Acustica*, **9**(1954), 430.
- [30] 梁之安, 声学学报, **2**(1965), 20.
- [31] M. Jane Collins, John K. Cullen, Jr, *J.A.S.A.*, **63**(1978), 469—473.
- [32] 桑野园子、难波精一郎, 日本音响学会志, **34**(1978), 5号, 316—317.

ON THE RELATIONS OF ACOUSTIC RATIO-REVERBERATION TIME-SPEECH ARTICULATION, EXPERIMENTS AND THEORADICAL CALCULATION

RAO Yu-an*

In the earlier studies on the speech articulation in reverberant sound field in room covered only the relations between the articulation score S and the reverberation time T .

The role played by the acoustic ratio R is investigated on a quantitative basis. We have designed the experiment in a special laboratory to obtain a normalized experimental curve of the articulation vs the product RT . The curve is expressed by a form of efficiency of acoustic ratio-reverberation time K_{rR} as

$$K_{rR} = 1.0 - 0.22 \lg (TR)$$

$$\text{or } S\% = 98.7 - 21.7 \lg (TR)$$

It is surprising to note that form of this physical law is analogy to, but more complicated than Weber-Fechner's law, because "reverberation" itself is a complicated process including space and time.

Comparison of our results to the classical S - T curves of Knudsen, we have found good agreement, if the decrease of R with increasing reverberation time is taken in to account and decrease of S with the increase of T should be slowed.

Bolt had calculated the articulation for reverberant masking on the model of equal probability of the occurrence of speech amplitude.

On the basis of previous experiments on acoustic ratio, we have calculated the relation for both equal probability and real first probability distribution of speech sound, the shape of the curve derived from theoretical calculation coincides with that obtained from experiments.

The binaural hearing helps to decrease the effective acoustic ratio and to increase the articulation. This experiment is confirmed to this viewpoint. And we recommend that articulation be expressed as

$$S = 98.7 K_l K_n K_{rR} K_{d(t)} K_e$$

*Now at Institute of Physiology, Academia Sinica.