

中华人民共和国国家标准

GB/T 15485—1995

声学 语言清晰度指数的计算方法

Acoustics—Methods for the calculation
of the articulation index of speech

1995-07-03 发布

1996-02-01 实施

国家技术监督局 发布

中华人民共和国国家标准

声学 语言清晰度指数的计算方法

GB/T 15485—1995

Acoustics—Methods for the calculation
of the articulation index of speech

1 主题内容与适用范围

由于语言清晰度试验繁复费时,现已发展了一种与语言可懂度高度相关的可计算的物理度量,象组织一组发音人和听音人来进行言语感知试验一样的评价语言可懂度。这个物理度量称之为清晰度指数,简称 AI。AI 是在给定的语言通道和噪声条件下,表示对一个听音人获得语言可懂度可利用的正常语言信号的有效比率(部分),它是一个加权的分数,由测量或估计的语言频谱和在听音人耳边存在的噪声的有效掩蔽谱可计算 AI。

本标准描述了计算清晰度指数 AI 的方法和步骤并给出 AI 与音节清晰度之间的函数关系。

本标准所述的计算方法的根据是男女成年正常发音人,以汉语普通话发音和男女成年正常听音人的平均结果。本方法中的数据不适用于儿童。

本标准所述计算清晰度指数 AI 的方法,部分采用美国国家标准 ANSI S3.5—1969《语言清晰度指数的计算方法》。

2 术语

2.1 清晰度指数 articulation index

通过大量语言清晰度试验,导出的具有频带可加性的用来计算给定的语言传递系统语言可懂度的一个指数。它取值在 0 与 1 之间。

2.2 长时均方根声压谱 long-term root mean-square (rms) pressure spectrum

语言信号的声压振幅与频率的函数关系。可采用多种方法来测量。在 1/1 倍频带和 1/3 倍频带语言分析中,对于正常连续语言,采用 1 s 的积分时间,可得到稳定的长时频谱。所测得的均方根值的数字表示称为长时均方根声压,以区别于采用较短积分时间,比如一个语音的平均时长 1/8 s 测得的均方根值。

2.3 谱级 spectrum level

信号在某一频率的谱级是以该频率为中心 1 Hz 带宽内测得的声压级,以分贝表示,基准声压 20 μ Pa。

当测量滤波器带宽为 Δf (Hz)时,连续语言的声压谱级等于测得的频带声压级减去 $10 \log_{10} \Delta f$ 。

2.4 1/1 倍频带和 1/3 倍频带谱 octave and one-third octave band spectrum

当测量采用 1/1 倍频带或 1/3 倍频带滤波器时,以分贝计的频带声压级与 1/1 倍频带或 1/3 倍频带的中心或边界频率的函数关系,称为 1/1 倍频带或是 1/3 倍频带谱。

注:① 一个滤波器的均方根声压级数值,是指在它的几何中心或中心频率。

② 一个滤波器通带的上下截止频率是指滤波器正弦信号响应最大值频率上下低于最大响应 3 dB 处的频率。

③ 本标准要求所用滤波器衰减曲线的斜率不小于每倍频程 18 dB。

国家技术监督局 1995-07-03 批准

1996-02-01 实施

2.5 语言峰值 speech peaks

1/8 s 平均时间语言信号的均方根值有百分之一超过长时平均值 12 dB 或更多。采用长时平均有效值加 12 dB 作为对语言可懂度有贡献的语言峰值振幅。

2.6 总声级 overall level

用 C 计权测得的声压级。长时总均方根声压级近似等于语音平衡词表各个词的峰值声压级的算术平均减去 3 dB。测量时声级计置于通信系统传声器所在位置,要是对面交谈则置于听音人耳所在位置。声级计测量时应在安静条件下,放在慢档和 C 计权。读语音平衡词时应在正常条件下发音。

2.7 截幅 peak clipping

当输入于一个放大器的瞬时电压,超过某一限定的线性增益值时出现的效应。

2.8 连续谱声音的听阈 threshold of audibility for sounds having continuous spectra

在安静环境中采用连续谱声音进行听力测试,50%的试验能引起听觉感受的最小信号声压级。

2.9 语音平衡词试验 PB(phonetically balanced) word test

语音平衡词试验使用单音节词表,每表 75 个音节(词)。这些音节是经过精选的,它们所包含的基本语音的比例与日常口语大致是一样的。

2.10 频带感觉级 band sensation level

频带声压级与该频带的听阈之差,以分贝来表示。

3 计算方法

计算 AI 可采用下述两种方法。

20 频带法:本方法基于在 20 个相邻的等清晰度频带(见表 1)中测量或估计的语言谱级和噪声谱级。在安静环境中,语言峰值的谱级超过听阈 30 dB 或更多,每一频带中的语言成分对语言可懂度的贡献都是相等的。

表 1 汉语 20 个等清晰度频带

Hz

频带号	频带边界	频带宽度	中心频率
1	200~400	200	300
2	400~550	150	475
3	550~730	180	640
4	730~900	170	815
5	900~1 020	120	960
6	1 020~1 150	130	1 085
7	1 150~1 270	120	1 210
8	1 270~1 400	130	1 335
9	1 400~1 520	120	1 460
10	1 520~1 700	180	1 610
11	1 700~1 900	200	1 800
12	1 900~2 100	200	2 000
13	2 100~2 400	300	2 250
14	2 400~2 700	300	2 550
15	2 700~3 000	300	2 850

续表 1

Hz

频带号	频带边界	频带宽度	中心频率
16	3 000~3 400	400	3 200
17	3 400~4 000	600	3 700
18	4 000~4 700	700	4 350
19	4 700~5 700	1 000	5 200
20	5 700~7 000	1 300	6 350

1/3 和 1/1 倍频带法:它是由 20 频带法导出的,只是要求测量或估计语言和噪声是在 1/3 或整个倍频带。

注:① 下面章节将详述用上述方法计算 AI。与计算有关的函数——理想语言谱、听阈、不截幅语言的最大允许声级,在图 1 和表 5 中给出。这些函数的根据是:汉语标准频谱和国家标准:GB 4983 自由场纯音标准等响曲线。表 5 中对于 12 dB 和 24 dB 的截幅语言的最大容许声级的曲线分别在不截幅语言的最大容许声级之上 6 dB 和 12 dB(参照图 1)。

② 1/1 倍频带法对语言和噪声谱的变化不如 20 频带法和 1/3 倍频带法灵敏,因而也不精确。当掩蔽噪声的能量明显集中在一个倍频带或更窄的范围内时,1/1 倍频带法不可使用。这时应采用 1/3 倍频带法,最好采用 20 频带法。

③ 在文献中报告结果时,一定要说明所使用的 AI 计算方法,即 AI(20 频带),AI(1/3 倍频带)或 AI(1/1 倍频带)。

3.1 20 频带法

3.1.1 步骤 1

在计算图(见图 1)上绘出已知的或估计的听音人耳处的语言峰值谱级。

语言峰值谱级可用下列诸项代数和近似得到。

3.1.1.1 所评定的系统的频率响应,以分贝计。每一频率的频响是听音人耳处的声压级与发音人在传声器处在该频率上的声压级之差,以分贝计。

注:必须注意保证频率响应完全反映整个系统的发送和接收换能器的特性。

3.1.1.2 理想化语言频谱的确定

a. 图 1 理想化频谱是在长时总均方根声压级为 65 dB 时测得的。当测得的或估计的长时总均方根声压级与 65 dB 不同时,按两者的差值上下移动曲线。

注:图 1 中理想化频谱严格应用在距发音人唇部 1 m 远基本上无混响和无噪声的环境中。在这些条件下,在发音人唇前 2.54 cm 到 1 m 之间测量语言频谱时,这个频谱的形状都相当精确。因此,语言声级可在靠近发音人唇部处测量(或估计),所得数值可依平方反比定律换算成 1 m 处的值,并且假定一个等效点声源是在发音人唇后 0.6 cm 处。这个导出值就可以与 65 dB 来比较,以得出理想化语言频率谱级所需的调整值。

b. 当语言是在一非消声室,非自由声场中以扬声器放声时要按表 2 进行校正。

表 2 对混响或半混响房间中的扬声器所放出的语言声级的校正

语言总声级, dB	应从语言声级中减去的校正量, dB
85	0
90	2
95	4
100	7
105	11
110	15

续表 2

语言总声级, dB	应从语言声级中减去的校正量, dB
115	19
120	23
125	27
130	30

注: 表 2 中给出的校正量不适用于语言通过耳机发声或在自由声场中以扬声器发声。

这个校正量是根据实验得到的, 它表明在这种情况下, 语言声级的提高会引起可懂度降低。例如, 假定在一混响房间中, 一扩声系统的长时语言总声级为 95 dB, 就要向经过按 3.1.1.1 条所述系统频率响应特性校正的理想化语言频谱加上 26 dB, 即 95 dB-65 dB-4 dB(表 2 校正量)。从而得出这个扩声系统的有效语言谱。

3.1.2 步骤 2

把到达听音人耳处的稳态噪声校正谱级画到图 1 上。几种噪声源, 例如, 听音人所在的环境噪声和通过语言传递系统到达听音人的噪声, 均方根声压要加在一起。

3.1.2.1 校正噪声谱

当噪声的频带感觉级超过 80 dB 时, 噪声掩蔽的有效性要比正常比率增加得更快。在 AI 计算中这种增加的掩蔽可对噪声声压级增加一个校正量来加以考虑。一旦噪声感觉级在一个频带的中心频率上(图 1 中的那些垂直线)超过 80 dB, 则在这一点上的噪声声压级便按表 3 所给的数值增加一个适当的量。

噪声感觉级可由噪声谱级减去听力阈限谱级来确定。

表 3 对掩蔽非线性增长的校正量

频带感觉级, dB	应加到噪声声压级的校正量, dB
80	0
85	1
90	2
95	3
100	4
105	5
110	6
115	7
120	8
125	9
130	10
135	11
140	12
145	13
150	14

3.1.3 步骤 3

在计算图(图 1)上绘出噪声有效掩蔽谱级。在任一频率上的有效掩蔽谱是由噪声谱、校正噪声谱或

者在该频率上的延伸掩蔽噪声谱中的最大的一个来决定。掩蔽噪声谱的延伸由下述方法画出来。

3.1.3.1 首先确定噪声掩蔽谱延伸,在最右边纵轴上找出比噪声谱或校正噪声谱的每个最大值低3 dB的点,再由最高的点作水平线与噪声谱相交。这些点称为“起始点”。

注:① 噪声谱只有一个最大或峰值的时候那就只有一个起始点。

② 要是噪声谱在7 000 Hz或更高频率有峰值,则起始点就定在比5 700 Hz低3 dB处。

3.1.3.2 从每一个起始点垂直下降57 dB,再以10 dB/oct的上升斜率向左画一直线,这直线便是掩蔽谱的低频部分。

3.1.3.3 从每一个起始点向右画水平线到一定长度后再以一定斜率下降。水平部分的长度与下降的斜率、每个起始点所在的频率以及在该频率上的噪声最大谱级,见表4。这些条线代表延伸掩蔽谱的高频部分。

男女声平均20个等清晰度频带中心频率(见表1)

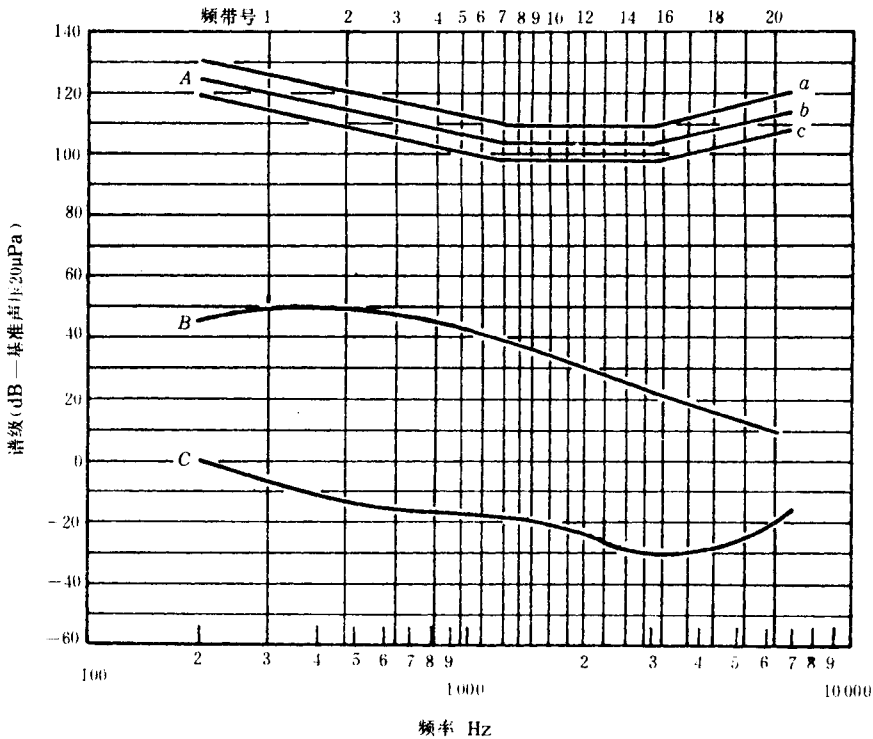


图1 20频带法AI计算图

- A 最大容许声级;a—语言峰值截幅24 dB;
b—语言峰值截幅12 dB;
c—不截幅语言。

比上述曲线高的语言峰值对语言可懂度无贡献。

- B 汉语普通话理想化语言频谱——男女声
长时频谱(rms)+12 dB(长时rms总声级为65 dB)。
- C 连续谱声音的听阈谱级。

表 4 掩蔽谱的高频段——掩蔽的上延伸

噪声最大谱级或 校正后的谱级, 用其中较高者 (基准声压 20 μ Pa) dB	起始点频率 ³⁾									
	50~800 Hz		800~1 600 Hz		1 600~2 400 Hz		2 400~3 200 Hz		3 200 Hz~7 000 Hz	
	A ¹⁾ Hz	B ²⁾ dB	A Hz	B dB	A Hz	B dB	A Hz	B dB	A Hz	B dB
96	250	10	500	8	1 000	5	1 500	3	3 000	0
86~95	200	15	500	13	1 000	10	1 500	5	3 000	0
76~85	200	20	400	18	800	15	1 500	10	3 000	0
66~75	150	25	250	23	500	20	1 000	15	2 000	5
56~65	75	35	150	30	300	25	500	25	800	20
46~55	50	45	100	40	200	35	200	40	200	40

注: 1) A 从起始点向右画水平线的频率(Hz)数;

2) B 从水平线右端点往下画斜线的斜率(dB/oct);

3) 起始点频率值的确定参见 3.1.3 条。

3.1.4 步骤 4

在 20 频带每个频带的中心频率均指示在图 1 上, 确定出语言谱级和有效掩蔽谱级的分贝差值。当差值为 0 或小于 0 时, 就定其为 0, 当语言谱级超出有效掩蔽谱级 30 dB 时, 就定其为 30。

注: ① 在计算图(图 1)上听阈曲线高于有效掩蔽谱的部分就当作最小噪声谱。

② 当语言峰值曲线超出图 1 中所标的最大容许声级时, 则最大容许声级就作为语言峰值声级。

3.1.5 步骤 5

将在步骤 4 中所求得的 20 个差值加起来除以 600, 所得到的数值便是给定通讯系统在噪声条件和对所给定语言声级下的清晰度指数。

3.1.6 举例

用 20 频带法计算 AI 的一个说明性的例子示于图 2。

男女声平均20个等清晰度频带中心频率(见表1)

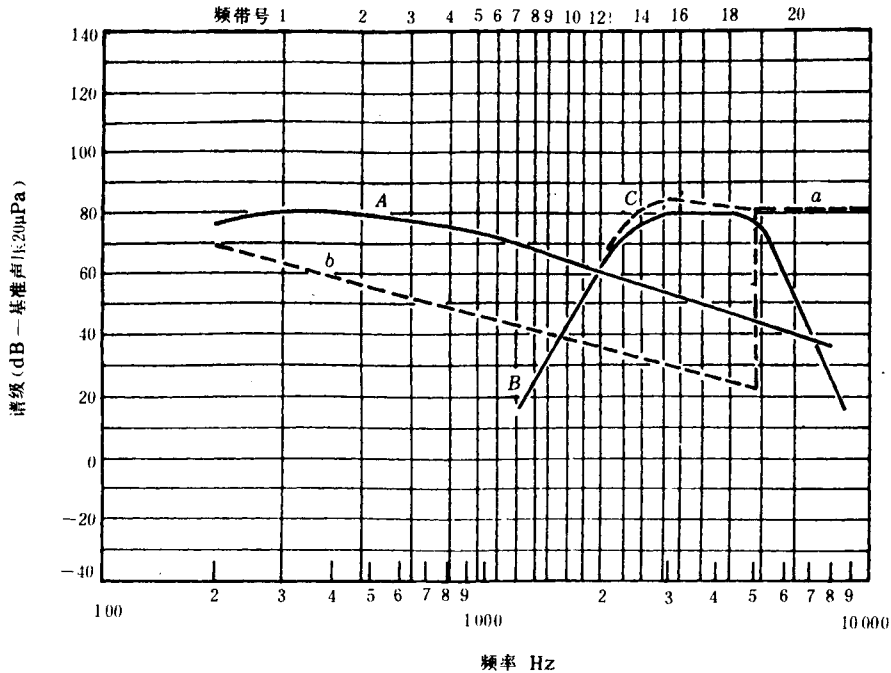


图 2 20 频带法计算 AI 示例

- A 语言峰值谱级——长时频谱(rms)+12 dB
(长时 rms 总声级 95 dB)。
- B 噪声谱级(总声级 113 dB)。
- C 噪声掩蔽级(见表 3):
 - a 掩蔽向上延伸;
 - b 掩蔽向下延伸。

3.2 1/1 倍频带法和 1/3 倍频带法

3.2.1 步骤 1

根据所用的带通滤波器测定出到达听音人耳处的语言声压级。

注:1/3 倍频带和 1/1 倍频带滤波器的中心频率在表 5 中给出。

语言峰值的频带声压级可用下列数值的代数和来近似:

频带号	语言峰值与噪声或听阈掩蔽的差
1	17
2	23
3	25
4	26

频带号 语言峰值与噪声或听阈掩蔽的差

5	26
6	26
7	26
8	26
9	25
10	20
11	9
12	}
13	
14	
15	
16	
17	
18	
19	
20	

$$\Sigma = 239$$

$$AI = 239 / 600 = 0.40$$

图 2AB 中心频率处读值与 AI 计算

3.2.1.1 被评价系统以分贝表示的频率响应特性。在每一中心频率处的频率响应是听音人耳处的频带声压级与在发音人传声器处该频率的声压级之差(应注意保证频率响应要完全反映整个系统的发送和接收换能器的特性)。

3.2.1.2 计算理想化语言频谱

a. 按着测得的或估计的语言长时总声级(rms)与 65 dB 之差来移动理想语言频谱,即在表 5 中相应的频带的理想化语言频谱数值加或减去这一差值。

b. 在一个非自由声场或非消声室中以扬声器放声时,语言长时总声级要按表 2 指示的数值来降低(表 2 中给出的校正值不适用于通过耳机呈现语言或在自由声场中使用扬声器放声)。

注:表 5 中的理想化语言频谱严格地适用于在基本上无混响无噪声环境中距离发音人唇部 1 m 处。

3.2.2 步骤 2

计算出到达听音人耳处的稳态噪声的频带声级。从几个声源,例如,听音人所在的环境噪声和通过语言传递系统到达听音人处的噪声,其有效声能要加在一起。

注:当在一个对可懂度有贡献的频带的中心频率上 1/1 倍频带或 1/3 倍频带的感觉级超过 84 dB 时,语言和噪声频谱就要变换成谱级的值(见 2.3 条),并将其结果画到图 1 上。这样就要采用 3.1.1 到 3.1.5 条所述的 20 频带法来计算 AI。变换成谱级的目的是为了能在计算 AI 时考虑到非线性延伸掩蔽效应。这种延伸掩蔽效应在频带

感觉级超过 84 dB 时就变得显著。

表 5 理想化语言频谱+12 dB、有效听阈频谱和对连续谱声音不截幅语言的最大容许声压级

20 个等清晰度频带中心频率的				1/3 倍频带中心频率					1/1 倍频带中心频率					
谱 级				频带声压级					频带声压级					
中心频率	A	B	C	中心频率	A	B	C	W	中心频率	A	B	C	W	
Hz	dB	dB	dB	Hz	dB	dB	dB		Hz	dB	dB	dB		
1	300	50.0	-7.0	114	200	67.0	16.0	138	0.000 4	250	72.5	19.0	140	0.002 4
2	475	50.0	-14.0	109	250	68.0	14.0	135	0.001 0	500	74.0	14.0	136	0.004 8
3	640	48.0	-16.0	105	315	69.0	12.0	132	0.001 0	1 000	68.0	12.0	129	0.007 4
4	815	45.0	-16.0	103	400	70.0	10.0	131	0.001 4	2 000	62.0	10.0	131	0.010 9
5	960	43.0	-16.0	101	500	68.5	9.0	129	0.001 4	4 000	57.0	10.0	140	0.007 8
6	1 085	41.0	-16.0	100	630	66.5	7.0	127	0.002 0	—	—	—	—	—
7	1 210	39.0	-17.0	99	800	65.0	8.0	124	0.002 0	—	—	—	—	—
8	1 335	37.0	-17.5	99	1 000	64.0	8.5	122	0.002 4	—	—	—	—	—
9	1 460	36.0	-19.0	99	1 250	62.0	8.5	122	0.003 0	—	—	—	—	—
10	1 610	34.0	-20.0	99	1 600	60.5	8.5	123	0.003 7	—	—	—	—	—
11	1 800	32.0	-22.0	99	2 000	59.5	5.5	125	0.003 7	—	—	—	—	—
12	2 000	30.0	-24.0	99	2 500	58.0	3.0	128	0.003 4	—	—	—	—	—
13	2 250	28.0	-26.0	99	3 150	56.0	1.0	130	0.003 4	—	—	—	—	—
14	2 550	26.0	-28.5	99	4 000	53.0	5.0	132	0.002 4	—	—	—	—	—
15	2 850	24.0	-29.0	99	5 000	51.0	12.0	135	0.002 0	—	—	—	—	—
16	3 200	22.0	-30.0	99	—	—	—	—	—	—	—	—	—	—
17	3 700	19.5	-30.0	100	—	—	—	—	—	—	—	—	—	—
18	4 350	16.5	-29.0	102	—	—	—	—	—	—	—	—	—	—
19	5 200	13.0	-26.0	104	—	—	—	—	—	—	—	—	—	—
20	6 350	10.0	-20.0	106	—	—	—	—	—	—	—	—	—	—

注：A 为理想化语言频谱+12 dB； B 为有效听阈频谱； C 为连续谱不截幅语言的最大容许声级； W 为计算 AI 的加权值。

3.2.3 步骤 3

计算每一频带的中心频率上语言峰值频带声压级与噪声的频带声压级的分贝差值(D)。要是这一差值为 0 或小于 0,就定差值是 0。要是语言频带声压级超出噪声频带声压级 30 dB 或更多,则定此差值为 30。

注：① 当听阈曲线超过噪声频带级的地方就把听阈当作最小等效噪声频带级。

② 语言峰值超出最大容许声级的地方,最大容许声级就作为语言峰值声级。

3.2.4 步骤 4

根据表 5 W 列所列计权值(W)与 3.2.3 条步骤 3 中得到的差值(D)相乘。

3.2.5 步骤 5

将 $D \times W$ 的数值相加。得数便是这一语言系统在给定的噪声条件和指明的语言声级下工作的 AI。

4 各种因素对 AI 的影响

4.1 AI 所评价的因素

语言清晰度试验得分受施加于所传递的语言信号上的诸多条件的影响,至今仍不完全清楚也不能完全定量。然而,在运用 AI 原理时,有一些因素是可以定量估计的。它或是个别出现或是结合在一起,除下述 4.2 条以外。这些因素如下。

4.1.1 稳态噪声掩蔽

AI 可预测宽带连续噪声的影响,以及频率范围在 200~7 000 Hz 频带宽度大于 200 Hz 的噪声的影响。

4.1.2 非稳态噪声掩蔽

非稳态噪声对语言可懂度只在一部分时间内有影响称为作用周期。当噪声不是稳态的并且起-止作用周期是已知的,可以把噪声当作是稳态的来计算,以后应用图 3 对所得到的 AI 值加以校正就可得到有效的 AI 值。这种方法只适用于噪声在作用周期中起始期的噪声级高于停止作用时期 20 dB 以上的情况。

4.1.2.1 噪声中断率

对一有着确定的起-止作用周期的噪声的通讯系统,按照 4.1.2 条和图 3 的方法算出的有效 AI,还要进一步根据图 4 中所给的函数来做调整。图 4 上横坐标是噪声的每秒中断次数,纵坐标是在给定参数下的有效 AI。

4.1.3 语言信号的频率畸变

频率畸变是信号传递时对不同频率的增益不相等。它对语言可懂度也是有影响的。这些影响利用 AI 可以考虑得相当精确。计算 AI 时对下列因素给以不同的侧重。

- a. 语言信号的高频成分;
- b. 低频成分;
- c. 中频成分。

可是,要是语言信号的长时频谱很不规则,也就是频谱有一系列峰和谷,平均的峰谷斜率超过 18 dB/oct 时,AI 对语言可懂度的估计就不太可靠了。

4.1.4 语言信号的振幅畸变

计算的 AI 可以估计明显的对称截幅对语言可懂度的影响,可采用下列步骤。

4.1.4.1 步骤 1

由图 5 来确定由一特定的截幅值和在一个系统中截幅后的放大量所造成的语言长时均方根声级的提高。

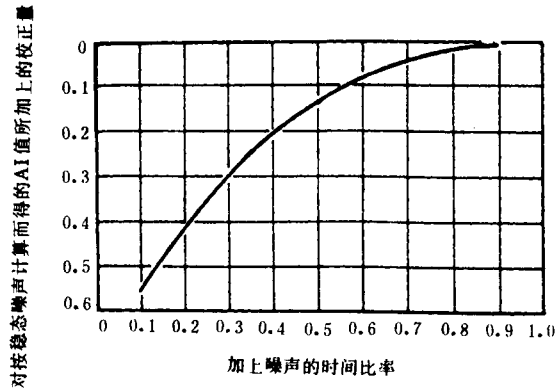


图3 对不同噪声持续因数的校正

注：纵坐标表示对不同噪声时间分数按照稳态噪声掩蔽所得到的 AI 值应加的修正量，AI 修正值不能超过 1.0。

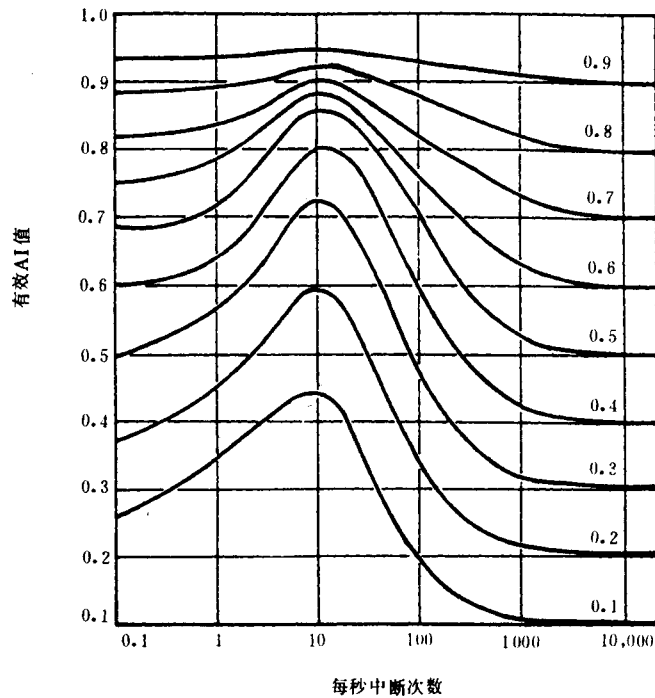


图4 有效 AI 与中断频率的关系(从图 3 读 AI 校正量)

注：该图表示有效 AI 值与掩蔽噪声中断次数的关系曲线上的系数是对稳态噪声掩蔽计算所得的 AI 值按图 3 对噪声持续时间予以修正后的结果。

4.1.4.2 步骤 2

将步骤 1 的结果加到语言峰值声级(非截幅语言的长时均方根+12 dB)这便是听音人耳处非截幅语言和截幅后放大语言的声级。

注：截幅后放大定义为对通信系统所加的放大量,使得峰-峰振幅达到与非截幅语言的振幅相等。语言的峰值振幅

定义为有百分之一时间超过的振幅。如果截幅后放大量,以分贝计,不等于对语言信号所加的截幅值,在步骤 1 中得出的长时均方根值的提高量就要降低一个量。这个量等于峰值截幅与截幅后放大量之差。

4.1.4.3 步骤 3

将步骤 2 的结果绘于 AI 计算图上,依次按照以前所述计算 AI。注意 AI 计算图上的最大容许声级对于截幅语言要比非截幅语言高。

注:一般来说,峰值截幅仅当语言信号在截幅前相对地没有噪声而在收听处有噪声或者在截幅后有噪声混入时才使用。

4.1.5 混响

房间中的混响将造成语言可懂度降低。降低的程度与房间的混响时间有关。混响时间定义为声源停止后一个稳态声降低 60 dB 所需的时间。

当混响时间已知时,可利用图 6 来校正对一个给定通讯系统求得的 AI 这个校正量要加到由表 2 修正后所得到的 AI 值上。

4.1.6 说话声级

一个说话人声级很弱或很强都将降低语言可懂度。在其他因素保持不变的情况下,要是说话声级平稳保持在 50 dB 到 85 dB(距说话人唇部 1 m 远测得长时均方根声压级)之间,一个给定的 AI 值可以是精确的。要是通讯系统中使用很弱或很强的说话声级,在 AI 计算图上就得画出有效声级而不是测得的语言声级。实际语言声级和有效语言声级之间的关系示于图 7。

4.1.7 视觉线索

由观察说话人的双唇和面部而得到的视觉线索对语言可懂度有很大帮助,特别是在有噪声存在的时候。可是,可以改变或调整一个 AI 使之成为一个“有效 AI”以反映一般没受过“唇读”技术训练的听话人的视觉线索对语言可懂度的影响(见图 8)。

4.2 AI 不能估计的因素

有不少影响语言通讯系统的因素在现行的 AI 中不能估计,特别是下述几项。

4.2.1 说话人性别

如前所述本方法是根据男女发音人进行清晰度试验的平均结果而设计的。因而,当个别男或女说话人使用语言通讯系统时的语言可懂度可能与估计值略有出入。

4.2.2 多传递通道

听话人从说话人直接接收的与同时还从一个扬声器接收的混合语言信号对语言可懂度的定量影响还不清楚。因而,AI 可能不适用于这种系统。

4.2.3 多因素组合

当几种失真综合出现,如截幅、加中断噪声、加混响,目前还没有充分试验足以说明在可能的因素组合时,确切 AI 值应如何。

4.2.4 非对称截幅,频率偏移和衰落

本标准只适用于通讯系统的非对称截幅不超过 3 dB,信号的频率偏移不超过 50 Hz,并且没有明显的衰落变化。

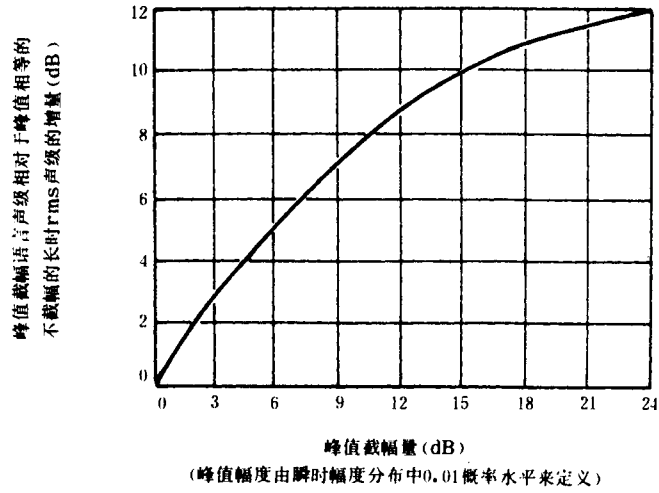


图 5 语言声级与峰值截幅的关系

注：该图表示当截幅级上升到截幅参考级时，rms 语言级增量与截幅量的关系。

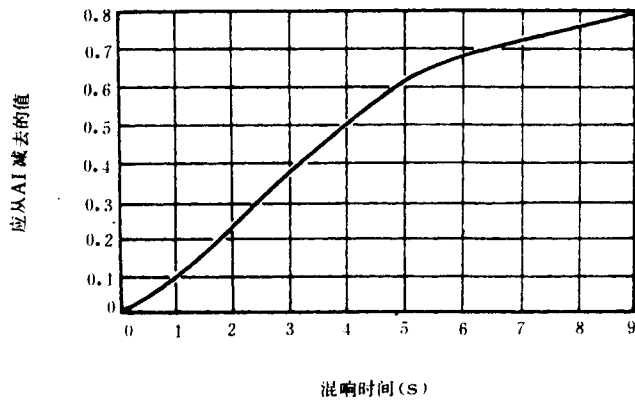


图 6 AI 的混响时间校正

注：该图表示 AI 的校正量，AI 不能小于 0.0，当校正后的 AI 小于 0.0 时按 0.0 计算。

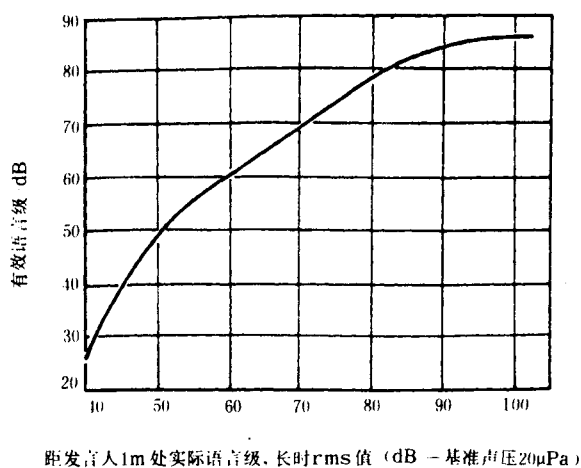


图 7 发音人实际语言级与有效语言级的函数关系

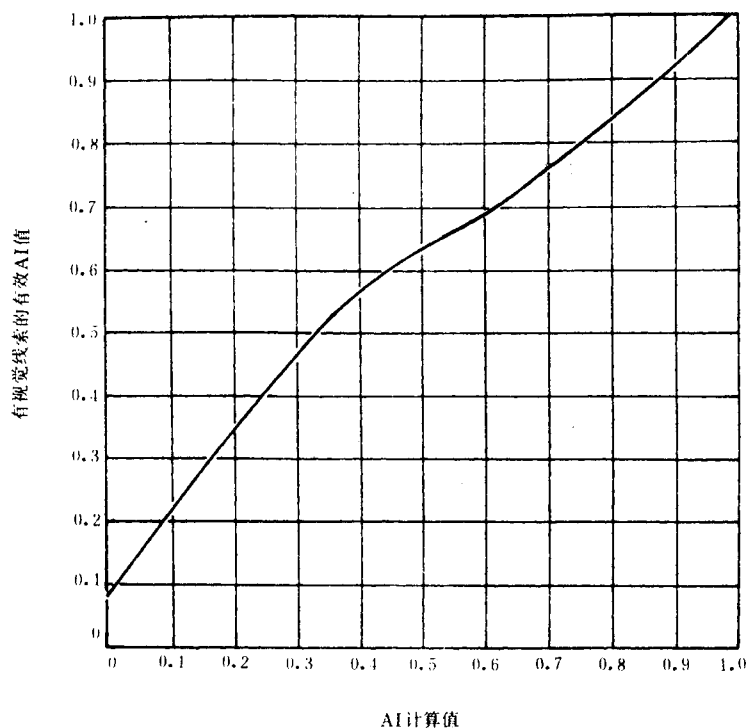


图 8 在听音人可看到发音人唇与面部的通讯系统中 AI 计算值与有效 AI 值的函数关系

4.2.5 单耳呈现与双耳呈现

对于有正常双耳的人,单耳或双耳倾听,测得的语言可懂度是可以比较的,并且也是可用 AI 预测

的。然而,当听话人用双耳倾听来自分离声源的语言和噪声时,由于在给定的各频率上双耳对于语言和噪声产生的相位关系是不同的,这样用双耳听到的语言可懂度就会与用单耳听到的不同。除了在 4.2.1 到 4.2.4 条中所提到的情况之外,AI 是可以预测在各种双耳倾听和所有单耳倾听条件下的相对可懂度的。

5 AI 与语言可懂度的关系

利用图 9 中所给的曲线,就可以把 AI 转换成估计的语言可懂度。图 9 中所给出的曲线仅对一组给定的发音人和听音人是有代表性的。应该特别注意的是,清晰度得分与对被传递的消息所加的约束有很大关系的。典型的约束包括语法结构和上下文(例如,在句子当中),词汇量大小,以及单词中的音节数。没有一个单值的 AI 可用来作为“可接受的”通讯的标准。图 9 中所示的通讯效率,也就是清晰度试验得分与 AI 的函数关系,与所传递的消息的范围和发音人与听音人的熟练程度有关。当使用一个给定的通讯系统在给定的条件下,用不同的发音人和听音人小组或者同一个发音人和听音人小组在不同的时间,得到的清晰度试验得分有差别,这并不表示 AI 不可靠。

如前所述,AI 是用于预测通讯系统的相对性能的。通讯系统是在一定条件下工作,测试时使用指定的典型的发音人和听音人小组,并且训练时听音人-发音人熟练因数保持恒定,或者对于不同的测试,要由恰当的试验程序来加以控制。AI 已经利用这种试验证实是可靠的了,其可靠性是依据在一次指定的试验中随着 AI 值的提高测得的清晰度试验得分也稳定一致地增长。

当然,对一指定的通讯系统所要求它的性能水平如何是与多种因素有关的。这些因素的重要性,只能由通讯系统的使用情况来评定。

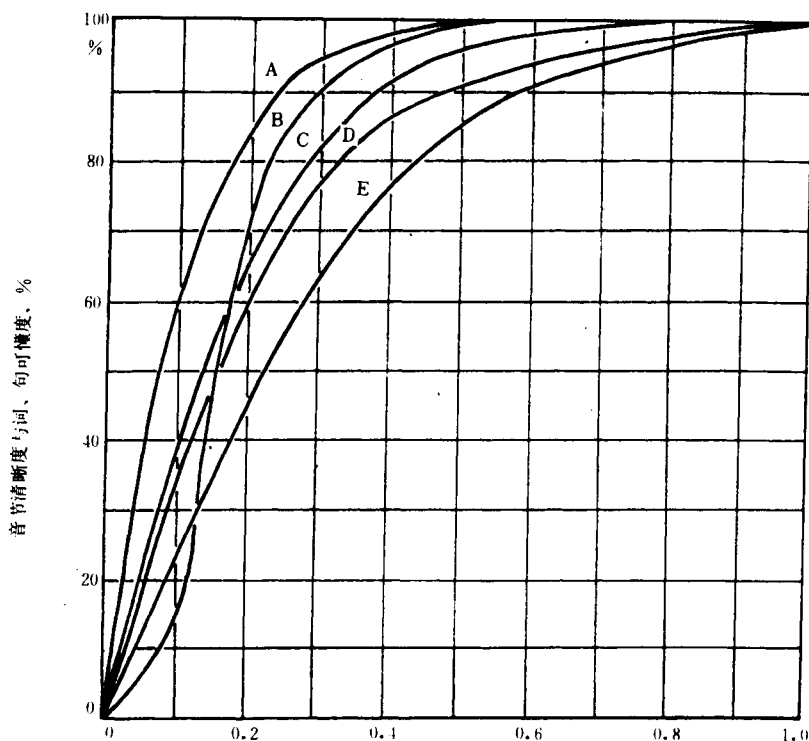


图 9 音节清晰度、单词可懂度和单句可懂度与 AI 的函数关系(男女平均)

A—三音节词(语音平衡词表中 48 个词); B—单句; C—双音节词;
D—单词(包括不同词的长度); E—音节