# 2 Introduction to large room acoustics

"The science of sound is called **acoustics**, a word derived from the Greek ***akoustos***, meaning "hearing". Beginning with its origins in the study of mechanical vibrations and the radiation of these vibrations through mechanical waves, acoustics has had important applications in **almost every area of life**" (from Encyclopedia Britannica).

The ultimate design goal of any concert-hall is of course good acoustics; or rather appropriate acoustics for the purpose of the building. It is by general standards relatively easy to distinguish between ”bad” and ”good” acoustic conditions. Especially for rooms intended speech or reinforced music, it is in general relatively easy to characterize the acoustic behavior of the room by means of acoustic measurements. There are for instance fairly reliable ways of measuring speech intelligibility.

It is however not possible to distinguish between ”good” and ”excellent” acoustics for music by any objective means. Halls may show the same measured values but sound entirely different.

Acoustics can in general be evaluated in two ways:

- By objective means, that is by means of physical measurements or model simulations. The result of this will be more or less single number quantities, as for instance reverberation time or clarity.
- By subjective means, that is by response from the performers and audience in the form of acoustic quality descriptors.

In the following, a short introduction to room acoustics and room acoustic parameters is presented.

## 2.1 Basics of room acoustics

### 2.1.1 Impulse response

From a room acoustic point of view, sound can be thought of as sound particles, reflected of the boundaries of the room. This analogy is acceptable for high frequencies, or rather for frequencies that have a short wavelength as compared to the dimensions of the room.

The perhaps easiest way to grasp the behavior of a sound field on in a room is by looking at the impulse response. If a short pulse is sent from a source and the response is recorded at a receiver point, one will first see the direct sound, followed by a number of spaced reflections. After a while, reflections will come at such a high density that it is impossible for track the individual reflection, hence the sound field can best be described by statistics.

This is called the impulse response or when measured with an “impulse like” source, the reflectogram. As all the measurements described in this work are done with measurement systems giving the actual impulse response, only the actual impulse response will be used from here on. It is however important to realize that some of the parameters described later can not be calculated from the “reflectogram” but only from the actual impulse response, whereas most energy parameters also can be calculated from the reflectogram.

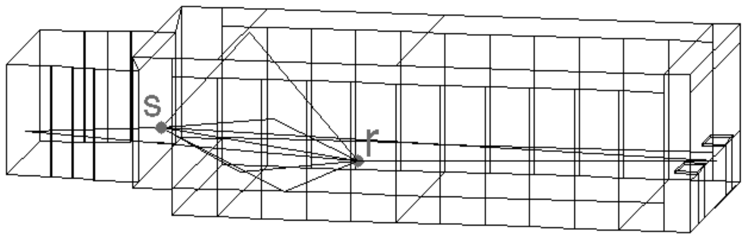


*Figure 2.1: 1st order reflections in a small hall*

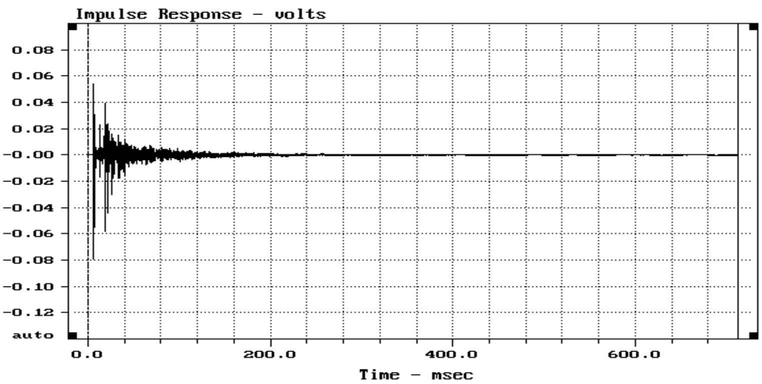


*Figure 2.2: Example of a measured impulse response*

The impulse response can be used to describe most acoustic factors guiding room acoustics and all of the parameters presented later are calculated from the impulse response, either from a monoaurale (one channel) response or from a binaural (two channel response).

In general, the impulse response can be divided into three parts:

- Direct Sound: normally defined as sound arriving within the first 10 ms of the first wavefront

- Early reflections: sound arriving within the first 50-80 ms of the first wave front. These are important both for the spatial impression as well as for the “coloration” of the sound field.

- Late reflection or late field: Sound arriving later than approx 80 ms from the first wave front.

## 2.1.2 Sound reflection

When a sound wave hits a surface it will reflected, absorbed or scattered, or actually a combination of all, see fig 2.3 – 2.5. The strength of the individual reflection will be determinated by the acoustic characteristics of the surfaces from which they are reflected as well as by the length they have traveled.

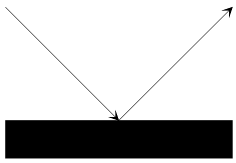

*Figure 2.3: Specular reflection of a perfectly smooth surface*

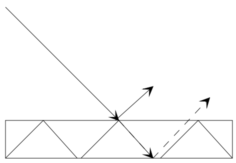

*Figure 2.4: Reflection of a soft absorbing surface*

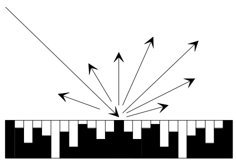

*Figure 2.5: Reflection of a diffusing surface*

When a sound wave hits a perfectly smooth surface it will be reflected in a specular manner. All in-regulariteies in the reflecting surface will however influence the reflection. If reflecting surface has absorbing characteristics, some of the energy will be absorbed but the reflection pattern will also be influenced by the arriving angel of the sound wave. This is in particular true for reflection of a thin plate, whis will be put in vibrating motion by the sound wave, thus the angel of reflection will variate.

Surfaces that are not smooth, will scatter the energy of the sound wave. A scattering surface can regular, for instance curved, or in-regular, like for instance the diffusing surfaces used in studios. From a design point of view it is important to remember that all scattering from surfaces is to some extent frequency dependant. This means that the size of the in-regularities must be in some relation to the wave length of the sound waves. This also implies that a surface that is scattering in one frequency range can produce focusing or very strong specular reflections in another frequency range.

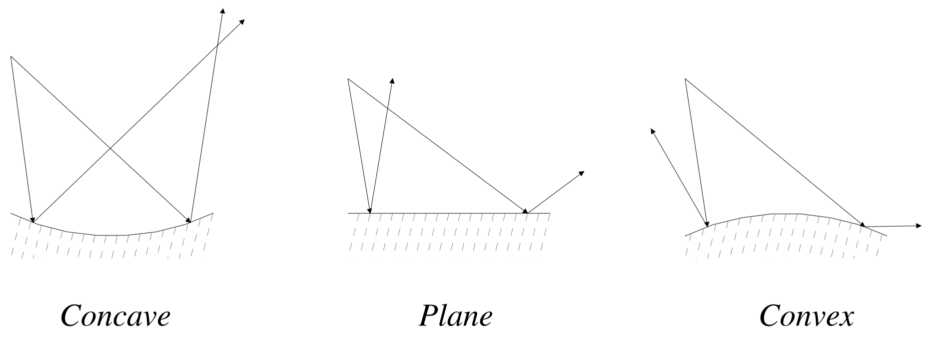


*Figure 2.6: Reflection off different shaped surfaces*

Another special reflection pattern is the so called ”Cats-eye” reflection, which accures when sound is reflected in a perpendicular corner. This is one of the reasons why horizontal detailing is needed on the side walls in halls to provide lateral reflections.

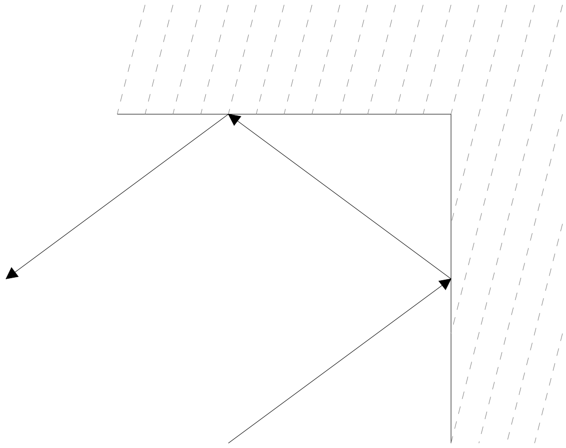

*Figure 2.7: ”Cats-eye” reflection.*

Typical reflection pattern related room acoustic problems are focusing, echoes and flutter echoes.

Especially concave surfaces of large scale are likely to produce focusing effects, which leads to both an uneven sound field and the risk of echoes. It is important to distingues between echoes and reverberation. Echoes are discreetly audible sound events or reflections whereas reverberation is a sound field.

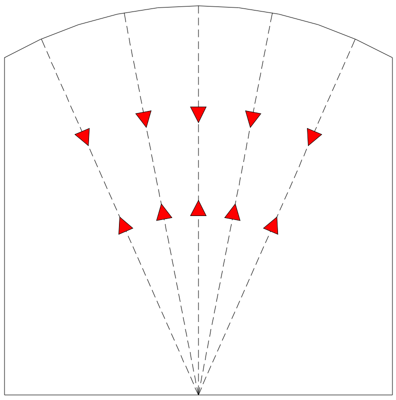

*Figure 2.8: Focusing*

Flutter echo is a short echo, reflected many times for instances between parallel walls. It is important to realize that the worst promlems with flutter echoes are not neccesaily between parallel walls, but other also a situation involving several surfaces or one focusing surface. Contreary to common belief, flutter echoes can be very hard to predict and very hard to get rid of without using absorbing materials.

### 2.1.3 Sound absorption

When wave strikes a surface, a part of the energy will be absorbed by the surface. Typical absorbing surfaces are porouse surfaces, as for instance mineralwool or heavy fabric, thin plates which will absorbe bass energy, as for instance gypsum boards or wood boards, and different kinds of resonator constructions. Typical resonator constructions are perforated or slited panels.

The absorbtion of different materials is given in percent, that is the amount of energy that will stay in the material and is weighted from 0 to 1. The absorbtion characteristisk of different materials will allways vary withe frequency. In general, porouse materials will absorb higher frequencies, wheras both resonator constructions and plates will absorb around the respective resonance frequency. It is important to realise that the room acoustic behavior or absorbtion behavior of material can not be compared to the vibration behavior of materials, in other words materials which makes good intruments are not neccesarily good room acoustic materials.

In concert hall acoustics, the major absorbing areas are the audience areas, and essentially absorbtion is avoided on other surfaces.

*Tabel 2.1: Absorbtion characteristics for some materials*

| Material | 63 | 125 | 250 | 500 | 1000 | 2000 | 4000 | 8000 Hz |
|---|---|---|---|---|---|---|---|---|
| Orchestra with instruments on podium | 0.27 | 0.27 | 0.53 | 0.67 | 0.93 | 0.87 | 0.80 | 0.80 |
| Empty chairs, upholstered with cloth cover | 0.44 | 0.44 | 0.60 | 0.77 | 0.89 | 0.82 | 0.70 | 0.70 |
| Audience, heavily upholstered seats | 0.72 | 0.72 | 0.80 | 0.86 | 0.89 | 0.90 | 0.90 | 0.90 |
| Audience, medium upholstered seats | 0.62 | 0.62 | 0.72 | 0.80 | 0.83 | 0.84 | 0.85 | 0.85 |
| Audience, lightly upholstered seats | 0.51 | 0.51 | 0.64 | 0.75 | 0.80 | 0.82 | 0.83 | 0.83 |
| Smooth concrete, painted or glazed | 0.01 | 0.01 | 0.01 | 0.01 | 0.021 | 0.02 | 0.02 | 0.02 |
| Ceramic tiles or marbel | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Wooden floor on joists | 0.15 | 0.15 | 0.11 | 0.10 | 0.07 | 0.06 | 0.07 | 0.07 |
| Parquet fixed in asphalt, on concrete | 0.04 | 0.04 | 0.04 | 0.07 | 0.06 | 0.06 | 0.07 | 0.07 |
| Minerawool 30mm, 70 kg/m$^3$ | 0.01 | 0.05 | 0.31 | 0.72 | 0.95 | 0.95 | 0.95 | 0,95 |
| Minerawool 50mm, 70 kg/m$^3$ | 0.05 | 0.20 | 0.84 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| Cotton curtains (0.5 kg/m2) | 0.30 | 0.30 | 0.45 | 0.65 | 0.56 | 0.59 | 0.71 | 0.71 |
| Cotton curtains (0.2 kg/m2) | 0.05 | 0.05 | 0.06 | 0.39 | 0.63 | 0.70 | 0.73 | 0.73 |
| Curtains of close-woven glass mat | 0.03 | 0.03 | 0.03 | 0.15 | 0.40 | 0.50 | 0.50 | 0.50 |
| 13 mm plasterboards on laths* | 0.30 | 0.30 | 0.12 | 0.08 | 0.06 | 0.06 | 0.05 | 0.05 |
| 2*13 mm plasterboards on laths* | 0.15 | 0.15 | 0.10 | 0.06 | 0.04 | 0.04 | 0.05 | 0.05 |
| 16-22 mm wood facing on laths* | 0.25 | 0.25 | 0.15 | 0.10 | 0.09 | 0.08 | 0.07 | 0.07 |
| 6 mm wood fibre board on laths* | 0.30 | 0.30 | 0.20 | 0.20 | 0.10 | 0.05 | 0.05 | 0.05 |
| 22 mm chipboard on laths* | 0.12 | 0.12 | 0.04 | 0.06 | 0.05 | 0.05 | 0.05 | 0.05 |

* at least 50 mm void filled with mineralwool

When calculating the absorption of the audience in a hall, one has to take into account that the audience area is not flat, the chairs will normally be between 0,6 m and 1 m high. To compensate for this, the absorbing area of audience is normally calculated as the area of audience + a 0,5 m wide strip around all extents not being ´next to a wall.

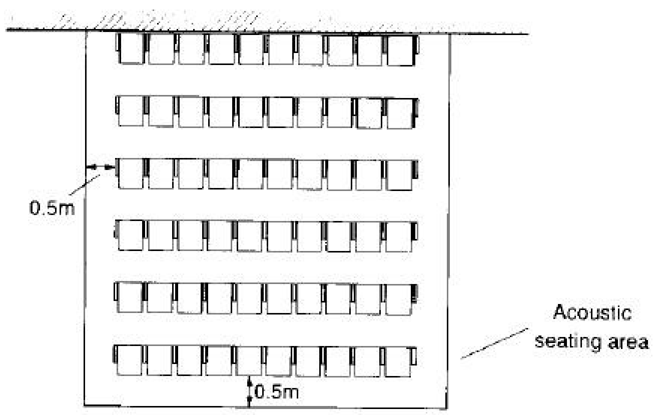


*Figure 2.9: Equivalent absorption area of audience (BARRON)*

### 2.1.4 Reverberation time

This is defined as the time it takes for a sound field to decay 60 dB after the source has been switched off. As it is difficult to achieve an audible 60 dB decay in practice, the measurement is normally done by measuring the decay time from –5 dB to –35 dB and multiplying this time by two or by measuring the decay time from –5 dB to –25 dB and multiplying this time by a factor 3.

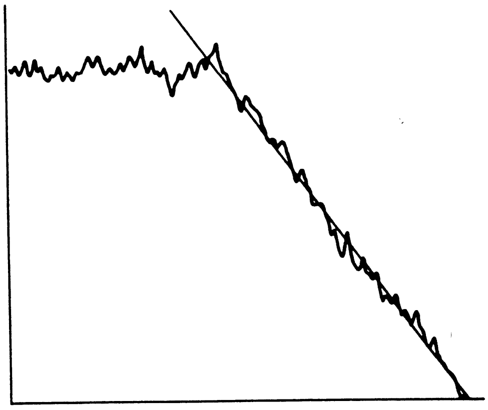


*Figure 2.10:*

The reverberation time can easily be estimated from the room volume and the amount of absorbing material in the room by Sabine's original equation:

$$T = \frac{0{,}161V}{\sum S\alpha + 4mV}$$

Where: T is the reverberation-time in seconds

V is the volume in $m^3$

S is the absorbing surface in $m^2$

$\alpha$ is the absorption coefficient of the surface (see table 1.1)

m is the absorption of air

### 2.1.5 Sound level

Under anechoic conditions, that with just direct sound, the sound level produced by any sound source, will depend only on the sound power of the source and of the listener distance from the source. In a room with sound being reflected of the surfaces of the room, the sound level will obviously depend on the amount of absorption. Traditionally the sound level has been calculated from:

$$L = L_o + 10\log\left(\frac{Q}{4\pi r^2} + \frac{4}{A}\right)$$

or, from an omni directional source, rewriting using Sabines equation:

$$L - L_o = 10\log\left(\frac{100}{r^2} + 31200\frac{T}{V}\right)$$

Essentially this means that the reverberant/reflected part of the sound field is constant throughout the space. Measurements shows that this not the case. In 1973 Barron presented his ”Revised Theory” on sound levels. This states that also the reflected part of the sound field will decay with distance from the source.

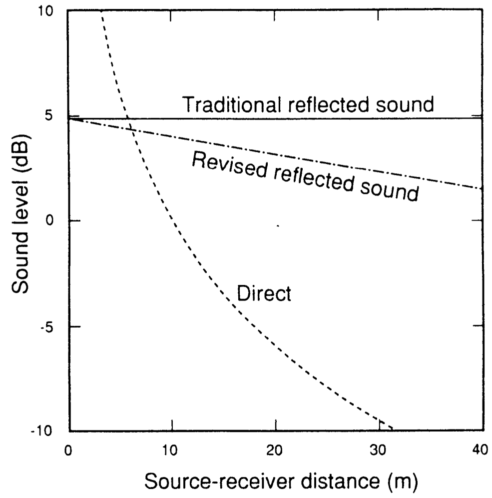


*Figure 2.11: Sound level distribution according to traditional formulaes and according to the Revised Theory (BARRON)*

The revised theory splits the sound field into 3 parts: direct sound d, early reflected sound $e_r$ (upto 80 ms from direct sound) and late sound l (later than 80 ms from direct sound). These parts can be calculated as follows:

$$d = \frac{100}{r^2}$$

$$e_r = (31200 \cdot T / V) \cdot e^{-0,04 \cdot r / T} \cdot \left(1 - e^{-1,11/T}\right)$$

$$l = (31200 T / V) e^{-0,04 r / T} \cdot e^{-1,11/T}$$

From the the total sound level at a give distance can be calculated as:

$$L - L_0 = 10 \cdot \log(d + e_r + l)$$

Also the revised theory can be used to predict early/late ratio of the sound field:

$$C_{80} = 10 \cdot \log[(d + e_r)/l]$$

Research has however shown that even this will in some case give a wrong estimate of the sound level.

## 2.2 Development of room acoustics descriptors

The purpose of all measured acoustic parameters is in general to find single number descriptors of the listener’s impression of the halls. In general the physical or objectively measured parameters will try to be a physical measure of a subjective parameter. But still we are in a situation were the correlation between measured objective room acoustic parameters and subjective parameters are weak.

About 100 years ago the Harvard University professor W.C. Sabine took the first steps to estimate the acoustic conditions of concert-halls by means of objective measures. He described the connection between reverberation time and the amount of absorbing materials in the space. He used this knowledge in the design procedure of the Boston Symphony Hall. The reverberation time is defined as the time it takes sound in a space to decay 60 dB. The reverberation time was measured by first creating a steady sound field, then shutting of the source and then measuring the time it took for the sound to decay below the hearing threshold. The parameter can however also be measured from the impulse response or from the reflectogram by performing a backwards integration first and then calculating the reverberation time from the envelope of the response.

What makes the reverberation time parameter still today so usable is that it can be calculated easily from the room volume and total absorption area of the room.

It was however realized soon after Sabine presented his definition of the reverberation time that the reverberation time alone could not be used as a sole descriptor of room acoustics. From the 1930:ies onwards a great deal of other room acoustic descriptors was presented. Common for all this descriptors was that they were monaural, that is for the most part measured with an omni directional microphone. Examples of these descriptors are clarity and strength (not loudness). One of the more important discoveries was that the ear has a “integration time” that is that sound that arrives within a time frame of up to about 80 ms of the direct sound, will change the aural impression of the direct sound.

It was only in towards the end of the 1960:ies that it was realized that the binaural parameters or stereophonic parameters of a space are by far more important, or at least as important, as the monaural parameters. Most of the monaural parameters dealt with the relationship between early and late energy that is with the strength of the early reflections, as it was understood that the perceived sound quality, or aural impression, was closely related to the characteristics of the early reflections. In 1967 Harold Marshall proposed that not only the strength of the early reflections but also their angel of arrival was important. This presented a major break with traditions that had ruled acoustics design, but at the same time present a possible solution to the question of why the acoustics in Filharmonie Berlin was so much better than in Philharmonic hall in New York. Throughout the 1970:ies this was studied intensively, in particular in Germany, (Cremer and Müller and Plenge), in New Zealand (Marshall) and in Great Britain (Barron). This led to the understanding that the spatial attributes of the sound field, or Spatial Impression (SI) are properly the most important factors of good acoustics in concert-halls. This showed the width; or rather narrowness of the old halls was one of the keys to why their sound was superior to most new halls.

Spatial impression can be described as “width of the source” and “being surrounded by music”. A simplified expression could be that in a hall with good special impression, stereophonic hall, one is “sitting in music” whereas in a hall with low spatial impression, monophonic halls, one is “looking at music”. In 1989 Morimoto et al suggested that spatial impression was a combination of at least two different factors: the Apparent Source Width (ASW) and Listener Envelopment (LEV). Bradley and

Soulodre concluded that the perceived source width depended mainly on the early lateral energy, that is energy arriving from the side within 80 ms of the direct sound, whereas the listener envelopment depends mainly on the late lateral energy. Later work has shown that there is a good correlation between some of the traditional measures of spaciousness and ASW, whereas the dependency of LEV and late energy is still not quite clear.

Spatial Impression is however also dependant of the source level. This can be demonstrated in principle with any sound system; when you increase the level, it feels like the source is broader, to a certain level. This effect was reported by Veneklasen and Hyde [5] where, in an "auditorium synthesis" listening system, the acoustical image was found to focus on the source during quiet passages while it broadened as the passages became louder, creating a feeling of "envelopment" by the sound, and a broadening of the source image.

Most of the commonly used measurement parameters are found or defined in the ISO standard ISO-3385

## 2.3 Objective descriptors

### 2.3.1 Reverberation time

As mentioned above, the first objective room acoustic descriptor used was the Reverberation time, $RT_{60}$ or T. For many applications, this very simple equation gives a good and reliable estimate of the reverberation time in space.

One problem with the Sabine equation is that the equation assumes that all absorption in the space is distributed evenly on all surfaces. This obviously is not the case in concert halls, where the only principle absorbing surfaces are the audience (and the musicians). After Sabine's original equation was presented, there have been a number of attempts to improve the equation, to take uneven distributed absorption areas. One example is the Fitzroi-equation, which divides the absorption into X,Y and Z directions:

$$RT = \frac{X}{S}\left(\frac{0{,}16V}{\alpha_x S + 4mV}\right) + \frac{Y}{S}\left(\frac{0{,}16V}{\alpha_y S + 4mV}\right) + \frac{Z}{S}\left(\frac{0{,}16V}{\alpha_z S + 4mV}\right)$$

where: S is the total surface of the room

X, Y and Z are the dimentions in the respective direction.

The Eyring equation is in some particular cases found to be more precise than the Sabine equation:

$$\mathrm{RT} = \frac{0{,}161V}{-S\ \ln(1-\bar{a})}$$

where: $\bar{\alpha}$ is average absorbtion coefficient of the surfaces

Also Arau-Puchades has presented an improved formular, based on the Fitzroy and Eyring formulars, where the Mean Free Path (l) is taken in account. The Mean Free Path is defined as:

$$l = \frac{4V}{S}$$

Beranek and Hidaka has proposed some correction terms to the Sabine formular to account for uneven distribution, by splitting the surface areas into Audience areas with edge effect correction, orchestra areas, aisles and "other areas"

A slightly different approach was suggested by Hyde and Möller. This is a correction term to the Sabine equation to take into account uneven distribution of the absorption in the space.

$$\beta = S_L / (S_s + S_a)$$

where:

$S_L$ is the "lumbed" areas, that is all areas which are not seating or aisles
$S_s$ is seating areas
$S_a$ is aisle areas.

Essentially this correction can be used to show that the less even distributed the absorption is in a hall, that is the larger β is, the more volume per listener is required to achieve reverberation.

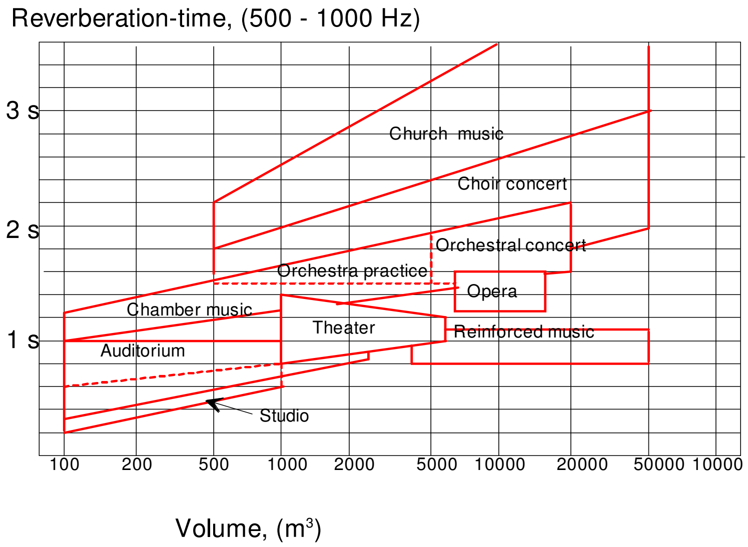


*Figure 2.12 :* *Traditional reverberation-time recommendations for different room types.*

It was soon realized that the reverberation time $RT_{60}$ did not correlate very well with the perceived sense of reverberation. The Early Decay Time, EDT, was therefore introduced as a better measure for perceived reverberation. The EDT is defined as the decay time from 0 dB to –10 dB multiplied by a factor 6. The idea is that in music one seldom hears the full decay of the hall, whereas the initial decay structure is heard essential between "all notes".

The reverberation time related parameters are defined from the decay of a steady state sound field and not from the impulse response. It is however possible to calculate them from the impulse response, using for example backwards integration or other ways to smoothen out the decay, and for the most part, this is how the measurements are done nowadays.

### 2.3.2 Early/Late energy ratios

These are parameters that describe the relationship between early energy, that is energy that arrives at the listener within a short time from the direct sound, and energy that arrives later. Essentially the parameters are either percentage early to total energy ratios or logarithmic ratios between early energy and late energy. These parameters are all defined from the impulse response and cannot directly be calculated from the steady state decay.

For concert hall acoustics, Clarity, $C_{80}$, is the most usual. This is defined as 10 times the logarithm of the ratio of the energy arriving within 80 ms of the direct sound and the energy arriving later:

$$C_{80} = 10LOG\left(\frac{E(0,80ms)}{E(80,\infty ms)}\right)dB$$

or

$$C_{80} = 10Log\left(\frac{\int_{0ms}^{80ms} p^2 dt}{\int_{80ms}^{\infty} p^2 dt}\right)dB$$

The parameter $C_{50}$ is essentially the same, only the time limit is 50 ms instead of 80 ms. This parameter has been found to correlated with speech intelligibility (REF Hem).

Another often-used parameter is the Deutlichkeit or Definition, defined as the ratio between the energy arriving within 50 ms of the direct sound and the total energy:

$$D_{50} = \left(\frac{E(0,50ms)}{E(0,\infty ms)}\right)$$

In some text when Clarity is given to both $C_{50}$ as well as $C_{80}$ and Definition is used as the name for both $D_{50}$ as well as $D_{80}$. The C-index parameters are related to the D-index as follows:

$$C_{te} = 10LOG\left(\frac{D_{te}}{1-D_{te}}\right)dB$$

For concert hall acoustics also the Center Time, $T_s$, is sometime used. This defined as the sum of energy weighted by its arrival time, divided by the sum of the total energy.

The parameter will give a time within which half of the energy has arrived. As can be expected, this parameter is very highly correlated with the reverberation time.

### 2.3.3 Strength

Strength is a parameter that describes the relative strength of a sound field in a room. This is not directly proportional to the reverberation time, as might be expected. Just as is the case for the above-described early/late ratios, also the strength is highly influenced by the shape of the room.

The parameter is defined as the ratio of the energy measured in the room, compared to the energy measured at 10 m distance from the same source in a free field:

$$G = \frac{\int_0^\infty p^2(t)dt}{\int_0^\infty p_{10m}^2(t)dt} => \frac{E(0,\infty)}{E_{1m}(0)} + 20 => L_{pl} - L_{pl:1m} + 20$$

For practical purposes the parameter is usually measured as the ratio between the energy measured in the room compared to the direct sound level measured at 1 m distance from the source, with 20 dB added to compensate level change from 10 m to 1 m. The 1 m measurement is furthermore compensated for the influence of floor reflections etc.

In newer literature, the strength parameter has also been divided into early and late strength, essentially making it a "calibrated" Cte measurement. These can also be calculated from the G and $C_{80}$ values as:

$$G_e = G_{80} = G - 10Log\left(1 + 10^{(-c_{80}/10)}\right)$$

and

$$G_l = G - 10Log\left(1 + 10^{(+c_{80}/10)}\right)$$

Also lateral strength measurements are being used, these will be described in the following.

### 2.3.4 Spatial impression

As stated earlier, it was realized around 1970 that the spatial attributes of a sound field was likely to have a great influence on the perceived sound impression. Very soon it was proposed to measurement of the spatial attributes of the sound field with a combination of omni-directional microphones and figure-of-eight microphones. This is essentially the same technique as is used today [Reichardt 76, Bradley].

Most of the proposed parameters were essentially variations over the same theme; the relation of lateral energy arriving within 80 ms of the direct sound and omni-directional energy. The main differences between the parameters are the time division of the energy. The parameters were for instance the Raumindrücksmass as proposed by Reichardt & Lehman [Reichardt 1976] and the Lateral Efficiency proposed by Jordan [Jordan 1980]. The most common parameter today is the Lateral Fraction, or LFC :

$$L_{fc} = \frac{\int_{0ms}^{80ms} h^2(t)\cos(\theta)dt}{\int_{0ms}^{80ms} h^2(t)dt}$$

For practical purposes this is reformulated to the following, when using a figure-of-eight microphone:

$$L_f = \frac{E_{fig8}(5{,}80ms)}{E_{omni}(0{,}80ms)}$$

Lately, with the increased availability of binaural recording or microphone systems, new parameters to estimate spatial attributes has been suggested. The idea of binaural measurements, that is measurements with microphones in the ear channels was first proposed by Ando in 1985 [ANDO1]. The *interaural cross-correlation function* IACF is a binaural measure of the difference in the sounds at the two ears, produced by a sound source on the stage. It is defined as [Hidaka et al. 1995]

$$\Phi_{l,r}(\tau) = \frac{\int_{t_1}^{t_2} p_1(t)p_r r(t+\tau)dt}{\sqrt{\int_{t_1}^{t_2} p_l^2 dt \int_{t_1}^{t_2} p_r^2 dt}}$$

where the subscripts *l* and *r* designate the sound pressures measured at the left and right ears.

The maximum possible value of the IACF is unity, resulting from identical signals arriving synchronously at both ears. The variable t denotes the time difference between the ears. The values for IACF are usually calculated in the range of (–1 ms < t < 1 ms) , which corresponds to the maximum natural sound delay between the ears. In order to obtain a single number measuring the maximum similarity of sound arriving at the ears within the time integration limits and the range of t , the parameter interaural cross-correlation coefficient IACC is defined as

$$IACC = \max(|\Phi_{l,r}(\tau)|)$$

$$, -1\ ms < \tau < 1\ ms$$

There are several standard IACC parameters with different integration periods:

| | | |
|---|---|---|
| $IACC_A$ or $IACC_{Total}$ | $t_1 = 0$ ms | $t_2 = 1000$ ms |
| $IACC_{Early}$ | $t_1 = 0$ ms | $t_2 = 80$ ms |
| $IACC_{Late}$ | $t_1 = 80$ ms | $t_2 = 1000$ ms |

The $IACC_{Early}$ parameter, together with the LF parameter, is also considered to be a measure of the *apparent source width* ASW, and the $IACC_{Late}$ parameter is a measure of the *listener envelopment* LEV. The time scale is fixed to the arrival of the direct sound at $t$ = 0 ms.

Within the last few years, it has been realized that not only the early part of the impulser response is important for spatial attributes, but also the later part. On of the parameters proposed as a measure for the Listener Envelopment, LEV, is the Late Lateral Strength, $G_{L,Late}$ also denoted GLL. As Barron and Bradley has shown, the correlation between $G_{L,Late}$ and LEV is however not very high. $G_{L,Late}$ is defined by:

$$G_{l,late} = 10\log\left(\frac{\int_{0,08}^{\infty} P_F^2(t)dt}{\int_0^{\infty} P_A^2(t)dt}\right), dB$$

At around the same time, Keet [6] published a paper showing a clear and definitive relationship between sound level, ASW (in degrees), and what he called the “incoherent lateral energy fraction,” being the crosscorrelation function of a binaural signal from a dummy head measurement. Barron has shown [3] a direct linear relationship between Keet’s crosscorrelation factor and his units of spatial impression thereby yielding a direct relationship between LF and the change in sound level at ΔLF=1,6%/dB (or ΔLevel/62,5). This translates to a change in LF of approximately 5% for each 3 dB change in level. Barron has since suggested [4] combining LF with level in defining a “degree of source broadening” (DSB) = LF+G(early)/*k*. He and Marshall [6] have also shown that work by Morimoto and Iida [7] yields a value of *k*=57. Combining the above, the expression DSB=LF+G(early)/60 is suggested. With all due respect to Barron, for the purposes of this paper we are using the descriptive term “effective spatial impression” or ESI = LF+$G_e$/60. The subjective response to the image broadening aspect of SI is therefore proposed as being a factor of both LF and $G_e$ more or less to this degree.

It has been shown, initially through observation and confirmed through listening simulations that Spatial Impression also depends on the source level. In other words, the louder the sound is, the more spatially broaded it is. It can be shown (Morimoto, Möller&Hyde) that the relationship between perceived Spatial Impression, lateral energy and sound level is:

ESI = LF + G(early)/60

Where: ESI is Effective Spatial Impression

LF is the measured Lateral Fraction

G(early) is the early Strength (0 – 80 ms)

In other words, basically a 3 dB increase of sound strength, is equivalent to a 5 % change of LF. Obviously this works within certain limits, at some point the sound strength will put the hall into “saturation” or “overload”.

Also resent research presented in [SOULODRE et all 1003] and refined in [BERANEK 2008] and shows that the listener envelopment in a hall can be estimated from the late strength and the late lateral fraction:

$$LEV_{calc} = 0{,}5G_{late,mid} + 10Log\left(LF_{late,mid}\right)dB$$

Instead of using the $LF_{late}$ parameter, Beranek suggest to use the $IACC_{late}$ parameter, thus giving the formular:

$$LEV_{calc} = 0{,}5G_{late,mid} + 10Log\left(1 - IACC_{late,mid}\right)dB$$

As stated ealier, the $G_{late}$ parameter can be calculated from the overall strength, G, and the clarity, $C_{80}$.

### 2.3.5 Stage parameters

Stage acoustic was not a subject of major interest until the early 1980:ies. Then the acoustic conditions for musicians on stage was thoroughly investigated by AC Gade [GADE]. One of the results of these investigations were a set of measurement parameters which were found to correlated with the musicians subjective evaluation of the acoustic conditions on stage. The parameters can be divided into three groups: acoustic support for the own instrument, ease of ensemble and acoustic connection to the audience chamber.

**Early Decay Time on Podium**

The Early Decay measured on the podium has been found to correspond well with the musicians sense of reverberation. The value is in general about 30% lower than the value measured in the audience chamber and is denoted EDTP. In some references also the frequency balance of the EDTP is defined as:

EDTF= (EDTP(250 Hz)+EDTP (500 Hz))/(EDTP(1000 Hz)+EDTP(2000 Hz))

This is believed to correlate with the musicians impression of timbre, that is the frequency response on stage.

**Clarity on stage**

Clarity on stage, CS, has been found to correlate with the musicians feeling of reverberance, like the EDTP. Like the support parameters, also the CS is measured with a omni-directional microphone at a distance of 1 m from the sound source. It is defined as:

$$CS = 10Log\left(\frac{E(0{,}80ms)}{E(80ms,\infty)}\right)$$

**Support parameters**

The support parameters, ST1 and ST2 (or in most cases $ST_{Early}$ and $ST_{Total}$) are measured with a omni directional microphone placed 1 m from the source. ST1 or $ST_{Early}$ have been found to correlate with the musicians ability to hear each, that is a measure of ensemble. The parameter is defined as:

$$ST_{early} = 10Log\left(\frac{E(20{,}100ms)}{E(0{,}10ms)}\right)dB$$

$ST_2$ is defined as $ST_1$ but with the integration time of 20 ms to 200 ms. In most cases however $ST_{Total}$ is used instead of $ST_2$. These parameters has been found to correlate with the musicans feeling of support for their own instrument. $ST_{Total}$ has an integration time of 20 ms to 1000 ms:

$$ST_{total} = 10Log\left(\frac{E(20{,}1000ms)}{E(0{,}10ms)}\right)dB$$

## 2.4 Subjective descriptor or acoustic vocabulary

One of the main tasks of when designing a concert space is of course fulfilling the needs of the musicians. To understand each other, one must develop a “common language” or an acoustic vocabulary. In the following typical subjective descriptors of acoustic conditions are described.

**Spaciousness and envelopment**

In a good hall the listener has the feeling of “sitting in music” as opposed to “looking at music”. This is properly one of the most important criteria’s for good room acoustics. As stated above, this is most likely a combination of at least two different components, ASW and LEV. But where ASW correlated nicely with the physical measure Lateral Fraction LF, there is no good objective measure for LEV.

It is also important to remember that ASW or spatial impression is highly dependable on the sound strength, ie a louder sound will be felt as more spatial.

**Warmth**

A sound with sufficient bass energy is heard as a warm sound. The acoustics should therefore provide sufficient bass energy, but too much gives the hall a “boomie” sound. This is closely related to the reverberation time and strength at bass frequencies.

**Reverberation or Liveliness**

Halls should be lively, that is the halls should have a clearly heard reverberation tail. The sense of liveliness is closely related the reverberation at mid and high frequencies so that the listener perceives it.

This is also described as the fullness of tone; long reverberation time will give a full tone whereas a short reverberation time gives a thin tone.

**Intimacy**

A hall should be intimate. Even in a hall seating 2000 people, the ideal is that the music sounds as played in a smaller hall. This is closely related to spaciousness. Also according to Beranek this is related to the Initial Time Delay gab, that is the time between the direct sound and the arrival of the first reflections.

**Acoustic glare**

An evenly distributed sound field is normally the ideal. This also implies a sound field with a degree of diffusion. In a hall with large flat and smooth panels and placed to give early reflections, the sound produced by them will usually be hard or harsh. Also later arriving clearly discrete reflection will influence the perceived sound field in a negative way. Physically this can be estimated in particular from the density of the early parts of the impulse response.

**Balance and blend**

This describes the balance between the orchestra sound and the sound of the individual instruments. In some “bad” halls, strong single reflections may cause “phantom images” of some parts of the orchestra or cause the orchestra to be perceived as “different” orchestras, not as a connected sound source.

**Ensemble and support**

This describes how well the individual members of the orchestra can hear each other and haw the hall supports the sound of their own instrument. The subjective parameter ensemble correlates with the ST1 parameter and the support with the ST2 parameter.

**Brilliance**

A certain amount of brilliance, which is clearly perceived reverberation at high frequencies, is good. But not too much, as will often be the case with sound reinforcement systems.

**Loudness**

The acoustic conditions should provide sufficient “acoustic gain” so that the music can be heard at appropriate strength in all seats. In smaller halls to problem is usually the other way; design care should be taken that the sound of an orchestra will not be too high.

**Clarity**

The sound should be clear enough that all parts of the music are heard, even in fast passages. But a hall with too high clarity will blend the sound of the orchestra into a whole but present each voice individually.

**Background noise level**

The background noise level of the hall should be so low that it does not interfere with the music. Or in other words, even Pianissimo should be heard without disturbing background noise.

Table 2.2 presents an often-used “acoustic vocabulary”, that is subjective criteria of acoustic performance and their descriptors.

*Table 2.2: Acoustic vocabulary; criteria and descriptor scale*

| Criteria | Descriptor | |
|---|---|---|
| Clarity | Muddy | Clear |
| Reverberance | Dead | Live |
| Envelopment | Expansive | Constricted |
| Intimacy | Remote | Intimate |
| Loudness | Loud | Quiet |
| Balance: treble re mid frequencies | Weak | Loud. |
| Balance: bas re mid frequencies | Weak | Loud. |
| Balance: Singers/soloists re orchestra | Weak | Loud. |

In the following subjective parameters for the listeners as well as for the musicians are presented as well as the objective parameters that correlate with them

*Table 2.3: Connection between objective and subjective parameters.*

| | |
|---|---|
| Reverberance | Early Decay Time |
| Liveness | Early Decay Time, Reverberation Time |
| Fullness of tone | Reverberation time |
| Spaciousness, | Early Lateral Energy Fractions, InterAural-Cross-Correlation (early) |
| Apparent source width | Early Lateral Energy Fractions, InterAural-Cross-Correlation (early) |
| Envelopment | Late Lateral Energy Fractions, InterAural-Cross-Correlation (late) |
| Intimacy | Early Lateral Energy Fractions, InterAural-Cross-Correlation (late) |
| Clarity | Clarity $C_{80}$ |
| Blend | Details of the initial part of the impulse response |
| Warmth | Strength at bass frequencies, Bass Ratio |
| Brilliance | Strength at high frequencies, Treble Ratio |
| Timbre | Frequency dependency of parameters |
| Stage Support | Support $ST_1$ |
| Hall response | Late Support $ST_{late}$ |
| Ensemble | Clarity and Early Decay Time on stage |

## 2.5 Specific characteristics in smaller halls

Not too much has been publicized about acoustic design for smaller halls and what has been publicized has first of all been directed energy designs and also in general just scaled down requirements from large rooms. Designing small rooms does how ever give both some more degrees of freedom compared to large rooms but also imposes some extra difficulties.

When one looks a concert halls around the world, it would appear that halls with less than 1000 seat are “recital halls”, implicating that only 6 of the halls in Finland and only 3 of halls in this study would qualify as concert halls. But if a concert hall is defined as a hall where there are regular symphonic orchestral concerts, the definition for concert hall implies a much lower seat count. In the study presented here the lower limit for seat are 350 and furthermore the stage should be large enough to accommodate a symphony orchestra. It is however clear that in an acoustic sense most of these halls are small halls. In the analysis the halls are divided into small halls, medium halls and large halls. From acoustic design point, both the small and the medium sized halls are small.

On of the main problems on encounter in small halls is the due to the limited volume the sound level produced by an orchestra will quite high. In my opinion there are two main mistakes; designing the stage to project sound to the audience and designing small halls using the volume per listener criteria's from large halls. First of all there is absolutely no reason why the sound of a symphony orchestra, or even a smaller ensemble, needs to be enforced in a 500-seat hall. Especially in small halls the stage should be designed only for the support of the musicians. Secondly, the traditional design criterion for concert halls is a volume of 10-12 $m^3$ per seat. But when this same rule of thump is applied to small halls, it can give situations where a 500-seat hall with a stage for a 90-person symphony orchestra and 100-person choir, has a total volume of 5000 $m^3$ implying that it is going to be loud. As such the seat count maybe realistic in a smaller city, but music is never going to be at the right volume. Therefore a criterion of 100 $m^3$ volume per musician makes a lot more sense for small halls. This means larger volumes and hence the absorption characteristics of the surfaces must be designed to match the reverberation time, but this is a far smaller compromise than the excessive sound level.

Small halls or narrow halls are not necessary halls with high envelopment. It is clear that even a small fan shaped hall will be monophonic. What perhaps is more surprising is that even rather narrow halls are no guaranty for a high level of envelopment and stereophonic sound. What accomplishes the strong early lateral sound fields is a combination of narrow size and balconies or balcony simulating surfaces, directing the sound field down towards the audience. So in other words, even narrow halls with smooth sidewalls, or rather sidewalls without horizontal details will have quite low ASW.

## 2.6 Acoustics for speech

In multi purpose halls in general it is important that it is possible to achieve good speech acoustics in the "theater" configuration. Normally good speech acoustics will also imply good conditions for electronically reinforced sound. There I will make a short presentation of speech acoustics.

Speech is a succession of utterances that produces a wave whose frequencies and amplitude change rapidly with time, [8]. The utterances can be divided into vowel and consonant sounds, where the vowels are created by the vocal cord.  Air movement through the mouth and over the tongue and lips creates most consonant sounds.  The frequency spectrum of all speech sounds is shaped by the resonance cavities in the mouth and throat.  From this it follows that the speech power of the vowels lies at low frequencies, while the consonant sounds lies mostly at higher frequencies, [8].

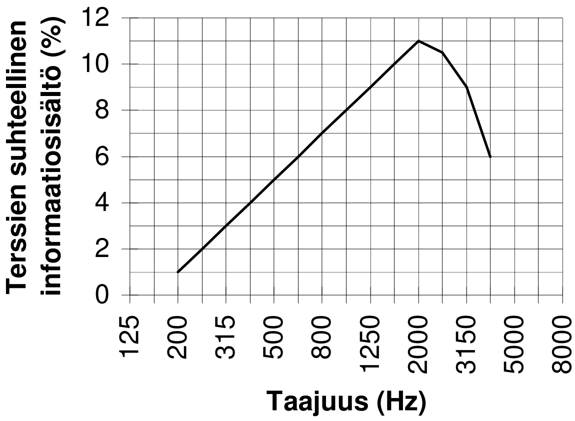


*Figure 2.13* *Content of information in different frequenciesbands*

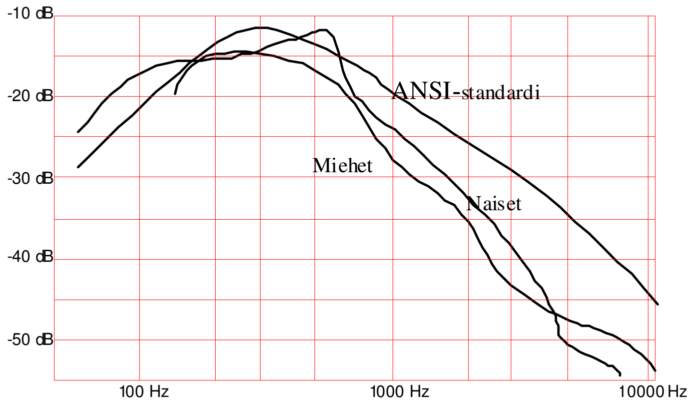


*Figure 2.14:* *Standard frequency spectrum of typical Male and Female speech as well the standard ANSI-spectrumi*

The factors considered to influence speech intelligibility are background noise, reverberation time and the reflection pattern of the room, which is caused by the room's geometrical shape and the position of the absorbers in the room, [44]. In order to achieve good speech intelligibility, all of these factors must be optimized.

From a planning point of view speech intelligibility problems can be divided into two categories; indoors, where it is normally possible to make acoustic treatment and outdoors, where it is normally not possible to change the acoustic environment but where sound reinforcement is used to improve or optimize speech intelligibility.

### 2.6.1 Background noise

The most obvious cause of low speech intelligibility is a poor signal to noise ratio, i.e. excessive background noise. Even though it has only been proven for severe cases of noise exposure [10], it is logical that this has a direct influence on the speech intelligibility even at lower background noise levels.

The normal speech level of a lecturer is about 60 dB measured at I m distance, even though some literature suggests that teachers will normally speak at a raised voice of about 70 dB. It is clear that in order to obtain a reasonable signal to noise ratio even in a small lecture room, the background noise level must be low.

### 2.6.2 Reverberation time

The reverberation time of a room intended -for speech communication is a compromise between speech power and clearness. A very short reverberation time will demand a lot of speech power to achieve a reasonable speech level, while an excessive reverberation time will spoil the clearness. It is also important that the reverberation time does not vary greatly with the frequency.

Often good speech intelligibility is in conflict with good musical environment, for instance the suggested reverberation time for a 500 person concert hall is around 2 s while a reverberation time less than 1.6 s is recommended for a 500 person lecture hall, [63].

### 2.6.3 Room geometry

In designing for good speech intelligibility, the shape and the layout of the room are as important as the background noise level and the reverberation time. These factors are more a question of architectural planning than a question of acoustical treatment, for instance placing the speaker over 'the crowd'.

As can be seen from figure 2.15, the source is localized from the first arrived sound, even though a sound arriving later has higher intensity, within certain limits. The maximum intensity that later arriving sounds can have without causing disturbance is about + 5 dB at a delay of 18 ms compared to the direct sound. This is called the Haas effect. This is related to the integration time of the hearing system. Surveys has indicated that energy arriving within the first 80 ms is treated as direct sound, or information, by the ear.

This shows that for rooms intended for speech communication it is important to have strong early reflections while late reflections should be damped. In practice this is done by placing reflecting areas around the lecturer, and by damping the back wall of the room. Furthermore it is common practice to have reflecting roof right over the teacher and down the middle of the room and absorbers on the rest of the ceiling.

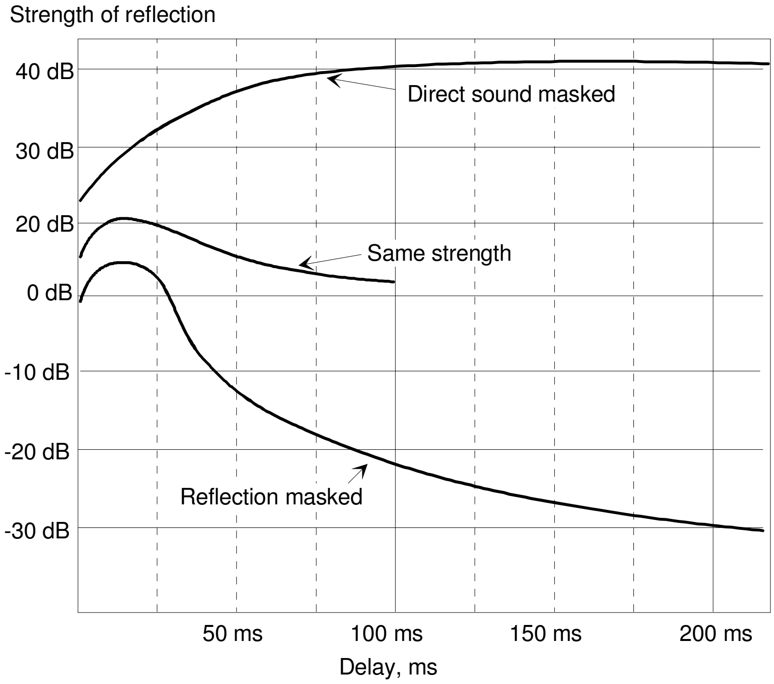


*Figure 2.15:* *Subjective perception of a side reflection, $90^0$ compared to the direct sound, as a function of relative strength and delay. Above the upper curve the direct sound is masked and below the lowest curve the reflection is masked. The middle curve corresponds to equal intensity of the reflection and the direct sound. [631*

### 2.6.3 Measurement of speech inteligibility

The most simple measurement parameter of speech inteligibility is Percentage of Articulation Loss of Consonants ($\%AL_{cons}$) The parameter basically descripes how many percent of the consonants are un-intelligible .The parameter can be calculated in accordance with :

$$\%AL_{cons} = \left( \frac{200 d_2^2 T_{60}^2}{VQ} \right)$$

where: $d_2$ is the distance between the speaker and the listener

$T_{60}$ is reverberation time (normally either broadband or mid-frequencies.

V is the volume of the space

Q is the directivity of the source

$\%AL_{cons}$ does not in general give a very precise estimate of speech inteligibility as it is based only on statistical parameters, and does not take for instance relflection patterns into account. The advantage of the parameter is that it provides a method of making an estimate at a very early stage of the design process.

A more precise parameter is the Speech Transmission Indes (STI). Tarkempia parametrejä ovat Speech Transmission Index, STI. This parameter is based ona an analysis of the Modulation Transfer Function, (MTF). The STI expresses the ability of a communication channel to carry across the information contained in speech. It is an indirect measure of speech intelligibility: it predicts how channels (systems, rooms,...) affect the intelligibility of speech. STI was developed in the beginning of the 1970's, it is a machine measure of intelligibility whose value varies from 0 = completely unintelligible to 1 = perfect intelligibility. This scale is actually based on "percentage correctly understood nonsense syllable" , a measured and method used for actual testing of phone lines. On this scale, an STI of at least .5 is desirable for most applications.

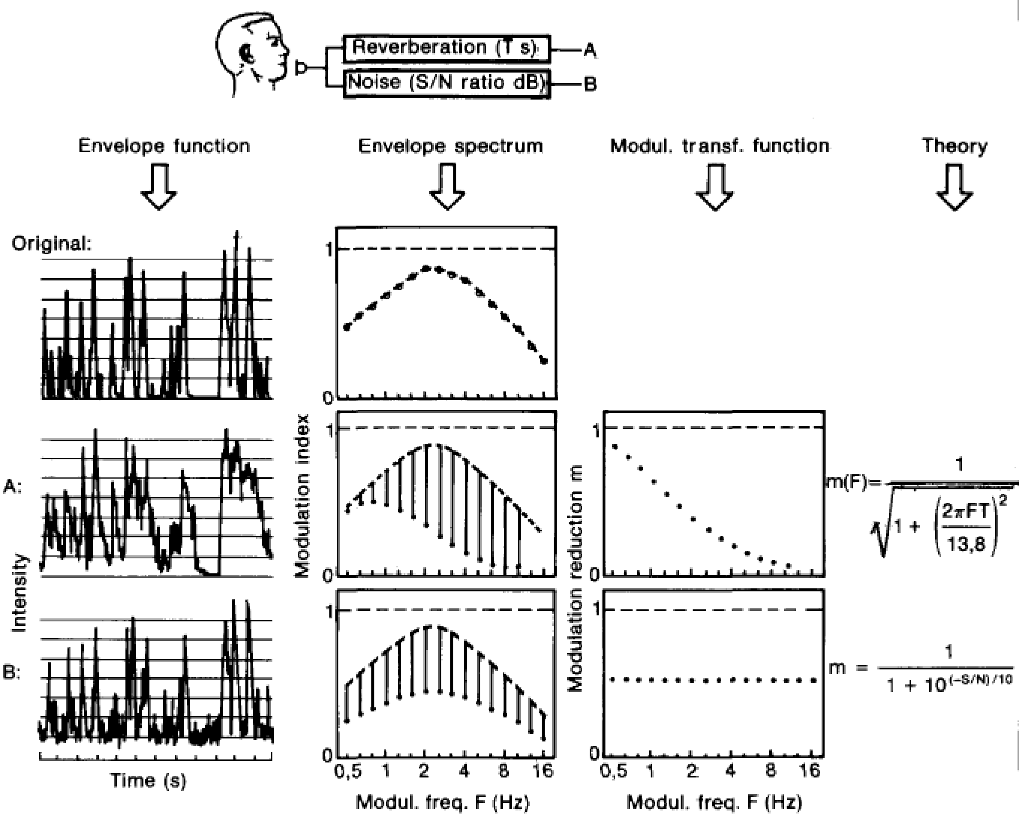


*Figure 2.16: Analysis basics to determine the STI*

Speech can be described as a fundamental waveform that is modulated by low-frequency signals. Therefore STI employs a complex amplitude modulation scheme to generate its test signal. At the receiving end of the communication system, the depth of modulation of the received signal is compared with that of the test signal in each of a number of frequency bands. Reductions in the modulation depth are associated with loss of intelligibility.

There are also simplified versions of STI developed for use in specific situations. RASTI ´(Room Acoustics Speech Transmission Index, or Rapid Speech Transmission Index) and STIPA (Speech Transmission Index for Public Address Systems). Both ot these measures are based on known signals, modulated noise signals, as input to the speech transmission system, and are designed as easy test methods for field use. The RASTI method was mainly used in the 1980:ies and 1990:ies and is not much used today. STIPA on the other hand is one of the criterias stated in the EN60849 standard for evaluation of sound systems for emergency purposes.

*Table 2.4: Comparison of STI/RASTI/STIPA and %ALcons values and their subjective equivalent.*

| | **unintelligible** | **poor** | **fair** | **good** | **excellent** |
|---|---|---|---|---|---|
| STI | 0 - 0.3 | 0.3 - 0.45 | 0.45 - 0.6 | 0.60 - 0.75 | 0.75 - 1.0 |
| Alcons | 100 - 33% | 33 - 15% | 15 - 7% | 7 - 3% | 3 - 0% |

All MTF based parameters can be estimated from the impulse response function with acceptable precision.