Proceedings of The Institute of Acoustics

RECENT DEVELOPMENTS IN SOUND SYSTEMS DESIGN, TECHNOLOGY AND INTELLIGIBILITY ASSESSMENT

Peter Mapp

Consultant

INTRODUCTION

The past few years has seen some major developmens in the field of electroacoustic technology and sound systems design. In the UK a perceptible trend has been developing both towards the use of more sophisticated sound systems and towards the use of constant directivity loudspeaker configurations in both entertainment, conference and industrial environments.

Bessel function and line source loudspeaker arrays have also been developed better to control the radiation and dispersion characteristics of multiple drive loudspeaker systems giving rise to more complete coverage and hence potential intelligibility and feedback margin.

Digital electronics and audio technology are also now beginning to make a substantial impact on sound system design and implementation, ranging from signal delay lines with full bandwidth delay capabilities of several seconds with resolutions of twenty microseconds or less to single cable, 60 channel audio communication and control circuits.

The aim of this paper is to review and discuss a few of these developments concentrating on the acoustic rather than electronic aspects.

It is hoped that some of the points raised will provoke further discussion in the intelligibility workshop.

LOUDSPEAKERS

Traditionally in the UK and in most of Europe, the column loudspeaker has generally been used either to provide directional sound coverage or speech and light music coverage of large or difficult acoustic areas. The techniques date back to the early 1950s when Parkin developed an 11 foot column loudspeaker for St. Paul's Cathedral.

An advantage of the column loudspeaker is that it is generally relatively simple to construct and manufacture. This ironically can also be a distinct disadvantage as a multitude of brands are now on the market with only a very few having been actually acoustically designed and tested. The result is that most column loudspeakers commonly available exhibit very poor frequency responses and pattern control. Furthermore, many of the manufacturers cannot provide even the most basic of acoustic data describing the parameters needed for acoustic design. There are of course exceptions to this.

A fundamental characteristic of the column or line source loudspeaker is that at the higher frequencies, where the wavelengths become comparable to or less than the length of the column, or equivalent to the spacing between drive units, the vertical dispersion (coverage) angle decreased and constructive and destructive interference occurs. The result is a narrowing of the coverage angle, the formation of radiation lobes and a degradation of the frequency response.

The narrowing of the vertical dispersion angle is of course a desirable feature, being the primary reason for using a column format. However, as the effect progressively increases with frequency, the overall result is a highly non-uniform coverage pattern in the vertical plane - resulting in poor or uneven sound distribution within the potential coverage area.

A number of techniques have been developed to overcome this effect including power and/or frequency tapering, the use of shorter sub-columns within te main enclosure fed via a crossover network, and the use of additional high frequency devices such as tweeter units or small horns.

Whilst a correctly designed column technique can produce a well preserved and controlled high frequency response, this is seldom encountered within current designs. The use of a separate high frequency device is however more frequent, and whilst this can substantially improve the 'on axis' frequency response, the dispersion characteristics of the resultant column become far from constant.

A number of specialist manufacturers do exist and can provide both a reasonably well controlled response and dispersion. Frequently this is implemented by either a form of 'articulated array' design or by a form of Bessel function tapering of the column - a useful feature of the latter technique is that not only can a very much more linear and controlled sound pattern be produced, but the radiated beam can be electronically 'steered' or bent, potentially providing a very flexible approach to sound systems design.

Another form of column loudspeaker is the line array - whereby one complete dimension of the room is used to form a true 'line source'. If correctly implemented, such a technique can produce a very even coverage of the space with good intelligibility in resonant areas.

A number of such systems have now been designed and commissioned by the author and provide a uniformity of coverage within 2 dB at 2KHz.

Although exponential, multicellular and radial horns and drive units have been manufactured in this country for over 50 years, they have only found fairly limited application within the UK - particularly in high quality sound reinforcement applications, though they are more extensively used in the USA where the column loudspeaker for example is more rarely used.

The multicellular orn also exhibits a non uniform dispersion characteristic - but is probably better able to provide controlled high frequency coverage than the simple column. Multicells are also highly amenable to stacking and splaying and can therefore be used as a basic building block to create a system providing a closer match to the desired coverage requirements.

RECENT DEVELOPMENTS IN SOUND SYSTEMS DESIGN

The invention of the constant directivity horn by Electro Voice in the 1970s overcame many of the coverage problems, and for the first time enabled sound system designers to aim and predict precisely the likely performance of loudspeaker systems. The constant directivity horn concept has been adopted and adapted by a number of manufacturers so that there is now quite a range of formats to select from. Recent designs provide totally controlled dispersion from around 500 or 600 Hz up to 16 KHz or beyond within a single horn/compression driver combination. Typical coverage formats are 90° x 40°, 60° x 40° and 40° x 2° (H x V).

Low frequency constant directivity horns following the above format have also been developed - and typically provide effective control from 250 Hz upwards.

The technique of using constant directivity, high Q, single source (centre cluster) loudspeaker systems has been well researched and has been widely used throughout the USA for many years and is now slowly becoming accepted practice within the UK and Europe.

The ability of such a system to produce a very high ratio of direct to reverberant sound enables high levels of speech intelligibility and clarity to be achieved in highly reverberant environments. The freedom from lobing, combined with the well controlled coverage patterns ensures minimum excitation of the reverberant field and of discrete echoes.

Perhaps the most recent and surprising development within the field of loudspeaker design and application, is the development and use of omni-directional loudspeakers for sound reinforcement. The substantially wider dispersion available enables very much greater areas to be covered from one unit, reducing substantially, for example, the number of loudspeakers required within a system. Whereas their advantage within acoustically well behaved environments is obvious, their use within more difficult areas, where traditionally a high Q directional device would be used is not! It is anticipated that some interesting developments will occur within this area which may well result in a re-examination of traditional methods of sound system prediction and design.

SPEECH INTELLIGIBILITY - PREDICTION AND MEASUREMENT

Although much research has been carried out in the fields of speech intelligibility and the effect of room acoustics on intelligibility 0 it is only more recently that the intelligibility of sound systems in their own right has been investigated.

The percentage loss of consonants (% ALCONS) developed by Peutz and subsequently modified by Klein is probably the most widely used prediction method. It relates in a simple way the basic acoustical factors of Reverberation Time (T), Distance (D() and room Volume (V) to intelligibility and to the Q of the loudspeaker.

$$\% \text{ ALCONS} = \frac{200D^2 \; T^2 \; (n + 1)}{Q \; V \; m}$$

The factor m is introduced as a critical distance modifier to take into account the common discrepancy between the average absorption coefficient of the room and the higher absorption of the audience at which the loudspeaker system is directed. The (n + 1) factor is introduced to take account of sound sources (loudspeakers) other than those directly contributing to the intelligible signal heard by the listener at a distance D from the loudspeaker.

Subjective correlation with the percent Alcons value produces a scale on which 15% is generally regarded as being the maximum permissible loss for intelligibility of fairly simple speech, a loss of 10% or less being required for more complex messages or difficult text. (Recent research shows that this is also the probable minimum requirement/expectation for a mid 1980s sound system.) A loss of 5% should result in very good or excellent intelligibility.

When calculating potential intelligibility, few people seem to take into account or stop to consider the likely accuracy of their calculations. (The author was recently presented with a system calculation producing a predicted intelligibility of 14.8% Alcons and was asked why, as this was within the 15% limit, was he worrying?)

An analysis of the likely errors involved in providing the predictive base data for the above formula results in a conservative estimate of between 30 and 40% likely accuracy. So that if a value of 15% is predicted, this really means that the potential intelligibility is likely to lie within the range 10-20% Alcons, subjectively corresponding to a range of good to poor/unacceptable.

The percent Alcons formula assumes a linear relationship between the Q of a loudspeaker and the resulting intelligibility, and similarly the number of loudspeakers not directly contributing to the early sound field. Whereas there is little doubt that there is a strong relationship between the parameters the linearity of the relationship is currently being questioned.

Where measurements can be directly made on a system in its environment a btter idea of the likely intelligibility can be obtained. By plotting intelligibility contours, potential trouble spots can be immediately identified.

Two objective measurements essentially are available apart from the more traditional reverberation time and background noise level checks. (For sound system the EST via the system is a better predictor of intelligibility than normal RT measurements.) The methods are STI or RASTI - based on the modulation transfer function and impulse response testing to provide a measure of direct to reverberant ratio and possibly direct to noise ratio.

The STI or RASTI method, which is now available as a commercial measuring system, has begun to take intelligibility testing very much more out into the field and into the hands of everyday practitioners and so should soon begin to provide much new data and provide direct feedback into the initial calculation and design procedure. Care must however be taken when interpreting the results of a RASI measurement as being a rapid measurement only two of the five octave bands of the full STI test are employued. Furthermore, the generation of a single number index - to two places of deciamls - may perhaps mislead the inexperienced user or uninformed client. Due to the statistical nature of the modulation transfer measurement and the acoustic environment in which the measurement is performed, a fairly wide range of possible values exist, rather than just the one number shown on the numerical display. Typically the potencial accuracy of the measurement is within 0.1 to 0.15 on the RASTI scale i.e. equivalent to one subjective scale graduation. Again, as with te percent Alcons calculations, perhaps a likely intelligibility range should be produced rather than a single value result. The system does however allow rapid comparisons to be made or intelligibility contours to be produced. It should be remembered that the intelligibility and not the subjective quality of the system is being measured.

A useful feature of the RASTI system is the calculation of the equivalent EDT and diagnostic readouts to help check whether poor intelligibility is a result of excess noise of reverberation.

The ability either to set the noise floor to the real condition or to measure a system without the normal ambient noise and then enter this manually into the unit for recalculation is extremely useful.

In practice it has been found that a doubling of the EDT decreases the STI by a factor of 0.15 or 1 complete subjective step. Similarly increasing the background noise by 3 dB reduces the STI by a factor of 0.1.

The STI or RASTI value can be predicted from the basic acoustic data - though this at present is only possible for the human voice or for a single low Q loudspeaker.

The measurement method which, according to Bradley's recent work, shows the highest correlation with subjective intelligibility testing is the Direct + Early to Late + Reverberant energy ratio.

There is still, however, considerable debate as to what the exact cut off time for the useful early energy should be, e.g. 35, 50, 80 or 95 mS. There does however appear to be good general agreement that the effect of background noise to speech signal ratio flattens out at around 15 dB, and that this value may be used as an optimal condition.

The impulse response also provides much other useful acoutic information - the formation of reflection sequences and echoes etc. Furthermore the STI can be directly computed from the impulse response - though until a suitable commercially available analyser becomes available - this is likely to remain very much a research tool.

A method of intelligibility prediction which is rapidly gaining acceptance is to predict the direct to reverberant ratio and direct to background noise ratios, and from these data make an assessment of the likely intelligibility. Figure 1 relates the direct to reverberant ratio information to % Alcons values. The accuracy of the method very much depends on the ability to predict accurately the reverberant field - procedures other than the straightforward Sabine approach generally therefore have to be adopted. This method also has the advantage that the gain before feedback margin of the system can also be estimated.

SYSTEM FREQUENCY RESPONSE AND CORRECTION

The foregoing discussion on speech intelligibility measurement assumed, as do the measurements themselves, the sound system to exhibit a notionally linear frequency response. However, in practice, this is unlikely to be the case, either due to the characteristics of the loudspeakers themselves or due to the interaction of their acoustic power outputs and the acoustics of the room or space in which they are operating. Equalisation has for many years been used to 'tune' a sound system to its acoustic environment.

Graphic equalisers are most commonly used on the basis that their filter adjustment (cut/boost) potentiometers direclty mimic the action of the equaliser. This, in fact, is an erroneous assumption for all but non combining filters.

Many sound systems still suffer from either inadequate or complete absence altogether. One-third octave equalisation is rapidly becoming adopted as the minimum requirement, the resolution of octave or two-third octave units being too coarse for the majority of applications. For many situations the additional flexibility and resolution of fully tuneable parametric equalisation is also necessary in order to optimise moth the overall response of the system or the desired feedback margin.

Industrial and commercial public address systems as well as the more sophisticated leisure and entertainment sound systems now incorporate broad band one-third octave equalisers as standard in order to enhance both intelligibility and quality.

OTHER FORMS OF SIGNAL MANIPULATION

Other, more sophisticated, forms of signal processing are also becoming increasingly used - again both in industrial, commercial and leisure/entertainment systems.

Although the application of audio signal delay lines dates bask to at least the early 1950s (e.g. Parkin at St. Paul's Cathedral) it is only within te last few years that the cost of such devices has made them a viable system in abll but the most prestigious and costly systems.

The ability of current designs to achieve resolutions down to 10 or 20 microsecond has opened up a new area of appliction to compensate for the short path length variations between component units in multiple loudspeaker clusters - enabling both more combined output to be achieved and the desired polar response to be preserved due to the preservation of the correct phase relationships.

After equalisers, compressor/limiters are probably the most commonly used forms of signal processors enabling a more constant speech signal with a higher average power component to be achieved.

With the exception of the delay line, most signal processing continues to be carried out in the analogue domain with one unit being daisy chained to the next. It is anticipated that within the near future this trend will decrease, being replaced by a single digital processor capable of performing all of te above functions enabling the sound system designer total flexibility and greater precision in achieving the desired result.

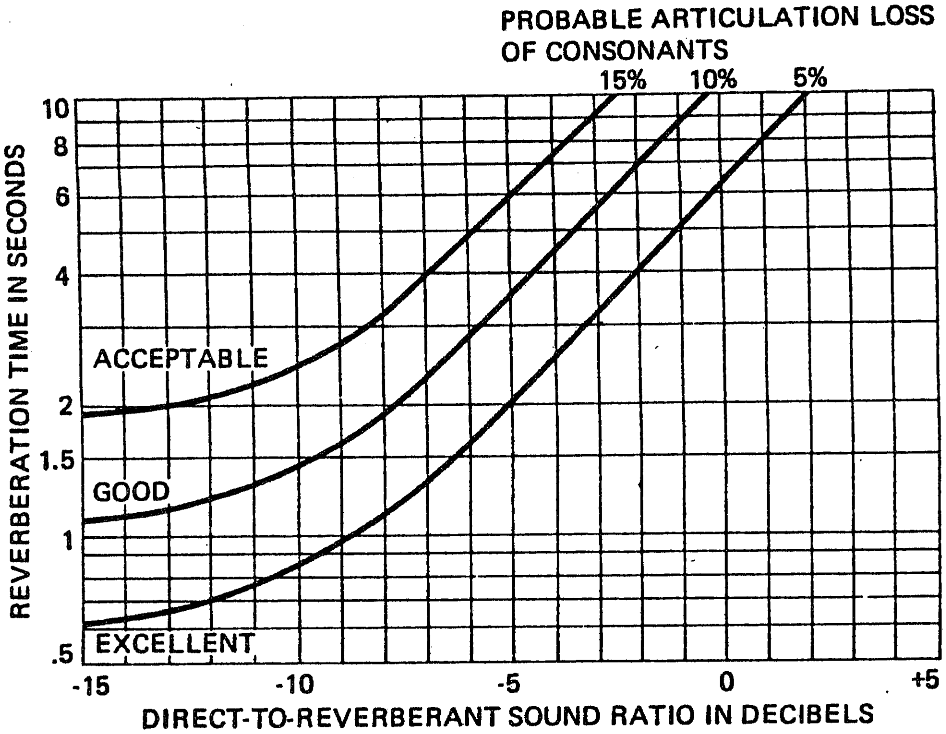

Chart developed from data published by V.M.A. Peutz in
Journal of the AES, Vol. 19, December, 1971

Probable Intelligibility as a Function of Reverberation Time and Direct-to-Reverberant Sound Ratio