Proceedings of the Institute of Acoustics

A SIMPLE MEANS OF IMPROVING THE QUALITY OF SPEECH FROM A DIVING HELMET

C.D. Mathers *and* M.D.M. Baird

BBC Research Department, Kingswood Warren, Tadworth, Surrey, England.

1. INTRODUCTION

Two methods are commonly used to supply a diver with air. The simpler and older is to pump air continuously from the surface via a hose; excess air escapes from under the diver's helmet or bell. This is known as the free-flow system. The alternative is the SCUBA system (Self-Contained Underwater Breathing Apparatus) which consists of one or more air bottles carried by the diver, and a 'demand valve' which supplies air only when the diver inhales; this is needed to conserve the finite quantity of air available to the diver.

SCUBA diving has the advantage of self-sufficiency, and is the usual choice of system for BBC broadcasts. Unfortunately, the demand valve generates a considerable amount of acoustic noise, whose level greatly exceeds the typical level of the diver's speech; in addition, the air noise is of a subjectively obtrusive nature which is detrimental to viewers' enjoyment of the programme, and fatiguing to programme staff.

Some years ago, at the request of the BBC Natural History Unit at Bristol, an air-noise suppressor circuit was developed [1]. This operates on the signal from the diver's microphone, using a tunable filter to identify the demand-valve noise, which chiefly occupies the upper octave of the audio band. It needs continuous adjustment of the filter frequency, because the characteristics of the noise in the helmet depend both on ambient pressure (proportional to depth) and on the remaining pressure in the air bottle. The unit has been used very successfully in a number of programmes during the past five years, but suffers from one major problem: namely, the difficulty of discriminating between speech and high-frequency noise. As might be expected, sibilants are the most difficult speech sounds to distinguish from unwanted 'hissing' noises. In particular, the recent adoption of a 'bubble' helmet, which provides an unobstructed view of the diver's face, rather than the previous type of helmet and mask (which resembles a war-time gas mask) has brought the unforeseeable consequence that the noise spectrum more than ever resembles that of speech.

Direct access to the air inlet system offers the advantage of much more positive discrimination between speech and air inlet noise, and forms the basis of a very simple noise suppressor which is compact enough to be fitted inside the helmet, thus ensuring that unpleasantly high noise levels never appear in any signal path. This offers practical and economic advantages over the alternative strategy of more sophisticated signal post-processing.

2. CHARACTERISTICS OF DEMAND VALVE

In the SCUBA system, the air must be bottled at high pressure so that a reasonable amount can be carried in a bottle of manageable size. The bottles are filled to 205 bars, and the pressure at any later time is of course a direct indication of the amount of air remaining. To ensure the consistent operation of the demand valve, it is fed from the air bottle via a 'first stage' regulator at an approximately constant pressure of 8 bars.

The demand valve is a passive mechanical device which ensures that the the pressure inside the helmet does not differ from that outside by more than about 0.005 bars. It therefore admits air to the helmet on detection of the very small relative pressure drop caused by the diver's inhalation. The valve is operated by a pressure-balanced piston, driven via a system of levers by a diaphragm of about 50 mm diameter, one side of which is exposed to helmet pressure, the other to external pressure. At peak flow rates of at least 10 litres/sec. (corresponding to a typical full inhalation in 0.5 sec., the maximum air velocity in the valve is very high, so that turbulent flow would be expected to generate very high sound pressure levels in the air inlet ducting.

From the demand valve, the air is admitted to the helmet via two tubes of length about 150 mm and of cross-sectional area about 3 mm^2: see Fig. 1. The tubes are intended to direct the air stream against the front surface of the helmet 'bubble' to prevent condensation; they have in additon three significant effects from the viewpoint of the work presently described. First, by providing an acoustic mismatch, they greatly attenuate the noise in the helmet that would otherwise be generated by the demand valve. Second, the acoustic mismatch causes equally effective attenuation in the reverse direction, so that the diver's speech is similarly attenuated on its route to the demand valve. Third, because of the viscosity of air, they cause a 'static' pressure drop of up to about 1 bar, so that the pressure at the output of the demand valve can temporarily exceed the helmet pressure by that amount.

By using a microphone connected pneumatically by a sealed tube to the demand-valve outlet, it is possible to achieve a high degree of discrimination between valve noise and speech. The noise pressure generated by the valve close to its output varies monotonically with flow rate, as shown in Fig. 2, which also indicates the static pressure drop in the tubing. As expected, the sound pressure levels are very high indeed, varying from 120 dB on light breathing to more than 150 dB on full demand; such levels exceed those of the diver's speech by at worst about 30 dB, and are therefore very easy to identify by the use of simple circuitry.

As an aside, it is interesting to note how evenly spaced on a logarithmic scale are the authors' subjective perceptions of breath intake –*very light, light, normal, hard, very hard, maximum*. These comments were added only to provide some intuitive idea of air flow rate. The rate described as *'light'*, corresponding to a sound pressure level of 130 dB as perceived by the noise microphone, was chosen as the trigger level for the noise gate.

3. NOISE SENSING MICROPHONE AND ITS ENCAPSULATION

The noise sensing microphone should preferably be of small size, and must be capable of operating at extremely high sound levels; it must also withstand sudden pressure changes of up to one bar without damage and without momentary loss of output.

To determine the conditions of noise and static pressure as shown in Fig. 2, a semiconductor pressure sensor was connected by a tube to the outlet side of the demand valve. The sensor chosen has a bandwidth extending from DC to about 5 kHz; its dynamic range extends to 2 bars gauge pressure, far in excess of that for any known microphone. The use of a Fast Fourier Transform analyser for the measurements allowed the convenient display of pressure as a function of both time and frequency. The helmet is normally sealed around the diver's neck by a very close–fitting rubber collar, which is uncomfortable to wear in room conditions, and very difficult to fit and remove. The measurements were carried out by inverting the helmet, and breathing into it through a large funnel which was first inserted into the collar from inside the helmet, then pulled from outside by its spout until the collar gripped tightly around its cone profile, forming a good seal. The demand valve then of course 'thought' that the helmet was occupied.

Fig. 3 shows a typical plot of pressure against time for a breath intaken as sharply as possible. In Fig. 4 the noise spectrum is shown in third–octaves for the brief period of peak flow. The noise levels plotted in Fig. 2 represent the highest recorded amplitude in any third–octave band; this is typically at a centre frequency of 1.6 kHz.

Even though the noise–sensing microphone was to be sealed from water ingress, it was felt that if possible, a low–impedance (probably dynamic) type would be preferable to, say, an electret type which depended on very high internal impedances; also, as the two other transducers in current use (earpiece and speech microphone) are passive, a passive noise sensor would enable all electronics to be placed outside the helmet if required, using a three–pair underwater connector. (Because of electrolytic effects, underwater connectors carrying DC can cause noise in nearby connectors carrying small signals, and electret microphones typically incorporate their own unity–gain preamplifiers which require a DC supply but still provide only very low signal levels.)

The first transducer to be tested as a noise detector was an earpiece identical to that used for surface–to–depth communication; a reduction in the number of *different* components in any system is of course always welcome. This device is of the moving–iron type; unfortunately, the iron diaphragm was forced against the magnet assembly by the static pressure, so that the dynamic output tended suddenly to disappear at high static pressures.

A brief investigation suggested that a subminiature moving–iron pressure microphone manufactured by Knowles Inc. might be suitable; with a response falling at about 12 dB per octave below 1 kHz, it seemed unlikely that it would be affected by the sudden rise in 'static' pressure shown in Fig. 3; the only matter for concern was its possible behaviour in the presence of extremely high sound pressure levels within its nominal frequency range of 1 to 5 kHz. In fact, the lower frequency limit is set by a hole punched in the diaphragm; no problems therefore occur due to ambient pressure changes, however large, provided that some limiting rate of pressure change is observed; although unknown, this rate apparently exceeds that of Fig. 3, which was readily tolerated by all of the individual transducers tested.

No problem was expected due to the sharp rise in pressure. Less predictably, the response of the microphone at high noise levels proved excellent; it remains linear up to about 140 dB SPL, and suffers a drop in sensitivity of only 6 dB at the maximum SPL of 154 dB. As previously mentioned, the chosen switching level for the noise gate is 'light' inhalation (130 dB SPL), corresponding to 0.5 V p–p microphone output; maximum rate of inhalation (probably limited by the maximum flow rate through the demand valve, and generating about 154 dB SPL) corresponds to 4 V p–p.

Problems can however occur in the provision of a hermetic seal against moisture ingress for any noise sensor required to work under ambient pressures varying from 1 to 5 bars. The first attempt to achieve this was based on the supposition that about 50 % of the volume inside the

A SIMPLE MEANS OF IMPROVING THE QUALITY OF SPEECH FROM A DIVING HELMET

microphone capsule would be occupied by solid components such as the magnet, leaving 50 % as airspace. The microphone measures about 8 x 6 x 4 mm; sound is admitted via a port of diameter about 0.25 mm, located in the centre of one 8 x 6 mm face. A bubble of thin polyethylene was sealed around the microphone, enclosing about 6 times the microphone's estimated free internal volume. A ready source of suitable 'bubbles' is an air-cushioned polyethylene packing material commercially available in sheet form.

The microphone is contained in a metal housing, which in turn is sealed into the box containing the related electronics. To check whether the bubble was of sufficient volume to avoid complete collapse at 5 bars, with consequent blocking of the microphone port, a special housing was machined from acrylic material to permit observation. As an additional check on whether the ingress port was blocked, the electrical impedance of the microphone was measured at 1 kHz. This reflects the mechano-acoustic impedance seen by the microphone; at 1 kHz under normal conditions it is largely inductive, reflecting the mechanical compliance of the diaphragm.

When the port is blocked, the inductance changes from about 600mH to 100mH at atmospheric pressure. Fig. 5 shows plots of inductance against pressure (a) for the microphone on its own, and (b) with a bubble seal. The sharp discontinuity in curve (b) shows that the port is obscured at quite a low external pressure; this occurs because the bubble assumes an irregular wrinkled shape, an event which was unforeseen and which proved the value of the transparent housing. The solution adopted was the partial replacement of the air in the bubble by a substance of lower limiting compressibility; in this case felt was chosen. The inductance curve (c) is for a microphone with a bubble containing a small cylinder of felt; it follows much more closely that of the microphone on its own, curve (a), showing that it is possible to obtain a hermetic seal without greatly affecting the microphone's acoustic properties.

The details of Fig. 6 show the physical arrangement finally adopted. The use of felt alone caused problems because when forced by ambient pressure against the front face of the microphone, the felt covered the port, greatly increasing its acoustic resistance. A layer of woven-wire gauze placed between the felt and the microphone face ensures that the port remains unblocked by the felt at pressures of more than 7 bars, the maximum that could be conveniently obtained with the available facilities.

4. NOISE-GATE CIRCUITRY

4.1 User requirements

In previous diving work, the signal from the speech microphone had been relayed to the surface by cable without amplification, and some crosstalk had occurred between the microphone conductors and those carrying the relatively high-level signal to the diver's earpiece. It was considered by the users that a 30 dB gain stage would provide a convenient signal level, and essentially eliminate crosstalk and cable noise problems, even with the proposed doubling of the present 200m cable length.

Another requirement of the circuit design was that it should be powered by its own battery, both to minimise the number of cable cores and to avoid the noise problems mentioned above which can occur with underwater connectors carrying both DC and audio signals. A PP3 (9V) battery was chosen because of its small size, universal availability, and convenient voltage. The typical life is about 100 hours to 'exhaustion' at 7V, so there is no need for an on/off switch with its attendant waterproofing problems. In practice, a new battery is fitted at the start of each day's diving, so that the probability of failure during a broadcast is extremely small.

Experience gained in using the previous noise suppressor had shown that it was undesirable to suppress the demand–valve noise completely, because this led to unnatural silences; indeed, the objective has always been to preserve verisimilitude as far as is compatible with comfortable listening. The optimum degree of suppression seems to be around 20 dB, which is the value chosen for the new equipment.

4.2 Circuit operation

The circuit diagram of the equipment is shown in Fig. 7. It is based on a quad CMOS operational amplifier chip, which is compact, economical on power, and has an output voltage swing capability of close to 100% of supply voltage. The noise signal is amplified, peak–rectified, and compared to a reference voltage. At a predetermined noise level corresponding to 'light' inhalation, the FET TR2 is switched off, reducing the gain of the output stage by 20 dB. Some hysteresis is provided so that the circuit does not 'dither' on the point of switching. No absolute voltage reference is provided, so that the switching threshold is proportional to battery voltage; but a drop from 9 V to 7 V changes the noise level at which switching occurs by only 2 dB, which for the required purpose of the equipment is negligible. At the request of the users, an unswitched output stage was provided as a backup channel. The output stages are driven from a speech microphone preamplifier TR1 which provides a low–noise front end (CMOS amplifiers have poor noise performance at low input impedances). The output transformers have a stepdown ratio of 5:1, which makes it easy to drive low–impedance loads from low–current output stages; the equipment will readily drive 400 m of commonly available cable, which appears as a capacitive load of about 30 nF.

The overload level of the speech channels is about 23 dB above the signal level corresponding to 'normal' speech from the diving helmet microphone, a Shure type SM 10A.

5. CONCLUSION

A very simple piece of equipment has been designed which can be accommodated within a typical diving helmet, and which combines a microphone amplifier with a noise gate to attenuate air demand–valve noise to an acceptable level. The device has potential applications both in broadcasting and in commercial diving.

6. REFERENCE

[1] D.J. MEARES *and* K.F.L. LANSDOWNE, 'Improvements to speech quality from a diving helmet' *BBC Research Department Report* No. 1981/9 (1981)

ACKNOWLEDGEMENT

The authors wish to thank the Director of Engineering of the BBC for permission to publish this paper.

A SIMPLE MEANS OF IMPROVING THE QUALITY OF SPEECH FROM A DIVING HELMET

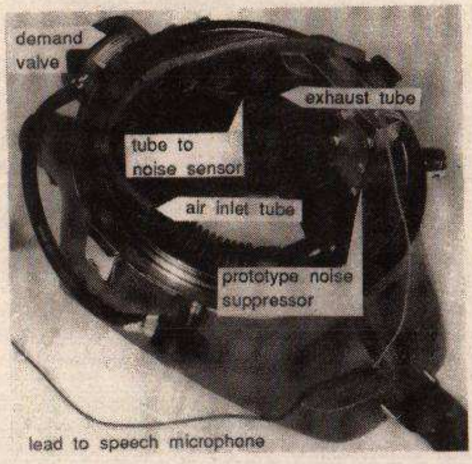


Fig. 1 Photograph of diving helmet with 'bubble' removed

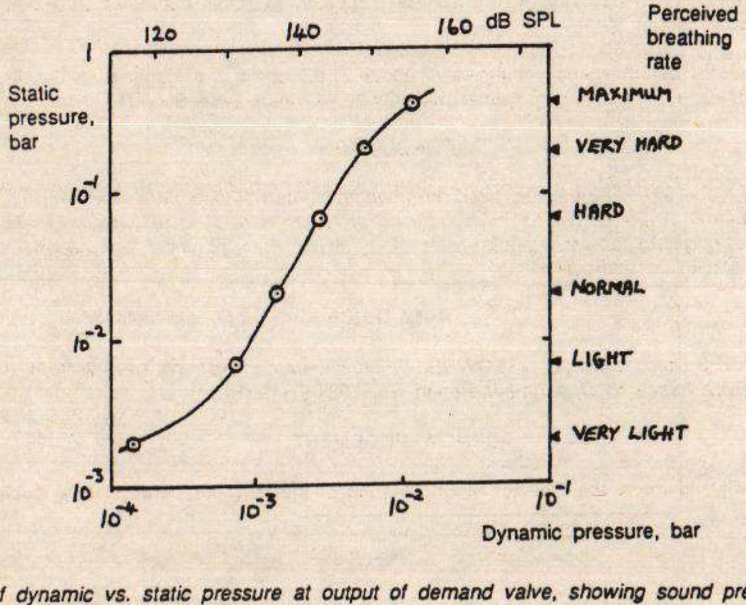


Fig. 2 Plot of dynamic vs. static pressure at output of demand valve, showing sound pressure levels and perceived inhalation rates

A SIMPLE MEANS OF IMPROVING THE QUALITY OF SPEECH FROM A DIVING HELMET

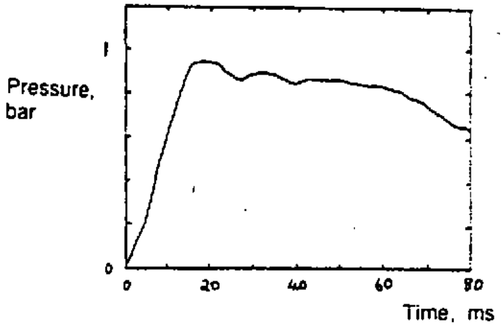


Fig. 3 Plot of static pressure vs. time at demand valve output for very sharply intaken breath

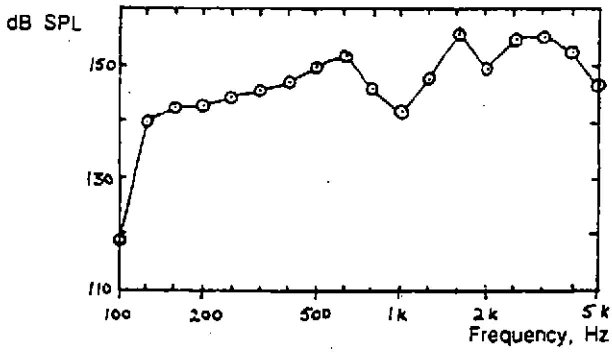
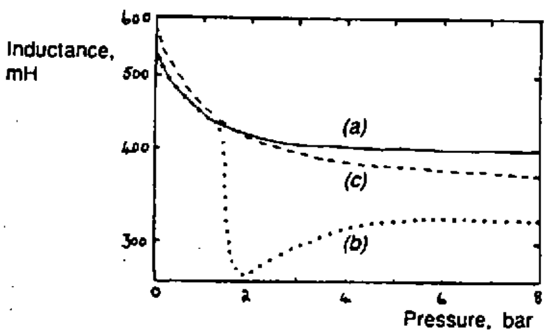


Fig. 4 Peak noise spectrum in one–third octaves at demand valve outlet for very sharply intaken breath



(a) Without seal

(b) With bubble seal only

(c) With bubble seal partly filled by pellet of felt

Fig. 5 Plot of microphone inductance vs. pressure for various sealing arrangements

A SIMPLE MEANS OF IMPROVING THE QUALITY OF SPEECH FROM A DIVING HELMET

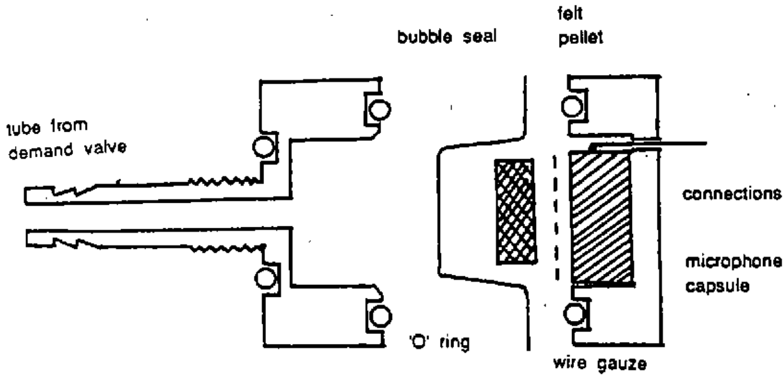


Fig. 6 Assembly diagram of housing for noise microphone capsule, showing details of hermetic seal

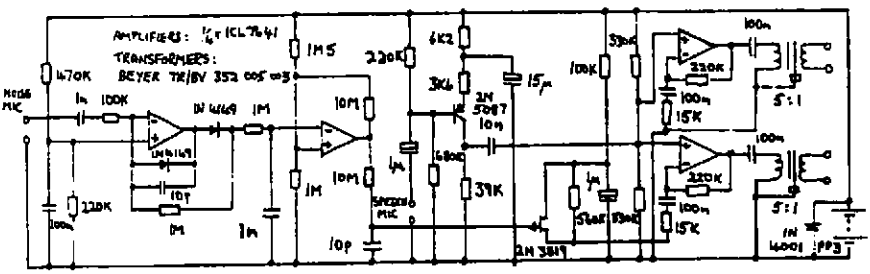


Fig. 7 Circuit diagram of noise suppressor

Proceedings of the Institute of Acoustics

RECENT ADVANCES IN SPEECH INTELLIGIBILITY AND SOUND SYSTEM DESIGN
Part I: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

N. Abdulrahim, P. Doany

Dar Al-Handasah Consultants (Shair & Partners), London

ABSTRACT

After a brief review of the latest advances in computer-assisted acoustical design packages developed by others, this paper describes A-CADI, a programme specifically tailored for the performance predictive analysis of acoustically difficult venues, where user-defined parameter flexibility is enhanced, with the application of a multitude of methods of analysis. These features include aspects relating to system design, coverage, reverberation time, time/frequency analysis (naturalness/annoyance), intelligibility predictions, gain-before feedback, in addition to overall system equalisation and control. Only a brief description of the data processing capability of objective measurements is given, these being discussed in more detail in the second part of this paper.

1. INTRODUCTION

Various PC-based computer programmes have been developed in recent years to assist in the analysis and design of sound reinforcement systems. This introduction describes some of the recent releases.

Altec Lansing's *AcoustaCADD*, is described by Lanphere [1]. Special features of the programme include general acoustics and sound system design, with specific facilities for cluster systems, further to powerful graphics and statistical ray-racing. The system can handle up to 20 clusters, with up to 40 loudspeakers in each. This is an extremely powerful programme, which has only been recently released.

JBL's Central Array Design Program, *CADP*, is described by Eargle & Kalmanson [2-3]. It is particularly useful in central-cluster system design, and is capable of direct-field calculations, direct-to-reverberant ratio calculations, and intelligibility predictions, in addition to mechanical aspects of array design. Up to 20 loudspeakers may be handled by this programme.

Bose Programmes, *Modeler* and *Speaker-CAD* are described by Birkle and Jacob [5-6], including acoustical space modelling, sound field calculations, intelligibility predictions and reverberation time analysis.

Prohs [4] describes the latest generation of *the PHD* programme. Other programmes and computer analysis techniques are described by McCarthy [7] and Anhert [8], amongst others [9-11].

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

2. OBJECTIVES OF A-CADX

In the course of carrying out the design of a complex sound reinforcement system with severe constraints, comprising several sub-systems of different characteristics with a total of over 3000 loudspeakers of various types, the Authors were confronted with the following problems:

(a) lack of a clearly defined set of *"quantitative" criteria*: in one sub-system for example, this related to analysis of intelligibility, involving numerous discrete echo contributions under near-free-field conditions. Although there appears to be general agreement on the effects of noise and reverberation on intelligibility, the effects of echo are not at all clear [12-13], and it was necessary therefore to retain some flexibility in the design criteria, with continuous refinement.

(b) several complex sub-systems involving *large numbers of loudspeakers*: in one sub-system for example, well over 100 loudspeakers needed to be considered at a time, so that the iterative calculations were far too laborious, due to the excessive degrees-of-freedom involved.

It became apparent very early on in the project that a specially-tailored programme needed to be developed, as the accuracy of manual methods was simply impossible to sustain, in addition to the fact that other available packages at the time were unable to provide the flexible features required (work actually carried out in 1987).

In this context, *Dar Al-Handasah Consultants* developed a programme entitled *A-CADX (Acoustic - Computer Assisted Design - XBASIC)* to provide the necessary features. Figure 1 gives a basic schematic diagram of the system functions, comprising the following components:

(a) *Main-menu*: this offers the user a choice of features, essentially comprised of "predictive analysis" and "objective measurement analysis" functions.

(b) *Template*: this feature enables the user to obtain appropriate "over-lays", based upon the scale of architectural plans being used, whereupon a proper choice of loudspeakers may be made, in addition to their tentative positioning. The latter parameters can then be refined by use of the main programme.

(c) *Data-entry*: user enters the above data, as appropriate. A memory-resident set of data is provided, for the loudspeaker models generally used, and the user can store additional data as necessary.

(d) *Plan generation*: a plan of the study area is generated here, giving location of each loudspeaker, with the aiming point being indicated.

(e) *Predictive analysis*: several features are included, such as:
- uniformity of coverage (direct-field)
- reverberation time
- time/frequency domain analysis (echo annoyance, naturalness criteria)
- intelligibility analysis
- gain-before-feedback analysis
- other features (optional loudspeaker parameters): operating level assessment, electronic time-delay, and equalisation).

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

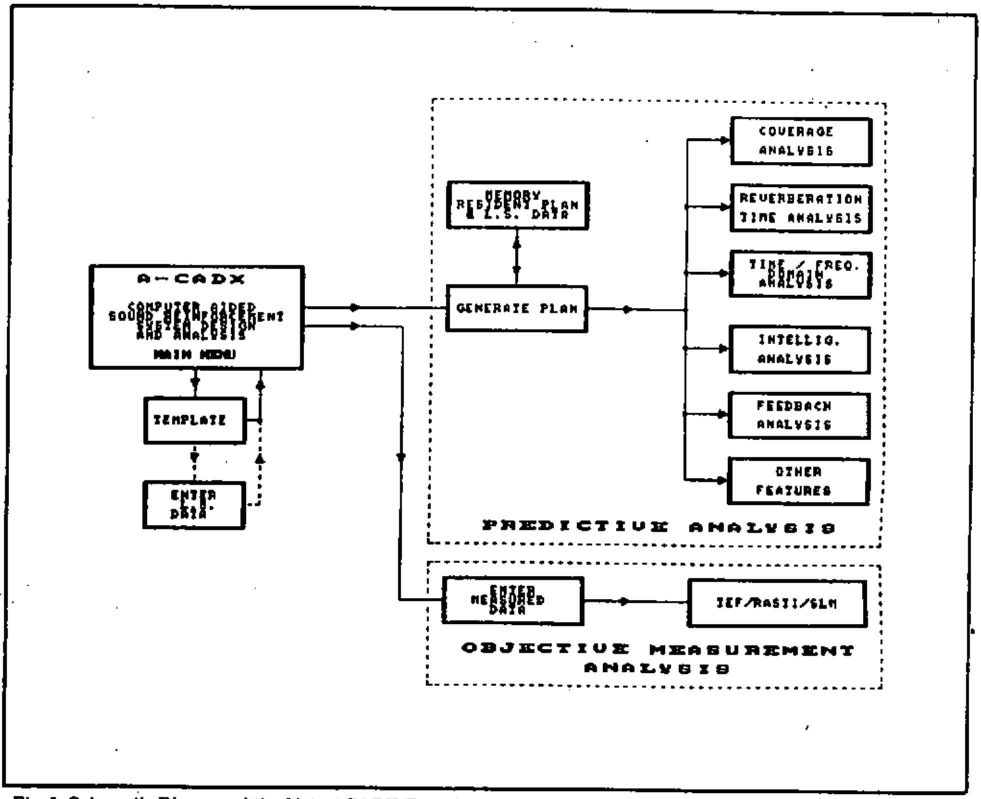


Fig. 1 Schematic Diagram of the Main ACADX Functions

(f) *Objective measurement analysis*: this includes analysis of measurement results noted by TEF/RASTI/SLM, mainly relating to parameters of intelligibility analysis.

A-CADX has been written in XBASIC, and runs on IBM PC, PC XT, PC AT, PS/2 or 100% compatible machines running MS-DOS. The original version of A-CADX was written for Hercules graphics, but recent versions are also available for EGA and VGA color graphics.

3. TEMPLATE DESIGN AID

"Template" is a simple graphical method developed by Tappan [14], which is useful in arriving at suitable choices of loudspeakers based upon available dispersion characteristics, with tentative definition of positioning and orientation, to achieve some coverage criteria. This is based upon the use of a transparency "overlay" which shows the loudspeaker aiming angles, the attenuation of sound pressure level (SPL) due to inverse square law, and angular dispersion. The difficulty in the application of this method is that a different graph is needed for each scale and loudspeaker mounting height encountered. With the help of A-CADX a template printout can be obtained virtually immediately and a transparency/photocopy made to suit any scale and mounting height being used. Figure 2 gives a sample printout.

T. Uzzle [15] describes another method, which may be easily incorporated. This is based upon spherical mapping (light projection method). The authors have opted for the template method.

4. DATA ENTRY PARAMETERS

The basic entry parameters are:

- *choice of loudspeaker* (polar characteristics in various bands, normally 500Hz, 1kHz and 2kHz),
- *loudspeaker sensitivity*, and frequency response: SPL (1W,1m), in each of the above octave bands.
- *loudspeaker operating power* (with individual attenuation as a user-option).
- *choice of mounting position* (co-ordinates)
- *choice of aiming point* (orientation)
- *options* (artifial time delay and equalisation)

Based upon the layout of the coverage area being studied, a plan is generated on the screen, with further allowance for the inclusion of *environmental data*. Figure 3 gives a sample plan for a simple area (ten loudspeakers), with menu options of that specific mode being indicated.

The user also enters *ear-height*, enabling assessment of various conditions. Auditory directional characteristics may be considered through the microphone sub-programme. Other user data entry relates to reverberation analysis and microphone characteristics.

5. SYSTEM PERFORMANCE EVALUATION FEATURES

Having entered the parameters, the programme is now ready to carry out any of the user-defined tasks.

5.1 Calculation Subroutine (Direct Field Calculations)

The main subroutine calculates the direct SPL (dB) and associated acoustical time delay for each loudspeaker contribution at the defined observation points. Each calculation consists of obtaining the inverse square law attenuation, adding the off-aim dispersion loss, further to the air-absorption loss in the various octave bands. This procedure is repeated for every loudspeaker present in the system, resulting in a multi-dimensional array of size equal to the number of loudspeakers for every observation point considered.

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

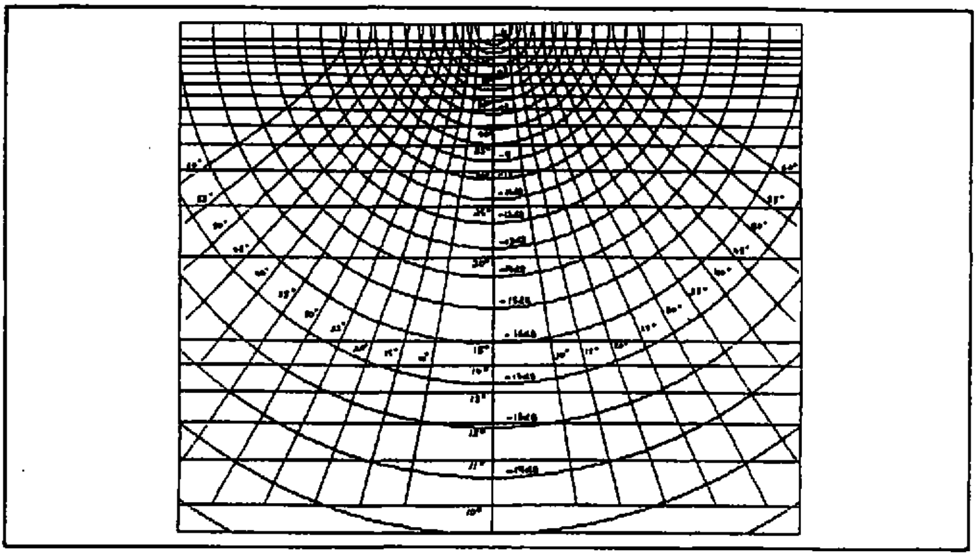


Fig. 2 Sample Template Print-out

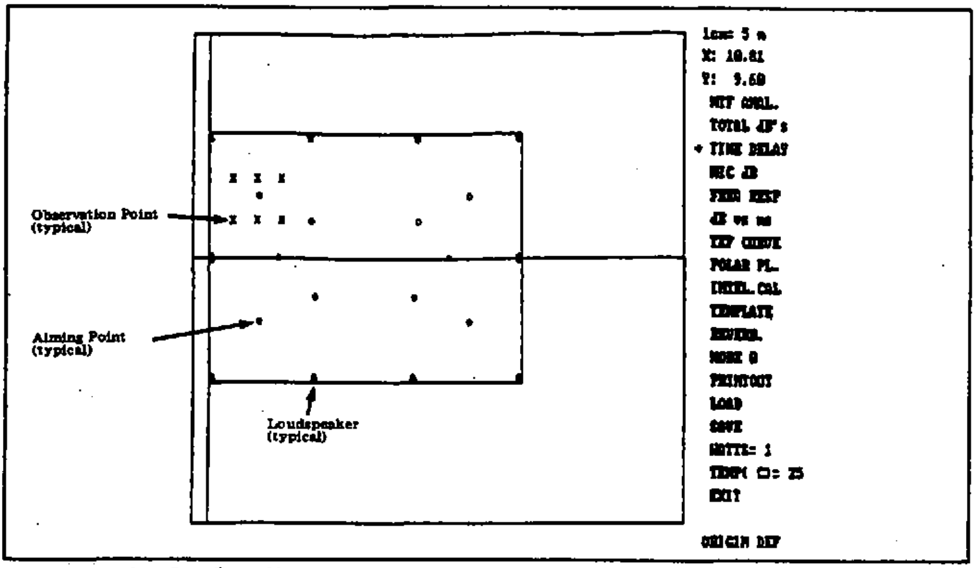


Fig. 3 Sample Plan/Data-Entry Menu

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

5.2 Uniformity-of-Coverage

The total direct contribution at any specific point consists of integrating the direct SPL at each observation point, as defined by the user. From these results, it is possible to check overall uniformity, in addition to checking the frequency response flatness criterion (calculations performed at 500Hz, 1kHz and 2kHz, with further extension to the bands from 250Hz to 8kHz, as required). This also allows the user to check whether or not equalisation will yield the desired performance criteria, under near-free-field conditions.

5.3 Reverberation Time Calculations

The basic models used in these calculations relate to the classical formulae of Sabine and Fitzroy [16], with allowance for other imperical results, such as Friberg [17]. In some of the sub-systems of the study, it has been necessary to develop other models of analysis. This has been necessary in cases where the sound reinforcement system is expected to have a marked effect on reverberation, further to difficulty in coping with a virtually "boundless" volume. For the latter, a three dimensional coupling method [18] was applied, which is the subject of a paper to be published shortly. The programme gives these options to the user, as suited to the problem being studied.

5.4 Time-Frequency Analysis

This feature of the programme provides the user with analysis options in both the time domain and the frequency domain. These are essentially suitable for near-free-field conditions only, with special provisions being required for diffuse-field conditions by consideration of reverberation time and direct-to-reverberant ratio, D/R.

(a) Time Delay/Early Integration: this feature of the programme is a simple extension of the fundamental direct field calculations. It helps locate the time delay annoyance areas and identify which loudspeakers are responsible for this annoyance. The calculations are performed for a single point at a time, and the output is given in the form of a table. Figure 4 gives an example for a complex study area comprising 98 loudspeakers. To assist the user in the analysis, integrated values are given for three durations: 0-50ms, 50-100ms, > 100ms, further to the full tabulated results. These parameters can be easily changed by the user, for analysis as required.

(b) Energy Time Chart (ETC): the above tabulated results are also presented in the form of a bar chart as shown in Figure 4. This consists of plotting the SPL contributions integrated over defined intervals (1-10ms) as a function of time. Superimposed on this plot is a 20% annoyance curve which indicates the maximum allowable delayed level to attain a 20% listener disturbance performance, (based upon a user-selected pseudo-reverberation time, applied to the results of Nickson at Al [19] and Haas [20]). The background noise entered by the user in this example is 55dB. Shorter integration periods than the indicated 10ms can be used for improved accuracy, at the expense of processing time.

(c) Energy Frequency Chart/Curve (EFC): this calculation is done in 1Hz steps. The polar responses of the loudspeakers in the regions between the 500Hz, 1kHz and 2kHz octaves are approximated by interpolation. The output can be presented in bar-chart form (Figure 5a), where the resultant SPL value is given for each frequency increment, or as a condensed graph (Figure 5b).

(d) Energy Frequency Time (EFT) Curve: the energy frequency time graph is a 3-dimensional plot of SPL against time/frequency. Figure 6 gives a sample plot.

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

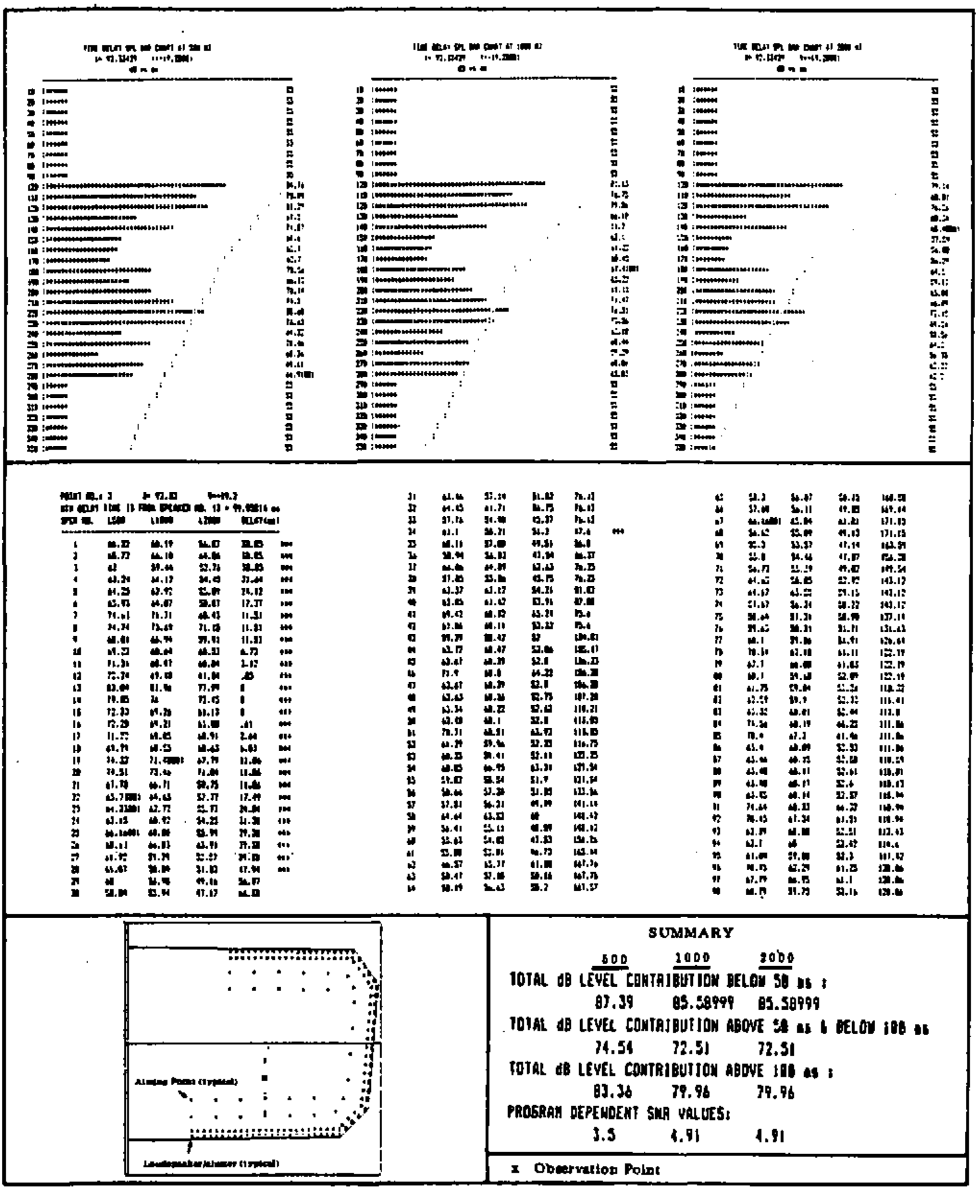
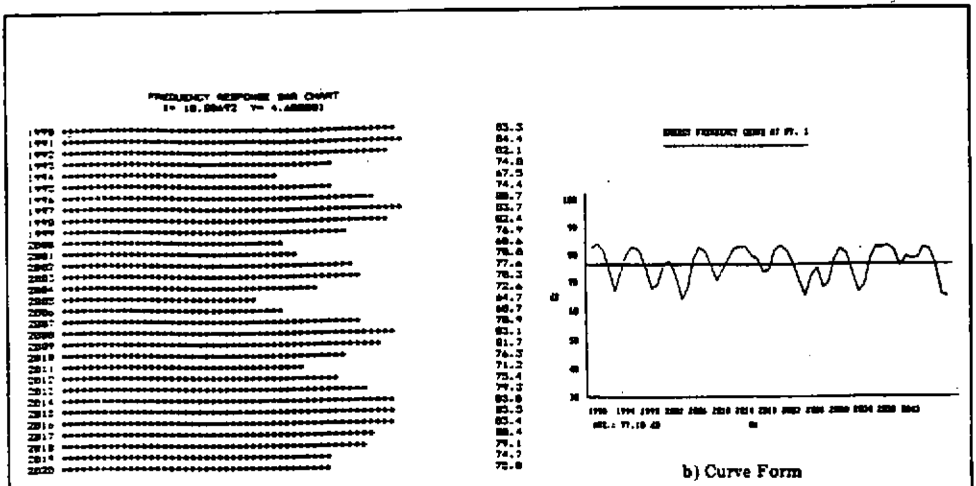


Fig. 4 Sample Result Illustrating ETC and Expected Annoyance Region

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS



a) Bar Chart Form

b) Curve Form

Fig. 5 Sample of Energy Frequency Chart/Curve (EFC)

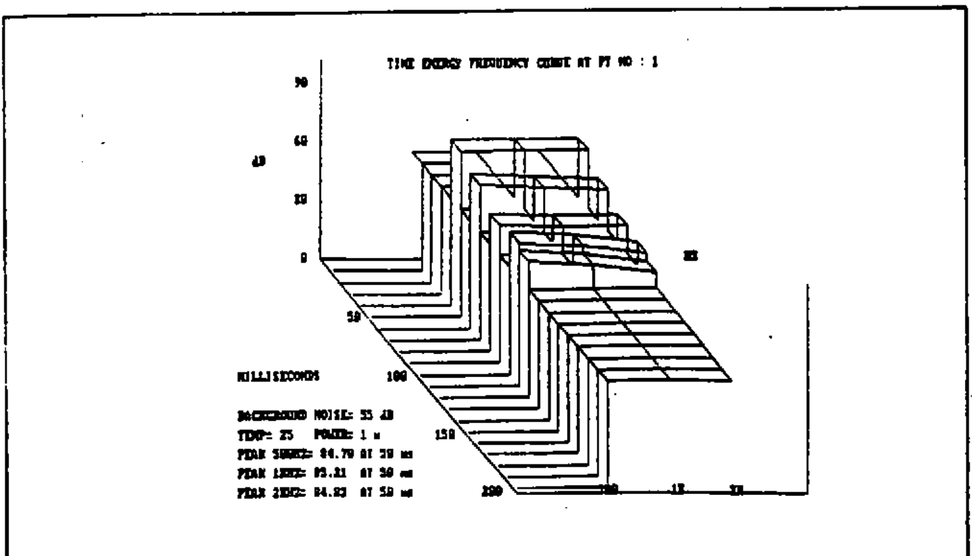


Fig. 6 Sample of Energy Frequency Time (EFT) Curve

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

5.5 Speech Intelligibility

The various methods of speech intelligibility analysis are outlined by Smith [13], amongst others. These methods are further expounded by Doany [21] and Mapp/Doany [22], in the context of this paper, including ALcons, STI and the modified-S/N method, as briefly described below:

(a) ALcons Method: ALcons, the Articulation Loss of Consonants method is based upon the work of Peutz [12], where ALcons is given as a function of the D/R, S/N and reverberation time. The diffuse field is computed as a function of room constant with various modifiers as appropriate [11]. Alternatively, the programme may also perform the calculation as a function of acoustical power and loudspeaker efficiency. The calculation is performed in the 2kHz band.

(b) STI Method: STI, the Speech Transmission Index, is based upon the work of Steeneken & Houtgast [23]. A-CADE offers two methods here. The first is based upon the work of Steeneken and Houtgast, but considers D/R, as derived by Doany [21], noted to yield results fairly close to those by Peutz [12]. The second method is based upon a direct computation of MTF, the modulation transfer function, based upon the specular components, somewhat similar to the work of Van Rietschote and Houtgast [24]. For both choices, the user can also select between 1/1 octave MTF (0.5-16Hz), 1/3 octave MTF (0.63-12.5Hz), and RASTI. A sample of a typical output is given in Figure 7, for a 2kHz band (1/3 octave). Similar plots may then be derived in other bands of interest, further to addition of noise "floor" values [23], as required by the user. It is also possible to carry out weighting in the frequency domain, if factors similar to those used in the Articulation Index (AI) are considered appropriate.

(c) Modified-S/N Method: This method is based on the Articulation Index (AI). In this method the speech peak-to-rms noise ratios are weighted and summed to obtain an index [25-26], from which intelligibility may be easily determined. This however, is strictly applicable to "additive" noise. For reverberant conditions, a modified version of this method can be applied, where the early part of the system square-impulse response (reflections/contributions) is treated as "signal", and the remaining part as "noise", thereby defining a programme-dependent S/N ratio [27]. This is normally implemented for an assumed exponential decay.

5.6 Gain-Before-Feedback Analysis

This feature of the programme assists in the analysis of gain before feedback. It takes into account various microphone types, and is mainly intended for near-free-field conditions. Under reverberant conditions, a simple calculation is done as a function of direct-to-reverberant ratio and reverberation time, with a frontal-to-random parameter, as appropriate.

A choice of various microphone types is given to the user, including omnidirectional, cardioid and figure-of-eight. Pick-up levels are computed by noting the direct-field contributions from each loudspeaker, in the various octave bands of interest, after applying the necessary microphone attenuation. Where long delays are present, the user may also obtain an ETC of the pick-up level, for analysis. A sample is given in Figure 8, for 44 loudspeakers, with one open cardioid microphone. Integrated levels are also given in the 50-100ms range.

Based upon the speaker level and choice of an appropriate margin, the operating level can then be set as required, thereby defining the expected SPL values in the coverage area. This routine proved particularly useful in cases where the microphone was required to be located in the midst of the listening audience, with the further constraint of an excessive talker-to-microphone distance.

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

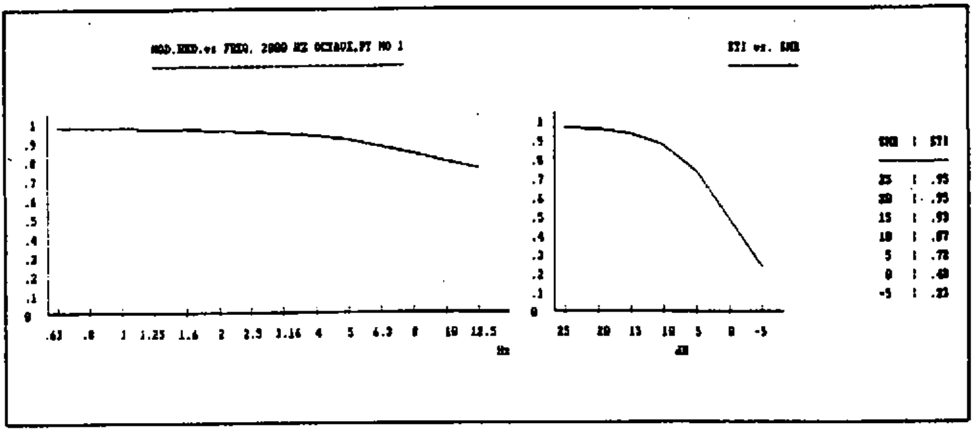


Fig. 7 Sample of MTF & STI Print-out

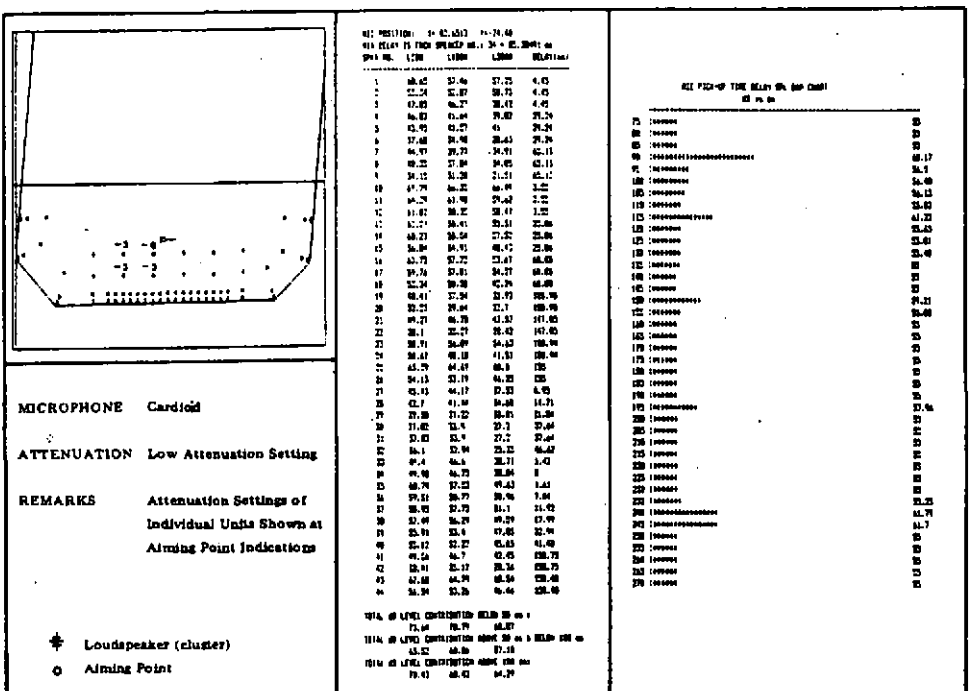


Fig. 8 Sample of Microphone Pick-up Level Analysis

5.7 Loudspeaker Parameters

(a) Loudspeaker Level Attenuation: it is possible here to attenuate the operating level of some loudspeakers, as suited to specific requirements. In feedback analysis, for example, this may be advantageous due to relative microphone proximity. Figure 8 gives an example, where the selected settings are actually shown, for a specified open microphone location.

(b) Electronic Delay: it is possible here to incorporate artificial delay for a group of loudspeakers, the parameter being changed to obtain some desired objective. This feature proved useful in some study areas, where delay-annoyance was inevitable, and thereby enabled its reduction to a reasonable minimum. This could also be useful in the speech intelligibility analysis of some systems, as suited to the method of analysis selected by the user.

(c) Equalisation: it is possible here to include octave equalisation settings to a set of loudspeakers, which is sometimes useful in the analysis of distributed systems under near-free-field conditions, involving highly frequency-dependent overlapping patterns. This enables an assessment of the expected quality of frequency response in the coverage area, and as to whether or not equalisation can actually achieve the flatness response required.

6. CONCLUSION

A number of computer-assisted packages are now available on the market, and these should be seriously considered by professionals involved in the analysis and design of sound systems, prior to considering in-house development of their own programmes.

In-house development of such facilities must be justified. Involvement in projects with severe constraints may provide an incentive towards such a consideration, however caution should be exercised against "over-designing" such systems.

A-CADX has been developed with the express aim of enabling a multitude of methods of analyses to be considered by the user where it is appropriate to do so. The plans for the future include designing an interface to a powerful graphical package, in addition to statistically ray-tracing. Other features being considered for future incorporation include: subjective articulation test analysis; multi-way loudspeaker analysis and optimisation; noise and vibration analysis.

REFERENCES

[1] LANPHERE, J.E., "AcoustaCADD: The Next Generation of Sound System Design Software", *AES 6th Int. Conf.*, 4E, (1988, May)

[2] J. EARGLE & M. KALMANSON, "Design by Computer", *Sound & Video Contr.*, Vol. 2, No. 7, pp 32-38 (1988, July)

[3] J. EARGLE, "Recent Updates for CADP", *AES 6th Int. Conf.*, 4.B, (1988, May)

[4] J. PROHS, "The New Generation of the PHD Array-Design Program", *AES 6th Int. Conf.*, 4C, (1988, May)

[5] T. BIRKLE & K. JACOB, "Modeler Design Program V3.0", *Sound & Video Contr.*, Vol. 7 No. 8, pp 38-50, (1989, Aug.20)

[6] K.D. JACOB & T.K. BIRKLE, "The Latest Advances in Predicting Sound System Performance in Real Spaces ...", *AES 6th Int. Conf.*, 4.D, (1988, May)

SPEECH INTELLIGIBILITY: COMPUTER-ASSISTED PREDICTIVE ANALYSIS

[7] T.G. McCARTHY, "Loudspeaker Arrays - A Computer Assisted Method of Designing", *AES 6th Int. Conf.*, 4.A, (1988, May)

[8] W. AHNERT, "Complex Simulation of Acoustic Sound Fields by the Delta Stereophony System (DSS)," *J. Audio Eng. Soc.*, Vol. 35, No. 9, (1987, Sep)

[9] M.R.SCHROEDER, "Computers in Acoustics: Symbioses of an Old Science and a New Tool", *J. Acoust. Soc. Am.*, Vol. 45, No. 5 (1969)

[10] B.S. ATAL et AL, " Evaluation of Acoustic Properties of Enclosures by Means of Digital Computers", *J. Acoust. Soc. Am.*, Vol. 40, No. 2 (1966)

[11] D. DAVIS & C. DAVIS, "*Sound System Engineering*", pp 235-260, pp 535-537, Howard W. Sons & Co., USA, 2nd Edition, (1987)

[12] V.M.A. PEUTZ, "*Speech Reception and Information*", App. X pp 639-644, in above ref [11]

[13] H. SMITH, "Acoustic Design Considerations for Speech Intelligibility", *J. Audio Eng. Soc.*, Vol. 29, No. 6, pp 408-414, (1981, June)

[14] P.W. TAPPAN, "A Graphic Method for Choosing and Aiming Loudspeakers for Reinforcement", *J. Audio Eng. Soc.*, Vol. 34, No. 4, pp 269-277, (1986, April)

[15] T. UZZLE, "Loudspeaker Coverage by Architectural Mapping", *J. Audio Eng. Soc.*, Vol. 30, pp 412-424, (1982, June)

[16] P. MAPP & G. GOOSSENS, "*The Audio System Designer Technical Reference*", Klark Teknik, England, (Undated)

[17] R. FRIBERG, "Noise Reduction in Industrial Halls Obtained by Acoustical Treatment of Ceiling and Walls", *Noise Control and Vibration Reduction*, pp 75-75 (1975, Mar.)

[18] L. CREMER & H.A. MULLER, "*Principles and Applications of Room Acoustics*", Vol. 1, Applied Science Publishers, (1982)

[19] NICKSON, A.F.B., R.W. MUNCEY & P. DUBOUT, "The Acceptability of Artificial Echoes with Reverberant Speech and Music", *Acoustica*, Vol. 4, pp 515-518 (1954)

[20] H. HAAS, "The Influence of a Single Echo on the Audibility of Speech", *J. Audio Eng. Soc.*, Vol. 20. No. 2, pp 146-159, (1972, Mar.)

[21] P. DOANY, "Speech Transmission Index (STI) Applications in a Reverberant Environment: I. Effect of the Direct Field", Paper Submitted to *J. Audio Eng. Soc.* for Approval, (1989, July)

[22] P. MAPP & P. DOANY, "Speech Intelligibility Analysis and Measurement for a Distributed Sound System in a Reverberant Environment", *AES 87th Conven.*, N.Y., (1989, Oct.)

[23] H.J.M. STEENEKEN & J. HOUTGAST, "RASTI, A Tool for Evaluating Auditoria", *B & K Technical Review*, No. 3, pp 13-39 (inc. Appendix), (1985)

[24] H.F. VAN RIETSCHOTE & T. HOUTGAST, "Predicting Speech Intelligibility in Rooms from the Modulating Transfer Function ...", *Acustica* 53, pp 72-78, (1983)

[25] "American National Standard Methods for the Calculation of the Articulation Index", S3.5-1969, *American National Standards Institute Inc.*, (1969, Jan)

[26] C.M. HARRIS, "*Handbook of Noise Control*", pp 14-1 to 14-9, McGraw-Hill, 2nd Edition, (1979)

[27] H.G. LATHAM, The Signal-to-Noise Ratio for Speech Intelligibility-An Auditorium Acoustics Design Index", *Applied Acoustics*, Vol. 12, pp 253-320, (1979, July).