OBTAINING INTELLIGIBLE PUBLIC ADDRESS AT STATIONS

JOHN W EDWARDS - LONDON UNDERGROUND LTD, SCIENTIFIC SERVICES DIVISION.

INTRODUCTION

London Underground is in the process of a fundamental change. It is setting the needs and aspirations of the customer at the centre of its business. The customers' requirements are the driving energy for the business.

One subject that is sure to arouse interest and comment by customers is the quality of the Public Address. In a recent survey of customer priorities, the quality of station public address ranked of equal concern with crowding on trains.

A customer can pay less than £50 to receive perfectly acceptable quality speech and music at home from a radio station miles away. He fails to understand why the announcements are very often incomprehensible when he's sitting on the platform of his local railway station.

Public address is of course only one means used to convey information to customers. Visual systems such as dot matrix displays as well as fixed signs and notices, complement the public address systems.

For some situations Public Address is clearly the most appropriate information system to use and for one of these, station emergency evacuation, it is of the utmost importance that the message is heard, is intelligible, and understood.

The fundamental aim behind the work, a part of which is described in this paper, was to examine the factors influencing sound quality performance of the Public Address Systems in London Underground. Then, using the knowledge gained from this, to set out in objective and measurable terms the requirements for installing and maintaining Public Address Systems to a specified performance standard.

HISTORIC SITUATION

Overall, the scale of Public Address in London Underground is large. It covers some 245 stations having 700 platforms and uses a total of approximately 20,000 loudspeakers. Whilst there is individual control of the systems at local (i.e. station or platform) level, there is some networking of systems that enables announcements to be made to all stations on an individual line from one announcement point.

Historically the same basic equipments have been provided at all sites, largely disregarding their particular acoustic environment, and the design of the Centralised Public Address is now some 15 years old.

Set-up and maintenance of the systems has been carried out in the main by subjective judgement of whether the system was or was not operating (i.e. output from all the loudspeakers) and whether the output was loud enough. This approach coupled with the variable quality of the announcer's input to the system and failings in the system design itself has resulted too frequently in an unsatisfactory performance of Public Address in London Underground.

The lack of definitive standards for set-up maintenance and operation of the Public Address systems was identified at the onset of the work and it was clear that a key parameter requiring objective measurement was that of intelligibility.

MEASURES OF INTELLIGIBILITY

There is a wide variation in the acoustic characteristics of stations on London Underground ranging from near free field conditions at some surface stations through to highly reverberant spaces found in many of the deep tube stations. The use of the Speech Transmission Index (STI), and its developed form RASTI, as an objective measure of intelligibility is now very well established and indeed is now incorporated in BS 7443 as a specified performance parameter for sound systems for emergency purposes. Nevertheless we considered that for our application, and in particular for the deep tube stations, it was necessary to confirm the validity of STI/RASTI as a suitable measure of intelligibility. Accordingly experiments were carried out measuring STI and RASTI against a specially tailored series of word score tests. The results showed generally good agreement although in the cases of the most reverberant deep tube stations where an RT_{60} (2kHz) of ~2.5s was measured, the reliability of the RASTI measurement when distant from the loudspeaker became more doubtful.

The overall conclusion however was that in general RASTI could be used as a satisfactory gauge of system performance with respect to intelligibility.

FACTORS AFFECTING SOUND QUALITY PERFORMANCE

Within the communication chain (Figure 1) the two end sections, that is the human input to the system and the loudspeaker to listener, present the most difficult technical challenges. The elements of transmission and amplification are, technically of very high performance in modern Public Address systems and in general can be regarded as the strongest links in the communication chain. However problems can still be experienced in these elements, the susceptibility to electromagnetic interference on the long transmission lines between stations being a particular example.

INPUT TO SYSTEM

Considering the input to the Public Address system, it is worthwhile examining the factors that are critical in contributing to the overall performance quality of the system.

There are two principal points, the message/announcement structure and its content, and the inputting of this message to the system by the announcer. Both of these areas contain a great number of interrelated factors, and to address them all in detail lies beyond the scope of this paper. However a number that have particular importance are considered.

Having prepared a suitable message, the manner in which it is delivered into the system is of crucial importance to its eventual intelligibility. There are a number of aspects which must be examined here.

Two which are commonly cited but which are not necessarily the most critical are dialect and accent. Dialect relates directly to vocabulary and language use. It is clear that for immediate understanding of announcements the use of commonly understood words is crucial. The effect of different dialects in terms of sentence construction is unlikely to cause any great intelligibility difficulties unless the word order is dramatically strange. Within reasonable degrees accent too is often not critical. If however it seriously deviates from standard pronunciation then it is likely that the number of people able to understand the announcement will be limited. Nevertheless some variation in accent is desirable as it avoids the tendency of creating a bland, stereotyped "Public Address Voice", an aspect which is lost with many of the auto-announcers or even worse with direct speech synthesis systems.

Two factors which are critical to sound quality are the speed of delivery and diction. For the first of these there is often a problem created by the announcer being remote from the area into which the message is being broadcast. If, for example this is a highly reverberant station at which a speaker would naturally tend to slow his/her delivery, then an announcer in a small and acoustically dead office, removed from the feedback that would be experienced on the platform, will tend to speak at a speed that produces a garbled result to the listener. Control, by means of training, is therefore crucial for announcers to learn how to pace their delivery.

Diction is a critical factor since the majority of the information is carried in the consonants of words. If diction is poor then consonants are missing and ambiguity or a general reduction in intelligibility occurs. Training on this aspect can greatly improve an announcer's performance.

Also important to the overall perceived quality of a Public Address system is the style of the announcement and the image that is conveyed. Thus although diction and speed of delivery may be good, care in presenting a professional image is an aspect that demands appropriate training and continuous encouragement.

SYSTEM OUTPUT

At the end of the communication chain two principal factors influence intelligibility: the frequency response of the system, including the effect of the acoustic conditions of the space into which the system operates, and the background noise level.

Of direct importance to the frequency response is the relationship between speech and intelligibility. This is shown in a simplified form in Figure 2 and it is clear from this that consonants which have the majority of the speech information are contained in a frequency band between ~1 KHz and 8 KHz. On the other hand the bulk of speech energy is in the vowels which are contained in the 50Hz-1KHz band (male speech having a fundamental frequency of ~130 Hz and female speech ~220 Hz). System design based only on the speech energy spectrum therefore would be unlikely to give good intelligibility. Although it is the overall frequency response which is of importance, the limitations occur mainly in the loudspeakers and the acoustic transmission paths to the listeners. Distortion generated in the loudspeaker can seriously degrade the system quality and attention must therefore be paid to ensure that both appropriate quality loudspeaker units are used and that they are operated within their linear region of power handling.

Disregarding electrically induced noise, acoustic noise enters mostly in the path between loudspeaker and listener. Noise comes from number of sources which include the general movement of passengers. By far the most important source however is trains. Background noise levels can vary from around 40 dB(A) with no train present up to around 100 dB(A) as a train enters or leaves a platform. A major effect of noise is to mask consonants and since these carry the majority of the information, their loss will cloud the intelligibility of the announcement.

The effect of reverberation has a strong deleterious effect upon intelligibility. Fig 3 provides an example of this. It shows the sound energy decay curve at a deep tube station having an RT_{60} = 2.5s. It can be clearly seen that the Early/Late energy ratio is low and that there are significant reflections which arrive at the listener more than ~50 ms after the direct sound from the loudspeaker. These late arrivals will have a detrimental effect upon the direct and early reflections and will act to reduce the intelligibility of speech.

Many of the strong late reflections result directly from the symmetric cylindrical geometry that characterises the majority of the deep tube station platform areas and which, coupled with the use of hard reflecting materials, give rise to very long reverberation times when judged against other architectural spaces of similar volume.

The low ratio of early/late sound energy illustrated in Figure 3 lies at the root of the difficulties in achieving good speech intelligibility from Public Address systems at many of the sites on the Underground. It also clearly indicates the reason why merely increasing the system output level will not improve intelligibility.

SYSTEM SPECIFICATION

For the loudspeaker element in the communication chain, we studied the influence of loudspeaker position on RASTI, within the constraints of where loudspeaker placement would be practical. This work showed that in general a reduction in speaker separation improved intelligibility and as a result for the deep tube station platform areas a 5 metre spacing was eventually chosen as being the best compromise between performance, cost and aesthetics, with the additional recommendation that the loudspeaker should be sited on the tunnel wall as close to passengers' head height as possible.

In this way the ratio of early to late sound energy was increased without at this stage any modification to the station's acoustic environment.

The effect of background noise on intelligibility was also investigated. Other sources [1] have shown that in measuring STI or RASTI, degradation of intelligibility commences when the S/N ratio falls below 15 dB. Our results showed that for the deep tube stations a reduction in RASTI only became apparent when the S/N was < 10 dB. This is understandable bearing in mind the high reverberation levels encountered at these sites which give rise to an equivalent S/N of close to 0 dB. The influence of the background noise therefore is of less significance.

With the loudspeaker positioning and sound output level defined, it has been possible to specify a RASTI value for the different types of station platform environments that exist in London Underground. At present, where these modifications have been tested, RASTI values range from 0.4 - 0.8 but the work has identified two distinct areas where there is potential for improving the conditions such that a system performance standard of RASTI ≥ 0.7 could be expected to be achieved even for the most difficult environments.

The first of these is the control of the acoustic conditions prevailing at many sites by reducing the reverberation time. Work is currently being planned to investigate the acoustic properties of the deep tube platform areas in order to establish the optimum type and positioning of absorbent material in order to reduce effectively the reverberant acoustic field. This will have a twofold benefit of improving speech intelligibility as well as reducing background noise levels to which both passengers and staff are subjected.

The second area that has been shown to hold potential for improvement is the design of loudspeakers having characteristics that are more appropriate to the specific application in London Underground. The loudspeakers used in the experimental work were deliberately chosen as the same units that had to be installed in all the deep tube stations to satisfy recent safety requirements following the Kings Cross disaster in November 1987. The timescale of that installation programme prohibited any major development of loudspeakers having specifically designed directional characteristics but this is a matter which is now being pursued. By using such loudspeakers, sound energy over the frequency range encountered in speech may be more precisely directed towards the passengers rather than into the general space of the station where it only serves to increase the reverberant sound field. Furthermore, speakers of this type will find application at surface stations where currently problems are encountered with acoustic spill from platform areas to residential property neighbouring the station.

It is considered that an upper limit on Public Address sound level should be set at ~ 90 dB(A). As improvements are made to reduce reverberation at stations, background noise will become the principal factor limiting STI. Setting the requirement for a 15 dB S/N headroom implies therefore that intelligibility will begin to degrade when background noise levels exceed 75 dB(A). The conclusion drawn from this is that in deep tube stations, Public Address cannot be effective when trains are entering or leaving platforms which as mentioned earlier can generate noise levels exceeding 100dB(A). To some extent, the optimum headroom over background noise can be catered for by the use of an automatic level control system built into the Public Address. However this of course will only be effective up to the upper limit of 90 dB(A).

PERFORMANCE SPECIFICATION

A System Performance Specification incorporating objectively measurable parameters was seen as a central aim of the work. To date we have been able to define a number of elements. These are

- a defined method measuring background noise.
- a minimum but achievable value for RASTI (0.4 for deep tube, 0.8 for open air stations.
- an operating Public Address sound output level that will enable the RASTI value to be achieved.
- a stipulation of loudspeaker placement, particularly in respect of the deep tube environments.
- Limit values on frequency response and distortion.

These elements have been used not only in the initial set-up of Public Address systems but also in specifying the routine maintenance.

As an adjunct to the work, we canvassed other rapid transit systems operators in Europe and North America for their response to the challenge of obtaining intelligible Public Address.

The response we received indicated the awareness of the need to specify objective performance standards for Public Address systems particularly with regard to intelligibility and frequency response.

SUMMARY

The work that has been carried out in investigating the factors influencing sound quality performance of Public Address in London Underground has established that:

- RASTI provides good objective indicator of sound quality performance provided that the reverberation time is below ~2 seconds. At greater reverberation times its correlation with perceived intelligibility is less good but nevertheless it provided a satisfactory means for monitoring changes in system performance.

- The ratio of direct to reverberant sound energy is improved by using more loudspeakers and decreasing the separation between them.

- Given the station environments as they exist at present and the Public Address equipment provided, loudspeaker separation of ~5 metres for deep tube gave the best balance between performance, cost and aesthetics.

- Critical to the clear transmission of speech to the listener are clarity of consonants and allied with this a Public Address sound level that is 10-15 dB above ambient noise levels.

- Without acoustic treatment of some underground stations to reduce reverberation time and thereby ambient noise levels, Public Address intelligibility cannot be significantly improved at these sites.

- There are positive benefits to be gained from the use of more directional loudspeakers both in improving intelligibility in deep tube stations and in reducing the risk of acoustic spill from open stations in environmentally sensitive areas.

- Last but in many ways the most crucial point is that the total performance quality of a Public Address broadcast will depend ultimately on the quality of the input. Within the foreseeable future many important announcements will be made with human announcers. Appropriate selection, training and guidance is required. Both the message construction and its delivery are therefore of great importance.

REFERENCE

1. Houtgast,T et al., "Predicting speech intelligibility in rooms from the modulation transfer function". Acoustica 46 (1980) pp 60 - 72.

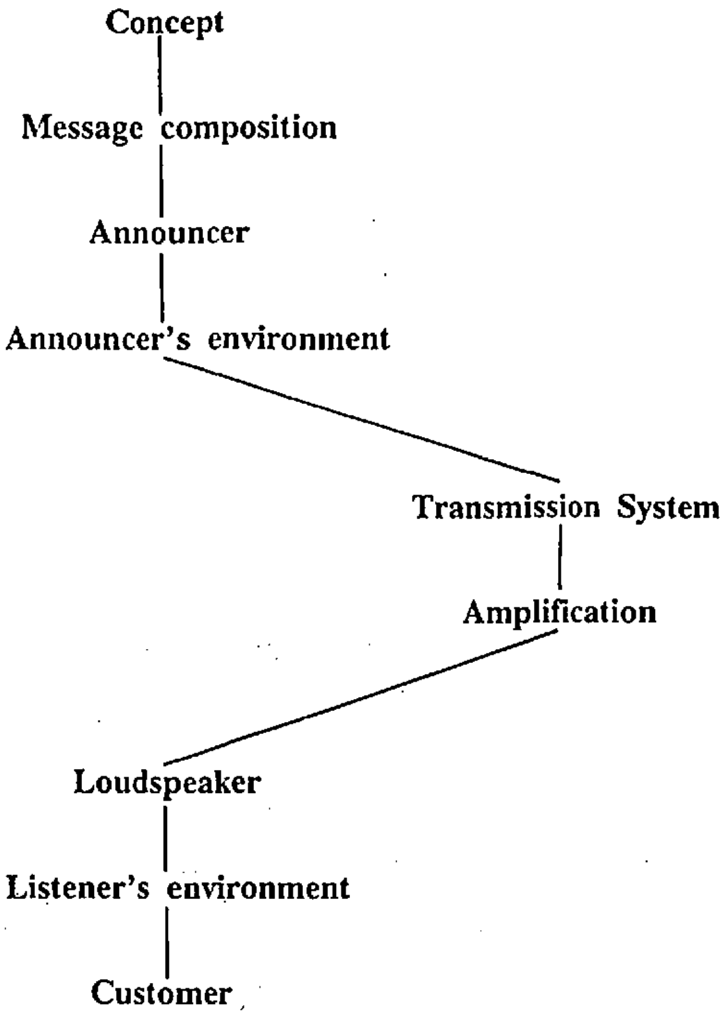

FIGURE 1

Communication Chain

Frequency (Hz)	50	500	1000	8000
Speech Energy		60%	35%	5%
Speech Information		5%	35%	60%
		vowels	consonants	

FIGURE 2

Speech Energy/Information vs. Frequency

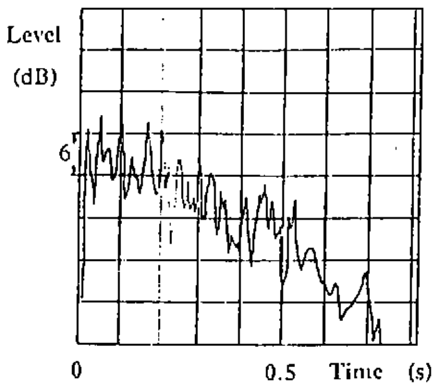

FIGURE 3

Sound Energy Decay curve for a deep tube platform area