1. Intelligibility and the PA system

The PA system

According to Peter Mapp, Designing for speech intelligibility, Handbook For Sound Engineers (2002): The fundamental reason for a speech reinforcement system; "is to deliver intelligible speech to the listener". He goes on to state that "A surprising number of systems however fail to achieve this basic goal. There can be many reasons for this, ranging from inadequate signal to noise ratio to poor room acoustics or inappropriate choice or location of loudspeaker". All of these factors will need to be considered when analysing the current PA system at Kedleston Road Atrium and the installation method of the proposed beam steering system.

Before examining the factors which affect speech intelligibility it is important to recognise what is meant by "intelligible speech". Just because somebody can be heard speaking does not mean that what they are saying is intelligible. P Mapp (2002) states that "a common mistake, often made when discussing intelligibility, is to confuse audibility with clarity". This basically means that just because something is heard does not mean it is understood, according to Dictionary.com (2011), Intelligibility literally means "capability of being understood". When considering what Public Address systems are used for it easy to see why intelligibility is so important. PA systems are used in a wide variety of places but they are all used as a medium for information, although there main purpose may vary it is common place to use a PA system in case of emergencies such as fire. Some PA systems may just broadcast an alarm but others broadcast information on what to do or where to go. If a PA system were to be used in such way it would be fair to assume that sufficient testing has taken place to make sure the system is intelligible. A typical example of poor intelligibility to which most people can relate is the infamous train station PA; Information about platform changes or delays is regularly announced over PA systems in train stations and is usually met by a chorus of "what did that just say?", this can be infuriating for most travellers let alone those who are hard of hearing. Obviously train stations have quite high noise floors which contribute towards their poor intelligibility, wordinfo.info (2011) gives the SPL of a speeding train to be 100 decibels, when compared to the value of 60 decibels which it gives for conversation at one metre it is easy to understand how audio clarity can be impaired.

However, the excuse of speeding express trains or even stationary diesel engines is not always available; According to cambridge-news.co.uk (2011); Cambridge train station is the third worst for announcements with "one in ten Cambridge station users saying they have frequently missed changes to train times and cancellations because they could not hear the announcements properly". The most interesting aspect of this article though is that "The results also reveal that 50 percent of respondents blamed a muffled voice and 20 percent the echo within the station for not being able to hear the service". That is a figure of 70 percent which effectively blame the PA system and not the noise inside the train station.

The examples of muffled voice and echo are an excellent jumping point for examining what the differences might be between an intelligible and a unintelligible PA system. P Mapp (2002) states that "it is quite possible to have a poor sounding system that is highly intelligible(limiting the frequency response) or alternatively a high quality system that is virtually unintelligible (Hi-Fi loudspeakers in an aircraft hanger)"
In order to fully understand how a voice can sound muffled it is important to identify what makes up a voice, or in other words, what makes up speech? All sounds have a frequency and according to Choudhary, hypertextbook.com (2004) the range of human hearing is argued to range from between 15 to 20Hz and 18 to 20kHz. Using the widest range possible as an example it can be determined that the frequency of human speech occurs between 15Hz and 20kHz, specifically though P Mapp (2002) states that "speech covers the frequency range from approximately 100Hz to 8kHz".

Figure 1 is the graph used by P Mapp (2002) to visually demonstrate the average speech spectrum and the different levels that each octave band contribute. The graph shows that most of the sound in regards to SPL occurs in the lower frequency which according to P Mapp (2002) correspond with the vowel sounds. This does not mean that intelligibility depends on these sounds, as stated previously it is important not to confuse audibility with clarity. In this case, just because you can hear the vowels does not mean you can understand the word.

Figure 2 is the graph used to illustrate the different octave bands contribution to intelligibility and actually show that 75% of what makes a word intelligible occurs between 1 kHz and 4kHz which correspond to the consonants. These sounds are made by the lips and tongue, it is these sounds that define syllables and shape words. This is also stated in Beranek, Speech Intelligibility (1954) who writes "The vowel sounds are not as critical to speech intelligibility as the consonant sounds. It is unfortunate that the consonant sounds are so weak and, therefore, are easily masked by noise."

Although Beranek (1954) and Mapp (2002) Agree that intelligible sound is dependent on the mid to high frequencies, Mapp (2002) does not claim that the "Vowels are not as critical" and instead focuses his research on explaining why the consonants are so critical. This could be because neither side of the spectrum (low frequencies on one side, mid to high on the other) could provide intelligibility alone, especially when considering that the vowels are responsible for the majority of a words energy. Beranek(1954) finishes his statement by writing "In some languages, such as Hebrew, no vowels are written, only consonants" This may be true but according to jewfaq.org (2011) the Hebrew language does have "aids for pronunciation", these are a system of dots and dashes called "nikkud", these nikkud are used to "indicate vowels" and for all intents and purposes are the equivalent of a vowel written in English. With this information it could be argued that Barenek's statement is incorrect and when considering speech intelligibility, vowels and consonants are just as critical as each other, especially if true intelligibility is to be achieved.
Another point which Beranek (1954) makes is that the consonant sounds are easily masked by noise, this point can be compared to figure 1 where we can see that this true and indeed, quite severe, with a roll off, of around 6 dB per octave from 1kHz upwards, which yields a difference of around 25dB between 250Hz and 8kHz.

According to meyersound.com (2011): "High-quality speech systems need to cover the frequency range of about 80 Hz (for especially deep male voices) to about 10 kHz (for best reproduction of consonants, which are crucial to intelligibility). Response below 80 Hz must be eliminated to the extent possible: not only do these frequencies fall below the range of the speech signal, but also they will cause particularly destructive masking at high sound levels.

It’s important, also, for the system response to be reasonably flat throughout its range. The gradual high-frequency roll off that many reinforcement professionals favour for music applications will tend to de-emphasize consonants, which are already as much as 27 dB less loud than vowels. Likewise, prominent peaks or dips in the response can cause either self-masking or loss of consonant articulation."

Although this information is critical to the success of this investigation it is important to remember that having the correct frequency range and adequate settings on an equaliser are redundant if the front end of said system is inadequate. P Mapp (2002) states "Many potentially adequate sound systems are often let down by employing a cheap or restricted bandwidth microphone at the front end of the system. In the author’s experience, even on a basic paging system employing restricted bandwidth speakers, e.g., re-entrant horns, the difference between a microphone with a reasonably wide and well controlled frequency response can always be readily identified over one with a restricted response, even if it exceeds the response of the loudspeakers themselves. Rubbish in equals rubbish out"

**Summery**

By using the data which has been explained in this chapter and baring in mind the research carried out by established professionals in the field of intelligibility, several considerations need to be implemented in order to create a successful PA system and all of them are equally critical. These considerations are defined below;

- **Correct frequency range**: The frequency range defined by Meyer on meyersound.com (2011) for a high quality speech system is 80Hz to 10kHz, this matches the research of Mapp (2002) and Beranek (1954) who both say that the makeup of Word spoken in English is covered within this range.

- **Flat frequency response**: meyersound.com (2011) also states that "it’s important, also, for the system response to be reasonably flat throughout its range". As explained previously and illustrated in figure 1, the higher frequency's roll off at around 6dB per octave from 1kHz, by using an equaliser it is possible to reduce the lower frequencies and increase the higher ones'.

- **Good enough “front end”**: As stated by Mapp (2002) the front end is as important as the rest of the system, the microphone used on any PA system should, at the very least, match the speakers frequency range. The frequency response of the microphone should also be as flat as possible within the desired range, compromising any aspect of the system will only hindrance its overall performance, if excellence is to be achieved then excellence is to be installed.
References from chapter 1