Intelligibility of TV and Cinema Sound

Peter Mapp Peter Mapp Associates, Colchester, UK.

peter@peternapp.com

ABSTRACT

In recent years there has been a significant increase in the number of complaints concerning the intelligibility of both Film sound tracks and TV productions – both in Europe and the USA. The paper reviews the background to dialogue intelligibility and looks at a number of mechanisms that may be responsible for the growing trend of dissatisfaction. The transmission chain is reviewed and new measurements and data concerning domestic listening conditions, as well as some cinema data are presented. The results of a pilot measurement program show that the in-situ frequency response of the TV systems, operated by many domestic listeners, is far from ideal with response variations of 10-15 dB being common. Speech Transmission Index (STI), Reverberation Time and Clarity data are presented that suggest that the room acoustic conditions of the domestic listening environment should not, in themselves, significantly degrade the received signal. Sound colouration however may well be a contributory factor.

1 INTRODUCTION

Over the past few years there has been a growing number of complaints concerning the dialogue intelligibility of both TV programmes / and made for TV films as well as mainstream film soundtracks. Examples of UK TV programmes that have received particular criticism and adverse press comment include 'Jamaica Inn' & more recently 'Happy Valley'. Both series have received adverse press comment with respect to the intelligibility of the broadcast dialogue with the BBC having to publicly defend the productions. One reason for the poor intelligibility, given by the programme producers, was the poor set up or quality of the viewers TV sets [1]. Mainstream films that have been criticised for the poor intelligibility of their dialogue include Interstellar, The Councillor, The Social Network and Batman Vs Superman. The paper begins by reviewing the main facets of speech and dialogue intelligibility before discussing potential causes of the current growing trend of listener dissatisfaction. The results of a pilot survey into UK listening room acoustics and the acoustic quality and reproduction of TV sound are reported. The survey, apart from measuring standard acoustic parameters such as room reverberation time, also included measurements of the Speech Transmission Index and Speech Clarity (C7) as well as the Centre Time of the reproduced sound at typical listening positions. The frequency response of the TV system loudspeakers was also measured at typical listener positions. The current trend of locating TV loudspeakers in the base or in the rear of a TV set is explored and the resultant colouration of the sound that occurs as the set is located near to a boundary is reported. The effect of such a loudspeaker location on the transmission of speech is discussed and illustrated. The variation in acoustic conditions, measured in a number of typical listening rooms / environments, is also reported and provides new data and insights into the problem. The TV – listener acoustic transmission chain is shown to be a potentially significant factor and contributor to the current problem. It is shown that the reverberation time in most listening rooms is relatively low (and lower than some previously reported data) enabling high STI and Clarity values to be achieved. The frequency response of the reproduced sound however was found to be highly variable with variations of 10 to 15 dB occurring over the speech range in some rooms / set ups. The effects of typical TV set tone controls, boundary correction filters, equalisers and user settings (e.g. 'clear speech', 'booster' and 'simulated stereo') were also investigated with the measurements providing some unexpected results. Whilst many sets have minimal user control over the quality or tone of the sound, it was found that those that did provide this facility, could indeed result in a loss of subjective speech intelligibility – though finding a suitable measure to illustrate the effect often proved elusive.

2 DIALOGUE INTELLIGIBILITY

Many factors can affect the reproduced intelligibility and perceived clarity of film or TV sound. The factors relate to the three main processes involved, Sound Capture, Processing & Production and Reproduction. However to this physical chain two further factors have to be added (1) The articulation and clarity of the person talking and (2) the hearing acuity of the listener. It is probable that these two latter factors provide the largest variable particularly, as they are not linear entities. The main factors influencing dialogue intelligibility and perceived clarity are listed below in their respective groups.

Sound Capture

- Microphone response & performance
- Location & alignment of microphone
- Microphone directivity & sound rejection
- Presence, level, spectrum & direction of background noise
- Room / set reverberation
- Sound reflections
- Distortion

Processing & Production

- Equalisation & signal processing
- Dynamic signal processing
- De-nosing
- Addition of background music & effects
- Audio formatting, e.g. mono, Stereo, or 5.1 surround etc

Audio Reproduction

- Frequency response of loudspeakers
- Location of loudspeakers
- Calibration of system
- Reverberation time / reflectivity of listening / playback room
- Presence of reflecting surfaces near TV / loudspeaker(s) and listener
- Distance of listener to loudspeakers & axial position
- Background noise level
- Sound level of audio playback
- Distortion (nonlinear) of audio reproduction system

Actor / Talker

- Articulation & Clarity of speech
- Modulation
- Speech Rate
- Accent
- Familiarity / fluency of language
- Voice level

Listener

- Hearing Acuity
- Attention / Alertness (Distraction)
- Familiarity / fluency of language

The above factors and transmission chain are summarised in figure 1.

[It is worth noting here that there is a subtle but significant difference between the meaning of ‘speech intelligibility’ and ‘speech clarity’. Whilst intelligibility can be objectively measured, it is a relatively crude measure – perhaps being distinguishable into just 5 perceptual categories (e.g. Bad, Poor, Fair, Good, Excellent) whereas speech clarity is very much more subjective and capable of finer resolution. For example, whilst some microphones have been described as providing better intelligibility than others (of a similar type) objective measurements cannot discern this, although audibly improved clarity has resulted].

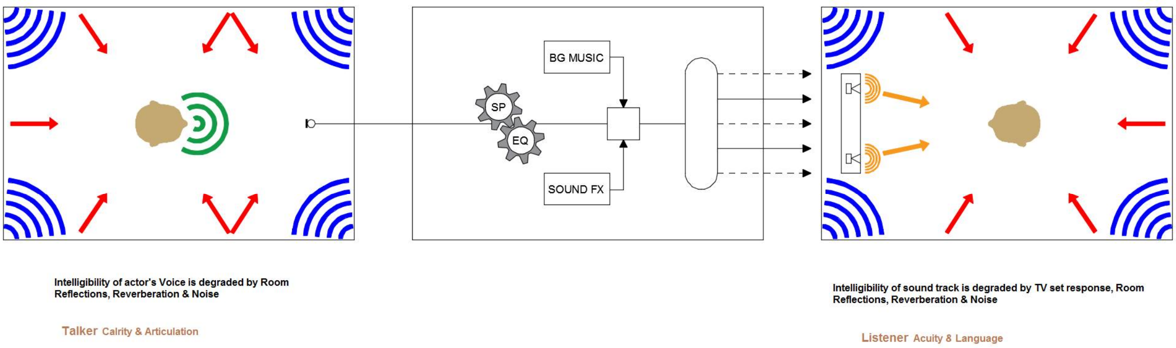

Figure 1 – Sound Reproduction Chain

‘Ease of listening’ or Mean Opinion Scores can also provide greater resolution than current objective intelligibility measures e.g. [2]. Speech Intelligibility is highly non-linear and can for example, under some circumstances, produce little change for a given variation in interference whereas a similar apparent change can bring about a large change in perceived intelligibility under slightly different conditions. (1 dB for example can produce a significant improvement in intelligibility under certain conditions and no effect under others). Figure 2 shows the effect of signal to noise ratio on sentence intelligibility. Whereas at low signal to noise ratios, the effect on intelligibility of a change in SNR is relatively linear, as the intelligibility improves, a similar change in SNR produces a much larger change in perceived intelligibility.

The most widely used objective measure of audio system intelligibility is the Speech Transmission Index (STI). This internationally standardised method has been existence for over 45 years and provides a validated (calibrated) objective measure albeit with a number of limitations [3, 4]. Although no specific guidelines exist for TV / cinema sound intelligibility, based on previous work, a preliminary minimum criterion of 0.70 STI is proposed. This ties in with band ‘B’ in IEC 60268-16, suggested as being suitable for complex speech and unfamiliar words with application to Theatres, Speech Auditoria, Courts and Assistive Listening Systems. Whilst band ‘A’, (equivalent to 0.74 STI) would be better and should be strived for, it was felt that band ‘B’ (0.70) should provide a useful lower limit. (Whilst values > 0.70 may be required for hearing impaired / more elderly viewers, such a value should enable listeners with mild or even moderate hearing loss to readily follow well-articulated dialogue).

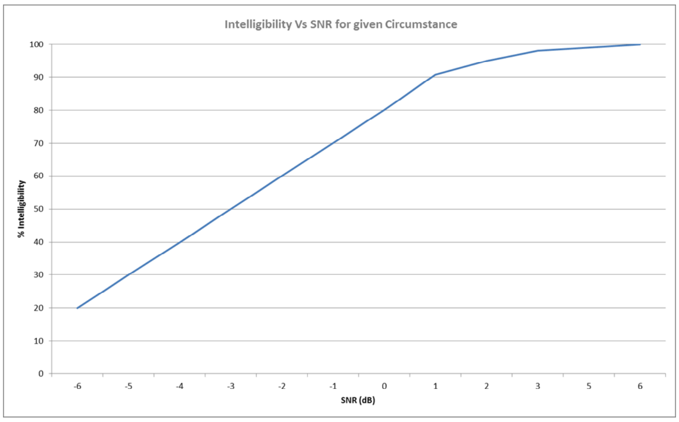

Figure 2 – Non-linear intelligibility behaviour

The topic of dialogue intelligibility is remarkably complex and can also be affected by many non-acoustic factors such as being able to see the person talking (Many viewers subconsciously lip read and use other body language cues to assist with understanding Film and TV dialogue). Due to years of listening experience & training, we are able to recognise sequences of speech sounds pronounced in quite different ways or affected by various, remarkably severe, distortions. Listeners subconsciously review what is being said and predict what they are expecting to hear next. Therefore, thanks in part to this running contextual assessment of the preceding and following speech segments, speech can be considered to be highly redundant. (For example, the most important identifying feature of a word can be masked (e.g. by noise) or even removed, yet the word can often be correctly identified). However, this continual processing of the dialogue is a cognitive load that under poor or degraded conditions can require considerable concentration resulting in listening fatigue and distraction. As a significant proportion of the UK population (around 14%) have a noticeable hearing loss, clear dialogue is required in order for this group of the population as well as those with even a mild hearing loss (defined as a loss of 25 to 40 dB HL) to be able to readily follow what is being said. Analysis of complaints received by the BBC from viewers complaining about poor intelligibility identified the following issues [5, 6].

60% of viewers had some trouble what was being said in TV programmes.

Issues complained about were :

Background Music – 11 %
Background Noise – 13 %
Accents & Dialects – 19 %
Mumbling & Poor Diction – 14%
Talking too fast - 11%

The result of these complaints was that in 2011 the BBC issued new guidelines and information for program producers and editors to produce clearer sound. [7, 8]

It is interesting to note that there is a significant variation in typical listening levels that directly correlates with listener age [9]. Figure 3 shows this relationship, where it can be seen that there is a range of 15-20 dB in preferred / typical listening level. Although not tested in the survey, this suggests that harmonic or non-linear distortion could become an additional factor as the loudspeakers in most modern TV sets are not capable of cleanly delivering such high SPLs.

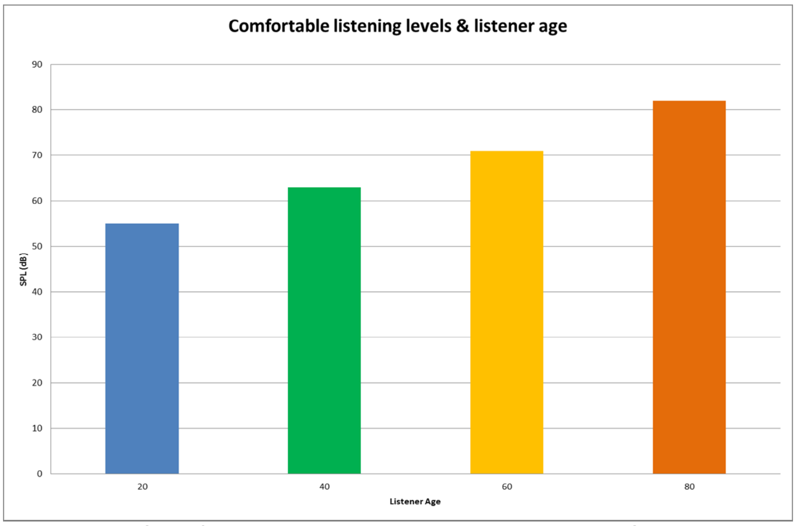

Figure 3 – Comfortable listening levels based on Coren [9]

3 TV – ROOM MEASUREMENTS

The in-room TV performance measurements were conducted in order to assess the potential contribution that this part of the reproduction chain may have on the loss of intelligibility perceived and complained about by listeners. The performance of 12 different TV sets in 12 different rooms was measured together with the prevailing acoustic conditions. The rooms were typical British Living Rooms. The average volume was 38.5m^3, with a range of 24 to 144 m^3.

3.1 Reverberation Time & Temporal Measurements

Figure 4 summarizes the reverberation time characteristics for 10 of the rooms measured. As can be seen, the majority of the rooms exhibit reverberation times of between 0.3 and 0.4 seconds, though there are two outliers. The upper curve is for a larger than average room with a preponderance of hard surfaces that has an average mid frequency reverberation time of 0.90 seconds. The lower curve is the author's acoustically treated listening room which exhibits an average Tmf of 0.13 seconds. Excluding the larger room, the average mid frequency reverberation time for the sample is 0.26 seconds. (Whilst it is debatable if the concept of reverberation time is applicable to a small, dead room, it is still a useful measure of the room's decay time and sound absorption properties).

Whilst there is a paucity of measured listening room reverberation time data, the above values are significantly lower than has previously been reported in the literature [10-12] as shown in figure 5.

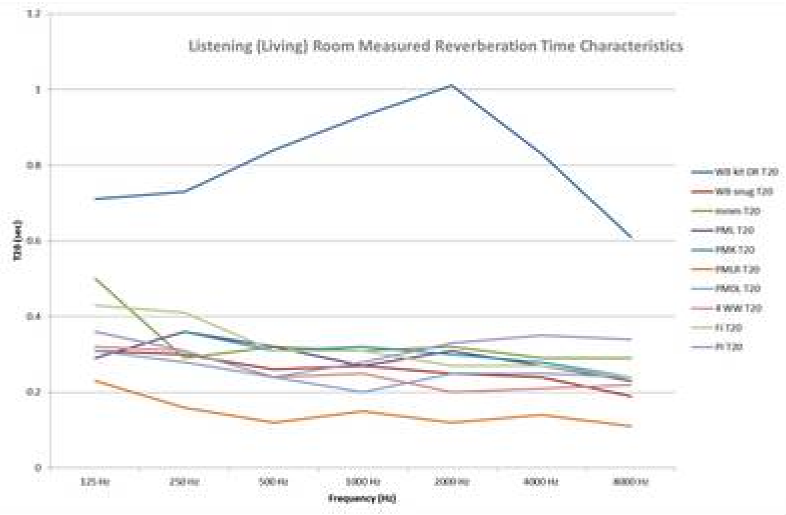

Figure 4 – Measured Reverberation Time data

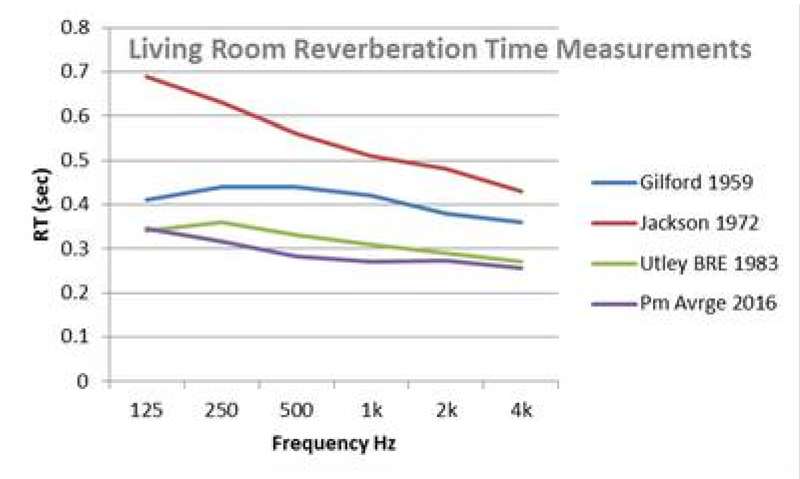

Figure 5 – Previous British Living Room RT data

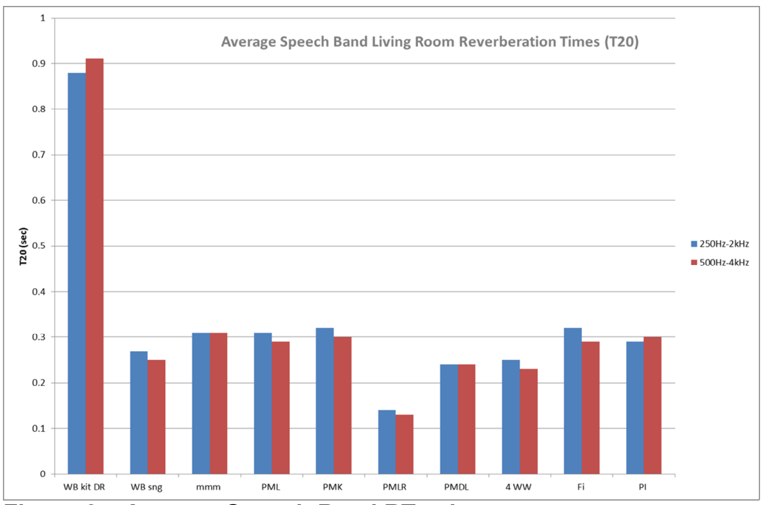

Figure 6 – Average Speech Band RT values

With such low values of reverberation time it can be expected that intelligibility should be high. (After all this is one's normal experience of speaking and listening to speech in a living room).

The reverberation time measurements were made using the TV set loudspeakers in each room and were derived from the measured in-room impulse responses. This technique also allowed the system frequency response and other measures to be obtained. Three measures related to speech intelligibility were computed from the impulse responses. These were the STI (Sound Transmission Index), Clarity Index (C7, C35 and C50) and also the Center Time. Whereas the STI and C50 measures are well known measures in room and sound system acoustics, Center Time is less well known but provides a useful adjunct to the clarity measures enabling the time to be established at which the direct and reflected energy within the room are equal or balanced.

The STI measurements were computed with male speech weighting factors and direct STIPA measurements also made. As expected from the reverberation time results, all the STI values were high, the average being 0.83 for the normal sized / furnished living rooms. Figure 7 summarizes the results of the STI measurements. As can be seen, all the rooms bar one achieve the nominal 0.70 STI criterion. Indeed these also all achieve the higher 0.75 target.

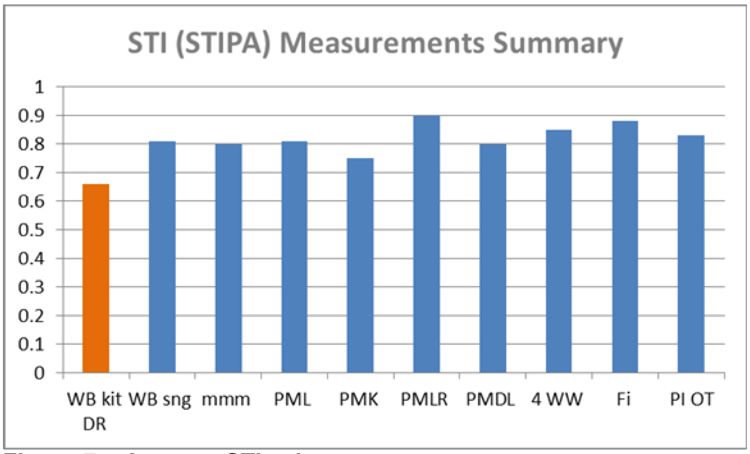

Figure 7 – Average STI values

The STI measurements were made at typical viewing seat positions, which typically varied from 2.5 to 3.5m distance from the TV. In all cases the background noise was low (< 35 dBA) and so the reduction in STI is purely due to reflections and room reverberation and loudspeaker / electronic losses.

The Clarity measures align closely with the STI values. Figure 8 shows a typical set of measurements that include frequency dependent data for C7, C35, C50 and C80.

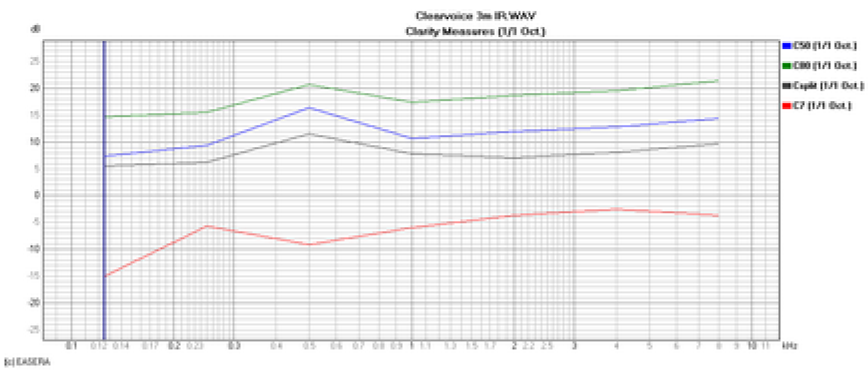

Figure 8 – Typical set of Clarity (C) data

There is no calibrated scale for these measurements. C50 was originally invented for the assessment of concert halls and large auditoria where there would be a significant Initial Time Gap (ITG) and so is not really applicable to small rooms. However, the C7 and C35 provide useful insights. The C7 compares the sound energy arriving in the first 7 milliseconds to the remaining energy (and is often considered to be a measure of the direct sound) whilst conversely the C35 does this for the first 35 milliseconds – which is often related to intelligibility and the integration time of the ear. Figure 9 summaries the Clarity measurements for the various rooms. The values shown are the averages taken over the main speech intelligibility range of 500Hz to 4 kHz.

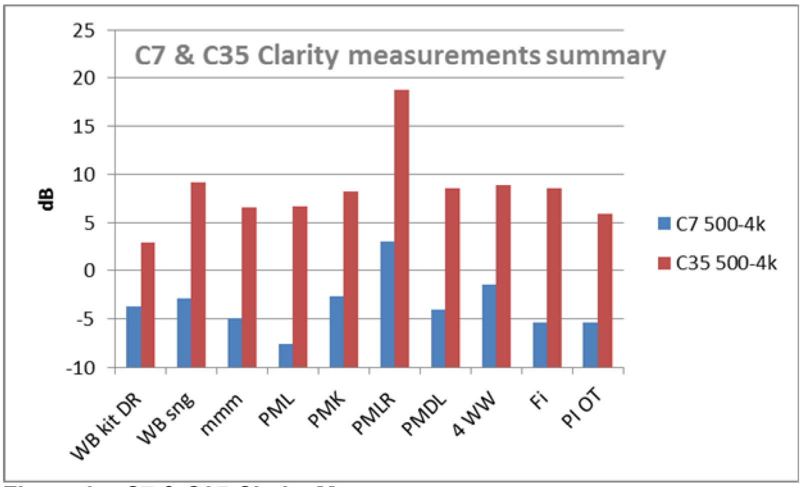

Figure 9 – C7 & C35 Clarity Measurements

In all cases, except the PM Listening Room, the C7 is negative, meaning that more energy arrives after 7mS than within this initial period. As would be expected for a small room, the C35 is positive. The averaged values for the group of rooms are -4.3dB and +7.8 dB for the C7 and C35 measurements respectively. Clearly somewhere between 7 and 35 milliseconds, the energy ratio turns from negative to positive. Where this happens can be investigated by calculating the Center Time. This value shows where the cumulative sound energy arrivals balance. Again this is highly frequency dependent but for comparison purposes, the values averaged over the range of 500Hz to 4 kHz are presented (fig 10). Here the lower values indicate the reduced involvement of the room to the perceived sound. The pattern is therefore similar to the reverberation time data shown in figure 4. The graph also shows the difference that occurs when looking at the lower speech band of 250Hz to 2 kHz which is therefore nominally dominated by the vowels as opposed to the consonants in the higher frequency band. The average center time, for all the rooms (excluding the reverberant outlier), is 18-20ms.

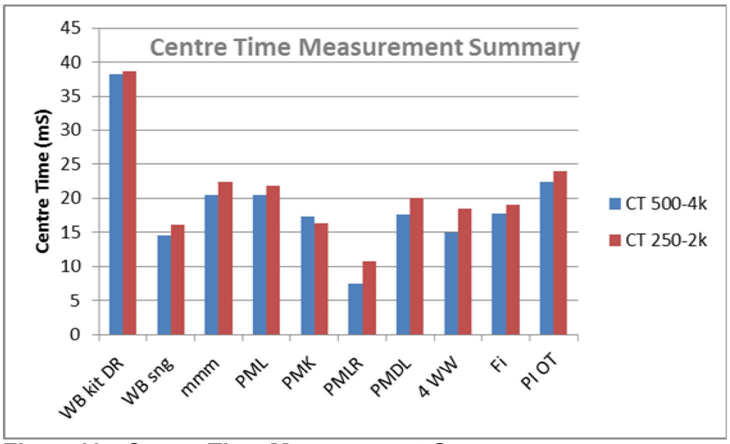

Figure10 - Center Time Measurement Summary

From the reverberation time and room acoustic intelligibility measurements, it is clear that there should not be a problem in perceiving highly intelligible sound in most living rooms. The cause of the reported intelligibility problems therefore must lie elsewhere.

3.2 Frequency Response Measurements

From anecdotal evidence, as well as objective assessment, it is clear that the frequency response and relative frequency balance of a loudspeaker system can have a significant impact on perceived intelligibility. STI is well documented as being flawed in this respect as it does not adequately take account of frequency response imbalances [3, 13]. In-room frequency response measurements were made at typical seating positions for each TV set up. Some TV sets were found to have a range of controls that affected the overall frequency response. These ranged from simple tone controls to octave band equalizers and 'speech enhancers'. Some operating modes were also found to affect the perceived intelligibility (e.g. 'Movie' mode on one set, 'simulated stereo' on another). Three basic loudspeaker formats / positions were found. (1) Front facing loudspeaker (2) Downward facing loudspeaker (3) Rear / upwards facing loudspeaker. Figure 11 illustrates these arrangements.

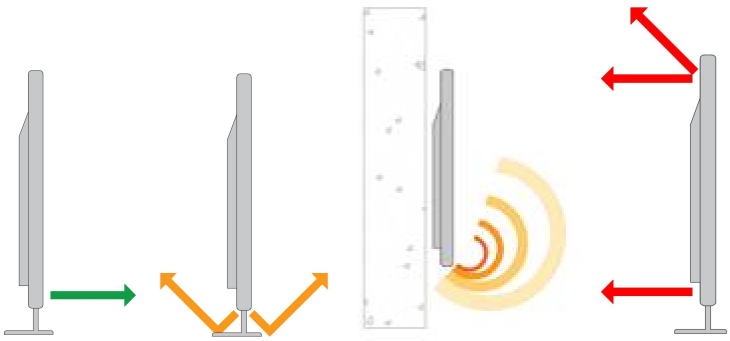

Figure 11- TV loudspeaker positions

Frequency response measurements were made using both a standard Real Time Analyzer and time selective (windowed) impulse response measurements. A selection of typical 1/3 octave TV-Room response measurements is shown in figures 12 - 17.

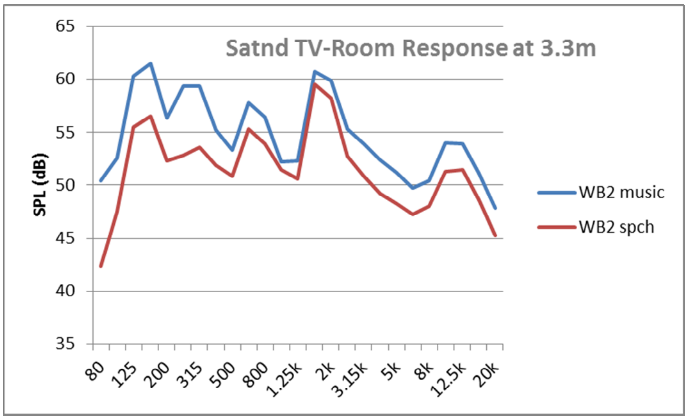

Figure 12 - stand-mounted TV with speakers under

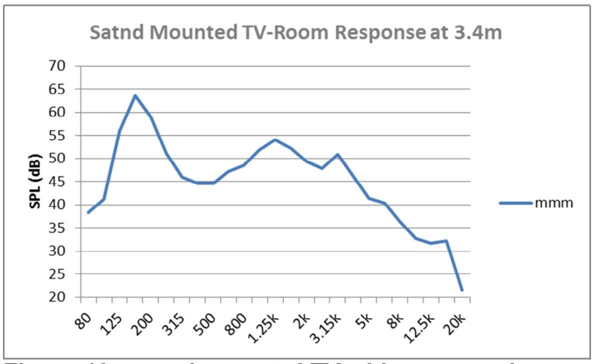

Figure 13 - stand mounted TV with rear speakers

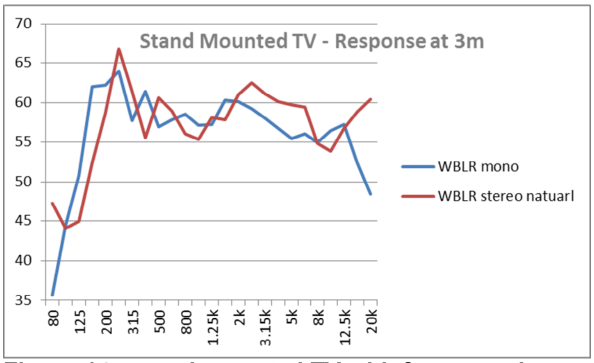

Figure 14 - stand mounted TV with front speakers

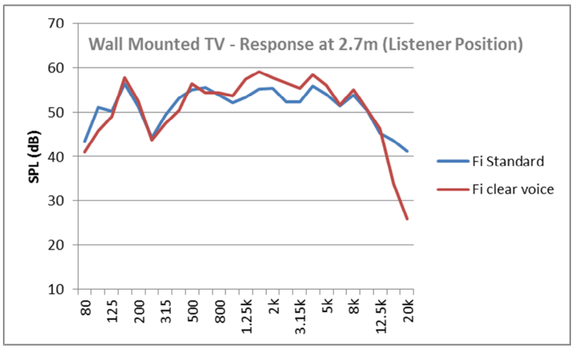

Figure 14 - Wall mounted TV with speakers under

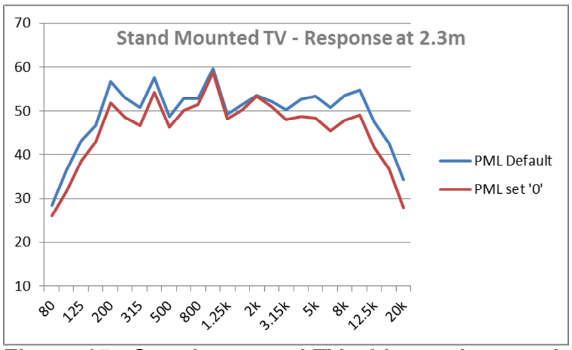

Figure 15 - Stand mounted TV with speakers under

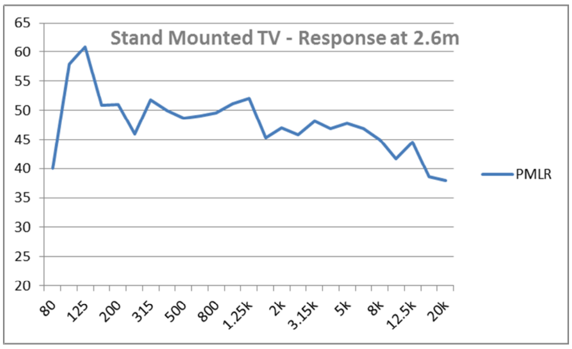

Figure 16 - Stand mounted TV with speakers under

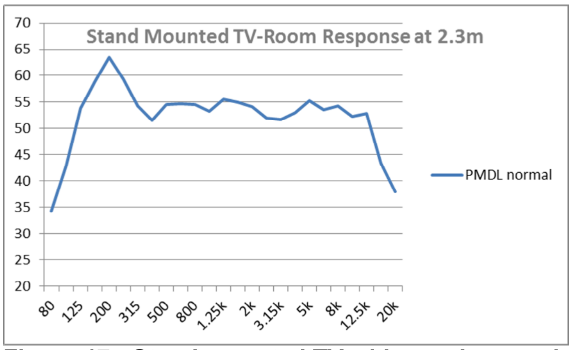

Figure 17 - Stand mounted TV with speakers under

Most current sets have a number of set-up options varying from a speech setting, movie setting, stereo, booster and boundary distance correction etc.

Figures 18 - 22 show how these settings can affect the response at the listening position.

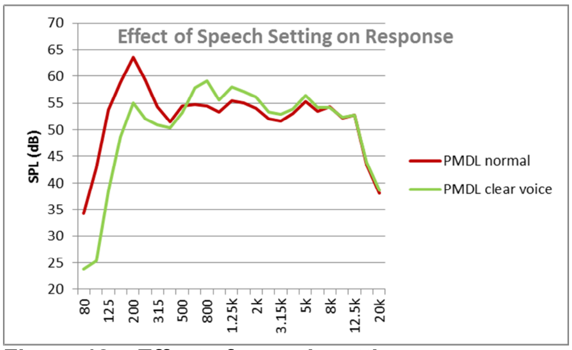

Figure 18 – Effect of speech setting

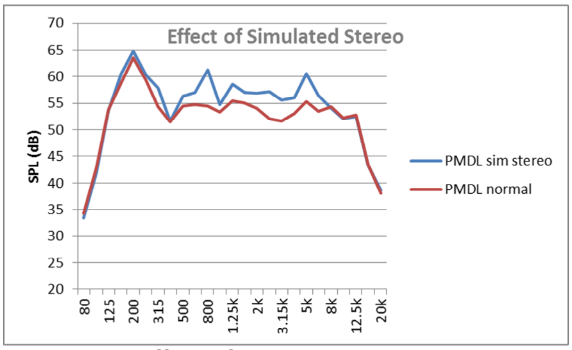

Figure 19 – Effect of simulated stereo setting

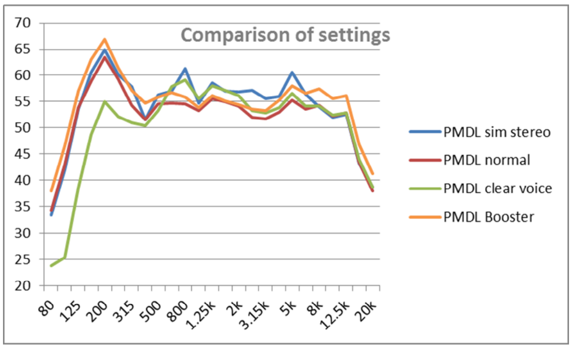

Figure 20 – Comparison of various setting

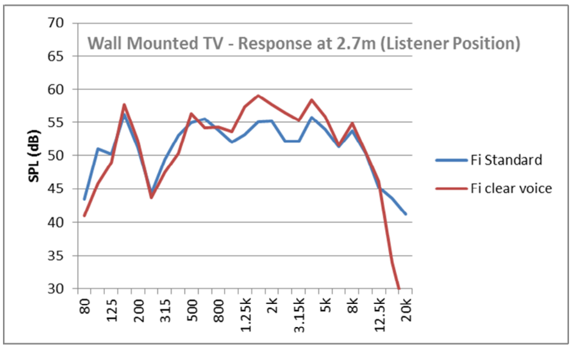

Figure 21 – Effect of speech setting

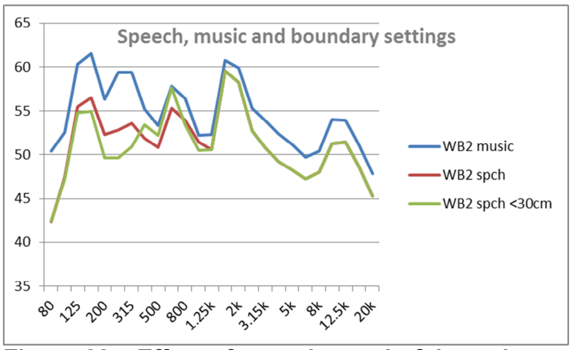

Figure 22 – Effect of speech, music & boundary correction settings

As can be seen form the frequency response plots, a wide range of responses can be obtained, many of which are not conducive to facilitating speech clarity & intelligibility. In many cases, the TV loudspeaker positioning combines with room mode excitation to produce strong, low and mid frequency colorations. These can be clearly seen in figures 12, 13, 14, 16, 17, 18 & 19. Figure 20 shows the response of the TV – Room effect shown in figure 13 but this time produced from an impulse response rather than steady state pink noise measurement.

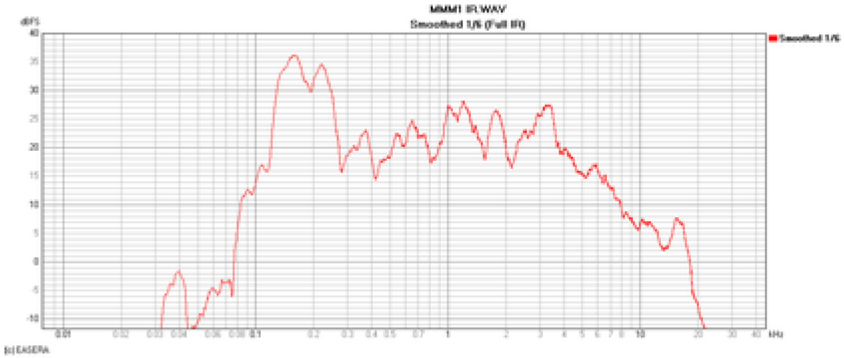

Figure 23 – Frequency response derived from impulse measurement (see figure 13).

As can be seen from figures 18, 21 & 22, employing the speech setting reduces the low frequency ‘boom’ and adds some useful boost for the consonant frequencies. Many of the responses however remain far from smooth. The popular method of mounting the loudspeaker below the screen and therefore just above a reflecting surface when adopting a stand / pedestal mounted position can give rise to strong comb filtering and resonances, particularly if the TV set is mounted near a corner or room boundary – as shown in figure 15 and with figure 24 being a higher resolution plot of this.

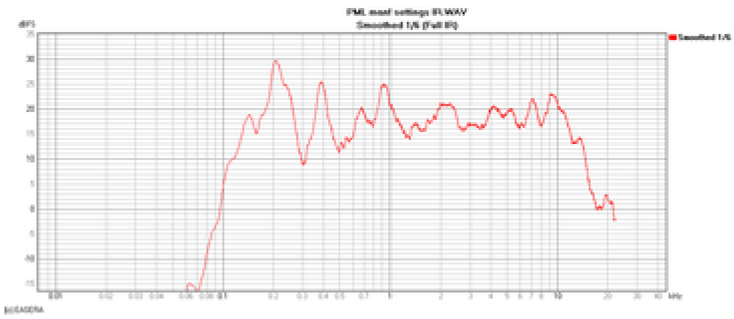

Figure 24 – Comb filtering & resonance effects

3.3 Comment on Room Acoustic and Frequency Response Measurements

At the current time there is no objective / electroacoustic measure that can be used to assess the effect of distortions of a system’s frequency response – though it has been well established that this can have a significant effect on perceived speech intelligibility and clarity. The STI method, whilst accounting for gross variations in response has been shown to be insensitive to smaller response changes that have been proven to reduce intelligibility and increase listener fatigue. [13,3]. The mean value of 0.83 STI for the TV/ Rooms, whilst good, means that there is not too much room for maneuver, particularly when additional signal chain losses are taken into account. From experience in many venues, complaints can be expected if the STI falls below about 0.65 and certainly if below 0.60. (This assumes a good talker and normal hearing listeners). This only leaves about 0.2 STI in hand. [Note an STI value of ~0.75 would be required for older listeners or listeners with mild hearing loss].

The measured electroacoustic losses and do not take account of talker (actor) articulation. If, for example, the actor is a non-native talker or has a strong accent or regional dialect, then a penalty of 0.1 to 0.15 STI needs to be applied [4]. Equally, if the listener is listening in a non-native language, a similar penalty applies. This would then take the cumulative losses / resultant STI to below 0.60

STI. Music soundtracks or sound effects embedded in the transmission can also be disturbing and as noted earlier are source of considerable complaint. Work by Ben Shirley as well as other previous BBC studies have examined these effects and have given rise to the Clean Sound Project and allied research [14]. Whereas no programme maker would deliberately set out to produce material with poor dialogue intelligibility, it may well be that they are not aware of an intelligibility issue in that :

- They will be listening to the program on high quality loudspeakers that have been set up in a good acoustic listening environment.
- They will have the script in front of them and will be very familiar with the dialogue so it would be almost impossible for them not to be able to correctly hear what is being said.
- Research has shown that musicians & conductors are better at picking out and deciphering speech in noisy environments as they are trained to listen and distinguish one sound (instrument) from another [15].

4 TALKER ARTICULATION

As noted above, not only can strong accents or non-native speech have a significant effect on perceived intelligibility but natural talker articulation varies enormously. Peutz for example quantified this effect in terms of % ALCons [16]. Converting this data into equivalent STI values suggests that talker articulation alone can account for a loss of 0.13 to 0.52 STI. Adding the talker losses to the cumulative losses that may occur in the electroacoustic chain, shows that the resultant intelligibility might vary from around 0.60 down to 0.1 STI. Voice articulation also varies significantly with the prevailing acoustic conditions, such that some voices judged to be good under good (e.g. quiet or non-reverberant conditions), become degraded more readily than others when the conditions deteriorate [17].

5 CINEMA DIALOGUE

Although cinema dialogue shares many similarities with television dialogue, it also exhibits a number of significant potential differences that should potentially reduce the problem (e.g. the talker is generally on screen and the dialogue is normally emitted by the centre channel with music and effects allocated to the Left, Right and Surround channels thus potentially providing some spatial release of masking). Against this, the listener cannot adjust the overall sound level and many complain that many sound tracks are played back too loudly).

Whilst it is not possible in this paper to discuss cinema sound and dialogue in any detail it was surprising to the author to find out how poor many cinema systems are and the almost total lack of electroacoustic design that apparently goes into them. This latter aspect was highlighted when the author contacted two leading manufacturers of cinema loudspeakers and asked them for polar plots and predictive data files. One replied that they didn’t have these and have never been asked for them before and the other eventually provided some limited data after a delay of several weeks and stated that they had never provided these before !.

A brief examination of a particular issue can be seen in figure 25. Here diagrammatically, the coverage of the centre channel in a typical multiscreen venue is shown. Clearly the front side seating areas are not properly covered. An acoustic computer model of the venue was built (Fig 26) and a number of scenarios investigated.

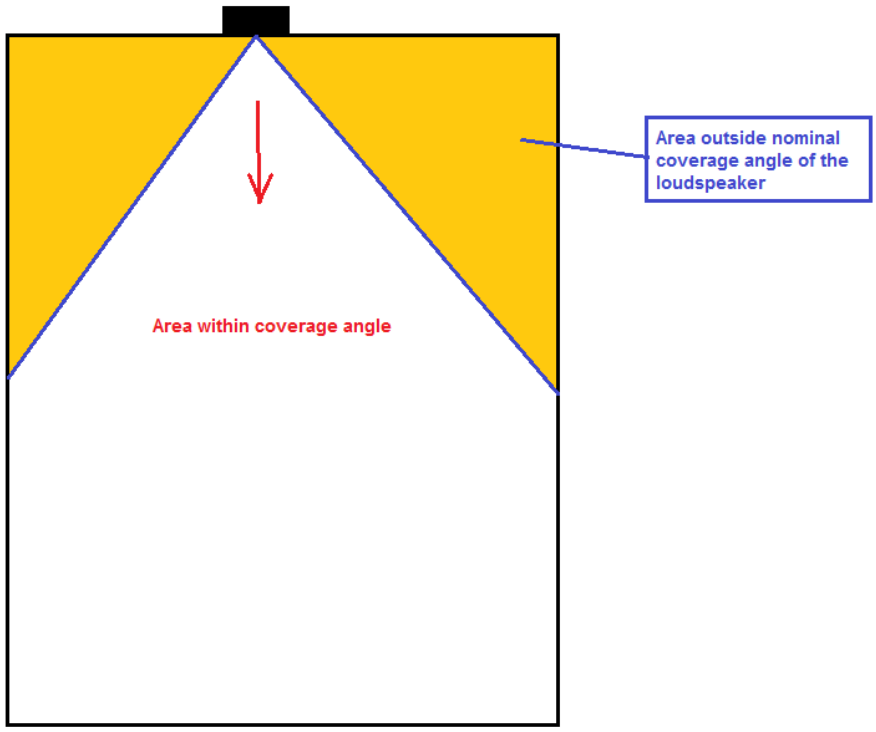

Figure 25 – Idealised coverage by centre channel in typical cinema / movie theatre

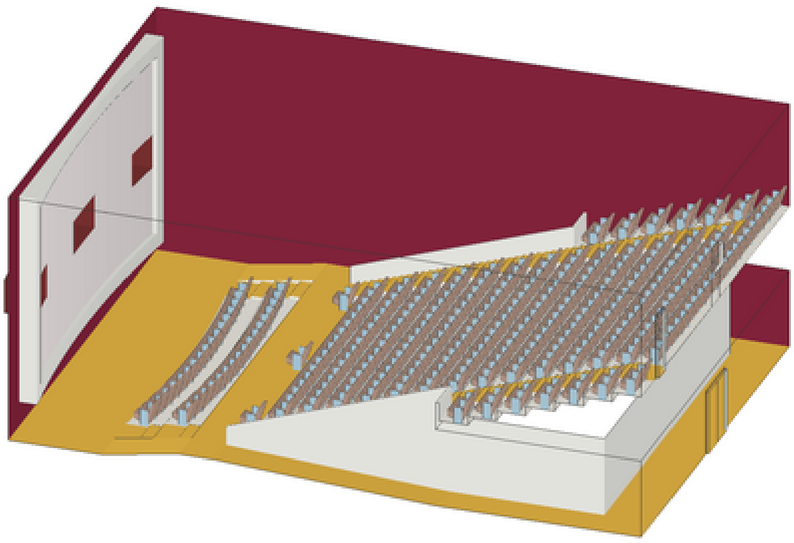

Figure 26 - Computer model of typical movie theatre (20m x12m x8m)

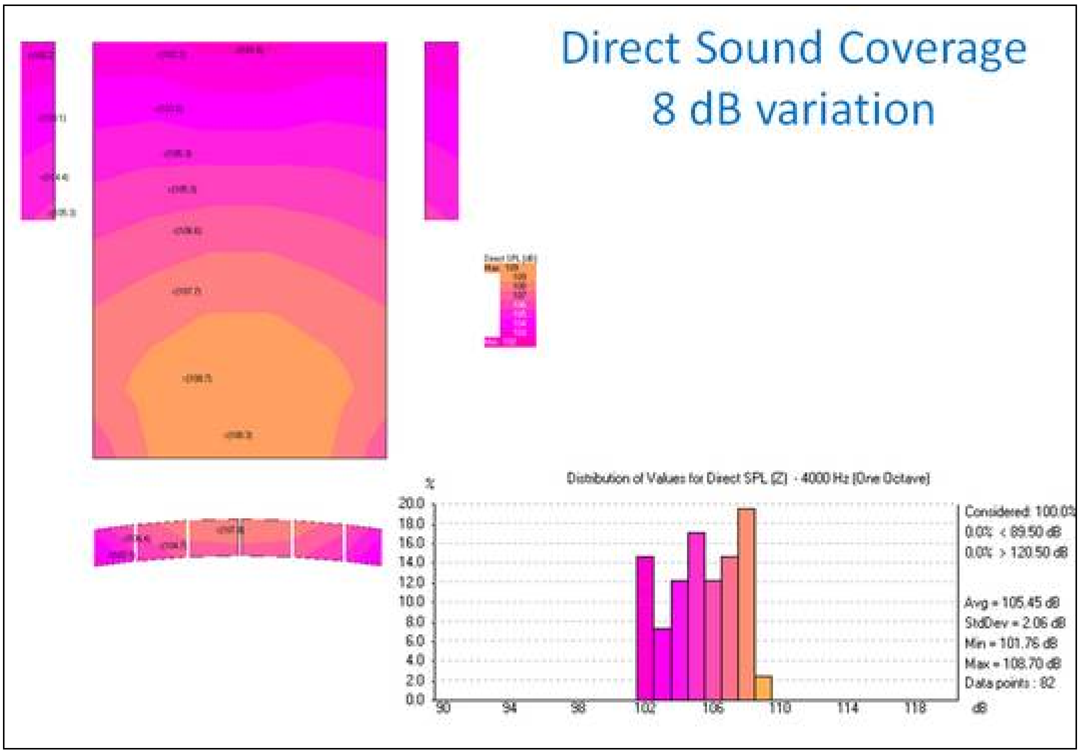

Figure 27 – Predicted Direct sound coverage for a typical venue set up as per SMPTE / Dolby requirements. Note the 8 dB variation in coverage from this state of the art loudspeaker.

From figure 27 it can be seen that there is a variation in the direct sound high frequency coverage of around 8 dB over the audience seating. This is significantly more than would be the case for a modern, well designed sound reinforcement system. Not only does the model show that some seats will be subjectively twice as loud as others but listeners sitting on the Left and Right hand sides will suffer poorer than intended dialogue to music / effects signal to noise ratios as these will have been mixed for a centre position.

6 COMMENTS & CONCLUSIONS

The measurement survey shows that the majority of British Living rooms provide an acoustically benign environment that should support good speech intelligibility. The measured reverberation times were lower than those previously reported in the literature but the room volumes may also have been lower.

The mean STI value of 0.83 suggests that the listening conditions are generally good and so are unlikely to be a cause of intelligibility complaints. However, the STI technique does not adequately take account of the TV set frequency response irregularities under such conditions.

Measurement of the TV frequency response, at typical listening distances, showed that wide variations (10-20dB) can occur and were of a nature such that experience suggests that this could affect perceived intelligibility. Unfortunately there is no currently agreed measure to assess this effect.

Measurement of different modes or setups of the TVs showed that some settings could improve the potential intelligibility by attenuating lower mid peaks and accentuating the speech consonant region (~1 to 5 kHz). It is conjectured though as to how many viewers are aware of these settings.

The frequency response problem is exacerbated by the current trend of locating the loudspeakers below the screen at the bottom of the set where they are directed downwards rather than outwards to the listener. With pedestal mounted sets this was shown to often give rise to strong comb filtering.

Whereas with good source material (i.e. clear speech) the frequency response anomalies may not be sufficient, by themselves, to significantly reduce the perceived intelligibility, however, with some voices or more particularly with poorly articulated speech, these effects may become more audible & problematic.

It has been asked by viewers as to why the problems with speech intelligibility are not picked up and corrected at the production stage. It is likely that the acoustic conditions and loudspeaker frequency responses that the broadcasters / producers use would not exhibit the typical degradations that the survey measurements have highlighted. Furthermore, familiarity with the programme material can have a very significant effect on the perceived intelligibility, which cannot be ‘listened through’ or compensated for – unless it is extremely poor.

In all the rooms surveyed, the background noise level was low (< 35dBA). However, in higher noise environments additional masking of the speech and further loss of intelligibility would occur.

The work highlights the need for an objective / electroacoustic method to assess the effects of poor or irregular frequency response on perceived intelligibility and listening fatigue under nominally good acoustic conditions.

As all the impulse responses have been stored, it is hoped that some additional analyses can be carried out together with auditioning via auralisation under controlled conditions.

Some of the findings may also be applicable to cinema audibility and intelligibility – though the acoustic conditions may not fully scale. The higher SPLs employed in commercial cinemas would also need to be accounted for and could well be a further contributory factor as well as the significant spatial variations in performance that are likely to occur.

An advantage of viewing DVD versions of films at home is that the L & R channels and surrounds can be turned down to reduce the masking of the centre channel which tends to be mainly dialogue, a benefit not available on most TV transmissions as few are in a 5.1 format. However this benefit is not available to mobile viewers, who now account for a significant proportion of the market.

It will be interesting to see if object based broadcasting offers solutions to some of the programme based issues but such technology cannot address the acoustic conditions of the listening environment and TV loudspeaker performance.

7 REFERENCES

[1] Daily Mail Article, 1 February 2016, “Can’t hear mumbling TV actors? You’ve tuned your TV set wrong, says TV Chief”

[2] R Hammond, P Mapp, A Hill “The Influence of discrete arriving reflections on perceived intelligibility and Speech Transmission Index Measurements” AES 141 Convention Los Angeles (2016).

[3] Mapp, P, “Modifying STI to better reflect Subjective Impression” AES 21st Conference, 2002.

[4] IEC 60268-16 - 2011 “Objective rating of speech intelligibility by speech transmission index”.

[5] BBC Research Department WHP 190, 2011

[6] Voice of the Listener & Viewer, “VLV’s Audibility of Speech on Television Project will make a real difference”, VLV News Release 06/11, March 2011.

[7] BBC, “Editorial Policy Guidance: Hearing Impaired Audiences”, March 2011

[8] Clear Sound: best practice tips”, BBC College of Production web site, March 2011

[9] Coren, S. “Most comfortable listening level as a function of age”, Ergonomics, 1994, vol.37, No. 7,

[10] Burgess MA, Utley WA, “Reverberation Times in British Living Rooms”, Applied Acoustics 18, 1985

[11] Jackson GM & Leventhal HG, “The Acoustics of Domestic Rooms”, Applied Acoustics, 5, 1972

[12] Gilford CLS, “The acoustic design of talks studios and listening rooms”, Proc Inst Electrical Engineers, 106, 1959.

[13] Leembruggen, G, Hippler M, Mapp, P “Exploring ways to improve STI’s recognition of the effects of poor spectral balance on subjective intelligibility”. Proc IOA 2009 Vol 31, Pt 4

[14] B Shirley and R Oldfield, " Clean Audio for TV broadcast: An Object-Based Approach for Hearing-Impaired Viewers", Journal of the Audio Engineering Society Vol. 63, No. 4, April 2015

[15] Musician advantage for speech-on-speech perception D Kent and E Gaudrain JASA Express Letters 2 March 2016

[16] Peutz VMA, “Articulation Loss of Consonants as a criterion for speech transmission in a room, JAES,V19, No 11, Dec 1971.

[17] Mapp P, “Dialogue Intelligibility”, AES 57th International Conference, The future of Audio Entertainment technology, Los Angeles, March 2015.