



University of
Salford
MANCHESTER

The clean audio project: Digital TV as assistive technology

Shirley, BG and Kendrick, P

Title	The clean audio project: Digital TV as assistive technology
Authors	Shirley, BG and Kendrick, P
Type	Article
URL	This version is available at: http://usir.salford.ac.uk/id/eprint/34322/
Published Date	2006

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.

The Clean Audio Project: Digital TV as Assistive Technology

Ben Shirley and Paul Kendrick
Acoustics Research Centre
University of Salford
Salford
UK
M5 4WT

Tel: 0161 2954524
Fax: 0161 2955145
Email: b.g.shirley@salford.ac.uk

Abstract

Technology used in Digital TV has the potential to enhance the viewing experience for millions of hard of hearing people. The Clean Audio project commissioned by the Independent Television Commission (ITC), and continued by Ofcom, looks at methods by which the extra information contained in 5.1 surround sound broadcasts may be used to improve the intelligibility and enjoyment of television audio for hard of hearing viewers and shows that audio processing can effectively turn a digital TV set top box into an assistive device to make digital TV more accessible. Listening tests were carried out which showed benefits in clarity and in perceived overall sound quality for hard of hearing participants by altering levels of centre and left and right channels. Further testing has shown average improvements in intelligibility of up to 9.4% by using surround sound equipment with a discrete central loudspeaker compared to stereophonic reproduction.

1. Introduction

There are estimated to be nearly 9 million people who are deaf or hard of hearing in the UK. Of these around 8.3 million suffer from mild to moderate deafness [1] and would benefit from any improvements that may be made in television audio. The Clean Audio project was set up by the Independent Television Commission (ITC) in response to viewers' complaints. The ITC received many complaints from hard of hearing people about the quality of sound on television, primarily that the dialogue is unclear and hard to understand owing to the level of background "noise". This noise consists of background music, sound effects and speech and it can have the effect of masking the dialogue and making it difficult or impossible to understand. Digital TV and surround sound has the potential for much improved TV sound quality and could therefore be of great benefit to hearing impaired viewers.

2. Literature Review

There is little previous research into television sound for hard of hearing people other than that focussing on the use of subtitles and other non-audio queues. For this reason the research has been informed by work aimed at hearing aid development and more general work on speech intelligibility and clarity.

Hearing aid design has used a number of approaches that could have application to television sound. Turner and Hurtig [2] investigated using frequency compression as an aid to intelligibility and found some improvements but concluded that it was less effective than high frequency amplification in most subjects. In a smaller study Mazor et al [3] found that frequency compression actually reduced intelligibility in most cases. Roch et al [4] discuss the benefits of frequency compression for some

listeners with sensorineural hearing loss and propose a pattern recognition system to compensate for the material dependent nature of this method. The research found that voices with different fundamental frequencies required different degrees of frequency compression to attain the best intelligibility improvements.

Multichannel amplitude compression solutions have been investigated and have shown superior benefits to conventional linear hearing aids (Moore et al [5] [6], Laurence et al[7]) although this is not universally accepted. Plomp [8] argues that fast acting multichannel amplitude compression has a negative effect on speech intelligibility and the subject has been the source of much debate. Humes et al [9] also compared conventional linear hearing aids with 2 channel, wide dynamic range compression (WDRC) aids and used a longer test period to allow for acclimatisation effects. This research utilised the Connected Speech Test designed by Cox et al [10] and found benefits to both types of hearing aid but with greater improvements being shown using WDRC, particularly for lower speech levels. Moore and Glasberg [11] compared the performance of single channel and two channel compression in hearing aids and found benefits to both but significantly better results from the two channel system in noisy situations. Barford [12], on the other hand, found multichannel compression to have less intelligibility benefits than an optimally fitted linear hearing aid. It is important to state that the characteristics of these multichannel aids are individually tailored to each individual and may therefore be of limited benefit in developing any “hard of hearing output” for digital television. However Moore’s research [13] indicates that compression can be beneficial even when not aiming to match the characteristics of an individual’s hearing loss.

The Dolby Digital 5.1 surround sound format may in itself bring advantages for hearing impaired and other television viewers. Some research suggests that there may be intelligibility benefits for television sound by the addition of a central loudspeaker, as is used in 5.1 surround sound systems, compared to a central “phantom” stereo image. Holman [14] suggested that the addition of a central loudspeaker made the material easier to understand although may not actually produce greater intelligibility. This effect, leading to an apparent difficulty in understanding, is a result of acoustical crosstalk (Holman [15]) that occurs when two identical signals arrive at the ear with one slightly delayed compared to the other. This produces a comb filtering effect that cancels out some frequencies in the audio. Other work shows actual intelligibility gains by using a central loudspeaker. Recent research found average improvements of word recognition in babble of up to 9.4% when comparing a central loudspeaker with a phantom stereo image [16]. These tests were carried out utilising an adapted Speech Perception In Noise test based on the SPIN test developed by Bilger et al [17] and was carried out with non-hearing impaired subjects. Methods have been proposed to enable the cancellation of this crosstalk by Cooper and Bauck [18] and Bauck and Cooper [19] but these may be impractical in the context of television viewing as they rely heavily on the listener being in the ideal listening position. It is suggested by Dressler [20] that the downmix process, whereby a 5.1 surround sound audio stream is converted for 2 channel playback, may distort the mix in such a way as to reduce intelligibility.

In addition to the intelligibility of the audio information there is an issue of the degree that visual cues can influence understanding of test material. Grant et al [21] found great variability between subjects in their ability to utilise audio/visual integration to

improve understanding of material but estimated potential improvements using visual content of up to 26% in some individuals. Any test procedures incorporating visual material must therefore be carefully designed to eliminate any bias resulting from greater or lesser visual cues.

3. Server Side or Client Side Solution

An important element in the planning of any research into broadcast accessibility solutions is inevitably the appropriate point in the route from producer to viewer at which change should be implemented. Changes in appropriate legislation, recommendations and guidelines can be implemented with a “top down” approach; this can be carried out at an international level and so retain or improve compatibility between the broadcast systems of different countries. Standards committees and professional bodies can be influential in bringing accessibility and inclusivity issues to the fore and in promoting solutions. In collaboration with the major audio companies in the world they are responsible for publishing the standards by which all of these companies should comply.

An alternative approach is to bring about improvements in the set top box (STB), at the viewers’ end of the chain. This approach may be capable of providing more in the way of a “quick fix” solution; an add-on to a set top box could perform any audio processing and be fitted to existing equipment, STB manufacturers can re-programme the software of much equipment and there is a potential for solutions in viewers altering settings and choosing equipment based on its accessibility and on their needs. Much is possible but it is sometimes difficult to persuade industry to commit funds to benefit what they see as a niche market.

The Clean Audio project's aim is to produce recommendations for hard of hearing viewers so that they may make appropriate choices in, and get the most from, their digital TV equipment. It will also generate broadcast and production guidelines where appropriate with the aim of improving the clarity of dialogue on TV for hard of hearing people. The first phase of the project commenced in April 2003, phase 2 started in October 2004 and is currently in progress funded by Ofcom, the UK communications watchdog.

4. Digital TV and Surround Sound Broadcast

One of the features of digital audio broadcast is the capability of a far greater dynamic range than analogue broadcast, the difference in level between the quietest sounds and the loudest can be far greater. This capability is being utilised to the full by producers, not least because more and more viewers are listening to their TV sets connected to hi-fi or home cinema equipment which can cope with reproduction of a greater dynamic range than TV loudspeakers. This increase in dynamic range has obvious implications for viewers suffering from loudness recruitment and can make understanding much more difficult for a range of hearing impairments.

Alongside the roll out of digital TV, although some way behind, is the growth of surround sound broadcast with the most common surround sound format for digital TV currently being Dolby Digital 5.1 Surround Sound. At the heart of the Clean Audio project is the premise that by using the extra information contained within the Dolby Digital format it should be possible to improve the clarity of TV sound for hard of hearing viewers.

5. Possibilities Offered by Surround Sound

Broadcast

Surround Sound Broadcast offers a number of potential solutions to create “clean audio”. There is additional audio data and there is additional data about the audio. These may both be utilised in an attempt to improve dialogue clarity. For additional detail about the Dolby Digital 5.1 surround sound format see Appendix C for an overview of Dolby Digital Surround Sound.

5.1 *Multi-channel Audio*

The first of these solutions to be investigated was the simplest; in the vast majority of films implementing Dolby Digital Surround Sound the entire dialog resides in the centre channel and emanates from a loudspeaker very close to the television screen.

Almost all sound effects, music and other peripheral audio is contained within the left and right front channels, coming from the front left and right loudspeakers, and in the rear surround channels, coming from the rear left and right loudspeakers. As mentioned earlier it is possible to make the dialogue clearer by reducing the level of the surround channels relative to the dialogue channel although the effect of this on the enjoyment and perceived sound quality for non-hearing impaired people was not clear. Details of an investigation into this possible solution are covered later in this paper.

5.2 *Hearing Impaired (HI) Audio Channel*

The AC3 stream has the capability to contain an audio channel intended as an aid to hard of hearing people. The HI channel is intended to be used as a single mono-aural audio channel containing dialogue processed so as to make it more intelligible for

hearing impaired viewers. Other than a statement that the HI channel should contain processed dialogue there is little guidance as to how this improved intelligibility should be gained. It is hoped that the Clean Audio project can bring some much needed clarity to this subject and this may be of benefit in applications such as DVD production where bandwidth is not a major consideration. In the broadcast environment however bandwidth is severely limited and a separate audio feed for hearing impaired people is unlikely to provide a solution that will be taken up by broadcasters. In the Clean Audio project a decision was made to concentrate on solutions that would not increase the bandwidth and therefore the relative cost to broadcasters.

5.3 Metadata

In addition to the extra audio channels available in the AC3 format the bit stream also contains information about the audio. This metadata is primarily concerned with performing three main functions.

- Allowing changes between programmes and channels with no sudden changes in level.
- Controlling the downmix of the 6 channels in 5.1 surround for stereophonic or mono-aural reproduction.
- Determining how the programme material is compressed for playback in less than ideal listening environments.

The first of these is accomplished by the use of a value within the metadata that gives an average level based on the level of the dialogue in the programme material. This value, known as the dialogue normalisation level, or dialnorm, gives a reference in order that broadcasters can ensure a standard level between programmes and between

channels. This reference level is based on the average level of dialogue, not on the average level of the audio content overall. The capability to downmix the 5.1 surround audio to stereo or mono is vital in order that material can be played back on non-surround reproduction systems without requiring an additional audio channel to be broadcast. An additional audio channel would use up valuable bandwidth and so be more expensive and therefore impractical to implement. The metadata contains parameters that determine the level of rear surround channels compared to the dialogue channel and also the level of front left and right channels. The information contained within the metadata is known as the Bit Stream Information (BSI) or the Extended Bit Stream Information depending on whether some more recent optional parameters are implemented. The metadata contained within the AC3 stream has the potential to help provide a solution with no extra bandwidth required for broadcasters. Any processing or downmixing implemented at the STB end of the broadcast chain could potentially be controlled by values in the metadata. This potential is to be explored more fully in phase 2 of the Clean Audio project which commenced in October 2004. This use of metadata, and particularly the dialnorm parameter, relies heavily on producers and broadcasters using the metadata appropriately and research has shown this not to be the case. According to Dolby Labs *Guide to Metadata* [22], “The consumer’s Dolby Digital decoder reproduces the program audio according to the metadata parameters set by the program creator, and according to settings for speaker configuration, bass management, and dynamic range that are chosen by the consumer to match his specific home theater equipment and environmental conditions.” “This control, however, requires the producer to set the metadata parameters correctly, since they affect important aspects of the audio—and can seriously compromise the final product if set improperly.” Dolby’s own research [23]

reveals that only 1 out of the 13 digital services surveyed in one area had set the Dialog Normalisation value correctly and, as a result, the audio level for these services varied by as much as 16dB, much higher than the “comfort zone” defined by Dolby based on listening tests. This misunderstanding of the importance and use of metadata has serious implications for the implementation of any metadata controlled processing at the STB and will be addressed in guidelines to broadcasters as part of the Clean Audio project. A more detailed breakdown of metadata parameters is contained in Appendix A.

6. Clean Audio Phase 1

6.1 Introduction

This phase of the Clean Audio project aimed to ascertain what detrimental effects, if any, there were in lowering the level of the surround channels relative to the dialogue channel in 5.1 surround sound. There are obvious benefits in intelligibility to raising the relative dialogue level compared to background sounds but it was unclear as to the effect this would have on the viewing experience of non-hearing impaired viewers sharing a TV and how the lack of sound effects, music and other peripheral sounds would affect the perceived sound quality and enjoyment of hearing impaired people. Some results from this research are published in the proceedings of the 116th AES Convention [16].

6.2 Aims

The aims of Clean Audio phase 1 were as follows:

- To assess the effect of attenuating left and right channels in a 5.1 surround sound system for hearing impaired viewers.

- To assess any benefits of 5.1 surround sound compared to downmixed stereo.
- To assess the effect of this remix for non-hearing impaired viewers.
- To produce recommendations for hard of hearing viewers as to how they may improve their viewing experience.
- To produce guidelines for broadcasters.

6.2 Methodology

Listening tests took place in a listening room that conformed to ITU-R BS.1116-1 multi-channel stereophonic sound with and without accompanying picture recommendations [24]. The test methodology adopted was that of blind AB comparison listening tests. Subjects were asked to watch excerpts of video material with a Dolby Digital 5.1 encoded soundtrack. Each excerpt was split into two similar sections with a different process being carried out on each section. The subject was then asked to assess which of the two sections was preferred based on three criteria:

- Overall sound quality.
- Their enjoyment of the section.
- The clarity of the dialog.

Subjects were also asked to show how much better their preferred section was for each of these criteria. There was no option for the two sections to be assessed as being the same. All AB and BA comparisons were assessed by each subject, with the order of the processes changed for each subject so ensuring that every process was carried out on every video clip

6.3 Processes Assessed

Processes assessed in this phase of the project, agreed between the ITC and the research team, were as follows:

- Centre channel, plus left and right channels at standard relative levels set using reference tones.
- Centre channel, plus left and right channel at -3dB.
- Centre channel, plus left and right channel at -6dB.
- Centre channel only.
- LtRt Stereo downmix.

6.4 Test Material

The test material consists of a series of 20 video clips with a Dolby Digital 5.1 soundtrack. Each clip was split into 2 sections and each section treated with a different process on playback. To this end it was considered important that the amount and type of left and right side channel audio was consistent throughout the clip so that like was compared to like. Clips were introduced by a title reading “Clip x” (where x was the number of the clip), each section of the clip was introduced with a 3 second title reading “Section A” or “Section B”.

Clips were chosen according to the following criteria:

- Length of between 1 min and 1 min 30 seconds.
- Moderate amount of side channel audio that could possibly mask dialogue in the centre channel.
- A variety of types of side channel audio including background speech, music and sound effects.
- Mix of off-camera and on-camera speech; does the camera remain on the persons face while talking? Is the mouth clearly visible? Is lip reading possible?

It was thought important in choosing the clips that each clip should appear to be complete in itself, i.e. at the end of a clip the subject matter is brought to some sort of conclusion. This was seen as vital in order to avoid influencing the “enjoyment” factor for each clip/process combination and in order to reduce potentially irritating breaks in the video sequences. Within this limitation, the length of each clip was standardised as far as possible. This avoided results being unduly influenced by the concentration span of the participants.

6.5 Analysis of Subject Group

The group was composed of 41 subjects with a range of ages and hearing impairments ranging from severely deaf to non-hearing impaired. Subjects were recruited by a number of means; by advertisement, via the University of Manchester Age and Cognitive Research Unit, through Hearing Concern and from the student population. All subjects signed a consent form indicating that they were willing to take part in the test and that their data could be kept on record (see Appendix B). Each subject’s hearing was assessed prior to tests by means of an audiogram carried out by the research team using Bekesley Pulsed Audiometry. Ethical approval for the research was granted by the University of Salford Research and Governance Ethics Committee.

The profile of subjects was as follows.

6.51 Age Profile

The age profile of the subject group is shown in the lower row of Table 1. Subjects were chosen in order to include a range of age groups.

6.52 Profile of Hearing Impairments by Age Group

The level of hearing impairments for each age group is also shown in Table 1. The degree of hearing impairment is shown in the far left column with the age groups of subjects being shown in the top row. The number of subjects in each hearing impairment category can be read in the appropriate cell in the table. Hearing loss was categorised using a pure tone audiogram with the hearing level threshold levels averaged at 250, 500, 1000, 2000 and 4000 Hz¹. When differentiating between hearing impaired and non-hearing impaired subjects a threshold of 20 dB (HL) was used.

6.6 Results

6.61 Data Analysis

During the listening test subjects indicated their preference section by use of a tick box and marked on a scale in order to indicate how much more they preferred it. Three scales were used in order to indicate preference based on dialogue clarity, the overall sound quality and their enjoyment of the clip. An example scale can be seen in figure 3. Each scale was unmarked and labelled “Slightly Better” at one end and “Much Better” at the other, with no option for the sections to be equally rated. This scale was defined and used differently by each subject so the results were normalised. Normalisation was carried out using each subject’s minimum and maximum values

1

Audiometric descriptor of loss	dB Hearing Loss
Mild	20 – 40
Moderate	41 – 70
Severe	71 – 95
Profound	> 95

Taken from The British journal of Audiology, 1988, 22, 123, Descriptors for pure-tone audiograms. [3] When differentiating between hearing impaired and non-hearing impaired subjects a threshold of 20 dB (HL) was used in order to remain consistent with the above recommendations.

rather than mean and standard deviation owing to the non-Gaussian distribution of results. The normalisation was done using the following formula:

$$pref_{norm} = \frac{pref_{original} - \min(pref_{original})}{\max(pref_{original}) - \min(pref_{original})}$$

This produced a value between 0 and 1; this was given a positive value to indicate when the process was preferred and a negative value to indicate when it was not. Processes were assessed in terms of perceived clarity of dialogue, overall sound quality, and how enjoyable the process made the clip. These results are summarised in figure 1. The most striking result from the tests was the low rating of the LtRt stereo downmix when compared with all other conditions. All other conditions were preferred to LtRt stereo for dialogue clarity, overall sound quality and enjoyment by both hearing impaired and non-hearing impaired subject groups. As can be seen from the graphs (figure 1), reducing the level of the side speakers from default, to -3dB, -6dB and then removing side channels entirely has very different effects on the perception of hearing impaired and non-hearing impaired subjects. Unsurprisingly, the clarity of the dialogue is perceived by both groups as having improved as the level of side channels is reduced. For the non-hearing impaired group this had an inversely proportional effect on the perceived overall sound quality and on their enjoyment of the video. It may be clearer but it sounded of poorer quality and was not as enjoyable. For the hearing impaired group however, the reverse was true; their enjoyment and the perceived sound quality were directly proportional to the clarity of the dialog. Data was analysed using a multivariate ANOVA with Tukey HSD with all comparisons between processes being tested. Figure 1 shows the average opinion with error bars showing 95% confidence limits. There is a marked statistical significance to most of the combinations tested with hearing impaired subjects giving a high degree

of confidence in these outcomes. The non-hearing impaired results show less significance, possibly as a result of the lower number of subjects.

6.7 Conclusions and Recommendations

The poor rating of the stereo downmix is as a result of the downmix process where the 6 channels in the AC3 audio stream are remixed for 2 channel stereophonic reproduction. The LtRt downmix is derived from a mix of all 5 full range channels including left and right rear surrounds. The inclusion of the rear surround information in this mix reduces the relative level of the centre channel, usually used for dialogue, and so affects the clarity of the dialogue compared to the other mixes, none of which include rear surround audio. Although this result could have been predicted the LtRt derived 2 channel stereo is often the default stereo on decoders and so provides an interesting reference between what a viewer may be listening to now and what improvements could be possible with surround sound equipment. For the hearing impaired subject group, perceived overall sound quality and enjoyment was directly related to the clarity of dialogue. The ratings of the other processes indicate that hearing impaired viewers may benefit from reducing the level of surround channels, maximum benefit being gained by muting side speakers entirely. For the non-hearing impaired subject group the reverse was true; although clarity was enhanced by reducing surround channel levels, this detracted from the perceived sound quality and enjoyment of the material. The conclusions of phase 1 can be summarised as follows:

- 5.1 surround sound broadcasts have the potential to dramatically improve the viewing experience of hard of hearing people.
- Hard of hearing viewers can significantly improve the dialogue clarity of Dolby Digital 5.1 programme material on television by listening to centre

(dialogue) channel only. This can result in a perceived improvement in sound quality and may enhance their enjoyment of the programme material.

- Hard of hearing viewers sharing a television with non-hearing impaired viewers can benefit from lowering the level of the surround channels. This can be less detrimental to the enjoyment of non-hearing impaired viewers than removing surround channels completely but can still improve dialogue clarity.

7. Current and Future Research

7.1 The Dialogue Channel and the Centre Loudspeaker

Phase 1 of the Clean Audio project demonstrated that surround sound broadcast can provide distinct benefits for hard of hearing people. The attenuation of rear surround and front left and right channels can be implemented using parameters within the Extended Bit Stream Information as part of the process of downmixing the AC3 audio stream to stereo. This can be done with no extra broadcast overhead of higher bandwidth requirements; parameters can be set at the broadcast end of the chain and implemented, or not, depending on the needs and preferences of individual viewers. What is not clear is the degree to which these benefits are dependent on having a discrete and separate dialogue loudspeaker as opposed to listening to a “phantom” stereo image between 2 loudspeakers. In stereo reproduction a centrally panned sound source, such as film dialogue, is obtained by replaying the same sound equally from both left and right loudspeakers. This produces a central sound source whose perceived location is determined to some degree by the position of the viewer relative to the loudspeakers. In the first phase of Clean Audio reproduction of dialogue was via a separate central loudspeaker positioned immediately above the TV as is used in surround sound reproduction systems.

More recent research carried out at the University of Salford [25] has investigated the effect of using a discrete central loudspeaker compared to stereo. The research focussed on dialogue intelligibility, as opposed to perceived dialogue clarity, using a test adapted from the Speech Perception In Noise (SPIN) test [10]. The listening tests involved playing a number of sentences each ending in a keyword which was to be identified; the total number of keywords identified was used to score the intelligibility of the playback method. The tests, carried out with non-hearing impaired subjects, showed measurable improvements in intelligibility of up to 9.4% using a separate central loudspeaker when compared with a phantom image between a pair of stereo loudspeakers. The full results of this research will be published in the near future.

7.2 Compression

It is likely that a form of compression may provide some benefits to hard of hearing viewers although opinions on the benefits of compression techniques for hearing impairments are mixed. Villchur [26] and Moore [7] [5] have found some intelligibility benefits to compression whereas Plomp [8] has found compression to be detrimental to intelligibility.

Compression has the effect of raising the level of quiet sounds relative to louder sounds and this can be helpful in bringing softer speech sounds within hearing range; it can also distort the envelope fluctuations in speech and so be detrimental to intelligibility. Band limited compression can be used to compress frequencies differently in different frequency bands. The most common hearing loss is concentrated in the mid to high frequencies and it is possible to compress and give gain to frequencies where helpful and to let hearing function normally where it is not.

Many hearing aids work on the principle of multiple band compression systems and these can be effective particularly where the aid is tailored for an individual user.

There are however side effects to increasing the number of frequency bands used; the speech envelope tends to be distorted to a greater extent and this can reduce intelligibility. It is possible that some form of compression may provide some of the answers to dialogue clarity issues experienced by hard of hearing viewers and future research will investigate this possibility.

7.3 Further Processing

In addition to the possibilities offered by compression systems of various degrees of complexity, it may be possible to implement more complex adaptive filtering methodologies to gain clarity of speech. Techniques already used in mobile communications technology and other fields may provide beneficial effects although these techniques have not been designed for people with hearing impairments and the effects on this group have yet to be established. Further research will be needed in order to test and adapt some of these techniques if they are to be useful in this application.

8. Conclusions

It has been shown that there are distinct benefits for hard of hearing people in the growth of digital TV and particularly in surround sound broadcast. As surround broadcast is introduced to the UK these benefits can be applied across the TV schedules. The implementation of phase 1 findings using metadata is currently under investigation to see if it is possible to incorporate control of a “hard of hearing” output from STBs by altering relative levels between channels. Dolby cinema processors already incorporate a hard of hearing output for induction loop systems using

attenuated surround channels to increase the relative level of the dialogue and this type of arrangement could readily be implemented in STBs. Other options such as those described earlier are under investigation however any processing based on dialogue level will be heavily dependent on production and broadcast decisions. It is vital that broadcasters and producers correctly use parameters in the metadata such as the dialnorm value in order that many of these processes can effectively aid hearing impaired viewers. It is likely that some element of a top down approach to the problem will be necessary in order to ensure standardised use of metadata and Ofcom are ideally positioned to undertake guidance of this sort. In parallel with this activity there is a viewer centred approach. There will be steps that can be taken by hard of hearing viewers in order to improve dialogue clarity once the potential benefits of surround sound systems become more widely understood. Any recommendations or guidelines that are likely to benefit hearing impaired viewers can have an immediate effect in influencing product choice as surround sound broadcast is rolled out. Product choice can influence product development and this will encourage developers to revise their opinion of the scale of this supposed niche market. In an ageing population the balance of consumer power is shifting and inclusive design is the key that will enable companies to unlock its potential.

9. Acknowledgements

The Clean Audio Project was funded by the Independent Television Commission and is now funded by Ofcom.

10 References

- [1] Hearing Concern, *Advice and Information, About Hearing Loss, Deaf Awareness*. 2004.
- [2] Turner, C.W. and R.R. Hurtig, *Proportional frequency compression of speech for listeners with sensorineural hearing loss*. Journal of the Acoustical Society of America, 1999. **106**(2): p. 877-886.
- [3] Mazor, M., et al., *Moderate frequency compression for the moderately hearing impaired*. Journal of the Acoustical Society of America, 1977. **62**(5): p. 1273-1278.
- [4] Roch, M., et al. *Towards a Cohort-Selective Frequency-Compression Hearing Aid*. in *International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences*. 2004. Las Vegas.
- [5] Moore, B.C.J., *Design and evaluation of a two channel compression hearing aid*. Journal of Rehabilitative Research and Development, 1987. **24**: p. 181-192.
- [6] Moore, B.C.J., et al., *Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss*. Ear and Hearing, 1992. **13**: p. 349-370.
- [7] Laurence, R.F., B.C.J. Moore, and B.R. Glasberg, *A Comparison of Behind-the-Ear High-Fidelity Linear Hearing Aids and Two-Channel Compression Aids, in the Laboratory and in Everyday Life*. British Journal of Audiology, 1983. **17**: p. 31-48.

- [8] Plomp, R., *The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function*. Journal of the Acoustical Society of America, 1988. **83**(6): p. 2322-2327.
- [9] Humes, L.E., et al., *Application of the Articulation Index and the Speech Transmission Index to the recognition of speech by normal-hearing and hearing-impaired listeners*. Journal of Speech and Hearing Research, 1986. **29**(4).
- [10] Cox, R.M., G.A. Gray, and G.C. Alexander, *Evaluation of a revised speech in noise (RSIN) test*. Journal of the American Academy of Audiology, 2001. **12**: p. 423-432.
- [11] Moore, B.C.J. and B.R. Glasberg, *A Comparison of Two Channel and Single Channel Compression Hearing Aids*. Audiology, 1986. **25**: p. 210-226.
- [12] Barford, J., *Multichannel Compression Hearing Aids: Experiments and Consideration on Clinical Applicability*. Hearing Impairment and Hearing Aids, 1978. **6**: p. 315-340.
- [13] Moore, B.C.J., *Speech Processing for the Hearing-Impaired: Successes, Failures, and Implications for Speech Mechanisms*. Speech Communication, 2003. **41**: p. 81-91.
- [14] Holman, T., *New Factors in Sound for Cinema and Television'*. Journal of the Audio Engineering Society, 1991. **39**(7): p. 529-539.
- [15] Holman, T. *The number of audio channels*. in *100th Convention of the Audio Engineering Society*. 1996. Copenhagen.

- [16] Shirley, B.G. and P. Kendrick. *ITC Clean Audio Project*. in *116th Convention of the Audio Engineering Society*. 2004. Berlin: Audio Engineering Society.
- [17] Bilger, R.C., et al., *Standardization of a test of speech perception in noise*. *Journal of Speech and Hearing Research*, 1984. **27**: p. 32-48.
- [18] Cooper, D.H. and J.L. Bauck, *Prospects for transaural recording*. *Journal of the Audio Engineering Society*, 1989. **37**(1): p. 3-19.
- [19] Bauck, J.L. and D.H. Cooper. *Generalised transaural stereo*. in *93rd Conference of the Audio Engineering Society*. 1992. San Francisco: Audio Engineering Society.
- [20] Dressler. *A Step Toward Improved Surround Sound Making the 5.1-Channel Format a Reality*. in *100th Convention of the Audio Engineering Society*. 1996. Copenhagen: Audio Engineering Society.
- [21] Grant, K.W., B.E. Walden, and P.F. Seitz, *Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition and auditory-visual integration*. *Journal of the Acoustical Society of America*, 1998. **103**(5): p. 2677-2690.
- [22] Dolby Labs, *Dolby Metadata Guide issue 2*. 2003.
- [23] Riedmiller, J.C., S. Lyman, and C. Robinson. *Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?* in *115th Convention of the Audio Engineering Society*. 2003. New York: Audio Engineering Society.

- [24] ITU, *ITU-R BS.1116-1: Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*. 1997, International Telecommunication Union.
- [25] Shirley, B.G. and P. Kendrick. *Measurement of speech intelligibility in noise: A comparison of a stereo image source and a central loudspeaker source*. in *118th Convention of the Audio Engineering Society*. 2005 (forthcoming). Barcelona: Audio Engineering Society.
- [26] Villchur, E., *Multichannel Compression in Hearing Aids*, in *Hair Cells and Hearing Aids*, C.I. Berlin, Editor. 1996, Singular Publishing Group. p. 113-124.

Appendix A

Dolby Digital Metadata Parameters

(Parameters in italics are part of the Extended Bit Stream Information)

Program Configuration

Program Description Text

Dialogue Level

Channel Mode

LFE Channel

Bitstream Mode

Line Mode Compression

RF Mode Compression

RF Overmodulation Protection

Center Downmix Level

Surround Downmix Level

Dolby Surround Mode

Audio Production Information

Mix Level

Room Type

Copyright Bit

Original Bitstream

Preferred Stereo Downmix

Lt/Rt Center Downmix Level

Lt/Rt Surround Downmix Level

Lo/Ro Center Downmix Level

Lo/Ro Surround Downmix Level

Dolby Surround EX Mode

A/D Converter Type

DC Filter

Lowpass Filter

LFE Lowpass Filter

Surround 3 dB Attenuation

Surround Phase Shift

Taken From *Dolby Metadata Guide vol 2* published by Dolby Labs [22].

Appendix B



Information Sheet

ITC Research Project into Improving Television Sound

We would be very grateful if you could help us in an important research project about the quality of sound on television.

The University of Salford Acoustics Research Centre is carrying out a series of tests on how to improve the quality of sound from television. Of course you are under no obligation and do not have to participate, but it would be extremely valuable if you could assist us in this study.

The work is sponsored by the Independent Television Commission (ITC) and your responses to the tests will help us develop better sound for television in the future.

We need your consent to:

- Retain some background information on you (name, age, gender, contact details);
- Carry out a hearing test and retain an audiogram showing your hearing ability.
- Carry out a series of tests where you listen to speech and other TV programme content and we ask a series of questions intended to assess how well you have heard and enjoyed the recordings. We need your permission to retain the results.

All information will be kept confidential. The work will be used to help improve the quality of TV sound. No individuals will be identified in the results of the research. As some of this data is held on computer, some is covered by the data protection act, and you will be able to see a copy of it on request.

Both hearing and hard of hearing people are required for the tests though we are particularly interested in contacting hard of hearing people who may wish to participate. If you would be willing to participate or if you know of anyone else who may be interested my contact details are as follows:

Ben Shirley
Lecturer
Acoustics Research Centre
University of Salford
0161 2954524
b.g.shirley@salford.ac.uk

www.acoustics.salford.ac.uk

Consent Form

I have read and understood the information sheet and this consent form. I understand that participation is voluntary and I am free to withdraw at any time.

Name: _____

Signature: _____

Date: _____

Appendix C Overview of Dolby Digital Surround

Sound

Dolby Digital 5.1 is the format chosen by Sky™ for their current surround sound broadcasts in the UK and, with around 28 million Dolby Digital receivers in use throughout the world, it looks set to continue as a market leader. The Dolby Digital format minimises bandwidth by using data compressed audio and currently allows for the use of up to 5 full frequency range audio channels and 1 low frequency effects channel. Loudspeakers are arranged with one central front channel (normally used for dialogue), front left and right loudspeakers and rear left and right surround loudspeakers arranged as shown in figure 1. The audio is broadcast as an AC3 bit stream and it is the format and content of this bit stream that may enable us to implement changes beneficial to hard of hearing viewers. The AC3 bit stream consists of between 1 and 6 discrete channels of audio, and metadata. AC3 metadata can be described as data about the audio data. The audio is compressed in the encoding process and AC3 streams of various bit rates encompass multi-channel and single channel formats. Additional audio channels can be included for multiple language support and there is the potential to include Hearing Impaired (HI) and Visually Impaired (VI) audio channels for viewers with sensory impairments. The metadata contains information about these audio channels, their format, how they are to be decoded, downmix parameters required to convert from 5.1 to stereophonic or mono-aural and the type of audio compression that should be applied if any.

Unlike some surround sound systems, the AC3 format maintains a separation between audio channels in the encoded bit stream, in other words, there are 6 discrete and

separate audio channels present in a 5.1 encoded AC3 stream. This in itself means that we can easily change relative channel levels and attenuate or amplify each channel independently. In much 5.1 encoded material the centre channel is used as a dialogue channel so gains in dialogue clarity can be made by attenuating the level of the surround loudspeakers relative to the dialogue.

Tables

Table 1: Degree of hearing impairment for each age group

	<30	30-44	45-59	60-74	>75
Non-Impaired	4	7	0	4	2
Mild	1	0	1	2	7
Moderate	0	1	2	3	4
Severe	0	1	0	0	0
Profound	0	1	0	1	0
TOTAL	5	10	3	10	13

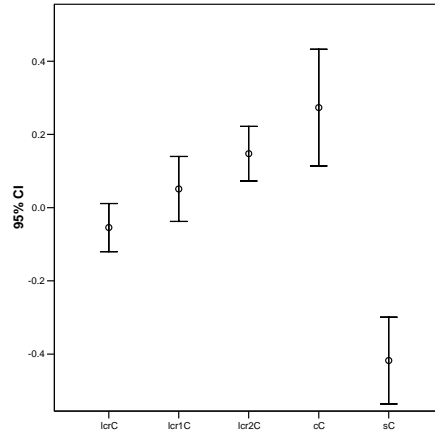
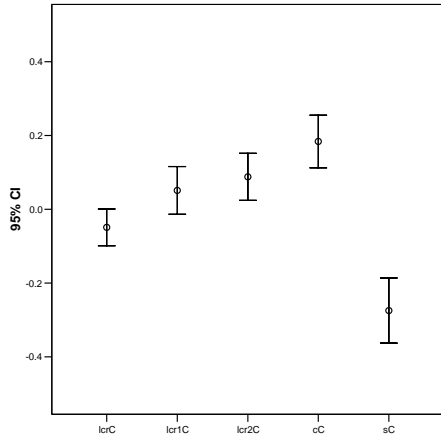
Figures

Figure 1: Results Showing Confidence Error Bars

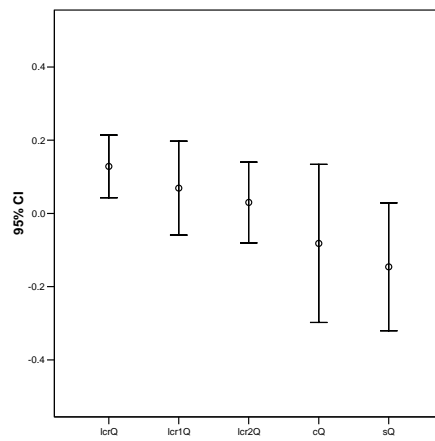
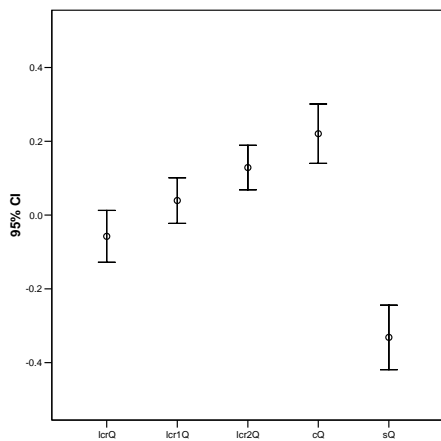
Hearing Impaired Group
29 Subjects

Non-Hearing Impaired Group
12 Subjects

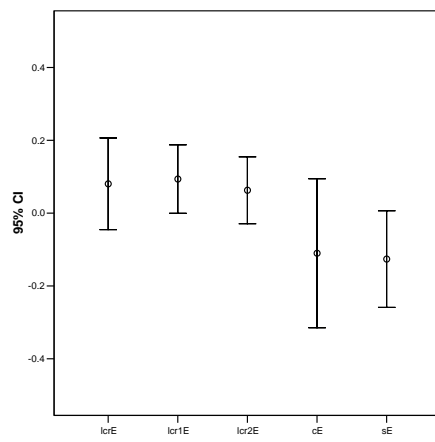
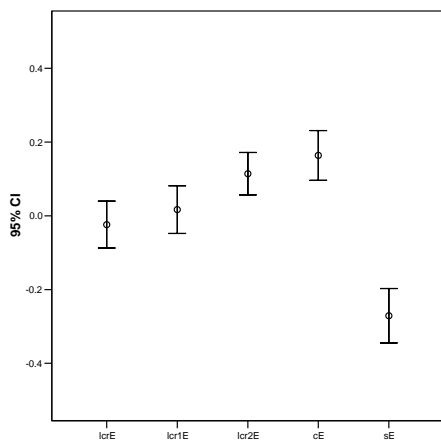
Clarity



Overall Sound Quality



Enjoyment



Key: lcr reference levels, c centre channel only
 lcr1 left and right -3dB s LtRt stereo downmix
 lcr2 left and right -6dB

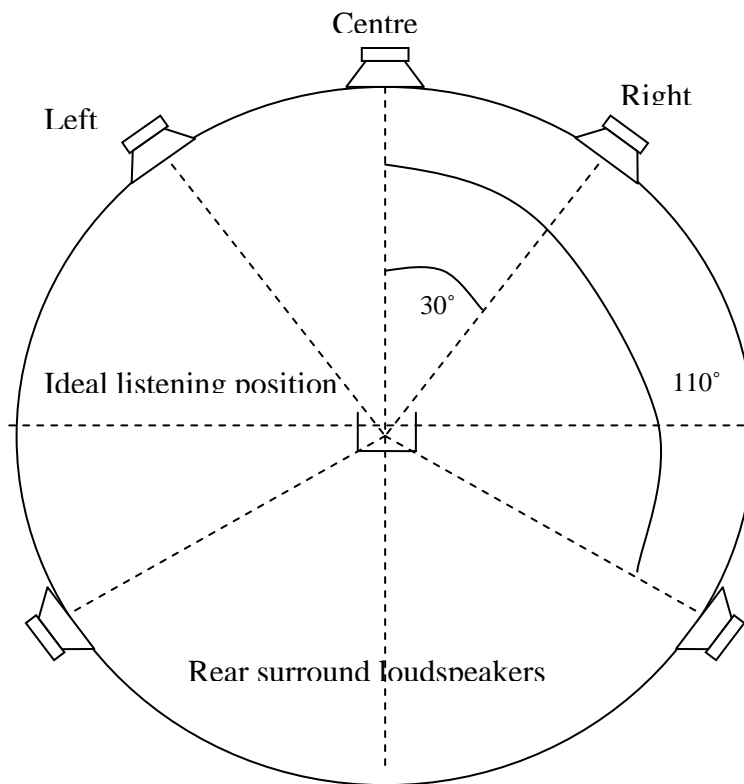


Figure 2: Loudspeaker Setup for 5.1 Surround Sound

Example

Which section do you think had the best sound quality.

Section A

Section B

How much better was your choice?

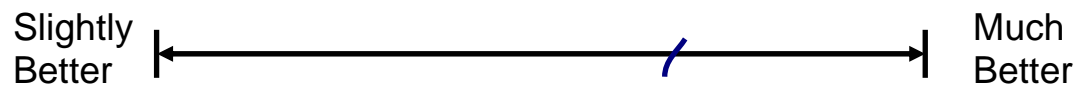


Figure 3 Example scale from questionnaire