

The Audibility of Speech on Television. Analysis of Problems

**Programme reviews, detailed
assessments and report by:**

**David Walker CEng MIET
Louise Willcox MIBS (Partner DWR Associates)**

The Audibility of Speech on Television. Analysis of Problems

Audience Surveys

Three television audience surveys were conducted between 14th August and 23rd August 2010, the BBC Pulse internet-based survey, the Voice of the Listener and Viewer diary-based survey of non-internet users over the age of 65*, and returns from a survey published in the magazine of the Royal National Institute for the Deaf. These surveys identified television programmes having some problem with the audibility of spoken dialogue.

From those programmes, 21 were subjected to detailed review and analysis. This detailed work was carried out by Louise Willcox, (DWR Associates) a highly knowledgeable Senior Sound Supervisor with many years experience of recording, balancing and dubbing television programmes in the studio and on location, and by David Walker, a senior audio, electronics and acoustics engineer. DVD's of the programmes were provided by BBC and Channel 4 for their networks, ITV1 and Channel 5 programmes were recorded off-air, ITV from Freeview and Channel 5 analogue terrestrial.

Of the 21 programmes reviewed, no problem was identified with 3 (Who do you think you are: Rupert Perry-Jones, Inspector Morse, The Queen), leaving 18 with identifiable faults.

An additional programme from outside the survey was also reviewed and commented upon by request but has not been included in this analysis.

Some Background about Hearing Sound

- 10 to 20% of what we hear is confirmed by watching people's lips, especially in noisy environments.
- A pause as short as a breath before uttering the next sentence, or cutting to the next item is also used as a thought pause, which allows the brain processing time. It is as if the brain is saying "that's just finished, that's what it is about, now on to the next thing". Thus, fast –paced editing can actually make things difficult to hear, no matter how well recorded, if the brain has not been given enough time to process the information.
- Popular music usually has a much smaller dynamic range than classical music, because it is compressed. Compressed music sounds subjectively louder than wider dynamic range music even if the peak level is the same on a PPM.
- Music manipulates the emotions, which is why it is used in drama. Fast music generates a feeling of urgency. Slow music has the opposite effect.
- Background music with lyrics needs to be at a lower level than without because it is a further distraction to the listener, who will often tune in to the lyrics rather than the spoken word.
- Dramatic speech normally has a greater dynamic range than for instance, the dialogue in a quiz show.
- Male voices are usually louder than female voices. Testosterone makes a man's vocal chords longer usually resulting in a deeper tone, and gives a man one third more muscle power than the same build woman, increasing the power of the diaphragm, the muscle that pushes air through the vocal chords. Ergo, women actors will loose out in a two-hander with a man unless they project more.

* The VLV survey was initiated by Dick Bates, Peter Meneer and David Walker, assisted by Louise Willcox, who all gave their time free of charge; facilitated by Voice of the Listener and Viewer; funded by Channel 4, Widex (a hearing aid manufacturer, and a private donor).

- A female voice has most power in the higher frequencies. High frequencies are most commonly lost when we lose our hearing so many of those who are hard of hearing find female voices more difficult to understand than male.
- Consonants convey the meaning in speech and are transmitted in the higher frequencies. With high frequency loss (whether hearing or in the recording), speech is difficult to decode.
- The ear is sensitive to defects in sound. It will, for instance, notice a quite short gap in the sound, and have difficulty “bridging” deficiencies.

Detailed Analysis.

For the analysis, Louise used a television, modified to reproduce mono sound (both sound channels summed together). Some programmes were also assessed on a stereo television for confirmation. David used a Panasonic plasma stereo television with good quality sound, and a Hitachi LCD stereo television with average quality sound.

The listening level was set by adjusting the volume in the first 30 seconds, and then left unaltered for the remainder of the programme.

Identified Problems.

Edit/ Dub Related.

Music masks voice: in 6 programmes (33%), added music occasionally masked the voice of the presenter or actor. In one programme the opening and theme music was perceptibly louder than the general overall level.

(Programmes affected: The One Show, Weakest Link, Vexed, Coast, The Bill: Tombstone, Come Dine with me Down Under, Great Railway Adventures with Dan Cruikshank).

Discussion: This is probably the most commonly complained about cause of poor audibility, and was found in a significant percentage of programmes reviewed. In many ways it should be the most easily avoided as the music is usually added in the final dub when conditions should be completely under control. (An exception to this is for instance, “The Weakest Link” where it is understood that the music bed and sting are recorded “as live” or The One Show, whose opening titles and bed are live, but VT music is prerecorded).

Music can be added for many reasons; in these reviews, to provide a background infill when picture action was recorded without synchronous effects (ambient sound), and to create mood or add dramatic effect. At times it was felt to be an annoyance, and this tends to condition the viewer to be less sympathetic if dialogue gets difficult to understand. Music was sometimes used as “wallpaper” not adding to the viewer experience and in the opinion of the reviewers not strictly necessary.

The type of music was a significant factor. Percussive music punches through the dialogue and can be quite obstructive even if the average level is reasonably below the level of dialogue. Drums, cymbals, plucked strings, piano were particular examples here. Anything with a vocal can also distract from the dialogue and needs to be treated with care.

The tonal quality of the music was also found to be a factor. Where music has significant content in the vocal range, the masking effect is magnified and relatively low levels clash with the voice and interfere with understanding.

Where a programme had strong signature music or theme it was found that this tended to set the listening volume. This can cause difficulty if the average programme level is lower.

Dipping the level of music as a presenter starts to speak worked well provided that the dip was quick and sufficiently low. In some instances the dip was slow and the first word(s) were concealed.

Recording Related

Poor microphone placement: 6 programmes (33%) suffered from poor microphone placement resulting in indistinct dialogue.
(Programmes affected: Five Days (Ep 1), The Deep, Coast, Great Railway Adventures with Dan Cruikshank, Our Drugs War).

Discussion: Whilst it is not always possible to site a microphone in an ideal position, a trained sound operator will know how to mitigate the effects of non-ideal placement. Unfortunately, untrained researchers, cameramen or other persons are sometimes used for sound recording. The problems encountered included the following:

Personal microphones under clothing. Personal microphones should give good coverage, but there were examples where they were placed under clothing, resulting in a lack of presence (the higher speech frequencies which convey the consonants and therefore meaning), and increased clothing noise, which masked the speech. This cannot be corrected in the dub as any correction to the speech is also applied to the overlying clothing noise. Unfortunately, Producers will often insist that microphones are unseen, not understanding the full consequences of this. Further, there were examples where the level changed as the artist turned their head away.

Use of camera microphones. Camera microphones are useful for providing synchronous effects (background sound recorded synchronously with dialogue), but are rarely satisfactory for picking up speech at any distance. There were examples where ambient noise and room acoustic competed with the speech, and the sound perspective changed when the camera panned. A separate, good quality microphone with appropriate characteristic will almost always be superior.

Microphone too distant. If the microphone is not close enough to the action too much ambient noise and room acoustic will be recorded with the dialogue. In controlled situations use of a boom is preferred. The situation will be exacerbated if the actor is giving a quiet vocal performance (see below).

Sync effects mask voice: in 5 programmes (28%) synchronously recorded effects masked the vocal performance. In one instance this may have been due to poor planning and operator error.

(Programmes affected: Casualty, Athletics London Grand Prix, Vexed, Coast, Great Railway Adventures with Dan Cruikshank).

Discussion: Where the microphone picks up ambient sounds at a level similar to the voice this may be a combination of a low voice level and the microphone being too distant, or a noisy environment. It is not possible to separate the effects and voice as whatever is done to one also affects the other. Microphone techniques are available to deal with noisy environments, and the recordist should always feel able to ask the contributor to speak up (see Low Voice below).

Applause/laughter masks voice: in 3 programmes (17%), live audience applause or laughter continued at a sufficiently high level so as to mask the start of the following contribution. In one programme the general level of the audience effects were higher than the presenters.

(Programmes affected: John Bishops Britain, University Challenge (only one instance), Have I got a bit more News for You).

Discussion: In a show with a live audience, audience reaction can be an essential part of the viewer's enjoyment. Skilled artists will allow time for the audience reaction to subside before continuing, but with a fast moving show the audience needs to be carefully managed, and the level controlled so that words are not missed. The studio audience is listening to foldback of the contributors and has only the live effects which surround them. They may hear the action better than the viewing audience. In the programmes reviewed the audience reaction over spilled the comment and was at too high a level. Some audience spill may be on the presenters mic. Close microphone placement with careful foldback levels and multi track recording allowing an audio dub, can eliminate some of these problems.

Artist Related.

Low voice: 7 programmes (39%) were identified as having artists or contributors whose vocal performance was at too low a level to be picked up adequately by the microphone. 6 (30%) instances were female, 4 (22%) instances were male; 3 (17%) programmes had both male and female artists giving low voice performance. In one instance this resulted in a large mismatch between the voice levels during an exchange between a male and female actor. (Programmes affected: Casualty, Weakest Link, Five Days (Ep 1), The Deep, Vexed, Crime Scene CSI, Our Drugs War).

Regional or national accent: 4 programmes (22%) had regional or national (American) accents which did not cause the reviewers undue difficulty but might have been a problem to some viewers. (Programmes affected: Coast, Crime Scene CSI, Raol Moat: Inside the Mind of a Killer).

Rapid speech: 3 programmes (17%) were identified as having artists or contributors who spoke so rapidly that it was difficult to follow what they were saying. They were all female. (Programmes affected: Coast, This Morning Saturday, The Bill: Tombstone).

Poor diction: 3 programmes (17%) had actors whose poor diction made it difficult to follow what they were saying. (Programmes affected: Coast, The Bill: Tombstone, Our Drugs War).

Muttered comment/ mumbling whispered asides: in 2 programmes (11%) the actors mumbled, made muttered comment or whispered asides that could not be understood and were part of the plot. One programme had both whispered asides and muttered comment. (Programmes affected: Casualty, Crime Scene CSI).

Discussion: A large proportion of the programmes suffered from defects which stem from the artists performance. With some programmes this applied to particular actors for the majority of their performance. Sometimes this can be corrected in the dub, but often correction is a compromise and is not always possible. In the majority of these instances it is likely that this was realised at the time of recording but no one seems to have felt able to request a better performance.

It has been suggested that there is a school of thought that it is not necessary for an actor to project when working in television. This is not true. Vocal performance is necessary if the voice is to be separated from extraneous noise (from clothing, footsteps, props, etc.).

With pressures on time and budget, rehearsal and setup time is limited and work may have to be done under less than ideal conditions. It takes courage to request a retake, particularly as many recordists are freelance. Producers need to be more aware of vocal performance and feel able to ask an actor to project, or ask the sound supervisor if other measures are possible. Sound supervisors need to feel sufficiently comfortable to request a retake; it is not their fault. The sound department will have some very talented and expert listening skills with high quality listening devices. Directors Editors and Producers will rarely comment on sound, but often on pictures. They need to be more aware that it is in their interest to have the best recorded sound.

Lessons Learnt.

1. Percussive music will punch through dialogue and must be added at a lower level than might be expected from its apparent loudness.
 2. Care needs to be taken with music having a significant content in the speech frequencies as it will clash with the voice and have a masking effect greater than may be expected from its level.
 3. Where music is used as a substitute for synchronous effects, care needs to be taken that it is appropriate in type and level. Inappropriate music will annoy the viewer and make them
- The Audibility of Speech on Television

less sympathetic if dialogue gets difficult to understand. Not using incidental music as “wallpaper” must be a serious consideration.

4. The level of theme or signature music tends to set the listening volume for the remainder of the programme. If its loudness is relatively high this may give difficulties with the remainder of the programme.

5. A qualified sound recordist should always be used. If one is not available, the person responsible for sound must have training in basic sound recording techniques but the end result is unlikely to be as good as if a professional had been used.

6. Placing of personal microphones under clothing should be avoided at all costs, but when it is the only practical solution (eg drama) a qualified sound recordist must be employed.

7. Camera microphones should only be used for synchronous effects. A separate, high quality, microphone should be used for dialogue.

8. Where a live audience is present, the audience effects must be at a lower level than the contributors and must be managed to avoid overspill.

9. When an artist or contributor is giving a quiet vocal performance the producer should take the initiative in requesting more volume. He/she should coordinate with the sound supervisor in seeking a satisfactory solution.

10. The sound supervisor must feel able to request a retake if the artist or contributors performance is not adequate. More people in the audio chain need to feel able to raise comment or feedback.