

ITC Guidance  
on  
Standards for Subtitling

February 1999

<b>INTRODUCTION .....</b>	<b>3</b>
<b>FOREWORD .....</b>	<b>4</b>
1 GENERAL REQUIREMENTS FOR SUBTITLE DISPLAY .....	6
1.1 <i>Basic Text Display</i> .....	6
1.2 <i>Colour</i> .....	6
1.3 <i>Control Characters</i> .....	6
1.4 <i>Formatting</i> .....	7
1.5 <i>Line Breaks</i> .....	8
1.6 <i>Positioning Subtitles on the Screen</i> .....	10
1.7 <i>Timing and Synchronisation</i> .....	11
1.8 <i>Leading and Lagging</i> .....	11
1.9 <i>Shot Changes</i> .....	12
2 SPECIAL TECHNIQUES .....	12
2.1 <i>Emphasis and Phrasing</i> .....	12
2.2 <i>Tone of Voice</i> .....	13
2.3 <i>Speaker Identification</i> .....	13
2.4 <i>Off-Screen and Off-Camera Voices</i> .....	14
2.5 <i>Dialogue Techniques</i> .....	14
2.6 <i>Sound Effects</i> .....	15
2.7 <i>Music</i> .....	16
2.8 <i>Silence</i> .....	16
2.9 <i>Failure of Subtitles</i> .....	17
2.10 <i>Acquired Programmes</i> .....	17
2.11 <i>Other Techniques</i> .....	17
3 SUBTITLING FOR THE INTENDED AUDIENCE .....	19
3.1 <i>Introduction</i> .....	19
3.2 <i>Children's Subtitles</i> .....	19
4 REAL-TIME SUBTITLING .....	22
4.1 <i>Text Entry</i> .....	22
4.2 <i>Advance Preparation</i> .....	23
4.3 <i>Choosing Shortforms</i> .....	24
4.4 <i>Subtitle Composition</i> .....	25
4.5 <i>Subtitle Presentation</i> .....	25
4.6 <i>Guidelines for Real-time Subtitling</i> .....	26
5 DIGITAL SERVICES .....	28
6 OTHER MATTERS .....	29
6.1 <i>National Emergencies</i> .....	29
6.2 <i>Apology Captions</i> .....	29
6.3 <i>Further Information</i> .....	29

## INTRODUCTION

The Broadcasting Act 1990, Section 35, requires Channel 3 licensees to provide minimum amounts of subtitling for deaf and hard-of-hearing people and to attain such technical standards in the provision of subtitling as the ITC specifies. These requirements are reflected in the licences.

This document provides guidance on the technical standards which are to be attained in the production and presentation of “closed” subtitles. It does not apply to “open” or “in-vision” captions. The guidance has been drawn from the experience gained in the production and usage of subtitles since the early 1980s. Much of the detail has been taken from an earlier handbook produced by the Independent Broadcasting Authority following research conducted at Southampton University. The ITC is grateful to the authors of that handbook for updating their work and for giving permission to use it. Further information on the use of subtitles has also been included as a result of a research project conducted on behalf of the ITC and BBC during 1990 and 1991 by Bristol University and the ITC research into Subtitling for Deaf Children 1996.

This document has been revised to take account of digital television services and acquired programmes. Further amendments may follow in the light of new techniques becoming available. In this respect investigative work is being conducted in the fields of veiling (see paragraph 5.9) and use of colours.

**To assist licensees and others reading this document, the specific rules and main points of guidance have been printed in bold type. These, however, cannot be properly understood in isolation from the rest of the text. Licensees are urged, therefore, to study the entire document carefully.**

The technical standards for subtitle transmission are covered separately by the ITC Technical Performance Code. The current UK standard for subtitle transmissions via teletext is ITU-R (CCIR) Teletext System B.

## FOREWORD

Subtitle editing is different from most other forms of text editing. Not only is the medium more dynamic than most others, but also the target audience is not a homogenous group. Although the common voice of deaf and hard-of-hearing people is that they require access to television, they vary in terms of literacy, hearing loss and socio-economic factors. With the increasing size of an aged population, the preferences of the elderly in particular should be taken into account.

The following are some of the factors which contribute to the distinctive nature of subtitling:

- i) Comprehension and enjoyment of normal text are enhanced for the reader by the ability, if necessary, to re-scan the article. Comprehension of a television programme, however, is typically gained only at the time of viewing.
- ii) Readers normally have only text to absorb. By contrast the user of subtitles must take in simultaneously the action within the television image as well as the information provided by the subtitling.
- iii) The pace of programmes sometimes means that subtitles cannot reasonably be expected to convey the full range of information contained in the television image. However, deaf and hard-of-hearing people naturally expect to receive as much as possible of the information which is available to the general audience.

Careful and sensitive editing is therefore needed in order to produce subtitles which will suit the intended audience, while still conveying the full meaning of the dialogue or commentary within the limitations set by the pace of the programme.

### Subtitling Priorities

The priorities for effective subtitling can be summarised as follows:

- 1) Allow adequate reading time. (See **1.7 Timing and Synchronisation**).
- 2) Reduce viewers' frustration by:
  - a) attempting to match what is actually said, reflecting the spoken word with the same meaning and complexity; without censoring
  - b) constructing subtitles which contain all obvious speech and relevant sound effects; and
  - c) placing subtitles sensibly in time and space.
- 3) Without making unnecessary changes to the spoken word, construct subtitles which contain easily-read and commonly-used English sentences in a tidy and sensible format.
- 4) In the case of subtitles for children, particular regard should be given to the reading age of the intended audience.

The craft of the teletext subtitler is to meet these objectives while working within the constraints imposed by the teletext medium. The remainder of this document presents guidance on ways in which this can be achieved.

# 1 GENERAL REQUIREMENTS FOR SUBTITLE DISPLAY

## 1.1 Basic Text Display

Subtitle legibility studies result in the following requirements:

- i) **Teletext characters should be displayed in double height and mixed (upper and lower) case.**
- ii) **Words within a subtitle should be separated by a single space.**
- iii) **Text should normally be presented in a black box.** (See **2.3 Speaker Identification** and **2.6 Sound Effects** for other background colours).
- iv) **To aid readability, text can be justified left, centre or right depending on speaker position.** (See **1.4 Formatting** and **1.5 Line Breaks** for examples of justified text).
- v) **The standard punctuation of printed English should be used. Punctuation gives valuable clues to syntactic structure and must be carefully displayed in order to be effective.** One means of enhancing the effectiveness of punctuation is by the use of a single space before exclamation marks and question marks, after commas, colons, semi-colons and mid-subtitle full-stops, on both sides of dashes (but not mid-word hyphens), before opening brackets and inverted commas and after closing brackets and inverted commas.

## 1.2 Colour

The teletext specification currently used in the UK is limited to the availability of seven different text colours, including white; and eight different background (boxing) colours, including black and white. **For normal subtitling purposes, black background should be used. Some early teletext decoders do not display coloured background and instead default to black. Therefore, if coloured background is used, a text colour should be chosen which will remain legible on a black background.**

The majority of text/background colour combinations are not satisfactory for subtitling, being insufficiently legible. The most legible text colours on a black background are white, yellow, cyan and green. Use of magenta, red and blue should be avoided.

Of the combinations with coloured background, the most legible are blue on white, white on blue, red on white, white on red, cyan on blue and blue on cyan. Of these, white on red, white on blue and cyan on blue are preferable, because certain older decoders will reduce these combinations to highly legible white on black, or cyan on black.

The principal ways of using colour in television subtitling are discussed in Sections **2.3 Speaker Identification** and **2.6 Sound Effects**.

## 1.3 Control Characters

The use of double-height boxed coloured text generally requires six control characters in the teletext line, or eight control characters if coloured background is used. Thus, the maximum space available for subtitle text is only 32 or 34 characters per line.

## 1.4 Formatting

A maximum subtitle length of two lines is recommended. Three lines may be used if the subtitler is confident that no important picture information will be obscured. (See Section 1.6).

Ideally, each subtitle should also comprise a single complete sentence. Depending on the speed of speech, there are exceptions to this general recommendation, as follows:

- a) Real-time subtitling (see **Section 4**).
- b) Short sentences may be combined into a single subtitle if the available reading time is limited (see **Section 2.5**). Additional reading time is gained in this way because the viewer's gaze needs to be directed to the subtitle area only once, rather than two or three times if two or three short sentences are displayed on consecutive subtitles.
- c) Very long sentences which are too long to fit into a single two-line subtitle. There are two procedures for dealing with such cases:

### Example (i)

It may be possible to break a long sentence into two or more separate sentences and to display them as consecutive subtitles eg 'We have standing orders, and we have procedures which have been handed down to us over the centuries.' becomes:

We have standing orders  
and procedures.

They have been handed down to us  
over the centuries.

This is especially appropriate for 'compound' sentences, ie sentences consisting of more than one main clause, joined by coordinating conjunctions 'and', 'but', 'or';

This procedure is also possible with some 'complex' sentences, ie sentences consisting of a main clause and one or more subordinate clauses joined by subordinating conjunctions such as 'since', 'when', 'because', etc or by relative pronouns such as 'who', 'that': 'All we wanted was a quiet chat just you and me together, but you seemed to have other ideas.' becomes:

All we wanted was a quiet chat  
just you and me together.

But you seemed to have  
other ideas.

It is sometimes also possible to break single main clauses effectively into more than one subtitle; eg 'I saw a tall, thin, bearded man with the stolen shopping basket disappearing into the crowd.' becomes:

I saw a tall, thin, bearded man  
with the stolen shopping basket.

He disappeared into the crowd.

### **Example (ii)**

If such sentence breaking procedures are inappropriate, it might be necessary to allow a single long sentence to extend over more than one subtitle. In this case, sentences should be segmented at natural linguistic breaks such that each subtitle forms an integrated linguistic unit. Thus, segmentation at clause boundaries is to be preferred. For example:

When I jumped on the bus...

..I saw the man who had taken  
the basket from the old lady.

Segmentation at major phrase boundaries can also be accepted as follows:

On two minor occasions  
immediately following the war,...

..small numbers of people  
were seen crossing the border.

There is considerable evidence from the psycho-linguistic literature that normal reading is organised into word groups corresponding to syntactic clauses and phrases, and that linguistically coherent segmentation of text can significantly improve readability.

Random segmentation such as

*On two minor occasions  
immediately following the war, small...*

*..numbers of people, etc.*

must certainly be avoided.

In the examples given above, sequences of dots (three at the end of a to-be-continued subtitle, and two at the beginning of a continuation) are used to mark the fact that a segmentation is taking place. Many viewers have found this technique helpful.

## **1.5 Line Breaks**

Similar linguistic considerations should guide the subtitler in deciding how to format a single multi-line subtitle. Subtitle lines should end at natural linguistic breaks, ideally at clause or phrase boundaries. However, since the dictates of space within a subtitle are more severe than between subtitles, line breaks may also take place after a verb. For example:

We are aiming to get  
a better television service.

Line endings that break up a closely integrated phrase should be avoided where possible. For example:

*We are aiming to get a  
better television service.*

Similarly:

He said it would increase  
the number of shareholders.

*He said it would increase the  
number of shareholders.*

**Line breaks within a word are especially disruptive to the reading process and should be avoided.** Ideal formatting should therefore compromise between linguistic and geometric considerations but with priority given to linguistic considerations.

Line breaks must be carefully considered when using left, right and centre justification for speaker position. Justified subtitles should balance linguistic considerations with eye movement:

### **Example (i)**

Left, right and centre justification can be useful to identify speaker position, especially in cases where there are more than three speakers on screen. In such cases, line breaks should be inserted at linguistically coherent points, taking eye-movement into careful consideration. For example:

We all hope  
you are feeling much better.

This is left justified. The eye has least distance to travel from hope to you.

We all hope you are  
feeling much better.

This is centre justified. The eye now has least distance to travel from are to feeling.

We all hope you are feeling  
much better.

This is right justified. The eye has least distance to travel from feeling to much.

### **Example (ii)**

Problems occur with justification when a short sentence or phrase is followed by a longer one. In this case, there is a risk that the bottom line of the subtitle is read first.

Oh.  
He didn't tell me you would be here.

Oh.  
He didn't tell me you would be here.

This could result in only half of the subtitle being read. Allowances would therefore have to be made by breaking the line at a linguistically non-coherent point:

Oh.  
He didn't tell me you would be here.

Oh. He didn't tell me  
you would be here.

Oh. He didn't tell me you would be  
here.

When the subtitler is forced to make a choice between formatting a subtitle into one long line or breaking it into two short lines, the decision should be made on the basis of the background picture. In general, 'long and thin' subtitles are less disruptive of picture content than are 'short and fat' subtitles, but this is not always the case.

Furthermore, in dialogue sequences it is often helpful to use horizontal displacement in order to distinguish between different speakers (see **Section 2.3**). 'Short and fat' subtitles permit greater latitude for this technique.

## 1.6 Positioning Subtitles on the Screen

The normally accepted position for subtitles is towards the bottom of the screen, but in obeying this convention it is most important to avoid obscuring 'on-screen' captions, any part of a speaker's mouth or any other important activity. Certain special programme types carry a lot of information in the lower part of the screen (eg snooker, where most of the activity tends to centre around the black ball) and in such cases top-screen positioning will be a more acceptable standard.

Subtitles should be displayed horizontally in the direction of the appropriate speaker, or source of sound effect (See 2.3 and 2.6).

When consecutive subtitles have boxes of similar size and shape and the second directly over-writes the first, it is useful to position them slightly differently on the screen. This makes it easier for the viewer to perceive that the subtitle has changed.

## 1.7 Timing and Synchronisation

It is crucial that subtitles are displayed for a sufficient length of time for viewers to read them. The subtitle presentation rate for pre-recorded programmes should not normally exceed 140 words per minute.<sup>1</sup> In exceptional circumstances, for example in the case of add-ons, the higher rate of 180 words per minute is permitted.

Presentation rates will depend upon the programme content. For example, real-time subtitling documentaries where the speaker is not on screen, or chat shows which have a higher text complexity than drama.

A fundamental function of television subtitling is to reduce frustration caused to hearing-impaired viewers by being faced with silent moving mouths. Therefore, all obvious speech should have some form of subtitle accompaniment.

Eye movement research shows that hearing-impaired viewers make use of visual cues from the faces of television speakers in order to direct their gaze to the subtitle area. If no subtitle is present, the resulting 'false alarm' causes considerable frustration. Further research into eye movement has shown the following pattern developed by hard of hearing viewers:

- i) Change of subtitle detected
- ii) Read subtitle
- iii) Scan picture until another subtitle change is detected.

Therefore, subtitle appearance should coincide with speech onset. Subtitle disappearance should coincide roughly with the end of the corresponding speech segment, since subtitles remaining too long on the screen are likely to be re-read by the viewer, ie another kind of 'false alarm'.

The same rules of synchronisation should apply with off-camera speakers and even with off-screen narrators, since viewers with a certain amount of residual hearing make use of auditory cues to direct their attention to the subtitle area.

## 1.8 Leading and Lagging

The target point for synchronisation should be at naturally occurring pauses in speech-sentence boundaries, or changes of scene. However, there are bound to be cases where this is either impractical or inapplicable. Recent research indicates the following:

- i) Monologue Material For hard-of-hearing people viewing programmes which consist mainly of monologue, research has shown that perfect synchronisation is not an absolute necessity and delays of up to six seconds do not affect information retention. The same is true of leading subtitles (providing that the first subtitle of a long speech is in synchrony). It should still be recognised, however, that some viewers use subtitles to support heard speech and will require synchronisation. Therefore, the technique should not be over used.

---

<sup>1</sup>140wpm (138wpm) corresponds to 690 characters per minute and uses 2 seconds and 15 frames per line. These speeds are increased to 180wpm when add-ons are used (increasing the reading speed by one quarter again).

- ii) Dramatic Scenes
- iii) For drama and programmes with continuous changes of shot, subtitles which lag behind dialogue or commentary by more than two seconds should be avoided.

## 1.9 Shot Changes

Besides the general recommendation for subtitle/speech synchronisation, there are certain other aspects of the television picture which influence subtitle timing. **Subtitles that are allowed to over-run shot changes can cause considerable perceptual confusion and should be avoided.** Eye-movement research shows that camera-cuts in the middle of a subtitle presentation cause the viewer to return to the beginning of a partially read subtitle and to start re-reading. In practice, it is recognised that the frequency and speed of shot changes in many programmes present serious problems for the subtitler. A subtitle should, therefore, be 'anchored' over a shot change by at least one second to allow the reader time to adjust to the new picture. Shot changes normally reflect the beginning or end of speech. The subtitler should, therefore, attempt to insert a subtitle on a shot change when this is in synchrony with the speaker.

General rules for dealing with camera-cuts are as follows:

- i) Avoid inserting a subtitle less than one second before a camera-cut and removing a subtitle less than one second after a camera-cut.
- ii) Attempt to insert a subtitle in exact synchrony with a camera-cut.
- iii) A decision to segment a single sentence into more than one subtitle, to be placed around a camera-cut, should depend on whether the sentence can be segmented naturally and on whether the resulting subtitles can be allowed sufficient display time.

Camera fades and pans do not produce the same perceptual effect as camera-cuts, and accordingly need not influence the subtitler in the same way.

Major scene changes can cause the same problems as shot changes within a scene. A particular difficulty arises when a speaker's last line in a scene, especially a vital punch line, is followed instantaneously by a scene change. In this case, the subtitle should be removed before the scene change to avoid visual confusion.

Some film techniques introduce the soundtrack for the next scene before the scene change has occurred. If possible, the subtitler should wait for the scene change before displaying the subtitle. If this is not possible, the subtitle should be clearly labelled to explain the technique.

JOHN: And what have we here?

## 2 SPECIAL TECHNIQUES

### 2.1 Emphasis and Phrasing

Text in upper case characters can indicate an increase in volume, while emphasis of an individual word can be achieved by a change in colour.

i) Overlays

ii) Subtle nuances of phrasing are difficult to deal with effectively, but special techniques can be used, for example:

No...No... But that isn't what I asked for.

A more powerful effect is achieved by the 'overlay' technique. This involves, for instance, turning the above example into two subtitles, by first displaying No ... and then adding the second part of the utterance after the pause and without deleting the No... . This dynamic method of simulating speech timing and phrasing can be very effective, but should be reserved for time and space emergencies because multiple overlays can result in jerky presentation and clogged screens.

## 2.2 Tone of Voice

There are no adequate resources for portraying tone of voice in teletext subtitles. Hearing-impaired people necessarily make use of facial cues in day-to-day communication, and this is a further important reason for allowing the viewer time to read each subtitle and to watch the associated picture. The devices mentioned in **Section 2.1** for expressing emphasis and phrasing can provide additional help.

Where tone of voice is particularly critical to meaning, and facial expression and body language are inadequate to convey the tone, the use of '(!)' and '(?)' immediately following speech can indicate sarcasm and irony as shown below:

No, no. You're not late (!)

## 2.3 Speaker Identification

People who have gained experience with teletext subtitles find that the use of colours to identify individual speakers is helpful. Suitable colours are recommended in **Section 1.2**. Where possible, therefore, each speaker should be identified by a single colour consistently throughout the programme.

Alternatively, programmes can be treated in terms of scenes. Seldom does any single scene on television include more than about four speaking characters. Thus, particular text colours can usually be associated with separate speakers within any single scene.

It is also helpful to use screen position and justification to support speaker identification. Each subtitle can be displaced horizontally towards the appropriate speaker. The main problem here is when characters move about while speaking. In such cases, the caption should be positioned at the discretion of the subtitler to identify the position of the speaker as clearly as possible.

A blue background with white text can also be useful to indicate a different quality of voice such as a robot or ghost.

## 2.4 Off-Screen and Off-Camera Voices

**When the source of off-screen/off-camera speech is not obvious from the visible context, special techniques should be used.** The best way to determine the effectiveness of visible context is to watch the scene without sound and then insert arrows to indicate the off-camera speaker. These can be either the teletext left-pointing and right-pointing arrows, or the visually more effective combination of the ‘greater than’ (>) or ‘less than’ (<) symbols.

When off-screen speech is employed throughout the programme, eg as in narrative documentaries, the common approach is to centre subtitles without symbols.

Other situations where the source of speech is not immediately apparent include telephone voices, radios, tannoy announcements, etc. It is helpful to accompany the first subtitle from these sources with a labelled (See **2.9 Other Techniques**). caption:

LOUDSPEAKER:  
"Fasten your seat-belts, please."

Character-name labels are sometimes necessary for clarification, eg for crowd scenes or scenes enacted in the dark.

JOHN: What's happened to the lights?

## 2.5 Dialogue Techniques

All obvious speech should be accompanied by subtitle information, but under conditions of rapid dialogue, several short subtitles displayed in rapid sequence can result in staccato or ‘machine-gun’ effect. There are two possible solutions for this:

### i) Double text

Here the speech of two different speakers is displayed simultaneously. However, double-text can be an infringement of the synchronisation recommendation: the appearance of each subtitle should coincide with the beginning of the corresponding speech segment. Using double text may also result in the extension of the black box which could obscure other visual information.

### **Double text techniques**

Double text is normally used when more than two characters speak simultaneously and contradict one another:

Have you had lunch?

Yes.                      No.

Both people's speech is contained within one subtitle.

### ii) Add-ons or cumulative titles

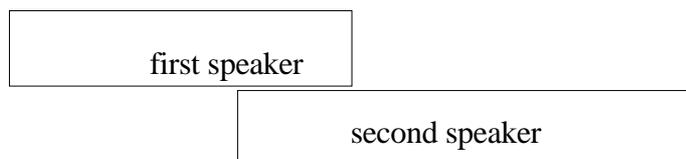
The second solution is to use ‘dynamic’ text known as ‘add-ons’ or ‘cumulative titles’. This is most effective when the two subtitles fit naturally together, for example in a question and answer sequence, or providing the punch line of a joke. In this technique the second part of the title is added on to the first part. Thus, the appearance of the second part, or ‘add-on’ of the title can coincide with the onset of the second utterance, while the subtitle corresponding to the first utterance remains on screen. In this way, the staccato effect is diminished while still preserving a natural relationship between speech onset and subtitle presentation.

A further advantage of add-ons is that they appear more natural when the two corresponding speakers are not shown in the same camera shot.

However, additional reading time is gained only if the gaze of the viewer remains on the subtitle area throughout its presentation. If the gaze of the viewer returns to other parts of the picture before the add-on appears, as can occur among unusually fast readers, the subsequent add-on can cause some perceptual confusion. In particular, the first part of a cumulative title may be re-read, resulting in a frustrating ‘false alarm’. For this reason, cumulative titles or add-ons should be used with care.

### **Add-on techniques**

In both double-text and add-ons, the second part should normally appear on the line immediately beneath the first part. If the second speaker is positioned to the screen right of the first speaker, then the second part of the double-text is displaced towards the right:



A general guide to the use of double-text and add-ons may be stated as follows:

- i) Double-text can be used when two characters or more speak simultaneously.
- ii) Add-ons should normally be preferred when two or more characters speak consecutively and time does not allow individual subtitles.
- iii) The total length of either double-text or add-on sequence should never exceed four lines.

### **2.6 Sound Effects**

Any relevant sound effect not immediately obvious from the visual action should be subtitled. This includes sound effects that become apparent in the subsequent action, eg the telephone ringing before it is picked up, an explosion occurring outside before everyone dives under the table.

Descriptive statements are normally preferable to onomatopoeic spellings for sound effects. But context and genre (cartoons versus drama, for example) must be taken into consideration.

GUNSHOT is usually preferable to BANG!!!

Sound effect subtitles can also be used judiciously to create the background atmosphere for a scene:

ROAR FROM AUDIENCE

LIVELY CHATTER

The use of background colour (eg white text on a red background) and upper-case text, provide a distinction between sound effect subtitles and speech subtitles.

In general, each sound effect subtitle should be displaced towards the source of sound. Flashing text can be used to direct attention towards particularly important sound effects, but this should be done only if the subtitle is to be displayed for more than three or four seconds. The importance of particular sound effects can best be determined by watching the programme without sound.

## **2.7 Music**

At the very minimum, the title of the music playing should be given. Where possible the words of a song should be included. This is especially important where the programme is to be viewed by younger people. Pop programmes, opera and songs connected to the story line are particularly important areas.

Song lyrics should be subtitled verbatim; but, if the pace of the song is very rapid, whole couplets or verses may be omitted.

The lyrics of a song should be made obvious by means of a sign. The ‘hash’ or ‘sharp’ sign is most commonly used. Where speech and song are interspersed, care should be taken to signpost each title correctly.

Provision of an occasional subtitle for mood music, if it is significant to the plot, can be very effective:

# FRENCH PROVINCIAL MUSIC

Such subtitles should be used only sparingly.

Occasionally, consecutive scenes are enacted in pitch darkness, and scene changes are signalled entirely by changes of incidental music. In such cases, if time permits, the subtitler should use subtitles such as:

# LIVELY DANCE BAND MUSIC

Then, when the tempo of music changes dramatically, it is followed by:

# MOVES INTO SLOW DANCE MUSIC

Thereby deaf viewers are made aware of the scene change.

## **2.8 Silence**

Long speechless pauses in programmes can sometimes lead the viewer to wonder whether the teletext system has broken down. It can help in such cases to insert an explanatory caption such as:

INTRODUCTORY MUSIC

LONG PAUSE

ROMANTIC MUSIC

## **2.9 Failure of Subtitles**

Losing subtitles is as frustrating for the hearing-impaired viewer as losing sound is for the hearing viewer.

**If subtitle insertion does fail, prompt transmission of an appropriately worded apology caption is extremely important and, if restoration of transmission is delayed, an early explanation is to be given.**

## **2.10 Acquired Programmes**

Trials in the conversion of subtitles, which have accompanied certain acquired programmes from the U.S.A., have been shown to produce acceptable results. It is possible to improve the subtitles to a satisfactory standard at reduced cost compared to the provision of new subtitles. **The conversion process must ensure that the subtitles achieve the minimum standards set out in these guidelines except in two areas where the ITC has agreed a limited degree of flexibility, as follows:**

- a) **Synchronisation – the requirements in section 1.7. do not have to apply, however gross synchronisation errors between dialogue and visual material must be eliminated.**
- b) **Presentation rate – however excessively fast dialogue must, where practicable, be edited to within the maximum limit set in section 1.7. (i.e. 180 words per minute).**

## **2.11 2.11 Other Techniques**

- i) Single quotes ‘...’ can indicate non-synchronous speech, eg a voice-over or thoughts or nothing on screen visibly connected with speech.
- ii) Unmarked text in upper and lower case indicates synchronous speech, ie the speaker is visible (most of the time) and titles follow.
- iii) Double quotes “...” can suggest mechanically reproduced speech, eg radio, loudspeakers etc or a quotation from a person or book.
- iv) Text in brackets can indicate whispered speech or asides.
- v) Brackets can also be used to indicate the way in which a person speaks:

(SLURRED) He wasn't there.

- vi) The source of speech can be labelled by using capital letters followed by a colon:

LOUDSPEAKER:

"Fasten your seat-belts, please."

JOHN: What's happened to the lights?

- vii) Uncommon abbreviations, such as SFX, should be avoided.

### **3 SUBTITLING FOR THE INTENDED AUDIENCE INCLUDING CHILDREN**

#### **3.1 Introduction**

The typical pace and complexity of subtitling can exclude a minority of less able readers within the deaf community. For many pre-lingually deaf children, preliminary experiments suggest that a subtitle presentation rate of 70-80 words per minute is appropriate. Similarly, the subtitler may consider adjusting the speed of presentation if the programme is primarily intended for profoundly deaf viewers.

Care however should be taken not to give the impression of spoon feeding, as this can be resented. In these instances these three main editing devices should be remembered:

- i) Reduce the amount of text by reducing the reading speed and removing unnecessary words and sentences.
- ii) Represent the whole meaning.
- iii) Increase the use of three-line subtitles and reduce the number of add-ons.

In case i), the subtitler must choose which sentence to present verbatim to the exclusion of other sentences. In case ii), the analysis of the whole meaning is made and the subtitler should then attempt to retain as many 'idea units' as possible. (See **Section 4.4**). For case iii) the use of three-line captions allows the material to be on screen for a longer period, thus helping the development of scanning techniques. These guidelines represent a level of subtitling suitable for the 'average' hearing-impaired viewer.

Subtitles should accurately reflect the spoken word and as such should not be censored.

#### **3.2 Children's Subtitles**

Special care is needed for the provision of children's subtitles.

Many deaf children over 11 years benefit from subtitles as they are currently provided. Recent ITC research has shown, however, that children under the age of 11 years need simpler subtitles.

The following guidelines are recommended for the subtitling of programmes targeted at children below the age of 11 years.

- i) There should be a match between the voice and subtitles as far as possible.
- ii) A strategy should be developed where words are omitted rather than changed to reduce the length of sentences.

For example:

**Dialogue:** Can you think why they do this?  
**Simplified:** Why do they do this?

- Dialogue:** Can you think of anything you could do with all the heat produced in the incinerator?  
**Simplified:** What could you do with the heat from the incinerator?

Difficult words should also be omitted rather than changed.

For example:

- Dialogue:** First thing we're going to do is make his big, ugly, bad-tempered head.  
**Simplified:** First we're going to make his big, ugly head.
- Dialogue:** All she had was her beloved rat collection.  
**Simplified:** She only had her beloved rat collection.

- ii) Where possible the grammatical structure should be simplified while maintaining the word order.

For example:

- Dialogue:** You can see how metal is recycled if we follow the aluminium.  
**Simplified:** See how metal is recycled by following the aluminium.
- Dialogue:** We need energy so our bodies can grow and stay warm.  
**Simplified:** We need energy to grow and stay warm.

- iii) Difficult and complex words in an unfamiliar context should remain on screen for as long as possible. Few other words should be used.

For example:

- Dialogue:** Nurse, we'll test the reflexes again.  
**Simplified:** Nurse, we'll test the reflexes.
- Dialogue:** Air is displaced as water is poured into the bottle.  
**Simplified:** The water in the bottle displaces the air.

- iv) Care should be taken that simplifying does not change the meaning, particularly when meaning is conveyed by the intonation of words.
- v) Numbers, including price, weight etc are easier than words to read and remember. They should, therefore, be presented in numerical form.

For example:

**Dialogue:** It takes about one kilojoule of energy to lift someone off their feet.

**Simplified:** It takes 1kJ of energy to lift someone.

**Dialogue:** We took a gram of margarine and placed it in a dish.

**Simplified:** We put 1 gram of margarine in a dish.

### **Additional Note For Schools Programmes**

Often, the aim of schools programmes is to introduce new vocabulary and to familiarize pupils with complex terminology. There are a number of options when it comes to subtitling schools programmes:

- i) Introduce complex vocabulary in very simple sentences and keep it on screen for as long as possible.
- ii) Provide teachers with scripts before the programme is broadcast, so that they can introduce and explain any new vocabulary beforehand.

## 4 REAL-TIME SUBTITLING

The production and transmission of subtitles in real time can present considerable problems for both the subtitler and the viewer. Current subtitling techniques, particularly for live broadcasts, do not provide the same high-quality service expected from pre-prepared scripts. **Such techniques should be limited to occasions when there is insufficient time to prepare subtitles using other methods.**

The use of scrolling word-by-word captions for live news can present particular difficulties, especially in the area of retention of information. It is not just a problem of speed, but that the text placed upon moving lines is working against the readers' natural reading strategy. Here the relationship of the displayed text to particular screen shots is uncontrolled and errors can occur in separating adjectives from nouns and articles from nouns.

During an unscripted live broadcast, however, the subtitles must be composed, entered, formatted and transmitted in a single pass through the programme. It may be necessary to subtitle one or more speakers who are delivering information at speeds up to 200 words per minute. In practice, the problems are usually less severe. For 'live' programmes, advance preparation may be possible. Portions which are scripted can have subtitles prepared in advance for manual cueing during transmission. Narrative-type presentations of public events and sports commentaries lend themselves particularly to the shortform approach. Programmes such as the News are less easy due to the large number of proper nouns that could be used, most of which may occur only once or twice in the bulletin. Background research can enable the subtitler to anticipate the likely subject matter, and then to take advantage of techniques such as 'shortform' abbreviations to speed up text input (see **Section 4.3**).

Automatic formatting, boxing and positioning of subtitle text is almost essential during real-time subtitle preparation. Without such facilities, the text throughput rate is considerably slowed. If prepared subtitles are being manually cued, it is also essential to have rapid random access to any point in the subtitle list should the running order be altered. A means of rapid text input is also required. The remaining problems for the subtitler concern the choice of subtitling strategy appropriate for the type of programme. Under truly live conditions, this will involve real-time subtitle composition and text entry. During real-time programmes it is possible to employ two or more operators working together. The technique is known as Dual Inputting.

Although not classified as 'live' subtitling, advances in newsroom technology have led to the development of software which allows the formation of subtitles direct from the newsroom computers. Any necessary editing and rearrangement of running order can be undertaken on a separate personal computer prior to transmission. However, it is difficult to achieve the high standards required by these guidelines using this form of inputting. It should, therefore, only be used in extreme emergencies.

### 4.1 Text Entry

Three methods of text entry have been used for real time subtitling. The first two employ fast-keying techniques. Such systems are designed to produce verbatim transcripts, this can be inappropriate for television subtitling if the word rate is very fast. The operators must therefore be retrained to edit the soundtrack, or to work in conjunction with an editing

interpreter. A more serious disadvantage is that deaf viewers can be confused by inaccuracies in the spelling of the subtitle output. This problem is likely to reduce as the development of transcription systems continues. It is desirable to integrate the machine-shorthand writer into the pre-production process, for example by making scripts available for preview. Lastly, of course, the technique relies on the availability of trained operators.

i) Phonetic Keyboards

The first involves a special phonetic keyboard designed for verbatim transcription, such as the Palantype system or the Stenograph system. A trained operator uses the keyboard to enter a series of phonetic codes representing speech, and a computer decodes this information to produce, so far as possible, a conventional transcription. Due to ambiguities in the phonetic coding, operator errors and spelling complexities, the spelling and word-boundary identification in the transcribed output is not always accurate. Depending on the size of the dictionary in the transcription computer, and on the error correction techniques available, an average output accuracy of between 75 per cent and 95 per cent is generally achieved, at speeds of up to about 200 words per minute.

ii) Velotype Keyboards

The second transcription method uses the Velotype syllabic chord keyboard, which can attain a speed of around 100-140 wpm with a trained operator. It is lower cost than the phonetic machine shorthand systems and does not require extensive dictionaries, although shortforms are essential where unusual spellings are likely to occur.

iii) Qwerty Keyboards

The third method of text entry for real time subtitling uses an ordinary Qwerty keyboard as an alternative to machine shorthand. The problems of non-standard spellings are largely overcome (except for occasional typing errors), but the maximum rate of text input is much reduced. A maximum subtitling rate of about 80 wpm is typical. The use of shortforms provides a valuable means of speeding keyboard input and reducing the likelihood of spelling errors.

## **4.2 Advance Preparation**

Having decided that a particular live programme is to be subtitled, it is necessary to choose an appropriate strategy.

The choice is based on an assessment of the likely format of the programme, the availability and reliability of scripts and the expected presentation speed. In most cases, a hybrid approach is probably necessary. This involves switching between manual cueing of prepared subtitles during scripted portions and live inputting.

Flexibility is an important feature of live subtitling equipment, since switching between different transmission modes must be achieved rapidly and straightforwardly.

It is helpful to draw up a running-order or programme plan based on information available in advance about the likely content of the broadcast. Liaison between the subtitling service and the programme production team is of value during the planning phase. Scripts available in

advance can be edited and typed into the subtitling system for storage in memory and/or on disk. Two points are worth noting here:

- i) Pre-stored subtitles should be accessible in groups chosen to distinguish subject boundaries. This makes random access easier should the running order change during the broadcast.
- ii) Pre-stored subtitles should be limited to two lines, since three-line texts may obscure foreground detail and there is little time to reposition subtitles during subsequent live cueing.

In addition to pre-stored subtitles produced from scripts, background research should enable a number of general 'fallback subtitles' to be prepared relating to the expected programme material. This approach involves making available a subtitled commentary which can be used independently of the soundtrack. Editorial discretion is required when integrating these standby subtitles with conventional commentary-based material.

### **4.3 Choosing Shortforms**

An additional and important aspect of advance preparation concerns a method for speeding keyboard input by using abbreviations, or shortforms, in place of words or phrases expected to occur in the broadcast. The technique was first developed to reduce the burden on the operator of a conventional qwerty keyboard when working under pressure.

Prior to transmission, the subtitler assigns two or three character shortform abbreviations or mnemonics to selected words or phrases. Experience indicates that proper nouns such as the names of people, places, buildings, bridges, boats etc can usefully be abbreviated in this way. When subtitling sport, terms relating to the particular game can also be stored.

Then, when a shortform is typed in an ordinary sentence, the subtitling computer automatically detects and expands it to its full form, using a predefined shortform dictionary. Such dictionaries can be stored on disk and recalled for later use.

Four guidelines have been found to be valuable when using shortforms:

- i) The shortforms should be chosen by the person who is eventually going to use them at the keyboard.
- ii) Where possible, some consistent abbreviation technique should be developed, eg taking the first three letters of single words, and the initial letters of multiple-word sequences. This assists in recalling the shortform, and may enable it to be deduced if forgotten.
- iii) It is important to ensure that no shortform can also be a valid word, otherwise there can be an erroneous expansion when shortforms and ordinary words are mixed.
- iv) An easily visible list of shortforms and their expansions should be available to the subtitler during the broadcast - for example posted on the wall. This makes it possible to look up quickly and check an item should it be necessary.

The advantage of the shortform technique is in its flexibility, since it is entirely up to the operator how the abbreviations are chosen and used. Considerable typing time is saved, and the possibilities of spelling error are reduced. For example:

thol = The House of Lords  
wbr = Westminster Bridge  
tbr = Tower Bridge  
ve = Victoria Embankment  
gm = Greenwich Meridian  
sou = Southwark

#### 4.4 Subtitle Composition

**The construction of subtitles for informative subjects such as news should convey the whole meaning of the material.** This need not mean using the same amount of words. Research into this area has described the concept of 'idea units'; that is where a proposition or key information is given. These units should be distinct with minimal repeats, and relate to the original information.

If the programme speech is too rapid either for the viewer or for the chosen means of text entry, the speech must be edited 'on the move' before entering it as a subtitle.

Such editing must be performed very rapidly to avoid long delays between speech onset and the appearance of a subtitle. This is a skilled task, and its degree of success depends on the type of programme and the editor.

Narrative-style programme commentaries given during major live outside broadcast events are relatively easy to edit in real time. The commentator is not visible, thus reducing synchronisation problems, and the salient points of what may be a leisurely commentary pace can readily be picked out and subtitled. In contrast, the subtitling of a live news broadcast presents severe difficulties. Information is presented in compact form, and the rate of delivery is usually rapid. In addition, it may be almost impossible to edit politically sensitive material without distorting it. Between these extremes are situations in which a trained editing interpreter can work with varying degrees of success.

#### 4.5 Subtitle Presentation

Each real-time subtitling system suffers from the problem that both composition and text entry impose a delay between the start of an utterance and the appearance of the corresponding subtitle. The delay varies from one to three seconds for verbatim phonetic machine shorthand, to around five seconds for edited input using qwerty. Delays can be reduced by 'scrolling' presentation methods, but in teletext this can be difficult to follow. This can cause difficulties to the viewer, especially if the programme soundtrack is in obvious synchrony with the visual material as, for example, during an on-screen interview.

##### i) Word-by-Word Display

Two methods of word-by-word displays are currently available: a screen which overwrites when it reaches the bottom line or a screen which scrolls, ie jumps up, pushing the top line of text out of the text window.

Although of value for live subtitling, the use of a word-by-word display can create problems for the reader because of the speed of speech output and possible confusion in eye-movement. Its advantage, however, is in the provision of near-verbatim text.

ii) Standard Format or Block Text

The subtitles are presented in complete phrases or sentences similar to those prepared subtitles associated with recorded programmes. Whilst the ‘on-screen’ appearance of this form is often slower because of the longer wait for complete syntactical sentences, the ability of the reader to flit from pictures to words assists certain deaf viewers in understanding the programme.

Current research indicates that both methods of live subtitling are accepted by approximately equal proportions of deaf viewers.

#### **4.6 Guidelines for Real-time Subtitling**

Early research indicates that first attempts were considered to be too fast. Although the rate of subtitling is driven at the rate of the presenter/journalist, there is still a need to focus carefully on reading variables. In such situations, although preparation time is limited, efforts must be made to adhere to at least the following:

- i) Subtitles should contain a reasonable percentage of the words spoken.
- ii) ‘Idea units’ or key facts should appear as a good percentage of the spoken message (see **Section 4.4**).
- iii) Avoid ‘idea units’ which are unnecessary or different from the original.
- iv) Where possible, avoid non-linguistic line breaks (splitting verbs etc).
- v) Attempt to avoid overrunning shot changes (synchronisation).
- vi) Where possible avoid dynamic displays; that is to say blocks are considered more acceptable than the scrolling/word-by-word format.

The following are offered as more detailed guidelines during the preparation of subtitles in real time:

- i) Maintain a regular subtitle output with no long gaps (unless it is obvious from the picture that there is no commentary) even if this means subtitling the picture or providing background information rather than subtitling the commentary.
- ii) Aim for continuity in subtitles by following through a train of thought where possible, rather than sampling the commentary at intervals.
- iii) Produce complete sentences even for short comments because this makes the result look less staccato and hurried.

- iv) Bear in mind that a subtitle specific to a particular scene can often be phrased sufficiently broadly to 'survive' a sudden camera cut without having to be abandoned. If pre-stored specific subtitles are used, ensure that they are cued at appropriate times.
- v) Send an apology caption following any serious mistake or a garbled subtitle; and, if possible, repeat the subtitle with the error corrected.
- vi) Do not subtitle over existing video captions where avoidable (in news, this is often unavoidable, in which case a speaker's name can be included in the subtitle if available).
- vii) Do not start subtitling 'cold'. A short rehearsal should be conducted just prior to transmission.

When cueing prepared texts for scripted parts of the programme:

- i) Try to cue the texts so that they closely match the spoken words in terms of start time.
- ii) Try to include speakers' names if available where in-vision captions have been obliterated.
- iii) Do not cue texts out rapidly to catch up if you get left behind - skip some and continue from the correct place.

## 5 DIGITAL SERVICES

The advent of digital television offers the possibility of visual and technical enhancement of subtitles. One such way is the introduction of a clearer and more legible font and the use of symbols. The purpose of these changes is to increase the user's enjoyment of television. They should not be introduced in a way which detracts from the prime purpose of the subtitling service, which is to facilitate the understanding of programmes by deaf and hard of hearing people.

The guidelines set standards on converting existing subtitle files for use on Digital Terrestrial Television (DTT). They also begin to set best practice for subtitles specifically prepared for DTT. It is not intended that these guidelines should inhibit future enhancements, but rather to control progress so that changes that deviate radically from current practice should be made only after careful consideration and consultation.

**The Tiresias font shall be used for all subtitles.**

**The nominal size of subtitles shall be 24 television lines for the capital 'V'.**

**For reasons of latency use of intensive four-line subtitles, to include those within "add-ons", shall be avoided. (Latency means the presentation of subtitles to the screen. Intensive subtitles mean those where the subtitle rows contain a considerable amount of text).**

**Subtitles converted from existing EBU 3264 files or from pre-recorded live 335 tapes shall as closely as possible retain the positioning and line breaks of the original.**

**The range of colours shall be limited to 12 that shall closely replicate the range associated with analogue teletext delivery.**

**Line breaks on conventional aspect ratio receivers (4:3) and widescreen (16:9) receivers must retain the original emphasis of the subtitle.**

**Subtitles shall be placed within the 'Safe Caption Area' of a 14:9 display. The location of text across the screen, which helps the identification of speakers, should remain intact.**

**The present practice of using '#' to indicate music shall be changed to use of two semi-quavers as part of the Tiresias set.**

It is proposed that the veiling of either foreground or background may be used to prevent completely obscuring large areas of the background picture. This is a particularly useful feature when it is difficult to position the subtitles without obscuring vital picture detail. However, the effect of reduced contrast, and possibly legibility, must be considered when introducing veiling. The amount of veiling, which might be used, is currently under review.

An italicised form of text may be used to indicate emphasis within a subtitle.

## **6 OTHER MATTERS**

### **6.1 National Emergencies**

Deaf and hard-of-hearing people need to be kept informed about national and local emergencies.

**When information about emergencies is being broadcast, the same information, including relevant telephone numbers, should be provided, preferably in open captions leaving sufficient time to write the details down.**

### **6.2 Apology Captions**

Where practicable an appropriate apology or explanation is to be transmitted as soon as possible after any loss of subtitles. A caption stating that “There are no subtitles for this programme” is not acceptable.

### **6.3 Further Information**

The ITC and BBC co-funded research into the use of subtitles was conducted by the Centre for Deaf Studies at Bristol University and the findings presented in the report ‘Switched On’.

Research on children’s subtitles, ‘Dial 888; Subtitling for Deaf Children’, was conducted by Dr Susan Gregory. Research on news subtitles, ‘Good News for Deaf People’ was conducted by IFF Research Ltd. Copies of findings are available from the ITC.