



*National Disability Coordination Officer Programme*  
AN AUSTRALIAN GOVERNMENT INITIATIVE

# Guidelines for Designing Accessible Instructional Media

Prepared by  
Len Bytheway  
*BTW Consulting*



---

Guidelines for Designing Accessible  
Instructional Media

First edition 2007  
Second edition 2015

A project funded by  
National Disability Coordination Officer Program

An initiative of the Australian Government

Prepared by  
Len Bytheway  
*BTW Consulting*

[www.btwconsulting.com](http://www.btwconsulting.com)

---

## Table of Contents

<b>Table of Contents</b> .....	<b>3</b>
<b>Preface to the Second Edition</b> .....	<b>5</b>
<b>Introduction</b> .....	<b>6</b>
<b>Accessibility Technologies</b> .....	<b>7</b>
Recording Video Content .....	7
Providing Access to Students who are Deaf or Hard of Hearing.....	7
Real-Time Captions.....	8
Open Captions .....	9
Closed Captions .....	10
Language modified captions .....	11
Live Speech Recognition Systems.....	11
Auslan in Videos.....	12
Access for Students who are Blind or have a Visual Impairment.....	13
Collaborative On-line Education Technologies .....	15
<b>Guide for Captions Styles and Presentation</b> .....	<b>16</b>
Choosing Accessibility Design Tools.....	16
How We Read Captions .....	16
Presentation of Captions .....	17
Number of Lines on the Screen.....	17
Line Breaks in Captions .....	18
The use of Punctuation in Captions .....	20
Choice of Font .....	22
Length of Caption Lines .....	22
What to Caption.....	23
Positioning Captions on the Screen .....	25
Caption Colours and Background .....	26
Colour Choice.....	26
<b>Creating Captions and adding them to Videos</b> .....	<b>29</b>
Sources of 'ready-made' captions .....	29
Material already captioned on the Internet. ....	29
Downloadable caption files.....	29
Extracting captions from a DVD. ....	30
Using .srt files with your video.....	30
Creating your Own Captions. ....	31
On-line captioning services. ....	31
Use YouTube to create or edit your captions.....	31
Use commercial or open source captioning software editors and encoding software.....	33
<b>Conclusion</b> .....	<b>35</b>
<b>APPENDIX A – Summary of Captioning Style Guidelines</b> .....	<b>36</b>

---

## Table of Figures

Figure 1. Program with open captions .....	9
Figure 2 Program with closed captions .....	10
Figure 3. Auslan interpretation added over the video.....	12
Figure 4. Another method of showing Auslan.....	13
Figure 5. Descriptive audio track.....	14
Figure 6. Too many caption lines .....	18
Figure 7. Title Safe Guides (as displayed in Final Cut Pro).....	25
Figure 8. White captions may be difficult to read.....	27
Figure 9. Black captions can also be difficult to read .....	27
Figure 10. Adding a text outline improves readability.....	27
Figure 11. Captions on a solid contrasting background are clear .....	28
Figure 12. Adding a transparent background adds image detail.....	28
Figure 13. YouTube caption editor workflow .....	32
Figure 14. Caption creation workflow .....	34

---

## Preface to the Second Edition

Since this guide was first published in 2007 there have been significant changes in the technologies available to create and deliver accessible instructional media accompanied by an increased awareness of the need for these media. For example the range and sophistication of caption creation tools has improved significantly. There is now a wide selection of Open Source freeware programs and commercial software that is now more useable and affordable.

The availability of YouTube caption creation and editing tools has brought captioning to the attention of many educators and within the grasp of non-expert content creators. This has been a double-edged sword. It is unfortunate that many tertiary educators believe that because YouTube offers an option to auto-generate captions in real-time that they have fulfilled their obligations to provide accessible resources. Anyone who has experienced the variable accuracy of these auto-created captions will understand that the high error rates render these captions unreliable and unhelpful. Put simply until the technology is radically improved auto-generated captions should not be relied on as a source of captioned media and it certainly does not meet accessibility standards.

There are many who advocate the use of YouTube's caption timing feature to convert audio transcripts of videos into closed captions. This has been lauded as a quick and simple way to get media captioned quickly. Auto-timed captions present the transcript text on the screen so that the voice and words are generally in synchronisation – job done? While it is true that timed captions are created, the visual presentation of these captions is not optimised for ease of reading and comprehension. Captions are broken according to line length rather than grammatical or meaningful caption breaks and reading them presents new challenges.

While this second edition updates techniques and tools used for captioning, the general caption style guidelines from the first publication are largely unchanged and still valid today. The introduction of easier to use captioning tools should also make well-designed and 'crafted' captions even more readily achievable.

As this edition goes to print new technologies will continue to emerge and improvements will continue to be developed to make these new technologies accessible. For that reason this guide has been kept largely technology neutral with suggestions about how to search for current versions rather than specific named applications. In my captioning practice I tend to use a mix of commercial and freeware captioning tools. It is probably advisable to try freeware first and only move to commercial versions if they offer additional and valued features. Often the paid products offer little more than the free ones.

As always the best way to judge the value of accessibility features is to consult with the end user, the person actually using your audio descriptions or captions and see what is important to them. Finally remember that one size does not fit everyone when it comes to accessibility requirements – your consumer consultation should seek a range of views and preferences.

---

## Introduction

Increasingly higher education is discovering new ways to deliver content. The traditional ‘stand-and-deliver’ lecture with overhead transparencies and white boards are being replaced by more media-rich content and delivery modes that might include on-line collaborative systems and media delivered via the Internet. This increased use of digital media has the capacity to enrich the learning environment, but it also has the potential to create even greater barriers for students with sensory disabilities. Through the creation of accessible instructional media the objective is to remove or reduce these barriers allowing full participation and equal educational access.

This guide provides an overview of some of the accessibility issues that confront many students in the tertiary education sector in Australia. It includes a guide to assist those producing or modifying educational media to improve access for all students, especially those with sensory disabilities.

It is often the case that implementing educational technologies that improve accessibility for students with a disability also makes them much more effective for all students to use. For example students whose first language is not English may improve their understanding if their educational media is captioned in English, they can see (read) and hear the content.

Unfortunately students with a disability often feel they are set apart from their peers as “accessible practices” may highlight their differences. Ideally practices used to provide educational access should offer minimal intrusion or disruption to the education process, remaining as transparent as possible.

Where possible mainstream technologies should be used to develop and deliver educational media. Where supplementary accessibility features need to be added, they should be set as options that the student could choose as they see fit.

---

## Accessibility Technologies

The ultimate objective of educational accessibility is that all educational instruction should be delivered in a form that allows all students to make full use without adaptation or modification – we should be aiming for *universal design*. Unfortunately this is often not the case and a large proportion of educational media are not created with accessibility features before they are distributed, requiring such features to be added post-production. This incurs additional cost and delays.

### Recording Video Content

Whenever content is created specifically for instructional purposes it is wise to consider and design the accessibility features at the time the audio and video is being captured and recorded. Some simple ideas could include:

- Ensure that any critical visual information is framed within the video so that future captions will not obscure important visual information. For example a sports science video that demonstrates a foot technique that cannot be seen once the captions are superimposed.
- Any critical information that relies solely on visual or auditory information will exclude either a blind or deaf viewer. If possible it is better to present the information both verbally and visually. For example imagine a graph of population growth, the lecturer could say “this graph proves my point”, or she could say “this graph shows how population growth has slowed in Australia since 1955, which proves my point.”
- Try and create links between audio and visual content so that viewers who only have access to one channel understand the context. For example a visual image of a river flowing with a voice over that talks about population growth, even if captioned accurately can be quite confusing. Don’t assume the metaphor will be obvious, the audio track might say “As a river flows from the mountains to the sea and grows larger along the way, so population growth changes gradually.”

As there is a vast amount of prepared educational media readily available in most cases instructors will use this material rather than create their own. This means we will need to adapt that media for accessibility

### Providing Access to Students who are Deaf or Hard of Hearing

In most interactive communication exchanges, such as information sessions, lectures and tutorials, the majority of information is conveyed verbally rather than visually. Even

---

with content that is rich in visual symbolism, such as mathematics, the verbal explanation usually provides vital information that explains the symbolic processes.

Students with who are Deaf or are hard of hearing may miss the environmental and metadata that is carried within a communication. For example a science experiment may demonstrate a cause and effect relationship between events, yet the effect may be the sound of say a small explosive “pop”. If that “pop was not heard by a student the whole point of the experiment may be lost. Likewise the meaning of a dramatic dialogue may rely on vocal stress and intonation to modify meaning. It is common for the literal meaning of a sentence to be reversed or shifted through selective vocal stress in dialogue’s delivery. A verbatim transcript of a conversation may miss this subtlety.

Subtitles or captions may be used to provide a visual support for student who may not get full access to the auditory components of an information source. Note the terms *captions* and *subtitles* are used interchangeably in this guide.

Some forms of captions:

#### **Real-Time Captions**

Real-time captions are often used in real-time situations such as classrooms or televised live events such as news or sports. They are created by a real-time captioner who translates conversations to text on a computer at the time the words are spoken. Typically these are created using a stenotype system similar to those used to record court proceedings.

Real-time captions require a highly skilled operator with expensive equipment who is usually physically present during the lecture; the captions are delivered as the conversation occurs so are ideal for “live” events; real-time captions typically have a relatively higher error rate as the words are recorded phonetically rather than spelt out in full. Errors occur particularly around jargon, technical terms, and names of people and places which may not pre-recorded in the captioning system’s dictionary or where two words sound the same but are spelt differently.

Finally real-time captions are usually delayed so they may not be synchronised with lip patterns and other visual and auditory information. For example a presenter may point to an object and make a comment, but the caption could appear on screen 3-6 seconds later when the presenter is pointing to a different object.

---

## Open Captions

Open Captions are added to a video after a program has been recorded. The captions are permanently “burned” into the video for all viewers to see – there is no option to turn the captions off. Such captions are often seen on foreign language films, for example those shown on SBS or world movies. See Figure 1. Below. The advantage of open captions is that they do not require a decoder or other special technology to display them. Unfortunately open captions may obscure parts of the visual image; may distract some viewers who don’t need them; and are not able to be varied to suit the needs of different viewers.

On the positive side it is easier to control the on-screen presentation of closed captions as their presentation is not subject to variations in the decoding between display devices. For example the font, size and backgrounds of televised teletext captions vary between television sets.



*Figure 1. Program with open captions*

---

## Closed Captions

Closed Captions are not permanently displayed as part of a video program, they are encoded within the video signal and can be selectively displayed or withheld from the screen using the displaying device. On a television there may be a subtitle button on the remote control handset. Smartphones have video apps in which there may be an on-screen subtitle option.

One additional benefit of closed captions is that can be coded with multi-language options. Some video media may include English captions, but also English SDH caption tracks (subtitles for Deaf and hard of hearing). The later includes sound effects, speaker identification, and other non-verbal audio information within the caption. For example the sound track accompanying a movie could add dramatic effect so may be captioned as “Haunting music”. These subtitles may identify speakers off screen, may add in sound effects such as “baby crying” or “rapid gun shots” that assist in the understanding of the on-screen visual action. See Figure 2. for an example of closed captions for hearing impaired viewers, note the speaker is identified, as opposed to Figure 1.



*Figure 2 Program with closed captions*

In addition to closed captions in broadcast television, Blu-Ray and DVD disks, an increasing volume of video-on-demand content is now captioned. For example the ABC iView video playback service now includes closed captions,

---

TED talks on the Internet are all captioned, as are all the video content available in Australia on the Netflix video streaming service. Some YouTube and Vimeo video content also includes closed captions, although this will depend on the content author.

Digital media is increasingly becoming distributed as video files in formats such as MPEG or DIVX, and again these all have the capability to carry closed captions. Such content can often be displayed on wide range of devices, from desktop computers, data projectors, laptop computers, tablet devices and smartphones, to name a few.

The Internet can often carry video content that is encoded in a variety of ways including the new HTML5 standards and viewed using a web browser (or similar). These have been updated to include very versatile and creative closed caption capabilities using WebVTT standards. Unfortunately the ability to display all the features of HTML5 are not fully supported by all browser types or devices and may cause compatibility issues displaying closed captions in some web browsers. As HTML5 develops the future of captioning options and presentation is extremely exciting.

#### **Language modified captions**

Language modified captions may be used to assist the reader. For example some caption writers will modify the linguistic level and reading speed required for captions aimed at young children.

#### **Live Speech Recognition Systems**

Some lecturers have 'discovered' speech recognition software that will translate spoken words to text in real-time. The technologies involve computers and smart devices and are improving in their accuracy. To be used effectively the recognition software should either be trained to understand a particular speaker (for example Dragon dictate) or the speaker modifies their delivery for ease of recognition (for example Siri). As the presenter speaks their verbal delivery is 'recognised', converted to text and displayed on a screen. Of course any comments, questions or audio information from videos will not be translated by the speech recognition system, and so is not accessible to the Deaf or hard of hearing student.

Perhaps the greatest downfall is that this is a one-way information channel – from presenter to student. If an interpreter is present they can relay a Deaf student's question or comment to the presenter or the class, but this channel is not available using speech recognition.

---

Live speech recognition technology does not deliver an equitable and accessible communication alternative and is not recommended at this time

### Auslan in Videos

Many Deaf students use Auslan (Australian Sign Language) as their first language and would prefer to have an interpreter translate verbal content into Auslan. Auslan is a nationally recognised Australian language with its own vocabulary, structure, grammar and culture and value system linkages. It is a rich and efficient language capable of conveying complex information.

Auslan interpreters may be provided “live” at a lecture or tutorial. A recent development in disaster management reporting on television has been to have a sign language interpreter within the video frame of the person providing emergency management information. This is often supplemented by real-time closed captions, offering the Deaf and hard of hearing viewers a choice of communication channels.

Video material may also have Auslan added post-production by a superimposed overlay on top of the image. There are a number of methods for displaying Auslan on a video clip. Figure 3. shows the traditional method of superimposing a cameo image of the interpreter over the image. Note this example also includes the closed caption option.



*Figure 3. Auslan interpretation added over the video.*

---

Because the cameo image is quite small this can be difficult to read the signs, and important visual information can sometimes be obscured behind the overlaid layer.

An alternative that allows both the interpreter and the video to be displayed in full is to reduce the video image to make way for the interpreter and the captions. This makes the original image a little smaller, as shown in Figure 4.



*Figure 4. Another method of showing Auslan*

There is no current technology that allows for an interpreter box to be turned on and off in the same manner as closed captions. It is of interest that the interpreter can be either Deaf or hearing as the Auslan interpretation is added post-production.

### **Access for Students who are Blind or have a Visual Impairment**

Students with a visual impairment may have difficulty accessing visual components of a presentation, both live and in videos. Graphs and charts, formulae, models, demonstrations, photographs, PowerPoint presentations are all rich in visual information that is usually not replicated fully in the accompanying spoken words.

Perhaps the most powerful and effective ‘technology’ to make this information accessible is to improve the skills of the presenter. A presenter who can clearly describe what is being demonstrated, displayed or written as part of their commentary will be a far better communicator for all students, especially those with a visual impairment.

---

Video material can be enhanced post-production by the addition of a descriptive audio track. Digital television, Blu-Ray disks and DVDs allow multiple audio soundtracks to be added to any video program. Most digital media viewing programs also support alternative audio tracks.

To create audio description tracks a trained commentator records a sound track to describe important visual information on the video source that supplements the standard audio channel. The descriptive information is added in the pauses and silence in the standard sound track. For example:

Prof Brown *“This shows that Australians females tend to have a longer life expectancy to males”*

Description track *“Prof Brown points to life expectancy bar chart on overhead...”*

Or

Video voice over *“Life in Vietnam offers little hope for the young.”*

Description track *“Video shows dirty brown river with floating detritus and children swimming on the banks”*



Figure 5. Descriptive audio track

---

Australia's national broadcaster, the ABC, has undertaken trials delivering some programming with descriptive audio tracks on broadcast television and their iView streaming service. Netflix is also delivering some content with audio descriptions.

### **Collaborative On-line Education Technologies**

One rapidly emerging 'disruptive' educational technology involves the use of collaborative software or systems for remote learning. With the increasing availability of affordable broadband Internet services it is now viable for many students to learn from homes or other convenient learning sites. This has spawned an education revolution, particularly for part-time students. Universities are employing collaborative systems, such as Blackboard Collaborate, to connect students and presenters that are physically dispersed.

Students log-on via their computers and are typically presented with a complex screen that presents PowerPoint slides, an audio track from the presenter, list of online participants, video screens and a chat window for questions, answers and comments. Delivering equitable access in these complex environments pose significant challenges for students with disabilities.

Course convenors should look for captioned content, offer real-time captions for capturing the content and possibly engage Auslan interpreters to help moderate and facilitate interactions. The accessibility solutions need to be well thought through before hand – ideally involving educational designers, disability support staff, the student themselves and, of course, the content deliverers.

At the extreme end an international trend to deliver MOOCs (Massive Open On-Line Courses) to tens of thousands of students across the globe make these access challenges significant, yet important as the way education is delivered is rapidly changing.

---

## Guide for Captions Styles and Presentation

### Choosing Accessibility Design Tools

There is an inherent dilemma in preparing any guide that describes computers and software used to create any media for distribution. Any attempt to make the information 'technology neutral' will result in a guide that provides no guidance. It risks becoming so generalised that it does not offer directions that will genuinely assist the reader.

Conversely references to specific computing platforms, operating system versions and particular software applications will certainly mean that the guide will be out of date by the time it is published.

In preparing this guide we hope to provide the broader principles for creating accessible media while referencing current computing platforms and software applications that may assist those wishing to engage in media adaptation or creation.

Regardless of the platform used to create the media it is essential that the material created must be able to be viewed by students using any of the commonly available computer systems without incurring additional cost or inconvenience.

Over recent years many new tools have been developed for the creation of accessibility features, captions in particular. There are many choices of freeware programs that are fully featured and relatively simple to use. Commercial versions are also available for most computer operating systems that have dropped in price significantly. Both will allow the creation of professional standard captions with a minimum investment in money and training.

### How We Read Captions

Most audio-video content, for example television and movies, do not simply deliver a message through the dialogue with supporting visual images. Both the visual and audio component provide separate meaning components that add up to provide the full story. The sound channel without the video, or the video without the sound do not convey all the content information. Captions can help deliver audio information in a visual manner for viewers who are Deaf or hard of hearing. This means there is a lot more visual information to process as viewers are required to read the text while watching the video action. Some caption users report the increased fatigue as a result.

Spoken passages can be delivered at very high speeds. A news reader often delivers text at speeds exceeding 200 words per minute, while excited storyline characters in a movie could exceed 250 wpm. These speeds are easy enough to listen to and

---

comprehend, but when converted to text it is more challenging to read at those speeds, especially if there is also live action on the screen at the same time.

A caption viewer must process the audio story as written text and still have time to follow the visual storyline at the same time. As the caption text obscures some of the visual story this can be even more challenging.

In creating captions we need to be aware that we are placing an extra visual load on our viewers. By 'crafting' captions so they present information in logical blocks of meaning and keeping information together on the screen can ease the reading burden and facilitate better comprehension.

## **Presentation of Captions**

The conventions used to display spoken words and sounds as text are not universally agreed. A seasoned caption viewer can often make an educated guess as to the origin of captioned sequences by observing the way captions are created. For example many captions created for television in the USA are all white text on a black background - in all capital letters. The BBC in Scotland specifically this country and not the UK? often uses three line captions, while in Australia, Red Bee Media or Ai-Live use colours and text justification to position speakers. Sounds effects, music, poetry and emphasis are all captioned in a variety of ways. None are 'wrong' - just different.

Given that Australian viewers will be most familiar with television captions delivered on local television, and that these largely use the standards and conventions pioneered by the Australian Caption Centre in the 1980s, this guide draws heavily on those conventions. In this way the captions created for educational media will follow familiar and effective standards that are common on Australian television.

### **Number of Lines on the Screen**

When creating captions a balance must be struck between the information that is displayed as text and the amount of video information that is obscured by the captions. Captions that have three, four or more lines of text cover a significant part of the screen and impede the video information. Figure 6. shows how too many lines of captions can obscure the visual information and lose important meaning.



*Figure 6. Too many caption lines*

If we use a series of single line captions to present the spoken text it would reduce the amount of screen area covered up, however there are some unwanted implications:

- Rapid single line captions tend to ‘flash’ on the screen making them distracting;
- Using single lines means that the captions must change more rapidly, this reduces the available time to scan the visual storyline;
- The use of single line captions provides less space to construct meaningful ‘chunks’ of information. It is easier to construct meaning if the idea is conveyed using text in a single visual block of information shown on the screen all at once, rather spread over sequential screens.

In most circumstances an effective compromise is to use captions displayed on two lines on the screen. This should provide enough space for meaningful ‘chunks’ of text without obscuring too much of the screen. There will be times in a program when the amount and timing of the sound track will make a single line caption appropriate.

**Guideline 1:** *For most content use two lines of captions. Single line captions can be used when there is only a small amount of text or slower timing of the spoken content.*

### **Line Breaks in Captions**

Most readers don’t read word for word but use short phrases to grasp meaning. By phrases we mean a string of words that collectively convey a unit of meaning. It is easier to understand captions if those phrases are on a single line, or at least

---

on the same caption screen. Thus line breaks can provide grammatical clues and aid reading.

Here are a few simple rules to assist keeping meaningful blocks of text together:

- a. If the caption is short, use the lower line.

John was feeling sick

***Difficult to read***

John was feeling sick

***Easier to Read***

- b. Shorter lines are easier to read than long captions. Split captions into two lines or use more caption screens

John and Mary were always busy

***Difficult to read***

John and Mary  
were always busy

***Easier to Read***

- c. Do not split names or people's titles across captions.

She was working with Mr John  
Smith during the day

***Difficult to read***

She was working with  
Mr John Smith during the day

***Easier to Read***

- d. Keep descriptive words (adjectives –for example 'blue', 'cooked', 'aboriginal', 'young') on the same line as the word they help describe.

Every day she drove her pink  
car to work

***Difficult to read***

Every day she drove  
her pink car to work

***Easier to Read***

- e. Keep position words (prepositions – for example 'at', 'in', 'on', 'under', 'beside') on the same line the word they are describing.

She found her sunglasses in  
the car glovebox

***Difficult to read***

She found her sunglasses  
in the car glovebox

***Easier to Read***

- f. Do not break a line after a joining word (conjunctions – for example ‘and’, ‘or’, ‘nor’, ‘either’). Breaks before conjunctions are better.

She was working with  
Mr John Smith during the day

***Difficult to read***

She was working  
with Mr John Smith during the day

***Easier to Read***

- g. Keep helping verbs on the same line as the verb they help (helping verbs – for example ‘was’, ‘has’, ‘could’, ‘not’, ‘should’).

His eldest daughter was  
leaving on the morning train

***Difficult to read***

His eldest daughter was leaving  
on the morning train

***Easier to Read***

Guideline 2: *line captions use the lower line.*

Guideline 3: *Split long captions over two lines where possible.*

Guideline 4: *Do not split names or people’s titles across captions.*

Guideline 5: *Keep adjectives on the same line as the word they help describe.*

Guideline 6: *Keep prepositions on the same line the word they are describing.*

Guideline 7: *Where practical break a line before a conjunction.*

Guideline 8: *Keep helping verbs on the same line as the verb they help*

### **The use of Punctuation in Captions**

Punctuation provides timing and meaning to our written words. People don’t speak punctuation. We don’t say “I was interested in what he had to say comma but didn’t have time to stop and chat full stop” or “So that was the girl you met question mark”. This isn’t necessary in conversation as we use the intonation in our voice, the sentence structure and timing to imply punctuation. Like the spoken word, captions are also time based and the timing and line breaks of captions can help provide punctuation information.

---

New sentences always start on a new line and a new caption. If there are very short sentences it may be permissible to have two short single line sentences in a single caption. A new sentence starts with a capital letter.

He was shocked. His eldest daughter was leaving

***Difficult to read***

He was shocked

***Easier to read – 1st caption***

His eldest daughter was leaving on the morning train

***Easier to read – 2nd caption***

It is recommended that full stops are not used in captions, rather that full stops are shown by a line break and are usually followed by a new caption. (see sample above).

Question marks are commonly used in captions but should always be followed by a line break or a new caption. Avoid exclamation marks where possible and never use multiple exclamation marks.

Leaving Wednesday? You have plenty of time!!!

***Difficult to read***

Leaving Wednesday?  
You have plenty of time!

***Easier to Read***

Avoid the use of all caps (all capital letters) and italics for conversational text. All caps should be reserved for identifying sound effects and identifying speakers. Italics should be reserved for songs, poems, dreams and reflections, or special effects voices such as radio conversations or a voice over a public address system.

JACK: Are you planning to use THAT thing?

***Difficult to read***

JACK: Are you planning to use that thing?

***Easier to Read***

---

## Choice of Font

The font or typeface will impact on the clarity and readability of captions. For broadcast closed captions the display font will be set by the decoder used to display them. For other captions such as open captions or closed captions used in web formats such as QuickTime, Windows Media files and Flash captions, the caption writer can select the font to use.

The general rule is to keep it simple. Fancy fonts make terrible captions.

It is recommended that a simple sans serif font be used for captions. Serifs are the little embellishments added to the characters, such as the 'feet' at the top and bottom of capital 'I' for times text. Sans Serif simply means no serifs and includes common typefaces such as Arial and Helvetica.

**Guideline 9:** *Use a sans serif font such as Arial or Helvetica for captions.*

## Length of Caption Lines

There is no hard-and-fast rule about the number of characters in a line for captions. Teletext captions used in Australia's superseded analogue television system allowed for up to 40 characters in a single line, although this is a function of the videotex technology used to create the captions.

Today's widescreen formats, high definition screens and larger televisions allow for longer lines of captions to be displayed clearly.

Today's subtitled material may also be delivered by the Internet or on mobile devices. The images may be presented in much smaller screen sizes. This will impact on the readability of the captions. Open captions that permanently displayed as a part of the video image are particularly problematic as they may not resize clearly.

Closed captions are generated and displayed at the time of viewing and they are generally able to be rescaled without loss of clarity or being cropped on narrow screens. As closed captions are scaled, if the video is scaled down to a small screen, the captions will also be scaled proportionally. While the text will be displayed as sharp images they may be too small to read comfortably. It is always advisable to test your captions at all sizes they may be displayed to check for readability.

---

The safest way to ensure the size and placement of captions are appropriate is to experiment with different versions. Test the readability of captions after compression and on the smallest screen that might be used to view the captions and adjust the size to suit.

Guideline 10: *Captions must be of a size that are readable yet do not cover too much of the screen.*

Guideline 11: *Prepare captions for the smallest screen size that may be used to display the video.*

### What to Caption

Where possible all captions should use the exact wording and phrasing of the original spoken words. For example if a speaker stumbles or stutters this should appear in the captions as it may provide information about the confidence or emotional state of the speaker. Captions for adults should never be “dumbed down” by simplifying the language or changing vocabulary. This impacts on the meaning.

When deciding what sound information needs to be captioned, remember to include:

- All words spoken by characters
- Words spoken by a narrator
- The words to any song
- Identification for off screen speakers
- Descriptions of sound events that impact on the story or meaning.
- Captions of sound events should describe any sounds that convey information that assists understand the story. This may include music that ‘sets the scene’, sound effects, sounds such as public address systems or radios, etcetera. For example use “GUNSHOT” rather than “BANG” as bang could be a balloon bursting, a car exhaust

---

Guideline 12: *Guideline 1. Captions for viewers with a hearing impairment should be verbatim and include:*

- *All words spoken by characters (including stuttering, etcetera)*
- *Words spoken by a narrator*
- *The words to any song*
- *Identification for off screen speakers*
- *Descriptions of sound events that impact on the story or meaning*

Guideline 13: *Prepare captions for the smallest screen size that may be used to display the video.*

Guideline 14. *Captions must make it clear that they are showing spoken words, narrator, songs, speaker identification or descriptions of sounds. The presentation convention for each caption should be consistent and clearly identifiable.*

Guideline 15. *Spoken words are presented in mixed case (upper and lower case text). If coloured they should be a light colour on a black or dark background.*

Guideline 16. *Words spoken by a narrator should be displayed in italics.*

Guideline 17. *Words to a song should be preceded by "SINGS:" followed by the words in italics.*

Guideline 18. *Off screen speakers should be identified by preceding their spoken words with their name in all-caps (all capital letters) followed by a colon.*

Guideline 19. *Captions of sound events should describe any sounds that convey information that assists understand the story.*

Guideline 20. *If colour is available present sound descriptions as white all-caps text on a cyan background.*

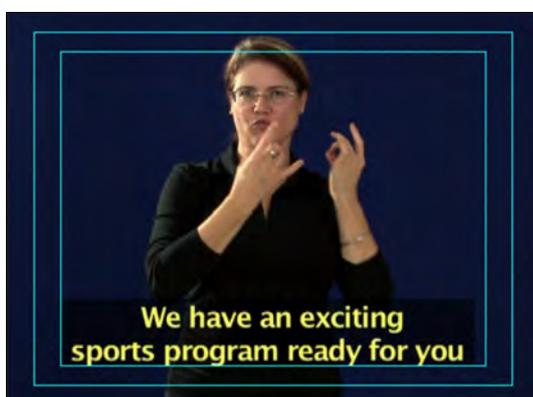
*If no colour is available show sound descriptions as all-caps enclosed by brackets [ ].*

---

## Positioning Captions on the Screen

The technology used to create captions may limit the ability to accurately position and align captions on the screen. Some captioning systems create captions at the lower third of the screen that are either left aligned or centred. If the only option available is caption alignment text that is aligned left or left justified is easier to read than centred text.

Given the ability to choose caption placement it is a common convention to present captions in the lower third of the screen. If important information is obscured captions can be related temporarily in the upper third, but this can be distracting and should be used in moderation. Professional video editing programs, such as Final Cut Pro (Macintosh) and Adobe Premier (Mac and Windows), offer guides that show you the ‘title safe’ areas that should be used to guide where the captions can be placed. In figure 7. the outer blue guide shows the image safe area, while the inner blue guidelines delineate the area where it is safe to place text.



*Figure 7. Title Safe Guides (as displayed in Final Cut Pro)*

Some caption writers may position sound description captions on the screen in a way that indicates the position of the source of the sound. For example “Gunshot” may be placed in the upper right corner to indicate the gun was fired from that direction. This is a matter of choice and may be distracting.

Caption alignment may be used to help identify the speaker. For example if there are three characters in a dialogue the speaker to the left of the shot should be presented with left justified captions, the centre speaker with centred captions and the right side speaker with right justified text. This also helps identify who is speaking when the speaker is out of view of the camera.

Changing the justification of captions can be difficult with some captioning software and may increase to time to create captions.

---

Guideline 21: *It is preferable to place captions in the lower third of the screen.*

Guideline 22: *The upper third of the screen may be used if captions at the bottom cover important information, however this should be avoided where possible.*

Guideline 23. *Where possible, multiple line captions should be aligned at the left (left justified) unless alignment is used to indicate speaker positions.*

### **Caption Colours and Background**

For captions to be readable they must stand apart from the screen image. At the same time they must not obscure or distract from the image too much. The choice of caption colours and backgrounds becomes important.

The ability to colour captions and their background varies depending on the captioning software or technique used. Some applications allow a great deal of flexibility, while others are limited to a single coloured text.

Colours can also be used to identify speakers, this is particularly useful if the person speaking is not always visible on the screen. By assigning a colour to a speaker for the duration of a scene it makes it easier for a caption viewer to know who is talking at throughout the sequence by the colour of their captions.

### **Colour Choice**

Colours should be chosen to allow for maximum contrast against the background or screen image. It is impossible to read white captions against a white background, or black against most dark colours. Unless the screen imaged is fixed so that the bottom third of the screen is always a set colour it is difficult to pick colours for captions that will always be readable.



*Figure 8. White captions may be difficult to read*



*Figure 9. Black captions can also be difficult to read*

Some captioning software will allow for fonts to be outlined, for example white text with a fine black outline for each character. An outline will mean that the caption can be read against any background colour. Outlines tend to thicken the fonts and making them a little crowded and sometimes harder to read.



*Figure 10. Adding a text outline improves readability*

---

Another option that is often used for closed television captions in Australia is to surround the text with a solid contrasting colour. This technique provides captions that are much easier to read. There is a greater probability that boxed captions could obscure visual information in the original image.



*Figure 11. Captions on a solid contrasting background are clear*

An increasingly popular technique that is becoming more common with digital television closed captions is to place a semi-transparent box behind the captions. This improves the contrast but allows the image behind the captions to show through. Semi-transparent boxes provide a softer, less intrusive appearance to the captions.



*Figure 12. Adding a transparent background adds image detail*

---

## Creating Captions and adding them to Videos

### Sources of 'ready-made' captions

Creating captions from scratch may require a significant investment of time and effort, and this may not be practical in some educational settings. Too frequently a last minute decision to use video material in a lecture or tutorial leaves little time for generating captions. With a little research you may be able to source ready made captions to save time and meet that deadline.

Some sources of ready-made captions include:

#### Material already captioned on the Internet.

Many on-line video services such as YouTube and Vimeo may already have closed captions. In fact there are many similar services on the Internet, a recent search found around 40 on-line video services. All Ted Talk videos are captioned in multiple languages and offer a wide range of educational content.

*A word of caution:* at the time of writing YouTubes' auto-generating captions are not accurate and should never be used to provide access for Deaf and hard of hearing viewers. Check the settings icon (a small cog) to make sure you are not relying on auto-generated captions. Often in the left hand upper corner of the video, "English auto-generated" may appear, and this is a good indication that the captions are not of good quality. It is good practice to take a minute and preview the video with captions enabled before you show it.

#### Downloadable caption files.

There are many sources of caption files that can be downloaded and added to your video to make them accessible. These vary in quality as they are created and distributed freely through a community-sharing model. There are software search programs that will assist you locate these files, but the easiest way is to simply put into your search engine (for example Google) "video title + .srt + download". Take care as some 'free' sites will try to encourage you to download software or subscribe, which should be avoided.

*A word of caution:* Caption files, typically in Subtitle Resource Track or .srt format, are timed to synchronise with the video. Videos can be presented at different frame rates, that is the number of frames per second (fps). In Australia most movies and television programs are distributed with 25 fps, video sourced from USA television may be at 29.97 fps, other frame rates are also available. You may find that your captions go out of synch and that could be as a result of a

---

frame rate mismatch. You can get lost in the technicalities, but sometimes trying a few different frame rates will resolve your problem.

### Extracting captions from a DVD.

You may have video material on a DVD that you now wish to convert to a digital file that frees you up from disk players, and that may contain closed captions. Unfortunately captions are usually stored as bitmap images rather than text files on DVDs, so extracting them is not straightforward. Recently a number of programs have emerged that let you extract these captions as .srt files. They 'find' the caption bitmap files and use optical character reader algorithms (OCR) to convert these images to text. Such extraction applications include SubRIP, but your search engine will identify a number of options, try searching "Open Source DVD subtitle extractor", or you could add in you computer operating system by adding "OS X" or similar.

*A word of caution:* Commercial videos may be subject to copyright restrictions and may have copy protection systems that prevent copying the video and captions, also OCR is not perfect and some text may be misread, so captions should be checked for errors be using them.

### Using .srt files with your video.

In order to show your closed captions you will need to prepare your video and caption files. Most video presentation applications provide the ability to display closed captions. In some cases simply placing the movie file and the .srt file in the same folder, making sure they have identical files (eg *movie\_name*.mov and *movie\_name*.srt) will be sufficient, but not always. The safest way to ensure your video's closed captions can be displayed is to encode the captions into the movie file. This also allows you to check that the captions display properly and are synchronised.

Following these steps will help add the caption files to the video file:

Download a software to add your closed captions, there are many choices, try searching "add closed captions open source software" or search on-line software stores to purchase commercial products.

- Open the software program and load the source video file.
- Add the caption file (.srt) to the video
- Check that the timing is correct and the captions synchronise with the audio track. You may need to offset the captions, that is move their start time back or forward until they line up.

- 
- If the audio becomes increasingly out of synchronization then you probably have a caption file the wrong frame rate.
  - Once it all works and you are happy with it, you can export the combined caption and video file. Some programs allow you to export to a file format and/or size that is optimized for different playback systems or devices such as smartphones, tables or games consoles.
  - All done!

### **Creating your Own Captions.**

Years ago creating captions was a 'closed shop' that required specialist hardware and software, skilled operators and a lot of time. Thankfully captions can now be created with relative ease.

The first task is to get a transcript of the audio track of the video you wish to caption. If you have access to a 'gun typist' they can prepare a transcript for you. Another option may be to get students to transcribe the audio for you. Lastly there are quite a number of commercial services that will transcribe your audio into a text file from a video or audio file. These can cost around \$1 to \$2 a minute, have a high degree of accuracy, and will return your file in one or two days.

There are a number of options you may like to consider:

#### **On-line captioning services.**

There are a number of on-line captioning services that allow you to use web based tools to create captions. These include Amara, DotSub and CaptionTube (which is the YouTube system). The YouTube captioning system is examined in more detail below.

#### **Use YouTube to create or edit your captions.**

After setting up your free YouTube account you can upload your video clip and use YouTube's caption editor to generate captions. The YouTube auto-generated captions are generally not accurate enough to be of use. You can either type in your captions in the editor, which is tedious, or you can simply upload a transcript of the video. YouTube will be able to automatically synchronise your transcript with the audio track of your video. As you might expect there are a number of YouTube instructional videos on the use of the YouTube captioning editor.

Figure 13. below shows the basic workflow used to create captioned videos using YouTube's build in caption editor.

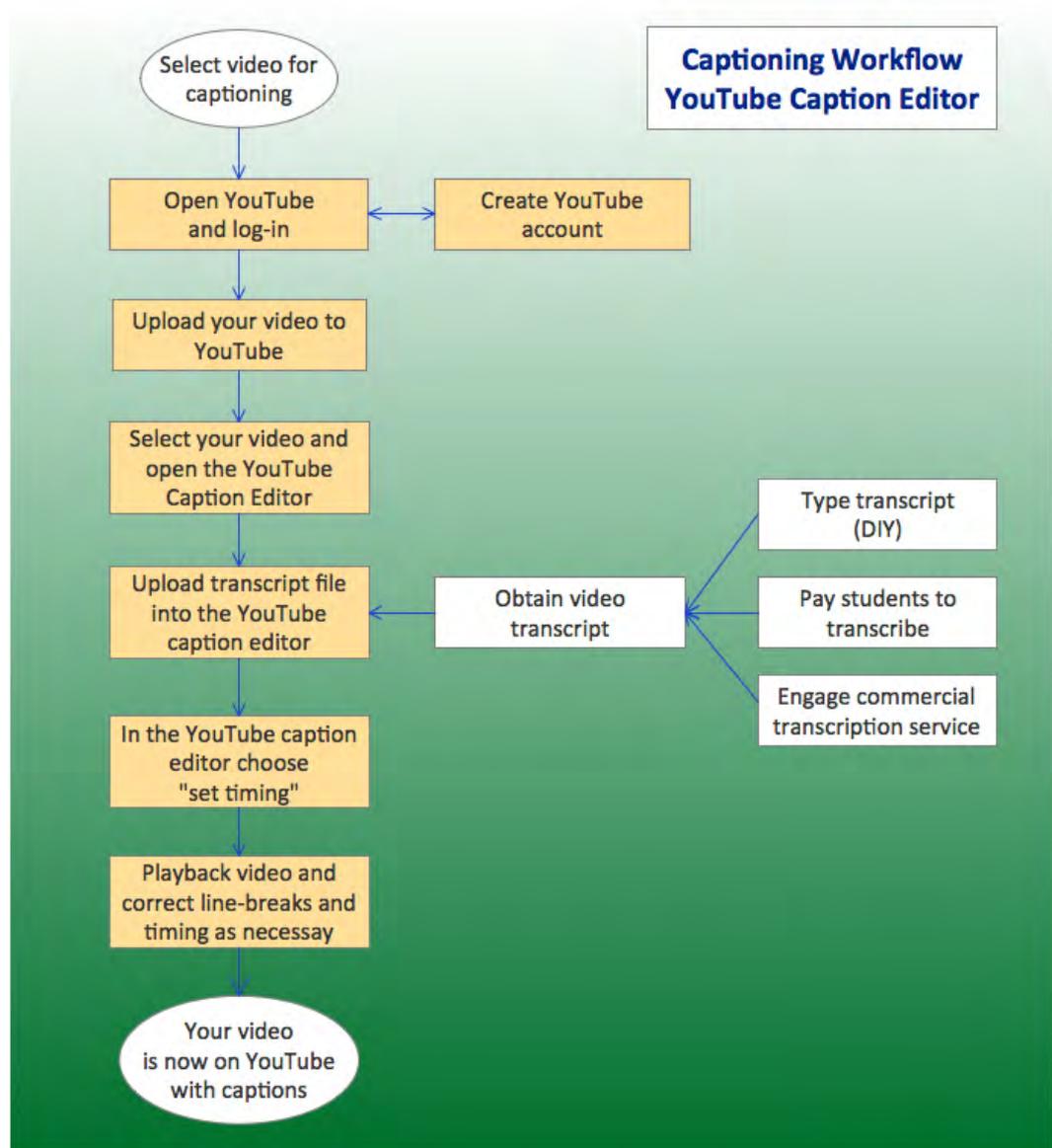


Figure 13. YouTube caption editor workflow

It is worth noting that this system assumes that you are showing your video using YouTube. If you want to display the video where there the Internet is not available, is too slow or has been blocked in a corporate network (some education networks block access to streaming video services), then these captions are not accessible. It is possible to export the caption file from YouTube but they are saved in SubViewer format (.sbv) that cannot be opened by most caption encoding programs. There are converters available.

A word of caution: If you simply upload a transcript and ask YouTube to create and time captions you will get a barely useable captioned outcome. Unfortunately it will not make intelligent decisions about where to break the lines and you will

---

almost certainly get a captioning product that is difficult to read. To improve these captions you should go back and re-edit the captions and timing using the guidelines contained in this manual. Unfortunately many educational designers believe captions created and timed by YouTube creates “accessible educational media”, but they fall well short of a professionally acceptable and equitable captioning product.

**Use commercial or open source captioning software editors and encoding software.**

There are a great number of readily available free or inexpensive caption creation tools that will work on contemporary computer systems. Try searching the Internet for “create closed captions open source software” or search on-line software stores to purchase commercial products. As you might expect these vary in their ease of use and the level of utility they provide. Generally there are three steps in the captioning process:

- Prepare the audio transcript, this can be done in a text editor such as Word. It is best to create the line-breaks to ‘craft’ the captions at this stage. This creates a text .txt file.
- Time the captions in a caption editor. Export the time-coded caption file as a Subtitle Resource Track .srt file
- Encode the source video and embed the captions, then export as a captioned movie in one of many available movie file formats (For example .mp4 or .mov)

Some captioning creation programs can be used to achieve one, two or three of these functions. Most of the free Open Source software is usually only capable of a single step.

The caption creation workflow is summarized in Figure 14 below.

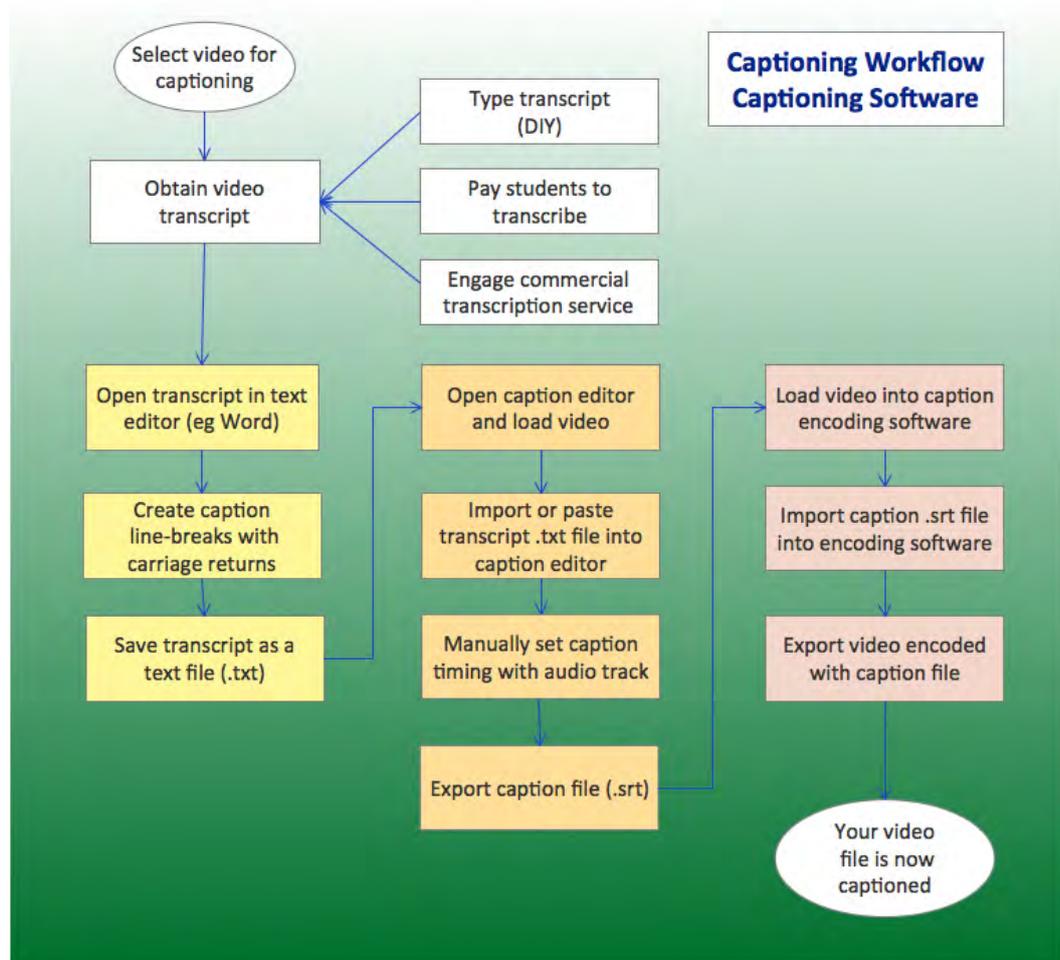


Figure 14. Caption creation workflow

---

## Conclusion

The pace of change in tertiary education is increasing exponentially – fuelled by communications technology innovation. With each step forward we are offering new opportunities for all students, including those with disabilities, to participate and learn. As instruction becomes more sophisticated, leaving chalk-and-talk behind, we must step up to our responsibilities to ensure our digital teaching resources are equally accessible to all. Luckily the tools that are used to create these accessibility features are also developing at a rapid rate. More than ever software is freely available to create professional quality captions and subtitles.

Instructional designers and teachers must take care not to fall for the traps of quick-and-easy accessibility ‘solutions’ for the sake of expediency and apparent ‘compliance’ to disability discrimination requirements. Accessibility provisions such as captioning must maintain the quality and craftsmanship of design that creates true accessibility.

By applying the guidelines offered in this guide the accessibility of digital teaching resources will substantially enhance the educational opportunities, and the life benefits that flow from them, to all students – including those with disabilities.

---

## APPENDIX A – Summary of Captioning Style Guidelines

- Guideline 1: *For most content use two lines of captions. Single line captions can be used when there is only a small amount of text or slower timing of the spoken content*
- Guideline 2: *line captions use the lower line.*
- Guideline 3: *Split long captions over two lines where possible.*
- Guideline 4: *Do not split names or people's titles across captions.*
- Guideline 5: *Keep adjectives on the same line as the word they help describe.*
- Guideline 6: *Keep prepositions on the same line the word they are describing.*
- Guideline 7: *Where practical break a line before a conjunction.*
- Guideline 8: *Keep helping verbs on the same line as the verb they help*
- Guideline 9: *Use a sans serif font such as Arial or Helvetica for captions.*
- Guideline 10: *Captions must be of a size that are readable yet do not cover too much of the screen*
- Guideline 11: *Prepare captions for the smallest screen size that may be used to display the video*
- Guideline 12: *Captions for viewers with a hearing impairment should be verbatim and include:*
- *All words spoken by characters (including stuttering, etcetera)*
  - *Words spoken by a narrator*
  - *The words to any song*
  - *Identification for off screen speakers*
  - *Descriptions of sound events that impact on the story or meaning*
- Guideline 13: *Prepare captions for the smallest screen size that may be used to display the video*
- Guideline 14. *Captions must make it clear that they are showing spoken words, narrator, songs, speaker identification or descriptions of sounds. The presentation convention for each caption should be consistent*

---

*and clearly identifiable.*

Guideline 15. *Spoken words are presented in mixed case (upper and lower case text). If coloured they should be a light colour on a black or dark background.*

Guideline 16. *Words spoken by a narrator should be displayed in italics*

Guideline 17. *Words to a song should be preceded by "SINGS:" followed by the words in italics*

Guideline 18. *Off screen speakers should be identified by preceding their spoken words with their name in all-caps (all capital letters) followed by a colon*

Guideline 19. *Captions of sound events should describe any sounds that convey information that assists understand the story.*

Guideline 20. *If colour is available present sound descriptions as white all-caps text on a cyan background.*

*If no colour is available show sound descriptions as all-caps enclosed by brackets [ ]*

Guideline 21: *It is preferable to place captions in the lower third of the screen.*

Guideline 22: *The upper third of the screen may be used if captions at the bottom cover important information, however this should be avoided where possible.*

Guideline 23. *Where possible, multiple line captions should be aligned at the left (left justified) unless alignment is used to indicate speaker positions.*