

JOE CLARK MEDIA ACCESS

Comments on CAB captioning manual

This document provides Joe Clark's comments on "Canadian English-Language Broadcasters' Closed Captioning Standards and Protocol" ("the Manual"). The version I used carries the document ID CABDR5_MS04.doc and is dated September 09, 2002. Page numbers refer only to that version.

Table of contents

All on this page:

1. [Severe](#) errors
2. [Moderate](#) errors
3. [Minor](#) errors
4. [Observations](#)

Every heading has an anchor and can be individually linked. A full table of contents that will make such links possible may be added this week after I alphabetize the contents.

Change history

2002.09.22 19:00
Posted.

I will continue to debug small copy errors over the week, and may reorder elements so they are alphabetized. All notable changes will be documented in this section. (An example of a change that would not be notable is correcting *teh* to *the*, or an HTML fix.)

Where to find this document

These comments can be found, in various file formats, at joeclark.org/access/captioning/cab/.

Contact information

Joe Clark • joeclark@joeclark.org • 416 461-6788

Notes

I was invited to provide these comments by Liz Chartrand, Sarah Crawford, Harvey Rogers, and Susan Wheeler, to whom I extend heartfelt thanks. By agreement with Susan and Harvey, my function is to identify mistakes. I have tried to emphasize research citations and rational argumentation rather than advancing unsubstantiated opinion.

But please understand the ramifications of my brief. I was charged with finding as many errors as possible in the Manual. Accordingly, these comments are all critical by definition. I am not, in fact, being “too negative”; “being negative” was the objective.

Note that, in technical documentation, critical readings like this are par for the course. In engineering documentation, for example, it’s possible for a single round of consultations to produce hundreds of action items for updating and correction. This document discusses some 70 topics – not atypical for a 50-page technical document.

As in technical documentation, the whole point is to flag every error so a decision can be made to either correct it or live with it. Readers are requested not to hold it against me that I was this thorough in doing what I agreed to do.

Improving future practice

The exact goals of the Manual need to be more clearly stated. Don’t be afraid to take a stand!

If the Manual is to improve future captioning practice, contributors must accept that *we have to outlaw what captioners are doing wrong now*. We should not be a little bit pregnant. It’s nice to “strongly discourage” this habit or that, but certain practices must be banned outright. In other words, captioners could not claim adherence to the Manual if they engaged in such practices.

It should also be pointed out that, save for cases where the content itself allows a certain discretion, any style guide should not be treated as a shopping list. It’s all or nothing. We have as many varieties of captioning as there are captioners because captioners have picked and chosen from predecessor styles and also made things up as they went along.

In the future, Canadian offline programming should not, for example, carry captions from house X or shop Y. Programs should not carry certain kinds of captions; they should simply carry *captions*, all of which look and act the same. If we really believe that captioning is an integral part of a television program, we have no choice.

Monoculture and recruitment

A special note on my frequent reference to the Canadian captioning monoculture: Having followed captioning for 20 years, I am not familiar with any country whose captioning industry is so dominated by women in their 20s with liberal-arts degrees. The “women” part is observably true and is mentioned for the sake of completeness but *is not the problem*, so please don’t call me sexist. The *other* parts of the description are the problem. Young captioners have not had enough time to read and write extensively and tend to lack wide general knowledge. Of course there are exceptions, and of course not everyone working in captioning in Canada is a young woman with a liberal-arts degree. But that is nonetheless the *trend*.

This lack of diversity in the captioning workforce – this monoculture – can explain why certain errors keep being made even across companies. Lacking the years under their belts and without a variety of educational and literary backgrounds, captioners simply do not know better.

I advocate better training for existing captioners, weeding out captioners who are not really suited to the job, and drastically expanding the diversity of applicants recruited into captioning. Our American friends have had good luck with a variety of approaches. I can suggest that the CAB start work on new ways to attract smart, seasoned, talented, and qualified people of all sorts to the field of captioning. (And audio description, for that matter.) As a question of human-resources development, this is the sort of thing the feds should also help out on, but we shouldn’t necessarily wait on that.

Recommendation and announcement

The Manual is not ready for distribution and implementation yet. I believe that will not come as a surprise to readers.

However, I can announce that the University of Toronto’s [Adaptive Technology Resource Centre](#) and the [Centre for Learning Technologies](#) at Ryerson University have received their first funding for Open & Closed, a worldwide accessibility training and standardization program that I originated. Industry Canada’s Canarie program has funded an initial project in online captioning standardization and training (that is, captioning video on the Web).

The goal of the Open & Closed project (expected to last three years) is summed up in its slogan: **Uniting global knowledge of accessible media**. Some of our objectives:

- a. Amass all documented standards in Web accessibility and in the four standard techniques of accessible media – captioning, audio description, subtitling, and dubbing.
- b. Produce recommended standard practices in all those fields, which may vary by technology (e.g., Line 21 captioning vs. movie captioning) and country (e.g., World System Teletext captioning vs. Line 21 captioning). Yes, there would finally be unified standard recommendations for each field.
- c. Deliver training for all fields in various media, including printed books, video and disc, distance education (mostly Internet-based), and classrooms.

As part of the process, we will attempt to license every captioning style guide in existence. It may make more sense for the CAB and its members to fold the current project into Open & Closed.

As an independent body affiliated with two respected academic institutions, a standardized English-language captioning style guide may be more widely implemented if it emanates from Open & Closed. A

French-language manual could also be created so that it does not conflict with the English-language manual. Keep in mind that our entire project involves research, training, and standardization; we can and will work on topics like the CAB manual day in and day out, while Committee members all are busy putting out captioned programming or managing entire stations. Subcontracting the current Manual to the Open & Closed project may make sense administratively.

Quite apart from that possibility, we are eager to discuss funding arrangements to carry out our research, standardization, and training work in various fields, of which two, captioning and audio description, are of immediate regulatory interest to broadcasters.

The Open & Closed project begins work 2002.10.01. A formal announcement will be made by one or both university partners and others.

“Canadian English-Language Broadcasters’ Closed Captioning Standards and Protocol”

Comments by [Joe Clark](#)

Severe errors

Style Guidelines preface

[p. 14] ¶ The goal of the following guidelines is to provide Canadian off-line caption editors and on-line caption stenographers with options and tools to use with discretion when making their decisions, so that there is uniformity and consistency across the Canadian captioning industry. These guidelines are the result of carefully considered research, observation, and experience and should be applied by Canadian closed captioning providers to the greatest extent possible.

- The goal of the guidelines must be to *limit* options. We have *too much* variation, mostly caused by one captioner’s efforts to be different from every other captioner or by untrained, inexperienced (and generally very young) staff deciding “this is how we do things here and that’s that.”
- I don’t see any “tools” in the Style Guidelines.
- Discretion is overrated. Give people “discretion” and they’ll come up with a tenth new and ill-advised way to handle a captioning problem that was already handled much better by someone else years ago. Give untrained, inexperienced (and generally very young) captioners too much discretion and *they will usually guess wrong*.
- To guarantee “uniformity and consistency” (redundant), the guidelines must be applied *wholesale*. This isn’t a shopping list. The role of this Manual is to *standardize* Canadian captioning, not to conveniently list several “options” that untrained, inexperienced (and generally

very young) staff can then pick and choose from.

- The References section barely accounts for a third of the captioning research I have in my own possession. The claim that the Manual is based on “research” is possibly exaggerated. I also dispute the role of “observation” as implied in this Manual: It seems to boil down to “I saw Captioner X do that the other day, and I personally like it, so let’s put it in the Manual.”

Four basic principles

[p. 14] ¶ 1. Put content first. Verbatim transcription is always the goal. Communicating the meaning and intent of the program must always take precedence over stylistic or aesthetic considerations.

I would say “stylistic or aesthetic considerations” need to be explained. By this reasoning, we would eliminate pop-on captioning altogether because roll-up captioning can caption verbatim nearly all the time while pop-on often cannot.

2. Edit responsibly. Adhere to established guidelines for editing and bear in mind that the only justification for deviating from the verbatim is to ensure sufficient reading time.

- There are no established guidelines for editing. Nor are Canadian captioners per se qualified to write them, because they are not linguists nor have they any grounding in psychology of reading.
- It is not true that “the only justification for deviating from the verbatim is to ensure sufficient reading time.” Taken to the extreme (not a hypothetical case: One captioner already captions in the extreme every single day), all programming would be edited to some magic presentation speed for no other reason than misguided and uninformed company policy. The Manual’s principle provides explicit license to edit every program. Moreover, the only incontrovertible reason to edit is buildup time, and even then only in pop-on captions. The explanation “We could not transmit all the words spoken” is much more persuasive than “We decided deaf people could not read all the words spoken,” which is what this discussion boils down to.
- The principle here also assumes that Canadian offline captioners are even competent at accurate transcription; with all-too-limited exceptions, they are not.

3. Be consistent. Document your style decisions and technical methods and apply them consistently.

This principle openly authorizes the establishment of competing style manuals. (Where else would “style decisions” be “documented”?) The goal of this Manual is to eliminate that outcome. How do Canadian captioners “document” style decisions when they don’t have the training, education, or research citations to do so?

4. Keep descriptions simple. Caption scripts should not be cluttered with excessive or complicated descriptive captions.

I don’t think this is important enough to become one of the Four Commandments, as it were. In any

event, actual audience research has clearly called for *more* notation on non-speech information.

A competing list: At a recent meeting of U.S. Department of Education–funded captioners and their consumer advisory panels (held 2002.09.14 in Washington), four competing goals of captioning were enumerated, and they are much more solid. (I had these dictated to me over the phone. When I get the actual transcript, I'll update the page, with appropriate notations.)

1. Transcribe accurately.
2. Make it clear who is speaking.
3. Notate non-speech information.
4. Make captions comprehensible.

This is a better point of departure than the Manual's current list.

Editors vs. writers

I submit that the Manual, and the entire Canadian captioning industry, must cease referring to captioners as caption *editors*. They are in fact caption *writers* (the Caption Center terminology). They write captions.

There's already too much editing in Canadian captions. Let's not provide implicit authorization through job titles.

"Captioner" is a word I use to mean person doing captions or firm doing captions, depending on context. It too is better than "caption editor."

Note that caption "editors" may feel they are of higher prestige than caption "writers." Having been both a writer and editor, I cannot support that position.

Editing

The Manual's examples of edited captions [p. 29] are in fact textbook examples of how *not* to edit captions, but fairly represent the skill level in evidence at present.

"He bought a wrench and he bought a hammer and a screwdriver, and a drill" is most aptly edited down to something like:

He bought a wrench and a hammer
and a screwdriver and a drill.

(15 words down to 13.) The suggested eight-word edit –

He bought a wrench, hammer,
screwdriver and drill.

is a sentence that would never be spontaneously uttered outside an ESL class. (At least add the word *a* before "drill.") It is condescending and patronizing to subject a viewer to that kind of editing, though at least one captioner would demand that *every* such sentence be so edited.

This advice –

[p. 31] ¶ Divide captions into very short phrases and distribute them evenly throughout a scene to average out caption durations to the acceptable maximum of three seconds per line.

– turns a slow scene into a techno music video. Very slow speech comes up so seldom that it does not require this degree of reconstructive surgery. Could it be that slow speech is slow for a reason, and tends to have built-in pauses that can be accurately mirrored in captions?

Why the laboured discussion about one- and three-line captions? Use as many lines as are necessary to accommodate the flow of the dialogue and reading speed. *Pulp Fiction* requires a different approach than *On the Road Again*. I am not convinced that a set of short captions is really better than a single big one in all cases.

Misunderstanding of eye-gaze in caption viewing; blink rate

[p. 15] ¶ Therefore, all obvious speech must have captions. Research shows that captioning consumers watch for visual cues from the faces of television speakers to direct their eyes to the caption area of the screen. If there is no caption, it creates a false alarm and considerable frustration. [7][vii]

[p. 32] ¶ Improperly divided text stops the reader, so that he or she spends time re-reading instead of reading quickly and then scanning the picture.

No, there are two eye-gaze studies that confirm what should be obvious to you if you reflect on your own caption viewing: You spend most of your time watching the captions. (In scrollup captions, you spend nearly all your time doing so.)

See Carl J. Jensema et al., “Eye-movement patterns of captioned-television viewers,” *American Annals of the Deaf*, 2000. “[S]omeone accustomed to speechreading may spend more time looking at an actor’s lips, while someone with poor English skills may spend more time reading the captions.... [T]here is some preliminary evidence to suggest that higher captioning speed results in more time spent reading captions than on a video segment.”

For segments with no captions, eye movements tended to zip around the screen. But for segments with captions, the preponderance of eye gaze dominated at the bottom of the screen. “The addition of captions apparently turns television viewing into a reading task, since the subjects spend most of their time looking at captions and much less time examining the picture.”

Rather interestingly, subjects were re-tested with the same videos a few days later. Subjects spent slightly more time looking at the picture than reading captions. That was also true for the hearing subjects, who had initially had less experience watching captions (“only the deaf subjects watched it regularly”). “Viewers read the caption and then glance at the video action after they finish reading.”

In other words, add captions to a program and people spend most of their time reading them. The effect holds irrespective of speed: 80, 106, 122, 197, and 220 wpm were used, with no significant difference in

results. Experienced captioning viewers – even if the experience comes from the first phase of this eye-tracking experiment and nothing else – spend less time focusing on captions.

This isn't the only evidence. See Jensema et al., "Time spent viewing captions on television programs," *American Annals of the Deaf*, 2000. "It was found that subjects gazed at the captions 84% of the time, at the video picture 14% of the time, and off the video 2% of the time. Age, sex, and educational level appeared to have little influence on time spent viewing captions. When caption speed increased from the slowest speed (100... wpm) to the fastest speed (180 wpm), mean percentage of time spent gazing at captions increased only from 82% to 86%." Four silent custom videoclips, captioned at 100, 120, 140, 160, and 180 wpm, were used in the study. All 25 subjects were deaf. On average, subjects spent 84% of their time looking at captions. The range was 82–86%. Variation in caption speed has no serious impact on the time spent watching the rest of the screen.

It just is not true that we watch the screen and then occasionally look down at captions. Now, that is certainly the case when a long stream of uncaptioned video is interrupted by a new caption, but that is an unusual event.

From the Manual:

A viewer watching action will need a change in caption shape or positioning to detect that a new caption has occurred. The pattern seems to be that a viewer first detects a change of caption, then reads it, then scans the picture until there is another caption change. [8][viii] Therefore, if two sequential captions have the same shape and placement, a change of captions may not be detected. It is therefore important to vary the placement of sequential captions when their shapes are identical.

How does one vary the placement of sequential captions? There are right and wrong ways. What are they?

In any event, what is stated above is not, in fact, "the pattern" of eye-gaze in caption viewing, according to Jensema's research.

It is not sufficient to argue that a change in caption shape or positioning will cue the viewer that "a new caption has occurred." For the better part of ten years, several but not all Canadian captioners have engaged in protracted negligence in creating captions with a nonstandard *blink rate*, the number of blank frames between captions. (Some Canadian caption writers have even claimed not to understand the concept of blank frames between captions.)

The best number, used for 20 years by every professional captioner in North America until VoiceWriter spread itself into Canadian captioning, is two frames. Vitac has used a blink rate of one frame for a couple of years, initially for talkier programs like *The Practice*, and it's OK in general; not coincidentally, Vitac was also the first captioner to commit to verbatim captioning whenever possible. A blink rate of zero (widely used in Canadian captioning) makes it too difficult to tell when captions have changed.

Higher blink rates make caption changes too conspicuous. A blink rate of four frames – the default in VoiceWriter and widely seen elsewhere – makes captions blink on and off like a turn signal on a dashboard. They are nonstandard captions. But worse yet, too-high blink rates force captioners to edit.

Each frame of video can carry up to two visible caption characters. (The number is not exactly two in all cases due to control codes and similar issues.) If a captioner uses a blink rate of four versus two, after three captions have appeared and disappeared, six additional frames have been lost in blink rate. That's

12 characters – two five-letter words plus their trailing spaces. After every three captions, captioners have no choice but to remove two words from subsequent captions just to maintain presentation-speed parity with captions everyone else has used for 20 years.

Also, be aware that captions remaining on the screen too long are likely to be re-read by the viewer, causing confusion. [9][ix]

It's pretty uncommon for a caption to stay visible for too long. I generally see that after an encoding error where the screen-blank command (a single pulse) is missing.

It's good practice to sow screen-blank commands through extended periods of no captions (even five seconds is long enough to do so). Among other things, people channel-surf, causing their caption decoders to pick up caption characters that are then displayed with the next full caption. Sending out clearing pulses from time to time reduces or eliminates that problem.

Reading speed

Reading speed – that is, *speechnate* (write it as two words if you wish) or *presentation rate* – is a much less contentious issue than Canadian captioners seem to believe. (How contentious is it? I got yelled at for 80 minutes in a meeting on this very topic. Why? It was claimed that big-D deaf people cannot read faster than 150 words a minute, hence every program must be edited to that speed, even when using roll-up captions. Note well: My use of the term “yelled at” is not hyperbole.)

There's solid evidence for the following:

1. Deaf viewers don't want captions edited to a maximum reading speed derived from presumptions about literacy; such editing is rightly viewed as condescension and a form of inequality.
2. Speeds of 170 words per minute are comfortable for long periods. Reading at up to 200 or more words per minute is quite possible.
3. Caption reading is heavily influenced by visual acuity. In fact, whenever deaf people complain that captions are hard to read, your first course of action is to find out how recently they had their eyes checked and how far away from the television they're sitting; altering captioning practice shouldn't even be discussed at that point.

See Jensema, “Viewer reaction to different television captioning speeds,” *American Annals of the Deaf*, 1998. “Participants used a five-point scale to assess each segment's caption speed. The ‘OK’ speed, defined as the rate at which ‘caption speed is comfortable to me,’ was found to be about 145 words per minute (wpm), very close to the 141 wpm mean rate actually found in television programs.... Participants adapted well to increasing caption speeds. Most apparently had little trouble with the captions until the rate was at least 170 wpm. Hearing people wanted slightly slower captions. However, this apparently related to how often they watched captioned television. Frequent viewers were comfortable with faster captions.”

See Jensema et al., “Closed-captioned television presentation speed and vocabulary,” *American Annals of the Deaf*, 1996: “The average caption speed for all programs [surveyed] was 141 words per minute, with program extremes of 74 and 231 words per minute.... The percentage of script edited out ranged from 0%

(in instances of verbatim captioning) to 19%.”

When it comes to editing captions as extensively as was done on *The Captioned ABC News* in the 1970s, “almost everyone now considers [that] overediting.... Deaf viewers wrote letters to caption companies indicating they wanted access to whatever was spoken on the audio and that captioners should not play the role of censors. According to conversations with captioning company officials, caption companies have tended to interpret this as meaning deaf people want straight verbatim captioning.”

The study documents the reality that caption speeds in excess of 150 wpm are common and that deaf viewers not only never asked captioners to alter the dialogue allegedly to suit them, they asked for the exact opposite.

See Captioned Films/Videos Program, *Captioning Key: Preferred Styles and Standards*, 1995: “CFV adult special-interest videos require a maximum presentation rate of 150–160 wpm. Special-interest videos include adult education, self help, and how-to-do materials.... CFV theatrical (movie) videos are captioned at near-verbatim rate. However, in general, no caption should exceed 225 wpm.”

See the *Television Broadcasting Services (Digital Conversion) Act 1998 Draft Captioning Standards* (Australia), 1999: “A caption should stay as close as possible to the original wording while allowing the viewer enough time to absorb the caption’s contents and still watch the action of the program. To meet the second criterion, programs for adult viewers are captioned at a reading rate of 180 words per minute.... Programs should be captioned at an appropriate reading speed for the intended audience. For adults this is 180 words per minute.”

Foreign-language subtitling is widely understood to use slower presentation rates than captioning – in many cases, maddeningly slower rates for hearing people used to captioning. But even the one and only book on subtitling in print advocates speechrates near 170 wpm. See Jan Ivarsson and Mary Carroll, *Subtitling*, 1998: “[Film distributors] have established a norm, which may be expressed as follows: 2 lines = 80 characters = 8 feet of film = 128 frames = 5½ seconds. This results in a reading speed of around 175 words per minute.”

The Manual’s advice to stay under 200 wpm is fine (though much of its advice is contradictory and misunderstands how captions work). But more explanation is necessary, because at least one Canadian captioner is willing to scream at you for an hour about the necessity of editing captions to a childish reading speed.

I would suggest a rewrite of this section. In particular, advice like “Leave two-line captions onscreen for three seconds” is too rigid and too vague all at once. Is that one line of dialogue plus a single-line speaker ID? Is that two 16-character lines? Is that two lines of edge-to-edge speech with italics and quotation marks and scientific vocabulary?

Further:

[p. 29] ¶ Off-line caption editors are reminded that they are likely to have a faster caption-reading speed than the average viewer. Therefore, they should not rely on their own reading skills to judge the pace of words.

In fact, Jensema’s evidence shows that experience reading captions makes it *easier* to read faster captions. Captioners merely have a lot of experience reading captions. (Actually, that is untrue. Not many captioners in Canada watch captioned TV at home. Squinting at a little MPEG window in Swift as a horribly distracting cavalcade of 30-point caption text streams by alongside does not constitute “reading

captions” in my book.) But caption *viewers* have a lot of experience reading captions. The advice above is exactly wrong: Captioners’ own reading *is* a valid basis of comparison.

Moreover, no one seems to have acknowledged that most households now have VCRs. If some deaf subgroup has a hard time reading captions, well, there’s probably no speed they could reasonably keep up with, and maybe they should just start taping programs and rewatching them rather than expecting all captions to be edited their level.

For this reason, if you’re captioning a program that you know will be released on home video first, you might be able to get away with faster presentation speeds *if it is genuinely necessary to represent the program*. The viewer can always rewind. If you’re given a TV series to reformat for DVD (heaven help us when a Canadian captioner is given that many tools to misuse), recaption nearer verbatim.

Visual acuity

Another overlooked issue in evaluating all aspects of caption readability is the visual acuity of the audience.

Sondra Thorn, a deaf optometrist *who used to work in captioning* (I corresponded with her in the 1980s!), has written two papers on caption-reading abilities with impaired vision. In addition, the Royal National Institute for the Blind has surveyed the television viewing habits of visually-impaired people (though I don’t have a research reference).

Conclusions? People sit too far from their TVs to read captions; their vision is often not corrected properly (though some vision deficiencies make it hard to read text on a screen anyway); and, most importantly, even tiny amounts of blur make it impossible to read fast captions.

Thus, captioning viewers should be encouraged to sit closer and get proper glasses before insisting that captioners do things their way. Captioners, and contributors to this Manual, are advised to take the requests of individual viewers, irrespective of hearing status, with a grain of salt. In some other cases, I have heard of captioners rewriting their style books in response to a single complaint, and in a couple of those cases the rewrite violated standard English orthographic rules.

Reference: Thorn, F., & Thorn, S., “Television captions for hearing-impaired people: a study of key factors that affect reading performance” *Human Factors*, 1996.

False starts and utterances

This section makes no sense at all:

[p. 15] ¶ Non-verbal utterances, repeated words and false starts to sentences are not generally included in real conversation. However, if they contribute to the understanding of a dialogue, dramatic effect, joke, or personality, they must be included.

- “Nonverbal” (it isn’t hyphenated) merely means “without words.” Clearing your throat is an utterance, as is a groan, but neither of them is “verbal” even though we would be very inclined to

caption them anyway.

- I don't know what "are not generally included in real conversation" means. It seems to be an indirect way of saying "Proper ladies and gentlemen would hardly do such things."
- If the passage is supposed to mean "Human speech is replete with nonverbal utterances, repeated words, and false starts," keep in mind that Noam Chomsky said the same thing in the 1970s and it was taken as gospel for years until somebody actually bothered to listen to real human beings for five minutes and realized it was nonsense. Later studies of actual human conversation (the field of linguistics known as pragmatics) disproved the theory that people have such poor command of their native languages that they can't even talk in proper sentences. (Cf. [p. 9]: "People involved in real conversations rarely use grammatically correct sentence structure. They use improper grammar, incomplete sentences, run-on sentences, slang, vernacular expressions, and so on.")
- Anyway, captioners cannot impose value judgements on the speakers they represent. If someone coughs or is repetitious or is too nervous to get a sentence underway properly half the time, those habits govern our captions. I thought our goal was verbatim captioning.
- What this section attempts to explain – exactly which nonverbal utterances should be captioned and how *narrow* our caption transcription should be – needs to be dramatically expanded.

Onscreen information

[p. 15] ¶ Captions should not cover graphics or keys, characters' eyes or lips, or areas of sports action.

Captions always have to cover something. Very large graphics will end up being covered. In certain unusual cases, captions will cover keys (a term not defined to this point); more commonly, captions cover keys because captioners don't bother to move the captions.

The prohibition of covering characters' lips is overrated. The common example of one character sitting at a desk and another standing behind that character forces captions to alternate between top of screen for sitting person and bottom of screen (at opposite side) for standing person – hardly worth the trouble, considering how small the sitting person's lips are.

In closeups, where this rule makes at least some sense, your options may be to cover lips *or* cover eyes. (Or nose, I suppose.) Which do you choose?

I'm not sure we can afford to be all churchy about covering certain parts of the screen. Covering the screen is what captioners do. There are times when we have no choice but to cover things; it's not always unrecoverable (we can move the caption a moment later, for example, in the case of a full-screen photo, map, or illustration) or even always important.

Justification

The Manual fundamentally misunderstands the role of justification in captioning. Justification refers to the placement of margins, and there are really only two kinds, flush left (left-justified, quad left) and flush right (right-justified, quad right). Centred captions – that is, captions without true centring but no flush margins, either – are not said to be justified. That may come as a surprise if, in your word

Centred roll-up captions

[p. 15] ¶ Roll-up captions may revert to a 16-character line with left justification, and may be positioned at the far left of the screen or at the centre of the screen.

- Why a 16-character line? Why not just whatever width is advisable to avoid obscuring, for example, sports scores and clock?
- Why no mention of the fact that you can use any tab stop, or indeed any character position, as the left margin?
- Centred roll-up captions must be outlawed save for brief display uses, as in the captioner credit. I had never seen such a creature in over 20 years, but suddenly last year they began to pop up. (More novices entering the field who figure they know better, I presume.) Such captions are very hard to read: You then have to track two dimensions at once – upward caption scrolling and an ever-changing left edge.

Missing words

[p. 15] ¶ If spoken words or lyrics are different from a textual graphic, for example, when there is talking over end credits, full captions must be included and moved so as to interfere as little as possible with the essential visual elements, using one of the above techniques.

Also:

[p. 16] ¶ Do not start or end a caption in mid-sentence assuming that a textual graphic will be read in the correct sequence to complete the captioned sentence. For example, the complete phrase, “Tonight’s program is brought to you by Sunnybrand Detergent” should appear in captions, even if the Sunnybrand Detergent logo appears as a full page graphic while the captions are displayed.

- The Manual should state that it is illegal to *refuse* to caption over end credits (because, for example, you think it looks too “cluttered,” or the idea simply scandalizes you, or that is simply the arbitrary policy of your shop).
- If captioner credits must appear at the end of a show, but dialogue continues during that time, under battle conditions it can be necessary to remove dialogue. But remember, you can always display captions on two parts of the screen at once (in pop-on, anyway); double up if necessary.
- When speech exactly duplicates onscreen text, usually there is no reason to place a caption. If a speaker ID is necessary, use it. If words are missing, fill them in to the extent possible, or just recaption the whole phrase. Examples:
 - Key: The quicker picker-upper
Speech: It’s the quicker picker-upper.
Caption: It’s...
 - Key: GM light trucks
Speech: 2% APR financing on GM light trucks. See dealer for details.

Caption: 2% APR financing on GM light trucks.
See dealer for details.

- Key: May cause drowsiness

Don't use if nursing Speech: Folderol may cause drowsiness. Don't take Folderol if you are nursing. Caption: Two lines, one at top, the other at bottom –

Folderol...

Don't take Folderol if you are nursing.

- If a host reads the entire text shown onscreen, ID the host's name if necessary, but do not caption, for example [Irshad reading text on screen]. (We figured that out!)
- Do *not* leave out an ellipsis or, in some cases, a colon when expecting the viewer to traverse from caption to onscreen text to complete an utterance. (One Canadian distributor is notorious for that practice.) A caption hanging by itself with no punctuation on either side nearly always means some utterance preceded and followed it in the same sentence.

The topic of handling subtitled passages or programs is not addressed in the Manual and needs to be. I have never seen a Canadian captioner handle it correctly.

Children's captioning

[p. 16] ¶ However, the reading rate for preschool children should be considerably slower than for adults: four to six seconds for each line of text. It is appropriate to edit text for preschool children using very short, concise sentences that do not fill a line. One-line captions are best.

These assertions are at best severely debatable and contradicted by practice. The only credible research basis for a decision of verbatim vs. edited captions is presently underway and will not be published for many months.

Children's programming is found in different forms. *Teletubbies* is famous for having been created specifically for kids too young even for *Sesame Street*. An enormous range of children's programming is not educational in nature and uses adult-level dialogue. (You may find the storylines clichéd and insufferable, but the characters talk like adults in a syntactic or pragmatic sense.)

Remembering that captioners have a responsibility to the program and to viewers, I don't see how it is defensible to aggressively edit all children's programming. In any case, that isn't what is presently done. (I wouldn't use Canadian-captioned programming on Canadian children's channels as examples given the poor expertise at work. Look at U.S.-captioned shows.) Even *Teletubbies* is captioned verbatim!

The only real-world experience we have with heavily-edited children's captions revolves around *Sesame Street*, which, to my knowledge, only NCI has captioned through its entire run. From day one, the show has been heavily edited, and uses a different approach to non-speech information (outright use of child-level onomatopoeia, like moo, cluck, meow). In recent years, captions are also closer to

characters' heads than ever before. But I am aware of no documentation whatsoever of this decade-long captioning practice. (It is typical of NCI not to document anything despite its nominal status as a captioning "institute.")

The other example is [Arthur](#), with near-verbatim grownup-style captions on CC1 and heavily-edited kids' captions on CC2 (with a different approach to NSI). It's an ongoing WGBH research project that will give us our first reliable evidence on this question. One outcome that WGBH confirmed is possible is as follows: Very young deaf kids cannot read and pretty much cannot follow whatever captions we give them. If true, then we have even stronger reason than ever to caption adult-style because adults are the only viewers who can read the captions.

One must also keep in mind that captioners are not trained grammarians. Canadian captioners are not really all that good at editing adult-level dialogue, let alone altering syntax to be more understandable by kids. I have a couple of very old papers providing guidelines on caption editing for kids, but to describe such papers as advocating "very short, concise sentences that do not fill a line" would be an absurd exaggeration.

It is also possible that roll-up captions might be easier for kids of certain ages to follow than pop-on captions, but that is supposition.

Case

Absolutely the most significant error in the Manual concerns case of captions. Real-time captions are very difficult to create in upper and lower case, though Waite & Associates manages it and it's the only way French and Spanish real-time captions are done. A case can be made that all-upper-case real-time captions are a necessary evil.

However, all offline captions *must use correct upper and lower case* just like the rest of the English language. I cannot overstate the importance of this requirement. For 20 years, we've enjoyed the full, glorious English language in print yet **HOURS AND HOURS A DAY OF SCREAMING, SEMILITERATE ALL-CAPITALS CAPTIONING ON TELEVISION**. It's time to grow up and start writing captions that actually follow English orthography.

Let's recap the facts:

1. The one *and only* reason why offline captions were typeset mostly in upper case *in the early days* is because of illegibility of fonts in the original two decoder generations. The original designers of the Line 21 spec were American engineers. Their knowledge of typography and the psychology of reading was nonexistent.
2. Original decoder fonts were made up of a 7-by-5-dot matrix with no descenders, a term captioners need to know like the backs of their hands. A *descender* is any part of a character that extends below the imaginary baseline on which characters sit. Lower-case letters with descenders are *gypqj*. Some other characters can descend, but they are not relevant to captioning.
3. Only one or two pixels (out of a grand total of 35, remember) differentiated letters like *e*, *g*, and *s*. As in all monospaced fonts, wide letters like *W* and *M* were scrunched and narrow characters like *l*, *I*, and *1* were too widely spaced.

4. Given all this, capital letters were deemed *less illegible* than mixed case.
5. But the English language *is not written in all capitals*. Certain limited applications may use all-caps setting (just who sends a Telex in the year 2002?), but 18 hours a day of captioning on every station on the dial do not constitute such an application.
6. Moreover, caption fonts improved. TeleCaption II fonts used a sharper dot matrix but still did not have descenders. Even so, the lower case was adequately legible. (I still own a TeleCaption II.)
7. EIA 608-compliant decoders can use any font a manufacturer wants. Please accept my word that I have spent a great deal of time since the enactment of the Television Decoder Circuitry Act in 1993 comparing caption fonts. I can tell you categorically that few televisions have fonts with illegible lower-case letters, and a great many of them have highly-legible fonts complete with descenders.
8. Again, I know this from experience: My television is an el-cheapo model with – get this! – no descenders. The font is more than adequately legible – much better than my old TeleCaption II. I read mixed-case captions with ease.

If mixed case is really “worse” in captioning, then why has it been so heavily used?

1. The Caption Center, Captions, Inc., CaptionMax, and Vitac have all used mixed case for NSI and speaker IDs since day one (or at least since before living memory). NCI has used mixed case for speaker IDs for most of the last decade. The Caption Center has always used mixed case for voices processed through telephones, communicators, and the like. NCI uses mixed case for whispering. That’s 90% of American network captioned shows and home video right there.
2. CBC used mixed-case captioning on *The National* throughout the 1980s. French captioning is mostly mixed-case; some backward French-language captioners cling to the horse-and-buggy era and caption in screaming capitals, which has serious implications for meaning in the French language. (Chrétien is still Chrétien even if you write it as CHRETIEN. But how about CHRETIEN TUE?)
3. The Caption Center has captioned essentially everything in mixed case since fall 2000. It is possible to do real-time comparisons between shouting capitals and mixed case just by channel-surfing. (Or by watching earlier and later episodes of the same shows, like *The X-Files*.) The oldest captioner in the world, one that is still a very large captioning provider, captions in mixed case, which in and of itself *proves* there is no problem.
4. Looking a bit beyond Line 21 for a moment, MoPix captions for first-run movies, and Tripod open captions, use mixed case. Nearly all DVD subpictures (except those created by our horse-and-buggy Quebec friends) are in mixed case.

The Manual itself countenances the use of mixed case for long passages and for NSI and speaker IDs. That amounts to yet another example of collapsing complex auditory phenomena onto relatively small and unexplained orthographic alterations.

But in the former case, what evidence can you provide that deaf and hard-of-hearing viewers immediately and transparently understand that a long passage of sudden mixed-case captions signifies, say, a literary reading? And in what other realm of English orthography would we write everything but literary readings in capitals?

Answer: We wouldn’t.

Now, please try to explain how the use of all-capitals setting for closed captions is still supportable. Save for real-time captioning, there is *no evidence whatsoever* supporting its use and an avalanche of evidence against it. It is an open-and-shut case, no matter how much it upsets Canadian captioners' apple carts.

Upper-case captions can and should (indeed must) be used as a differentiator in speaker IDs. Compare:

- Scully:
Mulder?
- SCULLY:
Mulder?

The Caption Center was eventually persuaded to adopt the latter style, and now it's working fine. All-mixed-case captions in which even speaker IDs use mixed case do not provide enough differentiation. However, mixed case can and should (indeed must) be used for non-speech information.

I've written two articles on this subject by now, one of them [current](#).

By the way, I'm not quoting the entire section on this topic [p. 17] because it is horrendously wrong from start to finish.

Also, converting to mixed case obviates the entire discussion of acronyms [p. 20]. Just write them in capitals (no, *not* italic capitals), unless the context suggests the acronym would be misread as a word.

That case does not come up often – I . R . A. is the only example I can think of. (Well, possibly A . D . for "audio description.")

Matching onscreen orthography

E-mail addresses do not have to be in mixed case. Domain names are case-insensitive by [RFC specification](#). (Try it sometime: Mix case all you want in domain names in a browser and in E-mail. Any permutation will work.) UserIDs can *theoretically* be case-sensitive but never actually are, except on ancient X.400 mail systems, of which there are almost none in existence. But there's no reason not to match the onscreen orthography in those cases.

[p. 17] ¶ To match on-screen case for proper names and titles, and use the spelling preferences of performers such as k.d. lang (not K.D. Lang) whenever possible.

Let's not get ridiculous. Should marchFIRST, KISS, *thirtysomething*, *sex, lies & videotape*, and other orthographic contortions be indulged? Hardly. A weak case can be made that captions must match onscreen type; a fair case can be made that a TV commercial must follow the corporate orthography because it is an example of flat-out corporate salesmanship.

A stronger general case can be made that examples like those listed above do *not* have the same orthographic etymology as intercaptioned names like WordPerfect or even MacDougall and are merely affectations promoted by questionably-literate marketers. It is not our job to abet such marketers.

Also note that sentences that end in URLs or E-mail addresses take standard end punctuation.

Colour

[p. 17] ¶ Captions usually appear as white text on a black background because this is the best combination for visibility. Colour captions have been successfully used as a special effect in music videos (not music segments within a program), and as an effect with certain voices in dramatic stories, but they have generally tested poorly both as an indication of speaker identification and as an indication of emphasis. [10][x] Colour captions can never be used as the sole indicator of who is speaking. Proper placement and speaker identification are always required because colour can be difficult to discern against the video background. Use of colour captions is discouraged until such time as research is conducted to develop proper guidelines for their use.

- Captions appear on a black background because engineers designed it that way (save for transparent spaces).
- White on black is not necessarily more visible than black on white, but it probably *is* more visible against luminous backgrounds, as TV images by definition are. (There's actually a great deal of research on legibility of screenfonts.)
- Colour captions test poorly because nobody has seriously tried to use them. Deaf viewers (particularly big-D deaf viewers) are notoriously adamant that captions in 2002 must look exactly the same as the first captions they ever saw way back in 1979. They will resist every effort to add colour until someone bites the bullet and uses it consistently and well on a program that merits it, of which there are arguably only a few (*Metrosexuality* is an obvious choice).
- The Manual fails to address a possible middle ground: Using yellow for all speech and white for speaker IDs and NSI, which would work fine in every case save for a bleep inside a word ("mother[beep]er"). (It wouldn't work there because turning a colour on or off requires the insertion of a visible space character, except at the beginning of a line.) The downside here is that the inserted space character reduces our line length to 31 characters. I am not sure this is not worth trying consistently on a couple of shows.
- "Colour can be difficult to discern against the video background" is untrue. To my knowledge, caption background colour must be black by default, but 608 decoders can optionally permit a different colour (usually translucent or clear). By default, colour captions will have a black ground.
- There is no discussion of colourblindness. Red is the worst foreground colour to use for a range of reasons (undifferentiable from green for protans and deutans; protans see a dark substitute colour, meaning red-on-black text may effectively disappear).

Italics

Recall that the Manual must be written to outlaw current mistaken, unjustifiable, or egregious practices of Canadian captioners.

Chief among these is promiscuous misuse of italics, including use for some applications the Manual

authorizes, like product names.

- Ridiculous on its face (the *only* print publication that italicizes product names is *Consumer Reports* – hardly a paragon of copy-editing), the practice becomes even more ridiculous given the huge range of topics covered in TV commercials. I have literally seen *Jamaica*, *lottery*, and *MS* italicized in commercials.
- Use of italics for product names is a misguided, half-baked, and self-compromising effort to make up for the inadequacies of all-upper-case setting. Contrary to common misunderstanding, we are perfectly able to discern that a word refers to a product without the use of italics. Switch to mixed case and the initial capital letter takes care of that problem handily.
- The practice involves outright toadying to advertisers: Who says that Tylenol is so important it deserves italics?
- In ordinary programming, it gets a bit ridiculous reading about **THE MINISTRY OF NATURAL RESOURCES** or a **HIS RUSTED-OUT TOYOTA TERCEL**. Is someone *shouting* those words? Are they *emphasized* in some way? (Don't lines of continuous capitals already represent SHOUTING?)

Use of italics must follow English orthography. Style manuals (you cannot get any better than *Chicago*) offer an exhaustive list of permitted, required, and prohibited uses.

This section mentions foreign phrases. Use italics unless the phrase is very well integrated into English, like *lederhosen*, *jihad*, *dépanneur*, or *rendez-vous*. I think the example given (*deja vu*) is sufficiently anglicized for roman setting (note the absent accents). Consult a recent Canadian dictionary and follow its advice.

There is no adequate coverage of use of italics for offscreen speakers (taken to absurd extremes by Captions, Inc.), thinking, inner voices, and narration.

Role and responsibility of captioning

[p. 5] ¶ The Manual asserts:

Television is recognized as the most popular source of information and entertainment in the world. By making television programs accessible with closed captions, Canadian broadcasters facilitate the involvement of Deaf, deaf, deafened, and hard of hearing people in popular culture. Closed captioned television programming provides these groups of people with accessible, screen-based, cultural, historical, and educational communications. Caption providers, therefore, bear an important responsibility to caption viewers. This is the most compelling factor in the creation of standards for closed captioning.

- Television does not limit itself to popular culture, nor to “cultural, historical, and educational communications.” I suppose someone’s trying to sum up TV in a grand adjective chain, but it is inadvisable to try.
- “[P]rovides these groups of people with accessible, screen-based, cultural, historical, and educational communications” is an attenuated and disengaged, not to mention incomprehensible

and clumsy, explanation of what captioning actually does. Try: “Captioning makes the full variety of television programming accessible to these groups.”

- Caption providers bear a responsibility to viewers *and to the program*. You are messing with *someone else’s show!* You have an obligation to art and posterity to provide intelligent, rational, accurate captions. The claim that only viewers are important gives outright license to captioners to do whatever they want to the program in the name of these viewers, many of whom they have never even met.
- I can’t even figure out what “this” in “This is the most compelling factor in the creation of standards for closed captioning” refers to. What is the antecedent? In any event, “the most compelling factor” is obviously the principle, backed up by laws as strong as the Charter of Rights and Freedoms, that people with disabilities have a right of access to television programming.

How captioning is treated

[p. 5] ¶ Canada’s broadcasters are committed to making television accessible to everyone and are therefore committed to treating closed captioning with the same responsibility and sensitivity to their audience as they treat the aural and visual elements of television. To achieve this objective, they have been instrumental in advancing technology so that captioning is of the highest possible quality.... Every year, the broadcasting industry invests significant resources in high quality program captioning to meet the needs of Deaf, deaf, deafened, and hard of hearing people.

- Use the term “auditory,” “audio,” or “audible” in preference to “aural,” a homonym with *oral*.
- Canadian broadcasters demonstrably are not committed to making TV “accessible to everyone” or else all programming would have been captioned and described by now. Were this actually true, programs would not be aired without captions and descriptions *at all*; that is the only imaginable manifestation of such a commitment. You’re either “making television accessible to everyone” or you’re not. Just how many uncaptioned and/or undescribed shows are broadcast on Canadian TV?
- As a rule, broadcasters take every shortcut possible, and avoid every possible expenditure, in delivering captioned programming. There is *no* broadcaster in Canada who treats captioning as seriously as main audio and video. If readers feel I’m mistaken, please prove it by naming a broadcaster that routinely airs programming *without* main audio or video, a circumstance comparable to airing a show without captions or descriptions. Now name a broadcaster that airs every program *with* captions and descriptions. Isn’t that what “treating closed captioning with the same responsibility” must mean?
- Canadian offline captioning quality is generally poor, a declaration that is easy to prove by watching similar or even identical shows with Canadian and U.S. captioning. If Canadian captioning is “of the highest possible quality,” why do we need this Manual? Wouldn’t there be no problems left to solve?

We are attempting to provide baseline training for captioners. This section is demonstrably untrue, distracting, and irrelevant to the Manual’s mission.

Use of the term Deaf

[p. 6] ¶ It is necessary to explain that some deaf people (that is, the signing deaf) contend they constitute an identifiable culture, and that they believe “deaf” should be capitalized when referring to that group and no other.

Big-D deaf people exist. But the Manual is not a sociological treatise. Specifically related to captioning, if certain practices must be engaged or avoided because of *demonstrable* rather than asserted characteristics of, say, prelingually-deaf viewers who use sign language as a preferred method of communication outside the written medium, that’s fine. Similarly, if other practices are indicated or contraindicated for, say, hard-of-hearing or hearing viewers or any other group, that’s fine, too.

I merely object to the Manual’s complete acquiescence to the philosophy that big-D deaf people are in some way more important than any other group. If they weren’t deemed more important, why would they be able to reuse a word and capitalize it?

Referring to the typical viewer base of captioning as “deaf and hard-of-hearing viewers” is sufficient. The overlong and hypercorrect phrase “Deaf, deaf, deafened, and hard-of-hearing viewers” is superfluous and pandering (and uses three of the same syllable all in a row). When specific issues require explanation based on the linguistic or other characteristics of a subpopulation, by all means use relevant terminology and go into all necessary detail.

Authentic language

[p. 7] ¶ Through captions, viewers can read authentic language and see how it is used in meaningful situations.

Given how much and how badly Canadian offline captions are edited (one broadcaster edits *everything* without exception down to 150 words per minute, with severely limited skill and finesse), it is untrue to claim that Canadian captioning viewers “can read authentic language.” Edited Canadian captions read like nothing a human being would ever spontaneously utter (and like nothing a screenwriter would ever spontaneously write). The author of these remarks has watched captions for 20 years, has published nearly 400 articles and a book, is an experienced editor, writes clean copy, and has experience working with seasoned proofreaders with decades of experience, so please accept my credentials in leveling this criticism.

Nonetheless, please keep in mind that all I need to do is transcribe the actual audio of 15 minutes of one or two Canadian shows and reproduce the corresponding caption text to put the lie to the contention that Canadian offline captions always or even usually represent “authentic language.”

Also, the phrase “and see how it is used in meaningful situations” is better recast as “and understand its usage in TV programming.” I would not go so far to say that every situation is meaningful, nor that TV programming represents actual English as used in conversation. There are too many levels or *registers* of English used on television for that to be true.

Hearing captioning viewers

⋮

[p. 7] ¶ Now that closed caption decoders are built in to most television sets, the audience for captioning has grown to include those who can hear, but who choose to watch a program in silence. They may choose to read closed captions when others are sleeping, on the phone or studying, or they may prefer reading captions to listening.

This paragraph seems to knowingly render invisible those hearing captioning viewers who watch captioned TV with the sound on. I am not the only one, and it is unwelcome to be rendered officially nonexistent in a captioning manual.

In fact, irrespective of whether or not they keep the volume turned on, hearing viewers are [probably the majority audience](#) now.

Succinct phrases

[p. 8] ¶ They break up the transcript into succinct phrases, which will either roll up on or pop on the video screen.

I wouldn't say [succinct](#) ("Characterized by clear, precise expression in few words; concise and terse"). This too implies an editorial function on the part of the captioner, whose presumptive job is to hack, delete, skim, lighten, and alter the source rather than represent it.

Captions can be quite lengthy. I have read innumerable three-line captions that are the right length, and none of them are particularly concise (since a three-line caption can contain roughly 96 characters, or about 20 five-letter words). Captions need not be "succinct."

Roll-up captions are not divided at all, really. (One can imagine a few exceptions, like the aberrant individually-placed scrollup captions that kids are adding to programs these days. I don't consider sentence-ending carriage returns "divisions.")

Try: "In the case of pop-on captions, they break up the transcript into individual segments [or chunks]. In the case of roll-up captions, they transcribe the audio in full and add linebreaks and other typographic divisions."

Caption error-checking

[p. 9] ¶ Preparation: Off-line caption editors must take the time to carefully research all names and unfamiliar words or phrases that occur.

Caption editors must strive for accuracy and are reminded that the ideal way is to have a second person screen for errors, and a third person do a final pass with the audio off.

- Canadian caption editors demonstrably *do not* research names and unfamiliar words. They can't even get familiar words right, like *its* and *it's*.
- I know that captioners like to pretend the best way to evaluate captions is with sound off, but it

simply is not true. We are hearing people; only hearing people can do captioning because we need to be able to hear to transcribe *and differentiate* auditory elements. We have access to more information than a viewer who cannot follow the audio. Endemic captioning errors, like failing to identify speakers (including narrators and announcers, disembodied voices, thinking, reading to oneself, and gender of all the above), are *untraceable* when viewed with sound off. You also cannot tell what speech or NSI has been missed; you can't hear it! And those are only *some* of the failings associated with error-checking with the sound muted.

- Also, people are not very good at spotting their own errors in any medium, particularly when reading off a computer monitor. Having your work reviewed by someone else in the office, who, in the Canadian experience, won't know any more than you do and makes all the same captioning mistakes, merely ensures that typical errors not spotted on computer screens will continue not to be spotted. You can't fix what you don't know is wrong, and applying consistent style when the style itself is maldesigned doesn't get you anywhere.
- Checking on a monitor is necessary, but it is not sufficient. One other person should watch the entire program *with sound on*; a third person should proof a printout of the caption text, which I guarantee you will result in spotting more errors than onscreen proofing alone.
- Capital letters are harder to proof than mixed case, yet another reason not to use all capitals for captioning.

Title area and broadcaster responsibility

[p. 12] ¶ Because captions usually cover the bottom three lines of the television screen, efforts should be made by broadcasters and producers not to put essential visual information in this area. Whenever possible, graphic information should be placed well above the safe title zone, so that there is room for both captions and graphics to display.

This isn't accurate and isn't very useful advice.

- Captions can occupy up to four lines at a time anywhere on the screen. If the totality of a day's captioned programming on a given station were considered, I doubt that most of it would "cover the bottom three lines of the television screen."
- Broadcasters should do nothing special in ordinary full-screen programming to accommodate captions. *We* have to accommodate the *broadcast*.
- You do realize that this advice would sequester essentially all onscreen Chyrons to the middle of the screen? We would go from having Chyrons at bottom and captions just above (where some portion of the screen is obscured, but only for captioning viewers) to having an unaccustomed, and usually more important, portion of the screen obscured for *all* viewers. Since no broadcast designer in his or her right mind would bother following this advice, why bother offering it?
- If this were, say, 1989, when full-screen positioning were impossible, the advice might make sense. But captions can now move anywhere on the screen to avoid onscreen titles. They don't have to stay there permanently, but we have all the tools we need to avoid clobbering onscreen Chyrons without getting laughed at by TV designers when we ask them to rejigger decades of evolution in television graphics to solve a problem we can solve ourselves.

Buildup time and first captions

[p. 12] ¶ Because caption data pulses occur before captions appear, the first caption of each program segment must occur at least 15 frames (half a second) into the program or its time code cue may be lost in the transition from commercial to program. Also, caption data must be blanked at least 15 frames before the end of each program segment so that irrelevant captions do not bleed into commercial breaks or other programming.

This isn't very clear, and is a bit too liberal.

- Seven frames is the generally-accepted standard. (You may not have heard of it, and it may not be done in your shop, but I don't know a U.S. captioner who doesn't follow this rule. They've been at it longer and they turn out more captioned programming; Canadian audiences perennially prefer to watch Canadian shows.) Captions may not begin to be transmitted until the seventh frame after the very beginning of the program and after any commercial break or similar clear disjunction in the tape. This adamantly *does not apply* to shot changes or scene changes.
- You wouldn't need to blank captions before the seventh-to-last frame in a program or before a commercial break or other disjunction. In practice, I don't see why you need to do this before the third-to-last frame (leaving two blank), but seven won't hurt.
- Fifteen frames are too much in both cases.

Ellipses

Ellipses must not be used between captions if no pause is present. In other words, don't add an ellipsis to the end of any caption that does not naturally terminate in punctuation, as Captions, Inc. does.

Hyphens

Do not soft-hyphenate words. Do not attempt to add hyphens to words. The only permissible hyphens are hard hyphens, that is, hyphens that must be used or the word would be misspelled.

Why? This too may come as a surprise to captioners, some of whom think a caption with nine words is ten full percentage points better than a caption with ten words, but it's been established for decades that adult readers do not read word-by-word. Instead, the eye bounces in *saccades* across several words at a time, landing in *fixations* that vary according to familiarity with the text, reading conditions, type size and linespacing, and other factors. Generally, only unfamiliar or hard-to-read words are deciphered letter-by-letter, the way we

learned to read in grade school.

With a 32-character line and monospaced fonts (in capitals most of the time, no less), soft-hyphenated words impede saccadic motion of the eyes. (How easy was it to read the preceding paragraph?) Hyphenation in print typography takes years to get right and still is a matter subject to some dispute. *Don't do it at all in captioning*, save for the suppassingly rare words that are too long to fit in 32 characters, like *supercalifragilisticexpialidocious*.

Here are some actual hyphenation atrocities perpetrated by a leading Canadian distributor:

1. ARTI-
CLE
2. UNDERSTAN-
DABLE
3. Emmett:
I CAN'T BE-
LIEVE IT'S
SATURDAY NIGHT
4. LYNNETE INSISTED MEL
AND I NOT DRAW AT-
TENTION TO OURSELVES.
5. THEY WERE WAIT-
ING FOR US
WHEN WE GOT HOME.
6. SUDDENLY EVERY-
BODY WANTS TO
INFANTILIZE ME.
7. I'D COUNT ON GET-
TING HITCHED IN
MISSISSIPPI FIRST.
8. WAS MORE INCEN-
TIVE FOR HIM
TO GET BETTER
9. DISCRIM-
INATION
10. WHAT-
EVER
11. YOU ARE SUR-
ROUNDED BY PEOPLE
WHO LOVE YOU.
12. THE JUNIOR PATHO-

LOGIST AT ASHFORD
HOSPITAL.

13. I GUESS IT PAYS
TO BE THE TEACH-
ER'S BOYFRIEND.
14. YOU'RE PLAYING
SOMETHING RIDICU-
LOUSLY ROMANTIC.

To repeat: No soft hyphens.

However, contrary to the Manual's advice, you can break a line after a hard hyphen. Why not? After hyphens is one place where the English language *can* break lines.

Also, the Manual's sections on "Dashes" [p. 21] are quite wrong. First, "parenthetical information" as found on TV is usually an *apositive*: "A construction in which a noun or noun phrase is placed with another as an explanatory equivalent, both having the same syntactic relation to the other elements in the sentence; for example, *Copley* and *the painter* in 'The painter Copley was born in Boston.'"

Second, the construct listed in the Manual –

Please take him
--the guy in red--

is, to use the scientific terminology, deeply weird and is unprecedented in English orthography.

Please take him--
the guy in red--

is the standard orthography.

Aspect in non-speech information

First, it seems to have been decided that NSI will be enclosed in parentheses without initial capitals unless proper nouns require them. A case can be made for that decision but has not been.

Even after 20 years, Canadian captioners still have not figured out the concept of *aspect* in writing non-speech information. It's a *grammatical term*: "A category of the verb designating primarily the relation of the action to the passage of time, especially in reference to completion, duration, or repetition."

(phone ringing) is in the *progressive aspect* (possibly the continuous aspect – it's a fine line). (phone rings) is in the *indicative aspect*.

The "Descriptive caption examples" section elides this issue entirely.

⋮

[p. 24] ¶ (engine revving)

(whispering)

(phone ringing)

(loud knocking)

(pager beeping)

(rapid gunfire)

(gunshots)

Isn't there a difference among (phone rings), (phone rings twice), (phone rings then stops), and (phone ringing)?

How about (belches) vs. (belching)? (clears throat) vs. (clearing throat)? (How long can *you* belch or clear your throat continuously?)

Moreover, the Manual fails to explain how to handle continued or interrupted NSI. An example is captioning the word (ringing) when we can see the phone and (phone continues ringing) later on.

This section requires much more elaboration. Along with speaker identification and absence of translation, non-speech information is the key criterion differentiating captions from subtitles and merits much wider treatment.

In particular, the Manual needs to outlaw the use of strictly limited and preordained structures in writing NSI. One Canadian distributor captions everything in an [Xing of Y] format (complete with errant spaces inside brackets):

- [Stopping of music]
- [Squeezing of onion]
- [Yell of joy]

Have you ever met anyone who talked like that?

Unclear speech

There is as yet no explanation why Canadian captioners constantly tell us (lyrics unclear) or (indiscernible). Trust me, lyrics are much less likely to be unclear than you think, and few, if any, conversations are indiscernible. *Distant* music or conversation is a different story. I'm talking about music or dialogue that I can not only understand but repeat out loud

right then and there. In fact, that's what I usually do, talking back at captions that lie to viewers like me by claiming the words could not be understood. *I* could understand them.

In fact, an ear for dialogue is a prerequisite for a successful captioner. It comes from deep-seated fluency in the English language and a wide vocabulary, aided by excellent audio quality. I've seen a lot of indistinct conversation captioned as such by Americans (because the director and sound engineer designed it that way; *nobody* could decipher the dialogue, and that is intentional), but I have seen one (count it: *one*) case of *unplanned* indecipherable foreground dialogue – a drill instructor screaming continuously at a recruit in a U.S. documentary. (`\indecipherable\`) was the caption, and it was true.

(The Caption Center would probably argue that disembodied overlapping voices in *September Songs: The Music of Kurt Weill* was another case, but such voices were meant to be indecipherable and fragmentary. I believe another counterexample could be given: An interview segment in which a helper pops into the room and mumbles something out of the coverage area of the microphone.)

I reiterate one of my shibboleths: If Canadian captioners weren't nearly all young women in their 20s with little life experience and insufficient experience reading, writing, and transcribing, we wouldn't have this problem. (One more time: Canadian captioners are a monoculture. Everyone has the same deficiencies and everyone blanks on the same issues. There is no safety net because there isn't a wide enough range of experience and knowledge.) And if captioners were perhaps paid better, they'd care more.

The explanation "Do not guess at indiscernible speech" [p. 25] is something of a motherhood issue. Who would *want* to guess? The real problem is an overdiagnosis of speech as indiscernible, a problem that traces itself to poor skills in the Canadian captioning monoculture.

Use of electronic databases

On the topic of transcription accuracy, I would note there is a tremendous reflex to Google everything. The assumption is that there is a page somewhere on the Web that can authoritatively answer any query on the correct way to render a word or phrase.

Well, that's not true. I'm a stickler for accuracy and I have myself left incorrect spellings of proper nouns online for months at a time. (Sometimes the original source can be incorrect. One article actually got its own byline wrong and I simply copied it. How could I have known? Fortunately, I later found the correction.)

Even printed sources are not always reliable. I read a book on current French cinema that consistently misspelled the surname of one of the directors the book itself interviewed. Further, foreign names can be transliterated in several ways, a phenomenon found in some languages more often than others – Russian and Ukrainian more than, say, Japanese.

While no specific reference *medium* is bulletproof, I would argue that a reliance on the Web is a recipe for trouble. Many captioners are unaware that, predating the Web, fee-charging electronic databases proliferated. They still rake in good cash even today; that's why, for example, Thomson has been selling off newspapers in favour of electronic databases. (Disclosure: Since I did not opt out of the court-defined class, I am party to a class-action lawsuit against Thomson Corp. over unauthorized reuse of copyrighted newspaper articles in fee-charging databases.)

The most famous American database service is [Lexis-Nexis](#), carrying an astonishing wealth of information. But specific industries and periodicals are also online. In the case of the fashion industry, for example, periodicals like *Women's Wear Daily* are online in full text; you just have to pay for access. Articles in trade papers like these are quite likely to spell the names of their subjects correctly.

If you're on a deadline and need to know the spellings of eight designers who showed in Paris four days previously, a subscription to one of these databases is well worth the money. Ten dollars in search time answers your question right then and there.

A case could be made that Canadian captioners should band together to negotiate group rates for database access. In any event, the costs are not particularly onerous.

Music

The manual flagrantly authorizes the full range of Canadian misdeeds in music captioning. Note on notation: The staffnote character is difficult but not impossible to typeset. The vertical-bar character / is used as a substitute by everyone but this Manual, which uses an underscore, a bad idea for several reasons: We *can* actually underline in captioning (do you mean you want that space character underlined?), and underscores bleed into each other, making it hard to count how many instances you intended. (Similarly, backslash \ is used to indicate the italic toggle.)

The Manual must explicitly outlaw the single most egregious and appalling habit of Canadian captioning, namely a caption like [| | |] slapped on the screen, up to a dozen times per program, as some indifferent, casual, nonspecific indication of "music."

1. Don't write spaces inside brackets. It's forbidden in English orthography. Brackets in caption fonts (yes, even nice fonts with proper descenders) look like the word *I*.
2. Captioners must not attempt to collapse complex auditory phenomena onto punctuation. *Write it out!* We are working with words here.
3. Captions must be pronounceable. Please pronounce [| | |] for me.
4. A typical use of this [| | |] abomination is as a placeholder when there's a tiny break between dialogue captions. Worse, certain real-time captioners will actually blank the screen to display the barbarism as a pop-on caption, destroying up to three lines of real-time captions to do so.
5. I've seen this barbarism used *in music videos*. Aren't music videos all about music? Wouldn't you expect music to be found in a music video? Isn't that why we're watching them? Couldn't you tell us a wee bit more about it than [| | |], whatever that means?
6. Another typical use? Opening theme songs. So those are entirely unimportant, then? (Won't Danny Elfman and They Might Be Giants be a bit surprised to hear that?)

Music is *important*. It is not unidentifiable, interchangeable, or inconsequential. It must be treated with much greater respect than Canadian captioners presently do. The [| | |] must never, *ever* be used again for any reason. Whenever you the captioner are tempted to slap it onto the screen, ask yourself:

1. Why is there music here?
2. What is the music like?
3. If I'm getting paid to turn sounds into words, why aren't I writing a little phrase to *explain* the music?

Wouldn't you say the following usage is a tad more sophisticated, literate, and understandable than robotically copping out from your responsibilities by slapping [| | |] onto the screen whenever anyone strikes a musical note?

1. (theme song plays)
2. (techno music in background)
3. (guitar solo)
4. (theme song: haunting woodwinds) – and actually here we see why it works better to use an initial capital for NSI so we can write a comprehensible phrase, viz. (Theme song: Haunting woodwinds)
5. (band starts playing)
6. (bass plays) followed by (drums join in) and then (Greg vocalizing)

In a similar vein, the Manual's authorization of the practice of setting staffnotes into the corner of the screen [p. 26], à la Captions, Inc., again collapses auditory phenomena onto punctuation. Better examples:

1. (music continues)
2. (music)
3. (background music)
4. (bagpipes play softly)

Use your heads. *Write it out!*

See also: [Punctuating music](#).

"No audio"

Please exercise caution in telling captioning viewers a segment has "no audio" [p. 26] (Cf. [eye-gaze](#)).

"No audio" means total silence. Now, what were the last five occasions in which you encountered dead silence on a videotape?

If mouths are moving but no voice is coming out, then that's what's happening: (no voice). (mouthing words) also sometimes works – though perhaps the caption (mouthing) is too unclear – when that is actually what's happening. But (mouthing

words) it's a way longer caption than (no voice), and such moments are usually so brief that a shorter caption is better.

If you really do find a segment with (no audio), by all means use it.

Inflections and accents

First of all, “inflection” is perhaps not an apt term here. The Manual states [p. 27] that “[m]any people speak with inflections or accents, use liaisons between words or leave endings off words, etc.”

Everyone speaks with an accent; accents are relative. Words are not articulated separately; speech is a continuous flow of sound. (Untrained people tend to violently disagree with that statement; it's like learning there is no such thing as Santa Claus. It remains true nonetheless: *There are no breaks between words in continuous speech.*)

I assume “leave endings off words” refers to the case of, say, *looking* vs. *lookin'*. A common misconception; nothing is being “left off.” Perhaps you refer to French speakers in English who leave off plural markers and verb tenses (“I walk back from the store after I pick up some tomato”).

No captioning style guide I've read has coherently addressed this topic. The result has been full-on miscaptioning of programming by *every* captioner I can name. Yes, *everybody* gets this wrong. One prominent case: It was never mentioned on *Star Trek: The Next Generation* that Jean-Luc Picard, while allegedly French, speaks with a British accent. I think it's fair to say that hearing viewers took note. In a more diegetic example, Carrie on *Sex and the City* makes a new friend who is, in fact, Australian. It comes as a big surprise in captions when he starts mentioning Sydney since we hadn't been told of his Australian accent.

This section requires expansion.

Moderate errors

Extensive experience

[p. 5] ¶ These standards are only now possible as a direct result of the extensive experience gained in English-language broadcasting in Canada.

It's been possible to write standards since the early 1980s when captioning began. The “extensive experience” noted here is the cause of the problem. The Manual attempts to explain right and wrong ways to caption; in many respects it seeks to deprogram Canadian captioners of their worst existing habits. We are trying to establish new standards, which will involve unseating existing standards, if they can be called that.

Degree of hearing loss

[p. 6] ¶ Deafened and deaf (lower case 'd') are terms that refer to individuals who have lost all hearing at some point in their lives.

[p. 49] ¶ *deaf/deafened*: Terms that refer to individuals who have lost all hearing at some point in their lives. These people use spoken language and rely on visual forms of communication such as speechreading, text, and occasionally sign language.

Surely it is inaccurate to say these groups have lost “all” hearing. “All usable” hearing may be more accurate, or simply “nearly all” hearing.

The term non-speech information

[p. 8] ¶ Trained off-line caption editors watch and listen to a videotaped program and create a transcript of the audio, including descriptions of sound and music.

Captioners render sound in text. Speech is one stream of the sound so rendered. It is tautological to state that captioners describe sound.

Music is not the only stream of non-speech sound that is rendered. The term is a bit strained and academic, but *non-speech information*, used in captioning research, better encapsulates the full range of sounds other than words that are rendered in text. NSI is the acronym.

The Manual means well in one section but makes a mistake:

[p. 23] ¶ The art of off-line captioning involves making creative and informed choices about what to include in a caption script, and descriptive captions should never be included at the expense of dialogue. Negotiating space and time limitations while simultaneously crafting the most accurate representation of the story possible is a constant challenge, and while descriptive captions can do a great deal to enhance a viewer’s understanding of a program, there are situations where their use is more appropriate than others.

Actually, if a cellphone starts ringing while Carrie and Samantha are chatting over brunch, and keeps ringing, and then finally, after 30 seconds, Carrie turns around and hollers “Are you going to answer that phone?” then you may well have to interrupt the dialogue while it is unfolding to indicate that the phone is ringing continuously.

Anyway, the rest of the paragraph is disingenuous given that pop-on captions can appear in multiple blocks. You can caption the dialogue and also the NSI. One has quite enough room in roll-up captions to do the same, presuming they are prewritten; you can easily fit in (phone ringing at other table) or equivalent.

Timecode assignment

[p. 8] ¶ The caption editor assigns a time code address to each caption as well as a position code.

“Timecode” is one word. Captioners demonstrably do not assign timecodes to each caption. Some captioners may be *forced* to do so by outdated software, but it is not necessary; proper software handles that drudgery for you. In a long monologue, for example, all the captioner needs to do is provide an in time for the first caption and an out time for the last and the software fills in the blanks.

Timecode is not even really manifest in the ultimate captioned submaster (not “sub-master”). Captions just appear and disappear at certain moments. Timecode really only exists up to the encoding stage. I’m not sure it’s that important.

Suitability of pop-on captions

[p. 9] ¶ Pop-on captions are most commonly used for dramas and sitcoms, movies, documentaries, and music videos.

It should be stated that pop-on captions are *suitable* for *any* programming. After all these years, the only programs for which I can see an exemption are prerecorded shows with continuous breakneck speechrates (*Iron Chef* could be a worst-case example) where buildup time would require so much editing that few captions would appear verbatim and it would require an entire workweek to caption the show (as mentioned later on [p. 9] of the Manual).

It is important to outlaw the use of roll-up captions for fictional narrative, arts, and music programming and require the use of pop-on captions for same unless there is an incontrovertible reason to use the former, which comes up rather less often than Canadian broadcasters would like us to believe.

[p. 9] ¶ In the off-line situation, roll-up captions are normally reserved for programs that have a live flavour, such as entertainment, sports and news magazines, awards programs, and soap operas.

You’re deliberately mixing up current practice with preferred practice. What does “live flavour” mean? Not many shows air live these days. *Entertainment Tonight* and its ilk certainly do not. Soap operas are almost always pre-recorded, with extremely unusual live segments, and in any event are fictional narrative programs that should not be captioned using roll-up unless there is no alternative. (I distinctly remember when soap operas started to be captioned. Pop-on was used, as it should be. Later, a combination of too many soap operas having to be captioned all at once, apparently tighter turnaround times, and apparently stingier production budgets forced the use of roll-up captions, including the entirely improper use of real-time captioning.) By the Manual’s own logic, use of roll-up for soap operas should be “strongly discouraged.”

Live-display encoding

[p. 10] ¶ This method can also be used when the program itself is pre-recorded and the text is created ahead of time, but there is not enough time for encoding in off-line mode.

I am not convinced that encoding time (the actual length of the program plus time to set up tapes) is the constraint.

Teleprompter captioning

First, by the Manual's own logic, you should be writing the word *Teleprompter* as TelePrompTer, the perverse corporate orthography. It is in any event not a generic word yet (even Xerox isn't xerox yet), so it needs a big T.

Meanwhile:

[p. 11] ¶ This type of captioning is mostly applied to news shows and soap operas that are 100% scripted. It is only appropriate when a script has been prepared and is available for an entire broadcast, and when there is no ad-libbing or improvising. Any other application is strongly discouraged.

You can't caption a news show with Teleprompter or ENR captioning and have it be legit under [CRTC regulations](#). All sizes of television station are required "to caption... all local news programming, including live segments, using either real-time captioning or another technology capable of producing high-quality captioning for live programming." Teleprompter or ENR captioning simply cannot meet that standard unless speakers never deviate from the script, nothing is ever ad-libbed, and speech is the only significant audio source. When are all those true at once?

Good real-time captioners (Canada does not have many real-time captioners who are *not* good, at least in the English language) can achieve over 90% accuracy when averaged over reasonable periods, like a one-hour show or a week of them. They also caption non-speech information. I challenge readers to point out a newscast where Teleprompter or ENR captioning can do the same. Its use for news should be outlawed in the Manual.

Preparation of real-time captioners

[p. 11] ¶ It is the caption stenographer's responsibility to prepare his or her dictionary, entering names and vocabulary that he or she can anticipate encountering during the captioning of various programs.

It is the broadcaster's responsibility to provide any available material such as guest lists, rundowns, key lists and so forth, which are necessary for the caption stenographer to prepare for a program.

Vocabulary that should be but generally is not included in steno dictionaries:

1. Street names in the city whose news you are captioning (sometimes difficult when captioning a far-off regional newscast, but you need to know the main streets in important towns)
2. Names of every country in the world, no matter how obscure, and adjectival forms, since U.K. newscasters tend to use them (“the Bangladeshi prime minister”)
3. Names of all Canadian prime ministers and U.S. presidents and all reasonably famous Canadian premiers, U.K. prime ministers, and foreign heads of state; names of municipal politicians in the regions whose news you caption
4. Typical foreign phrases, including French phrases not routinely translated by interpreters (*merci beaucoup* and the like – the list needs to be studied, actually)
5. Famous athletes of today and yesteryear. In practice, this means the entire league in whatever sport you specialize in, plus names that are apt to come up in colour commentary (Lafleur, Uinitas, Cassius Clay)
6. Scientific and technical terms used even occasionally, irrespective of length (“chlorofluorocarbon” is a perennial bugbear), including names of various disciplines and their practitioners (OBGYN, otolaryngologist, luger, philatelist)
7. Titles of popular films of today and yesteryear, including A- and B-list stars (Montgomery Clift, *Princess Mononoke*, Djimon Hounsou, *A Clockwork Orange*); very important to get Canadian titles and actor names right
8. All song titles and artists on the current *Billboard* Top 40, and important musical artists and composers of the past and present and their well-known works (Erik Satie, Trent Reznor, “Thus Spake Zarathustra,” “The Real Slim Shady”)
9. Musical styles and adjectives (gangsta, Beatlesque, ska)
10. Cultural terms, including food, decorating, and music terminology (batik, mesclun, sauteed, lieder)

That’s a better list than before, but it isn’t sufficient quite yet.

Emergency captions

[p. 13] ¶ In emergencies, call a stenocaptioner. Failing that, your station must have a plan with a chain of succession so that there is always someone whose job it is to type information in real time into an ENR or Teleprompter captioning system for display.

Sports captioning

[p. 16] ¶ For fast-moving sports, play-by-play commentary is not captioned. Instead the captions are blanked, giving time to see the play, then captions continue after the play is complete.

That’s not true. That is merely one method of captioning continuous play-by-play. There is no testing of viewer preferences on this issue that I know of, but I don’t consider the practice harmful. Nor is the practice of captioning continuously harmful (in fact, the absence of continuous captions was part of the [Vlug vs. CBC](#) human-rights complaint, so writers of the Manual might proceed with caution). I challenge readers to prove one method is really vastly better than the other, though if our goal is to caption

verbatim, I don't see how we can support the idea of deleting entire categories of speech from captioning.

Remember, we are not captioning for robots. People can ignore the captions if they want! If the sports action is more interesting, they'll watch the sports action.

During the Stanley Cup and the Winter Olympics, you can watch both methods yourself by comparing Canadian and U.S. hockey coverage airing at the same time, or Olympic figure skating.

Paint-on captions

There is no discussion of the use of paint-on captions as first or even also second captions when buildup time is tight. The thinking is that the letter-by-letter display makes it somewhat easier to read the whole caption in time.

Adding paint-on captions when a pop-up caption sits stationary onscreen (as in call and response of a music video or musical) works adequately well in my experience, though it is rarely seen.

The example under "Building captions" [p. 36] is too hypothetical and seems unwise anyway.

Underlining

There's a reason why we don't underline in English [p. 18], and it's called "italics." Captions should follow English orthography.

Moreover, to toggle underlining on and off in Line 21 (save for colour-plus-underline combinations) requires an inserted blank space. That means end punctuation will end up underlined:

... was added to the cast of Survivor.

Numbers

The discussion of numbers is inadequate.

- An extremely important captioning convention is consistently flubbed by Canadians: Utterances like "five to ten thousand" must be written as 5,000 to 10,000 and not five to 10,000 or 5 to 10,000 (because the lowest number in the sequence is 5,000, not five!). Hence \$3 to \$4, in the range of 15¢ to 25¢ a share.
- Hence "between a hundred thousand and three-quarters of a million dollars" is rendered as between \$100,000 and \$750,000.
- If a number below ten and other numbers above ten are found in the same utterance, use figures:

Hot dogs come in packs of 6, 12, and 24.

- Very large numbers can look ridiculous when written out in words at the beginning of captions and when they constitute the entire caption. Use figures.
- Approximate values should be written out in words:
 - thousands of dollars (not 1000s, 1,000s, or, God help us, 1000's)
 - a few hundred
 - a hundred good reasons
 - looks like a million bucks
 - at least it's not three hundred degrees out like yesterday
- Long phrases are a difficult case:
I wanted to lose 25 pound or so.
- Whole dollar amounts do not take a decimal and zeroes:
\$5, not \$5.00. (Check *Chicago*.)
- Commonly-uttered forms for dollar values ("two ninety-nine") should be written with dollar signs (It was only \$14.99 at Wal-Mart) unless the context sets up a difference, though I can't think of an example of the latter right now.
- The cents ¢ character exists and can be used: 25¢. If captioning a commercial and they use a different reasonable orthography, like \$.25, copy it. (Leading zeroes are nicer but not always necessary.)
- £ has always been part of the Line 21 character set, and Sterling amounts *must not be treated differently from dollar amounts*, as usually happens. Correct usage: £25, £43.25 (uttered often as "forty-three pounds twenty-five," a consistently-mishandled example).
- There is no euro or yen sign or any other currency symbol in Line 21 fonts. You must use phrases, but with decimal numerals: "two hundred and fifty-seven and twenty-three euros" would be written in print as €57.23 and in captions as 257.23 euros.
- Ordinals require care. Probably ordinals above "tenth" should be written as alphanumerics (125th, 2,000th). *nth* should be written thus (the nth degree).
- Fractions are mishandled. The $\frac{1}{2}$ fraction exists in the 608 character set ($\frac{3}{4}$ was eliminated) and should be used, but only in combination with a whole number or mantissa: 3 $\frac{1}{2}$ to 4 but not $\frac{1}{4}$ to 1. For other fractions, a space must be typed between whole number and fraction: 3 3/4, not 3-3/4 or anything else. Don't break a line at that space. (Note that $\frac{1}{4}$ and $\frac{3}{4}$ characters were replaced in EIA 608 decoders. See Gary Robson's [discussion](#).)
- Measurements expressed in words must in some cases be written in numbers: "A teaspoon and a

half" becomes 1½ teaspoons.

- In scientific discourse, there can be such a thing as a *eight and a halfth* occurrence. Best use alphanumerics:

8½th,

3 ¾th.

- Times can be stated in words:

half past 3:00,

a quarter to 7:00,

20 after 11:00,

half eight.

Punctuation

The overriding rule, as everywhere in captioning, is to follow English orthographic rules unless an incontrovertible case can be made otherwise. In punctuation, I can think of only a few such cases:

1. Quotations spanning more than one caption, covered in their own [section](#).
2. Speaker designation in roll-up captions: Use of >> (greater than–greater than–space) is unique to captioning.
3. En dash: Because we're dealing with monospaced fonts where every character is precious, the least illegible method of simulating an appositive en dash is nospace-dash-dash-space, as follows:

- Glawson's physical therapists concentrated his training

- on the quadriceps-- thighs-- and gastrocnemius-- calves--

- to compensate for wasting that occurred in hospital.

- In Inuvik-- formerly Frobisher Bay-- teen suicides have decreased

- a mere 5% since new measures were introduced in 1999.

(Liberties were taken with linebreaks. Spaces after the dashes can be elided by a linebreak, which is actually an advantage.)

To sum up the alternatives:

- Space-dash-space is too small (and is the wrong substitute for what is actually used in

print, space–en dash–space).

- Glawson's physical therapists concentrated his training on the quadriceps - thighs - and gastrocnemius - calves - to compensate for wasting that occurred in hospital.

- Space-dash-dash-space is too large (though it is the approximately correct substitute for nospace–em dash–nospace as used in print).

- Glawson's physical therapists concentrated his training on the quadriceps -- thighs -- and gastrocnemius -- calves-- to compensate for wasting that occurred in hospital.

- Nospace-dash-dash-space seems to work best as an appositive marker; it shows a break in sense without prompting viewers to read the dash as a word unto itself; it also prevents an en dash from being stranded at the beginning of a line.

4. Percentages: In print typography, only values above ten percent or non-integer values below that threshold tend to be written alphanumerically: 2.5%, three percent, ten percent, 12%. But in captioning, we tend to write everything alphanumerically.

5. Inverted exclamation point, as used in Spanish: It was never part of the Line 21 character set, one of many appalling inadequacies. In the days of the TeleCaption and TeleCaption II, captioners could be sure that a lower-case *i* would have no serifs and would pass for a *j* character, but that is no longer the case. We're stuck leaving it out (a spelling error) or using **!** (a childish error). Inverted question mark has always been part of the spec, apparently because engineers remembered their high-school Spanish (but not very well, since they forgot *j* and capital accented letters).

I believe these are the exceptional cases. *Everywhere* else, follow the rules. You are typesetting English in the captioning medium; you are not writing some new language with its own rules.

Follow the *real* rules of English as authoritatively documented in style guides like *Chicago* rather than any half-baked misconception of what the rules are. I have met Canadian captioners who believe that commas are required after every single adjective (e.g., “long, brown hair”), and everyone seems to think that spaces go inside brackets (they don't). Some captioners seem to think we use American double quotation marks with British logic, which in fact we do not.

Similarly, it may come as a shock to captioners to learn that *or*, *so*, *and*, *but*, and *then* used at the beginning of a sentence [p. 20] rarely require a comma right after (usually only when an actual pause is intended). However, a comma *cannot be omitted* before the name of an addressee. (It's an iron-clad rule.

Don't believe me? Compare the correct "Come on, Eileen" versus the rather incorrect "Come on Eileen.")

I return here to the youth and inexperience of Canadian captioners, nearly all of whom are a monoculture of young women with liberal-arts degrees and not enough life experience as readers and writers. Many, possibly most, captioners are trainable, but such training requires learning the established rules and following them.

Accordingly, the following advice is wrong:

[p. 20] ¶ It is essential that caption editors and caption stenographers make use of a variety of Canadian dictionaries and style guides to reach sound decisions about punctuation. Document your decisions and use them consistently.

You need a range of reference materials, but you must *not* pick and choose among style guides, which amounts to shopping for a justification for your own errors or biases. Canadian dictionaries are necessary (Gage and Oxford are the most recent), but U.S. style guides like *The Chicago Manual of Style* are authoritative and indispensable. I am not convinced Canadian offline captioners are so good at rendering the *spoken* word in the *written* that they know better than *Chicago*.

Spelling out words; enumerating numbers

Words spelled out should be hyphenated, as the Manual suggests [p. 21], but the "double-letter" construct deserves attention.

- M-I-S-S-I-S-S-I-P-P-I if every letter is uttered
- M-I-SS-I-SS-I-PP-I if doubled letters are uttered as "double-S," "double-P" (or, more literally, M-I-double-S-I-double-S-I-double-P-I)

But do *not* use hyphens in numbers. To

dial 911 is to

dial 911 and not

9-1-1.

Bearing 241 mark 73 is

bearing 241 mark 73 and not

bearing 2-4-1 mark 7-3. There are exceptions, like radio code:

10-4,

10-20. *NYPD Blue*-style station designations are hyphenated: I called McGuire

down at the 2-7. (You could use

Two-Seven, though it's a bit poncy.)

Extended quotations

The manual's coverage is not quite right. Here, "extended quotation" means a quotation spanning more than one caption.

Extended quotations in captioning – and note this excludes roll-up captions by definition – use a quotation mark at the beginning of every caption save for the last, which has only an ending quotation mark. (Periods and commas go *inside* quotation marks without exception; this is Canadian English, not British. Question marks and bangs rarely are set inside quotation marks unless part of the quotation: The question they ask is "What would Jesus do?" is correct, while What do you mean by "extra-virgin?" is not.)

Nested quotations follow the very same rules and continue alternating between single quotation marks (the apostrophe, used for the first level of nesting) and double. It's entirely possible to have two levels of ongoing extended quotations (on an interview with Maya Angleou, I witnessed three); it's also possible to have discrete quotations nested inside extended quotations. Don't use italics for nested quotations (an NCI error).

For comparison purposes, note that the rule in print is: Every paragraph gets an opening quote; only the last paragraph *also* gets a closing quote. We must alter this rule in pop-on captioning because pop-on captions are discrete units. No quotation marks? It isn't a quotation. Quotation marks at both ends? It's a *complete*, self-contained quotation. I don't know of any Canadian captioner who gets this right.

Note that roll-up captions are virtually never discrete captions the way pop-ons are. Hence the following usage, which I saw in the week before writing these remarks, is wrong in English orthography (unprecedented, in fact) and wrong in captioning:

The self is "not a very
"skilled and intelligent self,
"but allows animals to develop
"into the intentional, volitional,
"and cognitively selective creatures
they are."

In pop-on captions (setting aside justification and even line lengths for the moment), a correct usage is as follows:

1. I sat in his office
and he picked up this big sheaf of papers.
He goes,
"Let me read you what it says
in\Selleck vs. Globe."\
And he went on this big spiel.
I remember it like yesterday.

"The article complained of was called
"Tom Selleck's Love Secrets by His
Father,"

"and contained a number of statements
attributed to Tom Selleck's father

"which disparaged and downgraded

"the romantic character and capability
of his son.'

Well, I guess that's in the eye of the
disparager."

And that's the guy's
\defense\lawyer!

2. The self is "not a very skilled
and intelligent self,

"but allows animals to develop

"into the intentional, volitional,
and cognitively selective creatures they
are.

"Descarte's faith in the assertion
'I think, therefore I am'

"may be superseded by a more primitive
affirmation

"that is part of the genetic makeup
of all mammals:

'I feel, therefore I am.'"

3. He told me, "She helped with the financing.

"Armistead asked me,
'Who's going to read the narration?'

"And they said
'Lily Tomlin.'"

Then he groaned and tells me,

"The guy perks right up and says

"'No, it's fine.

"'She's going to use
this opportunity

"to come out and be honest.'

"So Armistead wrote
her coming-out speech after all.

And she cut it!"

I mean, you can only laugh
at something like that.

Editing sacred texts

The section on quotation marks avoids an important issue.

[p. 21] ¶ Where quoting a passage from a book, play or poem etc., it should be captioned verbatim.

If necessary, you can remove an entire sentence or, under battle conditions, an entire phrase. The same applies to sacred texts like Shakespeare (I just recently saw my first-ever edited Shakespeare captions: 25-year-old women earning \$13 an hour know Shakespeare well enough to edit him?) or any film with cult status, like *The Rocky Horror Picture Show*.

Other examples include historical speeches, song lyrics, and anything else that a viewer could actually look up.

Spaces

[p. 22] ¶ Do not leave more than one space in a caption. A single space may be used after a period, colon, or semicolon as necessary.

2. It is common to leave a single space before and after music notes and parentheses.
3. There should not be a space between parentheses and the text enclosed within them.

The case could be made that, since we're stuck with monospaced fonts, the orthography used in typewriting is applicable, namely the use of two spaces between sentences. Naturally, we use only one

space after other punctuation, like colons or semicolons; “as necessary” is meaningless and authorizes untrained captioners to write no space or a raft of spaces after such punctuation.

Sections 2 and 3 are correct but too vaguely written.

Punctuating music

The Manual is just a wee bit off here:

[p. 26] ¶ Punctuation in songs should be limited to occasional commas where phrasing requires it. Otherwise, there is no punctuation.

No. Use all necessary punctuation within song captions. (Why not? It’s English, right? But see below.) That includes quotation marks and extended quotations, which, believe, me, do come up:

```
| He'd say  
| "I know what I want |  
| "and I want it now |  
| "I want you |  
| 'cause I'm Mr. Vain" |
```

Commas, en dashes, colons and semicolons, you name it – if it’s necessary, use it.

What gets elided are caption-ending commas and periods. Some captioners (actually, only CaptionMax, to my knowledge) attempt to caption songs as though they were complete sentences, with full end punctuation. Sometimes they actually are full sentences, as in opera or musicals. Generally, though, song lyrics are fragmentary. It is not hard to find examples in print orthography of elided end commas and periods.

Question marks at ends of song captions must always be used. Exclamation points should be used with discretion; there’s so much EXCITEMENT! in music that one is tempted to end every caption with a bang.

Doggerel

The Manual is silent about utterances delivered somewhat in time to backing music without being entirely sung. Poetry set to music, you might call it. I can cite a reprehensible example that nonetheless fits: “Rico Suave” by Gerardo. Also nearly anything by King Missile or Meryn Cadell; the music video “Invalid Letter Dept.” by At the Drive-In is the best example yet seen.

Some kind of explanatory caption is desirable (I don’t have a ready-made example for you), then proceed to caption the doggerel as though it were music *but without staffnotes*.

Scene and shot changes

One very old Canadian captioning centre (derived from an old development agency) goes to eye-crossing lengths in slapping a new caption onscreen after nearly every shot change. Yes, they'll break a ten-word sentence into three captions. Yes, they'll break a caption after the word *the*; they'll break a caption anywhere at all. They're still doing it nearly 20 years on, demonstrating thoroughgoing misunderstanding of captioning and outright contempt for audiences, who will find reading 300 fragmentary captions harder than 100 intact captions.

I mention this because the Manual does not emphasize the important distinction between scene and shot changes. There are few cases where a caption can persist across a scene change; there are quite a few cases where a caption can persist across a shot change. (Indeed, in commercials, music videos, and twitchily-edited TV shows, you have no choice.)

I would also point out that experience shows the cleanest breaks around any kind of visible edit leave the last frame before the edit and the first frame after the edit blank. (Remember, we're using a two-frame blink rate, not zero or four.) Captions then look like an immutable feature of the program rather than something raggedly tagged on afterward, as Canadian captions tend to do.

The full topic here is timing of captions, not shot and scene changes. It's too important to be discussed in a single page, in part because nearly everyone in Canada gets caption timing so terribly wrong. The section in the Manual requires expansion.

Multiple captions

Setting multiple simultaneous caption blocks at different locations onscreen isn't the only way of denoting multiple (near-)simultaneous speakers. The Captions, Inc. approach (derived from subtitling and only subtly mishandled) also works. In fact, it probably works better than discrete blocks.

```
- Are you both there?  
- GEORGE: Yes.  
- MELINDA (whispering): Yes.
```

Beats the hell out of what everyone here is using, doesn't it?

Speaker IDs, non-speech information, and combinations

Given that we are now captioning in mixed case with speaker IDs in upper case, combinations of speaker ID and NSI use mixed case:

- JENNY (shouting):
- JENNY on answering machine:
- JENNY (clearing throat):
- JENNY and LINDA (singsong):
- JENNY, sarcastically:

There is too much variety in delivery styles and the sorts of information that must be notated to rely on a single style. Just keep everything on its own line whenever possible, though it sometimes is not, as in:

PRIME MINISTER CHRETIEN
(translated):

Hence, the following actual Canadian malapropisms are to be avoided: Too much punctuation, all of it nonstandard, the whole thing too difficult to read.

Shakira: [Singing]
["Objection"]

O-Town Group: [Singing]
["We Fit Together"]

Margaret: [On radio]
[Stammering]

Also, do not use slashes when more than one effect happens at once.

Police officers:
[Grunting/
chuckling]

Do you mean "and" or "or"? Are they grunting or are they chuckling?

Interpreters and translation

Another topic never handled correctly by Canadians. *Translators* work in the written medium; *interpreters* work in signed and spoken languages. Captioners could never caption a "translator."

Don't use an ID like **Voice of interpreter:** because what we usually caption are voices.

Here are some useful correct examples:

1. (Jorg speaking German)
INTERPRETER:
2. INTERPRETER:
3. FEMALE INTERPRETER: (useful when the speaker is male or vice-versa)
4. (asks question in Hutttese)
(replies in Hutttese)
(Hutttese continues)

Music in roll-up captions

The Manual ambiguously states:

[p. 38] ¶ If there are long music segments within a roll-up captioned program, music lyrics must be in pop-on style.

Lyrics can roll-up only when they are unusually fast, or when there are just a few lines of singing interspersed with conversation. Even if rolling up, lyrics must follow pop-on rules of division, and music notes must appear at the beginning and end of each lyric.

All sound and music descriptions should roll up, but follow pop-on rules of division.

The first edict is unnecessary and probably unworkable for nightly talk shows. More important is staffnote placement. I don't see any reason not to place beginning and ending staffnotes at exactly the same places you'd insert them were you using pop-on captions. The other option (inconsistently applied by the Caption Center) seems to be to use an opening staffnote at the very beginning and a closing one at the very end.

Note that one still requires >> and speaker IDs in music captioning. Hence the following would be correct:

```
(band plays "The Good in Everyone" )
>> PATRICK PENTLAND: | First off,
here's what you do to me |
| You get rough,
attack my self esteem |
| It's not much,
but it's the best I've got |
| And I thought you saw the good in everyone |
>> ALL: | Ooh, the good in everyone |
| You see the good in everyone |
| You see the good in everyone |
(instrumental break)
```

Blanking roll-up captions

Don't blank roll-up captions just to show a single pop-up caption. In other words, do not clear three lines of carefully-composed real-time captions to display a poorly-written NSI caption like

(Cheers/Applause) or, [worst of all](#), [| | |].

Minor errors

Document title

It's debatable whether "Canadian English-Language Broadcasters' Closed Captioning Standards and Protocol" really is possessive rather than descriptive. That is, I doubt we need the apostrophe.

Anyway, it's a very long adjective chain. Rewrite it as "Closed Captioning Standards and Protocol for Canadian English-Language Broadcasters." Among other things, it puts "closed captioning" first.

Programming day

[p. 7] ¶ The policy also requires that all such licensees close caption at least 90% of all programming during the broadcast day by the end of individual licence terms.

A lay reader of the term "broadcast day" will incorrectly assume that all 24 hours are covered. The Manual should mention that the CRTC-defined broadcast day ends at midnight. In fact, I'm not even sure when it actually starts – 6:00 A.M.?

Commitments

[p. 6] ¶ In its call for applications for licences for new digital, pay, and specialty television services, Public Notice CRTC 2000-22, the Commission stated that it expects applicants for new services to commit to close captioning at least 90% of their broadcast day by the end of their licence term.

However, the facts of the actual commitments are now [well established](#). Any captioning or description commitment by broadcasters undertaking Category 1 or 2 digital specialty services was rubber-stamped, ranging from no commitment to 90%. Just as an example, here is FashionTelevision's commitment: "In view of the still relatively sparse penetration of digital services and the corresponding likely his/her level of use of older recorded material, it is impossible to project the percentage of programming that will be captioned over the term of the license." FashionTelevision was *expected* to caption 90%, but the CRTC endorsed the applicant's reasoning.

The entire discussion of CRTC requirements is so incomplete as to be misleading. Nothing short of a "requirement" or "condition of license" has any enforceability with the Commission, which has never once fined or punished a broadcaster for failing to live up to captioning commitments. An "expectation" of 90% captioning or "encouragement" to achieve same has no practical relevance.

This entire discussion should perhaps be scaled back considerably or, if it's intended to be substantive, it must be *complete* even if the result does not entirely gladhand or congratulate the broadcast industry.

Program “hours”

[p. 8] ¶ Labour Intensiveness: It may take 18 hours or more to off-line caption a one-hour program depending on the complexity of the program, speaking rate, rate of scene change, and difficulty of topic.

An “hour” of television on commercial TV can occupy only 48 or fewer minutes – 20% shorter than a full clock hour.

Use of the term “steno”

[p. 10] ¶ A skilled caption stenographer listens to a program and writes the speakers’ words in what is called steno, a form of phonetic shorthand based on a special 24-key keyboard where common words are written in one stroke of the hands.

I am not sure why we can’t say that *steno* is a shortened form of the word *stenography*, and [define it](#) (“The art or process of writing in shorthand”). “Steno” seems to be used in a more sweeping sense to refer to the entirety of a stenographer’s theory, briefs, and lexicon.

Also, few words can be written in one stroke (pressing and releasing one or more keys and doing nothing else).

Real-time errors

[p. 11] ¶ Spelling: Mistakes in real-time captioning often look like spelling errors. However, they are usually the result of a mistranslation by the computer, or what is referred to as an “untranslate,” a phonetic rendering of a word that was not pre-programmed into the caption stenographer’s dictionary.

Not a very solid explanation. Mistakes in real-time captioning are often clear-cut operator error (like mishearing the word or not knowing which homonym to use). The intent here seems to be to explain *incomprehensible* real-time-captioning errors. I would expand this to more than a single paragraph or delete it.

Observations

Copy errors

The Manual is rife with copy errors. Hyphenation is particularly poor. I can edit a later version.

We have, in effect, the standard conundrum of typography: We're using the medium to discuss itself. A manual on accuracy and completeness of text transcription has to be very accurate and complete in itself.

I can heartily recommend that the Committee hire [Moveable Inc.](#) to proof the final report. It will cost good money, but you get what you pay for. I write clean copy, but I persuaded my publisher to hire Moveable to proof my [book](#) on Web accessibility. Only a few outright errors were caught, but Moveable made 2,000 suggestions concerning the 430 manuscript pages, the majority of which were enacted.