DOCKET FILE COPY ORIGINAL

Joe Clark
279a Danforth Avenue
Toronto, Ontario M4K 1N2
(416) 466-0784 *vox*
(416) 406-416-4 *fax*

RECEIVED
FEB 28 1996
FCC MAIL ROOM

26 February 1996

Office of the Secretary
Federal Communications Commission
1919 M Street, N.W.
Washington, D.C. 20554

Re: FCC NOI on closed-captioning and audio description, MM Docket № 95-176, FCC 95-484

Dear Secretary:

I am pleased to submit comments on the Notice of Inquiry on closed-captioning and audio description.

I am enclosing, for the perusal of all Commissioners, nine copies of my response, an appendix on captioning errors, the text of a 1989 presentation on the typography of HDTV captioning, and a selection of newspaper and magazine articles I have written on captioning. The appendix on captioning errors appears *on disc only*. On the same disc I include a WordPerfect 5.1 copy of this response.

If the Commission ever holds any kind of public hearing or meeting on captioning or audio description, I would be pleased to participate.

Sincerely,

Joe Clark

No. of Copies rec'd 0+8
List ABCDE

RECEIVED
FEB 28 1996
FCC MAIL ROOM

Response to FCC NOI on closed-captioning and audio description, MM Docket Nº 95-176, FCC 95-484

Joe Clark

A. Personal history

I am a hearing person who has been watching captioned TV for almost 18 years. I happened onto the medium of captioning by aimlessly channel-changing one night circa 1977 and found what was obviously ABC's *World News Tonight* with words superimposed on the screen. This, of course, was WGBH's *Captioned ABC News*, a weeknightly same-day rebroadcast of ABC's evening news show with open captions. Those captions employed an actual character generator (unlike the computer-esque characters in closed-captioning today) and were a greatly simplified and abridged version of the original dialogue. The captions thus were quite nice to look at but raised questions in hearing viewers' minds about the intent of the captions – were they in place to provide access to deaf viewers (who, according to prevailing stereotype, had too low a level of literacy to understand real English) or did the captions represent an additional track of information added to a television program, with a responsibility to effectively render the audio of a program in written form? Over the next 18 or so years, I would come to think rather closely about those issues, as I will describe later.

Captioning led to other interests. I wondered exactly how the characters were added to the picture, and had questions about some of the peculiarities of the font being used; after posing those questions to the Caption Center at WGBH, I began to learn more about typography and graphic design, which I have avidly followed and written about ever since. I was lucky enough to have been a verbal child with unusual reading and writing skills (I remember sitting at the front of my Grade 3 English class spelling words out loud for one student after another) and had something of an interest in languages. After graduating with a diploma in engineering from Dalhousie University in Halifax, I would put my linguistic acumen to use in earning a B.A. in linguistics from the University of Toronto via McGill University. I wrote a few linguistics papers on sign languages. Finally, the late '70s were the beginning of the microcomputer era, and the ability to use something as rudimentary as an Apple IIe computer to do word-processing suited my interest in type and writing; I've been using personal computers since that time and have written a number of articles on computer topics.

My interest in captioning led to an interest in disability issues in general, a topic I have also continued to follow and write about; I've also held jobs and done volunteer work in the disability field.

I have never worked in the captioning industry. I did apply for a job at the Canadian Captioning Development Agency circa 1987, but my considerable knowledge of captioning was deemed irrelevant ("You have no credibility in captioning," said Mardi Ferguson, a CCDA manager) and I didn't get the job. (That was all for the best, as it turned out, given the hideous captioning CCDA and its offshoot companies have produced.) I have also hit up a broadcaster or two for in-

house captioning positions, though in those cases I was more interested in managing captioning and access issues than in captioning actual programs. No jobs resulted.

Presently I am listmanager of an Internet mailing list on media-access topics (captioning, audio description, and the like), a relatively well-read forum with some 240 members around the world. Representatives of some of the bigger names in captioning subscribe to the mailing list, and I am not shy to upload comments about caption quality, which is by far the most pressing subject in captioning today. (Yes, more pressing than caption *quantity*.) To subscribe to the list, send the words *subscribe access* to `listmanager@hookup.net`. Partial list archives are at `gopher://trace.wisc.edu/00/ftp/PUB/TEXT/ORGS/MED_ACS/MAIL_ARC.TXT`.

Freelance writing has been my only source of income for about three years, though I've been writing since the late '80s and have more than 350 published articles to my credit in such periodicals as the *Economist*, *Popular Science*, *Technology Review*, *Details*, *BusinessWeek*, *Entertainment Weekly*, and about 30 others. The first article I wrote for pay was on the topic of captioning typography ("Typography and Television Captioning," *Print*, January/February 1989, enclosed); I've written about a dozen stories in total on captioning topics.

With nearly 18 years' experience watching captions (American and Canadian, English and French, closed and open), with extremely high literacy skills and an ironclad knowledge of the ins and outs of written English in many of its forms, with extensive expertise in typography, graphic design, linguistics, and computer software and hardware, and with (above all) fully functioning ears and eyes, I am in a unique position to evaluate captioning issues. People with my level of knowledge, expertise, and experience in captioning are rare (virtually nonexistent, in fact – I have met only one other hearing person who knew of captioning as a youth, and he now works for the Caption Center at WGBH), and it is from that vantage point that I offer my comments on the FCC's Notice of Inquiry.

The following sections relate to the NOI itself. I have not attempted to answer every question. I am also enclosing copies of relevant articles I have written on captioning.

B. Background issues in the NOI

¶3 and ¶6 in §II of the NOI ("Background") contain terminological and conceptual errors.

Captioning is *not* "similar to subtitles in that it displays the audio portion of a television signal as printed words on the television screen. To assist viewers who are hearing-disabled" – that, incidentally, is an atrociously bureaucratic term – "captions also identify speakers, sound effects, music, and laughter."

(i) The Commission must not make the mistake of equating closed-captioning with captioning. The current Line 21 closed-captioning system is not the only conceivable way to caption a film or video.

BOTTOM LINE → **The FCC must be very careful to draft regulations that do not in any way hinder the use of open-captioning (i.e., programming with captions always visible) and that do not place unnecessary restrictions on the use of technologies other than encoded Line 21 signals for captioning.**

(ii) Captioning is no more similar to subtitling than it is to the cast-and-crew credits shown at the opening and closing of feature films or the virtual ticker-tape continuously scrolling across the bottom of the frame on CNN Headline news.

- Subtitles are a translation. Captions are in the same language as the audio. (Some captioners may opt to caption passages in "foreign" languages via periphrasty – e.g., [SPEAKING ITALIAN] – rather than writing out the Italian words, but proper captions do not translate the audio.)
- Captions move to denote the location of who is speaking and/or to avoid obscuring titles or relevant onscreen action. (That's in the best case; some captioners are atrociously inept at positioning captions to show who is speaking, while at the same time the Line 21 technology imposes real limits on caption placement. Also, scroll-up captioning is very difficult to position for speaker identification.) Subtitles are fixed at the bottom centre of the screen, with rare exceptions.
- Captions denote sound effects (e.g., [PHONE RINGS])and indicate offscreen speakers by name (sometimes with hopelessly slipshod typography, other times elegantly, as in Megan: or Narrator:). Subtitles do neither of those things.
- Captions are nominally intended for deaf viewers; subtitlers assume that their audience can hear.
- Quality captioning renders repeated utterances, i.e., the same or nearly the same thing said more than once. Subtitlers might or might not render repeated utterances. For example, "Écoute-moi! Écoute-moi! Georges! Georges!" might be subtitled as "Listen to me!" followed by a blank screen; the thinking is that the hearing viewer will understand that the speaker uttered the same words twice, and will also recognize the name *Georges* without onscreen translation.
- Subtitling will often not give a deaf viewer sufficient information to *really* understand the film. (Then again, the same could be said for incompetent and/or reduced-*register* captioning, but more on that later.) A subtitled film may still require captions – for example, to caption the dialogue a subtitler assumes the audience will understand (see previous point) and to add speaker identification and notate sound effects.

FCC commissioners, like most Americans, likely have very little experience in watching "foreign"-language films, so the Commission is likely unaware that combination subtitled/dubbed films are quite common (try the recent Canadian film *Le Confessional*, with dialogue in English *and* French), and that some foreign-language scenes in English-language films do appear with subtitles (e.g., in *Four Weddings and a Funeral, Leaving Las Vegas, Nixon, Walker, Zentropa, Natural-Born Killers, Dances with Wolves*, and others). In all those instances, captions are still required.

- Subtitled or dubbed films often make no effort to translate songs, which typically remain in the original language. Captioning does render song lyrics, with variable success.

(iii) The Commission must make up its mind whether it intends to use its own term for audio description ("video description") or stick with the accepted industry term ("audio description"). The *descriptions* take place in the *audio* medium. Video is not the only thing that can be described – still pictures, live theatre, and film are all describable. While the FCC may have limited legislative purview (hence the focus on "video" description), that does not justify the deployment of its own malapropist terminology for a field whose name is already well-established.

C. Other topics in the NOI

The sections below respond to the cited paragraphs in the Notice of Inquiry.

11. **We... ask parties to submit information regarding the number of individuals in this country who can benefit** [from captioning and description].
12. **Closed-captioning can be an effective tool in teaching literacy skills for young children as well as for the estimated 23–27 million American adults who are functionally illiterate. It also provides a useful learning aid for the approximately 3-4 million Americans learning English as a second language.... Closed-captioning, for example, can be activated when the mute button on the television receiver is depressed or when noise levels in the viewing environment impede normal hearing (e.g., airport terminals, hotel lobbies, waiting rooms). We seek comment on the nature and extent of each of these potential benefits, including the most up-to-date data on the number of individuals who would utilize closed captioning and video description for these purposes.**

Other respondents will authoritatively estimate the number of deaf/hard-of-hearing and blind/visually-impaired persons who could benefit from captioning and description. I wish to address the topic of hearing viewers of captioning (and, peripherally, of sighted listeners of description).

(i) **The hearing and the deaf in the captioning audience**

A perennial fiction of the accessible-media movement is that hearing people so vehemently dislike captions that they adamantly reject their presence on television screens. I read a single document from the National Captioning

Institute (NCI) circa 1980/1981 claiming to have surveyed hearing viewers and discovered that, indeed, a majority of them reject captioning altogether. (I no longer have a reference for that study. Perhaps NCI, which consistently behaves as though it is a sort of pan-global combined Vatican and Kremlin of captioning, will cough up that study, which might have been undertaken in coöperation with PBS.) For this reason, a technically backward and unreasonably complex ghetto technology of *closed*-captioning was developed.

NCI's line made sense over a decade ago, or at least it made sense if you considered how rare decoders were (for a time, only Sears stores sold them), how infrequent captioned programming was, and the difficulty in bringing average people up to speed on the topic. If you stopped typical shoppers in a Sears store and asked them to watch a minute or two of captioned TV – in a bustling, noisy environment, standing up and looking down at a console TV, and without giving the subject a choice of which program to watch – *of course* hearing people are going to tell you they hate captioning. If that were my only exposure to captioning, I'd hate it too.

This point is central:

BOTTOM LINE → **We have a closed-captioning system in North American *purely because it was believed that hearing people resent open captions.***

In all my years reading about, watching, arguing about, and calmly discussing captioning issues, I have never been persuaded that this received wisdom, this stereotype, this *idée fixe*, is even remotely valid. However, even if the research were rock-solid and accurately reflected hearing public opinion at the time, we now live in a different world and a different approach is required.

Let's consider what the concentration on *closed*-captioning hath wrought. True, we now have a tremendous array of television programming and home-video items equipped with captions, from news and talk shows to documentaries to music videos. The majority of the U.S. broadcast networks' daily schedules is now accessible to captioning viewers, as are thousands of commercials, movies on home video, and music videos. Clearly, these are tremendous achievements.

However, by sequestering captioning in its own technological gulag – which until recently required the deliberate purchase of an overpriced, hard-to-find external decoder to turn the phantom signals into visible words – captioning has acquired a mystique. Until recently, very few hearing people *really* know what you're talking about when you mention captioning. ("Wait, that's, what, sign language or something?" "No, more like subtitles." "Oh, them. I hate them. I saw this subtitled movie once and jeez, the actors would be talking for half an

hour there and all the subtitles would say is, like, 'I disagree.' Is this captioning thing like that?") Because the captions were by definition hidden, average hearing people had no opportunity to get *used* to captioning. (This worked to the advantage of the less adept captioners, as I will come to soon enough.)

Moreover, the secrecy of captioning fueled what could be called a binary obsession of captioning: All people are really concerned about most of the time is whether an item was or was not captioned, with all other concerns – like quality and accuracy and comprehensibility and timing and typography and writing conventions – entirely out of mind. To this day, all some people are worried about is stuffing Line 21 full of signals so that (*a*) program producers or broadcasters can state that a program is captioned, thus giving the impression of meeting their civic duty to deaf people or (*b*) they will meet statutory requirements, of the sort we have in Canada, for minimum hours per week of captioned programming.

If we concentrate on closed-captioning as the only expression of captioning, we condemn closed-captioning to a minority backwater, something watched only by marginalized deaf people. Now, it was fortuitous for NCI and other decoder manufacturers that teachers of English as a second language discovered captioning and began to suggest that their students buy decoders to "improve their English" (see footnote 33 in the FCC NOI, which states that 40% of the external decoders sold in 1989 were to ESL students), since sales of decoders had apparently plateaued by the end of the 1980s. That plateau was to be expected; most deaf people who could scrape the money together to buy a decoder had done so, and the rest of the world, kept out of the loop due to the secret nature of coded captions, didn't care.

(ii) **The new reality of captioning**

Now, however, we live in quite a different world. The Television Decoder Act of 1990 required caption decoders in the vast majority of U.S. television sets. (In a rare case of the Tory régime's failing to immediately ape its American masters, the government of Canada passed no similar measure. However, because "split runs" of televisions intended purely for the Canadian market are rare, Canadians receive virtually the same TV models as Americans, meaning that decoder-equipped TVs are about as commonplace here as in the U.S.) As I show in an analysis based on publicly-available figures ("The hearing majority of captioning viewers: An analysis," enclosed), by the end of 1995 some 59.19 million decoder-equipped TVs had been sold in North America. Most of them were sold to hearing people, since hearing people are the majority in North America (and since most deaf people have decoders already).

If, in the American case, slightly more than 0.552% of those hearing people watch TV with captions on, then *hearing people become the majority audience*

of captioning in the U.S. If slightly more than 0.775% of hearing Canadians with decoder-equipped TVs watch television with decoders turned on, hearing people become the majority audience of captioning in Canada. It is reasonable to assume that fewer than one hearing household in 100 has been curious enough to turn captions on and leave them on.

BOTTOM LINE → **Hearing people are now the majority audience of captioning in North America, though that does not mean captions must not be optimized for deaf viewers.**

If those percentages, tiny though they are, strike you as unpersuasive, consider that some 26½ million decoder-equipped TV sets are sold every year in North America. As the years go by, an ever-smaller percentage of hearing people with caption-capable TVs will need to turn their decoders on in order for hearing people to become the majority audience of captioning. In a decade, or even in half that time, not only will hearing people outpace the deaf as captioning viewers in percentage terms, the former would massively outnumber the latter, with tens of millions of eventual hearing captioning viewers and a few million deaf captioning viewers.

Thus, the FCC needs to accept reality and acknowledge the following:

BOTTOM LINE → **Captioning exists to make audiovisual media accessible to deaf and hard-of-hearing persons and must always fully serve that audience, but for captioning to thrive and gain respect as a medium, it must also be seen as pleasing, relevant, helpful, and accurate to the hearing persons who make up the majority of the contemporary captioning viewership.**

Stating that hearing people are the main audience of captioning is rather like claiming that the sun revolves around the earth. It's so contrary to accepted wisdom that a first impulse may be to dismiss the idea altogether. But my numbers, and my analysis, are correct. The old NCI way of thinking – captions are an annoyance that must be hidden from sensitive hearing viewers at all cost – is hopelessly passé and does nothing but frustrate the improvement of captioning as a medium.

(iii) **Hearing viewers as enthusiasts of captioning – and as the captioning police**

While I have conducted no surveys of hearing viewers, I feel confident in the following estimates of the hearing population. In all these estimates, one assumes that captioning has been presented to viewers as something they might want to watch. These hearing viewers might then be classified into three groups.

- Some 40% of hearing television viewers are not particularly interested in captioning; of that group, a small portion actively dislike captioning and will never accept it in their homes.
- Another 40% rely on captions for language-learning purposes. If that seems like a large proportion, remember that typical hearing children fall into this category, as do ESL learners, adults and others with low literacy skills, and some learning-disabled persons, for whom captioning is as much an access technology as it is for deaf people.
- The remaining 20% are the group with the greatest relevance to the future of captioning. In this group we find sophisticated viewers able to assimilate a great deal of information at once – *very* highly literate viewers, younger adults who have grown up in the eternal presence of TV and computers, polyglots, cinéastes. For these people, captioning is not only acceptable but desirable and useful, since they are perfectly capable of following main audio and dialogue and all onscreen action *and captions* simultaneously.

Perhaps my numbers are wrong. Perhaps the proportions of the three groups are 65–30–5, or 90–8–2. However, the third group definitely exists, and even a tiny percentage adds up to large populations when 26½ million decoder-equipped TVs are sold each year. There are thousands (hundreds of thousands, millions) of hearing captioning viewers, most of them watching captions on decoder-equipped TVs, who actually *enjoy* captioning.

And those people have an advantage: They can hear the audio and compare the captions simultaneously. Thus, this group of hearing viewers amounts to an audience that, according to traditional captioning-industry dogma, does not exist, and one that stands as a kind of watchdog of the captioning industry. If a program is badly captioned – words misheard, characters misnamed, sentences mal-edited – deaf or hard-of-hearing viewers may never know the difference *because of* their hearing impairment. However, with rare exceptions, you can*not* fool a hearing captioning viewer. If a mistake has been made, we'll know.

Think for a moment about which strategy will have the greater success: Providing and marketing captioning only for deaf and hard-of-hearing viewers (a small minority in numerical decline), or providing and marketing captioning for deaf people *and* the millions of hearing viewers who enjoy captioning for whatever reason and can watch captioned TV thanks to built-in decoder chips?

BOTTOM LINE → **The future of captioning rests in marketing the medium to *hearing* as well as deaf viewers.**

(iv) **Captioning errors in the real world: Introduction**

Let's not pretend that there are no mistakes in captioning. The Commission asks in its NOI –

> 33. **We also request comment on the current accuracy of closed captioning on television programming. There have been reports regarding inaccurate closed captioning, especially for live programming where there is no chance to review and correct for errors. We seek comment on the extent of this problem and any steps that could be taken to promote accurate and appropriately formatted captions. We are particularly interested in proposals to address such concerns through private-industry initiatives in order to avoid unnecessary government regulation in this area.**

I have been watching captions for nearly 18 years, and I can confidently say that in 1996 there is no single captioner in the United States who can be relied upon to do work of consistently high quality for prerecorded programs. In broad terms, the Caption Center at WGBH is known for a commitment to quality captioning (and, in particular, for a commitment to super-accurate *timing* of captions, which can make a world of difference in simple enjoyment of a captioned program), but as I have been documenting over the last two years on the Media Access mailing list and elsewhere, the Caption Center has increasingly been coasting on its reputation; its standards have been nosediving. The perennial runner-up National Captioning Institute never did consistently good work in any sector of captioning *except* for commercials (in which NCI was no better or worse than the Caption Center, and both were generally excellent) and has also been nosediving. The bit players, such as Vitac and Captions, Inc., have never created credible captions (the latter particularly – Captions, Inc.'s typographic style is vastly different from nearly everyone else's in the industry but fails in being genuinely *better*).

In Appendix A of this response, I exhaustively document examples of captioning errors.

(v) **Captioning errors in the real world: Real-time or live captions**

In ¶33 of the NOI, the Commission seems to concentrate on the quality of real-time or live captioning. It's important for Commissioners to separate live and prerecorded programming in considering the issue of captioning quality. There is a tremendous demand for real-time captioning throughout North America; it is the *only* means of providing full access to a live TV program for deaf viewers. However, as the Commission is surely aware, real-time captioning is not done by typing on a standard QWERTY keyboard on an ordinary personal computer. Rather, a trained court reporter presses combinations of keys on a stenotype machine; the process amounts to verbatim phonetic shorthand in the computer medium. A personal computer interfaced with the stenotype keypad converts the shorthand into actual English words and sends those words out over Line 21 to be seen as captions.

However, in traditional handwritten ("penwriter") shorthand, and even in court reporting using a stenotype keypad *not* interfaced with a computer, homonyms like *to, two, too, tu* ("'Et tu, Brute?'"), *II, 2:00* and *2* can all be rendered with a single string of symbols. A human court reporter later reviews the shorthand and, based on the context of the words in question, types or handwrites or otherwise renders the shorthand into its correct English form. But with real-time captioning (and real-time court reporting), there is no human intervention: Once the captioner presses what he or she thinks are the right key combinations to produce a word, there's no turning back. (Actually, a backspace command to erase characters has been part of various versions of the Line 21 standard but is not universally understood by all decoders.) Thus, to accommodate homonyms and other linguistic details, real-time captioners have to memorize *different* keystrokes for all homonym variants. Sometimes the captioner will make a mistake with such homonyms.

In fact, since we're talking about human beings at work under stressful circumstances, mistakes of various kinds are inevitable, and since *content words* (major nouns, verbs, adjectives, and adverbs) require more finesse on the keyboard than *function words* (*and, of, the, a, an, with, but,* and the like), more often than not the mistakes we see in real-time captioning are with content words – words that strongly affect the meaning of a sentence. Even if a captioner maintains 98% accuracy, on average four words out of 200 will be wrong, and you can bet on the fact that those incorrect words will be major ones.

It is unlikely in the extreme that the "accuracy" of real-time captioning will be meaningfully improved. The more proficient real-time captioners are already doing the best work we can hope for. At the same time, the overwhelming demand for real-time captioning has forced caption firms to hire captioners with lower skills. Only some people are capable of keeping up with rapid-fire speech, and based on conversations with real-time captioners, the evidence suggests that such people are born, not made. The only way to find out if you *can* stenocaption a program at up to 200 words per minute is to undergo a difficult period of training and accumulate experience. Either you'll plateau at 120–140 words per minute or will excel and hit 200. However, there is enough demand for real-time work that even the captioners who can handle only 120–140 words per minute will be hired and put on the air despite the fact that such qualifications are far too low for real-world captioning.

The only way to improve the quality of real-time captioning is to increase the number of extremely-qualified court reporters, which will in turn require a very large influx of new people to sift through. That is clearly beyond the scope of the NOI and beyond the Commission's purview. Accordingly...

BOTTOM LINE → **There is nothing the Commission can directly do to improve the quality and accuracy of real-time captioning.**

(vi) **Captioning errors in the real world: Captions on prerecorded programs**

Captioning of prerecorded programming (sometimes called offline captioning) involves these steps:

- Working with a timecoded videotape,* transcribe the program, or scan or load a script and check it against the actual dialogue. Add notations for sound effects and IDs for speakers where necessary.
- While still working with the timecoded videotape, break the captions up into manageable blocks (editing where necessary), and position them in appropriate spots on-screen. Time the captions to match the dialogue and video.
- Check the finished work, ideally more than once and with more than one person.
- Create a computer file containing the caption text for mating with the master tape (the one with invisible timecodes). This *encoding* process can be done in-house by the captioner or at some postproduction house of the client's choosing. The result is a closed-captioned master tape that can be distributed and duplicated at will.

Sounds pretty simple, doesn't it? However, there are a number of caveats:

- The fact that the English language comes in spoken *and* written forms does not mean that an easy 1:1 correspondence between the two is always possible. Further, captioning is a kinetic medium: Blocks or lines of captions move, appear, and disappear, making the psychological process of reading captions meaningfully different from reading static text on a printed page.
- While notation of sound effects and musical interludes and explicit identification of speakers may be found in some uncommon genres of printed English (screenplays and scripts for theatrical plays, for example), the lessons learned from those genres are not necessarily applicable to captioning.
- Because captions are nominally intended for deaf viewers, it's necessary to pay attention to details we wouldn't think twice about in, say, written English or film subtitling. For example, if a character is seen on a TV program moving his lips but not actually saying anything, as occasionally happens when background music dominates the soundtrack, captioners need to note that fact with a caption like [NO VOICE]. Also, merely captioning the lyrics of music may not be enough to communicate the importance of the music to the

* One with visible timecodes burned into the videotape. Another master with invisible timecodes is held in abeyance for the *encoding* step.

narrative, as in PBS's *Tales of the City*, in which 1970s music was so prominent it had to be identified by name *and* captioned verbatim.

- Though captioning and audio description are supposed to be the last processes in production of TV programming, producers can and do alter the dialogue or audio after captioning is finished, making captions and audio differ.
- Since Line 21 closed-captioning consists of computer codes transmitted along with a picture for eventual decoding into comprehensible titles, a wide range of technical foulups are possible – everything from a VCR in a postproduction house that moves caption data out of Line 21 to an encoding machine that's slightly out of spec, causing garbled captions on some decoders while working just fine on others.

However, I am more concerned with captioning errors *directly caused by captioner incompetence*. The fact is that captioning work is not suited to everyone. In fact, it's not suited to anyone *other than* people with absolutely airtight knowledge of English writing conventions, huge vocabularies, near-perfect spelling, exceptional knowledge of homonyms and tricky words, uncommon listening skills, wide general knowledge, and an ability to write down, if not necessarily converse in, a broad range of English styles or *registers* – rabbinical, aristocratic British, urban black, parliamentary, Texan, medical, legal, urban gay, and other styles. I doubt that more than a fraction of the people doing captioning at the B-list firms are truly qualified for their jobs. Not everyone working at NCI and the Caption Center should be in the captioning business, and some captioners really ought to consider more appropriate careers.

Just as the rise of desktop publishing has led secretaries and office managers to believe that buying a computer, a laser printer, a few fonts, and desktop-publishing software will turn them into graphic designers, the impression widely held by captioning neophytes that captioning is "straightforward" and *conceptually* simple because it can be carried out on inexpensive personal computers has done more harm to the captioning medium than anything else. The Commission notes, in ¶19 of the NOI, that some 60 companies in the U.S. do captioning. That's probably true, and while no more than two of them (the Caption Center and NCI) are even worth talking about, those two are now producing errors of the sort we'd expect from semiliterate opportunists who mistakenly consider themselves qualified to caption television programming.

I am constrained at this juncture by the limitations of the print medium. The only way to properly discuss captioning is live and in person, with tapes rolling to show examples of good, bad, and mediocre captioning. However, I have invested a great deal of time in cataloguing captioning errors made by the Big Two captioning companies (and, to a lesser extent, the uniformly incompetent Canadian offline captioners and Captions, Inc., the American captioner with

extensive contracts in the home-video market). Appendix A consists of a large compendium of recent captioning errors. The Commission can plainly see that even putatively high-quality captioning bodies like the Caption Center and NCI can't do good work day in and day out.

BOTTOM LINE → **If the Commission decides to mandate a certain *quantity* of captioned programming, inevitably we will end up with poor *quality* for most of that quantity. Enforced quality standards are therefore necessary and overdue.**

Again, as in the case of real-time captioning, it's a human-resources issue: Schools just are not churning out graduates with airtight English literacy. Anyone who can pass a proofreading test (itself a misleading indicator of captioning success) can get a job in captioning, and some captioners' standards are so poor that even correctly-transcribed programming can still be a mess.

Let me give you an example of the latter point. In prerecorded programming, long sentences typically must be broken up into individual captions. It's important to make each caption an understandable conceptual unit,* a fact that bad captioners fail to acknowledge at all and which even good captioners have trouble with. For example, in the science-fiction TV movie *Alien Nation: Millennium* broadcast by Fox in January 1996, a narrator gives viewers new to the *Alien Nation* storyline a quick rundown:

> **As we look back from this final week of the 20th century, on the verge of the new millennium, surely the most amazing experience humanity shared was our first view of the huge Tenctonese spacecraft. Making a forced landing in California's Mojave Desert seven years ago, it brought a quarter-million beings bred as slaves to labor in any environment. Stronger than human beings, with keener senses, these alien Newcomers joined our society, including the L.A. Police Department. Welcomed by many, they have also faced much fear and prejudice from others. With no way to leave Earth, the Tenctonese Newcomers have become the latest immigrants to join the population of America.**

This monologue was delivered over relatively quick visual edits of the Tenctonese spacecraft, people gaping at it, Newcomers and humans mixing on city streets, and the series' principals. The monologue was captioned by the Caption Center (nominally the best in the business) thus:

---

* The terminology has to be vague here. In the linguistics and ESL literature, the term *chunking* is used to describe the process of subdividing sentences and phrases into comprehensible units. However, the purposes to which linguists and ESL teachers put chunked text are different from captioning's purposes; also, captioning is saddled with far more constraints than flashcards or other media used by linguists and ESL teachers. Accordingly, deciding where to divide long sentences into short captions is, like captioning itself, an art rather than a science, and one governed by heuristics or rules of thumb rather than ironclad rules. Still, as in the case of art or pornography, while it may be difficult to *define* sensible caption division it is not particularly difficult to spot good or bad division when it occurs.

NOW
[illegible] BACK
FROM THIS FINAL WEEK

OF THE 20th CENTURY

ON THE [illegible]GE
OF THE NEW MILLENNIUM

SUPER[illegible] AMAZING
EXPERIENCE HUMANITY SHARED

[illegible] VIEW OF
THE HUGE TERRIONESE SPACECRAFT.

MAKING A FORCED LANDING

IN CALIFORNIA'S MOJAVE DESERT

SEVE[illegible], IT BROUGHT
A QUARTER MILLION BEINGS

BRED A[illegible]LAVES,
TO LABOR

IN AN ENVIRONMENT,
STRONGER THAN HUMANS

WITH KEENER SENSES

THESE [illegible] NEWCOMERS
JOINED OUR SOCIETY

INCLUDING THE L.A.

POLICE DEPARTMENT.

WELCOMED BY MANY,
THEY HAVE

AND PREJUDICE FROM OTHERS.

WITH NO WAY TO LEAVE EARTH

THE TENCTONESE NEWCOMERS

HAVE BECOME
THE LATEST IMMIGRANTS

TO JOIN THE POPULATION
OF AMERICA.

(My fonts here do not permit me to show that all this text with the exception of Man: is set in italic. Centering is also somewhat unrepresentative here.) It must be noted that the Caption Center refuses categorically to end a caption with a comma, arguing (speciously) that in decoder fonts a comma looks like a period and can be confused for one. (Funny how commas and period can be used *within* captions, and that captions *can* end with colons or semicolons, two other characters containing periods and commas.) Combined with poor chunking or caption-division, this no-comma standard makes it rather difficult to follow *and fully understand* the dialogue.

As you can see, the Caption Center took very little care to make each caption a conceptual block; instead, the captioners at work on this program decided that *every shot change* or cut in the video would be given a new caption, sense be damned. (This is how most Canadian captioners operate, by the way, and it is *the* worst perversion the Canadians foist on viewers.) Some of the Caption Center's punctuation indicates that the captioners at work did not understand the sense of the words – what was spoken, for example, was "Making a forced landing in California's Mojave Desert seven years ago, it brought a quarter-million beings bred as slaves to labor in any environment. Stronger than human beings, with keener senses, these alien Newcomers joined our society, including the L.A. Police Department," not "(Making a forced landing) / (in California's Mojave Desert) / (seven years ago, it brought a quarter-million beings) / (bred as slaves, to labor) / (in any environment, stronger than humans) / (with keener senses) / (these alien Newcomers joined our society) / (including the L.A.) / (Police Department.)"

Issues like caption division are not entirely obvious and are highly prone to error. And it is in issues like these that mom-and-pop captioning companies can be relied upon to make mistakes – consistently. If the big players are doing it, certainly the little ones will.

Moreover, lately I have been catching NCI and the Caption Center in the act of committing the most basic captioning sins – mishearing the dialogue, getting characters' names wrong, simple typing errors – that no amount of boasting or excuse-making can justify. I don't care how much of a service you purport to provide to downtrodden deaf citizens; there is *no* excuse for such mistakes. We can expect those errors to become more common if captioning itself becomes more common.

(vii) **The need for captioning quality standards**
Captioning is not like the main audio or video of a program, which networks and other producers labour to perfect. After all, dropouts in the audio or staticky visual passages simply would not be tolerated by the audience *or* by the broadcaster. The same is not true for captioning.

BOTTOM LINE → **Network and production-house standards for audio and video quality per se are high. But the industry standard for captioning is low: The industry accepts captions if they are well-done *or* poorly-done.**

I believe that, over time, the vastness of the hearing audience for captions – and the ease with which hearing captioning viewers can phone up networks and producers and captioners and complain about misspellings, maladroit editing, and other faults – will force captioners to get their act together and do captioning right the first time.

BOTTOM LINE → **The Commission should impose *quality* standards if it legislates a certain *quantity* of captioning.**

The NOI in ¶33 states that "we are particularly interested in proposals to address [quality and accuracy] concerns through private-industry initiatives in order to avoid unnecessary government regulation in this area." I suppose that's a laudable Reagan-legacy sentiment, but the major players in the American captioning industry distrust each other so much that, after more than ten years, there is *still* no industry association. It is *unrealistic* to expect the captioning industry to police itself. How could it? Even good captioning firms are producing bad captions, and bad firms have always produced bad captions; everyone has gotten away with it.

Purely as a starting point, let me suggest some wording for a future discussion of quality standards.

BOTTOM LINE → **The FCC might declare that video, audio, audio descriptions, and captions are integral components of a program signal and require uniformly high standards. In particular, the FCC might require that captioning be of a quality comparable to that of the video and main audio of the captioned program. In captioning, broadcasters would be obliged to respect and accurately represent the original script and audio track using typographic and notational standards comparable in excellence to accepted standards in the print medium with the result that a program could be fully enjoyed and understood by relying solely on the content and presentation of its captions.**

I'm under no illusions that even the larger and more credible American captioners would accept a degree of regulation even as vague and permissive as the wording above. However, such a statement of intent could act as an overture to form an industry association for television captioning with the authority to impose standards – technical, typographic, linguistic, and otherwise – on member firms. If not, the Commission must accept its duty to regulate an industry that, while responsible for making thousands of hours of programming nominally accessible to deaf viewers, works with virtual impunity. While the captioning industry's absolute power has certainly not led to corruption in any widespread sense, there *has* been a certain degree of self-satisfaction and coasting on reputations, not to mention a tacit acknowledgement that even poor captions are better than nothing for the downtrodden deaf minority.

If the only way to improve caption quality is to place the industry under FCC regulation, so be it.

(viii) **Moral rights and copyright issues**

Copyright holders have the legal right to control the *derivative works* made from their property (so-called "moral rights"). For example, a film studio has the right to veto or approve a proposed German- or Chinese-language dubbed version of a movie. However, captioners and audio describers, with the rarest exceptions, do their work behind the backs of directors, screenwriters, editors, and creative people; while they have implicit license to create a derivative work (i.e., a captioned and/or described TV show or movie), it does not follow that captioners can edit dialogue however they wish and scatter captions over the frame as they wish (and, in the case of description, talk over dialogue, music, and sound effects as they wish) as though they had *explicit* clearances to do so *in detail*. But that's exactly how the industry works. Apparently Woody Allen, incensed at some typically ham-fisted editing on the part of NCI, now requires printouts of caption text in advance. If so, he's alone.

BOTTOM LINE → **The Commission needs to explicitly acknowledge that significant copyright issues arise in creating captioned and/or described versions of television and other programming. These issues have not been adequately addressed. In fact, they have scarcely been addressed at all.**

At the same time, it would be a mistake for copyright holders naïve about the demands of captioning and audio description to demand absolute veto power over every word, comma, and description; some specific and customized textual treatments are genuinely required by the captioning and audio-description media. But before a balance can be achieved between the competing demands of copyright absolutists and captioning/description absolutists, the relevant industries have to start *discussing* the topic in the first place.

---

11. **We ask commenters to provide data regarding the current availability of closed captioning of television video programming. Has the amount of closed captioning been increasing in recent years, or has it reached a plateau?** [...] **To what extent are each of the following types of programs closed-captioned (on the media within the scope of FCC jurisdiction): entertainment programs, local and national news, documentaries, public affairs programming, children's educational programming, other types of children's programming, sports, movies, cable public access programming, and live vs. pre-recorded programming?**

Captioning is widely available on broadcast network and cable TV, though the smaller networks like Black Entertainment Television and Comedy Central do little captioning. Other respondents will have more authoritative numbers. As for genres: I've heard of very little captioned "cable public access programming" (and seen very little of it).

The Commission must not overlook seemingly "marginal" areas of programming (unless restricted by law). I have written extensively about the virtues of captioned music videos, for example. TV commercials need captioning, including public-service announcements and the often-overlooked promo spots, teasers, and bumpers for *upcoming* television programming. (The only network I know of that captions such commercials is CTV in Canada.)

BOTTOM LINE → **Any television program worthy of telecast is worthy of captioning (and audio description). No genre of programming should be exempt from requirements to caption, with the exception of programming with extended dialogue in languages which cannot be written in the Roman alphabet.**

Wordless programming is virtually impossible to find. Even wordless TV commercials are rare. Any television program with dialogue or soundtrack should be

captioned. Obviously, since the Line 21 closed-captioning system, designed by linguistically naïve Americans, cannot accommodate even common "foreign" languages like French (still missing the character Ç in decoder fonts even after two revisions to the decoder character set), any programming in Chinese, Japanese, Korean, Farsi, Hindi, Hebrew, Russian, or another language not written in the Roman alphabet cannot be captioned at any reasonable level of acceptability. Programs with *brief* passages in such languages *can* be captioned – most non-Roman languages have romanization schemes that captioners can work with. (An early captioned program – the miniseries *Shogun*, with extended Japanese dialogue – was captioned with romanization throughout.)

---

11. **We seek comment on the current availability of video description, including its availability within each of the categories described in our discussion of the availability of closed captioning. Is this service confined to certain PBS programs and movies carried on NTN? Are there efforts by other video programming providers and producers to provide video description or at least experiment with this service? Are there particular program types that are more likely to carry video description, and, if so, why? To what extent are live programs video described?**

Audio description is rare. We typically find it on "prestige" programs on PBS, like *Masterpiece Theatre* (drama), *Nova* (documentary), and *Mister Rogers' Neighbourhood* (wholesome children's fare). The commercial networks haven't touched the audio-description medium at all, to my knowledge. There are a few explanations for this:

- Technological limitations: The only way to transmit an encoded audio track is to use one of the two audio tracks provided by the MTS standard. Thus you can have *either* stereo main audio and no descriptions *or* monaural main audio with descriptions but not both. Broadcast networks are in no hurry to give up the luxurious Dolby Surround Sound they laboriously add to programs via MTS. Worse, not all stations even broadcast in stereo in the first place.

  Moreover, there is no way to encode an audio track on a videotape. It is thus impossible to record broadcast shows with Descriptive Video Service descriptions such that the descriptions are hidden on the tape until you want to hear them. Any tape you make will give you either no descriptions whatsoever or descriptions all the time. Accordingly, it is impossible to produce exactly one home-video cassette of, say, a Hollywood film that can serve blind *and* sighted audiences in the way that hearing and deaf people can be served with one tape via closed-captioning.

  However, it *is* possible to encode a separate audio track on 12" video*discs* (they all come with a third audio track as standard equipment), a fact which purveyors of descriptions *must* capitalize on. (It's entirely conceivable for Hollywood studios to distribute a single videodisc per title containing main

audio, main video, closed-captions, and a separate audio track combining main audio and descriptions.) All those issues are of course beyond the FCC's legal jurisdiction.

- Novelty: Captioning has been around for two decades. Audio description as a concept, let alone a reality, is barely a decade old.
- Psychological displacement: Broadcast executives can well imagine how difficult it is to understand television without a soundtrack; simply hitting the Mute key on a remote control will induce almost immediate frustration without captions. However, it is relatively easy to follow at least the broad gist of a TV program without access to the video. The information loss suffered by deaf and by blind viewers due to their respective disabilities is not *strictly* comparable, and the means used to remedy the information loss are not comparable (captioning ≠ description), though those are not valid excuse for failing to provide access.
- Political organization: Deaf people (*sic* – here I really mean "culturally Deaf" people) are more effective at lobbying for their interests than blind and visually-impaired people. Average citizens are more likely to conceive of deaf and hard-of-hearing people as a separate social entity (actually two separate entities, but that distinction is lost on most people) than blind and visually-impaired people.

Broadcast/commercial networks, then, have little incentive to go to the trouble of providing description. The least they could do is provide description on a few episodes of a few shows (as was done in closed-captioning's prehistory).

Live programs are rarely described. It's tough to do: Typically one provides description only during pauses in dialogue, and in live events, there are very few pauses. And a certain lagtime comes into play: A moment after describers realize that they're hearing a pause, the pause might end; if the describers started talking, they might have to stop abruptly to let the main speaker be heard.

I have to concede the Commission's point in ¶33: "Is it necessary to require video description of a sporting event that already provides a play-by-play commentary or that is covered by a radio broadcast?" The Commission would be mistaken to allow one medium to rely on an entirely separate medium for access; in other words, a TV broadcaster should not be exempt from audio description because a radio broadcast is nominally accessible. TV programming must be accessible unto itself.

I agree that description for some types of programming could be seen as superfluous, sports play-by-plays being a common example. However, I am concerned that exempting broad classes of programming from audio description will encourage broadcasters to fudge a little and try to shoehorn *other* kinds of programming under that exemption.

Even a sporting event may not be suited to a *blanket* exemption from an audio-description requirement. I'll draw a parallel with captioning. It's entirely ordinary for play-by-play commentators to talk more or less nonstop through, say, a hockey game. However, not all those words are necessarily captioned. In the early days (the 1988 Olympics come to mind), one approach had the real-time captioner continue to caption all spoken words, but with those words collected into large pop-up captions and displayed chunk by chunk on-screen instead of the usual continuous scrollup stream of captions. You could ignore the chunks of captions or read them, as you wished; continuous scrollup captions, on the other hand, always call for attention.

Nowadays we're more likely to find no captions whatsoever most of the time during a sports event with continuous verbal play-by-play, and captions only after a break in the action (after a goal or foul, say). Obviously there are exceptions; occasionally a factual issue (like the score) will be captioned amid a running monologue of commentary. The goal here is to allow an assumed deaf viewer to actually watch the action rather than read about it a ridiculous number of seconds after the fact. (The play-by-play commentators are reacting *to* an event, which captioners *then* transcribe. The time delays can be considerable in a fast-paced hockey game.)

Similarly, in the case of an audio-described sporting event, one could imagine a lack of audio description through much of the play but *some* description at selected points (after a goal or foul, say). I do not see a way for the Commission to legislate such moderate, selective audio description created with informed understanding of the interaction between description and certain genres of programming. I am more or less resigned to the Commission's exempting programming where the main soundtrack broadly adopts the functions of audio description.

---

27. ***The application of mandatory requirements.*** **One issue we seek comment on is which entities should be subject to any mandatory requirements. Broadcast television stations and other delivery systems produce some of their own programming, such as news programs, but they often obtain programming from program producers or independent distributors of such programming. The legislation pending in Congress refers to "video programming providers or owners," and a House report states that it "is clearly more efficient and economical to caption programming at the time of production and to distribute it with captions than to have each delivery system or local broadcaster caption the program." Should mandatory requirements regarding closed captioning or video description be imposed on entities that produce or distribute programming that is aired on television? Or should they be confined to the "providers" of this programming, i.e., broadcasters, cable systems, wireless cable, SMATV operators, direct-to-home satellite services, and local exchange carrier/video dialtone systems? Should both providers and owners/producers be subject to any mandatory requirements that might be adopted, with parties being allowed to allocate responsibilities in this area by private contract? We seek comment on current industry practices for providing closed captioning and video description. Which entity – producer, distributor, or provider – normally provides these services?**

Mandatory requirements for captioning and description *should* be imposed. 99% captioning within five years and 75% description within eight years seem like reasonable numbers.

If outside captioning "agencies" are guilty of serious sins against captioning, then "in-house" captioners are even greater sinners. I have *never* seen *any* captioning created by the actual producer of the programming that was anything but appalling. I am not exaggerating.

I doubt it will be within the FCC's purview to dictate whether it is program producers or third parties who must do captioning. I do support the concept of assigning responsibility for captioning jointly to the broadcaster *and* program producer to be dealt with as they agree.

As for audio description: I support the same provision, but I anticipate that it will be some years before any audio-described programming on broadcast networks even comes into being and far longer before in-house audio-description becomes commonplace. (For PBS, the work of the Descriptive Video Service effectively amounts to in-house production of descriptions.)

In ¶28, the Commission asks about exemptions. I oppose exemptions in general, with the reluctant exception of exempting an audio-description requirement for programs whose soundtrack fulfils nearly the same purpose as description (e.g., sports play-by-plays – see above). Disabled persons have a right to the same information nondisabled persons take for granted.

I oppose exemptions due to "economic burden." Captioning – even good captioning – is not hugely expensive. Again, if a program is worth x dollars to produce, it's worth $x+c$ dollars to caption and $x+c+d$ dollars to caption *and* describe. Anything less just does not cut the mustard anymore; disabled persons no longer can be excluded from public life, which includes the capacity to understand broadcasts. Audio description will probably remain expensive for some time, and may be expected to lag behind captioning for a decade or more in ubiquity and technological advancement. Those facts do not, however, justify exemptions.

It may be necessary to provide an exemption due to leadtime constraints; it may not be possible to arrange audio description for late-breaking programs. (That's much less true for captioning: There's always the real-time technology.) Still, program producers are just going to have to get used to budgeting time and money for captioning and audio description, processes that can certainly happen in parallel (as long as captioners and describers work from identically-timecoded tapes). If the FCC provides an exemption for leadtime constraints, the exemption must be tightly-worded enough that unscrupulous broadcasters cannot use it as a blanket excuse to

avoid making their programming accessible. Again, industry associations could develop guidelines and assess penalties.

---

17. ***The Impact of Digital Television.*** **We also ask parties to comment on the impact that implementation of Advanced Television ("ATV"), and the use of digital technology, may have on the provision of closed captioning and video description on video programming carried by broadcasters and other program providers. Can this new technology facilitate the transmission of closed captioning or video description? Will it be possible to transmit over a digital signal closed captioning or video description data that is encoded in programming intended for the current analog transmission system?**

I gave a seminal presentation at the Deaf Way conference in Washington in 1989 on the typography of captioning for high-definition television (copy enclosed). I argued that, with only one or two worldwide standards for HDTV likely, we would be remiss in failing to provide for several streams of caption and subtitle text in various fonts and language scripts. (Perhaps downloadable fonts could be countenanced.) It is equally conceivable to provide for multiple audio tracks. Thus a single HDTV program could be broadcast in various countries, providing main audio and video plus captions and audio descriptions in the same language as the main audio, plus various closed-dubbing and -subtitle tracks.

The Commission must ensure that any ATV system be entirely backward-compatible with all EIA-standard Line 21 closed-caption codes from day one of captioning. Further, the Commission needs to require that any ATV system come equipped with a standard secondary audio track à la Second Audio Program on today's MTS transmissions. Further details are hypothetical at this point, though the multi-language, multi-track, multi-font model I have proposed is clearly desirable.

Joe Clark
279a Danforth Avenue
Toronto, Ontario M4K 1N2
(416) 466-0784 *vox*
(416) 406-416-4 *fax*

24 February 1996

The hearing majority of captioning viewers

An analysis

In the initial decade of closed-captioning, something like 300,000 external or set-top decoders were sold in North America. (An exact number is unavailable. The National Captioning Institute in the U.S. manufactured and sold nearly all those decoders but, unaccountably, has always refused to divulge accurate sales figures.) Assume 92% of those decoders – 276,000 – are in the U.S., the rest in Canada.

Let's assume that each of those decoders is watched by one person – a conservative estimate, since many households consist of more than one person. Let's also assume that all those people are deaf or hard-of-hearing – also a conservative assumption, since NCI propaganda stated that the largest single group buying decoders in the late 1980s and early 1990s consisted of people learning English as a second language, nearly all of whom are hearing, and since a small number of hearing people have been watching captioning since its early days.

Since July 1993, all U.S. TVs with screens 13″ or larger have carried decoder chips as standard equipment. U.S. sales of TVs, according to the Electronic Industries Association in Washington, are:

- 23,005,000 in 1993
- An estimated 24,820,000 in 1994
- An estimated 25,600,000 in 1995

Let's assume 90% of those sets are 13″ or larger, and that sales are more or less consistent month-to-month. Thus a reasonable estimate of the total number of decoder-equipped TVs in the U.S. at the end of 1995 is:

$$0.9 \times (0.5 \times 23{,}005{,}000 + 24{,}820{,}000 + 25{,}600{,}000)$$
$$= 55{,}730{,}250$$