COLORADO LAW
UNIVERSITY OF COLORADO BOULDER

**Clinical Legal Education Program**
Robert & Laura Hill Clinical Suite | 404 UCB
Boulder, Colorado 80309-0404
**303.492.8126** | 303.492.4587 fax
www.colorado.edu/law

November 26, 2014

*via electronic filing*

Marlene H. Dortch
Secretary, Office of the Secretary
Federal Communications Commission
445 12th Street, SW, Room TW-A325
Washington, DC 20554

**Re: Closed Caption Quality · CG Docket No. 05-231 · PRM11CG**

Dear Ms. Dortch,

Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI), the National Association of the Deaf (NAD), the Hearing Loss Association of America (HLAA), the Association of Late-Deafened Adults (ALDA), and the Deaf and Hard of Hearing Consumer Advocacy Network (DHHCAN) (collectively, "Consumer Groups"), and the Technology Access Program at Gallaudet University (TAP) write to respond to the *ex parte* filing of VITAC and several other caption providers in the above-referenced docket.[1]

The captioners' filing raises whether the accuracy of real-time captions should be:

1) Scientifically and objectively measured by comparing captions to the *actual* dialogue of the corresponding video program, as the Commission's recently-adopted best practices for captioners require;[2] or
2) Subjectively measured by comparing captions only to those words that the captioners determine that they *should have been able* to caption.[3]

We urge the Commission to stand by its well-reasoned decision to reject the second option in favor of adopting the first, and continue to require real-time captioning vendors following best practices to determine caption accuracy by comparing the words captioned to the underlying program's actual dialogue.[4] The Commission's best-practice metric

---

[1] *See Ex Parte of VITAC, et al.* (corrected version), CG Docket No. 05-231 (Nov. 14, 2014), http://apps.fcc.gov/ecfs/document/view?id=60000984347 ("*Captioner Letter*"); http://apps.fcc.gov/ecfs/document/view?id=60000984271 ("*Captioner Slides*").

[2] *See* 47 C.F.R. § 79.1(k)(2)(iv) (requiring real-time captioning vendors adhering to best practices to "[c]onsider 'accuracy' of captions to be a measurement of the percentage of correct words out of total words in the program, calculated by subtracting number of errors from total number of words in the program, dividing that number by total number of words in the program and converting that number to a percentage").

[3] *See Captioner Letter* at 1-2.

[4] *See Closed Captioning of Video Programming, Telecommunications for the Deaf and Hard of Hearing, Inc. Petition for Rulemaking*, Report and Order, CG Docket No. 05-231, PRM11CG, 29

sensibly and logically treats omissions, substitutions, and other errors in captions as inaccuracies that must be reflected in assessing the accuracy of captions, and is an essential component of the Commission's "overarching goal . . . to ensure that the captioning required by [its] rules fully and effectively conveys the content of [video] programming for people who are deaf and hard of hearing to the same extent that the audio track conveys this content to people who are able to hear."[5]

**The captioners' proposed "grade-inflation" metric would discount substantive errors that deny viewers who are deaf or hard of hearing equal access to video programming.**

The captioners nevertheless urge the Commission to reverse course and allow discounting various captioning omissions and substitutions from the best practices' accuracy metric—essentially adopting a policy of "grade inflation."[6] The captioners contend that this practice should be acceptable because under the proposed "grade-inflation" metric, "omissions [and substitutions] that affect meaning are **always** counted as errors."[7]

We urge the Commission to reject the premise that captions with substitutions can ever meet the goal of "fully and effectively convey[ing] the content of [video programming] for people who are deaf and hard of hearing to the same extent that the audio track conveys this content to people who are able to hear." The content of a video program's dialogue is defined by the *all* the words that are spoken, and by definition it is not possible for captions that omit or substitute words, intentionally or otherwise, to fully convey the content of a program. In the same way that cutting or overdubbing incorrect words into the audio of a program would impact a hearing viewer's ability to understand the program's content, it is obvious and fundamental that omitting or substituting words in captions impairs the ability of a viewer who is deaf or hard of hearing to do the same

Even "repeated words or phrases" and "false starts or stutters"—dismissed by captioners as non-essential—can convey important information about the language patterns, mannerisms, sophistication, tone, confidence, and background of a speaker that is often deliberately included by a program creator—sometimes to establish critical plot

---

FCC Rcd. 2221, 2259-60, ¶ 62 & n.229 (Feb. 24, 2014) ("*Caption Quality Order*") (distinguishing and rejecting captioners' practice of "consider[ing] missing words under completeness rather than accuracy").

[5] *Id.* at 2240, ¶ 26.

[6] *See Captioner Letter* at 1-2.

[7] *See id.* at 2 (emphasis original).

and character points.[8] For example, consider Robert De Niro's famous monologue from Martin Scorsese's *Taxi Driver*:

> You talkin' to me? You talkin' to me? You talkin' to me? Well, who the hell else are you talkin' to? You talkin' to me?[9]

This sequence "remains one of the best known sequences in film history."[10] Under the captioners' heuristics, a single instance of "You talkin' to me?" would have sufficed—but would have deprived a viewer who was deaf or hard of hearing of critical context about the attitude of De Niro's character, Travis Bickle, toward confronting an adversary. Similarly, the inclusion of verbal tics such as "I was," "You know," and "Like" conveys critical information about the sophistication and emotional state of the speaker.[11]

Moreover, the captioners do not specify a bounded set of verifiable and repeatable heuristics for determining what omissions and substitutions actually "affect meaning." Instead, the captioners offer various vague, subjective, and circular qualifications, such as whether an omission or substitution "change[s] the meaning of the spoken word," "deprives [a viewer] of important information," "leads to difficulty in reading the captions," or has "negligible effect on the content."[12]

Assuming for the sake of argument that these qualifications might have been workable, a brief perusal of the captioners' examples of "non-errors" demonstrates that captioners do not consistently follow the qualifications in determining accuracy. For example, the captioners believe substituting a synonymous word for a spoken word—such as "adviser" for "consigliere"—or misspelling a name should not qualify as an error if the spoken word is *not in the captioner's dictionary* or *not provided to the captioner in advance of the program*—even where the substitution plainly affects the content of the captions.[13]

Just because a real-time captioner is unable to properly caption a word does not mean that omitting it or substituting a different word will accurately convey the content. For

---

[8] *But see id.* at 2.

[9] Tim Dirks, *Filmsite Movie Review: Taxi Driver (1976)*, http://www.filmsite.org/taxi.html (last visited Nov. 24, 2014).

[10] *Id.*

[11] *But see Captioner Slides* at 8 (suggesting that dropping these words should not be considered errors).

[12] *Captioner Letter* at 2. The captioners also raise difficulties in calculating accuracy, such as how to treat situations where "people tal[k] over each other," that are not specific to either the best-practice or "grade-inflation" metrics. We welcome discussion of how to address such difficulties, but their existence does not warrant, and would not be mitigated, by changing the accuracy metric.

[13] *See Captioner Letter* at 2-3; *Captioner Slides* at 10.

example, the captioner's qualifications would justify obvious errors in critical lines, such as Rhett Butler in *Gone with the Wind* saying "*Candidly*, my *darling*, I don't give a *darn*" instead of "*Frankly*, my *dear*, I don't give a *damn*."[14] Or, using the captioners' own example, Michael Corleone's infamous line in *The Godfather*, "You're not a wartime *consigliere*, Tom" would hardly be recognizable as "You're not a wartime *adviser*, Tom."[15] Moreover, misspellings of names routinely lead to serious mistakes, such as news reports on "the war in *Iran*" rather than "the war in *Iraq*."

**The "grade-inflation" metric would perpetuate misunderstanding of the serious gap between real-time and offline captioning.**

More broadly, the captioners' proposed change to the Commission's best practice accuracy metric does not appear primarily aimed at permitting negligible omissions and substitutions that do not affect viewers who are deaf or hard of hearing—nor do such omissions and substitutions even exist. Rather, the proposed change appears primarily aimed at discounting errors that are inevitable from the *captioner's* perspective—in essence, advertising real-time captioning as though its shortcomings do not exist.

Indeed, the captioners' own statistics demonstrate that their proposal would serve to afford a substantial "grade inflation" for the accuracy metric. The captioners show that under the Commission's best practice metric, captioners tackling several difficult examples could only achieve an average accuracy of 90.66%—a rating that dips as low as 79.98%, meaning that *more than one in every five words is missing or wrong.*[16] Yet under the captioners' "grade-inflation" metric, the captioners achieved an average of 99.21%, an increase of nearly 10%, and never any lower than 98.68%.[17] For the most difficult example, in which more than 20% of the words were missing or wrong, the "grade-inflation" metric attributes an astounding 99.24% accuracy—an implausible rating that demonstrates that the underlying methodology is unsound, unscientific, and subjective.[18]

We understand and agree with captioners, of course, that present technology does not facilitate perfectly accurate real-time captioning.[19] But the fact that captioners cannot yet perfectly caption content in real-time captions does not justify pretending that real-time

---

[14] *See* Wikipedia, *Frankly, my dear, I don't give a damn*, http://en.wikipedia.org/wiki/Frankly,_my_dear,_I_don't_give_a_damn (last visited Nov. 25, 2014).

[15] Tim Dirks, *Filmsite Movie Review: The Godfather*, http://www.filmsite.org/godfC.html (last visited Nov. 25, 2014).

[16] *See Captioner Slides* at 5.

[17] *See id.*

[18] *See id.*

[19] *See Captioner Letter* at 2 ("[I]t is simply impossible to write verbatim [real-time] captions.").

captions are nearly perfect. Indeed, no one seriously disputes that it is possible to create 100% accurate offline captions using the Commission's best-practice metric, and that metric correctly reflects the significant gap between real-time and offline captions that is experienced every day by the millions of Americans who rely on closed captions to watch video programming.[20]

Declaring the gap between real-time and offline captions essentially non-existent because it is difficult to create perfect real-time captions would accomplish little more than ignoring a very real problem. Moreover, doing so would undercut the Commission's ability to make meaningful policy decisions on critical issues, such as the circumstances under which real-time captions are permissible. Doing so would also hurt the ability of captioners to differentiate their own services by pushing the envelope of accuracy; when poor-quality captions are rated "99% accurate" instead of "79% accurate," there is little incentive to make improvements.

**The Commission's best-practice "grade-inflation" metric will not impose any serious long-term burden on captioners.**

Moreover, as the captioners acknowledge, the Commission's best-practice accuracy metric does not require captioners to achieve any particular level of accuracy; it merely requires truthful disclosure.[21] Captioners nevertheless contend that the Commission should abandon the metric because programmers are insisting on levels of accuracy that are unattainable under the metric.[22]

In the long term, this concern is baseless because programmers demanding truly unattainable levels of accuracy will find no willing and able caption vendors to serve their needs. Because the Commission's rules require captioning, those programmers will have little choice but to ratchet down their demands to match what the market can supply.[23]

Captioners also contend that complying with the Commission's best-practice accuracy metric will "more than triple the amount of time necessary to review caption files and

---

[20] *See* 47 C.F.R. 79.1(k)(4)(i)-(iii) (requiring offiline captioning vendors to "ensure offline captions are verbatim[,] error-free[, and] punctuated correctly").

[21] *See Captioner Letter* at 1.

[22] *Id.* at 2.

[23] In the short term, we understand that some captioners may have entered contracts that require them to meet certain a certain level of accuracy that is unattainable under the best practice accuracy metric. It is not clear to us, nor does the captioners' *ex parte* filing contain any significant evidence of, the extent to which these contracts actually exist, how they interact with the best practices, or whether they are of sufficient length to warrant Commission intervention. We would object to any temporary deferral or waiver of the rules or other accommodation not supported by specific, verifiable evidence.

require additional manpower at a cost that is not supported by the current captioning market."[24] According to captioners, this is the case because the best-practice metric requires creating a baseline verbatim transcript to which the captions can be compared.[25]

To whatever extent the captioners' preferred method of computing accuracy—"the reviewer simply records errors"—takes less time and costs less, it does so at the untenable expense of meaningfully measuring caption accuracy. In particular, it affords reviewers the discretion to ignore a substantial portion of errors—on the order of 20% in some cases.[26] Whatever time and cost savings flow from following this methodology are false and irrelevant because the methodology does not legitimately measure the accuracy of the captions. Moreover, the best practices do not require every program to be evaluated, but merely "frequent and regular evaluations and sample audits"; it is unclear why the costs of evaluation could not simply be incorporated into the evaluation schedule and passed on to programmers purchasing captioning services.[27]

Finally, the captioners suggest that measuring accuracy under the best-practice metric will distract captioners with "the consequences of not writing every word" and lead to more inaccurate and misspelled words.[28] We disagree, because captioners will have an equal disincentive not to misspell or use incorrect words, actions that are still errors under the best-practice metric. At most, the captioners will simply have to internalize that both omitted and incorrect words will lower their accuracy ratings.

**Revising the best practices without a full notice-and-comment rulemaking would be procedurally improper.**

Finally, we understand that the captioners intend to seek a declaratory ruling that the Commission's best-practice metric in fact incorporates the captioners' reformulated metric. While the captioners have not spelled out specific arguments outlining how the Commission might do so, even a cursory reading of the best-practice metric makes clear that it is not compatible with the captioners' approach.

More specifically, the best-practice metric plainly requires accuracy to be calculated by reference to the "total number of words in the program," rather than the number of

---

[24] *Captioner Letter* at 2; *see also Captioner Slides* at 5-6 (noting that measuring accuracy using the best-practice metric took 3.32 times longer on average than with the "grade-inflation" metric).

[25] *Captioner Letter* at 2.

[26] *See* discussion *supra*.

[27] *See* 47 C.F.R. § 79.1(k)(2)(iii).

[28] *Captioner Slides* at 12.

words captioned, which the Commission expressly confirmed in adopting the metric.[29] The best-practice approach also explicitly requires the captioners to treat all "incorrect" and "missing" words as errors, an approach plainly incompatible with the "grade inflation" metric's approach of disregarding of certain incorrect and missing words.[30]

Because there is no way to reconcile the plain language and obvious intent of the best-practice metric with the "grade-inflation" metric, the Commission has no legal basis upon which to declare that the two metrics are compatible. Accordingly, the Commission could only adopt the "grade-inflation" metric on a long-term basis by launching a new notice-and-comment rulemaking and upending its well-reasoned conclusion in the *Caption Quality Order* adopting the best-practice metric and rejecting the "grade-inflation" metric.[31] We strongly discourage the Commission from undertaking such an ill-advised, unjustified, and time-consuming endeavor, or from taking any temporary steps that would pave the way for doing so.

* * *

We appreciate the captioners' efforts to engage seriously with the Commission's best practices, but cannot endorse the misguided "grade-inflation" approach and urge the Commission to reject it. Nevertheless, we remain open to further discussion and look forward to meeting with the Commission and captioners in the coming weeks.

Please don't hesitate to contact me if you have any questions regarding this filing.

Respectfully submitted,

/s/

Blake E. Reid

*Counsel to TDI*

blake.reid@colorado.edu
303.492.0548

---

[29] *See* 47 C.F.R. § 79.1(k)(2)(iv); *Caption Quality Order*, 29 FCC Rcd. at 2259-60, ¶ 62 & n.229.

[30] *See* 47 C.F.R. § 79.1(k)(2)(v) (treating "at a minimum, mistranslated words, incorrect words, [and] misspelled words . . . as errors."); *Captioner Letter* at 1-2 (conceding that the "primary difference" between the best-practice metric and the "grade-inflation" metric is "the way 'errors' are defined).

[31] *See Caption Quality Order*, 29 FCC Rcd. at 2259-60, ¶ 62 & n.229.

**Telecommunications for the Deaf and Hard of Hearing, Inc. (TDI)**
*Contact:* Claude Stout, Executive Director • cstout@TDIforAccess.org
8630 Fenton Street, Suite 121, Silver Spring, MD 20910
www.TDIforAccess.org

**National Association of the Deaf (NAD)**
Howard Rosenblum, Chief Executive Officer • howard.rosenblum@nad.org
*Contact:* Andrew Phillips, Policy Counsel • andrew.phillips@nad.org
8630 Fenton Street, Suite 820, Silver Spring, MD 20910
301.587.1788
www.nad.org

**Hearing Loss Association of America (HLAA)**
Anna Gilmore Hall, Executive Director • AGilmoreHall@Hearingloss.org
*Contact:* Lise Hamlin, Director of Public Policy, LHamlin@Hearingloss.org
7910 Woodmont Avenue, Suite 1200, Bethesda, MD 20814
301.657.2248
www.hearingloss.org

**Association of Late-Deafened Adults (ALDA)**
Dave Litman, President • aldaprez2014@gmail.com
*Contact:* Brenda Estes • bestes@endependence.org
8038 Macintosh Lane, Suite 2, Rockford, IL 61107
www.alda.org

**Deaf and Hard of Hearing Consumer Advocacy Network (DHHCAN)**
Cheryl Heppner, Vice Chair • CHeppner@nvrc.org
3951 Pender Drive, Suite 130, Fairfax, VA 22030

**Technology Access Program at Gallaudet University (TAP)**
*Contact*: Christian Vogler, Ph.D., Director • christian.vogler@gallaudet.edu
Department of Communications Studies
SLCC 1116, Gallaudet University
800 Florida Avenue NE, Washington, DC 20002
202.250.2795
tap.gallaudet.edu

CC: Karen Peltz Strauss, CGB
Greg Hlibok, DRO
Eliot Greenwald, DRO
Suzy Rosen Singleton, DRO
Caitlin Vogus, DRO
Diana Sokolow, MB
Susan Aaron, OGC
Heather York, VITAC
Jill Toschi and Darlene Parker, National Captioning Institute (NCI)
Adam Finkel, National Court Reporters Association (NCRA)