

November 15, 2019

VIA ELECTRONIC FILING (ECFS)

Marlene H. Dortch, Esq.
Secretary
Federal Communications Commission
445 Twelfth Street, SW
Washington, DC 20554

RE: **EX PARTE PRESENTATION**
*Misuse of Internet Protocol (IP) Captioned Telephone Service;
Telecommunications Relay Services and Speech-to-Speech Services for
Individuals with Hearing and Speech Disabilities*
CG Docket Nos. 13-24, 03-123

Dear Ms. Dortch:

On November 13, 2019, representatives of Hamilton Relay, Inc. (“Hamilton”) and CapTel Inc./Ultratec Inc. (“Ultratec”) met with Commission staff from the Consumer & Governmental Affairs Bureau (“CGB”), Office of Economics and Analytics (“OEA”), and the Office of the Managing Director (“OMD”). Participating on behalf of Hamilton were Dixie Ziegler, Beth Slough, and Jeff Knighton as well as Rachel Wolkowitz (outside counsel) and the undersigned counsel. Participating on behalf of Ultratec was Christopher Engelke. Commission staff present at the meeting were Diane Burstein, Robert Aldrich, Eliot Greenwald, Michael Scott of CGB; Virginia Metallo (by telephone) and Susan Lee (by telephone) of OEA; and David Schmidt (by telephone) of OMD.

During the meeting, the parties shared a supplemental report on trials Hamilton and Ultratec have been conducting examining Automatic Speech Recognition-only (“ASR-only”) and Communications Assistant (“CA”) performance in the context of Internet Protocol Captioned Telephone Service (“IP CTS”). Specifically, the new research demonstrates that caption accuracy “remains the only significant predictor of functional equivalence.”¹

Further, Hamilton and Ultratec reported that a trial testing equipment which included a CA/ASR button permitting users to switch between ASR-only and a CA was inconclusive.²

¹ See attached slide deck, A Trial of Automated Speech Recognition for IP CTS Calls: Supplemental Report, at 5 (Nov. 13, 2019).

² *Id.* at 9.

Federal Communications Commission
November 15, 2019
Page 2

During this trial, volunteers were given the option of switching between a CA and an ASR-only captioning source. The study showed that users switched between 0-130 times over the week, and between 0-28 times per call, and that users also reported being confused or distracted by switch operation when on calls.³ Hamilton and Ultratec believe that additional research is needed in this area.

This filing is made in accordance with Section 1.1206(b)(1) of the Commission's rules, 47 C.F.R. § 1.1206(b)(1). In the event that there are any questions concerning this matter, please contact the undersigned.

Respectfully submitted,

WILKINSON BARKER KNAUER, LLP

/s/ David A. O'Connor

Counsel for Hamilton Relay, Inc.

Enclosure

cc (via email):

Diane Burstein
Eliot Greenwald
Michael Scott
Robert Aldrich

Virginia Metallo
Susan Lee
David Schmidt

³ *Id.*

A Trial of Automated Speech Recognition for IP CTS Calls

Supplemental Report

Nov. 13, 2019



Copyright 2019 CapTel, Inc. Madison, WI



Team Participants

Christopher Engelke

Vice President, Ultratec, Inc.

Dixie Ziegler

Vice President, Hamilton Relay, Inc.

Beth Slough

Compliance Manager, Hamilton Relay, Inc.

Jeff Knighton

President, Hamilton Innovations, LLC.

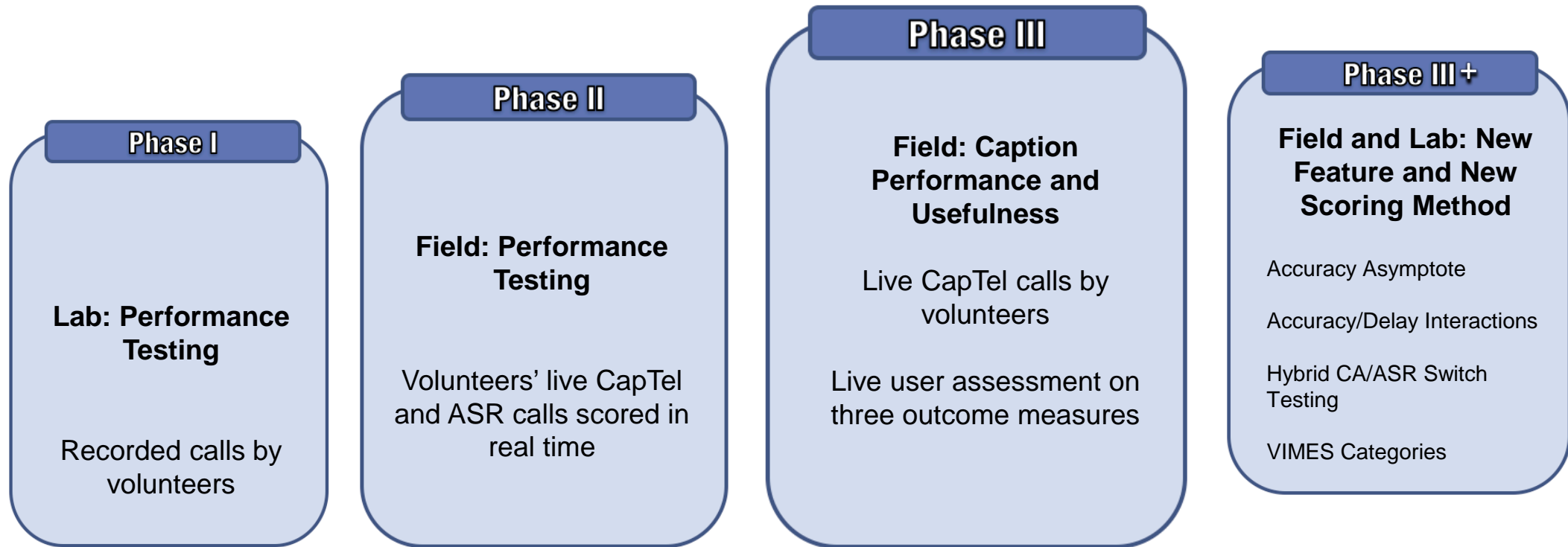
David O'Connor

Outside Counsel to Hamilton

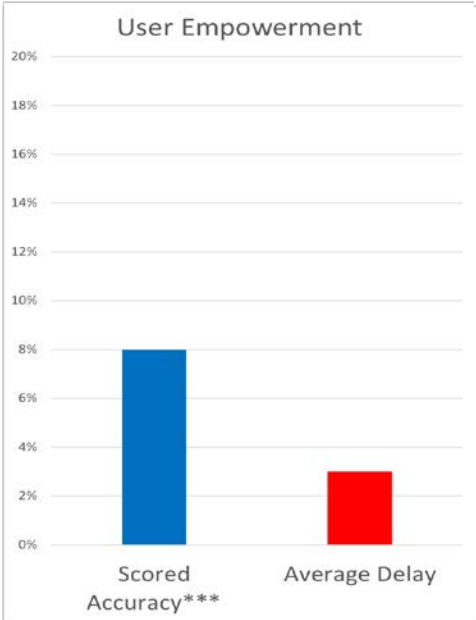
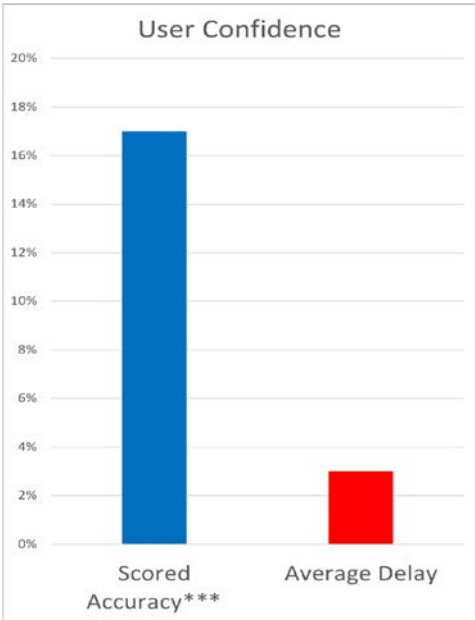
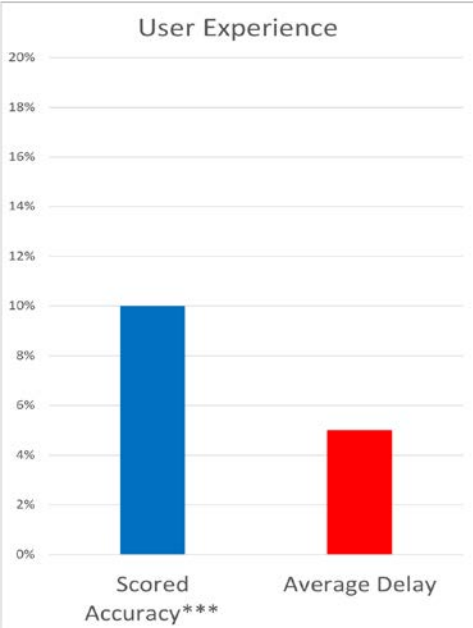
Rachel Wolkowitz

Outside Counsel to Hamilton

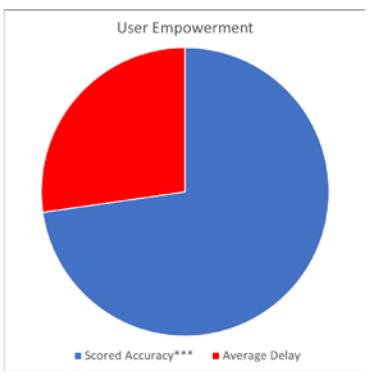
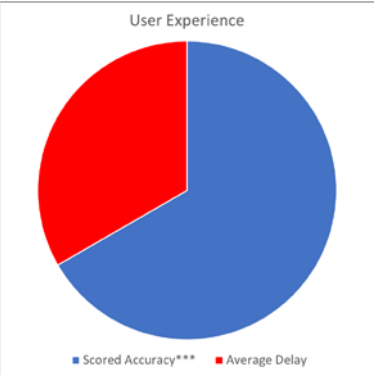
Three+ Phase Trial of ASR in IP CTS



Relative Impact of Scored Accuracy and Delay*



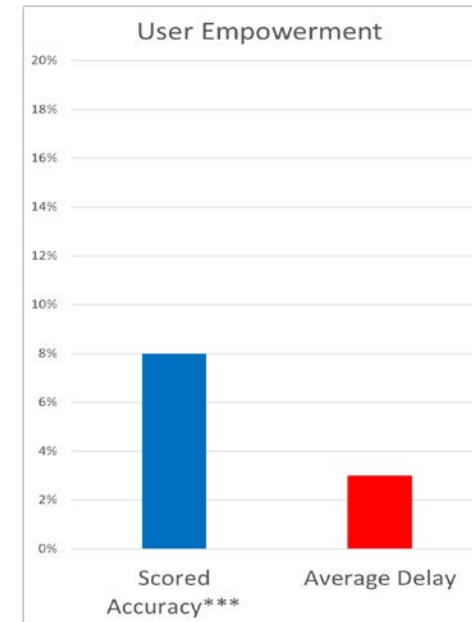
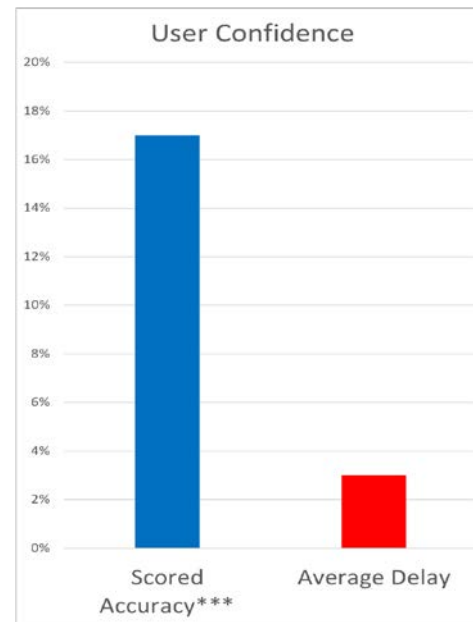
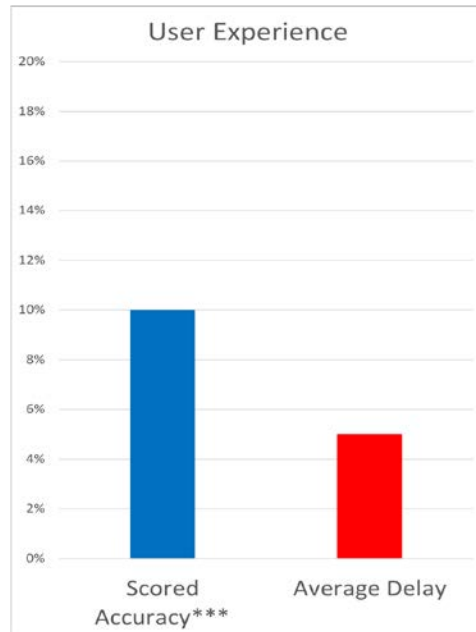
Relative Percentage of Variance Explained



*Comparison of Standardized Regression Coefficients (Predictive Power)



Phase III+: Interaction Effects of Accuracy and Delay



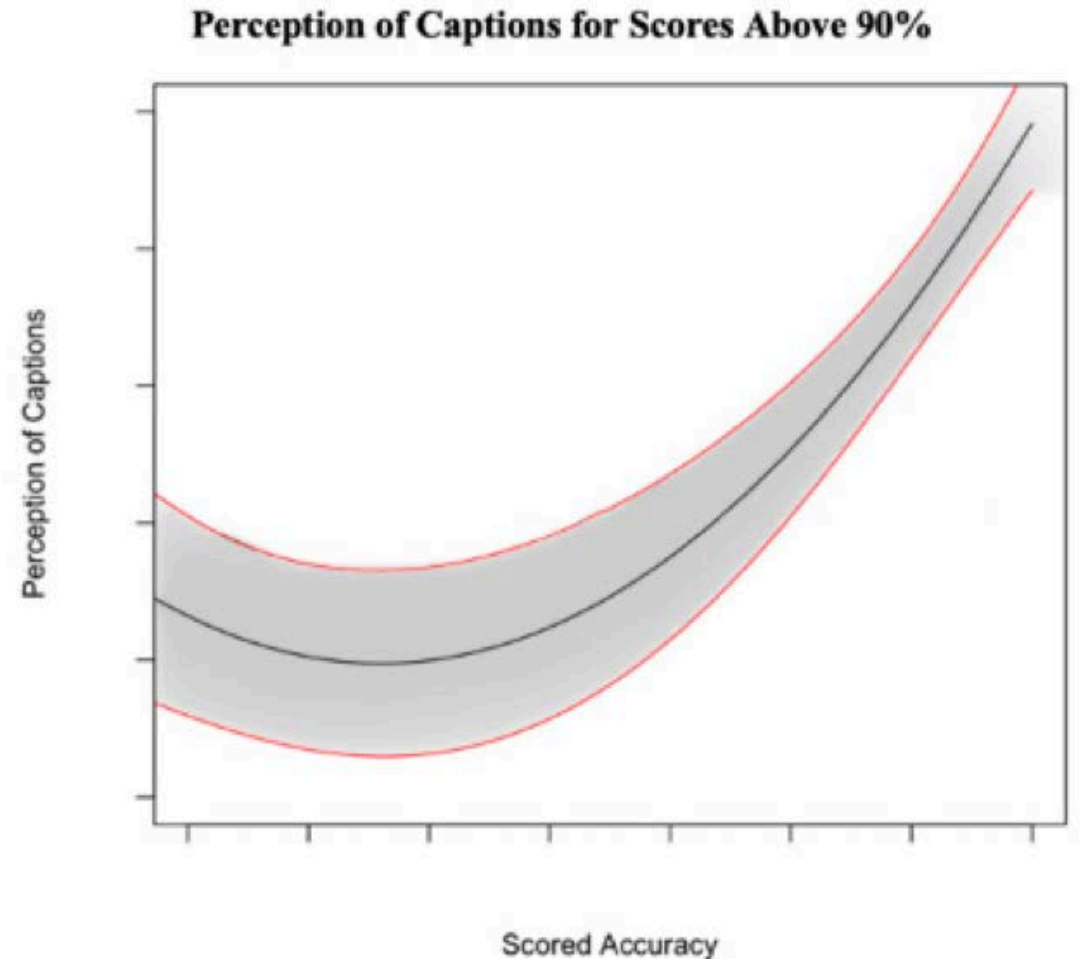
- We observed no significant differences in importance of accuracy at different levels of delay for any of the three domains.
- Caption accuracy remains the only significant predictor of functional equivalence.

Predictive Power of Scored Accuracy and Delay on Functional Equivalence

- Graphics show the predictive power of scored accuracy and delay on volunteers responses
- Accuracy was significantly correlated to all post-call response items ($p < .001$)
- Delay **did not** significantly correlate to any post-call responses (no better than random)
- Accuracy was 2x more powerful than delay in predicting User Experience
- Accuracy was more than 5x more powerful than delay in predicting User Confidence
- Accuracy was more than 2x more powerful than delay in predicting User Empowerment

Phase III+: Accuracy Asymptote

- Scored accuracy significantly predicted user experience across all three domains examined.
- The relationship between scored accuracy and user experience was relatively linear. No “leveling off” or asymptote was found.
- In the domain of “Confidence,” users did experience an inflection point, above which **the importance of accuracy intervals increased** compared to lower scores.
 - Different scoring methods are not interchangeable or transposable.
 - Because of the differences between how different scoring methods measure accuracy, inflection points will be unique to each scoring method.



Phase III+: CA/ASR Switch Button Research



- 40 volunteer participants, 1,055 calls
- Users started calls in either CA or ASR mode
- Volunteers were given the option of switching between CA and ASR caption source (CA and ASR were both live throughout call)
- Volunteers were trained in person on switch operation
- Caption source was hidden from user to prevent bias
- Volunteers were paid per questions answered (up to \$100/wk)
- Calls not reimbursed
- Accuracy, Delay, and Switch use measured in real time (no call content recorded)

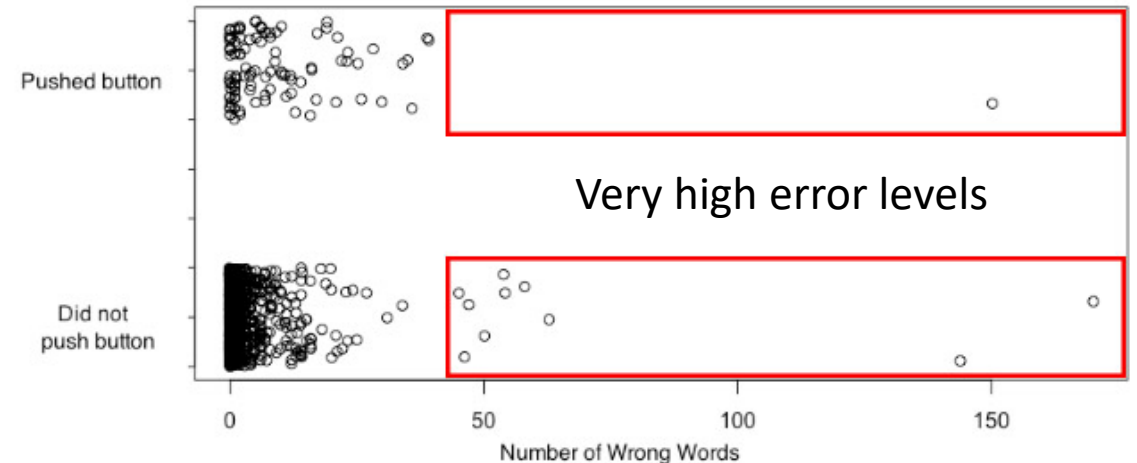


- **Double-Blind Design**



Phase III+: CA/ASR Switch Button Research

- Results from Switch Option testing were inconclusive.
- Users switched between 0-130 times over the week, and between 0-28 times per call.
- Neither Scored Accuracy nor Delay were significant predictors of switch activation.
- At very high levels of errors ($>\sim 50$), most users did not switch modes.
- Users reported being confused or distracted by switch operation when on calls.
- The more users switched, the more they felt the option greatly hurt their experience on the call and the less helpful they found the option.
- CapTel and Hamilton are continuing to test this option and will share details as they are available.



Phase III+: Scoring Methods and Impacts

- Based on Phase III findings, CapTel and Hamilton began investigating the different scoring methods in order to better represent user experience.
- This research includes both qualitative and quantitative research into:
 - Differences in how error types disrupt conversational flow and structure
 - Patterns of responsibility and accountability for error types
 - Post-call impact of error types
 - Users' perceptions of accuracy during a captioned telephone call
- CapTel has identified “V.I.M.E.S” categories that better capture user experience and is evolving this method.



Phase III+: Error Categories

- **Visible Errors** – Errors that cause confusion or reduce the grammatical or conversational flow of the captioned utterance.
- **Invisible Errors** – Errors that maintain the grammatical or conversational flow but change the meaning of the captioned utterance.
- **Minor Errors** – Errors that do not change the meaning or valiance of an captioned utterance.
- **Essential Errors** – Errors of material information that preserve the typical or identifiable form of the information (e.g. errors in phone numbers).
- **Severe Errors** – An invisible error that creates the potential for harm if not corrected.



Thank You
