Before the

Federal Communications Commission

Washington, D.C. 20554

In the Matter of Application of New and Emerging Technologies for Video Relay Service Use	)))))))	CG Docket No. 10-51

Comments of the Rehabilitation Engineering Research Center on Telecommunications Access

Christian Vogler, Ph.D.,
Co-Principal Investigator,
RERC on Telecommunications Access,
Director, Technology Access Program
and Norman Williams,
Technology Access Program
Gallaudet University
800 Florida Ave., NE, SLCC 1116
Washington, DC 20002
(202) 250-2795

and

Gregg C. Vanderheiden, Ph.D.,
Co-Principal Investigator,
RERC on Telecommunications Access,
Director, Trace R&D Center
University of Wisconsin-Madison
1550 Engineering Drive, 2107 ECB
Madison, WI 53706-1609
(608) 262-6966

Summary

The RERC-TA supports the goal of making off-the-shelf videophone equipment interoperable with VRS, and VRS-provided equipment. Not only would achieving this goal make innovations in a huge market available to VRS users, but more importantly, it also would lead to a huge improvement in functional equivalence. Deaf and hearing users would be able to call one another, from anywhere, at any time, with any videophone and voice phone equipment, respectively. Currently, we are nowhere near this ideal, because technological and policy barriers stand in the way of interoperability.

It needs to be stressed that these barriers can be overcome. With respect to off-the-shelf to VRS interoperability, they pertain to camera quality, access to 10-digit numbering, meeting TRS registration requirements, passing voice numbers through to VRS, potential problems with managing E911 and ANI via VRS, bypassing firewalls, and incoming call alerts.

RERC-TA proposes that transitioning VRS away from H.323 to SIP could do much to improve the interoperability of equipment, since SIP provides standardized mechanisms that could be exploited for overcoming most of the barriers. SIP also provides interoperable mechanisms for other desirable features, including call waiting and real-time text channels. In addition, RERC-TA recommends that off-the-shelf equipment should be allowed to make point-to-point calls to registered VRS users via their 10-digit numbers. Technologies will need to be developed to bridge off-the-shelf equipment with call alert systems suitable for deaf people, and minimum requirements for cameras should be set, so as to ensure that they support natural sign language

conversations. In a similar vein, minimum requirements need to be set for functional parameters, including frame rate, bandwidth, latency, data volume caps, and light levels.

There is a window of opportunity now to effect the changes required for integrating VRS technologies with mainstream off-the-shelf technologies, and to improve functional equivalence between deaf and hearing telephone users. These changes need to happen before deaf consumers invest in a new generation of expensive videophones that would pose backward compatibility problems for a long time to come.

Table of contents

I. Introduction

The Telecom RERC (RERC-TA) is a joint project of the Technology Access Program at Gallaudet University and the Trace Center at the University of Wisconsin-Madison. The RERC is funded by the U.S. Department of Education, National Institute on Disability and Rehabilitation Research, to carry out a program of research and development focused on technological solutions for universal access to telecommunications systems and products for people with disabilities.

The FCC's Public Notice on the application of new and emerging technologies for video relay service asks a set of questions on whether and how off-the-shelf video technologies and emerging video call technologies can be used with VRS. These questions cover the angles of what feature sets are required in videophone equipment, and what the technical prerequisites are for ensuring high-quality transmissions. They provide an excellent opportunity to examine where we currently stand with respect to interoperable videophone and video relay services, and to develop proposals for solving current interoperability problems and maintaining quality standards for video calls.

There is little doubt that we are entering a transition period in the area of video relay services. On the one hand, the era of obtaining free equipment for video relay calls is drawing to a close, and VRS-specific phones with mutually incompatible feature sets have proliferated. On the other hand, off-the-shelf options for videophones are multiplying, and some of their characteristics are becoming increasingly attractive for the purposes of video relay and point-to-point calls.

Some of the advantages of off-the-shelf equipment are:

- It is easy to propagate innovative features and advances in video compression technology, due to quick product cycles and larger potential manufacturer and user bases.
- It has integration of conferencing capabilities.
- It has potentially easier integration of split video with PSAPs and VRS interpreters, in NG 911 services, due to the previous point.
- It is widely available in the hearing world, so VRS users can easily borrow equipment on-site during travel or business calls.
- In emergencies, people can use any videophone nearby – wherever they are.
- For the same feature set, it is cheaper than VRS-specific equipment, due to volume savings.
- It is easier to integrate into organization or company-wide setups, and plays better with corporate IT policies, because no special setup would be required for VRS users that differs from the rest of the organization's equipment.

Disadvantages of off-the-shelf equipment at this time are:

- It does not provide some of the needed features for effective VRS use.
- It has poor interoperability with the existing deaf videophone base.
- It has no certification or labeling process in place for ensuring that the equipment meets the TRS rules[1].
- Some products may not meet the video and camera quality needs of VRS users.

[1] 47 C.F.R. §64.601 and following

Because we are in a transition period with respect to video relay services, there now is a window of opportunity to close the gap between VRS-specific equipment and off-the-shelf equipment, before VRS users start investing in new videophones. If the transition is handled correctly, functional equivalence can be greatly improved.

II. Functional equivalence as the overarching goal

The principle of functional equivalence between deaf and hearing people is central to the advances that have been made with respect to ensuring that deaf people can access mainstream telecommunications, since the passage of the Americans with Disabilities Act, and the set-up of relay services. Video relay services have contributed further to this goal by virtually eliminating communication delays in relaying deaf users' messages, and allowing bidirectional communication without the unnatural rigid turn taking inherent to TTY communication. There also exist some off-the-shelf products that function well for point-to-point signed communication; for example, when both users have compatible systems that produce excellent video quality.

Despite these advances, we are still far from the point where we can consider deaf and hearing people to be truly functionally equivalent in how they can use and access telecommunications. For example, a hearing person can pick up a phone, anywhere, anytime, dial a number, and complete the call or get called back at that number. A deaf person (or a hearing sign language user calling a deaf person) is unable to do the same, due to incompatibilities in videophone equipment, incompatibilities across video relay services, and barriers to setting up point-to-point calls (off-the-shelf equipment is not integrated into the TRS Numbering Directory).

RERC-TA maintains that integrating off-the-shelf equipment with VRS and point-to-point calls is a commendable undertaking, and is key to providing functional equivalence, as long as it is driven by the motivation to standardize VRS and point-to-point communications. Under ideal conditions, this could spur innovation and growth in the videophone market, and help lower the cost of the equipment required for VRS calls. The latter point is particularly important for deaf consumers in the low income segment, who otherwise would not be able to afford specialized VRS equipment. Unfortunately, as things currently stand, off-the-shelf equipment is not even close to interoperating with competitors' video communication products, and with the custom VRS-provided equipment and software that is in widespread use among the deaf population. In particular, the following unresolved technical and policy issues stand in the way (details are provided in the following section):

- The camera quality of much of the current off-the-shelf equipment is insufficient for clear and natural signing, especially under low lighting levels found in many home environments.
- Off-the-shelf equipment and non-VRS providers do not have access to the TRS Numbering Directory[2].
- There is no process for verification of eligibility to use VRS with off-the-shelf equipment, and meeting TRS registration requirements.
- There are no standardized provisions for passing voice numbers to VRS in the absence of integration with the numbering system.
- There are currently no provisions for passing on E911 information via VRS, and ANI callbacks, with off-the-shelf equipment.

[2] 47 C.F.R §64.613 (3)

- Firewalls currently block many types of VRS-related video calls, particularly incoming calls.
- Off-the-shelf equipment is incompatible with the systems that deaf consumers use for incoming call alerts.

Each of these is solvable through technical means or policy changes, or may be solved by normal advances in communication technology. But they need to be addressed before a switch to mainstream equipment is effected.

III. Features and functional parameters required for effectively using VRS

III.A. Details on technical and policy issues that need to be resolved

III.A.1. Camera quality

Good camera performance in a variety of lighting situations, including relatively low lighting conditions, is essential. For home and office use, cameras need to provide acceptable frame rates and exposure times at the typical ambient lighting levels in these settings. An added requirement for home use is that the camera must perform well even when no ceiling lights are available or turned on.

For clear communication, the maximum exposure time per frame has to be 40 milliseconds or less (with 20 ms being the optimum) at lighting levels as low as 30 lux[3], or else motion blur prevents the users from discerning the details of the handshape in signs. As the exposure time increases, it also becomes increasingly difficult to discern hand movements at normal signing speeds.

[3] cf. III.B.6

Many off-the-shelf cameras (such as the ones in laptops and videophones, as well as external webcams) do not currently meet the exposure time targets. Moreover, to date, there are large differences in quality under low lighting conditions even among the custom equipment provided by various video relay services.

There are some mainstream cameras that exhibit excellent performance with VRS calls, but finding them at the store is a hit-or-miss proposition for deaf consumers. An in-store environment is not suitable for testing whether a particular device meets a consumer's requirements, because lighting tends to be substantially brighter than what is encountered in a home environment. At the same time, store return policies frequently prevent VRS users from returning a device without a restocking fee if its camera turns out to be unsuitable, causing them to spend money unnecessarily. Possible solutions to this problem include providing lists of cameras that meet the requirements, so that users can make informed buying decisions, or providing a "VRS-approved" sticker that companies could put on the product packaging (voluntarily) for specific makes and models that pass the tests (preferably of an independent institution – especially for small brands).

III.A.2. No access to TRS Numbering Directory

Currently, off-the-shelf equipment by non-VRS providers does not have access to the TRS Numbering Directory. This restriction makes it impossible for one user's off-the-shelf equipment to determine the IP address of another registered Internet-based TRS user via that person's NANP telephone number. Consequently, callers with off-the-shelf equipment currently cannot use the ten-digit number of the receiving party to establish a point-to-point connection.

Instead, they have to go through the cumbersome process of having the receiving party determine its IP address in advance and transmit it to the calling party. This is a nontrivial task for nontechnical users, and furthermore, dynamically assigned IP addresses (which constitute the majority for home users) are subject to frequent changes.

Importantly, these restrictions also prevent hearing people who are fluent in ASL from calling deaf people directly on their videophones via their ten-digit number, and vice versa. As a result, the most convenient way to establish a connection between a deaf and a hearing party is through VRS, even when they could easily communicate directly in a point-to-point call, thereby wasting VRS interpreter minutes[4] and unnecessarily depleting the VRS fund.

Changes to the TRS numbering system policies are needed to maintain functional equivalence goals, and at the same time allow integration of off-the-shelf equipment. In particular, off-the-shelf equipment registered with mainstream operators needs to be able to establish point-to-point connections to users registered with the TRS system, and vice versa.

III.A.3. Verification of VRS eligibility and meeting TRS registration requirements

As per the registration requirements set forth in 47 C.F.R. §64.611 (a) and (b), all VRS users must register with a VRS service as the default provider before they can begin using VRS in the first place, which enables VRS providers to update the TRS numbering

[4] Telecommunications for the Deaf and Hard of Hearing, Inc. et al. (Aug 18, 2010). Comments in Response to Notice of Inquiry, In the Matter of Structure and Practices of the Video Relay Service, CG Docket No. 10-51. (Available at http://fjallfoss.fcc.gov/ecfs/document/view?id=7020708952)

directory and E911 information. Off-the-shelf equipment does not provide the necessary mechanisms for supporting the registration requirement, and other mechanisms for doing this are not provided. Some VRS providers currently allow callers to sign or fingerspell the registration information and let them proceed with an outgoing VRS call, even if they use off-the-shelf equipment for making that particular call. Although going through this process is acceptable for one-off calls when the caller borrows someone else's phone, it is too cumbersome and time-consuming for making regular calls from a workplace phone. Moreover, some VRS providers refuse service outright if off-the-shelf equipment is used.

Therefore, current practice fails to meet the functional equivalence criteria in two ways: First, deaf people, who have to use off-the-shelf equipment in the workplace as part of organization-wide policies, have to take additional steps to complete VRS calls. Second, if a deaf user borrows or uses someone else's off-the-shelf equipment for making a quick VRS call – similar to the way hearing people use someone else's phone for calls during site visits, business meetings, and so on – there are no assurances that the VRS provider will allow the call to go through.

For functional equivalence, there must be a standard way to register videophones, and to verify registration, as per 47 C.F.R. §64.611 (b), without additional user involvement. One possible solution would be to complete and verify registration based on an authenticated SIP identifier. In addition, to meet functional equivalence for one-off calls on borrowed equipment, there should be a requirement that all VRS providers must accept calls from off-the-shelf equipment, as long as the user has registered with a VRS provider before, and can sign or fingerspell the relevant information to the CA for verification.

III.A.4. Passing voice numbers to VRS

Many videophones provided by VRS offer the option of automatically passing the voice phone number that the originating party wishes to call. This mechanism greatly simplifies the calling process, as the user can simply pick the number from the videophone's directory or use a numeric keypad to type it in.

Because this mechanism is VRS-specific, it currently does not work with generic off-the-shelf equipment. In this case, the user needs to memorize or make a note of the number, and fingerspell it to the VRS interpreter, in order to proceed with the call. For regular calls, this can be considered an inconvenience, but for emergency calls it poses potential problems, for two reasons: First, valuable time is spent on communicating with the interpreter before the call can proceed; and second, the provider has no indication that a call to 9-1-1 is intended, so it cannot be prioritized and routed to the first available interpreter in the call queue irrespective of the order in which the pending calls arrived.

If a VRS provider allows connecting via SIP, one possible solution to the problem of passing voice numbers is to use addresses in the form of 2025551212@abcvrs.com, where the first part constitutes the voice number to call, and the second part constitutes the provider's domain. Although not quite as convenient as allowing simple voice numbers, this would provide a nearly functionally equivalent mechanism that works with both VRS-specific and generic videophones. For ease of use, this mechanism should be integrated with off-the-shelf phonebooks, such that the end user needs to enter only the ten-digit number, rather than the complete SIP identifier.

Another approach would be to allow the information to be passed through to the VRS provider using standard telephone touch-tones. Both physical and softphone provide the ability to transmit touch tones to allow control of IVRs, voicemail, and so on. The touch-tones could be captured, decoded and used to automatically make the call to the destination number to complete the call. For a user's own phone, the VRS number can be programmed in a speed dial to make the step easy. The second number could also be programmed – with the phone switching to touch-tones when dialing is done with a phone connection already in place.

With VoIP calls there should also be the ability to call the relay and 3rd party simultaneously. This eliminates the need for delays at VRS re-calling the 3rd party but would introduce problems when the 3rd party answers before the VRS operator. So calling the third party would have to wait until the VRS operator answered. For 911 calls, however, – where there is a priority answer by VRS – this might be an option and also allow the 911 center to see the video directly.

III.A.5. Emergency services

The local 10-digit numbering system for registered TRS users is not currently supported by off-the-shelf equipment, and there must also be a mechanism for registering the location information, as per 47 C.F.R. §64.605 (b)(4)(ii), which states that there must be at least one way to provide this information using only CPE. Since unimpeded access to emergency services is mandatory, features that support VRS calls to 9-1-1 are absolutely essential in off-the-shelf equipment.

In particular, there must be reliable standardized ways to register and transmit the E911 location information, and support callbacks from the PSAPs through VRS back to the caller via ANI. Such callbacks are especially important when a 9-1-1 VRS call is made through a cell phone, and the call drops.

If VoIP videophones are configured to dial 911 and a VRS in parallel, the normal VoIP location information would be available directly to 911 centers (as would a direct video feed) without special pass-through by the VRS provider. Doing parallel video calls, however, would require double the bandwidth to and from the caller – which may be a problem in the short run, but not likely in the long run.

Whatever the mechanism ends up to be, videophones, including the ones for VRS calls, should use standard call mechanisms as much as possible to prevent a failure in the mechanism to go undetected until an emergency call is made.

III.A.6. Bypassing firewalls and bridging.

Firewalls and network address translation (NAT) pose major obstacles for VRS and point-to-point communications. The H.323 protocol suite requires having a range of TCP and UDP ports open for handling calls, or video may end up missing on one or both endpoints. This requirement runs counter to the network security policies of many large organizations, and is a particular concern for off-the-shelf equipment and services. The reason is that these same organizations also are particularly likely to have settled on a standardized single off-the-shelf solution for all their communication needs. Some VRS provide their own servers to route calls properly, even when the user's equipment is behind a firewall, but this solution is not guaranteed to work with generic videophones.

However, this problem would not exist for SIP based communications if the SIP standard were supported. SIP provides for a proper negotiation for voice, video and text channels of communication on a call.

The approach being used (or migrated to) by most mainstream systems (including IMS) is SIP. This standard is supported by a wide range of off-the-shelf equipment, as well as some VRS providers. It alleviates some of the concerns with H.323, because the ICE[5] protocol and its cousins STUN[6] and TURN[7] can assist with the discovery of services, ports, and NAT traversal. However, SIP is not compatible with the existing installed base of H.323-based videophone equipment, so some point-to-point calls between off-the-shelf equipment and VRS-provided equipment will not work without employing a form of bridging between SIP and H.323, which would need to be set up and maintained by some party, for as long as H.323 still is in use by VRS users.

Since SIP is the overwhelming standard being adopted by mainstream telephony (including IMS), and IMS will in all likelihood be the backbone that will replace the PSTN over time, we recommend that SIP be considered for VRS interoperability formats, going forward. Other systems should be required to work with the mainstream SIP technologies. However, we must make sure that the installed H.323 base is not left out in the cold, until the transition has been completed. VPs are expensive to replace. Alternatively a decision to support the H.323 system in parallel permanently could be made but this would perpetuate the higher costs and complexity of supporting two systems by all parties.

---

[5] Interactive Connectivity Establishment, IETF RFC 5245
[6] Session Traversal Utilities for NAT, IETF RFC 5389
[7] Traversal Using Relays around NATs, IETF RFC 5766

III.A.7. Incoming call alerts

Hearing people are alerted to incoming calls via ring tones, which they can perceive even when they are not in the immediate vicinity of the phone or computer. Deaf people use visible and tactile signals, such as flashers and vibrators. In particular, there is a large installed base of house-and apartment-wide alerting systems that flash lights when an incoming call is detected. These systems are hooked up to the phone equipment via standard RJ-11 jacks. Some VRS-provided videophones also provide such jacks (although not all do), whereas no off-the-shelf equipment does. Neither are there currently any commercially available solutions for connecting the currently used incoming call alerting systems to computers.

If equipment does not provide capabilities to hook up home-wide visual alerting systems, the user has to be in the same room as the equipment to detect a visual incoming call alert, which clearly runs counter to the functional equivalence principle. Although it is possible to hook up equipment that picks up sounds to such alerting systems, they are not a practical solution for picking up ring tones, because they are too sensitive to picking up external noises and causing false alarms (such as babies crying, TVs running in the background, etc.).

For mobile access to VRS, it is also important to have functional cell phone-based alerting systems in place. Although cell phones have vibrators, a recent trend has been that they have become weaker as phones have slimmed down. This increases the risk of a deaf person missing a call on a cell phone. A Bluetooth-based standard for hooking up external vibrators would alleviate this issue.

It is recommended that the FCC launch an effort, or challenge, or both, to create a standard for connection of call alerts to new-generation VoIP devices. This might take the form of an IP based alert module (that could be registered to a phone) to allow remote alerting of calls to that phone, and a Bluetooth module that could be paired with a phone (without rerouting the audio from the phone) to allow remote alerting. These would allow modules to be created both for mainstream users and users who were deaf or hard-of-hearing. Some form of bridging to existing RJ-11-based alerting systems should also be provided, as home-wide systems are expensive to replace.

III.B. Details on functional parameters

Some of the following information (resolution, frame rate, lighting, bandwidth, and latency) was first filed in a joint ex parte comment by the Telecom RERC and the MobileASL project[8]. Where appropriate, it has been updated in response to current developments. Some data are based on the results of previous studies[9], as well as the information contained in the attached PDF file on the topic of "Camera requirements for video telephony with sign language."

[8] Harkins, J., Kozma-Spytek, L., Williams, N., Hellstrom, G., Vanderheiden, G., Ladner, R. (Jan 5, 2010). Ex Parte Comments of the Rehabilitation Engineering Research Center on Telecommunications Access and the MobileASL Project, In the Matter of Public Safety Issues Related to Broadband Communication To and From People with Disabilities, NBP #14, GN Docket Nos. 09-47, 09-51 and 09-137. (Available at http://fjallfoss.fcc.gov/ecfs/document/view?id=7020355298)

[9] Application profile – Sign language and lip-reading real-time conversation using low bit-rate video communication. ITU-T H-series Recommendations – Supplement 1, 05/99. (Available at http://www.itu.int/rec/T-REC-H.Sup1-199905-I)

III.B.1. Image Resolution

At least an effective QCIF resolution (176 × 144) is required; any resolutions below this number result in a continual decrease in usability. However, the CIF resolution (352 × 288) provides a much better user experience and, with currently available broadband speeds, provides a good tradeoff between quality and bandwidth requirements for VRS.

III.B.2. Frame rate

A minimum of 20 frames per second is required for clear and natural communication. Rates lower than 20 frames per second force the users to employ unnatural signing methods by slowing down, asking for clarification, repeating signs, and waiting for acknowledgment by the other party, especially for fingerspelling. This is a significant concern for access to emergency services, where time is of the essence.

III.B.3. Bandwidth

For CIF resolutions, a bidirectional bandwidth of 384 kBit/s has proven to be adequate, using H.263 encoding over an H.323 transport, which is currently in widespread use on VRS-provided equipment. Because various VRS providers have handed out such equipment free of charge to deaf consumers, it can reasonably be expected that H.263 over H.323 will need to be supported for a long time to come, both for VRS and point-to-point calls. If H.264 encoding were used and offered widely by VRS providers, bandwidth requirements could be cut in half (i.e., 192 kBit/s bidirectionally) for each call. In households or other places where multiple calls may occur at the same time, this bandwidth (bidirectionally) would be needed for each simultaneous call.

III.B.4. Latency and quality of service

The delay between the endpoints of a connection affects the quality of the conversation. If delays become too long, it disrupts the normal pattern of taking turns, to the point where people start talking or signing at the same time. To ensure a normal flow of the conversation, the total delay between end-to-end transmissions should be no more than 0.5 seconds, similar to what is considered tolerable in voice communications. Because encoding and decoding video takes substantial processing time by itself, the network latency between the endpoints should be no more than 0.25 seconds (i.e., no more than half of the maximum total delay).

III.B.5. Data volume caps

In order to maintain functional equivalence between hearing telephone and deaf videophone users, it is important to ensure that traffic due to VRS and point-to-point communications does not run into monthly data volume caps. For instance, a monthly 2 GB cap on a wireless data plan allows only 6 hours of total call time at a bidirectional transfer rate of 384 kBit/s before hitting the limit, which unfairly penalizes deaf VRS users. Hence, it is either necessary to establish generous caps for them, or to meter video calls by the minute, at the same rates as voice calls, instead of counting VRS and point-to-point video calls against the data transfer costs and limits.

III.B.6. Light sensitivity

Many systems are used in home environments where the lighting is significantly lower than in stores, offices, and public places. At a minimum, suitable cameras need to support the CIF resolution, a minimum of 20 frames per second, and exposure times no longer than 40 ms, under all of the following typical lighting conditions:

- 30 lux in low-lighting home environments (including situations where lighting is indirect, due to no ceiling lights)
- 100 lux in home desk environments
- 300 lux in office environments

Exposure times of 20 ms are optimal for sign language conversations under normal lighting conditions, and should be considered the goal to attain for all video calls, except very low lighting levels. Shorter exposure times (e.g., 10 ms) may result in sign language users perceiving doubled fingers.

Many consumer-level cameras offer no, or only very limited, control over exposure times. The standard strategy implemented by vendors for the auto-exposure mode is to increase exposure times and decrease the frame rate under poor lighting conditions. Although this results in better image quality for individual frames, it makes these cameras unusable for any kind of real-time sign language conversation in low lighting conditions. A better strategy for video calls would be to risk a slight increase in image noise levels under low lighting conditions, just so that the minimum exposure time and frame rate requirements are met.

These considerations imply that off-the-shelf cameras need to make the capabilities available to the OS driver for the following controls: overriding automatic exposure, overriding automatic frame rates, setting exposure times, and setting frame rates. This will allow the software or firmware to apply optimal settings specifically for sign language conversations.

III.C. Other desirable features

III.C.1. Call waiting

Call-waiting features that alert users to incoming calls while on the phone, and allow them to switch between calls at will, have become commonplace in voice telephony. Some VRS providers also offer call waiting functionality, but these are currently tied to specific videophones, and not interoperable across different VRS providers; that is, call waiting only works if the user calls a specific VRS via the specific equipment supplied by that VRS provider.

SIP already provides call-waiting functionality, so switching from H.323 to SIP for VRS calls would make this feature available to deaf consumers, as well, and interoperability with off-the-shelf equipment would be attainable.

III.C.2. Camera control

With stationary desktop-style videophone equipment, it is useful to have a camera that can zoom, pan, and tilt, so as to accommodate the individual preferences of users, especially in home environments (e.g., sitting on a couch as opposed to a chair in front of the videophone), and for switching between multiple parties that use the same camera on the same endpoint during the call. However, not every VRS-provided videophone has this functionality, and it is not required for functional equivalence.

Allowing the remote side to control the pan, tilt, and zoom of the local camera is of interest, because it could potentially provide PSAPs with another means of taking a look at the scene when the caller is unable to adjust the camera according to the responder's requests. Some VRS-provided equipment does provide remote control capabilities, but in

order to take full advantage of it during 9-1-1 calls, it will be necessary to provide a split view showing the caller, the VRS interpreter, and the PSAP responder during an emergency call – something that should be left to the NG 911 initiatives. For the purposes of these comments we note that the technical capabilities for such types of calls need to be available and interoperable.

III.C.3. Real-time texting during video calls

Most VRS providers make a chat mode available, which allows the parties involved in the video call to exchange text messages during the course of a call. With some providers, this takes the form of an ad-hoc IM channel, while other providers have integrated this functionality into their equipment or software. The idea is that some types of information are much easier to transmit and retain via a text channel than sign language. For example, many people struggle with writing down information at the same time as watching the other party sign, including web URLs, confirmation numbers, and order numbers.

Practice has shown that this feature constitutes a great time saver, and simplifies the jobs of the VRS interpreters, which also has the effect of reducing the number of minutes spent on a VRS call – thereby reducing the costs associated with such calls. For these reasons, providing a real-time text channel along with the video (for both VRS and point-to-point calls) should be assigned a high level of importance. RERC-TA strongly recommends considering it for inclusion in standardization and interoperability efforts. Note that SIP already provides facilities for both real-time text and video along with audio as a natural part of a single call. Having text as part of the call (rather than

requiring a separate parallel text communication call) is important for many users who are older, or for other reasons do not handle technical complexity well.

III.D. Interoperability with popular video chat and web-based services

Some VRS providers offer options to place calls via iChat and FaceTime. In addition, recently, web-based services have emerged that require nothing more than a computer with a webcam and a running web browser to make a VRS call. These options should be viewed as complementary to interoperable videophone and VRS call standards, not as a substitute for them. They are mainly useful in situations where a videophone is not available. Deaf users can borrow someone else's computer (who is not a VRS user) to place a call, and avoid spending time on downloading and installing additional software on it.

These services also can be used as a convenient way to hold high-quality point-to-point and conference video conversations among users who have downloaded and installed the respective software. Moreover, because they are not subject to the TRS rules – specifically the TRS Numbering Directory access restrictions –, they can be used for communication among mixed deaf and hearing participants without constraints. Yet, it needs to be stressed again that they are not functionally equivalent substitutes for interoperable videophones.

The downsides to using such services are similar to the interoperability hurdles described in the previous sections. Because they were developed for a general audience, rather than the videophone – or VRS – market, they cannot be used for point-to-point

calls to videophones (not even the web-based VRS service at the time of writing this document), and they lack the E911 and ANI features that VRS-supplied equipment provides. They also face issues with incoming call alerts, until modules for bridging PC software with alerting devices are developed.

IV. Conclusions

At present, substantial technical and policy hurdles remain in the path of making VRS equipment and off-the-shelf equipment interoperable, and providing true functional equivalence. Currently available off-the-shelf equipment does not meet functional equivalence requirements, and the lack of access to the 10-digit numbering system and a standard way to pass telephone numbers make it hard to fix these problems. Market forces will not drive deaf accessibility to these products for purposes of video communications. Yet, as we have shown in this document, none of these hurdles are insurmountable.

Mapping out a transition from H.323 to SIP as the standardized communication protocol of choice for VRS and point-to-point calls would clear out many obstacles at one fell swoop. SIP is emerging as the standard protocol for VoIP services and off-the-shelf videophone services, and switching to it would go a long way toward interoperability and also allow VRS, as well as off-the-shelf services, to take advantage of technological innovations quickly. Conversely, off-the-shelf-solutions must be integrated into the system for routing emergency calls via VRS, and passing on E911 and ANI information.

A switch to SIP implies that care must be taken to retain backward compatibility with the established H.323 base of videophones in the deaf community for VRS and

point-to-point calls. This could be handled via servers that bridge SIP with H.323 for as long as the latter protocol remains in use.

We are at the threshold of a transition period for sign language video calls, which may be painful, but promises a bright future for functional equivalence between deaf and hearing people in telecommunications. Now is an especially good time to take steps toward interoperability, before deaf consumers make substantial investments into videophone hardware. The RERC-TA looks forward to working with the FCC on achieving interoperability across videophones.

Respectfully submitted,

/s/ Christian Vogler

On behalf of the RERC on Telecommunications Access[10]:

Christian Vogler, Ph.D.,
Co-Principal Investigator,
RERC on Telecommunications Access,
Director, Technology Access Program
and Norman Williams,
Technology Access Program
Gallaudet University
800 Florida Ave., NE, SLCC 1116
Washington, DC 20002
(202) 250-2795

and

/s/ Gregg C. Vanderheiden

Gregg C. Vanderheiden, Ph.D.,
Co-Principal Investigator,
RERC on Telecommunications Access,
Director, Trace R&D Center
University of Wisconsin-Madison
1550 Engineering Drive, 2107 ECB
Madison, WI 53706-1609
(608) 262-6966

Date: April 1, 2011

[10] The contents of these comments were developed with funding from the National Institute on Disability and Rehabilitation Research, U.S. Department of Education, grant number H133E090001 (RERC on Telecommunications Access). However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.