



Human Factors (HF); Real-Time Text (RTT) in Multiparty Conference Calling (standards.iteh.ai)

[ETSI TR 103 708 V1.1.1 \(2022-08\)](https://standards.iteh.ai/catalog/standards/sist/06f5bd82-5e85-4206-8a8a-386d2c462cf4/etsi-tr-103-708-v1-1-1-2022-08)

<https://standards.iteh.ai/catalog/standards/sist/06f5bd82-5e85-4206-8a8a-386d2c462cf4/etsi-tr-103-708-v1-1-1-2022-08>

Reference

DTR/HF-00103708

Keywords

accessibility, HF, ICT

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://standards-portal.etsi.org/People/CommitteeSupportStaff.aspx>

If you find a security vulnerability in the present document, please report it through our

Coordinated Vulnerability Disclosure Program:

<https://www.etsi.org/standards/coordinated-vulnerability-disclosure>

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2022.
All rights reserved.

Contents

Intellectual Property Rights	6
Foreword.....	6
Modal verbs terminology.....	6
Executive summary	6
Introduction	7
1 Scope	8
2 References	8
2.1 Normative references	8
2.2 Informative references.....	8
3 Definition of terms, symbols and abbreviations.....	10
3.1 Terms.....	10
3.2 Symbols.....	10
3.3 Abbreviations	11
4 Real-Time Text (RTT)	11
4.1 What is RTT?	11
4.2 RTT is for conversation not messaging.....	11
5 Multiparty calls	12
6 User interface elements and use cases.....	12
6.1 Background	12
6.2 Types of media	12
6.2.1 Real-Time Text (RTT).....	12
6.2.2 Audio	13
6.2.3 Video	13
6.2.4 Total Conversation.....	13
6.2.5 Messaging.....	13
6.2.6 Documents, drawings and stored or streamed media.....	13
6.3 Use of media and user interface elements	14
6.3.1 Potential user interface components	14
6.3.2 User generated media presentation	14
6.3.3 Active and waiting user activity indication.....	14
6.3.4 User control of presentation and input.....	14
6.4 Use cases	15
6.4.1 Use case variations.....	15
6.4.2 Call using RTT within a small group of Deaf persons.....	15
6.4.3 Deaf person calling emergency service and using RTT.....	15
6.4.4 Hard-of-hearing user talking with hearing friends.....	15
6.4.5 Deaf user participating in conference getting transcription support	16
6.4.6 Deaf user participating in conference contributing by text-to-speech	16
6.4.7 Deaf-Blind user participating in remote meeting.....	16
6.4.8 Person in a critical situation making an emergency call by RTT.....	16
6.4.9 Person in remote group meeting in occasional noise	17
6.4.10 Relay service using multiparty technology	17
6.4.11 Using an RTT relay service to connect to a voice conference call	17
7 RTT text creation and presentation	17
7.1 Text creation and transmission.....	17
7.1.1 Variation in text creation means	17
7.1.2 Verification of produced text	18
7.1.3 Erasure	18
7.1.4 Input language and character set.....	18
7.1.5 Sending of emoji.....	18
7.1.6 Graphic rendition	18

7.1.7	Time from text entry to presentation	19
7.1.8	Support for user to send text when the situation allows.....	19
7.1.9	Reliability versus rapidity in transport of text.....	19
7.2	RTT text presentation.....	19
7.2.1	RTT text presentation basics.....	19
7.2.2	Visual presentation variants	20
7.2.2.1	Introduction.....	20
7.2.2.2	One column view	20
7.2.2.3	One column per contributing user	20
7.2.2.4	Text beneath a video image of the corresponding participant.....	21
7.2.3	RTT text presentation by screen reader with spoken output.....	21
7.2.4	RTT presentation by screen reader software with braille output	21
7.2.4.1	Notes on braille device user interfaces.....	21
7.2.4.1.1	Braille Devices	21
7.2.4.1.2	Braille Keyboards.....	21
7.2.4.1.3	Braille Displays	21
7.2.4.1.4	Braille Matrix Displays	22
7.2.4.2	Anatomy of a Braille Display.....	23
7.2.4.3	Example RTT Implementations with Braille Displays	24
7.2.4.3.1	The purpose of the following examples.....	24
7.2.4.3.2	Example: RTT using Reserved Status Cells.....	24
7.2.4.3.3	Example: RTT without Status Cells on Portable Displays	24
7.2.4.3.4	Example: RTT on a Braille Matrix Display.....	24
8	Control of RTT in multiparty calls.....	25
8.1	General aspects of control of RTT	25
8.2	Control and indications of the visible user interface components	25
8.2.1	Control of the screen space	25
8.2.2	Scrolling in RTT presentation.....	26
8.2.3	Searching and selecting the RTT contents	26
8.2.4	Varying the RTT presentation	26
8.2.5	Notifications and status information about participants and media	26
8.2.6	Selection between meeting views	26
8.2.7	Notifications by audio and visual means	27
8.3	Control via assistive technologies	27
9	Control of calls with RTT	27
10	Transport of RTT with multiparty contents in various technical platforms.....	27
10.1	General transport considerations.....	27
10.2	Multiparty RTT transport in centralized SIP conferences	28
10.3	Multiparty RTT transport in WebRTC.....	28
10.4	Multiparty RTT transport in PSTN	28
10.5	Multiparty RTT transport in PEMEA.....	29
11	Interaction with services.....	29
11.1	Services of specific importance for RTT users.....	29
11.2	Continuous real-time conversational services	29
11.3	Conference services.....	29
11.4	Emergency services.....	29
11.5	Relay services.....	30
11.5.1	The purpose of Relay Services	30
11.5.2	Functionality and connection alternatives.....	30
11.5.3	Delay caused by translation or interpretation	30
11.5.4	Queueing situations.....	30
11.5.5	Language and translation	31
12	Relations to specifications from other groups.....	31
12.1	General about support from other standardisation bodies	31
12.2	Relations to 3GPP specifications.....	31
12.3	Relations to GSMA specifications	32
12.3.1	Introduction.....	32
12.3.2	Analysis of GSMA NG.114, IR.92 and IR.94	33

12.3.3	Other GSMA profiles for conversational voice, video and RTT	33
12.4	Relations to emergency service specifications	34
13	Proposed information in EN 301 549	34
13.1	The need to update EN 301 549 to better address RTT	34
13.2	Changes to informative references in clause 2.2	34
13.3	Changes to definitions of terms in clause 3.1	35
13.4	Changes to existing requirements in clause 6	35
13.4.1	Continuous real-time conversation	35
13.4.2	RTT communication	35
13.4.3	Concurrent voice and text	36
13.4.4	Distinguishable display	36
13.4.5	Programmatically determinable send and receive origination	36
13.4.6	User identification	37
13.4.7	Visual indication of Audio with RTT	37
13.4.8	Additional clauses about RTT presentation	37
13.4.9	Interoperability	38
13.4.10	RTT responsiveness	39
13.4.11	Further requirements on RTT	39
13.5	Identification of actively speaking users	40
13.6	Identifying who wishes to communicate next	41
13.7	Awareness versus distraction	41
13.8	Programmatic determinability of the new notifications	41
13.9	Changes to requirements on services in clause 13	41
13.9.1	Access to relay services	41
13.9.2	Access to emergency services	42
Annex A (informative):	EN 301 549 extracts with proposed changes	43
A.1	Introduction	43
A.2	The EN 301 549 revisions and additions	43
Annex B (informative):	Bibliography	51
History		52

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org/>).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Human Factors (HF).

Modal verbs terminology

In the present document **"should"**, **"should not"**, **"may"**, **"need not"**, **"will"**, **"will not"**, **"can"** and **"cannot"** are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"must" and **"must not"** are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

Real-Time Text (RTT) is text communication sent as it is created without any specific sending action by the user. It is required in EN 301 549 [i.13], including its use in multiparty calls, but has lacked details on user interface requirements and how to accomplish the multiparty function.

The present document explains the use and characteristics of RTT including its multiparty aspects. It also provides an analysis of documents from other standardisation bodies, and briefly proposes modifications to them to make lower layers of RTT implementation support multiparty calling and to ensure that emergency services are able to support RTT with multiparty functionality. Finally, detailed change proposals for EN 301 549 [i.13] version 3.2.1 are included in the final clause of the present document and in Annex A.

Introduction

The present document was developed to support future upgrading of those parts of the EN 301 549 "Accessibility requirements for ICT products and services" [i.13] standard that relate to the accessibility of Real-Time Text (RTT) when used in multiparty/conference applications. Multiparty usage is of particular importance in ensuring that communication with Emergency Services will meet the accessibility needs of persons for whom voice communication is undesirable or impossible. The inclusion of multiparty requirements will also ensure that EN 301 549 [i.13] remains relevant to the design of interoperable RTT and Total Conversation services.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ETSI TR 103 708 V1.1.1 \(2022-08\)](https://standards.iteh.ai/catalog/standards/sist/06f5bd82-5e85-4206-8a8a-386d2c462cf4/etsi-tr-103-708-v1-1-1-2022-08)

<https://standards.iteh.ai/catalog/standards/sist/06f5bd82-5e85-4206-8a8a-386d2c462cf4/etsi-tr-103-708-v1-1-1-2022-08>

1 Scope

The present document establishes user interface guidelines for RTT conference call interfaces and identifies technical support needed to implement the guidelines. The present document:

- introduces RTT;
- addresses the types of media and user interface elements that can be associated with RTT;
- describes some RTT use cases, highlights issues particularly relevant in multiparty scenarios;
- explains technical support issues related to text creation and presentation, and the standards and related services associated with RTT transport and control.

The final and crucial part of the present document proposes some potential changes and additions to EN 301 549 [i.13] to ensure that it covers the additional accessibility issues that arise when considering multiparty scenarios.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

- [i.1] ETSI ES 202 975: "Human Factors (HF); Requirements for relay services".
- [i.2] ETSI TS 101 470: "Emergency Communications (EMTEL); Total Conversation access to Emergency Services".
- [i.3] ETSI TR 103 201: "Emergency Communications (EMTEL); Total Conversation for emergency communications; implementation guidelines".
- [i.4] ETSI TS 103 478: "Emergency Communications (EMTEL); Pan-European Mobile Emergency Application".
- [i.5] ETSI TS 103 479: "Emergency Communications (EMTEL); Core elements for network independent access to emergency services".
- [i.6] ETSI TS 123 167: "Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS) emergency sessions (3GPP TS 23.167)".
- [i.7] ETSI TS 123 226: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Global text telephony (GTT); Stage 2 (3GPP TS 23.226)".
- [i.8] ETSI TS 124 147: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Conferencing using the IP Multimedia (IM) Core Network (CN) subsystem; Stage 3 (3GPP TS 24.147)".

- [i.9] ETSI TS 124 229: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (3GPP TS 24.229)".
- [i.10] ETSI TS 124 371: "Universal Mobile Telecommunications System (UMTS); LTE; Web Real-Time Communications (WebRTC) access to the IP Multimedia (IM) Core Network (CN) subsystem (IMS); Stage 3; Protocol specification (3GPP TS 24.371)".
- [i.11] ETSI TS 126 114: "Universal Mobile Telecommunications System (UMTS); LTE; 5G; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (3GPP TS 26.114)".
- [i.12] ETSI TS 126 236: "Universal Mobile Telecommunications System (UMTS); LTE; Packet switched conversational multimedia applications; Transport protocols (3GPP TS 26.236)".
- [i.13] EN 301 549 (V3.2.1): "Accessibility requirements for ICT products and services" (jointly produced by ETSI/CEN/CENELEC).
- [i.14] GSMA PRD IR.92: "IMS Profile for Voice and SMS".
- [i.15] GSMA PRD IR.94: "IMS Profile for Conversational Video Service".
- [i.16] GSMA PRD IR.51: "IMS Profile for Voice, Video and SMS over untrusted Wi-Fi access".
- [i.17] GSMA NG.106: "IMS profile for Video, Voice and SMS over trusted Wi-Fi access".
- [i.18] GSMA NG.114: "IMS Profile for Voice, Video and Messaging over 5GS".
- [i.19] GSMA NG.115: "IMS Profile for Voice, Video and Messaging over Untrusted WLAN Connected to 5GC".
- [i.20] IETF BCP 47: "Tags for identifying languages", M. Davis, A. Phillips, September 2009.
- [i.21] IETF RFC 3261: "SIP: Session Initiation Protocol", J. Rosenberg et.al., 2005.
- [i.22] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications", H. Schulzrinne et.al., 2003.
- [i.23] IETF RFC 4103: "RTP Payload for Text Conversation", G. Hellstrom, P. Jones, 2005.
- [i.24] IETF RFC 5194: "Framework for Real-Time Text over IP Using the Session Initiation Protocol (SIP)", A. Van der Wijk, G. Gybels, 2008.
- [i.25] IETF RFC 6497: "BCP 47 Extension T - Transformed Content", M. Davis, A. Phillips, February 2012.
- [i.26] IETF RFC 8373: "Negotiating Human Language in Real-Time Communications", Gellens R., 2018.
- [i.27] IETF RFC 8825: "Overview: Real-Time Protocols for Browser-Based Applications", 2021.
- [i.28] IETF RFC 8865: "T.140 Real-Time Text Conversation over WebRTC Data Channels", Holmberg C. and G. Hellström, 2021.
- [i.29] IETF RFC 8866: "SDP Session Description Protocol", A. Began et al., 2021.
- [i.30] IETF RFC 9071: "RTP-Mixer Formatting of Multiparty Real-Time Text", DOI 10.17487/RFC9071, Hellström, G., (July 2021).
- [i.31] IETF RFC 9248: "Interoperability Profile for Relay User Equipment", DOI 10.17487/RFC9248, Rosen, B., June 2022.
- [i.32] ISO/IEC 10646:2020: "Information technology - Universal coded character set (UCS)".

- [i.33] Recommendation ITU-T F.700 (2000): "Framework Recommendation for multimedia services".
- NOTE: Available at <https://www.itu.int/rec/T-REC-F.700-200011-I/en>.
- [i.34] Recommendation ITU-T T.140 (1988): "Protocol for multimedia application text conversation".
- [i.35] Recommendation ITU-T V.18 (2000): "Operational and interworking requirements for DCEs operating in the text telephone mode".
- [i.36] ETSI TS 103 871: "Emergency Communications (EMTEL); PEMEA Real-Time Text (RTT) Extension".
- [i.37] W3C Recommendation 05 June 2018: "Web Content Accessibility Guidelines (WCAG) 2.1".

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

communications assistant (CA): person providing alternative communication for a two-party conversation, multiparty meeting, or broadcast

NOTE: Communication assistants include captioners/transcribers, sign language interpreters, spoken language interpreters, relay service operators, call handlers, telephone operators, etc.

conference floor control: functionality for conference administrators and attendees to manage a formal or informal communication queue or other communication resources

NOTE: Examples of conference floor control include a "hand-raising" feature and a "request to share screen" feature.

continuous real-time conversation: type of organization in conversation and discourse where one participant's contribution is made available to other participants while it is being made

programmatically determinable: able to be read by software from developer-supplied data in a way that other software, including assistive technologies, can extract and present this information to users in different modalities

NOTE: WCAG 2.1 [i.37] uses "determined" where this definition uses "able to be read" (to avoid ambiguity with the word "determined").

Real-Time Text (RTT): form of a text conversation in point to point situations or in multipoint conferencing where the text being entered is sent in such a way that the communication is perceived by the user as being continuous

relay service: electronic communications service that enables users of different modes of communication (e.g. text, sign or speech) to interact by providing conversion between different modes of communication, usually through a communications assistant

Total Conversation service: audiovisual conversation service providing bidirectional symmetric real-time transfer of motion video, text and voice between users in two or more locations (from ETSI TS 101 470 [i.2])

WebSocket: computer communications protocol, providing interaction between a web browser (or other client application) and a web server facilitating real-time data transfer from and to the server

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

3GPP	3G (mobile) Partnership Project
4G	4 th Generation (mobile networks)
5G	5 th Generation Mobile Networks
API	Application Programming Interface
BCP	Best Current Practice
CN	Core Network
EMTEL	Emergency Telecommunications
GSMA	GSM Association
HF	Human Factors
ICT	Information and Communication Technology
IETF	Internet Engineering Task Force
IM	Instant Messaging
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IR	International Roaming Expert Group
ITU-T	International Telecommunication Union - Telecommunication standardization sector
LTE	3GPP Long Term Evolution (4G)
NG	Next Generation
PEMEA	Pan-European Mobile Emergency Application
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RFC	Request For Comment
RTP	Realtime Transport Protocol
RTT	Real-Time Text
SIP	Session Initiation Protocol
SMS	Short Message Service
TC	Technical Committee
UE	User Equipment
UI	User Interface
VoIP	Voice Over IP
WCAG	Web Content Accessibility Guidelines
WebRTC	Web Real Time Communication
WLAN	Wireless Local Access Network

4 Real-Time Text (RTT)

4.1 What is RTT?

RTT, or Real-Time Text, sends text characters input by a user shortly after they are entered. This contrasts with most text chat services where the user has to perform a confirmation step before the composed message is sent, usually by tapping a "Send" button. At the receiving end of an RTT conversation, the recipient sees the words from other participants as the characters are entered. The immediacy of the sending and receiving allows a dialogue that has a flow closely approximating that of a spoken dialogue.

Persons who use RTT use it for conversations in the same way as persons who can talk and hear use voice communications (or Voice and RTT together).

4.2 RTT is for conversation not messaging

Users may choose to communicate using a stand-alone messaging service or a chat feature in a conferencing system. They are making an active choice to use this less immediate, asynchronous form of communication.

The provision of a messaging/chat capability cannot be seen as providing an effective means for persons who cannot effectively make spoken contributions to effectively participate in a real-time person-to-person conversation or a conference call.

5 Multiparty calls

On a multiparty call (where most participants are using speech), including conference calls/meetings, the participants should not ideally talk at the same time, but in practice it occurs. If two or more participants speak at the same time it can rapidly lead to the multiparty call becoming ineffective for most of the participants. Multiple participants simultaneously using RTT could also have a similar effect, but the fact that text stays on the screen and the fact that text from different sources is presented with separation makes it feasible that some simultaneous text communication from different sources can be manageable. To what degree this is possible depends on the situation and how the conference is managed.

An important factor that leads to simultaneous contributions is a lack of awareness of when other participants have stopped contributing and when it is a suitable time for them to contribute. With voice communication, it is possible for most participants to hear when other participants are contributing, but they may not be able to identify who is speaking.

When several RTT users are in a multiparty call they need to be made aware when other RTT users are contributing. They also want to identify the other RTT users when they contribute. These are issues that are addressed in the present document. When contributions are being made both by voice and by RTT, the awareness of when someone is contributing becomes a much more complicated issue as contributions may be being made using a means that the user cannot, or is not able to, continuously monitor. This is even more true when video is used for information, e.g. by using sign language. User interface techniques that support users in having maximum awareness of when another person is actively contributing is a significant part of the scope of the present document.

6 User interface elements and use cases

6.1 Background

Availability of RTT is essential or desirable in several different situations in calls with two or more participants. This clause lists several valid use cases, with different combinations of media, different use of RTT, different characteristics of the devices for generation and presentation of RTT, and different means for management of the expressions in the different media in the call. The purpose is to be a base for checks that the specified characteristics of the user interfaces fulfil the user needs in these use cases.

6.2 Types of media

6.2.1 Real-Time Text (RTT)

The important characteristic of RTT is that text is transmitted at the same rate as it is produced, so that the receiver can follow the senders' thoughts as soon as they are turned into words. There will be no excess waiting time for completely expressed messages, similar to spoken communication in audio and signed communication in video. All three real-time continuous exchanges enable rapid interaction in conversation. A commonly used coding and presentation standard for RTT clarifying the characteristics of RTT is Recommendation ITU-T T.140 [i.34].

RTT may be used so that all call participants are enabled to create and send RTT to the other participants, and the text being presented in readable chunks growing in real time with indication of source. The presentation should provide an approximate view of the relative timing of text from different parties.

RTT may also be used by one or more participants in a call, where other participants communicate by speech or signing. In such situations, a human-transcribed or automatic translation may be included between the users of the different media in both directions.

A third way to include RTT in a call is to include translation from other media only for creation of RTT, while all participants have means for presentation of multiparty RTT, and those who need or prefer to express themselves in RTT are able to send RTT.

6.2.2 Audio

Audio is often used in calls with two or more participants to convey speech between the parties. The most common configuration is that audio presented to each participant is mixed from all other participants, except from those whose audio is muted. Transmission from each party can usually be muted both by the party and by a meeting administrator. Reception to a party can be controlled by the party and sometimes by the meeting administrator.

Audio is different from all other media described here in that the media itself does not occupy any screen area. However, small screen areas are needed for visual indication of audio, and for speaker identification.

Speech can be reliably perceived only from one participant at a time. Only extremely brief expressions from more than one party can be perceived, and then with reduced reliability. Therefore, multiparty calls are managed so that for most of the time only one participant at a time talks. The management may be self-controlled by the participants or made automatically or made by technical and human means by a meeting administrator.

Audio may also convey other sounds than speech, e.g. important environment sounds from an emergency scene in an emergency call.

6.2.3 Video

Video in calls with two or more participants is most often used for showing the participants, in order to aid in communication (e.g. facial expressions, body language) and facilitate camaraderie.

Sometimes the video view is used for conveying sign language.

Sometimes the video view is used to show a document, slideshow, or other resource, but presenting these components separately is becoming more common.

Video for conversational use is rarely used alone, but usually combined with other conversational media.

6.2.4 Total Conversation

Total Conversation enables real-time interaction between two or more participants in a call with the opportunity to shift focus immediately between communication modalities in the three real-time conversational media available in the service.

6.2.5 Messaging

Multiparty calling services commonly provide ways for the participants to send messages to each other. Sometimes these messaging functions work independently of any real-time session, sometimes messages can only be sent while there is a real-time call. All users or specific users can be addressed by the messages and answers. Some services offer only plain text messages, while others can include images, files and video clips in the messaging channel. In some services translation can be requested.

6.2.6 Documents, drawings and stored or streamed media

It is very common that documents and drawings are presented in multiparty conferences as well as stored or streamed video, audio and text media. The transmission of this media is usually controlled by a meeting administrator or a presenting participant. The layout on the screen is usually controlled by the receiving participant, deciding how large the visual media is and to what degree other visual components in the user interface are combined on the screen.

On small screens, users would benefit from the ability to temporarily hide selected visual user interface components while watching meeting presentations. When a user interface component is hidden, it may also be desirable to get indications when there is something new in the hidden component.

6.3 Use of media and user interface elements

6.3.1 Potential user interface components

A multiparty multimedia call may contain many user interface components which compete for presentation space and the awareness, perception, creation, and management by the users. It is therefore relevant to consider all commonly appearing user interface components in a multiparty call, when designing the user interface for RTT.

6.3.2 User generated media presentation

Audio is unique in that it requires no screen space for its media contents presentation. All the other features listed in clause 6.2 require presentation on a screen for them to be perceived by sighted participants, and each feature will compete with the others for screen space. Participants will only be able to view contributions from other participants that are using video or text presentation if they currently have a part of the screen dedicated to displaying that media.

Video from each participant is often presented in subsections of the screen. The selection, size and placement of the video views of the participants can usually be controlled by the receiving user, and by a meeting administrator. Sometimes the size, style, or placement of the view of a participant is also influenced by sound or other activity from the participant, with the intention that the current speaker will be shown with a higher priority than the other participants.

6.3.3 Active and waiting user activity indication

Hearing participants will immediately be aware when another participant is speaking. Hard-of-hearing or Deaf participants may not be able to detect when one of the other participants is speaking, but often they are able to infer who is speaking by interpreting on-screen markers or placement of the participant information about the active speaker.

By default, users will only be aware that another participant is actively attempting to communicate with other participants if they are already monitoring the media channel that the participant is using, or if they infer it by interpreting on-screen markers or placement of the active speaker.

Knowing that someone is actively communicating is very important, knowing the means by which they are doing so is of secondary importance. It is, therefore, important that any means of identifying that a person is actively communicating should be available to all contributors at all times, irrespective of the means that they have chosen to use, or are only able to use. This will minimise the risk that someone using one communication method is not noticed by the other participants that are using a different communication method. As well as indication that a person is actively communicating, it is also important for other participants to be aware who is actively communicating at any time. The means of identifying that communication is taking place and the means of identifying who is communicating can often be the same.

It is more complex to automatically identify that video contains active communication than it is for text and audio, but it would be equally useful in an unmanaged conference with users mainly focusing on different media.

In some managed conference systems, there are controls that allow a participant to indicate that they wish to participate next (e.g. a "hand raising" system). This is sometimes supplemented by a system that allows other participants to view a queue of upcoming speakers. To ensure inclusive participation, it is important that all users, regardless of their chosen communication media type, can perceive the list of upcoming speakers.

6.3.4 User control of presentation and input

The sections above make it evident that there are many factors in a multiparty call that a user may need to control and be made aware of in an easily perceivable way. All the media types described in clause 6.2, except the audio media itself, may compete for screen space. In addition, the indications described in clause 6.3.3 may compete with each other and with the media presentation for the user's attention and screen space.

It is unlikely that a single interface design can present all the content from these competing sources, while satisfying all users in all situations. Users will need an easily handled means to select and configure the size and placement of what to present on the screen whilst simultaneously being aware that there are other contents available for presentation.

Controls will also be necessary to allow a participant to select which function (e.g RTT or text messages) will get text input, and which external devices will be used for audio, video, text and other input.

In addition to the fundamental complexity of presenting the correct content and status information at the right time, there will be a need for participants to check back and review the earlier RTT and message contributions. One reason is that not all participants will be able to reliably switch and refocus their attention to follow the flow of the conversations in real time.

6.4 Use cases

6.4.1 Use case variations

The following clauses contain descriptions of use cases that include RTT. The variations in use of RTT and other media are wide. The use case list can therefore not be comprehensive, but instead is a collection with sufficient variation to serve as examples for brief checking that a real or tentative implementation of RTT will be accessible and usable.

RTT is often indicated as a media of interest for communication with and between Deaf and hard-of-hearing persons and with persons with speech related disabilities, providing them with the same responsive way of communicating that is available to those able to use voice communication. The rapidity and lack of delays in RTT can also be of value in a written conversation between any users. Therefore, most of the use cases contain users with accessibility needs.

6.4.2 Call using RTT within a small group of Deaf persons

Four Deaf persons engage in a multiparty call with video, audio and RTT. Three of them use mobile phones, and one makes use of a small computer. They all prefer to use sign language, and therefore select a view where the three video images of the other participants take up most space in the user interface. They get into discussing a medicine, and in order to provide the name with exact spelling in a way that the others can keep, the person talking about the medicine types its name in RTT. The computer user gets the RTT field presented immediately and sees the name being typed. The ones with mobile phones get an indication in the user interface that there is new text in RTT to read, so they switch the view to one with place for the RTT content and can make note of the name before they return to the view with larger video images.

6.4.3 Deaf person calling emergency service and using RTT

A Deaf woman with a mobile phone gets into an emergency situation, so she calls the emergency number with RTT activated. The country she is in supports calls with RTT and audio to the emergency service. This is verified by a service discovery mechanism. The callers' language preferences are conveyed to the emergency service, so the call is routed to a suitable call taker. The call taker answers and carries out an initial conversation in RTT. It is decided to add a poison expert to the call, so the call is extended to a three-party call with RTT, and the poison expert takes over the main discussion, but the call-taker stays on the call observing the outcome by following the RTT flow, and then helps to dispatch an ambulance. During the call, audio is used so that the poison expert can listen to the Deaf woman's breathing and the expert can confirm the diagnosis. During the call, all parties see the text flow from the parties as it is typed, and the text gets presented in readable chunks and has brief labels indicating the role of the emergency service participants and the name of the woman in the emergency situation.

6.4.4 Hard-of-hearing user talking with hearing friends

A small group of friends have started sending text messages within the group discussing how to get together. One man in the group is hard-of-hearing. The alternatives and opportunities to arrange how to meet up get complicated so one of the parties takes the initiative to start a real-time multimedia call with audio, video and RTT, so that they can resolve more rapidly what is possible and what is not. By hearing and seeing the others talking, even the hard-of-hearing person manages quite well to understand what is said. But occasionally it gets hard to understand, and then he asks, "please type that!". The friends are used to this form of communication, so they divert to typing the hard part in the RTT connection. The hard-of-hearing man understands already what it was about by seeing the first half in text of what he missed in voice, and can say "thanks", and the call can continue mainly by speech.