Camilla Warnicke* and Charlotta Plejert

The use of the text-function in Video Relay Service calls

<https://doi.org/10.1515/text-2019-0174>

Received April 19, 2019; accepted January 13, 2021;

published online February 4, 2021

Abstract: The objective of the current study is to investigate whether and how the text-function offered in the Video Relay Service (VRS) is used and to demonstrate how its use affects the interaction of participants within this setting. The VRS facilitates calls between a person using signed language via a videophone and a person who is speaking via a telephone. An interpreter handles the calls and simultaneously interprets between the users and has direct contact with both users. All participants are physically separated from each other. The data consist of 12 recordings from the regular VRS in Sweden and the method used is Conversation Analysis. The findings show that typed text is used to: 1) conduct a repair; 2) pre-empt problems; 3) recycle text; and 4) overcome language differences.

Keywords: interpreter; signed language; spoken language; telephone; text; video relay service

1 Introduction

The Video Relay Service (VRS) is a service that facilitates calls between a person using a videophone and a person using a telephone. This kind of service is widespread in several countries around the world, such as Sweden, Norway, and the US (Haualand 2011). In the VRS, an interpreter works in a studio and handles calls by simultaneously interpreting between users who communicate with a signed language (on a videophone) and with a spoken language (on a telephone). Thus, all three interlocutors in the VRS are physically separated from each other (see Figure 1).

An option available in the visual space is the possibility to exchange written language between the user of the videophone and the interpreter by means of a

***Corresponding author: Camilla Warnicke**, University Health Care Research Center, Faculty of Medicine and Health, Örebro University, Sweden, SE 70182 Örebro, Sweden,
E-mail: camilla.warnicke@oru.se

Charlotta Plejert, Department of Clinical and Experimental Medicine, Division of Speech Language Pathology, Linköping University, Linköping, Sweden, E-mail: charlotta.plejert@liu.se

Figure 1: Flow chart of the VRS setting.

specific text-function in the VRS interface. The text-function has been more available to use over time as different types of videophones have been developed. The impact of using text in addition to signed and spoken languages in this setting has only briefly been previously investigated in a study of participants' co-creation of communicative projects within the VRS (Warnicke 2018). Increased knowledge about different communication options in video-based solutions and services for social interaction is important, not only for the possibility to communicate across languages (such as signed languages and spoken languages) via an interpreter, but also for the development of solutions for remote communication more generally.

In light of the above, the aim of the current study is to explore and describe, in detail, the use of typed text as part of the VRS exchange between the user on the videophone and the interpreter. The study is guided by the following research questions: 1) How and when is the text function used in VRS calls? 2) What purposes are the text-function used for, and how does the use of typed text affect the interactional organization of the VRS exchange?

The article is organized in the following way: In Section 2, an overview of the VRS is provided, focusing on the fact that the users of the service and the interpreter interact within and across different spaces, i.e., auditive and visual ones. This is followed by a presentation of relevant interactional research on interpreting under remote conditions. In Section 3, participants, data, and methodology are outlined. Section 4 is the analytical part of the article, where findings are presented. Section 5 offers a discussion of the key findings in light of the research questions and, finally, drawing on the findings, several implications for the future development of VRS and related services, are suggested as part of conclusion.

2 Literature review

In this section, the key aspects of the organization of the VRS are presented, including findings from some of the very few empirical studies that use interaction analytical approaches, such as Conversation Analysis (CA) when investigating VRS communication. A brief overview of the concept of repair in the VRS setting is also provided, since the achievement of mutual understanding is instrumental for the efficiency of the service, which the study at hand will also show. The VRS setting is somewhat more complex than mundane face-to-face interaction. Since participants are physically separated, the exchange is mediated by an interpreter, and participants are dependent on the resources that the technology provides.

2.1 The organization of the VRS

In the VRS, the shared relational space between the user on the videophone and the interpreter is visual, whereas an auditive space is shared between the interpreter and the interlocutor on the telephone. Signed languages, such as Swedish sign language (Svenskt teckenspråk, henceforth STS), rely on visual/gestural resources, whereas spoken languages in this setting rely on audio/oral resources. However, in face-to-face encounters, both signed languages and spoken languages naturally also rely on a wide range of resources, for example, bodily actions, facial expressions, and the use of objects. A difference in how the signed language is communicated in VRS compared to face-to-face interaction is that the setting of VRS is limited to a two-dimensional screen, whereas sign language is three-dimensional when all interlocutors share the same physical space. Interaction with a signed language on a videophone needs to be adapted to potentially limiting external circumstances (e.g., the technology), and may result in, for example, reduced signing speed and more explicit formulations (Keating et al. 2008; Keating and Sunakawa 2013). In addition, there are dissimilarities in spoken Swedish and STS in terms of linguistic features, such as differences in their structure and rules (Ahlgren and Bergman 2006). In fact, not only do STS and Swedish spoken language have a different syntax, but there are also natural differences between written language syntax (used in the text resource) and that of spoken language.

Earlier studies have demonstrated how technical devices that are essential for VRS calls affect both the interaction and the interlocutors in several ways, particularly in terms of turn-allocation and turn-organization (Marks 2018; Warnicke and Plejert 2012, 2016). One specific technical device in the visual space between the sign language user and the interpreter is the interpreter's headset (Warnicke and Plejert 2018). Although the headset is only handled by the interpreter and is visually

available to the interlocutor on the videophone, it affects all interlocutors and the organization of the VRS call, since the interpreter uses the headset for different purposes. For example, pointing towards the headset may be used as reference to a person or as a signal that someone is currently speaking. This action affects the possibility for the sign language user to take his or her turn. Another option available in the visual space, which is in focus of the current study, is the possibility to exchange text between the user of the videophone and the interpreter in the VRS interface.

2.2 Repair in the VRS

In a seminal paper by Schegloff, Jefferson and Sacks in 1977, repair is defined as a resource used to address “recurrent problems in speaking, hearing, and understanding” (1977: 363), and its overall function is to facilitate participants’ achievement of intersubjective understanding. In relation to signed language, a comparison has been made between spoken language repair, and repair in American sign language (Dively 1998). When it comes to research on interaction in the VRS, there is no study with a focus on repair *per se*. However, in most studies, repair features as an important resource for the turn-organizing machinery. For example, in Warnicke and Plejert (2012), the interpreter, or a user, initiated repair in order to signal that technology failed in some way (e.g., a frozen screen in the visual space).

In a similar vein, the interpreter may use pointing at the headset as a device to clarify who is currently holding a turn, or provide more specific information about a certain referent. The signed language user may also initiate repair by means of pointing, for example towards the screen, in order to indicate a misunderstanding by the interpreter concerning a particular referent (Warnicke and Plejert 2018). Just like in spoken language interaction, repair in the VRS is used by participants for the purpose of facilitating understanding and enhancing the progressivity of the exchange. However, in contrast to face-to-face interactions where all interlocutors can see each other, in the VRS repair may occur in one space between just two people of a triad, for example in signed language and by gestures between the user on the video-phone and the interpreter, without the user on an ordinary phone actually knowing or noticing (Warnicke and Plejert 2018).

3 Method and material

An initial sample of 25 authentic recordings of the participating interpreters’ computer screens in the regular VRS in Sweden was collected. The project of which

the current study is a part has been ethically approved by a Regional Ethical Board in Sweden (Dnr 2010/016), and all the participants have been provided information about the study in STS or in Swedish; and have signed a written consent form for participation. Approval comprised scientific publication of written extracts and screen shots from the data. All names used are pseudonyms. For the project, video extracts were accepted to be presented in scientific contexts and for educational purposes. In the present article, only written transcription excerpts are used. The corpus is stored on a server to which only the main author has password-protected access, and it is not open for researchers other than those connected to the project.

A preliminary analysis of the 25 recordings revealed that the text resource was used in the visual space in 12 of the recordings. Therefore, the final empirical corpus for the current study is 12 calls in which the text resource is used once or several times. The total length of the 12 calls was 2 h, 24 min, and 53 s. The participants were 13 videophone users, 18 telephone users, and 10 VRS interpreters. The experience of sign language interpreting by the participating interpreters ranged between 3 and 28 years, and their experience of VRS interpreting ranged between 1 and 16 years.

The recordings were analyzed using Conversation Analysis (Sidnell 2010; Sidnell and Stivers 2013) as developed into Multimodal interaction analysis (Mondada 2006). The choice of method was based on its suitability to study the complex interplay between different interactional resources used in the VRS, in this case signed language, spoken language, and written text, alongside resources such as gaze, facial expression, body movement, and gestures. The videos recorded from the computers were repeatedly watched, with a particular focus on instances where users and the interpreter employed the text resource. These instances were then transcribed in terms of what was signed, spoken, and written.

Making the transcriptions readable and usable for both languages (STS and spoken Swedish) at the same level of detail was cumbersome in several respects. As signed and spoken languages differ in the manner in which they are communicated; i.e., different modalities are used, simplifications of the transcripts were conducted to address the issue of equality of status in the different languages under consideration. From our perspective, these perceptions clearly illustrate the “written language bias” (Linell 2004) at play, which unfortunately continues to permeate much of the thinking concerning language structure worldwide. From our perspective, all forms of interactional resources used by interlocutors to achieve a sufficient degree of understanding are valuable in their own right and should be treated and analyzed as such. Therefore, transcribed STS may appear “simplistic” if viewed through the lens of the written-language-bias. However, conventions for spoken languages are well established (see below), whereas conventions for the transcription of signed languages are less developed (for STS,

however, cf. Mesch and Wallin 2015). Thus, it should be stressed that all means for interaction are assigned an equal status by the authors of the current study. However, for international publication purposes, all modalities are translated into English. Regarding this latter issue, the languages in the transcriptions were glossed: spoken Swedish with lower case letters and the signed language with upper case letters. In the transcriptions, symbols were also added in the margin for clarity: 🗣️ for spoken language and 🖐️ for signed language. In the English translation of the original language, the same conventions are used. A final aspect of transcribing the VRS interaction is that spoken, signed, and written languages may occasionally occur simultaneously. For example, the interpreter may speak and type text at the same time. The signing user and the speaking user may also express something while the interpreter is active. These kinds of simultaneous actions are marked in the transcriptions in accordance with Jefferson's convention of using brackets (Jefferson 2004; for further details, see Appendix).

There are, of course, other transcription annotation systems for signed languages, based, for example, on analytical work using the software ELAN (e.g., Mesch and Schönström 2018; Wallin and Mesch 2018). For the purpose of the current study, however, we used a more traditional means of CA-annotating, working directly with analysis of the video-data in parallel with transcribing in a Word-document following previously used conventions in work on the VRS (e.g., Warnicke and Plejert 2012). Although ELAN is a very useful tool for annotating different modalities, the somewhat more simplistic way of working in Word was chosen, since it is also immediately transferable into easy-to-read excerpts.

All transcribed excerpts in the current study involve an interpreter, abbreviated as IN. Some of the examples involve a videophone user who is signing, abbreviated as VP, and a telephone user who is speaking Swedish, abbreviated as IT. In one case (Example 5), an automatic answering machine was responding in the auditive space. In that particular case, the machine is abbreviated as AAM. In accordance with the focus of this article, all the examples involve the participants' exchange of written text. The text in the examples are typed either by the interpreter: TIN (Typed text from the Interpreter) or by the user of the videophone: TVP (Typed text from the Videophone). A list of abbreviations is provided below for clarity:

- IN: Interpreter
- VP: Interlocutor on the videophone
- IT: Interlocutor on the telephone
- AAM: Automatic answering machine
- TIN: Text from the interpreter
- TVP: Text from the interlocutor on the videophone

It should be noted that the interlocutors on videophones may be using different types of equipment. Therefore, text may be typed either letter by letter or typed and not displayed for the other partner in the visual space until the person who was typing pressed the “enter” key.

4 Analysis: use and function of typed text in the VRS

Analyses of the 12 calls in which the text-function was employed revealed that it was used for one overarching, and perhaps not surprising, purpose, that is, as a resource for ensuring the correctness of the information being exchanged between the person on the videophone (VP), the interpreter (IN), and the person on the telephone (IT). Specifically, the type of information in the corpus addressed names, numbers, or details about the aim of the call. The text-function was used both before and in connection with the IT, thus between the IN and VP only and during the interpreted interaction between all participants.

In the following sections, our analyses focus on how the exchange of text is used and what purposes it serves. The findings point to the main function of the text being employed to ensure correctness. However, when examining each case in detail, the procedures and foci vary to some extent. Therefore, the presentation below is divided into four sections in which the function of texting is highlighted by means of four interrelated practices: 1) conduct repair; 2) pre-empt problems by the use of text; 3) recycle text; and 4) overcome language differences.

4.1 Use of text to conduct repair

Interacting within the VRS is challenging since participants are restricted to some degree in their ability to access interactional resources and are dependent on well-functioning technical solutions. As with face-to-face interactions, it is important for participants to be able to resolve problems in producing and comprehending contributions to the interaction also in the VRS setting (cf. Section 2.2 on repair above). In this respect, our analyses reveal that the text-function provides an opportunity to initiate and conduct repairs. In Example 1, several repairs are performed to sort out the name of a street. These repairs are conducted in STS, followed by a repair that uses text; IN has requested the address of VP to send him information. VP attempts to provide IN his address; however, IN does not obtain

the exact address at once. The text-function is finally used as a resource to accomplish the aim of obtaining the correct address.

Example 1: I1: P12: The street number

Time code: 02.03-02.38

English translation

1. VP: [V]
2. IN: [((looks down))]
3. VP: STREET V-E-N-S-O-N-E-S STREET
4. IN: WHAT V
5. VP: V-E-N-S-O-N-E-S STREET
6. IN: V-E-[N STONE STREET
7. VP: [V-E-N-S-O-N-E-S STREET
8. VP: V-E-N-S-O-N-E-S
9. IN: ((nod))
10. VP: STREET
11. IN: ((looks down and writes on a sheet of paper))
12. VP: 9 9 9
13. IN: [((looks down))]
14. VP: [9 0]
15. VP: 9 0 CAN TYPE SAY((leaning forward and typing))
16. VP: CAN TYPE ((pointing towards the screen simultaneously))
17. VP: CAN SEE ((pointing towards the screen))
18. IN: YES

Swedish original transcription

1. VP: [V]
2. IN: [((looks down))]
3. VP: GATA V-E-N-S-T-E-N-S GATA
4. IN: VAD V
5. VP: V-E-N-S-E-N-S GATA
6. IN: V-E-[N STEN GATA
7. VP: [V-E-N-S-E-N-S GATA
8. VP: V-E-N-S-E-N-S
9. IN: ((nod))
10. VP: GATA
11. IN: ((looks down and writes on a sheet of paper))
12. VP: 9 9 9
13. IN: [((looks down))]
14. VP: [9 0]
15. VP: 9 0 KAN SKRIVA SA((leaning forward and typing))

16. VP: KAN SKRIVA ((pointing towards the screen simultaneously))
 17. VP: KAN SE ((pointing towards the screen))
 18. IN: J-A

Throughout Example 1, several kinds of repairs are performed, both to ensure the name of the street is correct and (from line 12) to determine the street number. In line 1, VP begins to spell his address. As IN does not look towards the screen at that time, she is not able to see this initiative, whereas VP makes a repair in line 3 and declares that the name of the street is “Venstone Street”.¹ However, since it may be difficult to perceive an unknown, spelled word in a two-dimensional screen, IN initiates an other-repair, requesting clarification (line 4). VP attempts to repair by once again spelling the name of the street (line 5). However, IN does not obtain the correct address in her next repetition of the address either (line 6). Again, VP further reiterates the name of the street (lines 7, 8, and 10), when IN looks down and writes something. This first part of the sequence, to sort out the address, is a negotiation with several attempts to repair information concerning an address. What comes next in the interaction is a negotiation concerning the number of the street (lines 12–18). Again, this negotiation is performed when IN looks down when VP provides the number (line 13). However, in this case, VP takes the initiative to use the text-function as a resource to type the number (line 15), which IN confirms (line 18).

As signed languages are three-dimensional and the screen within the VRS is two-dimensional, it may be difficult for VP and IN to see what is signed (cf. Keating and Sunakawa 2013; Keating et al. 2008; Warnicke 2018), and certain information, such as an address or phone number, may be more sensitive and must be exact, as illustrated above. Similar to Example 1, in Example 2, the typing of a phone number supports the understanding between IN and VP in addition to the signing of VP.

Example 2: I1.P16: Phone number

Time code: 00.22–00.50

English translation

1. IN: WHERE WANT CALL
 2. VP: YES YES CALL TO
 3. VP: 0 [7 7 1 (.) 5 6 7 (.) 5 6 7
 4. TIN: [0 7 7 1 5 6 7 5 6
 5. IN: LAST 5 4
 6. VP: [5 6
 7. IN: [0 7 7 1 (.) 5

¹ In the English translation, “stone” has been translated as the interpreter signs “STONE” in line 6 and does not spell it letter by letter.

8. IN: [6 7 5
 9. VP: [YES THEN 7
 10. VP: 5 6 7
 11. TIN: 7
 12. IN: YOUR ERRAND DEALS SOMETHING SPECIAL I PREPARE
 13. VP: I CALL TO TAX AGENCY
 14. IN: OKAY

Swedish original transcription

1. IN: VAR VILL RINGA
 2. VP: J-A J-O RINGA TILL
 3. VP: 0 [7 7 1 (.) 5 6 7 (.) 5 6 7
 4. TIN: [0 7 7 1 5 6 7 5 6
 5. IN: SISTA 5 4
 6. VP: [5 6
 7. IN: [0 7 7 1 (.) 5
 8. IN: [6 7 5
 9. VP: [J-A SEN 7
 10. VP: 5 6 7
 11. TIN: 7
 12. IN: DITT ÄRENDE GÄLLA NÅGOT SPECIELLT JAG FÖRBEREDA
 13. VP: JAG RINGA TILL SKATTE^VERKET
 14. IN: O-K

In Example 2, IN asks for the telephone number of a person who she would call (on behalf of VP). When IN looks at the screen and perceives the signed number from VP (line 3), she uses the text resource to type the number (line 4). However, IN does not obtain the entire number to type the first time and therefore initiates a repair, asking for the last few numbers (line 5). IN then uses the numbers that she has typed by reading from the screen and repeats them for VP (line 7 and 8). VP responds by providing a repair in lines 9 and 10, which IN then adds to the text. Thus, what can be observed in Example 2 is how IN utilizes the text-function to facilitate obtaining the correct telephone number. Thus, the typed text can be used either by the interlocutor on the videophone or by the interpreter to conduct a repair.

4.2 Pre-empt problems by the use of text

In the VRS, communication in spoken and signed languages is conducted in real time. As the interpreter needs time to see/hear what is expressed, comprehend

what is said, and formulate a rendition for the other user in the interaction, this processing takes time (Russell 2005) and influences the turn-organization of the concerned social exchange (Warnicke and Plejert 2012). The interpreter may handle the challenges of the discrepancy, the delay, in the interaction as well as the challenge to see what is signed. As the text is sustainable in the call because of the user interface and does not disappear after it has been signed or stated, potential problems may be pre-empted by using the text resource. In Example 3, how IN uses the text resource to render a precise telephone number from IP in the auditive space, to VP, in the visual space is demonstrated.

Example 3: I2:P24: Number for the switchboard

Time code: 07.26-07.50

English translation

1. VP: CALL [NUMBER ALSO PHONE NUMBER]
2. IN: [yes I need[]a phone number then]
3. IT: [mmm]
4. IN: if you have that
5. IT: yea it is just that you call our [switchboard there
then you you know then zero nineteen]
6. IN: [ONLY TO OUR
SWITCHBOARD JUST]
7. TIN: [019-]
8. IN: [zero nineteen]
9. VP: [YES]
10. IT: nineteen
11. IN: [mmm]
12. TIN: [19]
13. IT: eee fortytwo e zero zero
14. IN: [fortytwo]
15. TIN: [42]
16. IT: zero zero
17. IN: [zero zero]
18. TIN: [oo]
19. IT: mmm right
20. IN: okay
21. VP: ((thumb up))

Swedish original transcription

1. VP: RINGA [NUMMER OCKSÅ TELEFON NUMBER]
2. IN: [ja jag behöver[]ett telefon^nummer då]
3. IT: [mmm]

4. IN: om du har dä
5. IT: ja dä ä bara ni ringer till våra [växel där
du dä vet du dä noll nitton]
6. IN: [BARA TILL
VÅR VÄXEL]
7. TIN: [019-]
8. IN: [noll nitton]
9. VP: [J-A]
10. IT: nitton
11. IN: [mmm]
12. TIN: [19]
13. IT: eee föti tvåe noll noll
14. IN: [förti två]
15. TIN: [42]
16. IT: noll noll
17. IN: [noll noll]
18. TIN: [oo]
19. IT: mmm precis
20. IN: okej
21. VP: ((thumb up))

In Example 3, VP and IT have been discussing a complaint from VP, mediated by IN. IT has recommended that VP call another person for help with his aim. The transcript begins when VP requests the telephone number to call for assistance. IT begins to relay that the number that will be provided is only for the switchboard. Later, IT begins to state the number (line 5). However, IN makes a rendition that the number (only) is to the switchboard (line 6); thereafter, IN types the number for VP (lines 7, 12, 15, 18). When IN types the number in the visual space for VP, he simultaneously repeats the number vocally for IT as a confirmation in the auditive space. Thus, IN types the number instead of signing it for VP. This texted rendition is performed without negotiating the practice with VP or letting IT know.

As shown earlier, it may be difficult to see signed numbers in the VRS. In the case above, IN types text and responds to IT with a verbal confirmation (see lines 7 & 8, 10 & 12, 14 & 15 and 17 & 18). As shown in the above example, IN thus responds to IT himself. By means of this action, IN may use the text and self-select in response to IT, as a preference for progressivity of the conversation (Stivers and Robinson 2006). This self-selection by IN may avoid an instance of another initiated repair, i.e., pre-empting a potential source of problems (cf. Svennevig 2010). VP responds to the provision of the typed number with a thumbs up. In previous

research (Warnicke 2018; Warnicke and Plejert 2012), researchers have demonstrated how the interaction is influenced by cases when VP is about to write something on paper, such as a number. In such cases, when VP withdraws his/her gaze from the screen and therefore cannot see what IN is rendering, IN nonetheless needs to manage the situation somehow (cf. Example 1, lines 11 and 13). IN may inform the partner on the telephone that VP cannot see what is signed at that point. However, in Example 3, instead of signing what is occurring, IN utilizes the text-function, which facilitates the interaction since VP receives the exact telephone number via text. An additional aspect of the manner in which IN renders the telephone number here, i.e., by typing, is that the information is sustainable and can be returned to at a later stage. Thus, the use of text also has an impact on the auditive space in terms of how turns are organized (Warnicke 2018).

Interpreters in the VRS handle large amounts of calls on a daily basis and normally do not know what a call is about beforehand. Thus, the interpreter must handle unknown circumstances and facts at the moment they occur. This activity may be challenging (Warnicke 2018, 2019; Warnicke and Plejert 2012, 2016). As demonstrated in the example below, IN does not obtain the information from VP regarding where to call. The example demonstrates how the text-function may serve as an alternative resource to pre-empt further problems of misunderstanding. In Example 4, the interaction consists of several repair attempts concerning which institution to call before VP chooses to type the number instead of fingerspelling it. This use of text may function as pre-empting potential problems for the interpreter to obtain the correct number.

Example 4: I1:P17: Typing the phone number

Time code: 02.31–03.20

English translation

1. IN: WHERE WANT TO CALL YOU?
2. VP: YES NOW CALL TO MIGRATION BOARD
3. IN: OMNITOR?
4. VP: ((gesture: NO)) [MIGRATION M I G
5. IN: [M-I-G-R-A-T-I-O-N?
6. VP: YES
7. VP: NUMBER ((leaning forward and typing))
8. TVP: 0771 235 235
9. IN: ((nod))
10. IN: ((typing the phone number and call))

Swedish original transcription

1. IN: VAR VILL RINGA DU?
2. VP: J-A NU RINGA TILL MIGRATION

3. IN: OMNITOR?
4. VP: ((gesture: NEJ)) [MIGRATION M I G
5. IN: [M-I-G-R-A-T-I-O-N?
6. VP: J-A
7. VP: NUMMER ((leaning forward and typing))
8. TVP: 0771 235 235
9. IN: ((nod))
10. IN: ((typing the phone number and call))

In Example 4, VP takes the initiative to use the text resource and types such that IN can call, in this case, the national migration board (line 7). However, prior to the typing, VP has tried to tell IN what number to call (line 2), but IN does not understand what was signed and initiates a repair, requesting a confirmation of understanding (by asking about the name, Omnitor, which is the name of a video relay company in Sweden; line 3). VP makes a repair, stating that IN's understanding is incorrect and again attempts to sign who to call (line 4). IN responds via a request for confirmation of understanding (line 5). This time, VP confirms the understanding (line 6). After this repair sequence, VP types the number for IN to call by first clarifying that he will deliver the number and then types it for IN (lines 7 and 8). This example shows how the interaction is facilitated by the use of text and that the typed number causes no further repairs; instead repairs are avoided, and IN calls the correct number on VP's behalf.

4.3 Recycle text

The previous section has addressed how typed text in the VRS calls may pre-empt potential problems by interlocutors' use of the text function. In addition to this observation, in this section, we will further stress how text, as it is sustainable and may pre-empt potential problems, can be used several times during a call. This work facilitates the progress of the interaction.

In Example 5, VP has anticipated what information an answering machine will provide since he has called the same institution many times previously, which he notes later in the call.

Example 5: I1:P1: Personal identity number

Time code: 03.16-03.45

English translation

1. AAM: For faster service press your ten-digit personal identity number and square (.) otherwise wait and your call will be linked to a manager

(9.0)

2. VP: I WRITE PERSONAL IDENTITY NUMBER
3. IN: J-A BRA
4. VP: DONE
5. IN: ((nod))
6. TVP: 7111196992

Swedish original transcription

1. AAM: för snabbare behandling tryck ditt tiosiffriga personnummer och fyrkant (.) annars vänta så kopplas ditt samtal till en handläggare

(9.0)

2. VP: JAG SKRIVA PERSON^NUMMER
3. IN: J-A BRA
4. VP: FÄRDIGT
5. IN: ((nod))
6. TVP: 7111196992

Although VP has not received any rendition from IN regarding what the answering machine is requesting (note that line 1 is not rendered by IN), VP informs IN that he will type his personal number (line 2) to be prepared for the call (line 4). This action by VP may be a demonstration of his awareness of the potential challenges for IN to see the signed numbers; i.e., it is recipient designed (Sacks et al. 1974). In addition, this action could also show VP's potential meta-knowledge that IN needs to process the numbers to render them. VP may also anticipate a challenge for IN to render the numbers with a sensitivity towards the sequence in terms of when a potential question from IT may arise. However, approximately 10 minutes later in the call, VP is connected to the institution being called, as illustrated in Example 6 below. VP presents himself using his name (line 1) and points to the screen to provide his personal number (line 3).

Example 6: I1:P1: Pointing towards the screen

Time code: 18.50–19.12

English translation

1. VP: [HI ME P-E-T-E-R I CALL
2. IN: [e well hi my name is peter and I call

3. VP: MY PERSONAL IDENTITY NUMBER GIVE((pointing towards the screen))
4. IN: e my personal identity number it is seventy one eleven nineteen
5. IN: sixty nine ninety two
6. IT: seventy one eleven?
7. IT: n [nineteen?
8. IN: [ni
9. IN: yes
10. IN: [sixty nine[nin
11. IN: [69 [92
12. IT: [four last
13. IN: [sixty nine ninety two]
14. IN: [69 92]
15. IT: sixty nine ninety two
16. IN: mmm

Swedish original transcription

1. VP: [HEJ JAG P-E-T-E-R JAG RINGA
2. IN: [e jo hejsan jag heter peter och jag ringer
3. VP: MITT PERSON^NUMMER GE((pointing towards the screen))
4. IN: e mitt personnummer där är sjuttiet elva nitton
5. IN: sextinie nittitvå
6. IT: sjuttioett elva?
7. IT: n [nitton?
8. IN: [ni
9. IN: jaa
10. IN: [sextinie [nitt
11. VP: [69 [92
12. VP: [fyra sista
13. IN: [sextinie nittitvå]
14. IN: [69 92]
15. IT: Sextinie nittitvå
16. IN: mmm

VP has prepared the presentation to IT (see Example 5). When connected to the institution, VP introduces himself and states he will provide his personal number (line 3). VP points to the screen to IN in the visual space (line 3). VP's pointing is responded to by IN, who starts to render the personal number by reading it out loud for IT. IN performs this action with no further question or clarification to VP. Thus,

for VP and IN, it appears completely clear when and how to use the pre-typed text, although it was initially typed more than 10 min earlier in their interaction. This action could be viewed as a proof of a reciprocal agreement of how to render the personal identity number, i.e., to include the number in the rendition from IN to IT. Thus, the typed text is recycled in the interaction.

In several cases in the data, an exchange of text is used at the beginning of the call. This activity is, in certain cases, a consequence of the need to obtain consent from participants concerning recording their calls for research purposes. All the interpreters are instructed to request the addresses of the users to send a written document with information regarding the study and a consent form to be signed. However, it is not necessary to use the text resource for this purpose. In the corpus, there are examples when either the interpreter or the user of the service takes the initiative to use text on occasions where an address is needed, which may occur at the beginning as well as later in the call. However, in Example 7 below, the user of the service takes the initiative to use the text-function at the beginning of the call.

Example 7: I1:P1: E-mail address

Time code: 00.45-01.12

English translation

1. IN: I NEED YOUR ADDRESS OR E-MAIL OR SOMETHING YOU MUST
SIGN PAPER DOCUMENT[APPROVE]
2. VP: [YES YES]
3. IN: OKAY?
4. VP: OKAY
5. VP: I WRITE [E-MAIL WRITE
6. IN: [WRITE ((gesture)) CAN YOU WRITE
7. IN: YES THANKS
8. VP: YES
9. VP: [((looks down and types))
10. IN: [((touches the door behind her and arrange herself at
the chair))
11. VP: LOOK
12. TVP: E-mail: sinnet71@live.se
13. IN: J-A

Swedish original transcription

1. IN: JAG BEHÖVER DIN ADRESS ELLER MAIL*ADRESS ELLER NÅGOT
DU MÅSTE UNDERTECKNA PAPPER DOKUMENT [GODKÄNNA]
2. VP: [J-A J-A]
3. IN: O-K?
4. VP: O-K

5. VP: JAG SKRIVA [E^MAIL^ADRESS SKRIVA
6. IN: [SKRIVA ((gesture)) KAN DU SKRIVA
7. IN: J-A TACK
8. VP: J-A
9. VP: [((looks down and types))
10. IN: [((touches the door behind her and arrange herself at the chair))
11. VP: SE
12. TVP: E-mail: sinnet71@live.se
13. IN: J-A

In Example 7, IN notes her need to obtain VP's address (line 1), whereas VP initiates use of the text-function as a possible resource for providing IN with the information. As stressed above, using signed language and spelled words may be difficult for the party to perceive in the visual space (Warnicke 2018; Warnicke and Plejert 2012, 2016); however, by means of text, the information is precise, and as stated above, remains on the screen until one of the interlocutors has removed it. In cases when information is typed, the interlocutors in the visual space have the chance to gain sustainable access to the correct spelling (or to initiate a repair if the spelling is incorrect). The sustainability of the text also opens the possibility to use it several times at different occasions during the call. In Example 7, VP typed the e-mail address for IN at the very beginning of the call. However, the e-mail address also became relevant later in the same call when IT asked how to get in touch with VP. The recycling of the e-mail address, as typed text, is shown below.

Example 8: I1P1: May I?

Time code: 21.47–22.47

English translation

1. IT: if we need to get in touch [with you (.) how to contact you
2. IN: [IF CONTACT YOU HOW I?
3. VP: MAIL
4. IN: mmmajl could you eee send
5. IN: MAY I?
6. VP: [YEA GOOD
7. IT: [that's the best?
8. IN: m e [yea that's the best
10. IN: [YES
11. IT: yes

Swedish original transcription

1. IT: om vi behöver komma i kontakt [med dej (.) hur når vi
dej
2. IN: [O-M KONTAKT DIG HUR JAG?
3. VP: MAIL
4. IN: mmmej! kan man eee skicka
5. IN: FÅR JAG?
6. VP: [J-A BRA
7. IT: [de e de bästa?
8. IN: m e [jae dä ä dä bästa
10. IN: [J-A
11. IT: ja

In Example 8, IT asks how to contact VP if needed. VP says, in line 3, that e-mail is a possible means. IN has previously received the e-mail address as text (Example 7), and both IN and VP are aware of this. IN asks VP if she is allowed to deliver the e-mail address using what was typed previously (line 5). IN's request is sanctioned by VP (line 6). Thereafter, IN reads the e-mail address for IT. Hence, the use of the previously typed text becomes relevant at two separate occasions during the call, although the use and purpose of the text are different. In Example 6, the use of the text is to achieve precision in the information exchange between IN and VP. However, in Example 7, the typed text is recycled in the negotiation between IN and VP. The second use of the typed text (Example 8) makes it possible for IN and VP to co-construct IN's rendition for IT. A co-construction of how to render the e-mail address may allow the interaction to continue.

In sum, typed text can be recycled for various purposes over time during a call. The examples provided in this section demonstrate how the recycling of text facilitates the progression of an interaction. In certain respects, the facilitation is simplifying the participants' objectives since less effort needs to be made when information can be recycled.

4.4 Use of text to overcome language differences

In the VRS, text can be used as an alternative resource to other communicative means, such as a signed language, gaze and gesture, in the interaction between the interlocutors of the visual space when the interlocutors do not use the same signs. The combination of visual resources to overcome language differences is demonstrated in Example 9 below. VP uses text to define the aim of the call. In the example,

VP and IN are put on hold. These participants have been connected to each other and have occasionally talked to each other for approximately 15 min while waiting. In the interaction prior to Example 9, VP has used several signs that are not Swedish, such as the signs for PASSPORT and VISA. Throughout, VP and IN use different signs for these referents. VP, as previously noted, appears to have another first signed language than STS. When VP defines his aim, he uses the text-function as an additional (visual) resource to the visual contact and signed language.

It should be mentioned that Example 9 below was an exclusive case in the entire dataset. Nonetheless, we have included it since it displays how the text-function may facilitate a complex multilingual context, not just in terms of interpreter-mediated interaction between signed language and spoken language. Signed language users as well as spoken language users are at times multilingual (using several signed languages or several spoken languages). Such conditions, albeit unusual in the data at hand, may very well become common in a future of increased use of remote, video-based communication means. From this one case, it appears that the text-function resolves some challenges of a complex multilingual context. Also, it cannot be assumed that an interpreter of STS is also to master other signed languages.

Example 9: I1:P17: My passport Kap Verde

Time code: 15.18–15.45

English translation

1. IN: YEA TAKE OPPORTUNITY ASK BEFORE SWITCHBOARD ANSWER
(impossible to comprehend) ASK YOU?
2. VP: YES MY PASSPORT(not a Swedish sign for passport)IS
K ((start to type))
3. TVP: my passport kap verde
4. VP: ((pointing towards the screen))
5. IN: OH

Swedish original transcription

1. IN: J-O PASSA*PÅ FRÅGA INNAN VÄXEL SVARA (impossible to
decode)FRÅGA DU?
2. VP: J-A JAG MIN PASS(not a Swedish sign for passport)ÄR
K ((starts to type))
3. TVP: min pass kap verde
4. VP: ((pointing towards the screen))
5. IN: JASÅ

Example 9 begins at a point where IN takes the initiative to ask VP about the aim of his call (line 1). VP responds to this question by signing that it concerns his

passport. However, he uses a non-Swedish sign for passport, followed by a K, possibly the initial letter for “Kap Verde”² (line 3). VP makes a self-repair by continuing his turn-construction using text instead of signs, typing “my passport kap verde” (line 5). This self-repair is followed by a point towards the screen that is responded to by IN by means of a display of surprise (see Heritage 1998, for equivalents in English) (line 5). Thus, when the interlocutors do not share the same (signed) language and/or do not fully understand the other person’s language, difficulties can be overcome by using another, complementary modality, in this case typed text. Thus, the text-function supports the interaction and, as in the case above, sorts out referential challenges.

5 Discussion and conclusion

The current study is explorative and descriptive, and the aim has been to investigate how and when, and for what purposes the text-function in the VRS was used by participants. The findings point towards the fact that the overarching function of the opportunity to type text in the VRS interface, from how it is used, is to secure the correctness of information among interlocutors. However, as we noted in the introduction, the function was initially not always easy to use (for technical reasons) and has, over time, been refined. By contrast, what we can see from our recordings at the time of data-collection is that the text-interface now appears non-problematic for the users of VRS as well as for interpreters.

The study findings can be summarized as follows: text was used to conduct repairs when issues concerning the certainty of information arose. In these cases, the trouble source appeared to primarily occur in the visual space between VP and IN, when something in STS became unclear for some reason or needed to be corrected. At such points, the participants employed the text function to conduct an appropriate repair. In addition, and interestingly, the text function was also used to pre-empt potential difficulties or ambiguities and was therefore occasionally used ahead of signing and/or speaking. This practice was exemplified in Examples 3 and 4, where an interlocutor chose to use text instead of other interactional means. This practice is interesting in many ways, not the least in terms of how various interactional resources most of the time complement each other and how participants orient towards the different resources available to them and pick the one they judge to be the most apt for a particular purpose of interaction.

2 Kap Verde, in English Cape Verde, is not translated in the English transcript, as the interlocutor on the videophone signs the letter K as a potential initial letter, which is essential for the analysis.

In relation to this latter issue, one cannot help to reflect on the role of the sustainable nature of text as it has been designed in the VRS. As demonstrated, in contrast to signed and spoken languages, something typed in the VRS interface remains there permanently during the call, unless it is deleted by participants, and can be recycled and referred to later with precision and minimal effort. Thus, text can be typed regardless of the others' simultaneous gaze towards the screen, and all the interlocutors would have access to the written text, lasting over the duration of time relevant for the purposes of the interlocutors. This resource is, of course, less cognitively taxing than needing to actively recall something previously signed or said.

Thus, the sustainable nature of the text function is a likely explanation why it is used for certain purposes during VRS interactions. However, it is not claimed here that writing by any means is superior to other modalities, and it is clearly the case that the interlocutors in the VRS primarily use other modalities of communication (i.e., signed language, spoken language, facial expressions, and gesture), each of which has its own strengths and weaknesses. In our view, the interactional resources available to interlocutors can be adapted more or less well to a current context and specific tasks. The VRS may be viewed as unusual in some respects, not the least since participants are not present in the same physical space and are highly dependent on the technology offered to them in that setting. In addition, in a face-to-face encounter, the use of text would likely be perceived as cumbersome, in contrast to the immediate availability of a slot on a computer screen, which can effortlessly be typed in, as in the VRS.

A relevant observation is that within the 25 recordings initially used as a basis for the current study, the text resource was used in only half of the corpus. This may indicate that the users of the service either prefer communicating using other interactional means (e.g., signed language or spoken language), perhaps have not noticed or do not want to use the text function, or, potentially view it as a resource for, as we demonstrate here, an option to be used for certain circumstances, when it is necessary for information to be as unambiguous and clear as possible. As the use of typed text appears to facilitate the progressivity of interactions among all the interlocutors in the VRS, our findings imply that this resource could possibly be used more often, but it is also noteworthy that the interlocutors themselves orient to this textual resource as an option to be used only under certain circumstances, i.e., for the purpose of being exact.

From the perspective of the participants, and perhaps importantly, with regard to VRS-interpreters, the text-function may be viewed as an untapped resource that could be more explored and further used in the practical work by interpreters as well as the users of the VRS. In relation to other interactional modalities, such as signed language, and spoken language, this claim is, we are aware, not uncontroversial. We

need to stress, however, that we do not view any semiotic means as superior to another one. Instead, we suggest that an interplay between resources and opportunities for interaction (e.g., as made possible by technical solutions, signs, speech, or whatever) should not be discarded. In some ways, such issues resemble the debate concerning code-switching among multilingual people being viewed as “bad language use” not too long ago; something that has empirically proven to be entirely wrong. From the 1980s and onwards, it has become established that switching codes among multilingual persons is beneficial cognitively as well as socially (Auer 1998; Grosjean 1982). It is perhaps time for certain communities to accept, and benefit from, all available resources for communication that are out there (e.g., signed languages, spoken languages, written languages, non-verbal means, tactile means etc). It is a fact that text can be useful, for certain purposes, and decrease cognitive load, if it is designed as sustainable (as is the case in the VRS). It is therefore a resource that not only facilitates the interpreters’ work to coordinate the VRS interaction, but the interpreter as well as the users may also use the text with a sensitivity of interaction and re-use it when needed. In this respect, the use of text may be viewed as a useful “interpreting design” in the VRS calls.

As briefly presented as a rationale for the present article, knowledge regarding the use and the effect of the technical resources available to VRS-interlocutors is important for developers, users of the service, and interpreters. Since the setting is very complex, lacking many of the cues for checking and achieving understanding that are available when participants interact face-to-face in the same physical spot, it is relevant to understand what resources can be used for the VRS-call to proceed with a minimum of breakdown and repair and with as high degree of accuracy as possible. Based on the findings of the present study, it appears that the text-function seems to be one such facilitating device.

Competing interests: The authors declare that there is no conflict of interest.

Appendix: Transcription key. A selection of conventions and adaptation from Jefferson (2004)

swedish	lower case is used for spoken Swedish utterances
STS	upper case is used for the transcription of Swedish Sign Language (STS)
((gesture))	double brackets indicate a significant gesture
[overlap]	overlapping sign/speak is indicated by brackets

TWO PARTS one sign by a combination of two signs

utterance is in spoken Swedish
utterance is in CAPITAL LETTERS

References

- Ahlgren, Inger & Brita Bergman. 2006. Det svenska teckenspråket. SOU 2006:29 [The Swedish sign language. SOU 2006:29]. In Government Offices of Sweden (ed.), *Teckenspråk och teckenspråkiga. SOU 2006:29 [Sign Language and signers. SOU 2006:29]*, 11–70. Stockholm: Government Offices of Sweden.
- Auer, Peter. 1998. *Code-switching in conversation: Language, interaction and identity*. London: Routledge.
- Dively, Valerie. 1998. Conversational repairs in ASL. In Ceil Lucas (ed.), *Pinky extension and eye gaze: Language use in deaf communities*, 137–169. Washington, DC: Gallaudet University Press.
- Grosjean, François. 1982. *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.
- Haualand, Hilde. 2011. Interpreted ideals and relayed rights: Video interpreting services as objects of politics. *Disability Studies Quarterly* 31(4). <https://dsq-sds.org/article/view/1721/1769> (accessed 17 July 2020).
- Heritage, John. 1998. Oh-prefaced responses to inquiry. *Language in Society* 27(3). 291–334.
- Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. In Gene H. Lerner (ed.), *Conversation analysis: Studies from the first generation*, 13–31. Amsterdam: John Benjamins.
- Keating, Elizabeth & Chiho Sunakawa. 2013. “A full inspiration tray”: Multimodality across real and virtualspaces. In Jürgen Streck, Charles Goodwin & Curtis LeBaron (eds.), *Embodied interaction: Language and body in the material world*, 194–204. Cambridge: Cambridge University Press.
- Keating, Elizabeth, Terra Edwards & Gene Mirus. 2008. Cybersign and new proximities: Impacts of new communication technologies on space and language. *Journal of Pragmatics* 40(6). 1067–1081.
- Linell, Per. 2004. *The written language bias in linguistics: Its nature, origins and transformations*. London: Routledge.
- Marks, Annie. 2018. Hold the phone! Turn management strategies and techniques in Video Relay Service interpreted interaction. *Translation & Interpreting Studies: The Journal of the American Translation & Interpreting Studies Association* 13(1). 89–111.
- Mesch, Johanna & Krister Schönström. 2018. From design and collection to annotation of a learner corpus of sign language. In Mayumi Bono, Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, Julie Hochgesang, Jette Kristoffersen, Johanna Mesch & Yutaka Osugi (eds.), *Proceedings of the 8th workshop on the representation and processing of Sign Languages: involving the language community [Language Resources and Evaluation Conference (LREC)]*, 121–126. Paris: European Language Resources Association (ELRA).
- Mesch, Johanna & Lars Wallin. 2015. Gloss annotations in the Swedish Sign Language corpus. *International Journal of Corpus Linguistics* 20(1). 102–120.
- Mondada, Lorenza. 2006. Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In Hubert Knoblauch, Bernt Schnettler, Jürgen Raab &

- Hans-Georg Soeffner (eds.), *Video analysis: Methodology and methods*, 51–68. Frankfurt: Peter Lang.
- Russell, Debra. 2005. Consecutive and simultaneous interpreting. In Terry Janzen (ed.), *Topics in signed language interpreting*, 135–164. Amsterdam: John Benjamins.
- Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language in Society* 50(4). 696–735.
- Schegloff, Emanuel A., Gail Jefferson & Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53(1–2). 361–382.
- Sidnell, Jack. 2010. *Conversation analysis: An introduction*. Chichester, U.K.: Wiley-Blackwell.
- Sidnell, Jack & Tanya Stivers. 2013. *The handbook of conversation analysis*, vol. 121. Chichester, West Sussex: John Wiley & Sons.
- Stivers, Tanya & Jeffrey D. Robinson. 2006. A preference for progressivity in interactions. *Language in Society* 35(3). 367–392.
- Svennevig, Jan. 2010. Pre-empting reference problems in conversation. *Language in Society* 39(2). 173–202.
- Wallin, Camilla & Johanna, Mesch. 2018. *Annoteringskonventioner för teckenspråkstexter: Version 7*. [Annotation conventions for sign language texts: Version 7]. Stockholm: Department of Linguistics. Stockholm University.
- Warnicke, Camilla. 2018. The co-creation of communicative projects within the Swedish Video Relay Service (VRS). In Jemina Napier, Robert Skinner & Sabine Braun (eds.), *Here or there: Research on interpreting via video link*, 210–229. Washington: Gallaudet University Press.
- Warnicke, Camilla. 2019. Equal access to make emergency calls: A case for equal rights for deaf citizens in Norway and Sweden. *Social Inclusion* 7(1). 173–179.
- Warnicke, Camilla & Charlotta Plejert. 2012. Turn-organisation in mediated phone interaction using Video Relay Service (VRS). *Journal of Pragmatics* 44(10). 1313–1334.
- Warnicke, Camilla & Charlotta Plejert. 2016. The positioning and bimodal mediation of the interpreter in a Video Relay Interpreting (VRI) service setting. *Interpreting* 18(2). 198–230.
- Warnicke, Camilla & Charlotta Plejert. 2018. The headset as an interactional resource in video relay interpreting (VRI). *Interpreting* 20(2). 285–308.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/text-2019-0174>).

Bionotes

Camilla Warnicke

University Health Care Research Center, Faculty of Medicine and Health, Örebro University, Sweden

camilla.warnicke@oru.se

Camilla Warnicke holds a PhD and works at the University Health Care Research Centre (UFC) in Örebro County, Sweden. She is affiliated with the School of Health Sciences at Örebro University, Sweden. She is a certified interpreter of (spoken) Swedish and Swedish Sign Language. She also works as a trainer in the interpreter programme in Örebro. Her research interests are related to

accessibility and interaction in interpreted encounters (spoken/signed language interpreting), and she is primarily conducting her work within the frameworks of Conversation Analysis and dialogical theory.

Charlotta Plejert

Department of Clinical and Experimental Medicine, Division of Speech Language Pathology,
Linköping University, Linköping, Sweden

charlotta.plejert@liu.se

Charlotta Plejert, PhD, is an Associate Professor of Linguistics at Linköping University, Sweden. Her primary research interests are language and interaction in dementia, multilingualism and dementia; children and adults with speech and language impairments; atypical interaction; and interpreting within and outside of healthcare settings. She is the main editor of the volume *Multilingual Interaction and Dementia* (Plejert, Lindholm & Schrauf, 2017, Multilingual Matters) and co-editor of *International Journal of Interactional Research in Communication Disorders* (<https://journals.equinoxpub.com/IIRCD/index>).