Interpreting in small-group bilingual videoconferences: Challenges and adaptation processes

Sabine Braun, University of Surrey

Abstract

When interpreting takes place in a videoconference setting, the intrinsic technological challenges and the very remoteness of the interpreters' location compound the complexity of the task. Existing research on remote interpreting and the problems it entails focusses on remote conference interpreting, in which the interpreters are physically separated from the conference site while the primary interlocutors are together on site as usual. In an effort to broaden the scope of research in the area of remote interpreting to include other types and to address other questions, in particular that of the interpreters' adaptability to new working conditions, this paper analyses small-group videoconferences in which the primary interlocutors as well as the interpreters all work from different locations. The findings from an empirical case study (based on recordings of videoconference sessions as well as introspective data) are used to identify and exemplify different types of interpreter adaptation.

Keywords: videoconference interpreting, remote interpreting, simultaneous interpreting, dialogue interpreting, discourse comprehension and production, monitoring, adaptation, motivation, working conditions

1 Introduction

Interpreting is a highly complex cognitive and communicative task. When it takes place in a videoconference setting, the technological challenges and remoteness add yet another dimension of complexity and difficulty. The term videoconference (VC) is used here to cover any form of video-mediated remote communication. Different types of technology are available for videoconferencing (ISDN¹, Internet, Satellite), and they can in principle support a whole range of communicative situations, including both dyadic communication (e.g. peer-to-peer and small-group communication as well as panel discussions, workshops and seminars) and monologic or unidirectional communication (e.g. lectures, conferences or TV broadcasts).

The use of VC technology in connection with interpreting has two principal motivations. On the one hand, interpreter mediation may be required in any of the above-mentioned communicative events, taking place via VC connection because the primary participants (i.e. the speaker and the audience in monologic communication or the interlocutors in dyadic communication) are geographically separated. In this scenario, there are in principle three ways to integrate one or several interpreters: they can be located at one or all of the primary participants' sites, or at a separate site. On the other hand, VC technology can be used to support situations in which the interpreters are physically separated from the primary participants, who are together on site. While both scenarios allow for consecutive interpreting, some constellations also allow for simultaneous, especially where the interpreter is separated from the primary participants and additional equipment is available (e.g. additional sound channels in the VC connection and headsets for the primary participants).

In practice this leads to a wide variety of possible remote interpreting scenarios combining different types of technology, distributions of primary participants and interpreters, and modes of interpreting (cf. Braun 2004; Daly 1985; O'Hagan 1996). When analysing remote interpreting, it is important to emphasise that only some characteristics, such as the general impact of remoteness, are

common to all of the possible scenarios, while others are specific to particular subsets thereof (e.g. turn-taking problems in VC-based conversations).

To date, research on remote interpreting has focussed on one particular scenario, namely that of *remote conference interpreting*, where all primary participants (of a multilingual conference, meeting or debate) are in the same room and the interpreters work from a separate location. In an early experiment of this type, carried out by the European Telecommunications Standards Institute (ETSI) in co-operation with AIIC in 1992, a conference situation was simulated by presenting pre-recorded video material (speeches, debates) to the participating interpreters via an ISDN connection (cf. Böcker & Anderson 1993). In a joint project of the International Telecommunication Union (ITU) and the École de Traduction et d'Interprétation at the University of Geneva (ETI) in 1999 the technical feasibility of remote conference interpreting and its impact on human factors such as stress and fatigue were compared with traditional conference interpreting in a controlled-experiment, also using ISDN-based VC technology for the remote interpreting part of the experiment (cf. Moser-Mercer 2003). Since then a number of further experiments have been carried out in UN and EU institutions. Like the ITU/ETI study, these experiments involved remote interpreting of live events (in contrast to the use of pre-recorded material, as in the ETSI study), but they were partially carried out with proprietary VC technology, providing better sound and image quality than ISDN. An overview of the various studies, focussing on the technological setups used, is given by Mouzourakis (2006). Moser-Mercer (2005) provides a meta-analysis of the outcomes in terms of fatigue, motivation, role of vision and other factors.²

What emerges from these studies is a tentative overall conclusion that even under extremely good technical conditions, conference interpreters who are not in the same location as the speakers experience more fatigue and stress and have a number of physiological and psychological complaints. While Böcker & Anderson (1993) and other earlier studies reported in Moser-Mercer (2005) and Mouzourakis (2006) attributed such negative outcomes mainly to inferior technical conditions, Mouzourakis concludes that

it has become clear that interpreter complaints were not only due to the inferior technological conditions, but also the result of a number of physiological (sore eyes, back and neck pain, headaches, nausea) and psychological complaints (loss of concentration and motivation, feeling of alienation) stemming from the remote interpreting conditions. These complaints resurfaced in subsequent experiments, conducted in a variety of technical conditions and by a number of multilingual organisations; it would thus be difficult to attribute them solely to a particular technical setup or even to the working conditions provided by a particular organisation. (Mouzourakis 2006: 52)

This view partially coincides with findings from studies on monolingual communication via VC which suggest that remote work is more difficult because there is “more sense of what others are doing and what they mean when we are face-to-face than can be presented via even very good video channels.” (Olson et al. 1997: 170). A more detailed look at the findings of the above experiments, however, suggests that feelings of remoteness, alienation and loss of motivation, while prevailing regardless of the *technical quality* of the communication channels, are more closely linked to some of the *experimental conditions* than to others, especially to the viewing conditions (view of meeting room vs. view of the speaker, cf. Moser-Mercer 2005: 733; Mouzourakis 2006: 53). Moreover, the experimental conditions themselves – involving e.g. questionnaire surveillance, stress hormone tests etc. – may have

created a less than favourable framework for interpreting in this difficult as well as altogether unfamiliar situation.

The above experiments have addressed the multilingual communication needs of international organisations. This becomes most obvious in the more recent experiments (in the EU Parliament), which were a consequence of the EU enlargement and the increasing number of interpreters required, along with a shortage of interpreting booths in EU Parliament meeting rooms. Hence, as critically reported by Mouzourakis (2003, 2006), the idea of ‘outsourcing’ the interpreters from the main conferencing room gained momentum.

By contrast, the use of VC technology to enable *remote dialogue interpreting* (public service/liaison interpreting) – e.g. to mediate bilingual court hearings or medical conversations through an interpreter who is not on site – has only recently come under consideration (cf. Niska 2002), though an early study on the remote interpretation of medical conversations using an *audio* link was carried out by Hornberger et al. (1996). Comparing remote simultaneous interpreting with onsite consecutive interpreting, Hornberger et al. found the former to be not only preferred by the primary participants but also more complete and accurate. The interpreters, while preferring to work on site, felt that the interlocutors would benefit from the simultaneous mode and the interpreter’s remoteness.

The underlying idea of all the scenarios outlined above has been to link an interpreter from a remote site to an otherwise traditional communicative event. My own interest in remote interpreting has a different background. On the one hand, I wanted to investigate the *integration of an interpreter into a VC*. The motivation for this was to support small and medium-sized enterprises (SMEs) in their growing need for bi/multilingual communication with partners and clients at distant locations. In such situations interpretation has traditionally been delivered in the form of telephone interpreting (cf. Oviatt & Cohen 1992, Wadensjö 1999), but the introduction of affordable (ISDN- and web-based) VC technology has put the integration of an interpreter into VC-based ‘tele-meetings’ on the agenda. On the other hand, my initial observations of interpreting in VC-based conversations had revealed a number of adaptation processes, i.e. the development of novel strategies and/or the activation and application of strategies most appropriate in the context of the situation (including problem-solving strategies). This, in turn, made it worthwhile to investigate how such adaptation processes take place.

This paper focuses on the integration of an interpreter into a conversation via VC, and on the interpreter’s adaptation potential in this scenario. The specific nature of this scenario will be described in section 2. In section 3 a process-oriented model of communication will be outlined which accounts for the specifics of VC-based communication as well as those of dialogue interpreting, serving as a basis for the analysis of adaptation processes which will be presented in section 4.

2 Communicating and interpreting in small-group videoconferences

For a better understanding of the specifics of interpreting in VC-based conversations, it is helpful to first of all take a look at VC-based conversations themselves. I will do this briefly in section 2.1 before describing the solution to integrating an interpreter into a VC, adopted in the study described here. In section 2.2 the specific challenges of VC interpreting in this setup will be outlined.

2.1 From monolingual to bilingual videoconferences

The simplest and perhaps most common way of using VC technology is a so-called peer-to-peer VC involving two participating sites. Peer-to-peer VCs are typically done with a PC (or TV-size VC equipment) and allow for a small group of interlocutors to participate at each site. With the most basic equipment (PC and Webcam with fixed camera position) no more than one to two interlocutors per site can be involved if visibility and clarity are to be maintained, whereas more sophisticated hardware (especially mobile and high-resolution cameras, larger screens) offer more possibilities. Alternatively, a VC can be conducted as a multipoint conference connecting more than two sites.

The great achievement of VC technology is that it enables people who are geographically separated to communicate in real-time, as on the telephone, but with the added benefit of seeing each other and getting a glimpse of their respective communicative environments. However, the interlocutors in a VC have to rely on technical channels for communication, and the quality of these affects the way in which the interlocutors and the communicative situation are perceived. Firstly, in the equipment used for small-group VCs the video image tends to be small, incomplete and/or sketchy. When there is more than one interlocutor per site, speaker detection may be a problem. Second, direct eye contact is difficult to achieve, because the interlocutors in a VC must make a choice between looking at the screen (to see the video image of their remote interlocutors) or looking into the camera (to give the remote interlocutors the impression that they are being looked at). Third, the sound quality can be problematic, too, depending on the hardware and the transmission protocols used. Fourth, the interlocutors are exposed to different influences at their respective sites (e.g. background noise or disruptions). As a result of different environments and impediments in the visual and aural perception of the others, the social relationships between the interlocutors in a VC are usually weaker than in traditional face-to-face communication, leading to a feeling of reduced 'social presence' (cf. Short et al. 1976). It manifests itself, for instance, in incoherent discourse, as the interlocutors seem less focussed on what they want to say (cf. Braun 2004).

One factor which has received less attention is the potential distortion of the sequential perception in conversational exchanges via VC or the possibility of its being different at each site because of an audio/video transmission delay. In fact, even the slightest delay (according to Kraut & Fish (1997), one as short as 0.2 seconds) may impede the interaction (cf. Braun et al. 1999; O'Connell & Whittacker 1997). On the one hand, a transmission delay creates the impression of long pauses before the arrival of replies from remote interlocutors, which in turn creates uncertainty and produces the wrong signals (such hesitations usually being interpreted as inability to produce an appropriate reply or as disagreement; cf. Clark 1996). On the other hand, attempts to resolve 'deadlock' situations created by long pauses frequently result in overlapping speech, e.g. when an interlocutor starts to restate his/her question just as a delayed reply from the remote site arrives. Furthermore, a transmission delay causes backchannel signals to lose their effect or even to become a disturbance. Thus the interactional problems entail difficulties in discourse comprehension and production. They also reinforce the feeling of reduced social presence. As Olson, Olson and Maeder have observed, remote group work is all in all more laborious than face-to-face group work because it requires "greater process overhead" (1997: 170), i.e. the interlocutors spend more time coordinating the interaction and clarifying what they meant.

A *bilingual* VC of this type creates a number of additional problems. Due to the novelty of this scenario, no standard practice has as yet been established for integrating an interpreter into a VC-based

conversation. As pointed out in section 1, the use of VC technology in connection with interpreting has so far focused on settings where only the interpreter is at a different location, e.g. in court or healthcare settings as well as in conference scenarios. If in contrast the interlocutors are geographically separated from each other (the scenario I was interested in), the main question is how an interpreter can best be integrated, a question which boils down to whether it is possible to re-create the conditions for traditional dialogue interpreting or whether a specific solution is required or at least more appropriate.

In dialogue interpreting the interpreter is traditionally part of the group of interlocutors, and the delivery mode is mainly consecutive. In the interest of a smooth flow of the conversation, the interpreter usually keeps pauses between the end of a speaker's turn and the delivery of the target text to a minimum and sometimes even starts his/her rendition while the speaker is still completing his/her turn. Ideally the interpreter should be visible for all interlocutors and should be able to maintain eye contact with the interlocutors to handle turn-taking in this tripartite communication situation.

In connection with the study described in this paper, two types of initial trials were carried out. In one, an interpreter took part in a peer-to-peer VC and interpreted from one of the sites, working in the consecutive mode. In another, four multipoint VCs were conducted, with the interlocutors at two sites and an interpreter at a third, also working in the consecutive mode.³ In both settings it turned out that many important conditions for dialogue interpreting could not be met. In the former setting in particular, visibility of the interpreter and eye contact with all interlocutors were very difficult to achieve. Moreover, any kind of parallel speech of two speakers (i.e. also of an interlocutor and the interpreter) caused disruption and incomprehensibility. The interpreters were thus strictly confined to the consecutive mode. With the pauses added by the audio transmission delay, conversations became very cumbersome.

In view of these trial results it seemed to be more appropriate to adjust some of the features of traditional dialogue interpreting so as to allow for interpreting the VC conversations simultaneously. This solution required a dedicated VC workstation for the interpreter, similar to an interpreting booth. At the same time, compatibility considerations called for a solution based on standard, commercially available VC equipment at the interlocutors' sites. The solution adopted in this study is outlined in Figure 1.

Figure 1. Bilingual VC setup

The audio channel setup was such that a) the source text produced by the interlocutor in the speaker role was transmitted to the interpreter's station and b) the target text produced by the interpreter was transmitted from the interpreter's station to the interlocutor in the hearer role. At all times, there was furthermore an open backchannel over which both the interpreter and the speaker were able to receive

any backchannel comments made by the hearer. The interpreter's VC workstation included a (software) console with control sliders for volume, bass and treble, a microphone mute button and an output channel selector, enabling the interpreter to switch the direction of the interpretation. The video channels were set up in such a way that the interpreter was able to see both interlocutor sites at all times (on two adjacent screens). The interlocutors saw each other but did not see the interpreter. The technical basis for this solution was an ISDN connection.⁴

The case study carried out with this solution was based on 11 simultaneously interpreted bilingual VC sessions (English<>German and French<>German) of an average length of 30 minutes. The VC sessions consisted of role-plays of peer-to-peer and small-group conversations of two types. Half were job interviews where the interviewers came from Human Resources departments of German and Austrian companies, and the candidates were freelance language trainers (British native speakers) who were asked to apply for a job as language trainer, and students applying for various other jobs. The other half were information-gathering sessions in which German university students talked to informants from universities in the UK in preparation for a planned exchange semester. The interpreters were trained (conference) interpreters who – with one exception – had many years of experience of all forms of interpreting. All were native speakers of German. The interpreting was done by one interpreter per VC session.

2.2 Specific problems in the chosen bilingual VC setup

While the solution outlined above provided more satisfying results in terms of social presence and smooth running of the conversation than the trials with consecutive interpreting, it was not without problems, because the specific conditions of interpreting and interpreter-mediated communication compounded the problems characteristic of monolingual VC communication.

The problems can be analysed from both perspectives: that of the interlocutors and that of the interpreters. From the interlocutors' point of view the bilingual VCs were difficult because of the combined effects of VC communication and interpreter mediation. On the one hand, the physical/geographical distance of the interlocutors combined with a linguistic and cultural distance and on the other, the transmission delay added to the interpreter's time lag, creating yet longer pauses, causing uncertainty and further reducing the flow of conversation.

From the point of view of the interpreters it is the combination of a) the conditions of VC-based communication and b) the specific requirements of interpreting that leads to a number of problems. In the scenario described here, the following four main sources of difficulty can be identified:

- The feeling of reduced social presence affected the interpreters in the same way as it affected the interlocutors. The interpreters reported that it was more difficult to relate to the interlocutors and that fatigue set in earlier than usual.
- Due to the use of an ISDN connection and related restrictions on audio transmission, the sound quality was problematic. Although the highest ISDN frequency rate (7 kHz) was used since only this rate was felt to be acceptable by the interpreters,⁵ the sound quality nevertheless varied, depending on the hardware (especially the microphones) used at the interlocutors' sites.

- The interpreters' comprehension and production task was further complicated by the VC-specific shortcomings in the interlocutors' discourse, especially incoherent utterances caused by the interlocutors' own technically conditioned communication problems described in section 2.1.
- Because of the interlocutors' technical inability to solve interaction problems, the interpreters were required to adopt the role of a moderator, which posed a number of ethical and other problems.

Two examples will be introduced here to illustrate some of the characteristic problems in the interpreted VC sessions. The examples are taken from a VC in English and German in which a job interview was simulated. In the transcripts, source text (ST) and target text (TT) are time-aligned. Each section of the transcript represents five seconds, with variation in font width indicating variation in the speech rate. The uppermost line within each section presents the ST (in capitals). The middle line shows the TT. The lowest line is reserved for backchannel utterances from the hearer (in capitals; those utterances were audible for the interpreter and the speaker). The transcription conventions are explained in the Appendix. The main speaker in example 1 is the interviewer, a Human Resources manager from an Austrian company seeking to employ a language trainer.

Example 1: GE → EN, speaker=P1 until line 10; then EN → GE, speaker P2

1	P1: » ALSO ES WIRD SO SEIN, WENN SIE FÜR UNS ARBEITEN, DANN/ EHM EIN GRUND, WARUM WIR DIESES SPRACH- I: <i>Well once you would be working</i>
2	P1: TRAINING MACHEN WOLLEN, IST, DASS WIR MITARBEITER INS AUSLAND ENTSENDEN WOLLEN, DASS WIR VERMEHRT MITAR- I: <i>for us, then/ one reason for us to do such language courses would be that we want to send</i>
3	P1: BEITER AUS DEM AUSLAND NACH ÖSTERREICH KRIEGEN WOLLEN. UND DAZU IST ES MEINER MEINUNG NACH WICHTIG, I: <i>staff members into other countries and we want to attract those people from other countries to</i>
4	P1: ANDERE KULTUREN ZU VERSTEHEN, ZU WISSEN, WIE MAN IN EINEM FREMDEN LAND/ ALSO WIE MAN SICH HELFEN KANN, WO MAN I: <i>Austria more and more. And therefore it is important, from our point of view, to understand other cultures, to</i>
5	P1: HINGEHT. UND ICH GLAUB/ ALSO ICH <u>GLAUBE NICHT</u> , SONDERN ICH HÄTT' GERN, DASS ES MIT I: <i>know how to behave in other countries, how to try and find your</i>
6	P1: MIT DEM SPRACHTRAINING MITGEHT. (2.0) WAS/ WELCHE IDEEN HABEN SIE DAZU? WAS I: <i>ways and means to cope with the situation in the respective country. And to/</i>
7	P1: KÖNNT' MAN MACHEN? I: <i>so what are your ideas against this background? How could we try and</i>

8	P1:	.	.	.	.	(9.0)
	I:	<i>apply this for our purposes?</i>				
9	P1:	<u>ALSO EIGENTLICH MEINE ICH AUCH VORBEREITUNG</u> FÜR EINE ENTSENDUNG.				
	I:	⊥	⊥ <i>Well what I mean is also preparation for being sent</i>			
	P2:	<i>RIGHT EHM</i> <i>(3.5/2.5 after end of I)</i>				
10/	P2:					» P2: OK EHM
	I:	<i>to another country for example.</i>				⊥
11	P2:	YEAH	I'VE-	I'VE DONE QUITE A BIT OF TEACHING WITH THAT ...		
	I:	<i>Ja da hab ich schon ...</i>				

One problem in this passage is that the speaker (the interviewer) is not very well focussed on what she is saying. According to the comments of the informants who were involved in this study as interlocutors, such instances, which were frequent throughout the VC sessions, were a result of the weak relationship with their remote interlocutor(s). The incoherence is partially revealed by the large number of breaks in the German ST (marked with a slash in the transcript). In contrast to purely syntactic breaks, which are a characteristic feature of spoken discourse and often occur as a consequence of replanning *how* something is said, most of the breaks in the above example are accompanied by (minor) semantic changes. Another (more serious) semantic change occurs in lines 2/3 and 9 – without any obvious marking. In lines 2– 3 the speaker corrects the reason for the planned language training in her company from sending employees abroad to recruiting people *from* abroad, without signalling this change in any way. In line 9 she reintroduces her original idea of posting people abroad.

A second problem arises from the way in which the speaker reintroduces her original thought. It is only after a 9-second pause (starting in line 7) that she finally adds it to her utterance (in line 9). Many of the informants commented that such belated amendments, which were very common in the VCs conducted in the framework of this study, were the result of a prevailing uncertainty as to whether an utterance had been clear enough. This, in turn, was triggered by the long pauses which occurred when waiting for a reply. In example 1 the interviewer was obviously irritated by not receiving a reply for 9 seconds. That this has purely ‘technical’ reasons becomes clear on closer inspection of the timing, but was not normally transparent for the interlocutors in the VC at the time of conversing. One part of the 9-second-long pause is created by the interpreter’s time lag (here approx. 5.5 seconds, see lines 7-8)⁶ and another part by the transmission delay (0.5 seconds in each direction). This means that the applicant had actually started to reply within a relatively ordinary pause of 2.5 seconds after receiving the end of the interpreted question at her site.

The belated amendments usually had very disturbing effects. In example 1 the applicant had already begun to reply (*right uhm*, line 9) when the amendment was made, to the effect that the applicant’s and the interviewer’s utterances overlapped at the interpreter’s site (marked in italics in line 9). The interpreter, who had already switched the audio channel⁷ and was expecting to interpret

the applicant's reply from English into German needed to switch back – in a literal and figurative sense. The applicant needed to wait, and it took approx. 8 seconds until the situation was resolved and she finally succeeded in giving her reply (end of line 10).

In example 2 a similar instance of overlapping speech had even more serious consequences. On the one hand, the applicant (who speaks first, in lines 1-2) had to relinquish the floor to the interviewer. More importantly the interpreter's comprehension was thoroughly disturbed by the overlapping speech. He was only able to save the situation by rounding off the utterance with a very general comment.

Example 2: EN → GE, speaker=P2 until line 3; then GE → EN, speaker=P1

1	P2: ... I ALSO TEACH QUITE LARGE GROUPS, LARGE MEANING ANYTHING UP TO TEN STUDENTS. EHM	
	I: <i>Dann unterrichte ich auch relativ große Gruppen, das heißt bis zu zehn</i>	
2	P2: I LIKE CLASSES OF ALL SIZES FOR DIFFERENT REASONS. I FIND I CAN/	
	I: <i>Teilnehmern. Ich mag eigentlich alle Größen, ich hab da vier unterschied-</i>	
	P1: <i>Ahm, und</i>	
3	/P1: YES?	» P1: OK. WAS ...
	I: <i>liche/ und ich denke, das kann man auch sehr unterschiedlich handhaben.</i>	

Interpreting is in itself a cognitively and communicatively complex task requiring, amongst many other things, careful management of the interpreter's cognitive processing resources (cf. Gile 1991). In view of the additional challenges outlined above it may not be too surprising that the task of remote interpreting is more stressful than traditional interpreting, as was reported by the informants of this study and found by Moser-Mercer's (2003) comparative study.

However, apart from the problems characteristic of the overall situation, the findings of the case study described here also suggest that the interpreters and the interlocutors tried to adapt to this new and unfamiliar form of communication, and that with experience and practice, they acquired a better understanding of how to handle and exploit the communication channels available in the VC situation.

In accordance with this, my primary interest was to find out more about *how* the adaptation takes place in the scenario described here. My hypothesis was that whilst any act of communication (including interpreting) takes place under specific situational and/or technical conditions and whilst these are particularly prominent when the communicative situation is new and unfamiliar, communication (including interpreting) has a strategic orientation, and human communicative competence includes a basic *capacity for adaptation* to new, unfamiliar and even 'adverse' communicative situations. The model of communication adopted for the purposes of the analysis will be outlined in the next section.

3 A process-oriented model of communication and interpreting

To analyse the adaptation processes taking place in bilingual VC conversations, I have adopted a process-oriented model of communication in which communication is seen as a linguistic and cognitive process of discourse comprehension and production (cf. e.g. Brown & Yule (1983) and van Dijk & Kintsch (1983) from a discourse analytical perspective; Blakemore (1992) and Sperber & Wilson (1995) from a pragmatic perspective).

In *discourse comprehension* hearers aim to create a (coherent) mental model of the talk of others (cf. van Dijk & Kintsch 1983; Johnson-Laird 1983). They do this by forming hypotheses about the speaker's meaning and intentions, using their perception and the linguistic cues provided together with their prior knowledge (in the widest sense, i.e. linguistic, encyclopaedic, cultural), their situation-specific expectations and their ability to draw inferences (cf. Brown & Yule 1983; Wilson & Sperber 2004). To channel and delimit their cognitive processing efforts, hearers furthermore rely on 'pragmatic principles' of communication such as the Principle of Relevance suggested by Sperber & Wilson (1995).

Two characteristics of the comprehension process are crucial in the context of the present paper. Firstly, the hearer's ability to combine complementary resources for comprehension enables him or her, according to Brown & Yule (1983), to process utterances 'online', i.e. to start processing an utterance as soon as a communicative stimulus is received, but it also creates the need for a continuing revision of initial comprehension results in the light of subsequent utterances. Secondly, any comprehension result is of a hypothetical nature, and Brown and Yule claim that hearers (only) continue processing until they have formed a hypothesis that they deem stable enough for the purposes of a particular situation. This enables them to process utterances with varying depth, according to the requirements of the situation. I will return to both aspects below.

In *discourse production* speakers aim to enable the hearer to build a mental model on the basis of their utterances. Speakers therefore shape their utterances in such a way that their intended meanings are easily accessible for the hearer (termed *recipient design* by Sacks et al. 1974), e.g. by making assumptions about the knowledge they share with a particular hearer (termed *grounding* by Clark 1996) and about the hearer's ability to draw inferences. In dyadic communication, the production of utterances also needs to be co-ordinated, which is why interaction strategies play an important role here as well (cf. Clark 1996).

The model outlined here can accommodate both monolingual and interpreter-mediated communication. Interpreting involves comprehension and production under specific conditions. In simultaneous interpreting, which is of particular relevance to the VC scenario considered in this paper, the interpreter builds up a mental model on the basis of incoming ST segments and relays it in the target language according to the linguistic, discourse and cultural conventions of the target audience. As Kalina (1998) and Kohn and Kalina (1996) point out, this entails a number of peculiarities with regard to both comprehension and production, most prominently the interpreter's need to gain a deeper understanding of the ST than that of an ordinary hearer, and the interpreter's lack of semantic autonomy in TT production. All of this requires the interpreter to develop specific comprehension and production strategies (for a discussion of these, cf. Kalina 1998).

The core of the model proposed here, and a crucial point with regard to analysing adaptation processes, is a speaker's/hearer's and interpreter's use of **monitoring strategies** to control and if necessary revise their communicative performance. Monitoring has usually been discussed in

connection with speech and/or discourse production and has been portrayed as a device for output control or as the ‘component’ of production responsible for correction or repair (cf. Laver (1970, 1973, 1980) in a neurolinguistic framework; Levelt (1983, 1989) in a psycholinguistic framework; Krashen (1981) in second-language acquisition research; Petite (2005) in relation to interpreting). Following Morrison & Low (1983) and Kohn (1990), I have argued in Braun (2004) that monitoring is a more comprehensive process, controlling both production and comprehension performance.⁸ As Levelt (1983) shows, monitoring is closely linked to the ‘online’ nature of communication. This is most obvious in spoken production, but as was pointed out above, comprehension works ‘online’ as well; it starts upon first perceiving an utterance. Therefore the model suggested here presumes that processes which are similar to those of ‘output control’ also take place in the comprehension process, even if less directly visible, allowing hearers (and interpreters) to assess (and revise) their comprehension hypotheses.

Monitoring in this wider sense then can best be described as an evaluative process in which performance is assessed against a set of requirements, using different types of resources. Kohn (1990) emphasises that communicators have no access to ‘objective’ knowledge or norms against which to evaluate their performance. As Kohn showed empirically (for language learners’ production), the assessment is based on individual requirements which are established for one’s performance in a particular communicative situation (e.g. accurateness, depth and completeness of comprehension; accurateness, appropriateness, fluency and correctness of production). They are shaped by one’s communicative goals and intentions, the presumed knowledge and expectations of the others involved in the communicative situation and by socio-cultural conventions, technical and other constraints of the communicative situation, but it is important to note that they are individual and subjective.

In order to evaluate whether their actual performance matches their requirements, communicators use a variety of resources. Based on different types of resources, three monitoring ‘cycles’ can be distinguished in oral and dyadic communication: a) monitoring on the basis of one’s own linguistic, encyclopaedic and socio-cultural knowledge, b) on the basis of verbal and non-verbal feedback from the interlocutors, and c) on the basis of the co-text and subsequent utterances from the other interlocutors. These resources are not different in principle from the resources used for the comprehension and production of utterances in the first place. What is, however, crucial and was also shown in Kohn’s empirical research is that communicators are able to apply them with increased intensity in the monitoring process, once they focus their attention on a particular aspect of production or comprehension.⁹ This can happen because they have noticed a problem in their performance (e.g. a lack of comprehension) or because their requirements are such that they constantly focus attention on a particular performance aspect (e.g. fluency of TT production, appropriateness of choice of expressions in production). With regard to interpreting, Petite (2005), for example, found that interpreters not only corrected errors but also tried to improve the appropriateness of their output.

To satisfy different requirements, monitoring takes place in different phases of comprehension and production (cf. also Kohn 1990; Levelt 1983):

- *Anticipatory monitoring* occurs in connection with utterance planning in production and with the generation of expectations in comprehension.
- *Simultaneous monitoring* is going on while a message is encoded/decoded.
- *Retrospective monitoring* takes place after an utterance has been produced and after an initial comprehension hypothesis has been formed.

4 Monitoring and adaptation

I will begin the analysis with an example of simultaneous comprehension monitoring. It is taken from one of the job interview VCs. The speaker is the interviewer, a Human Resources manager of a medium-sized German company.

1	P1:	... WIR HABEN ABER INSGESAMT	VIERHUNDERT	ANGESTELLTE.
	I:		...	but
2	P1:	DAS HEIßT, ES GIBT NOCH EINEN ANDEREN STANDORT,	WO DER REST DANN	EHM
	I:	ehm all in all we have four hundred employees		ehm
3	P1:	ARBEITET,	UND DIE	EHM SOFTWAREGRUPPE BESTEHT
	I:	at another site	in Germany.	
4	P1:	INSGESAMT AUS FÜNFZEHN MITARBEITERN.	DIESE MITARBEITER SIND ALLE	
	I:	And	well your working group would	

I will focus here on the decision to translate *Softwaregruppe* (*software group*) with *your working group*. This interpreter commented on his decision as follows:

Interpreter's comment

Ich habe es nicht verstanden. Ich hab nur 'group' gehört, und dann musste ich auch erst mal kurz überlegen. 'Group' – könnte auch sein, dass sie sich jetzt irgendwie mit Anglizismen ausdrückt, könnte der Konzern sein. Aber das waren jetzt 15 Leute. Also du musst dann auch praktisch einen langen Abstand lassen, damit du dann nicht plötzlich sagst 'Konzern, fünfzehn, nee ...'.

[I didn't understand it. I only heard 'group', and then I just had to think about it. 'Group' could have meant that she was using an Anglicism, or it could have been the company. But then she talked about 15 people. Well, you have to stay behind quite a bit to make sure you don't say 'company – 15 people – nope'.]

The first point of interest is what the comment (together with the time-aligned transcript) reveals about the monitoring process itself. According to the comment by the interpreter his initial comprehension hypothesis was only tenuous (he heard /group/ in the German (!) source text and hypothesised it might be an Anglicism to denote the entire company). Thus, the initial comprehension hypothesis obviously did not match the interpreter's requirements with regard to the completeness of ST comprehension. As his comment shows further, he tried to improve his comprehension performance by activating additional resources. In particular, he exploited:

- the co-text of the utterance (he heard the ST segment indicating that the 'group' in question has only 15 employees before presenting his TT solution);
- his world knowledge (he knew that a 'group' in the sense of 'company' has a very large number of employees);
- his ability to draw inferences (if the 'group' in question has only 15 employees, it cannot mean 'company');
- his expectations about the meaning of the utterance (the applicant's prior question was about the unit in which he would be working).

Hence, what appears to be a generalisation in the TT (for reasons which do not become clear from an analysis of ST and TT alone) is in fact the result of simultaneous monitoring in the course of which an initial lack of comprehension was first identified and then overcome through the activation and use of additional cognitive and linguistic resources. The combination of these resources enabled the interpreter to come up with the solution *your working group*. From the interpreter's point of view this was clearly an improvement on his initial hypothesis.

The second interesting point is that this example provides some cues as to how monitoring processes led to adapted strategies. In his comment the interpreter not only indicated that he had activated additional knowledge for comprehension; he also pointed out that his time lag in the TT production needed to be rather long in order for him to be able to exploit resources such as subsequent ST segments (co-text). Both the intensified use of available resources for ST comprehension (which becomes apparent in the interpreters' comment) and a *consistently* increased time lag in the TT

production (which is very noticeable across all my data, and which is also known from other ‘difficult’ interpreting situations) can thus be regarded as adapted strategies. They emerge from simultaneous monitoring of initial comprehension hypotheses and the recurrent need to overcome listening comprehension problems.

Since ST comprehension is crucial for TT production in any interpreting situation, it follows that comprehension monitoring will be a very far-reaching process. While the above example demonstrated how it contributes to the development of adapted comprehension strategies, the following example will illustrate its role in creating adapted *production* strategies. This example is taken from a job interview with a British student looking for a student job in Germany. The applicant talks about her aims and her previous work experience.

Example 4: EN → GE, speaker=P2

1	P2:	... I'M LOOKING FOR EHM ANY POSSIBLE OPPORTUNITY	
	I:	... Und ich	
2	P2:	TO IMPROVE MY GERMAN, BUT I'M ALSO A BIT DESPERATE FOR MONEY EHM	
	I:	versuche alle Möglichkeiten wahrzunehmen um mein Deutsch zu verbessern, aber ich möchte auch ein	
3	P2:	MAYBE LOOKING FOR A JOB EHM WAITRESSING OR HELPING OUT IN ANY SCHOOL. ANYTHING	
	I:	bißchen Geld verdienen und suche deswegen nach einer Arbeit, vielleicht	
4	P2:	THAT'D HELP ME IMPROVE MY GERMAN. EHM I'VE HELPED OUT IN A SCHOOL BEFORE,	
	I:	eh als Bedienung oder irgendwo, wo ich meine Deutsch-	
5	P2:	EHM WITH CHILDREN AGED FIVE TO SIX, I'VE DONE WAITRESSING, AND	
	I:	kenntnisse verbessern kann. Ich kann/ ich hab auch schon als Kellnerin gearbeitet, ich	
6	P2:	SEVERAL JOBS, WORKED IN A SUPERMARKET ...	
	I:	habe auch schon Nachhilfeerfahrung, ich hab auch schon mal im Supermarkt ...	

The focus here is on the omission of *children aged five or six* by the interpreter (see ST line 5).¹⁰ The interpreter's comment makes it clear that the omission of the children's age was not caused by a lack of comprehension here but was a result of a strategic decision not to include this information in the TT:

Interpreter's comment

Also die Altersgruppe, die hab' ich zwar gehört, aber dann untern Tisch fallen lassen, weil ich zu dem Zeitpunkt noch Probleme hatte, ich musste mich erst noch ein bisschen 'reinhören in ihre Art und Weise zu sprechen, und ich hab' die Altersangabe dann untern Tisch fallen lassen, weil ich gedacht hab, die Hauptaussage, die ist rübergekommen, sie hat Nachhilfe gegeben.

[Well, I did hear the age range but I left it out because at that point I still had some problems, I had to get used to her way of speaking. So, I skipped the information about the age range because I thought that the main point – she had provided some private tuition – had come across.]

The interpreter's comment suggests that he had had general difficulties in understanding the speaker here. The time-aligned transcript shows that the interpreter reduced the simultaneity of ST comprehension and TT production, i.e. he (systematically) separated listening to ST segments from delivering TT segments by frequently pausing in the TT production while, in turn, using small pauses in the ST for delivering TT segments (see lines 4 to 6). The interpreters' comments on other VC sessions confirm that this strategy was frequently used to free up resources for comprehension. To apply this strategy successfully, the TT had to be condensed at times in order to fit into the short pauses in the ST. Hence, a strategy which originally resulted from *comprehension monitoring* led here to an adapted strategy in the TT *production*.

This, in turn, had consequences for *production monitoring*, because the condensation of the TT became an additional production requirement against which the production performance had to be evaluated. Condensation in the TT involves production planning, i.e. it primarily concerned anticipatory production monitoring. The interpreter had to assess the relative importance of the various pieces of information in the utterance context before producing the TT. In many cases this led the interpreters to simply omit what they considered to be non-essential details. In the above example, the omission was combined with a strategic rearrangement of the order of individual segments. The interpreter delivered the unproblematic information about the speaker's waitressing experience before the information about her work in a school (see lines 4-6; the segments in question are framed). A likely explanation is that the interpreter withheld the information about the speaker's work in a school until he could be sure that the children's age was not very relevant here.¹¹

By analogy to example 3 (*your working group*), in which an apparent generalisation in the TT turned out to be the result of simultaneous comprehension monitoring, it can be assumed for the above example that the apparent omission in the TT was a result of anticipatory production monitoring, in the course of which the interpreter established the 'local' production requirement to condense the TT, and was able to fulfil it. In other cases initial production requirements were spontaneously abandoned in order to carry out the interpreting task as a whole. The interpreters involved in this study pointed out, for instance, that they readily accepted the 'second-best' TT solution where the most appropriate solution, e.g. a particular lexical item, could not be activated immediately, in order to deliver the TT quickly and free up time and resources for comprehension. Both the anticipatory establishment of specific 'local' requirements and the spontaneous abandonment of initial requirements for the sake of completing the interpreting task as a whole can be regarded as adapted strategies to cope with the difficult VC interpreting situation.

Overall the data suggest that in the VC sessions considered here, production monitoring was mainly triggered by comprehension monitoring and that this type of monitoring – anticipatory production monitoring – usually resulted in a revision or reduction of initial production requirements. In contrast, monitoring in relation to other aspects of TT production, e.g. to control how the TT is presented and accentuated, could not be observed. Hence it can be assumed that the focus of attention was very much on comprehension.

The only other aspect of communication in the VC conversations in which the interpreters' production monitoring came to play a significant role was the coordination of the interaction. Two major strategies developed by the interpreters to coordinate the interaction in the VC sessions can be illustrated here by one final example. Taken from another job interview with a language trainer, it shows a passage around a speaker turn. In the first part the English-speaking applicant is talking. In line 3 the interviewer tries to take the floor (cf. examples 1 and 2 for problems which are likely to arise in this situation). After the applicant's turn completion (in line 4), there is the characteristic long pause before she actually hears the interpreted version of the interviewer's next question.

Example 5: EN → GE, speaker=P2 until line 5; then GE → EN, speaker=P1

1	P2:	... FOR EXAMPLE POLITE LANGUAGE, BUT ALSO VERY PRACTICAL THINGS EHM	
	I:	... das muß ich immer wieder bedenken, und sie auch daran erinnern, was die Alltagssprache	
2	P2:	THE MOST ESSENTIAL EHM ITEMS, SO 'HOW DO- DO YOU GO	
	I:	angeht. Aber auch sehr praktische Dinge also mal ganz	
3	P2:	SHOPPING', 'HOW DO YOU EHM BUY TICKETS FOR THE TRAIN'	
	I:	wesentliche Dinge, wie man zum Beispiel einkaufen geht, wie man	
	P1:		MH JA
4	P2:	AND THINGS. IS THIS WHAT YOU MEAN?	
	I:	<u>Fahrkarten</u> kaufen kann, für den Zug und so. Ja? Also ganz grundlegende Dinge eigentlich.	
5	/P1:	» P1: MHM. WISSEN SIE AUCH BESCHIED ODER- ODER IST ES EIN VORURTEIL ODER GIBT'S	
	I:	⊥	
6	P1:	DAS WIRKLICH EHM UNTERSCHIEDLICHE MENTALITÄTEN IN DER ART, MITEINANDER UMZUGEHEN.	
	I:	Well, (3.0) what about the different men- (8.0/9.0 after end of P2)	
7	P1:	WENN ICH ZUM BEISPIEL FRANKREICH MIT AMERIKA VERGLEICHE ...	
	I:	talities, how to deal with one another ...	

In this example the interpreter used two strategies to coordinate the conversation. The first strategy served to prevent the interviewer (in the hearer role) from taking the floor before the applicant's turn was completed. When the interviewer tried to gain the floor (in line 3), the interpreter very quickly increased the pitch in the TT (beginning of line 4, underlined word Fahrkarten), thus signalling to the interviewer that the applicant's turn had not been completed yet. The second strategy is about preventing the applicant, once she had completed her turn, from amending it while waiting for a reply. Example 1 illustrated that any belated addition to, or amendment of, a turn was a potential source of serious problems in dyadic VC-based communication. In the present example the interpreter, who still needed more information before he could begin to interpret the interviewer's next question, inserted

the discourse marker *well* (in line 6), which had the pragmatic function to signal to the applicant that the interviewer had in fact taken the floor.

This strategy constitutes an interesting adaptation to the VC situation. With a view to ordinary dyadic communication Sacks, Schegloff and Jefferson (1974) have argued that such signalling is not necessary because upcoming opportunities for speaker change are predictable, e.g. on the basis of the syntactic structure of an utterance, intonation pattern or pausing. In the VC sessions analysed here the predictability was sometimes suspended due to the interpreters' time lag, the transmission delay and compromises the interpreters had to make in their TT production (cf. the reduced simultaneity which led to 'unnatural' pauses in the TT production). In this situation signalling the status of the conversation turned out to be an efficient compensation strategy. Signalling strategies result from continuous anticipatory monitoring, enabling the interpreter to plan ahead and avoid potential turn-taking problems before they could arise at all. This is confirmed by an overall comment of the interpreter on this particular VC:

Interpreter's comment

Erster Eindruck: Es lief reibungsloser als die erste [Video]Konferenz, und ich habe mich auch schon ganz bewusst bemüht, da irgendwelche Füllwörter oder Signale an den jeweils Zuhörenden zu schicken, falls da noch was nachgeschoben wurde, nachdem schon eine Aussage gemacht worden war, die man möglicherweise als abschließend ansehen könnte, damit die Zuhörenden merken, dass es noch weitergeht.

[My first impression is that it went smoother than the first [video]conference. I tried very consciously to use fillers and to send signals to the hearers when an utterance that could have been understood as completed was being continued – to make sure the hearers realise that the utterance is being continued.]

The VC in question was the second VC in which this interpreter worked. The repeated participation of interpreters in some of the VC sessions led them to refine their strategies and made them cope better with the VC situation.

5 Conclusion

In this paper I have analysed the adaptation processes of interpreters who were involved in a series of small-group bilingual VC-based conversations. This situation represents a combination of traditional dialogue interpreting and VC-specific conditions. To mediate a conversation between interlocutors at two different VC sites, the interpreters worked remotely, at a third site and from a dedicated videoconferencing station, which was similar to an interpreting booth. This scenario shares some features with remote *conference* interpreting as described in section 1, but as I pointed out there, the rationale behind it, the focus on small-group communication and the greater amount of interactivity give this scenario its own specific profile.

I have argued that monitoring plays a central role in the adaptation process. The analysis of the interpreters' performance suggests that the early discovery of a (potential or actual) problem was the key to preventing or resolving it. Thus, adaptation to the specifics of the interaction in the VC (turn-

taking) was very successful as soon as the interpreters started to rely on *anticipatory monitoring* processes to predict and avoid potential interaction problems. Anticipatory monitoring processes also enabled the interpreters to adapt the TT production to the specific requirements (or constraints) which resulted from the difficulties encountered in the comprehension process. *Simultaneous monitoring* processes helped the interpreters to cope with a number of comprehension problems. An initially tenuous comprehension hypothesis – due to a *lack* of comprehension – could frequently be consolidated or improved by activating and/or more intensely using additional resources. In contrast, as the analysis of other data from the corpus shows, problems that could only be detected in *retrospective monitoring* processes, especially the discovery of comprehension *errors*, had fewer chances of being resolved. In this regard the initial adaptation hypothesis needs to be revised or refined. The findings emerging from the analysis of the interpreters' performance are corroborated by the interpreters' more general comments on the situation. These reveal that the interpreters were keen to adapt and, in spite of all problems, were mostly optimistic with regard to the feasibility of adaptation in the VC scenario discussed in the present paper.

Two aspects of monitoring seem crucial in the adaptation process. Firstly, the flexibility arising from the different options for dealing with mismatches between initial requirements and actual performance enabled the interpreters to focus on those aspects of their performance which they regarded as crucial in this VC scenario (e.g. comprehension of the ST and management of the conversation) while paying less attention to other aspects (e.g. presentation of the TT). The possibility to modify, reduce or abandon some of their initial performance requirements often helped them to cope *spontaneously* with the unfamiliar and difficult VC situation, whereas the repeated encounter with the same problem or a similar one enabled them in some cases to resort to other, more elegant ways of adapting. Secondly, since monitoring in principle took place in all phases of comprehension and production (especially anticipatory and simultaneous monitoring), the interpreters were in some instances able not only to discover actual problems but also to anticipate potential problems and solve them or avoid them before they actually occurred. This, in turn, seems to be the key to the development of *longer-term*, adapted strategies.

While the findings of the study presented in this paper create (moderate) optimism with regard to adaptability, especially if considered in connection with improving technical conditions (first and foremost better sound quality), it is clear that they need to be carefully counter-balanced with findings from other scenarios, e.g. from the scenarios of remote conference interpreting reviewed in section 1. Irrespective of the fact that these studies have focused on the problems of remote interpreting rather than on the interpreters' adaptation potential, making a comparison of results difficult, it would certainly not be valid to simply transfer the conclusions from the small-group bilingual VCs to remote conference interpreting. Some of the adaptation strategies emerging in the VC scenario investigated here, for example, are closely linked with dyadic communication and the interpreter's role of coordinating the conversation and would not be of help in mostly unidirectional communication, where the interpreter has a more 'passive' role.

Having said that, it is equally clear that new forms of interpreting will continue to emerge (cf. also Braun, 2006). This has implications for interpreter training as well as for the shaping of future workplaces. From a training perspective, further research into *how* adaptation takes place will provide a useful basis for the design of training methods and resources which facilitate adaptation to new and emerging tasks.¹² With regard to future working conditions, it seems worthwhile to explore what

various scenarios of remote interpreting have in common, to identify the adaptation potential they share and to single out the conditions under which adaptation can best be achieved.

One key issue that needs to be addressed further in this context is that of the effects of remoteness. Moser-Mercer (2003, 2005) and Mouzourakis (2003, 2006) have reported that the remoteness leads to a feeling of alienation or ‘being left out’ and to a decline in motivation. In the scenario discussed in this paper the condition of remoteness also affected the work of the interpreters, especially through the restrictions in aural and visual perception and through knock-on effects from the interlocutors’ discourse production problems (cf. section 2.2), but it did not lead to a loss of motivation. Among the many factors which contribute to these differences (see Braun 2004: 89-99), the most important ones are perhaps a) that in the scenario discussed here the interpreter as well as each of the interlocutors were at different sites, i.e. worked remotely, so that no single participant was in a situation of ‘being left out’, and b) that the idea of working for (truly) remote clients was acceptable for the interpreters in the context of meeting the needs of small and medium-sized enterprises. This may have provided a better motivation for the remoteness conditions than a scenario where the interpreters are simply ‘banned’ from a meeting room because of a lack of space.

In any case it will be important that the findings of the study presented here and of other studies are used to support the process of creating favourable working conditions and to make new and changing work environments (including new technologies) less of a threat for those involved at the forefront.

Notes

1. ISDN (Integrated Serviced Digital Network) is the digital telephone network. It provides higher transmission speeds and better transmission quality than the old analogue telephone system. This has made ISDN attractive for Internet access and videoconferencing.
2. Detailed reports on most of these studies are not publicly available.
3. The scenario in which an interpreter is integrated at each site of a peer-to-peer VC was not trialled because it was deemed unrealistic that in a bilingual VC conversation in the context of small and medium enterprises more than one interpreter would be used.
4. ISDN-based VC technology was the state-of-the-art technology and was available at the participants’ sites (SMEs). The VC systems used worked on the basis of the H.320 standard for audio and video encoding (G.722 and H.261 respectively). The technical setup was thus very similar to that of the ITU/ETI experiment described in Moser-Mercer (2003).
5. See also note 4.
6. The interpreter did not comment on the reasons for the lag here. One could argue that it was caused by the incoherence in the ST as well as by the direction of the interpretation from German into English. However, in other instances the interpreters attributed lag times to problems experienced with listening comprehension and incoherence in the ST, irrespective of the direction of the interpretation.
7. The channel switch to reverse the language direction was carried out with a mouse click that is audible in the recordings and is marked with ⊥ in the transcript.
8. In dyadic communication similar strategies are also used to control the performance of the other interlocutor(s). For a discussion of these, see Braun (2004).

9. The view that additional cognitive resources are only activated when the (subjectively felt) need arises is also supported by Sperber & Wilson's (1995) Relevance Theory, which suggests that hearers aim for the optimally relevant interpretation of an utterance that can be achieved without disproportionate effort. It seems plausible to hypothesise an equally economical use of resources for utterance production. Gile's (1991) Effort model for interpreting is also well in line with these assumptions.

10. The mistranslation of *I have helped out in a school before* (line 4) with *ich habe auch schon Nachhilfeeerfahrung* [I also have some experience with private tuition] (lines 5-6) will not be discussed here.

11. It is interesting to note that the reversal-of-order strategy can also be motivated by the interpreter's attempt to reduce his/her cognitive processing load (cf. Kalina 1998: 120). The example thus illustrates the methodological difficulties arising in connection with the analysis of performance data. Different strategies can lead to the same performance result. Moreover, the strength of strategic motivation behind performance results can vary (cf. Kohn 1990: 132ff).

12. In interpreter training, somewhat ironically, conditions which are similar to those of remote interpreting are common practice, as trainees are often subjected to working with pre-recorded video (or worse: audio) tapes. This is likely to increase with the use of virtual learning environments in interpreter training.

References

- Blakemore, D. (1992). *Understanding utterances*. Oxford: Blackwell.
- Böcker, M. & Anderson, B. (1993). Remote conference interpreting using ISDN videotelephony: A requirements analysis and feasibility study. *Proceedings of the Human Factors and Ergonomics Society, 37th annual meeting*, 235–239.
- Braun, S. (2004). *Kommunikation unter widrigen Umständen? Fallstudien zu einsprachigen und gedolmetschten Videokonferenzen*. Tübingen: Narr.
- Braun, S. (2006). Multimedia communication technologies and their impact on interpreting. In Gerzymisch-Arbogast, H. (Ed) *Proceedings of the Marie Curie Euroconferences MuTra: Audiovisual Translation Scenarios Copenhagen, 1-5 May 2006*. http://www.euroconferences.info/proceedings/2006_Proceedings/2006_proceedings.html.
- Braun, S., Kohn, K. & Mikasa, H. (1999). Kommunikation in der mehrsprachigen Videokonferenz: Implikationen für das Dolmetschen. In H. Gerzymisch-Arbogast, D. Gile, J. House & A. Rothkegel (Eds.), *Wege der Übersetzungs- und Dolmetschforschung*. Tübingen: Narr, 267–305.
- Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Clark, H. H. (1996). *Using language*. London: Longman.
- Daly, A. (1985). Interpreting for international satellite television. *Meta* 30 (1), 91–96.
- van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Faerch, C. & Kasper, G. (1984). Two ways of defining communication strategies. *Language Learning* 34 (1), 45–63.
- Finn, K., Sellen, A. & Wilbur, S. (Eds.) (1997). *Video-mediated communication*. Mahwah, NJ: Erlbaum.
- Gile, D. (1991). The processing capacity issue in conference interpretation. *Babel* 37 (1), 15–27.

- Hornberger, J., Gibson, C., Wood, W., Dequeldre, C., Corso, I., Palla, B. & Bloch, D. (1996). Eliminating language barriers for non-English-speaking patients. *Medical Care* 34 (8), 845–856.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge, MA: Harvard University Press.
- Kalina, S. (1998). *Strategische Prozesse beim Dolmetschen. Theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen*. Tübingen: Narr.
- Kohn, K. (1990). *Dimensionen lernersprachlicher Performanz. Theoretische und empirische Untersuchungen zum Zweitspracherwerb*. Tübingen: Narr.
- Kohn, K. & Kalina, S. (1996). The strategic dimension of interpreting. *Meta* 41 (1), 118–138.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Kraut, R. & Fish, R. (1997). Prospects for videotelephony. In K. Finn, A. Sellen & S. Wilbur (Eds.), *Video-mediated communication*. Mahwah, NJ: Erlbaum, 541–561.
- Laver, J. (1970). The production of speech. In J. Lyons (Ed.), *New horizons in linguistics*. Harmondsworth: Penguin, 53–75.
- Laver, J. (1973). The detection and correction of slips of the tongue. In V. Fromkin (Ed.), *Speech errors as linguistic evidence*. Den Haag: Mouton, 132–143.
- Laver, J. (1980). Monitoring systems in the neurolinguistic control of speech production. In V. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press, 287–305.
- Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition* 14, 41–104.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT.
- Morrison, D. & Low, G. (1983). Monitoring and the second language learner. In J. Richards & R. Schmidt (Eds.), *Language and communication*. London/New York: Longman, 228–250.
- Moser-Mercer, B. (2003). Remote interpreting: Assessment of human factors and performance parameters. *Communicate!* [Online] Summer 2003. <http://www.aiic.net/ViewPage.cfm/article879.htm> (accessed 17 June 2006).
- Moser-Mercer, B. (2005). Remote interpreting: Issues of multi-sensory integration in a multilingual task. *Meta* 50 (2), 727–738.
- Mouzourakis, P. (1996). Videoconferencing: Techniques and challenges. *Interpreting* 1 (1), 21–38.
- Mouzourakis, P. (2003). That feeling of being there: Vision and presence in remote interpreting. *Communicate!* [Online] Summer 2003. <http://www.aiic.net/ViewPage.cfm/article911.htm> (accessed 17 June 2006).
- Mouzourakis, P. (2006). Remote interpreting: A technical perspective on recent experiments. *Interpreting* 8 (1), 45–66.
- Niska, H. (2002). Community interpreter training. Present, past, future. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century*. Amsterdam: Benjamins, 133–144.
- O'Hagan, M. (1996). *The coming industry of teletranslation*. Clevedon: Multilingual Matters.
- O'Connell, B. & Whittacker, S. (1997). Characterising, predicting, and measuring video-mediated communication: A conversational approach. In K. Finn, A. Sellen & S. Wilbur (Eds.), *Video-mediated communication*. Mahwah, NJ: Erlbaum, 107–131.
- Olson, J., Olson, G. & Maeder, D. (1997). Face-to-face group work compared to remote group work. In K. Finn, A. Sellen & S. Wilbur (Eds.), *Video-mediated communication*. Mahwah, NJ: Erlbaum, 157–172.
- Oviatt, S. & Cohen, P. (1992). Spoken language in interpreted telephone dialogues. *Computer Speech and Language* 6, 277–302.
- Petite, C. (2005). Evidence of repair mechanisms in simultaneous interpreting. *Interpreting* 7 (1), 27–49.
- Sacks, H., Schegloff, E. & Jefferson, G. (1974). A simplest systematics for the organization of turn taking in conversation. *Language* 50 (4), 696–735.
- Short, J., Williams, E. & Christie, B. (1976). *The social psychology of telecommunications*. Chichester: Wiley & Sons.
- Sperber, D. & Wilson, D. (1995). *Relevance. Communication and cognition*. 2nd edn. Oxford: Blackwell.

- Wadensjö, C. (1999). Telephone interpreting and the synchronisation of talk in social interaction. *The Translator* 5 (2), 247-264.
- Wilson, D. & Sperber, D. (2004): Relevance theory. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, 607–632.

Appendix

Transcription convention:

P = participant (interlocutor); I = interpreter; each line of transcript = approx. 5 seconds

1	P2:	... 'HOW DO YOU EHM	BUY	TICKETS	FOR THE TRAIN'
	I:	... wie man zum Beispiel einkaufen geht,		wie man	
	P1:			overlapping speech (in italics) due to backchannel utterance from the hearer, here from P1	MH JA
2	P2:	AND THINGS. IS THIS WHAT YOU MEAN?			
	I:	<i>Fahrkarten</i> kaufen kann, für den Zug und so. Ja? Also ganz grundlegende Dinge eigentlich.		increased pitch	
3	/P1:	» P1: MHM.	WISSEN SIE AUCH BESCHIED ODER- ODER IST ES EIN VORURTEIL ODER GIBT'S		
	I:	speaker change	channel switch by I		
4	P1:	DAS WIRKLICH	EHM	UNTERSCHIEDLICHE MENTALITÄTEN IN DER ART, MITEINANDER UMZUGEHEN.	
	I:	Well,		(3.0) what about the different men-	
		(8.0/9.0 after end of P2)		length of a speaker pause (here after well)	
		length of pause at speaker change (at the interpreter's/former speaker's sites)			