Open Research Online

The Open University's repository of research publications
and other research outputs

The use of videoconferencing to support multimodal interaction in an online language classroom

Journal Item

oro.open.ac.uk

The use of videoconferencing to support multimodal interaction in an online language classroom

REGINE HAMPEL

*Centre for Research in Education and Educational Technology,
Walton Hall, The Open University, Milton Keynes MK7 6AA, United Kingdom
(email: r.hampel@open.ac.uk)*

URSULA STICKLER

*Department of Languages, Walton Hall, The Open University,
Milton Keynes MK7 6AA, United Kingdom
(email: u.stickler@open.ac.uk)*

Abstract

The introduction of virtual learning environments has made new tools available that have the potential to support learner communication and interaction, thus aiding second language acquisition both from a psycholinguistic and a sociocultural point of view. This article focuses on the use of videoconferencing in the context of a larger exploratory study to find out how interaction was influenced by the affordances of the environment. Taking a mainly qualitative approach, the authors analysed the written and spoken interaction in recorded videoconferencing sessions, alongside examining some quantitative data to reveal participation patterns. Exploring language learning interaction in a synchronous online medium allows us to show how this is a process mediated by interaction with experts and peers as well as by the artefacts used (e.g., technology) and how learners use and combine multiple modes to make meaning. Our findings illustrate how an online videoconferencing environment with its multiple modalities can be used in language teaching, how teachers and learners adapt to the multimodal online environment and how new patterns of communication emerge in the process.

Keywords: computer-mediated communication, online language learning and teaching, online classroom discourse, virtual learning environments, videoconferencing

1 Introduction

Communication plays a major role in language learning, and since the 1980s the concept of interaction has been used to explain second language acquisition both from a psycholinguistic and a sociocultural point of view (Ellis, 2003). Scholars following the psycholinguistic approach stress the importance of language input and of the ensuing interaction (between learners and other speakers or between learner and texts) for language development, focusing on the cognitive processes in the

learner's mind (e.g., Long and Robinson, 1998). Sociocultural theorists call attention to the lack of context that characterizes these views (e.g., Firth and Wagner, 1997; Hall, 1997) and argue that interaction does not happen in a vacuum but is mediated by a number of factors which go beyond the interlocutor's input. In sociocultural theory, learning is perceived as a process that happens through mediation and is 'inherently situated in social interaction, cultural, institutional, and historical context' (Wertsch, 1991: 86). The approach in this article is a socio-interactionist one (Mondada and Pekarek Doehler, 2004: 461), which stresses that 'learning is situated in learners' social, and therefore profoundly interactional, practices', and sees learning as the development of interactional resources (Brouwer and Wagner, 2004).

The project that this article is based on aimed at trialling and evaluating a number of new communication tools for language learning and teaching in order to increase interaction between learners at The Open University, a distance education institution in the UK. In 2005, the university adopted Moodle as the platform for its virtual learning environment (VLE), and a range of communication tools were introduced that have the potential to contribute to students' experience of studying in a distance education setting as well as improving learning in their particular subject area.

This university-wide development coincided with the Department of Languages planning a new generation of courses, thus offering the possibility of not just bolting the VLE on to existing courses but actually designing courses with the VLE integrated into them and taking into account pedagogical considerations. This is particularly crucial for language learning and teaching where the medium of learning is often also the content (i.e., language). So the authors devised a five-week project – an intensive German course run fully online. The target audience were students who had completed one of the university's mainstream distance learning courses at intermediate level (Common European Framework of Reference for languages, exit levels B1 or B2). Out of 472 students who were invited to participate in the free, non-credit-bearing and non-assessed course, 66 students expressed interest. After further details about the course and the evaluation tasks were sent out, a total number of 25 students enrolled in the course and they were divided into two groups taught by two teachers. Twenty of those students (13 females, 7 males) reached week five and completed the course.

To give students a maximum of flexibility and opportunities for synchronous speaking practice, the course combined asynchronous and synchronous tools. FlashMeeting, a videoconferencing system developed in-house[1], was the main platform for interaction, offering scheduled tutorial sessions, and it was also used for optional additional meetings amongst learners. In addition, forum, wiki, blog, quiz, and survey activities were built into the course (for more information, see Stickler & Hampel, 2010).

The following tools were used to collect data:

- One pre-course questionnaire and one post-course questionnaire (each consisting of three parts)
- Moodle feedback forms that invited students to send in their impressions at three stages throughout the course
- FlashMeeting recordings (10 sessions)

---

[1] see http://flashmeeting.open.ac.uk/home.html

■ Moodle activity logs
■ Interviews with individual students after the end of the course
■ A student focus group after the end of the course
■ A focus group with the two teachers and the course assistant after the end of the course.

For the purposes of this article we concentrate on the synchronous videoconferencing sessions, with the following aims:

■ To analyse the interaction that goes on using the main modalities available in this environment (namely speaking and writing)
■ To examine how this interaction is influenced by the affordances of the environment.

To achieve these aims, we analysed participation patterns, investigated functions and content of contributions and examined the interplay of spoken and written modes in online classroom discourse.

2 Multimodal computer-mediated communication and learning

Technological progress changes not only classroom practice but also research perspectives. This literature overview starts by examining the changes that using written CMC has on learner interaction before delineating research into the nature of communication in bimodal and multimodal language learning environments which include spoken communication.

For a number of years research has focused on communication and interaction in written online environments (both synchronous and asynchronous); these have been examined from cognitive as well as sociocultural perspectives. While some researchers see the effects of the introduction of CMC in education mainly as an increase in active participation, others claim that computer mediation actually qualitatively changes communication (as well as our actions and minds). In his introduction to a special issue of *Distance Education* Wertsch (2002) takes some of the contributors to task for suggesting that CMC only has quantitative effects:

> [I]t may be important to take into consideration the possibility that one cannot simply add asynchronous communication tools into an existing mix of social and psychological processes without changing them in fundamental, unintended ways and that this may be one of the most interesting aspects to consider in computer-mediated PBL [problem-based learning]. (*op. cit.*: 106)

Yet research does exist which highlights the qualitative changes that written CMC brings about compared to traditional classroom settings both within and without the languages context. This includes work done on interaction (Kreijns, Kirschner and Jochems, 2003), types of discourse (e.g., Lamy and Goodfellow, 1999; Weininger and Shield, 2003), and literacies (Lankshear, Gee, Knobel and Searle, 1997; Warschauer, 2000).

With technology moving on, there has been a shift in teaching towards using multimodal environments (Stockwell, 2007), including audio-conferencing as well as

desktop video-conferencing applications. However, there is a lack of research that examines the impact of this combined use of tools on interaction and analyses multimodal communication in an online language classroom. The applications available (e.g., Breeze, DimDim, Elluminate, FlashMeeting, Gong, NetMeeting, and Skype) are more or less sophisticated tools combining different modes – that is, 'organized, regular means of representation and communication' (Jewitt, 2004: 184) – such as spoken and written language, visual and graphic systems, spatial systems and body language.

Research on the use of text and audio for language learning includes Jepson's (2005) SLA-oriented study which compares the patterns of repair moves in synchronous non-native speaker text chat rooms in comparison to voice chat rooms. His conclusion is that '[i]nteraction patterns sway – and are swayed by – the unique social activity of the electronic environment' (*op. cit.*: 80). Other comparisons of interaction in face-to-face, telephone and computer mediated speaking environments have shown the influence of technological affordances on L2 communication, for example in the patterns of silence (Heins, Duensing, Stickler and Batstone, 2007; Stickler, Batstone, Duensing and Heins, 2007).

Because of easy and often free access to technology that combines text, audio and video, teachers are increasingly likely to use desktop videoconferencing environments. Some of the issues identified in earlier studies prevail (e.g., time lags, lack of lip synchronization, turn-taking – see Wang, 2004) while at the same time the new multimodal environments pose additional challenges. Wang, Chen and Levy (2010) tackle training for a multi-user virtual classroom that encompasses synchronous videoconferencing. Cunningham, Beers Fägersten and Holmsten (2010) focus on communication problems in a videoconferencing environment and identify compensatory strategies. In her study of FlashMeeting, Örnberg Berglund (2009) examines the quantity of communication and the quality of interaction. She concludes that while there were some examples of multimodal interaction, students' 'contributions often consisted of long monological turns' (*op. cit.*: 202) and the multimodal tool did not equalize communication in terms of quantity.

There is a small but growing number of studies that focus on the interplay of modes in synchronous CMC. Blake (2005) reports on the benefits of using an application that combines sound and text chat for negotiating meaning and highlights the socio-affective benefits particularly for distance learning contexts. Another study shows how while watching online lectures via a webcast, learners were able to use both audio and text chat to communicate with their peers (Scheffel-Dunand, 2006). Ciekanski and Chanier's (2008) study serves as an illustration of working with multiple modalities by combining them to make meaning and foster collaboration. In the context of an audiographic environment, they show the benefits of modality-switching, in this case between audio and text. '[T]he multimodal learning environment, due to its process-oriented and collaborative nature, helps learners focus more on the writing process than on the results of their writing. The integrated word processor affords a shared visible image of the writing process' (*op. cit.*: 180). The writing process is thus 'perceived as a complex and procedural activity and as a social event' (*op. cit.*: 179). Finally, Kenning (2010) specifically looks at the different impact of using voice and text chat in two different environments, identifying a

number of factors that affect discourse patterns: the level of integration with other devices and tools; tutor privileges; the number and function of icons; and the type of transmission, that is, full versus half duplex (*op. cit.*: 17).

As multimodal perspectives in analyses of classroom discourse are still rare (Flewitt *et al.*, 2009), we were particularly interested in how learners as well as teachers would use different modes to make meaning.

3 Research methodology

Language learning is a process mediated in interaction with experts and peers as well as with the artefacts used – artefacts that include technology (Lantolf, 2000; Lamy and Hampel, 2007). To explore the interplay of different factors in the online language learning situation presented here, a predominantly qualitative approach – advocated, for example, by Debski and Levy (1999), Warschauer (2000) and Mercer, Littleton and Wegerif (2004) – was devised. However, before examining in more detail all ten FlashMeeting tutorial sessions we interrogated the quantitative data to gain some statistical insights into turn taking, speaking dominance, or participation in text chat activities. One session was selected as exemplar for the purposes of this article (see section 4). Where relevant, comments from interviews and focus groups were used to support conclusions drawn on the quantitative and qualitative tutorial analysis.

FlashMeeting (see Figure 1) is a desktop videoconferencing tool; at the time of the project it offered the following communication modes:

1. Linguistic: spoken and written
2. Visual: icons (vote buttons yes/no/?, emoticons), still images (on whiteboard), moving images (via webcam); and display/scrolling of text
3. Gestural: via webcam

Communication in FlashMeeting was influenced by the following functionalities that were specific to the tool:

- Audio: half duplex (i.e., only one person can speak at a time)
- Text chat: no private channels available
- Joint creation of text and images (whiteboard)
- Parallel use of audio, text chat and whiteboard possible
- Video images (via webcam): large picture of speaker (relatively fast refresh rate) and small thumbnail images of all participants (slow refresh rate)
- A choice of name list or thumbnail video images view of participants
- Joint web browsing
- Queuing system for turn-taking with automatic allocation of turn (i.e., audio channel opened automatically to the next speaker in the queue)
- 'Interrupt' button
- Automatic recording of sessions (allowing participants later re-viewing)

Because of the relatively slow refresh rate for multiple video images body language is not an effective communication tool in FlashMeeting beyond contributing at a socio-affective level. (There is some evidence for this in the interviews.) Vote buttons or emoticons to confer additional meaning were hardly used by students in our

Fig. 1. FlashMeeting screenshot

study. Therefore verbal interaction via audio and text chat was chosen as the main focus of this study.

As a first step in the analysis, spoken interaction in all video sessions (five sessions per teacher, ten overall) was transcribed by the authors. These transcripts were combined with the parallel text chat log (synchronised by timestamps created by the software) and subsequently analysed. In the analysis we focused in more depth on the discourse functions of multimodal interactions. These functions are central for learning in all subjects, but they are particularly crucial in language learning where negotiation of meaning, that is, the modification of input (by using, for example, simpler grammar and vocabulary) and of interaction (by, for example, requesting clarification), has been shown to contribute to second language acquisition (Long, 1983; Varonis and Gass, 1985). We used the discourse functions which Sotillo (2000: 84) employed in her study of synchronous and asynchronous communication and which she defines 'as categories of behavior in electronic discourse, such as requests, responses, apologies, greetings, complaints, and reprimands'. Although Sotillo's research focuses on comparing text environments, her categories could be adapted to our purposes. When we started examining the data more systematically it became obvious that beyond these discourse functions relating to (1) social interaction and (2) on-task negotiating meaning, additional functions could be identified, namely (3) off-task conversations and (4) technical discussions.

The quantitative data presented in this paper relates to one tutorial session only and serves as an illustration of common usage patterns throughout the course.

4 Speaking and chatting: participation patterns

FlashMeeting was deemed to be important for real-time communication and inter-action in the course as well as keeping up student motivation and community building and it was used for weekly synchronous tutorial sessions with the two teachers Eva and Sylvia. Worksheets prepared students for each session which focused on speaking practice and a follow-up task prompted participants to reflect on their learning. Because FlashMeeting sessions are automatically recorded, students were able to use the replay of meetings for critical reflection or to catch up with the course when they had missed a tutorial session.

As a multimodal tool, FlashMeeting offers more than audio. The illustration in Figure 2 shows the detailed pattern of tool use in one particular session (Sylvia's tutorial 4) – which is typical of all tutorial sessions led by the two teachers.

The participants are listed on the left in the order they logged in (teacher SY and 7 students) and the timeline (in 5-minute intervals) is given at the top. The diagram includes information on participants using the audio channel (called 'broadcasting' in FlashMeeting), text chat, emoticons, and other buttons. The top row combines the audio contributions of all participants (each of whom is identified by a particular colour), while the other rows represent each participant's spoken and other contributions. So we can see that user 2, the teacher, spoke most and took turns very frequently, thus clearly dominating the meeting. The students contributed less frequently. In the post-course questionnaire one student remarked that the Flash-Meeting sessions were 'rather hard […]. The meetings did not seem to inspire any of us to contribute freely.' One of the teachers also commented on this: 'I think that was more a thing of being still very much reliant on the tutor, in my own group anyway, that you'd interact a lot over email and then the FlashMeeting'. This reliance on the teacher is supported by similar studies on interaction in online audio-graphic conferencing (see Stickler *et al.*, 2005).

Tutor dominance in terms of speaking and chat use was observed across all tutorial sessions. For session 4 a graphic representation can be seen in Figures 3 and 4 which give the proportion of audio and text chat contributions per person (shown by the angle of the pie segment) as well as the number of broadcasts and messages respectively (shown by the radius of the segment). After the teacher, the
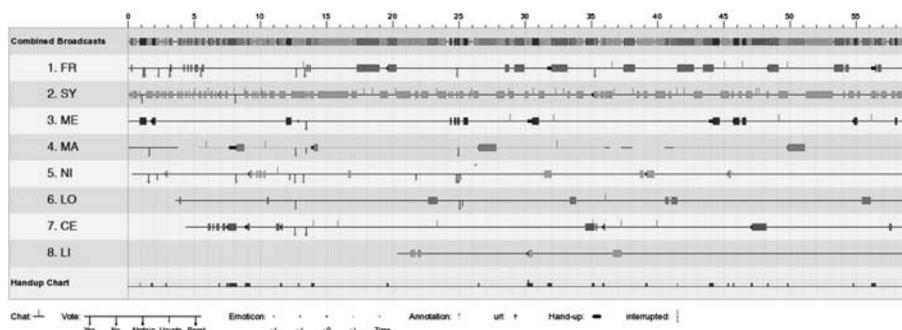

Fig. 2. Participation rates (teacher: Sylvia = SY)

Broadcast Dominance

Polar-Area Diagram
Arranged clockwise by increasing area
Segment angle shows the proportion of the total broadcast time (50 mins 40 secs)
Segment radius representative of number of broadcasts (max. = 77)
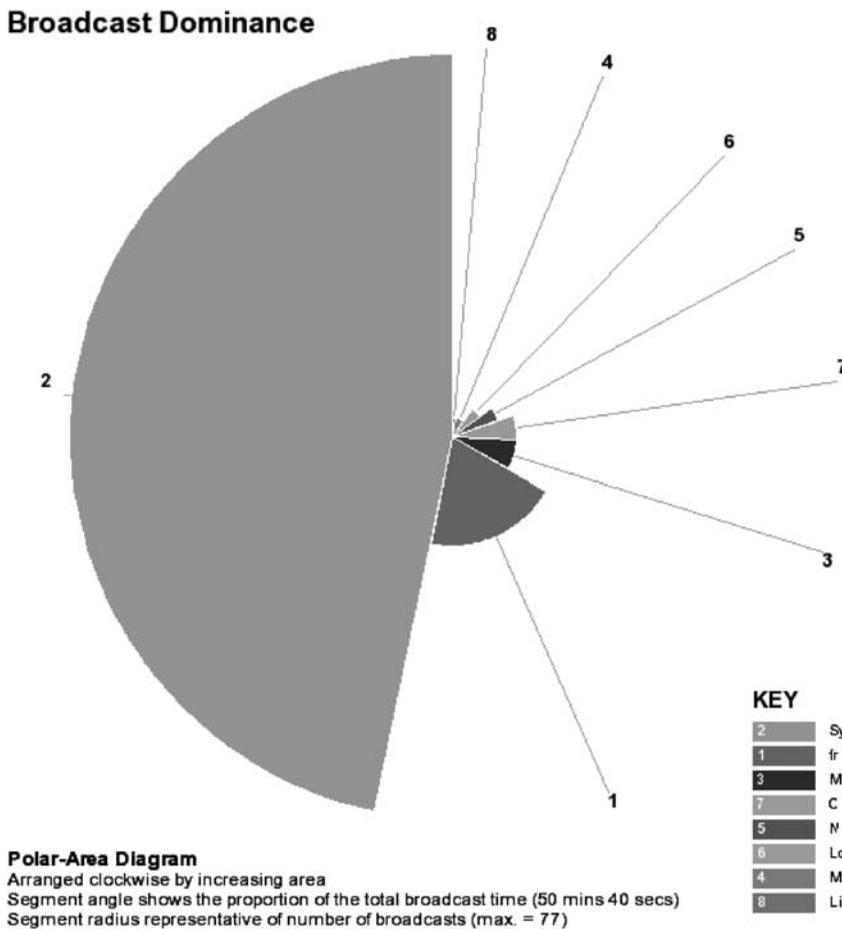
Fig. 3. Broadcast dominance (teacher: Sylvia = Sy)

next most active user was student Fr (Frida – all student names have been changed) who used the speaking function most amongst students and who broadcast 20% of the time of the meeting. Most students only spoke for an average of 5% of the time. It is interesting to note that while the teacher's proportional contribution is similar in speech and in text chat, the students' contributions in the spoken and the written medium differ – with Frida, for example, speaking more than chatting, and Caroline (C in the key) chatting more than speaking.

Table 1 shows the number of spoken turns in this tutorial session, together with the number of turns in the written text chat. It also gives additional information about silent turns, that is, students pressing the broadcasting button without actually speaking.

Thus the teacher took 77 turns to speak. Frida, the next most active user, took 22 turns to speak during this session, a considerable contribution. With her combined communication modes, Caroline achieved the third highest rate of turns taken in the tutorial session, after the teacher and Frida. However, because her audio was
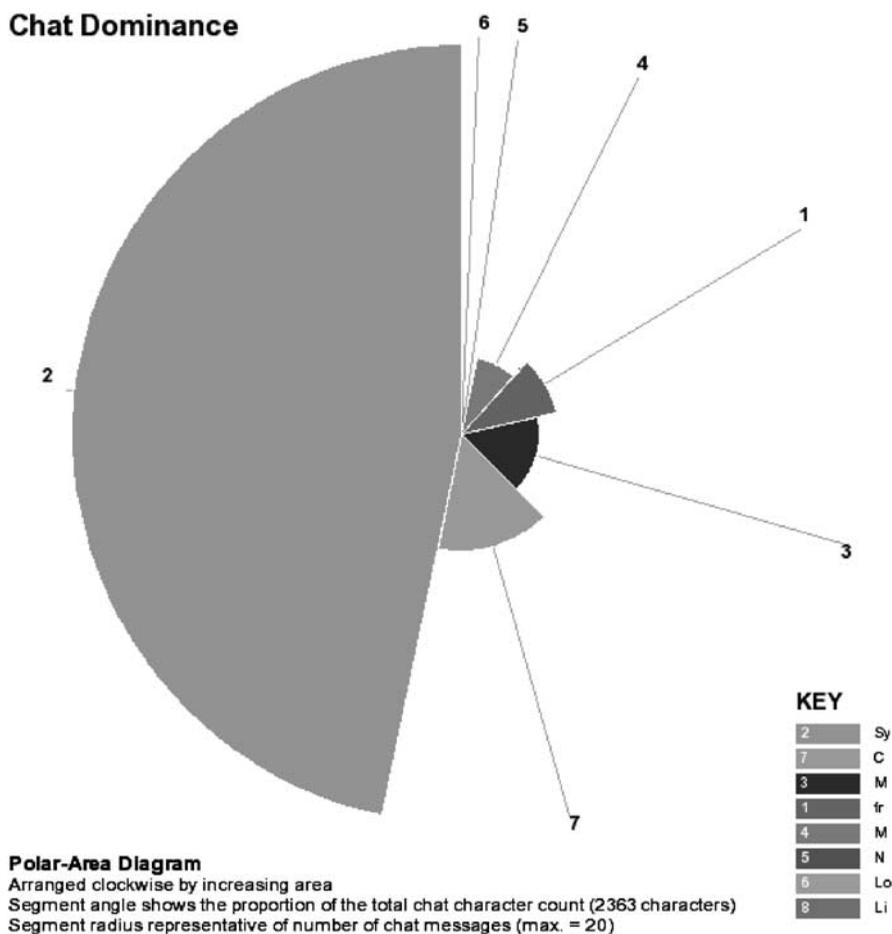
Chat Dominance

Polar-Area Diagram
Arranged clockwise by increasing area
Segment angle shows the proportion of the total chat character count (2363 characters)
Segment radius representative of number of chat messages (max. = 20)

Fig. 4. Chat dominance (teacher: Sylvia = Sy)>

Table 1 *Number of turns per participant in tutorial 4*

	Speaking turns	Silent turns	Text chat	Overall
Sylvia (teacher)	77	0	19	96
Frida	22	1	5	28
Megan	11	2	4	17
Norman	10	0	1	11
Lorna	7	0	1	8
Caroline	4	10	6	20
Mary	4	0	4	8
Linda	4	0	0	4

working only intermittently during the first part of the meeting, her spoken con-
tributions could not be heard by other participants and many of her turns were silent
(where she continued to broadcast but no speech was audible). The low number of

turns in Linda's case can partly be explained by the fact that she joined late (see Figure 2). This, together with the fact that she had missed the previous tutorial session, may have resulted in a lack of confidence in speaking. She did not use text chat at all.

5 Discourse functions of multimodal conversations

In order to find out how students communicated with written and spoken language, what the audio and the text chat were used for and how the two modes interplayed, the quantitative analysis was followed by a qualitative examination of all ten tutorial sessions. If we look at the main discourse functions of the audio channel (the primary channel of communication in all ten FlashMeeting sessions) and the text chat (the secondary channel), the following ones appear across both modes:

- Social conversations (greetings and farewells)
- Management of technology
- Negotiating meaning (related to the task)
- Off-task conversations amongst students
- Teacher feedback

5.1 Social conversations (greetings and farewells)

So why do participants use audio in some instances, and text chat in others? The greetings may give us an example of why students and teachers choose text chat over audio: While the audio often served for greeting others and saying farewell to them, the text chat tended to be used more by and for students who came in late or who had to leave early and where a participant wanted to say hello or good-bye without disrupting the flow of spoken conversation. FlashMeeting does not allow more than one person to speak at a time and the refresh rate of the thumbnail video images is very slow. As a result, the video is far from ideal for making meaning through body language (e.g., waving) if one is not already speaking. To avoid interrupting the speaker with small but crucial discourse features such as saying hello or farewell, the text chat (alongside the voting buttons and emoticons) is a useful tool, and in both groups students and their teacher seemed to develop a routine for doing this.

5.2 Management of technology

While the audio was the main tool for discussing technical issues at the beginning of each tutorial (e.g., in the context of a sound check), the text chat was used as a less intrusive tool for managing the technology at other points in the tutorials. In some extreme cases, it was used as a back-up when the audio channel failed for individual students. Extract 1 shows Norman using the text chat to comment on the low audio volume of Linda while Caroline is talking about where she comes from.[2]

---

[2] Please note that the text chat (indented and shown in bold) has been inserted into the audio transcript at the point where it appeared in the tutorial. […] indicates that the transcript has

Extract 1: Management of technology

Linda:	<inaudible>
Sylvia:	Ah, das ist interessant […]. Und Norman, kommen Sie aus dieser Gegend? *Ah, that's interesting […]. And Norman, are you from that region?*
Norman:	[…]
Sylvia:	[…] Und Caroline, wie ist es bei Ihnen […]? *And Caroline, how about you […]?*
Caroline:	Ja, ich komme aus Essex in Südostengland, ich wohne noch in Essex, aber nicht in die reichen *Yes, I'm from Essex in south east England, I still live in Essex, but not in the rich*

   Norman: Linda, ich kann Sie kaum hören
       Linda, I can hardly hear you

reichen Kaff, wo ich geboren war – uh, das war schlecht Deutsch. Also ich wohne jetzt in Great <inaudible> das ist Nord Nordost Essex, in der Nähe von Colchester.
rich place where I was born – uh, that was bad German. So, I now live in Great <inaudible> that is north northeast Essex, near Colchester.

5.3 Negotiating meaning (related to the task)

The visual representations of spoken participation in the FlashMeeting session (see Figures 1 and 3 above) have already shown the teacher dominance in this particular virtual classroom. So if we focus on the interaction in the audio recordings, it will not come as a surprise that patterns of communication can be observed that are often teacher-led and where the typical IRF pattern occurs, that is multiple series of initiation (teacher), response (student) and feedback (teacher). However, if we analyse the spoken interaction together with the text chat, a more complex picture emerges.

 Let us first examine multimodal interactions that helped participants to negotiate meaning. The recordings show how the text chat was used in conjunction with the audio channel for the following purposes:

■ To make assertions that contributed to the main spoken discussion
■ To agree with what somebody had said
■ To make requests for clarification/confirmation/explanation re vocabulary used in the audio
■ To challenge what someone had said

The following examples (extracts 2–5) will illustrate these points. Extract 2 includes a student making a written assertion in the text chat.

Extract 2: Assertion

Sylvia: Stimmen Sie da zu, Norman? Denken Sie, Wikipedia ist auch ein bisschen Anarchie im Internet?
Do you agree, Norman? Do you think, Wikipedia represents a little anarchy in the internet?

Norman: Vielleicht, aber kann man auch nicht sagen, dass das Internet anarchistisch ist?
Perhaps, but couldn't we also say that the internet is anarchic?

Sylvia: Ja, sicher, denn jeder kann veröffentlichen, wie er möchte. […] Gibt es hier eine Regel in der Wikipedia? Hat man einen Schutz z.B. vor Terrorismus sogar – Mary.
Yes, sure, for everybody can publish how he wants. […] Is there a rule in Wikipedia? Is there protection, e.g., from terrorism even – Mary.

Mary: Ja, ich, in einer encyclopaedia man erwartet – expect
Yes, I, in an encyclopaedia you expect – 'expect'

    Frida: **Anarchy ist ein positive wort für einige Leute die mit freie software arbeiten**…
    Anarchy is a positive word for some people who work with free software…

I mean – ein authority, das was es geschreiben ist mit ein bisschen authority
'I mean' – an 'authority', that things are written with a little 'authority'

This extract shows the teacher (Sylvia) addressing Norman directly with a question about Wikipedia. He answers with another question, which Sylvia answers in turn. Then she moves to another aspect of the topic, the regulation of the internet, and asks Mary a question. Mary answers, and while she answers a text chat message appears, posted by Frida who also wants to make her point.

In extract 3 we can see how the text chat is used for requests, in this case for a vocabulary item.

Extract 3: Confirmation request

Norman: Ja, ich sch-stimme Mary zu – ist das richtig? Ich suche nach einem Wort. Kann man nichts nichts anstößig sagen?
Yes, I a-agree with Mary – is this correct? I am looking for a word. Is it not possible to say offensive?

Sylvia: Wie bitte?
Pardon?

Norman: Ich schreibe
I write

Sylvia: […]

Caroline: […]

    Norman: **anstößig – offensive?**
    offensive – 'offensive'?

    Sylvia: **ja**
    yes

Norman is looking for a word to use in his agreement with Mary's point. He uses the correct German word ('anstößig') but – possibly because he is uncertain of its meaning or the correct pronunciation – says it in a way that Sylvia does not understand. So he resorts to text chat, announcing that he is going to write, and a few seconds later – while the conversation has moved on from the teacher to Caroline – the German word and the English equivalent appears with a question mark in the written chat. And in order not to interrupt Caroline's speech, Sylvia types 'yes' into the text chat to give feedback to Norman.

Extract 4 shows how a student uses the text chat to agree with the spoken contribution of another student.

Extract 4: Response (agreement)

Caroline: Das ist ganz schwierig zu sagen. In eine traditionelle Encyclopaedia kommt etwas – schreibt hoffentlich nicht Fachmänner, Spezialisten? Ich weiß es nicht das Wort, jemand, der etwas sehr tief studiert hat und so vielleicht ist das ein bisschen besser als Wikipedia, […] aber in Wikipedia – oh ja – in Wikipedia gibt es die Gelegenheit viele Spezialisten oder Fachmänner zu zu schreiben – ich meine

That is very difficult to say. In a traditional encyclopaedia something comes – hopefully it is not specialists who write, specialists? I don't know it the word, someone who has studied something very deeply and therefore this may be a bit better than Wikipedia, […] but in Wikipedia – oh yes – in Wikipedia there is the opportunity for many specialists to to write – I mean

    Frida: **Ich stimme dir zu**

            I agree with you

contribute, aber ich es nicht das Wort, so, zu schreiben

'contribute', but I don't the word, so, to write

In face-to-face communication this kind of agreement can be given through back-channelling (e.g., some affirmative noises) or body language (e.g., nodding). In FlashMeeting the half-duplex audio and the low video refresh rate prevent this and so Frida uses the text chat to react to Caroline's point about the merits of Wikipedia which has many specialists contributing to it. The teachers similarly used the text chat to respond to students, signalling agreement (see extract 7).

In extract 5, the teacher (Eva) is being challenged by one of the students in her group.

Extract 5: Challenge

Eva: […] Ihr solltet also alle diese Woche, wer's noch nicht gemacht hat, einen Blog selber kreieren […]. Und was ihr gesagt habt, […] dass wir versuchen, interessant zu schreiben […] Ist das okay für alle?

[…] This week you are all supposed to create a blog if you haven't already done so […]. And what you mentioned, […] that we try to write in an interesting way […]. Is everybody OK with this?

> *Neill*: **Eva wann mann will mehr interessant schrieben vielleicht kommt mehr fehler?**
>
> *Eva if one writes in a more interesting way perhaps there are more mistakes?*

Andy: Bitte habe ich das verstanden, wir werden einen Blog haben. Aber wann beginnt dieser Blog?

Please have I understood this, we will have a blog. But when does this blog begin?

Eva: Also die Blogs kann man jederzeit machen. […]

Well, you can create the blogs any time. […]

> *Neill*: **Eva Was sollten wir machen—mehr interressant oder weniger fehler?**
>
> *Eva What are we supposed to do – more interesting or fewer mistakes?*

Eva's group has been discussing blogs and Eva reminds students towards the end of the tutorial session that they are supposed to create a blog if they have not already done so. She also summarizes the points that the group had collated about using a blog, which include writing in an interesting way and being polite. This is being challenged by Neill, who uses the text chat in order not to interrupt her. He may also have chosen the written format because it is less daunting to challenge what the teacher says in the text chat rather than in speaking. When Eva continues talking about the blog and answers the spoken query of another student rather than responding to his written question, he repeats his text chat message, this time phrasing it in a more direct manner and asking what they should do.

5.4 Off-task conversations amongst students

Students' contributions consisted mainly of responses and enquiries; they did not initiate a new topic often – at least not in relation to the set course tasks. However, there were off-task conversations that usually occurred in the mode other than the one the teacher was using, that is, they took place mostly in the text chat. However, in Eva's group students used the teacher's speech pauses – which in the main occurred when she was typing in the text chat – for social conversations and technical peer support in the audio mode (see extract 6, where she takes almost three minutes to summarize some key points in the text chat).

Extract 6: Off-task conversation

Eva: […] ich schreibe das wieder in den Textchat, und ihr könnt vielleicht auch noch einmal überlegen: die Hauptunterschiede und die Haupt-gemeinsamkeiten. […] Gebt mir bitte eine Minute oder zwei Minuten um das zu tippen. Augenblick, bitte.

[…] I'll write this into the text chat again, and you can perhaps think once more about the main differences and the main similarities.[…] Give me a minute or two to type this up. One moment, please.

Andy: Wilma, Sie sind laut und klar […]
Wilma, you are loud and clear […]
[…]

Wilma: […] Vielleicht wäre es gut mal zu sehen
[…] Perhaps it would be good to see

> **Eva**: **Unterschiede Blog/Wiki: Objektivität, Selbstdarstellung versus Information, Zugang (Wiki über die Wikipedia Plattform, Blog weniger uniform), Schreibstil und Inhalt, Edition von Texten. Gemeinsamkeiten: beides ist über das Internet zu erreichen, links und Fotos, ec. sind inegrierbar, leicht zu aktualisieren, beide werden nicht nur von akademischen Spezialisten geschrieben**
> *Differences blog/wiki: Objectivity, self-representation vs Information, access (wiki via the Wikipedia platform, blog less uniform), writing style and content, editing of texts similarities: both are accessible via the internet, links and photos, ec. are easily integratable, easy to update, both are not only written by academic specialists*

was die settings sind für Audio. Hast du das mal ge- nachgesehen?
what the 'settings' are for audio. Have you had a look at that?

Andy: Jawohl, das habe ich gesehen. Vielleicht ist es ein ein bisschen besser als bevor, aber nicht sehr gut. Wilma, sind Sie in Belge, in Belgium?
Yes, I have seen it. Perhaps it is a little better than before, but not very good. Wilma, are you in Belgium, in 'Belgium'?

5.5 *Teacher feedback*

As the extracts above have already shown, the teachers also used the text chat. They did so to execute the following discourse functions in the context of giving written feedback:

- To respond to a student's spoken contribution
- To respond to a student's query
- To recast or model language in writing
- To summarize the spoken discussion

Extracts 3 and 6 above have already shown how the teacher uses the text chat to give positive feedback and to provide students with a written summary. The latter was done regularly by Eva who may have felt that the text chat was a less ephemeral modality than the audio, allowing students to go back to what she had written.

Extract 7 is an example of the teacher (Sylvia) responding to a student (Lorna) in the text chat.

Extract 7: Response (agreement)

Sylvia: […] Haben die anderen hierzu bitte noch einen Kommentar? Was ist Ihre Meinung? […]

>
> *Does anybody want to comment on this? What is your opinion? […]*

Lorna: Ich habe es sehr interessant gefunden, dass jemand – ich glaube es war Megan – hat allgemein Tipps gegeben für eine Sprache zu lernen, und das war für mich sehr interessant.

> *I have found it very interesting that somebody – I think it was Megan – gave general tips for learning a language, and I found that very interesting.*

Megan: Ja, bei mir sollte

> Yes, I think that

> > **Sylvia:** **ja, super Idee!**
> > **yes, great idea!**

ein Blog – es sollte auf dem Schirm elegant und einfach aussehen, mit gute

Weg – Wegweiser, ist das Wegweiser signposting? Wenn es Frage gibt, es wird Feedback sofort verfügbar.

> *a blog – it should look elegant and simple on the screen, with good sign – signposting, is that signposting 'signposting'? If there is a question, feedback will be immediately available.*

Sylvia: Ja, genau, ich denke auch, das ist ganz wichtig, das Aussehen, es muss attraktiv aussehen. […]

> *Yes, exactly, I think so too, this is very important, how it looks, it needs to be attractive. […]*

This shows a complex IRF pattern which combines teacher feedback in both the written and the spoken mode to two student contributions.

The last example, extract 8, shows how the teacher (Eva) uses the text chat to recast language and then models the use in the audio.

Extract 8: Recast and modelling

Anita: Und vielleicht auch uh, jede, uh entry, was ist entry? nicht zu lange sollte, als sondern es ist ein bisschen langweilig.

> *And perhaps uh every uh 'entry', what is 'entry'? shouldn't be too long as but it is a little boring*

> > **Eva:** **Eintrag = entry**
> > **entry = 'entry'**

Eva: Gut, also schoene kurze Einträge und keine langen Romane.

> *So, nice short entries then and no long novels.*

These examples illustrate how students and teachers succeeded in using the affordances of the technology for multimodal communication, by taking the tools and adapting their use to their purpose and to the context. We can see how the chat takes over some of the functions that back-channelling, paralinguistic cues or body language have in face-to-face contexts, thus compensating for the fact that Flash-Meeting does not have full-duplex audio, has a slow refresh rate and only a limited visual field. In this way the participants use the tools to interact at the task level (contributing to the language learning activities, co-constructing knowledge) and at

the socio-affective level (supporting one another, building relationships in the group), and they do so using multiple modalities.

6 Interplay of modes: complementation, compensation and competition

When we analyse the multimodal interactions in FlashMeeting, it becomes obvious that there were three ways in which the modes were combined: to complement each other, to compensate for shortcomings, or in competition with each other.

6.1 Complementation

Firstly, there were instances where the spoken and the written language complemented each other and where an input in one mode triggered a response in another. This includes those examples where a student used the text chat to make a written assertion, thus contributing to the discussion going on via the audio channel (extract 2), to make a request (extract 3), to respond to something that had been said (extract 4) or to challenge a contribution (extract 5). The teachers used this approach when they gave written feedback to student contributions (extracts 3 and 7), recast or modelled vocabulary (extract 8), summarized spoken contributions (extract 6) or gave links to websites related to the discussion. As can be seen in the latter two instances, the text chat proved to be useful as a more permanent record than audio. The other main reason for using text chat was the fact that students as well as teachers could contribute to the discussion without having to interrupt the speaker.

6.2 Compensation

Secondly, users compensated for an actual or perceived lack in one mode by using another – as Caroline did when she resorted to text chat because of audio problems. Also, as the refresh rate of the thumbnail video images was very slow and only the speaker's facial expression and perhaps some body language could be used to make meaning, the text chat took over some of the functions that body language carries out in face-to-face contexts (through paralinguistic cues such as nodding or smiling). It also compensated for the lack of full-duplex audio which prevents the back-channelling that we use all the time as supportive moves in face-to-face conversations. In extracts 4 and 7, Frida and Sylvia can be seen to use it to show their agreement with what is being said explicitly.

The half-duplex mode and the turn-taking tool and queuing system also make the spoken conversation more formal and less spontaneous. As a result, it was sometimes difficult to complete adjacency pairs – as tends to happen more frequently in synchronous written environments such as Instant Messaging. The following comment made by one of the students in their end-of-course questionnaire picks this up: 'the medium actually is not well suited to spontaneous exchange because of the time delays between speakers. The simultaneous text chat was usually much more lively and spontaneous than the real speaking.' Not all students displayed the skills of using multiple modes at the same time; familiarization with other, similar software undoubtedly helped – several of the students were acquainted with the university's

audiographic conferencing system used in some courses for the delivery of tutorials (see also Hampel and Hauck, 2004).

We can also envisage a user compensating for his or her perceived deficiencies in one modality (e.g., difficulties with typing or anxiety about speaking) by using another more frequently. The audio, for example, gives learners little time to rehearse their statements and can thus create anxiety, whereas the text chat allows users to reflect on their contributions. This was Frida's comment in the student focus group: 'I really, really hated FlashMeetings. […] I just really, really struggled to speak. I just was so nervous'. This preference for one mode over another could also be witnessed in students' use of tools – thus some students preferred the asynchronous writing tools such as blogs over the synchronous videoconferencing environment (see also Stickler and Hampel, 2010).

6.3 Competition

Thirdly, we were able to observe competing conversations going on in the spoken mode and in the text chat. In most cases, these parallel conversations consisted of discussions of the topic in one modality and advice or comments about the technology in the other channel or greetings of latecomers and early farewells. In extract 1, for example, one of the students used the text chat to comment on the low audio volume of another student. We saw Eva using the text chat like a whiteboard (a tool which is included in FlashMeeting but which the teachers did not find sufficiently user-friendly) and extract 6 shows how students were using the pauses that occurred while she was writing to talk about the technology and even engage in social communication.

These findings illustrate how users appropriate new modes of communication to suit their own purpose and how they start developing a new culture of interaction – even in the relatively short period of five weeks in the case of this project.

7 Implications for language learning and teaching

The quantitative analysis of communication in this study has provided us with some general information about user participation. While it has shown the danger of teacher dominance in synchronous oral environments – which is also in line with research on face-to-face classrooms (Ho, 2007) – it also shows the extent to which learners used the main modes to make their meaning and indicates patterns of user preference. In contrast to Örnberg Berglund (2009) we did not find learner monologues.

The qualitative data gleaned through observation of a virtual classroom has shown the impact of a multimodal environment combining text, voice, image and live video on communication and interaction. It illustrates that the interaction that goes on in this distance learning and teaching situation reflects in many ways the discourse functions in face-to-face classroom settings – social conversations, negotiation of meaning, and off-task conversations – plus frequent conversations about technical issues such as sound problems.

At the same time, this study has confirmed that teachers and learners need to take into account the functionalities of the medium which shape the affordances of the tool and mediate the interaction between participants (see also Jepson 2005). In the case of videoconferencing, this includes being aware of whether the audio allows only

one person to speak at a time or several. As this has an impact on turn-taking, interruptions and back-channelling, learners need to know that, for example, in FlashMeeting text chat may be used as back-channel. When using text chat it makes a difference to students if entries are visible to everyone or if private messages are allowed. As Wang (2004) has shown, video conferencing is often limited by a slow refresh rate which means that body language cannot be used extensively.

While relevant for all subjects, the impact on language learners (particularly at beginner or intermediate level) of not being able to use body language can increase already present levels of language anxiety (de los Arcos 2010) and diminish the impact of social presence (Satar 2011).

In the process of using the online resources over a period of time, successful users adapt to the functionalities of the tools and transform them for their own purposes. In our exploratory study we discovered a number of strategies that learners and teachers developed. Firstly, participants used the text chat to comment on what the speaker said while at the same time not interrupting them, thus making up for the lack of being able to use body language, or to make a point while somebody else was speaking. Secondly, the text chat was used to check and confirm vocabulary by learners and teachers, following a breakdown in communication. Thirdly, because of the lack of a private communication channel, learners turned to the modality less used at a particular time for asides. Fourthly, the teachers used the text chat in the absence of a blackboard.

The modes available in multimodal online learning environments interrelate in different ways; they can complement one another, they can compensate for any real or perceived shortcomings–of the tool (Cunningham *et al.*, 2010) or of the learner–and they can compete with one another.

8 Conclusion

We are aware of the limitations of the study. In terms of the research set-up we were using volunteers with self-perceived high CMC competence (as revealed in the pre- and post-course questionnaires, see Stickler and Hampel, 2010); and in this paper we singled out one tool which was an integral part of a complex whole. In addition, we have focused on the audio channel and text chat, have only referred briefly to the use of the video channel and have not taken into account the impact of using the voting and emoticons (although this was admittedly infrequent in our particular project). However, the findings substantiate a number of crucial observations regarding multimodal communication online and newly emerging discourse patterns in online student–teacher interaction.

Because of the additional technological mediation, online tools such as Flash-Meeting cannot replicate face-to-face interaction but they offer channels of communication that can bridge physical distance (Blake 2005). Their potential for language learning lies in offering learners multiple modes for making meaning, modes which can be used simultaneously for reinforcement, or to fit in with the learners' perceived strengths. In contrast to using telephone or written online environments as teaching media, audio and video conferencing offer a number of benefits to learners, especially in a distance setting: they allow for a combination of different modes and multiple parallel representations, thus increasing modal density (Norris, 2004) and media richness.

An online videoconferencing environment such as FlashMeeting with its multiple modalities can be used to foster language learning, catering for learner differences and preferences. For this potential to be realized, users have to acquire appropriate literacy skills, and explicit training in these strategies and coping mechanisms has to be provided to enable students as well as teachers to successfully use multimodal synchronous online tools. To communicate successfully, learners as well as teachers have to take control of tools and adapt their ways of interacting. However, more research is needed to identify the precise skills and strategies that can be taught to teachers and students in preparation for online language tutorials (see Hampel and Stickler, 2005). As Scheffel-Dunand (2006: 353) points out, 'literacy and orality are social practices whose forms and functions correlate with context of situation, personal involvement and audience.' This context includes the communication environment with its particular mix of modes which has an impact on how these are best used (Kenning, 2010). New patterns of communication emerge as a result, for example, the combination of modes in complex IRF exchanges. Goodfellow and Lea (2007) describe online environments as 'sites of practice' rather than 'tools for interaction'. In this project, we have seen how the environment shapes interaction, how users adapt the available tools to their purposes and how different modes can be used in a complementary, compensatory and competing manner.

Acknowledgements

References

Blake, R. (2005) Bimodal CMC: The glue of language learning at a distance. *CALICO*, **22**(3): 497–511.

Brouwer, C. and Wagner, J. (2004) Developmental issues in second language conversation. *Journal of Applied Linguistics*, **1**(1): 29–47.

Ciekanski, M. and Chanier, T. (2008) Developing online multimodal verbal communication to enhance the writing process in an audio-graphic conferencing environment. *ReCALL*, **20**(2): 162–182.

Cunningham, U., Beers Fägersten, C. and Holmsten, E. (2010) "Can you hear me, Hanoi?" Compensatory Mechanisms Employed in Synchronous Net-Based English Language Learning. *The International Review of Research in Open and Distance Learning*, **11**(1). http://www.irrodl.org/index.php/irrodl/article/view/774/1488

de los Arcos, B. (2010) *Emotion in online distance language learning: learners' appraisal of regret and pride in synchronous audiographic conferencing*. Unpublished PhD dissertation. Milton Keynes: The Open University.

Debski, R. and Levy, M. (1999) Introduction. In: Debski, R. and Levy, M. (eds.), *WORLDCALL: Global Perspectives on Computer-assisted Language Learning*. Lisse: Swets and Zeitlinger, 8–10.

Ellis, R. (2003) *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.

Firth, A. and Wagner, J. (1997) On Discourse, Communication, and (Some) Fundamental Concepts in SLA Research. *The Modern Language Journal*, **81**(iii): 285–300.

Flewitt, R., Hampel, R., Hauck, M. and Lancaster, L. (2009) What are multimodal data and transcription? In: Jewitt, C. (ed.), *The Routledge Handbook of Multimodal Analysis*. London: Routledge, 40–53.

Goodfellow, R. and Lea, M. (2007) *Challenging E-learning in the University*. Maidenhead: SRHE/Open University Press.

Hall, J. K. (1997) A consideration of SLA as a theory of practice. *Modern Language Journal*, **81**: 301–306.

Hampel, R. (2007) New literacies and the affordances of the new media: Using audiographic computer conferencing for language learning. In: Schneider, S. and Würffel, N. (eds.), *Kooperation and Steuerung: Fremdsprachenlernen und Lehrerbildung mit digitalen Medien* (Gießener Beiträge zur Fremdsprachendidaktik). Tübingen: Gunter Narr Verlag, 33–53.

Hampel, R. and Hauck, M. (2004) Towards an Effective Use of Audio Conferencing in Distance Language Courses. *Language Learning & Technology*, **8**(1): 66–82. http://llt.msu.edu/vol8num1/hampel/default.html

Hampel, R. and Stickler, U. (2005) New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, **18**(4): 311–326.

Heins, B., Duensing, A., Stickler, U. and Batstone, C. (2007) Spoken interaction in online and face-to-face language tutorials. *CALL (Computer Assisted Language Learning)*, **20**(3): 279–295.

Ho, D. G. E. (2007) *Classroom Talk: Exploring the Sociocultural Structure of Formal ESL Learning*. Bern: Peter Lang.

Jepson, K. (2005) Conversations – and negotiated interaction – in text and voice chat rooms. *Language Learning and Technology*, **9**(3): 79–98. http://llt.msu.edu/vol9num3/jepson/default.html

Jewitt, C. (2004) Multimodality and New Communication Technologies. In: Levine, P. and Scollon, R. (eds.), *Discourse and Technology: multimodal discourse analysis*. Washington, DC: Georgetown University Press, 184–195.

Kenning, M.-M. (2010) Differences that make the difference: a study of functionalities in synchronous CMC. *ReCALL*, **22**(1): 3–19.

Kreijns, K. P., Kirschner, A. and Jochems, W. (2003) Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, **19**: 335–353.

Lamy, M.-N. and Hampel, R. (2007) *Online Communication in Language Learning and Teaching*. Basingstoke: Palgrave Macmillan.

Lamy, M.-N. and Goodfellow, R. (1999) Reflective Conversations in the Virtual Language Classroom. *Language Learning and Technology*, **2**(2): 43–61. http://llt.msu.edu/vol2num2/article2/index.html

Lankshear, C., Gee, J. P., Knobel, M. and Searle, C. (1997) *Changing Literacies*. Buckingham and Philadelphia: Open University Press.

Lantolf, J. P. (2000) Second Language Learning as a Mediated Process. *Language Teaching*, **33**(2): 79–96.

Long, M. H. (1983) Linguistic and Conversational Adjustments to Non-Native Speakers. *Studies in Second Language Acquisition*, **5**(2): 177–193.

Long, M. H. and Robinson, P. (1998) Focus on Form: Theory, Research, and Practice. In: Doughty, C. and Williams, J. (eds.), *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press, 15–41.

Mercer, N., Littleton, K. and Wegerif, R. (2004) Methods for Studying the Processes of Interaction and Collaborative Activity in Computer-based Educational Activities.. *Technology, Pedagogy and Education*, **13**(2): 195–213.

Mondada, L. and Pekarek Doehler, S. (2004) Second Language Acquisition as Situated Practice: Task Accomplishment in the French Second Language Classroom. *The Modern Language Journal*, **88**(iv): 501–518.

Norris, S. (2004) *Analyzing Multimodal Interaction: A Methodological Framework*. New York: Routledge.

Örnberg Berglund, T. (2009) Multimodal student interaction online: an ecological perspective. *ReCALL*, **21**(2): 186–205.

Satar, H. M. (2011) *Social presence in online multimodal communication: a framework to analyse online interactions between language learners*. Unpublished PhD dissertation. Milton Keynes: The Open University.

Scheffel-Dunand, D. (2006) Bimodal Communication over Webcasts: From CSCL to CALL. *Computer Assisted Language Learning*, **19**(4+5): 341–355.

Sotillo, S. M. (2000) Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning and Technology*, **4**(1): 82–119. http://llt.msu.edu/vol4num1/sotillo/default.html

Stickler, U., Batstone, C., Duensing, A. and Heins, B. (2005) Distance and Virtual Distance. Preliminary results of a study of interaction patterns in synchronous audio graphic CMC and face-to-face tutorials in beginners' language tutorials. In: *Proceedings of the 2004 International Symposium on Applied Linguistics and Language Teaching*. Beijing and Shanghai: University of Northern Iowa, 740–750.

Stickler, U., Batstone, C., Duensing, A., and Heins, B. (2007) Distant Classmates: speech and silence in online and telephone language tutorials. *EURODL (European Journal of Open, Distance and E-Learning)* 2007(II). http://www.eurodl.org/materials/contrib/2007/Stickler_Batstone_Duensing_Heins.html

Stickler, U. and Hampel, R. (2010) CyberDeutsch: Language production and user preferences in a Moodle virtual learning environment. *CALICO Journal*, **28**(1): 49–73.

Stockwell, G. (2007) A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL*, **19**(2): 105–120.

Varonis, E. M. and Gass, S. (1985) Non-native/Non-native Conversations: A Model for Negotiation of Meaning. *Applied Linguistics*, **6**(1): 71–90.

Wang, Y. (2004) Supporting synchronous distance language learning with desktop videoconferencing. *Language Learning and Technology*, **8**(3): 90–121. http://llt.msu.edu/vol8-num3/wang/default.html

Wang, Y., Chen, N. S. and Levy, M. (2010) The design and implementation of a holistic training model for language teacher education in a cyber face-to-face learning environment. *Computers and Education*, **55**(2): 777–788.

Warschauer, M. (2000) On-line Learning in Second Language Classrooms: An Ethnographic Study. In: Warschauer, M. and Kern, R. (eds.), *Network-based Language Teaching: Concepts and Practice*. Cambridge: Cambridge University Press, 41–58.

Weininger, M. J. and Shield, L. (2003) Promoting Oral Production in a Written Channel: An Investigation of Learner Language in MOO. *Computer Assisted Language Learning*, **16**(4): 329–349.

Wertsch, J. V. (1991) A Sociocultural Approach to Socially Shared Cognition. In: Resnick, L. B., Levine, J. M. and Teasley, S. D. (eds.), *Perspectives on Socially Shared Cognition*. Washington, DC: American Psychological Association, 85–100.

Wertsch, J.V. (2002) Computer Mediation, PBL, and Dialogicality. (Special issue). *Distance Education*, **23**(1): 105–108.