ReCALL **22**(3): 293–312. 2010 © European Association for Computer Assisted Language Learning
doi:10.1017/S0958344010000170

The use of the webcam for teaching a foreign language in a desktop videoconferencing environment

CHRISTINE DEVELOTTE
INRP, 19 allée de Fontenay, BP 17424 – 69347, Cedex 07 Lyon, France
(email: cdevelotte@gmail.com)

NICOLAS GUICHON
Université Lyon 2, Bâtiment FILTRE, 5 avenue Pierre Mendès-France, 69500 Bron, France
(email: Nicolas.guichon@univ-lyon2.fr)

CAROLINE VINCENT
ENS-LSH, 15, parvis René Descartes, UMR 5191 ICAR, BP 7000- 69342 Cedex 07 Lyon, France
(email: caroline.vincent@gmail.com)

Abstract

This paper explores how language teachers learn to teach with a synchronous multimodal setup (*Skype*), and it focuses on their use of the webcam during the pedagogical interaction. First, we analyze the ways that French graduate students learning to teach online use the multimodal resources available in a desktop videoconferencing (DVC) environment to monitor pedagogical interactions with intermediate level learners of French in a North-American university. Then, we examine communicational and pedagogical aspects of this process which involves orchestrating different modalities and deploying various regulations for "semio-pedagogical" purposes. We define semio-pedagogical skills as the capacity to mediate a pedagogical interaction by combining or dissociating modalities (written, oral, and/or video) that are adapted to objectives and to the cognitive requisites of the task. We posit that these skills have to become part of the professional repertoire of future teachers, as they will increasingly be required to exploit the multimodal potentialities of online communication in their teaching.

The study draws on screen capture recordings of teacher trainee-student interactions and is completed by semi-directive interviews with teacher trainees (n = 5). It aims (1) to identify the importance of webcamming in the share of the pedagogical range available to teachers and (2) analyze the non verbal dimension of pedagogical communication via DVC.

The outcome of this study is the identification of five degrees of utilization of the webcam medium: there a certain gradation in the way webcamming is used (with a more or less significant use of image) when compared to other modalities. The different uses that are identified vary according to the perceived usefulness of webcamming to monitor teaching and to the teacher trainees' capacity to manage different workspaces.

Keywords: desktop videoconferencing, synchronous online teaching, webcamming, non-verbal communication, semio-pedagogical skills

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

1 Introduction

Synchronous online teaching requires a teacher to coordinate a set of complex operations in order to enable learning through communication via an audio-conferencing (Lamy & Hampel, 2007) or videoconferencing platform (Wang, 2004). This article will explore how teachers in training learn to teach with a synchronous multimodal setup. Both communicational and pedagogical aspects of these skills will be examined through an analysis of how students of education – currently learning to teach online – choose to use the multimodal resources available in desktop videoconferencing (DVC) environments in order to monitor online pedagogical interactions. The main focus of this study will be on the webcam image because of its crucial semiotic importance as a complement to speech and as a facilitator of comprehension thanks to its multimodal potential.

The possibility of maintaining a visual medium to see one's partner during videoconference mediated communication has spurred numerous descriptions of pedagogical experiences (Wang, 2004 & 2006; Jaraugi & Banados, 2008; Lee, 2007) and raised questions about the perception of a social presence in computer-mediated pedagogical communication (Yamada & Akahori, 2007) as well as about how to design the most appropriate tasks in a language learning situation (Wang, 2007). But in spite of these recent studies, Hrastinski and Keller (2007) point out that "there is a lack of research on non-text-based and synchronous media in computer-mediated communication (CMC) in education" and the gap is more obvious when it comes to training language teachers to use this specific tool. This article therefore proposes to contribute to the field of language education based on studies related to the learning process as well as social communication (Cosnier, 2008; Heath & Luff, 1992; Short, Williams & Christie, 1976).

As far as we know, little research has been conducted on the input and limits of webcamming for the purpose of learning a second language, beyond the psychological effects. This explains why our research, which aims to understand the cognitive elements at stake in webcamming, is limited to studies on cognitive psychology in multimedia (Mayer, 2005) and to drawing conclusions on studies concerning the input of video image in learning a second language and on pedagogical agents in a multimedia environment.

By analyzing the various techniques used by teacher trainees, this research aims to underline the diversity of articulations that operate between the different modalities offered by DVC – video, audio, iconic, textual. We posit that these "semio-pedagogical" techniques will become part of the professional skills required of future teachers who will be increasingly required to exploit the multimodal potentialities of online communication in their teaching.

2 The challenges of synchronous teaching via DVC

This section describes the online interaction situation and delineates the constraints it imposes on the teacher. We will then identify two types of challenges that teacher trainees have to face when they lead a synchronous multimodal teaching session: communicational challenges and pedagogical challenges.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press



Fig. 1. A teacher trainee's screen

2.1 The teacher trainees' work environment and its constraints

The context of this study is that of a French course at the university of Lyon 2 taught by graduate students (N = 10). These students are in their second year of Master's studies in Teaching French as a Foreign Language (TFFL) to American students from UC Berkeley (N = 16)[1]. As part of their training course, lasting eight weeks, teacher trainees were to develop their TFFL skills in an online learning environment involving a weekly session of individual computer-mediated videoconference with the UC Berkeley students who, in turn, were to practise their French language skills.

Teacher trainees managed different pieces of information during their synchronous teaching task, as shown in Figure 1. To avoid ambiguity, we distinguished between the teacher trainees' communicational tools[2] (Instant Messaging tools, blogs) and the different modalities they used (audio, text, webcam image, still image)[3]. *Skype* was

[1] For the details of this action research see Develotte, Guichon & Kern (2008).

[2] We propose a general definition of communicational tool as a technological device or facility used to exchange information through different modalities. Some communicational systems, like Skype, include several tools: written chat facility, oral chat facility and webcam.

[3] The authors are aware that "multimodality" is defined in a slightly different manner by such researchers as Kress & Van Leuwen (1996) but the present study is in line with the definition of "tool" and "multimodality" widely adopted among the French researchers working on communication (Lancien, 2000; Cosnier, 2008).

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press



Fig. 2. A teacher trainee at work

chosen as the platform for social communication and a collective blog[4] was used to present two categories of participants (learners from UC Berkeley and teacher trainees from Lyon 2) as well as to store documents (pictures or weblinks to online videos) to be used for the creation of language learning tasks.

In the example shown in Figure 1, the teacher trainee[5] organized her workspace in three main areas. The first area (area A), on the left, is a written file in which the teacher trainee listed different types of information that help monitor pedagogical interaction:

- the link to the blog where various resources that will serve as a basis for interaction are stored as well as the password to have access to them;
- detailed information about each task used during the teaching session (instructions, main issues, lexical and grammatical items).

The second area (area B) is composed of a textual *Skype* chat presenting vocabulary that will complement interaction between teacher trainees and learners.

Finally, the third area (area C) overlaps a little on the former and contains the *Skype* webcam image of the learner (top picture) and the smaller webcam image (bottom picture) of the teacher trainee.

Figure 2 illustrates a teacher trainee at work and underlines the complexity of her work environment. It shows that CMC integrates several communication tools (*Skype* and blog) into the limited space of a computer screen for various uses.

It is worth noting that the teacher must also use other types of offline tools in this pedagogical communication setup. Those are evenly distributed in the work space and are listed in Table 1 according to their modalities and their function during the teaching session.

[4] www.apprentissageenligne.blogspot.com

[5] The participants explicitly gave their consent to have the pictures of their faces used for the purpose of research and publication.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

Table 1 *Offline tools used by the teacher trainees*

Tools	Lesson outline + private notes	Documents (stimuli)
Functions	- Help manage the pedagogical interaction (i.e. list of instructions, key words, description of the activities, time indications…) - Take notes of linguistic errors for feedback	- illustrate cultural elements - trigger language activity
Modalities	Textual	Image Textual
Temporality	Prepared before and used during the synchronous interaction	

Articulating two spaces in the same work environment – the first on the screen and the second around the computer (papers, books, pens, etc.) – enables the teacher trainee to pay equal attention to simultaneous ongoing operations. In order to achieve multiple pedagogical objectives at a given time, the teacher trainee must moderate pedagogical conversation, monitor herself (see 2.3.1), manage tools, and manage resources. The temporal constraints, the risk of an unreliable internet connection, and the difficulty in managing several simultaneous tasks make it particularly interesting to observe and study the strategies adopted by the teacher trainees.

The teacher trainee must overcome numerous difficulties. These include multitasking, interrupted connections, and various distractions, as well as varying degrees of complexity and cognitive load linked to completing the task (Oviatt *et al*., 2004).

In their analysis of students' reading behaviors, Scollon *et al*. (1999: 35) refer to what they call the 'polyfocality' of attention. This characteristic is particularly significant in CMC and constitutes one of its intrinsic elements according to Jones (2004: 27) who contends that "polyfocality seems, in fact, to be part of the very ethos of new communication technologies".

2.2 The communicational challenges

2.2.1 The interactional framework of VDC mediated pedagogical exchanges. In the context of this complex working situation, this article proposes to focus on aspects more specifically linked to webcamming, and to question the impact that this situation can have on pedagogical communication, as well as on the challenges that confront teachers during its use. Webcamming is defined here as the interlocutors' voice and image accessible through two of the DVC windows allowing for potentially rich interactions (see section 2.3).

Basing his study on the theoretical framework previously developed by Goffman (1994), de Fornel (1994) shows that webcam-mediated interaction establishes a "virtual co-presence". He also underlines the distorted characteristics of the situation when compared to a normal face-to-face situation. De Fornel stresses the necessity to maintain a "common interactional ground" in order to see and to be seen.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

He views webcamming as "an interactional artifact that restructures interactional activity itself and ultimately the nature of the practical tasks that interlocutors are confronted with" (de Fornel 1994: 126).

Licoppe and Relieu (2007: 19) push Fornel's theory further when they claim that "learning the necessary skills to monitor webcam-mediated interactions leads to the adaptation of one's interactional skills to this kind of medium". Such skills involve choosing the distance to adopt from the screen and using the appropriate body language. Dwelling on these studies, we will therefore investigate how the communication tools are used by teacher trainees in the conduct of pedagogical tasks, with a particular focus on the use of the webcam.

2.2.2 The effects of presence. Several authors have chosen to explore the feeling of virtual co-presence specific to DVC. Thus, Weissberg (1999) wrote about "presence at a distance" phenomena which, he claimed, "do[es] not reproduce the performances we usually accomplish. [Presence at a distance] invents *another realm of perception* […]" (Weissberg, 1999: 14). He means that talking, seeing, and listening in such setups refer to different activities from those we usually experience in our daily lives.

In other words, one could hypothesize in line with Weissberg's research that a webcam mediated communication situation would eventually lead to the development of a specific interactional body language, one that is adapted to this other realm of perception.

Jones (2004: 23), on the other hand, stresses that the control of the presence at a distance facilitated by technology modifies the general context of communication: "What makes communicating with new technologies different from face-to-face communication is […] the different sets of 'mutual monitoring possibilities' that these technologies make available, the different ways in which they allow us to be *present* to one another and to be aware of other people's *presence*". In referring to research by Short, Williams and Christie (1976), Jonassen *et al.* (2005: 266) define social presence as "the degree of salience of the other person in the mediated interaction and the consequent salience of the interpersonal relationship". This point of view refers to how the image of others is received in this type of setup. In this study, it is also worth looking into the production of pedagogical discourse and into the way the teacher trainee uses webcams. Just like these degrees of salience, we believe that the degrees of investment in webcam use during pedagogical communication can be identified.

2.2.3 Non-verbal communication. The degree to which studies focused on face-to-face (f2f) interaction can be applied to DVC has yet to be assessed, but certain strands of f2f research appear to be directly relevant. Cosnier, for instance, studied the non-verbal dimension of dialogical situations. According to Cosnier (2008: 120–124), one can distinguish between the "gestures that accompany discourse" and the "gestures that monitor interaction".

The former refer to the "mimetic gestural activity that is linked to the genesis of the utterance to which it is integrated: first by the deictic or pointing gesture […], then by the illustrative body language that mimes the action or represents in space certain characteristics of the signified […] (e.g. "spiral staircase"). Other co-verbal gestures are directly linked to abstract thought and are therefore characterized as "metaphorical" ("a lot").

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

The second category specifies which gestures serve to monitor interaction. Indeed, gestures not only support discourse but maintain the interaction between the interlocutors, ensure mutual exchange and foster empathy. These gestures mostly rely on nods, smiles and eye movements. This coordination system is already recognized to be important in traditional pedagogical situations, and plays a determining role in online multimodal pedagogical communication.

2.3 Pedagogical challenges

The teacher's control of tools and graphics interface constitutes a basic skill that Warschauer (2002) called "computer literacy". An online teacher will also develop a specific pedagogical skill by mediating interaction with learners through modalities (written, oral, and/or video) adapted to objectives and to the cognitive requisites of the task (Guichon & Drissi, 2008). We propose to define this skill as a "semio-pedagogical skill" in line with Peraya's (1998) work.

The management of multimodal tools by teachers is not specific to CMC. But in a situation using videoconferencing facilities, the main novelty comes from the teacher's monitoring of pedagogical interactions, for instance by providing instructions, explanations or feedback, via the webcam. The notion of monitoring echoes that of cognitive guidance presented by Mayer (2005: 12), who claimed that the role of a teacher is not limited to feeding information but also contributes to guiding the learner.

This monitoring occurs in production when teacher trainees facilitate feedback operations (Develotte, Guichon & Kern, 2008) with smiles, nods, facial expressions or frowns. In reception, they give paralinguistic clues helping the learners' comprehension (Guichon, 2009).

2.3.1 Webcamming for language teaching/learning purposes. In this section, we will concentrate on the significance of the interlocutors' image in a synchronous videographic setup for both learning and teaching. After defining the inherent characteristics of videoconferencing, we wish to determine the informational input of image and the cognitive treatments it requires based on cognitive and psycholinguistic research. The interlocutors' image will be designated as "webcam image" in order to differentiate it from the video image of documentaries or films.

As a starting point, it is suggested that a webcam image presents the following characteristics:

- it takes place in a simultaneous fashion and is interactive with other semiotic modes on the limited space of a computer screen, thus creating a spatial and temporal contiguity with a positive impact on learning (van Merriënboer & Kester, 2005: 84).

But,

- it is ephemeral (unlike a video image, it is not repeatable unless it is recorded);
- its informational content is poor because it is restricted to a close-up shot of the interlocutor leaving all precious contextual information such as clothes, body language and the interaction context off screen (Zähner, Fauverge & Wong, 2000; Lamy & Hampel, 2007: 77);

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

- it is seldom fluid and sometimes creates micro-gaps in oral production and its visual completion, which is particularly harmful when learning a foreign language (Ruhleder and Jordan, 2001);

Until now, research on webcamming in the context of learning a second language has focused primarily on psychological aspects, based mainly on interviews or questionnaires asking students to evaluate the importance of image after the event. Webcam images were perceived to allow for an individualization of the relationship (Wang, 2004) and lessen the sensation of isolation often felt when learning takes place mainly online or at a distance (Carlson, 1997). This visual link between teachers and learners can also contribute to reinforcing the learner's motivation (Marcelli, Gaveau & Tokiwa, 2005) and creating more depth in oral exchanges (O'Dowd, 2006).

If Mayer's cognitive theory of multimedia learning (2005:31) is applied to a webcamming situation, it appears that a learner faced with an online learning situation in a synchronous videographic environment will complete the five following operations:

- select appropriate words from the chat text or the teacher's oral message;
- select the appropriate images from the teacher's webcam image and other visual information appearing on screen (smileys, pictures, etc.);
- organize the selected words in a coherent verbal representation;
- organize the selected images in a coherent visual representation;
- link visual and verbal information with former knowledge.

As this list suggests, several complex cognitive operations are required from the learner, and it is important to be cautious when analyzing the impact of webcamming on the comprehension process of learners. Furthermore, the informational input of a webcam image for a language learner must be determined in relation to the oral message and the necessary cognitive treatment appraised, whether the webcam image of the teacher be redundant (correspondence between the image of the teacher's face and the oral message – see van Merriënboer & Kester, 2005: 82) or distracting (it contradicts the oral message, either because it is senseless and adds nothing to the oral message or because it actually diverts the learner's attention).

This article examines the different modalities available to teacher trainees. These modalities are not to be considered as isolated but as an array of possibilities in the semiotic range that can be strategically combined or not, so as to be tailored to the learners' needs and the pedagogical strategies used. The impact of an individual screen webcamming platform on the semio-pedagogical mediation of teacher trainees must then be determined.

2.4 Synthesis

A teacher must develop particular semio-pedagogical skills in a videoconference situation. These skills correspond to the capacity to display a set of appropriate signs in a given time in order to facilitate learning. This is assuming that the teacher learns to coordinate her pedagogical action between the different means available to her (voice, facial expressions, gestures, images, text, etc.) and the different tools (webcam window, textual chat, and so on), and that she is capable of making choices (associating or

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

dissociating) between the different modalities according to her objectives. It is also worth noting that the pedagogical relationship is decisive in this type of communication because the protagonists interacting are not equal. In fact, a teacher is responsible for the comprehension of a message by the learner. Finally, the fact that the participants in this experiment are in the process of learning their profession makes it an ideal testing ground to observe the development of their semio-pedagogical skills.

The rest of this study proposes to analyze the way in which teacher trainees seize opportunities given by webcamming to teach a second language, and find answers to the following questions:

- What share of the pedagogical range available to teachers does webcamming take?
- How was webcamming explored by the teachers?
- What is the non-verbal dimension of pedagogical communication via DVC?

3 Methodology

3.1 Data collection

An empirical approach was adopted to ensure that the researchers' work would influence that of the participants as little as possible. It also appeared essential to preserve the trainees' work situation so that the instruments necessary for collecting data had as little impact as possible on training. Every teacher trainees's activity was thus recorded with a screen video recorder.

Several factors contributed to defining the coherence of our data collection:

- Our research concentrated on five teacher trainees (out of a total of eleven) teaching individually (the other six were teaching in pairs); the trainees were four women and one man, respectively aged 25, 26 (two trainees), 31 and 45.
- Screen captures were taken in Lyon 2 and in UC Berkeley in order to compare what was happening on the teacher trainee's screen and on the learner's; however, because the screen video recorders failed at times, data could only be collected from 60% of the total interactions.
- Our research focused on one of the last teaching sessions – when the teacher trainees were already familiar with their work environment – to study the use of multimodality; the sixth session (out of eight) was therefore chosen, as five synchronous interactions had already taken place and offered the opportunity of developing interpersonal knowledge between interlocutors.
- Because our research team wished to work on finer aspects of interaction, such as eye focus, the database was limited to a single task identical to all five teacher trainees. During this task, learners had to imagine a holiday destination, what they would do there and what they would take. The choice of this particular task was motivated by the fact that it implied having to go on the blog in order to look at pictures and having to work with one's partner to try to locate the place that was illustrated. The task was therefore liable to lead to a close analysis of monitoring behavior on screen and different pedagogical resources.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

The database chosen for the behavioral study of five teacher trainees represents 38 minutes 23 seconds of screen captures. A descriptive perspective was therefore adopted to frame our research, seeking answers to the research questions which were as precise as possible from a reduced database.

Moreover, a series of semi-directed interviews was held with each of the teacher trainees, in order to compare the analysis of the teacher trainees' behaviors with their perception of the way they teach. This added another two hours and 35 minutes of interviews to our corpus. Some of these data will be presented toward the end of this article in order to develop the analyses of the behaviors analyzed here.

3.2 Protocol analysis

We chose to observe how teacher trainees associated and combined the different tools (text chat, webcamming) and the different textual and iconic documents available to them. Because the teacher trainees were simultaneously in production and reception situations, our research distinguished between them, as indicated in the table below:

Table 2 *Modalities* in reception and production from the teacher trainee's point of view*

Modalities	Reception	Production
Audio only	The interlocutor's image does not appear on screen, the exchanges are only oral	The interlocutor's image does not appear on screen, the exchanges are only oral
Audio-video image of the interlocutor	Eyes on the interlocutor's image, meanwhile participating in oral exchanges	Eyes on the webcam or on the interlocutor's image, meanwhile participating in oral exchanges
Textual (chat)	Reception of written message	Production of written message
Window concealing the interlocutor's video image (textual, video, audio)	On learners' screens	On teacher trainees' screens

*We define modality as the type of semiotic representation (textual, aural and visual) used to exchange information. Multimodality makes sensory information accessible in diverse semiotic codes and offers the opportunity to comprehend information through different channels (Pudelko *et al.*, 2002: 42).

Teaching involves a double spatial and temporal constraint: on the one hand, there is pressure from the limited time of each 45-minute session and, on the other hand, the size of the screen imposes limits and forces teacher trainees to select which elements will appear on the screen or not. ELAN multimodal transcription software was used in order to study the way in which teacher trainees used the screen for each modality (video, audio and chat). This software allowed us to synchronize the different modalities and to show how teacher trainees used each one of them[6].

[6] Elan Linguistic Annotator Max-Planck-Institute for Psycholinguistics. (http://www.lat-mpi.eu/tools/elan/download)

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

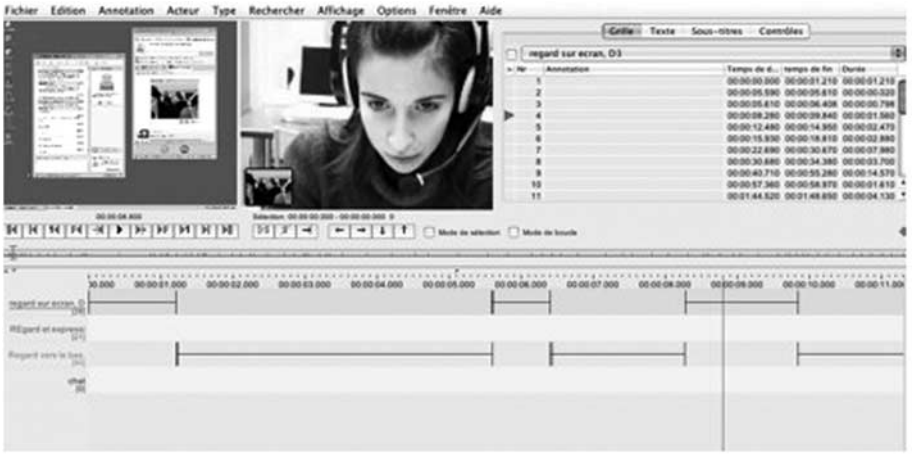


Fig. 3. Example of a multimodal annotation using ELAN

The image in Figure 3 shows two videos (top left) that allow the analyst to see what the teacher trainee is looking at and where her eyes focus (in this case on the computer screen). The bottom of the image shows a timer where we added markers to indicate where her eyes focused during the interaction and to provide precious information about the gaze duration.

4 Results

First, we will provide a general overview of the ways teacher trainees use the different communication tools. Then, we will closely explore the way in which webcamming is used. A series of interview extracts illustrating the teacher trainees' behaviors will supplement our presentation of these results.

4.1 Use of time for each medium

Table 3 represents how long the different teacher trainees used each modality in terms of percentage, except for the textual chat measured in terms of the number of times it was used. We decided to count the frequency and not the running time of chat use, as it would have called for the integration of variables (teacher trainees' typing speed, sentence length, breaks in writing) that would not have allowed a valid comparison of the use of chat, with the number of productions depending on each teacher trainee.

Three types of behavior were observed: (1) audio-oriented, (2) alternating audio and webcam, and (3) webcam-oriented.

- The first type of behavior is that of a trainee who favors an audio medium. Therefore, one of the teacher trainees (T1) only looks at his interlocutor's image for a quarter of the time and chooses not to look at the learners' image for the remaining three quarters.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

Table 3 *Time use for each modality by the teacher trainees*

		T1	T2	T3	T4	T5
Total time for task		7:46	9:40	8:47	6:00	6:10
Webcamming		2:32 (26%)	5:14 (54%)	6:32 (74.5%)	6:00 (100%)	6:10 (100%)
Audio only	Document concealing the interlocutor's image Reduced size video window	5:10 (74%)	4:26 (46%)	1:42 (19.5%)	0	0
Chat use frequency (number of times)		2 times	1	3	0	7

- The second type of behavior corresponds to a mixed use. Two teacher trainees are included in this category (T2 and T3): although they mainly use the webcam, they sometimes conceal the learner's image on screen (19.5% and 46% of the time respectively) with another tool (e.g., the blog). Those two teacher trainees thus use different tools in dealing simultaneously with their communication potential.
- Finally, the last type of pedagogical communication corresponds to an exclusive use of webcamming. Two teacher trainees (T4 and T5) used the webcam to its full potential and never concealed the image of their interlocutor to switch to another window.

Chat was used in very different ways by the teacher trainees in this task (0 to 7 times). Although the use of textual modality will not be developed further here (cf. Develotte, Guichon & Kern, 2008), it appears that the mere availability of a given tool does not necessarily imply its use.

While some teacher trainees appear to use *Skype* in a similar fashion as they would use a telephone, thus favoring audio communication, others push webcam potentialities further by exploiting the dynamic image they receive from their interlocutor and that they in turn send back to them. This begs the question of the consequences of these choices at an interactional and pedagogical level. Because webcamming remains a fairly unexplored subject, we chose to focus our analyses more specifically on the use of this medium by our teacher trainees.

4.2 The use of webcamming: five degrees of utilization

A graduated scale measuring the different degrees of utilization of a webcam medium was designed in order to understand more precisely the use made of webcamming by teacher trainees.

The video window on a computer screen can indeed be used in various ways. Some teacher trainees can choose not to look at it or not to appear on camera themselves, while others will choose quite the opposite and use the medium as much as possible by bringing gestures and facial expressions into play to back up their message.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

Below are the different degrees of utilization established, as illustrated by the screen capture images included in appendix 1:

- Degree 0: the teacher trainee does not appear on the video window, she is standing outside the camera focus or it is not possible to use a video medium.
- Degree 1: the teacher trainee does not look at the computer screen.
- Degree 2: the teacher trainee looks at the open video window on the computer screen.
- Degree 3: the teacher trainee looks at the open video window on the computer screen and she uses facial expressions and/or gestures to back up her message. This degree will be described later in more detail.
- Degree 4: the teacher trainee looks straight into the webcam, giving her interlocutor the impression that she is looking directly at her.

When this graduated scale of video window utilization is applied to the database, the results are shown in Table 4.

The first two degrees (0 and 1), labelled here as "non-utilization of the video window," refer to the time teacher trainees spent off-screen or on-screen with no utilization of the image. Two teacher trainees thus spent half their time without using the image representing their presence. In other words, these teacher trainees' American learners did not have access to information provided by the mimetic gestures of their teacher trainee. It is important to underline that the time period during which the teacher trainee does not use a video medium corresponds to a time period during which she concentrates on other things, such as taking written notes on mistakes for the final assessment, or reading lesson plans on paper.

The last three degrees (2 to 5) refer to a gradation in the utilization of webcamming. Degree 2 concerns two teacher trainees and lists all the moments when teacher trainees look at their interlocutors on screen without any particular expression.

Degree 3 is the most frequent utilization degree among three of the five teacher trainees and corresponds to the teacher trainee's wish to integrate the potentialities of webcamming into her pedagogical activity. The gestures and facial expressions linked to the utilization of a video window are used to their maximum capacity by two of the teacher trainees.

According to degree 4 of our graduated scale, the teacher trainee uses video by looking at the camera and not the screen, thus giving her interlocutor the feeling that she is looking directly at her. This would suggest that the affordance of webcamming has been integrated by this teacher, even if tentatively, in order to foster empathy.

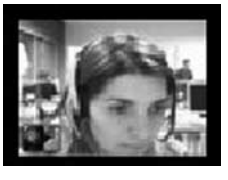
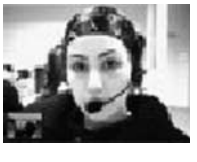
Graduation from zero to four seems to imply that the use of webcamming is at its full potential when reaching degree four, but the authors wish to make clear that each degree corresponds to a specific moment of the task as it is monitored by the teacher trainee at one point of his/her training.

4.3 Gestures and facial expressions in webcamming utilization

All teacher trainees use nods and inviting facial expressions to create empathy as well as to encourage or sustain the learner's speech. Teacher trainees also produce empathic expressions in order to monitor interaction, with facial expressions suggesting incomprehension, interrogation or the wish to speak, for instance.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

Table 4 *Distribution of teacher trainees according to their utilization of the video window*

		T1	T2	T3	T4	T5
Total time for the task		7:46	9:40	8:47	6:00	6:10
Degree 0 	Non utilization	1:20 (17%)	3:36 (37%)	2:35 (29.4%)	0	0
Degree 1 		2:00 (26%)	2:37 (27%)	1:45 (20%)	0:25 (7%)	0:50 (13.5%)
Degree 2 	Utilization	0	3:22 (34.5%)	4:25 (50.3%)	0	0
Degree 3 		3:38 (47%)	0:15 (2.5%)	0:02 (0.3%)	5:35 (93%)	5:20 (86.5%)
Degree 4 		0:48 (10%)	0	0	0	0

Only two teacher trainees (T4 and T5) use webcamming to illustrate the content of their speech with gestures, such as counting on their fingers, putting their hands on their heads to express feigned despair, pointing at one of their ears with a look of incomprehension, pointing at a garment to illustrate a vocabulary item.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

An analysis of the teacher trainees' gestures brings different pieces of information: facial expressions or gestures take on various empathic and interactional functions. They coordinate exchanges of speech (evident wish to speak) and their contents, guiding discourse (the teacher trainee's expression of surprise aims at bringing the learner to reformulate), encouraging speech (nods, smiles) or correcting it (expression of incomprehension) in turn. They thereby participate in the construction of discourse. Moreover, because aspects of gesture such as laughs and smiles are strong socio-affective indicators, they help construct an interpersonal relationship between the teacher trainees and their students.

4.4 Teacher trainees' perceptions of their behaviors

Interviews with teacher trainees help explain these behaviors. Although T1 favors the audio modality, she seems to use the webcam image in a very specific way. When asked if eye contact was important between the learners and herself, T1 replies: "*When I'm the middle of a task, I look at the documents that they the learners are working on or the notes I'm taking, I leave them aside a little. When they speak I don't necessarily look at them, I let them talk. Also, when I do look at them, it is necessary to look at the camera to give them the feeling I'm looking at them directly, which isn't easy. Sometimes I'd watch them on screen*". Although she uses the webcam image only 26% of the time during the interaction, she seems to give primary importance to video in production rather in reception. When she does look at her learners (see Table 4), it is especially for the pedagogical impact it will have on them. It is interesting to note that T1 is conscious of the presence effects created by the webcam, as she is the only teacher trainee who looks directly at the webcam rather than at the image of her interlocutors.

T2 and T3 were both classified as engaging in mixed behaviors, using the different affordances of the communication setup. Their image remained hidden from their interlocutors for a third of the total time allotted to the task. When they looked at the screen, they had no marked facial expressions and they used very few gestures.

Interviews revealed that T2 was conscious of the fact that she had not always looked at her interlocutors: *"Yes, they the learners looked at me all the time and I wasn't always looking at them because I had other things to manage on my side: the keyboard, the pedagogical record sheet with all our activities and the time slots allotted to each. I looked at them from time to time to reassure them"*.

T2 is thus conscious of the fact that her learners look at her and of the positive impact that her image can have on them. However, it is noticeable in her remarks that she doesn't realize the benefits it could have for her learners' production.

T3 looked alternately at the screen and elsewhere. When asked which *Skype* tool he most used, T3 replied that he used audio the most but still considers video as important: "*Video is important even if I don't look at it all the time, there are gestures. I am limited because of all the time I spend on audio when they the learners talk, although I do look at them when they're having a mental block and that's a way for easing mutual understanding*". Although he favors the use of audio in his monitoring of pedagogical tasks (74% of the interaction), T3 uses the webcam from time to time to solve any comprehension problems his learners may encounter.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

If tables 3 and 4 are compared, it appears that the two teacher trainees who completely favored webcamming over other tools (T4 and T5) were never off screen and used many facial expressions, and especially mimetic gestures, in front of the camera.

The following explanations to these types of behavior were given during the series of interviews:

T4 admitted she considered the multitude of available communication tools difficult to manage: "*I had trouble looking at them the learners, listening to them, typing on the chat and taking notes at the same time so I used the chat very little, except when I was encountering sound difficulties*". T4 acknowledged her preference for image whether it was used for production or reception: "*To me video is essential and is akin to the way we behave in life. I personally can't talk to people if I don't see their eyes and I don't see them. If I don't see them, I feel uncomfortable. That's why I had trouble using the chat, the pencil, etc. I have to look at them to see if they understand or not, if they have a question, if they hesitate. I can feel it if I see them*". Therefore, when managing multimodality, T4 adopts a strategy defined on the one hand by what appears appropriate for this type of pedagogy and on the other hand by what is possible to achieve in this demanding cognitive situation.

T5, however, said she had favored video to make the situation seem more natural and as close as possible to a face-to-face situation. When asked: "*Is the webcam image important to you?*" she replied: "*Many things are transmitted by your expressions, by your face. It is a much more natural way of communicating and when you see your students' faces you can tell if they understood or not, and they can tell if it's good or not if you smile…*". T5 appears to have integrated the technological setup into her pedagogical and communicational range. A comparison between the different behaviors observed and the teacher trainees' comments shows that there are as many points of convergence as there are differences: this may indicate that teacher trainees are acquiring a certain number of professional skills but that those skills have not yet been stabilized. It may also denote differences between environmental conditions and the communicational preferences of teacher trainees. Most interviews convey the idea that the more complex the teacher trainee's activity gets, the more important the cognitive cost will be because it requires alternating resources and multiple modalities.

5 Discussion

This article has identified a range of behaviors in the use of DVC-enabled multimodality for a given task (see Table 3). This range showed a certain gradation in the way webcamming was used (with a more or less significant use of image) when compared to other modalities (written or oral only) and according to the perceived usefulness of webcamming to monitor teaching and to the teacher trainees' capacity to manage different workspaces. Other variables play a role, such as personal communicational habits and professional experience.

As shown earlier, there are several degrees of utilization of webcamming (see Table 4), which can either be used by the teacher trainees to be seen by the learners or to see them. By separating the reception and production of the webcam image of teacher trainees, it appeared that all teacher trainees chose to stop using the webcam image at one moment or another in order to focus on an audio medium only. It appears

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

that the use of a webcam image is more important in terms of its availability as a possible resource in case of need than as a favored type of communication. It can even prove to be useless for some teacher trainees.

Webcam images are complementary when they contribute to the information contained in an oral message. Multimodality depends on the interactivity between, albeit qualitatively different, visual and verbal information. The achieved interactivity will contribute to comprehension as long as learners are capable of registering these visual and verbal representations (Mayer, 2005: 5). In the case of DVC, the supplementary information brought by the teacher trainee's webcam image is either psychopedagogical (feedback input), cultural (cultural elements inherent to a native speaker of the target language), or linguistic (a gesture adds to orally delivered information).

However, we can wonder whether webcam images take on additional value only if they are used. Indeed, the teacher trainee's webcam image may be distracting, either because it contradicts the oral message, because it makes no sense and adds nothing to it, or because it distracts the learner's attention. But this aspect needs further investigation as it has not been demonstrated in this study.

When webcamming is utilized, it plays a major part in the socio-affective dimension of pedagogical communication and in the development of interpersonal relationships. Webcamming creates presence at a distance, installs an obvious connection between the participants and, furthermore, develops the quality of the pedagogical relationship as suggested by the series of interviews presented in this article. Moreover, a detailed analysis of the mimetic gestures of teacher trainees shows how their semio-pedagogical range is developed and helps monitor interaction. The video window can be compared to a theatre stage that the teacher trainees use to enact their role: they learn to adapt their gestures to the size of the stage. Some teacher trainees consider there to be some 'staginess' to this particular tool, as mentioned earlier.

5.1 Limitations of the study

From a methodological point of view, the focus was placed on the use of multimodality by teacher trainees. However, it would also be pertinent to take the learners' expectations and preferences into account. The combination of each participant's point of view will make it possible to determine which pedagogical techniques are best adapted to each specific task. Following the model set by Foulon-Molenda (2000), it would also be interesting to isolate other elements such as the influence of the learner's and the teacher trainee's gender on the use of webcamming.

Furthermore, this study relies on a limited number of participants and on a single sample of a much longer pedagogical experiment. Maximum scientific use was made of non-normalized data usually found in a well constructed database developed for a specific purpose, such as that used for the experiments led by Yamada and Akahori (2008). It was necessary here to break down all elements of DVC so as not to minimize its complexity and focus more specifically on webcamming. This study illuminates the fact that a pedagogical communication situation is liable to generate different utilizations of webcamming according to the teacher trainees' preferences, the pedagogical potential they grant it, and the confidence developed as to the use of different resources and simultaneously generated modalities.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

6 Conclusion

This study measured the complexity of the teacher trainee's task in a foreign language training setup based on DVC. Five degrees of utilization of webcamming were identified in this type of online synchronous pedagogical interaction. Yet, this study alone does not allow us to make clear recommendations as regards the optimal use of webcamming for language teaching. The attention of online teachers should nevertheless be drawn to the fact that degree 4 allows for an especially intense interaction and augments the feeling of co-presence. Given the complexity of the task, it is impossible – and possibly not desirable – to sustain such intensity for a very long time. Over time, online teachers will have to become aware of the potential of the webcam for different types of regulation. By learning to adjust the tool to their objectives and to the relation they wish to establish with their distant students, online teachers will learn to develop their own semio-pedagogical competence and maximize the advantage of webcamming for online language teaching.

The results of this study allow teacher trainers to define three skill bases in terms of sociocultural theory (Lantolf, 2000): namely, other-regulation skills, self-regulation skills and media regulation (Henderson & Cunningham, 1994; Guichon, 2009).

Other-regulation skills correspond to the capacity to evaluate the effect one has on a distanced interlocutor, using the appropriate semiotic system to maintain a learner friendly environment and contribute to learning, and to measure the comprehension level of a learner in order to adapt one's pedagogical strategies.

When it comes to self-regulation skills, the series of interviews in this research shed light not only on the complexity of the task but also on the psychosocial aspects in play for teachers confronted with their own image. The following are classified in this category: the capacity to manage (and bear) one's image, the capacity to use various tools quasi-simultaneously and the capacity to deal with uncertainty due to an unreliable network and potential technical problems.

Finally, media regulation skills involve the capacity to utilize one's own image depending on the semio-pedagogical objectives that are targeted or the capacity to dissociate or combine the modalities according to the task or the learner's needs. As our typology suggests, online language pedagogy requires basic teaching skills that are formed by the technological potentialities of DVC, thus calling for future online teachers to be "taught online communication strategies, including non-verbal skills such as the use of body language and facial expressions in order to optimize the effective use of DVC" (Lee, 2007: 285). In defining these three skill bases, this article contributes to current debates on the training needs for teaching online as a future profession, a profession gradually defined by the advances of research on the limits and potentialities of the internet as a tool for language teaching.

Acknowledgments

The authors would like to thank Marie-Noëlle Lamy from the Open University (UK) for her valuable suggestions on reading drafts of this article as well as the two anonymous reviewers for their thoughtful comments. The study is part of ITHACA, a research programme funded by the French Agence Nationale de la Recherche,

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

project led by Yannick Prié (LIRIS, Lyon 1), Nicolas Guichon (ICAR, Lyon 2) and Mireille Bétrancourt (TECFA, Université de Genève).

References

Carlson, R. (1997) *Educating online: Creating the virtual classroom community*. ERIC ED 412 934.

Cosnier, J. (2008) Les gestes du dialogue. *La communication état des savoirs*. Auxerre: Sciences humaines, 119–128.

Develotte, C., Guichon, N. and Kern, R. (2008) "Allo UC Berkeley? Ici Lyon 2… Vous nous voyez bien?" Etude d'un dispositif de formation en ligne synchrone franco-américain à travers les discours de ses usagers. *Alsic*, **11**(2): 129–156.

ELAN Elan Linguistic Annotator developed by Max-Planck-Institute for Psycholinguistics Nijmegen, The Netherlands. (http://www.lat-mpi.eu/tools/elan/download).

de Fornel, M. (1994) Le cadre interactionnel de l'échange visiophonique. *Réseaux*, **64**: 107–132.

Foulon-Molenda, S. (2000) A-t-on besoin de se voir pour coopérer? Contribution théorique issue de la psycholinguistique. *Le Travail Humain*, **63**: 97–120.

Goffman, E. (1994) *Les cadres de l'expérience*. Paris: Minuit.

Guichon, N. (2009) Training future language teachers to develop online tutors' competence through reflective analysis. *ReCALL*, **21**(2): 166–185.

Guichon, N. and Drissi, S. (2008) Tutorat de langue par visioconférence: comment former aux régulations pédagogiques. *Les Cahiers de l'ACEDLE*, **5**(1): 185–217.

Heath, C. and Luff, P. (1992) Media Space and Communicative Asymetries: Prelminary Observations of Video-Mediated Interaction. *Human Computer Interaction*, **7**(3): 315–346.

Henderson, R. W. and Cunningham, L. (1994) Creating interactive sociocultural environments for self-regulated learning. In: Schunk, D. H. and Zimmerman, B. J. (eds.), *Self-regulation of Learning and Performance*. Hillsdale: Lawrence Erlbaum Associates, 255–281.

Hrastinski, H. and Keller, C. (2007) Computer-Mediated Communication in Education: A Review of Recent Research. *Educational Media International*, **4**(1): 61–77.

Jaraugi, K. and Banados, E. (2008) Virtual interaction through video-web communication: A step towards enriching and internationalizing language learning programs. *ReCALL*, **20**(2): 183–207.

Jonassen, D. H., Lee, C. B., Yang, C. C. and Laffey, J. (2005) The collaboration principle in multimedia learning. In: Mayer, R. E. (ed.), *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge university Press, 247–270.

Jones, R. H. (2004) The problem of Context in Computer Mediated Communication. In: Levine, P. and Scollon, R. (eds.), *Discourse & Technology multimodal discourse analysis*. Washington DC: Georgetown University press, 20–33.

Kress, G. and Van Leuwen, T. (1996) *Multimodal discourse: the modes and media of contemporary communication*. London: Arnold.

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

Lancien, T. (2000) *Multimédia : les mutations du texte*. Lyon: ENS Editions.

Lantolf, J.-P. (2000) *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Lee, L. (2007) One to one Desktop Videoconferencing for Developing Oral Skills: Prospects in Perspective. In: O'Dowd, R. (ed.), *Online Intercultural Exchange*. Clevedon: Multilingual Matters, 281–286.

Licoppe, C. and Relieu, M. (2007) Présentation. *Réseaux*, **144**: 9–22.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press

Marcelli, A., Gaveau, D. and Tokiwa, R. (2005) Utilisation de la visioconférence dans un programme de FLE: tâches communicatives et interactions orales. *Alsic*, **8**: 185–203.

Mayer, R. E. (2005) *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

O'Dowd, R. (2006) The Use of Videoconferencing and E-mail as Mediators of Intercultural Student Ethnography. In: Belz J. A. and Thorne S. L. (eds.), *Computer-mediated Intercultural Foreign Language Education*. Boston: Heinle and Heinle, 86–120.

Oviatt, S., Coulston, R. and Lunsford, R. (2004) When do we interact multimodally? Cognitive load and multimodal communication patterns. In: *Proceedings of the 6th international conference on Multimodal interfaces*. Pennsylvania State College, 129–136.

Peraya, D. (1998) Les campus virtuels. Principes et fondements techno-sémio-pragmatiques des dispositifs de communication et de formation médiatisées. Paper presented at the Conference «Dispositifs et médiations des savoirs», at the University of Louvain-la-Neuve.

Pudelko, B., Legros, D. and Georget, P. (2002) Les TIC et la construction des connaissances. In: Legros D. and Crinon J. (dirs), *Psychologie des apprentissages multimédias*. Paris: Armand Colin, 40–62.

Ruhleder, K. and Jordan, B. (2001) Co-Constructing Non-Mutual Realities: Delay-Generated Trouble in Distributed Interaction. *Journal of Computer Supported Cooperative Work*, **10**(1): 113–138.

Scollon, R., Bhatia, V., Li, D. and Yung, V. (1999) Blurred genres and Fuzzy Identities in Hong-Kong public discourse: Foundational Ethnographic Issues in the Study of Reading. *Applied Linguistics*, **20**(1): 22–43.

Short, J. A., Williams, E. and Christie, B. (1976) *The social psychology of telecommunications*. New York: John Wiley & Sons.

Van Merriënboer, J. J. G. and Kester, L. (2005) The four-component instructional design model: Multimedia principles in environments for complex learning. In: Mayer, R. E. (ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press, 71–93.

Wang, Y. (2004) Supporting synchronous distance language learning with desktop video-conferencing. *Language Learning & Technology*, **8**: 90–121. http://llt.msu.edu/vol8num3/wang/default.html

Wang, Y. (2006) Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL*, **18**(1): 122–145.

Wang, Y. (2007) Task design in videoconferencing-supported Distance Language Learning. *CALICO Journal*, **24**(3): 591–630.

Warschauer, M. (2002) A developmental perspective on technology in language education. *TESOL quarterly*, **36**(3): 453–475.

Weissberg, J. M. (1999) *Présences à distance*. Paris: L'Harmattan.

Yamada, M. and Akahori, K. (2007) Social Presence in Synchronous CMC-based Language Learning: How does it affect the productive performance and consciousness of learning objectives? *Computer Assisted Language Learning*, **20**(1): 37–65.

Yamada, M. and Akahori, K. (2008) Self-awareness effect on learner-centered communication in videoconferencing, *EuroCALL 2008*, 96–97.

Zähner, C., Fauverge, A. and Wong, J. (2000) Task-Based Language Learning Via Audio-visual Networks. In: Warschauer, M. and Kern, R. (eds.), *Network-Based Language Teaching: Concepts and Practice*. Cambridge: Cambridge University Press, 186–203.

https://doi.org/10.1017/S0958344010000170 Published online by Cambridge University Press