ReCALL **25**(1): 122–142. 2013 © European Association for Computer Assisted Language Learning
doi:10.1017/S0958344012000286

Multimodal language learner interactions via desktop videoconferencing within a framework of social presence: Gaze

H. MÜGE SATAR
Boğaziçi University, School of Foreign Languages, Güney Kampus, Bebek, 34342 İstanbul
(email: muge.satar@boun.edu.tr)

Abstract

Desktop videoconferencing (DVC) offers many opportunities for language learning through its multimodal features. However, it also brings some challenges such as gaze and mutual gaze, that is, eye-contact. This paper reports some of the findings of a PhD study investigating social presence in DVC interactions of English as a Foreign Language (EFL) teacher trainees. The case study approach involved the exploration of online interactions of five cases (pairs) within an interpretivist paradigm. Data collection included interviews, questionnaires and analysis of DVC recordings. The study emphasizes the importance of eye-contact in online multimodal communication to facilitate the establishment of social presence. Five types of gaze that were observed in learner interactions and participants' perspectives on eye-contact are reported. The conclusions include technical suggestions for the use of a webcam as well as pedagogical implications of online video interaction.

Keywords: Desktop Video Conferencing (DVC), Social Presence, Online Language Learning, Multimodality, Gaze

1. Introduction

Synchronous computer mediated communication (CMC) technologies have improved incredibly since the 1990s. Synchronous written interaction in the 1990s (e.g., MOOs and text-chat) was later supported by the development of audiographic tools which allowed synchronous audio, whiteboard, file sharing and desktop sharing. Audiographic conferencing has been considered effective for online tuition in distance learning environments (Hampel & Hauck, 2004; Coleman, Hampel, Hauck & Stickler, 2010).

1.1 Desktop videoconferencing (DVC)

Increasing bandwidth and ease of access to the Internet have enabled even more access to multimodal interaction, including videoconferencing. However, research investigating video interactions was relatively scarce until fairly recently. In a pilot

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

study conducted by O'Dowd (2000), videoconferencing was reported as a powerful medium to support intercultural interactions. However, between 2000 and 2007 there was little research on the use of video in CMC for language learning purposes. During this time Wang (2004a, 2004b, 2006, 2007, 2008) was the main author in the area exploring task design and negotiation for meaning via DVC. Her findings indicated that DVC supports a rich environment provided by text, audio and video modes (2006). She also concluded that facial expressions and gestures visible through the video facilitated task completion, and that learners' proficiency levels were observed to improve (2007). More recently, studies have begun to recognize the need for DVC tools developed specifically for the requirements of language learners and teachers (Guichon, 2010) and the skills and strategies that should be acquired in order to realize the full potential of DVC tools in online language tutorials (Hampel & Stickler, 2012).

Other studies have seen videoconferencing as a motivating factor in foreign language learning, but not without challenges, such as "cognitive load" and "lack of eye-contact" (McAndrew, Foubister & Mayes, 1996). Wang (2006) similarly warned learners and tutors that although DVC is a very rich environment, multimodality "can put pressure and strain on the users". Moreover, DVC can cause apprehension for some learners and reduced social cues might be preferred in these contexts. De los Arcos and Arnedillo Sánchez (2006) considered audio-graphic conferencing without video interaction as a comfortable environment especially for shy learners; as one of their participants suggested: "[h]ow can you make a fool of yourself if no one can see you?" (*op. cit.*: 91).

As communication technologies offer faster and more reliable interaction among people who are physically apart, so does their capacity to imitate face-to-face communication. Yet the interaction is still *mediated* via the technology. Certain features of face-to-face communication, which enhance feelings of being physically together with others, such as eye-contact (oculesics), touch (haptics), and physical distance (proximity), are still not available via DVC.

1.2 Social presence

Short, Williams and Christie (1976) first used the term social presence (SP) to describe the capacity of the media to transmit information. For instance, the telephone was considered as a lean medium that supports less social presence as it cannot transmit nonverbal cues. Rourke, Anderson, Garrison and Archer (1999) later defined it within a "community of inquiry" as the "participants' ability to present themselves socially and emotionally into the community". Social presence has a major role in creating an environment where learners are "willing to put themselves at risk" (Kehrwald, 2008) and feel free to make mistakes through expressing emotions and experiences.

All human communication is mediated via such tools as language, people, technology or cultural and institutional assumptions (Vygotsky, 1978; Leontiev, 1981). In computer-mediated contexts, tasks, participants and physical settings are also tools that mediate interaction (Lamy & Flewitt, 2011). Norris (2004) argued that there are two aspects of mediated interaction: expression and perception.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Interpersonal interaction is a key feature of contemporary online learning (Kehrwald, 2008) and mediation is an important determiner of learners' experiences. Communication technologies mediate learner interactions and thus can introduce some social and psychological distance between interlocutors. Social presence is pivotal in maintaining mediated interaction between language learners because social presence enhances connectedness and decreases psychological distance.

1.3 Gaze and social presence

Full gaze awareness is considered as "knowing where one is looking", partial gaze awareness means "knowing the direction of another's gaze" and mutual gaze is "usually referred to as eye-contact" (Gale & Monk, 2000). In face-to-face settings, gaze can be "a multimodal resource in organising the activity", i.e., negotiating meaning and turn-taking (Mondada, 2006).

Gaze, and more specifically eye-contact, is an element of social presence because it is one of the nonverbal cues for two main social psychological concepts of SP grounded in face-to-face interaction: intimacy and immediacy. Intimacy is related to the amount of eye-contact, physical proximity, topic of conversation and smiling (Argyle & Dean, 1965) and immediacy is defined as "communicative behaviors which enhance closeness to another" (Mehrabian, 1969: 203). Immediacy cues (e.g., eye-contact, physical proximity, smiling) are similar to intimacy cues. While intimacy research is more concerned with physical distance, immediacy is understood as the psychological distance between two people.

In educational research, verbal (e.g., humour, inclusive pronouns, encouraging participation and providing feedback) and nonverbal (e.g., gestures, facial expressions, touching, smiling, eye-contact, meaningful posture and intonation) teacher immediacy behaviours are believed to reduce physical and/or psychological distance between the teacher and the learner, and positively influence learner participation and attitudes (Bozkaya, 2008). Establishing eye-contact can improve the way teachers and learners interact face-to-face, whereas the lack of eye-contact, especially online, can be a big challenge for learners to maintain interaction and establish immediacy.

Until recently, most research has focused on the development of SP within text-based CMC (e.g., Lomicka & Lord, 2007) and few studies have explored SP in multimodal contexts (Satar, 2010; Ko, 2012). Focusing on the impact of eye-contact, Bondareva and Bouwhuis (2004) compared two traditional videoconferencing systems (VC), one with a slight adaptation of the set-up by installing a mirror to allow direct eye-contact. The results indicated significant differences in SP in favour of the eye-contact VC set-up group. Participants' comments also indicated that participants who used the traditional VC felt discomfort due to the discrepancy between looking at the camera to establish eye-contact and looking at the screen to see the reactions of their interlocutors. Participants who used the direct eye-contact VC set-up reported that "the communication was very natural and similar to the real life. … they did not feel as if they were in different rooms" (*op. cit.*: 8). A second study involved three experimental conditions with the addition of a face-to-face set-up (Bondareva, Meesters & Bouwhuis, 2006) and initial analysis indicated that there was no significant difference between the direct eye-contact VC set-up and face-to-face condition.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

The findings of these studies are significant in manifesting the importance of eye-contact for SP in VC settings and the discomfort perceived due to lack of it.

1.4 Gaze in DVC

Gaze is an important resource in face-to-face interactions, but it is difficult to establish in desktop videoconferencing (DVC) due to different positions of the webcam and the position of eyes on the computer screen (Grayson & Monk, 2003). Grayson and Monk (2003) recommended that for optimal mutual gaze, "the video camera should be placed as close to the image of the remote participant as possible" with a head and shoulders view (*op. cit.*: 241). They suggested that within this setup, people could learn to *interpret* gaze.

In language learning situations via DVC, two recent articles explored gaze, which I review in more detail here. Lamy and Flewitt (2011) analysed DVC interactions via MSN messenger in tandem learning. They described the online interaction of one pair based on Scollon and Scollon's (2003) idea of geosemiotics. They illustrated four different types of gaze: looking at the interlocutor, own image, camera, and chat window. Interviews with the participants indicated uneasiness when looking straight into the webcam and the impossibility of eye-contact when interlocutors looked at the webcam simultaneously.

Develotte, Guichon and Vincent (2010) investigated how language teachers learn to teach via DVC in Skype. Eleven French graduate teacher trainees learning to teach French as a foreign language (FFL) online were paired up with sixteen French learners from an American university for eight sessions. The gaze of five teacher trainees was analysed and semi-structured interviews were conducted. The findings indicated three different types of behaviour in the teacher trainees: preferring the audio mode and only looking at the learner's image a quarter of the time; mixed use dealing with multiple tools simultaneously; and exclusive use of the webcam making sure the learner's image was visible at all times. Teacher trainees were observed to use the webcam mainly to increase empathy supported by their facial expressions and less to communicate information. The socio-affective indicators, such as laughs and smiles, were found to "help construct an interpersonal relationship between the teacher trainees and their students" (*op. cit.*: 14). In addition, five degrees of utilizing the webcam were identified, 'zero' indicating no use of the webcam and 'four' indicating a direct look at the camera. Although the fourth degree was recommended to increase co-presence, the semio-pedagogical skills needed to determine when to use what degree of intensity were foregrounded in the conclusions.

Given this background, this paper aims to provide a deeper understanding of the impact of a webcam that mediates gaze in DVC. More specifically, it explores participants' strategies for using the webcam in a language learning setting, how the participants establish eye-contact in DVC, how they compensate for the lack of direct eye-contact, and how they feel about mediated gaze.

2. Research methods

This study was conducted as part of a PhD study at the Open University, UK that explored social presence in online multimodal communication and

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

proposed a framework to analyse learner interactions (Satar, 2010). Within this framework, gaze was one aspect of multimodality, which was observed as an emerging concept of social presence in addition to its existing components (affective: building immediacy; interactive: sustaining interaction; and cohesive: establishing intersubjectivity). The study belonged to a constructionist/interpretivist paradigm following a qualitative research tradition and computer mediated discourse analysis approach (Herring, 2001). The research design was based on an exploratory and instrumental case study method (Creswell, 2007; Richards, 2003; Yin, 2003).

2.1 Participants

Ten first-year teacher trainees, aged 19–22, studying English Language Teaching at three different universities in Turkey participated in the study. They shared the same native language and culture (Turkish) and used English as a foreign language for their DVC interactions. There were three reasons why this group of participants was approached: familiarity, cultural similarity and level of language. The perception of social presence depends very much on the "medium, knowledge of the other, content of the communication, environment, and social context" (Biocca, Harms & Burgoon, 2003: 469). Therefore, it was crucial that pairs in each case were not previously acquainted, shared the same culture and used the same tasks. The participants had sufficient level of language skills to complete the tasks. Sharing the same culture and language with the participants also gave the researcher an insider's view into the cultural assumptions and native language interference in the interactions between the participants.

The participants were approached via their lecturers at the university. The lecturers provided information about the project to their students and volunteer students then got in touch individually. Selection and pairing of dyads in each case was opportunistic and random. The physical conditions and hardware set-up as well as the Internet connection speed and quality of the participants varied. Table 1 provides a summary of the cases, their hardware set-up and the number of DVC sessions for each case.

2.2 Data collection

Several methods of data collection were used including DVC recordings to observe salient features of the interactions, questionnaires, interviews and stimulated reflection. DVC sessions were in English, with some instances of code switching. Interviews were conducted in Turkish and questionnaires were bilingual. A total of eighteen DVC sessions were conducted which resulted in about fourteen hours of recording. Participants' interactions were guided by open-ended tasks which could potentially stimulate interpersonal interaction and encourage optimum projection of social presence, such as talking about personalities or hometowns, or describing and drawing dream rooms.

For DVC interactions, ooVoo (http://www.oovoo.com) was selected as the online platform because at the time of data collection (2008–2009 academic year) it was

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Table 1 *Participants/Cases*

Cases	Participants (Gender)	Hardware set-up	No of DVC sessions/total time
Case 1	Deniz (M)	External webcam + desktop PC (internet café)	3 sessions/158 mins
	Zeynep (F)	External webcam + desktop PC (internet café)	
Case 2	Filiz (F)	Inbuilt webcam + laptop (home, own room)	3 sessions/132 mins
	Nil (F)	External webcam + desktop PC (home)	
Case 3	Defne (F)	External webcam + laptop (dormitory room)	4 sessions/144 mins
	Hale (F)	External webcam + laptop (shared flat)	
Case 4	Emre (M)	External webcam, fixed on desk + desktop PC (internet café)	4 sessions/165 mins
	Osman (M)	Inbuilt webcam + laptop (shared flat, own room)	
Case 5	Eda (F)	Inbuilt webcam + laptop (shared flat, own room)	4 sessions/230 mins
	Ali (M)	External webcam + desktop PC (internet café at dormitory)	

the only tool that was free, allowed seamless recording of the sessions and had good two-way audio and video quality.

2.3 Ethical considerations

Protection of privacy and ensuring confidentiality and anonymity are essential (Hammersley & Atkinson, 1995) in any research project. Yet due to the nature of the data, i.e., audio and video recordings, ethical considerations were especially significant in this study. To ensure privacy and confidentiality, the names of the participants as well as the real names of institutions and any personal information were anonymized. Informed consent was ensured through an information letter and a consent form. These forms included details on the aims of the project, what participation involved, any foreseen harms and benefits, how anonymity and confidentiality would be ensured and what would happen in case of withdrawal. The consent form asked for participants' agreement on the use of data "for educational or research purposes, including publication".

2.4 Data analysis

Data was transcribed and analysed using the qualitative analysis software Atlas-ti 6.0 (www.atlasti.com). Prior to the systematic analysis of the data, a general reading and annotation of the data provided an insight into the significant features of

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

each case. Salient extracts from the DVC recordings were also discussed and reviewed with experts in the field. Case study (Stake, 1995; Creswell, 2007) and grounded theory principles (Strauss & Corbin, 1998) for thematic analysis guided the analysis of questionnaires and interviews. Based on insights from the questionnaires and interviews, a multimodal analysis of DVC recordings was conducted, drawing on the principles of social semiotics (Kress & van Leeuwen, 2001) and interactional sociolinguistics (Gumperz, 1982, 2003).

3. Results and discussion

In this section, I first provide very brief introductions to each case and then describe participants' different strategies of using gaze and their perspectives on eye-contact. Although a case study method requires in-depth representations of the cases, the limitations of this paper allow me to report only their distinctive features. Yet I believe the following introductions will provide some useful background information for a better understanding of the following data analyses.

3.1 Case introductions

The participants in the Case 1, Deniz (M) and Zeynep (F), completed their DVC sessions in Internet cafés and they both used external webcams and desktop computers. Zeynep was mostly uncomfortable with the people around her and did not like the fact that she was attracting others' attention as she was speaking in a foreign language, i.e., English. Deniz was more relaxed but nevertheless, both displayed nonverbal signs of embarrassment such as face concealment. Their interaction was marked by jokes, banter and lots of giggles, some of which could be interpreted as flirtatious.

Both participants in Case 2, Filiz (F) and Nil (F), were at home during the sessions. Nil's computer skills were low and her gaze was distinctive; fixed towards the webcam, which I explain in more detail below. Nil was wearing a headscarf and did not smile much. Filiz perceived these as signs of a serious personality. Filiz was using a laptop and she looked more willing to continue the interaction than her partner by initiating new topics and asking questions.

Defne (F) and Hale (F), Case 3, had laptops with separate webcams and they both looked relaxed, using gestures freely. The pace of their interaction was moderate, allowing each other time to construct their sentences. They asked each other questions to continue the interaction and there were few overlaps. The participants reported that they felt psychologically close to each other and had a sense of familiarity and trust.

The participants in Case 4 were Emre (M) and Osman (M). Emre was at an Internet café, using a desktop and an external webcam which was fixed on the desk. This meant that his posture was restricted, as he could not move the webcam to suit the way he would want to project his image. In contrast, Osman was in his room, using a laptop with an inbuilt webcam. His posture was much more relaxed, leaning back in his chair. There were frequent overlaps in their interaction and some power struggle to take the floor, which resulted in Emre's submissive behaviour of leaving the floor mostly to Osman. Both reported instances of boredom and limited development of friendship.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

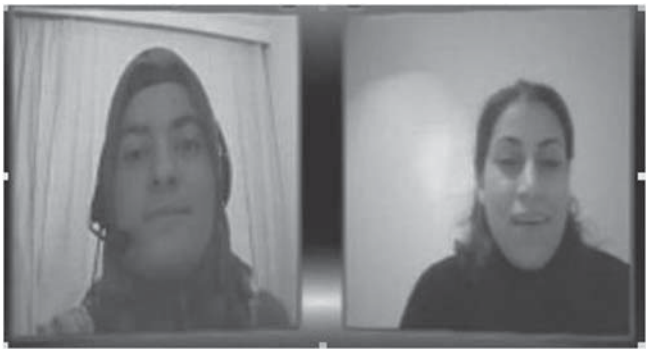
In Case 5, Eda (F) was at home and used a laptop with an inbuilt webcam, while Ali (M) was at the Internet café of his dormitory and used a desktop computer with separate webcam. Although they also had different physical settings, in comparison to case 4, the participants felt they supported each other and had an equal level of dominance in the interaction. Frequent silences were observed which were filled mostly with smiles, laughter and self-adaptors (Knapp, 1980; Richmond, McCroskey & Payne, 1991) that signal psychological discomfort and concern for self-representation. Eda's discomfort was perhaps linked to her uneasiness with video interaction, a feeling which she described as being watched by a stranger secretly. She reported that, especially when she was asked to look or when her interlocutor looked straight into the camera, she had a feeling similar to one she would have if she made eye-contact with a stranger while dancing.

3.2 Gaze in DVC

The multimodal analysis of DVC recordings revealed five gaze types used by the participants: fixed gaze, free gaze, strategic gaze, averted gaze and directed gaze.

3.2.1. Fixed gaze. Nil (Case 2) was the only participant who constantly looked at the camera, perhaps with an attempt to sustain eye-contact. Figure 1 shows the usual posture of the participants in Case 2. Nil (left) is looking straight into the webcam, while Filiz (right) used her gaze freely. Throughout the DVC interactions, Nil's gaze constantly stayed on the camera.

As the researcher, I had initially thought that she had positioned her webcam in a way that would allow her to alternate her gaze with little effort between the webcam and the screen where Filiz's image was displayed (for example, on top of or close to her screen). However, it only became clear that Nil was intentionally looking into the webcam when the participants showed each other pictures and Nil had to look at the screen to see the pictures (Extract 1).[1]



Nil (on the left) is looking at the camera, listening. Filiz (on the right) is talking.

Fig. 1. Fixed gaze

[1] In this paper, Jefferson's (1984) transcription notations are used.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Extract 1 Can you see?

	Verbal	Nonverbal
1	N: can you see?	Nil shows a photo of her sister; her gaze to her right (on screen checking how well she shows the picture). Filiz moves closer to screen
	 (line 1)	
…	(10 lines, Nil and Filiz talk about the picture)	
12	(1.0)	Nil removes picture; laughs
13	F: also my sister is here (you) see her err	Nil's gaze: to her right (screen); Filiz turns her head right taking pictures
14	she is my sister (.)	Nil's gaze: camera; Filiz puts picture close to camera
15	this one	Nil's gaze: to her right (screen); Filiz points to the photo
	 (line 15)	
16	N: yes, I saw it	Nil nods; Filiz looks at photo and points again with the other hand

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press



Fig. 2. Free gaze

When compared to her usual posture and direction of gaze in Figure 1, it is apparent in Extract 1, line 1 and line 15 that Nil had to turn her head to her right, almost to a 90 degree angle, to check how well the photograph she was showing was displayed and also to look at the photograph Filiz was showing her. The position of Nil's webcam and screen required additional effort to alternate gaze. Nil's fixed gaze and intentional use of the webcam was distinctive and unusual in the sense that she ignored her partner's image on the screen and preferred to look into the webcam, which in a way implied that she paid more attention to how she projected her presence via eye-contact, even if it meant losing important visual cues to interpret her partner's messages.

3.2.2. Free gaze. Some participants (Hale, Defne, Osman, Emre and Filiz) were more flexible in terms of their gaze in DVC sessions. They used gaze freely, finding it more natural to look at the screen to see their partners' reactions as opposed to looking into the webcam (Figure 2a). They did not hesitate to look around while speaking and to reinforce their gestures with their head movements and varied direction of gaze as they would when speaking to others face-to-face (Figure 2b).

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press



Fig. 3. Strategic gaze

3.2.3. Strategic gaze. In their interviews, Deniz (Case 1) and Ali (Case 5) stated that they intentionally looked at the webcam when they wanted to send a specific message. Although Zeynep (Case 1) usually avoided looking into the webcam, there were a few instances when she briefly looked at the webcam, too. For instance, screenshots in Figure 3 are taken from Case 1, session 1. Figure 3a and Figure 3b illustrate when Zeynep teasingly asked Deniz if he was crying. Deniz said no and showed his eyes to Zeynep, first getting closer to the screen instinctively and then strategically closer to the webcam realising that she could only see his eyes clearly if he looked into the webcam. Similarly, at some point during the session, Deniz adjusted his camera position, changing the way his image was projected. The new position of the webcam transferred a closer picture of his face. He then asked Zeynep how he looked from this new camera angle (Figure 3c). Zeynep said he looked handsome and Deniz, pleased with her response, looked at the camera and said, ''okay, I'll keep it in this position'' (Figure 3d). Lastly, Figure 3e shows Zeynep's usual direction of gaze and Figure 3f one of the few times when Zeynep directed her gaze towards the webcam. This was when she was acting out a character in a Turkish TV sit-com which was popular at the time. She used the character's catch phrase in Turkish gesturing towards the camera and looking towards the webcam. The gesture, gaze and the quote from the sit-com were used for a humorous purpose.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press



Fig. 4. Averted gaze



Osman is showing his height.

Fig. 5. Directed gaze

3.2.4. Averted gaze. Eda and Zeynep were the two participants who stated their uneasiness with trying to establish eye-contact via the webcam. Eda did not manifest any attempt to attain direct mutual gaze via the webcam. In Figure 4a, although Eda is talking, her gaze is away from the webcam, mostly looking downwards or around. She rarely attempted indirect eye-contact by looking towards the screen, as I will illustrate in Extract 2 (section 3.3.3).

Zeynep usually looked downwards when talking (Figure 4b). She rarely looked towards the screen (lifted her head up) to send a message (Figure 3f). During the DVC sessions, Zeynep was in a busy Internet café. She usually spoke with a quiet voice and was not comfortable. Thus, in addition to her expressed uneasiness with DVC communication, by avoiding direct gaze via the webcam and keeping her gaze downwards, she might be trying to block out her awareness of other people around her and the resulting discomfort. This self-consciousness has pedagogical implications in drawing attention to the physical context of the learners. While tutors and language learning partners might overlook the inaccessible visual field for themselves, the invisible background of the partner would certainly have an influence on the partner's interactions and thus on the conversation.

3.2.5. Directed gaze. Participants sometimes directed their partners' gaze to certain items (Figure 5). For example, Zeynep and Osman stood up to show their height;

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Table 2 *Types of Gaze in DVC*

Type of Gaze	Definition
Fixed Gaze	The speaker tries to establish direct eye-contact by looking at the webcam at all times.
Free Gaze	The speaker directs his/her gaze freely and naturally without paying any particular attention to the webcam.
Strategic Gaze	The speaker tries to establish direct eye-contact via the webcam when needed for meaning negotiation.
Averted Gaze	The speaker predominantly avoids direct attempts at eye-contact via the webcam.
Directed Gaze	The speaker directs the gaze of the listener to a particular item by controlling the visual field transmitted via the webcam.

Zeynep and Deniz showed each other their accessories (e.g., watch, fingernails, ring); Filiz and Nil showed each other photographs and pointed out people in these photographs; and others showed each other the room pictures they had drawn, guided by the task, and pointed out the objects in these drawings.

In this section, five different types of gaze that were observed in the data have been reported and exemplified. Table 2 summarizes the types of gaze with brief definitions of each type.

In this study, the frequency of each type of gaze was not calculated. There were two reasons behind this. First, based on the exploratory nature of this study within an interpretivist paradigm using qualitative methods, interest was focused on exploring different ways of utilising the webcam and understanding how and why these ways were employed by the speakers and interpreted by the listeners, rather than how much each type was used. Second, quantifying the use of different types of gaze in this study would not produce valid and reliable statistical information, because the quality of video transmission and the clarity of gaze were not constant across participants at all times. For instance, Ali's video image was frequently distorted. Moreover, one of the participants in Case 4, Osman, was wearing glasses in some sessions which made it much more difficult to accurately observe his gaze throughout. Furthermore, each participant represented a particular way of using one or two gaze types. Research exploring the frequency of gaze types would be most valid perhaps within more controlled settings employing tools such as eye-tracking.

3.3 Participant views on gaze via DVC

The thematic analysis of interviews and questionnaires revealed that participants thought mutual gaze was unnatural and perhaps impossible, which decreased trust online. Some felt that a direct attempt at eye-contact via the webcam was intimidating; this could be overcome by using an indirect method and simply looking at the screen. Participants also believed that DVC required manipulation of one's own image and gaze for better projection of themselves. Each of these perceptions is explained in detail below.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

3.3.1. Mutual gaze is unnatural and perhaps impossible. Although some participants (for example, Defne) stated that people might look at the camera to maintain eye-contact via DVC directly or indirectly, the general consensus was that mutual gaze was not really possible online (Hale, Osman, Emre, Eda, Ali, Defne, Deniz). Deniz said he could only see Zeynep's face partially from one side where the camera was positioned and could not look into her eyes. Ali expressed the inability to attain eye-contact due to the impossibility of looking at the camera and his partner's image on screen simultaneously. Likewise, Emre talked about the difficulty of alternating gaze between the webcam and the screen.

Nil, Defne, Osman and Emre expressed their preference for cameras built into the screen as they felt more relaxed, with flexibility in their posture, and felt nearly the same as if they actually had eye-contact. Thus, communication using an inbuilt camera or a camera attached to the top of the screen where little effort was needed to alternate gaze was perceived as more natural. Hale, Eda and Osman also suggested staying close to the camera in order to obtain a clearer image of the facial expressions.

Some participants (Hale, Defne, Deniz, Ali) mentioned occasionally looking at the camera. Deniz stated that he looked at the camera when he "wanted to give a direct message", including humorous exchanges. Hale and Defne could not specifically identify when they looked at the camera, but said they mostly preferred watching their interlocutors. Ali also preferred looking at the screen, and thus his partner, "90 percent of the time".

Eda reported in the final questionnaire that she "always looked at the screen" and never the webcam (Figure 4a). Similarly, looking at the screen was quite natural for Hale. In her interview, she said "I generally looked at the screen, not at the camera. The video call is on the screen, I can already see myself there. Defne is also there. The tasks are also there."

Participants reported that they looked at the screen mainly to see their partner's visual reactions and backchannels and understand her/him better. Specifically, video image helped Zeynep and Defne determine audio delay or problems and avoid interruptions. For example, in her interview, Zeynep said: "I looked to see if he was laughing when I made a joke, or if he understood but did not give any reaction. Also my audio was delayed, I looked to see if he heard me, listened to me. If he was still talking, then I thought he is talking and that's why he didn't hear me or my voice was delayed."

3.3.2. Direct attempt at eye-contact can be intimidating. Participants had varying views about how they felt in the multimodal online communication. For example, Eda and Defne reported uneasiness with using the camera. Eda's views were quite extreme (Figure 4a): in her final questionnaire she wrote: "I can't look at the camera and when somebody directly looks at the camera while speaking with me, I feel very anxious, sometimes nervous as if s/he is watching me secretly, without my permission. Even if I avert my eyes from the screen, I get mad, irritated". Defne expressed a moderate view: "Not everyone is used to the camera, especially if you haven't used it before you feel under scrutiny under the camera, like Big Brother, you don't know what to do, where to keep your hands or body".

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

This could imply that some people might perceive an attempt to attain direct gaze and eye-contact in DVC as unnatural or even as staring. Hence, pedagogically, intentional use of the webcam is highly relevant to issues of confidence and anxiety. For the online teacher and learner, an awareness of various ways in which intended eye-contact might be perceived is crucial to accommodate a variety of interlocutors, especially inexperienced users of DVC. It is also important to make an effort to look out for linguistic and paralinguistic cues of embarrassment or uneasiness whenever mutual gaze through looking directly into the webcam is pursued.

3.3.3. Gaze on screen might mean an attempt at eye-contact. Filiz and Defne felt that their partner was listening to and was interested in them when they thought that their partner looked at the screen and thus 'themselves'. Extract 2 is an example of how looking at the screen can replace looking at the webcam to make eye-contact. In the last session of Ali and Eda (Case 5), Ali was describing his dream room and Eda was trying to draw it on paper.

Extract 2 a lot of

Lines	Verbal	Eda's gaze
1	Ali: A lot of window	 looking down on paper, drawing
2	Eda: A lot of? (smiles)	 looking up at the screen

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

3	Ali: Yes	 looking down on paper, drawing

When Ali wants 'a lot of' windows in his room (line 1), Eda is surprised and requests confirmation of the information (line 2). Eda shifts her gaze towards the screen, but not towards the webcam. However, due to the nature of the message (repeating Ali's expression with emphasis and a smile) and the fact that she abandons her drawing task, it could be perceived as an attempt at eye-contact. Once she receives the confirmation, she resumes her task of drawing (line 3). Thus, in DVC, looking at the screen towards the other's video image could replace direct (and perhaps unnatural) eye-contact established via staring at the webcam, as long as both the interlocutors know how to read the attempted eye-contact.

3.3.4. Lack of mutual gaze can lead to a decreased sense of trust online. Deniz thought that in online interactions, compared to face-to-face, the "authenticity of emotions" was missing. He thought it was easier to lie online because "you did not really know and see (meet) the person". However, in face-to- face settings, he felt quite confident in predicting what the person really wanted to do and what s/he thought about him "by looking into their eyes". Moreover, Hale also underlined the fact that eye-contact was one of the main indicators of "honesty" in face-to-face interactions and lack of it in DVC made it difficult to trust people.

3.3.5. DVC requires manipulation of one's own image and gaze. Participants with inbuilt cameras usually had a wide frontal view of their portrait and did not complain about the lack of control over their image (Hale, Filiz, Eda). While some participants using separate cameras made no mention of it (e.g., Zeynep), others complained about the inability to move the camera as they wished (Nil & Emre). The individual variation seemed to depend both on hardware specifications and on the way learners would strategically use the webcam to project their image. Like a director adjusting the visual frame, participants tended to select camera angles to transmit their image or another object or person to their partners. For instance, Ali (Case 5) frequently moved the camera position to suit his own posture whenever he moved (reclining in his chair or sitting upright).

In Figure 6, although Ali's gaze is not clear, his awareness of the camera and how much is projected is evident. These screenshots were taken from the fourth session of

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

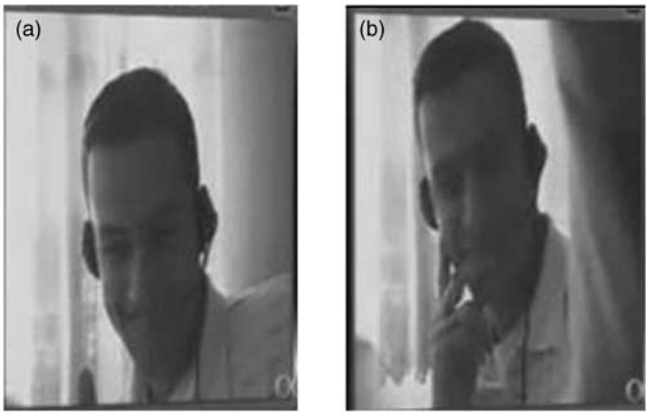


Fig. 6. Manipulating own image

Eda and Ali. When Eda says she likes playing football with boys on the street, Ali makes a 'thumbs-up' gesture to provide visual feedback, which appears at the bottom of his image (a). He notices that his gesture is not visible and rearranges his camera position, thus the projection of his image, so that his gestures, in addition to his facial expressions, would be visible (b).

Although it could be distracting in a face-to-face setting, the availability of one's own image was welcomed in DVC. Whilst this could be due to prior experience of using similar tools, it could also be a wish to control the projection of the self via the limited and defined visual field of the webcam. However, it is important to note here that although some learners might expect to see their own image in DVC and would be comfortable with seeing it (such as Hale, Deniz & Ali), others might avoid their own image if they lacked self-confidence or found it distracting and/or anxiety-provoking to see themselves.

4. Conclusions

The increasing availability and ease of access to DVC tools has resulted in an increase in their use in online and/or blended language courses as well as in independent language learner telecollaboration projects. Therefore, there is an increasing need to better understand the effects of multimodal elements, such as gaze, on learners' cognitive, affective and social interaction skills.

Further to the degrees of using the webcam reported by Develotte *et al.* (2010), the findings of the present study indicated that the participants had five different ways of utilising the webcam; manipulating gaze constantly, strategically, avoiding gaze totally, directing gaze and free gaze. Similar to the findings of previous research (McAndrew *et al.*, 1996; Bondareva & Bouwhuis, 2004; Lamy & Flewitt, 2011) eye-contact was mostly believed to be unattainable in DVC, at least with the technology that was available to the participants at the time. Participants preferred looking at the screen, finding it more natural, as they would miss their interlocutors' visual feedback if they only looked at the camera. Intentional use of the camera was utilized for varying purposes by the learners and interpreted with varying feelings.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Learners either pursued fixed or strategic eye-contact by directly looking into the webcam or tended to avoid eye-contact, perceiving direct gaze into the webcam as stare. Alternatively, looking at the screen was a compensation for direct eye-contact. Certain hardware features were influential in how participants could manipulate gaze. They suggested staying close to the camera, using inbuilt devices, and if that was not possible, attaching the external webcam on top of the screen to allow more flexibility. These recommendations echoed those of Grayson and Monk (2003). Thus, it is possible to suggest that an inbuilt camera which accommodates minimal change in gaze between the camera and the screen might provide the optimum set-up for less intrusive gaze.

Multimodal challenges influenced the way participants maintained interaction and established immediacy. The DVC environment could enhance trust, yet, for some participants it was still insufficient to permit the generation of immediacy due to disembodied and limited representation, delays and distortions in audio and video and lack of eye-contact. As new DVC tools are developed, it would be important to investigate the effects of the new multimodal features on the development of social presence online. Moreover, using the currently available DVC tools, it is not very easy to explore gaze by multimodal analysis of the recordings. Future research using eye-tracking technologies could reveal a more accurate understanding of gaze in DVC.

The existence of one's own image on the screen in addition to that of the conversational partner underscored two important issues for software and hardware design. First, the ability to adjust the position of the webcam or the availability of multiple webcams and selection of a preferred angle could change interlocutors' practices in terms of their awareness and strategic use of the camera position in projecting their presence. Second, future research could investigate the influence of the availability, size and refresh rate of the interlocutors' video images in language learning contexts to allow for individual variation and suitability to various tasks. Better quality video image would enable more detail to be transmitted and thus a clearer image would probably assist fluent interaction, especially when negotiating affective content where paralinguistic cues are important. Inequality in video size and frame rate of each participant would perhaps result in an inequality in terms of attention paid to each image. When selecting the tool for DVC interactions, all these features, and their relative impacts, should be taken into consideration in line with the specific requirements of the learners and the language learning tasks.

Finally, skills in interpreting mediated eye-contact are necessary for the online tutor and learners. The mismatch between *intended* and *perceived* attempted eye-contact via direct gaze into the webcam, that is, whether it is supportive eye-contact or a challenging stare, is an obstacle built into current DVC technology. Learners and teachers should keep in mind that although being a multimodal and rich context, DVC is a technology that mediates interaction and it is different from face-to-face communication. The most prominent differences are delays and distortions in audio and video, limited visual field and mediated eye-contact. Teachers and learners should learn to accommodate and manipulate these factors to *project* and *interpret* social presence online.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Acknowledgements

I would like to extend my special thanks to the students in this study for their participation and to the anonymous reviewers for their constructive comments.

References

Argyle, M. and Dean, J. (1965) Eye-contact, distance and affiliation. *Sociometry*, **28**(3): 289–304.

Biocca, F., Harms, C. and Burgoon, J. K. (2003) Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoperators and Virtual Environment*, **12**(5): 456–480.

Bondareva, Y. and Bouwhuis, D. (2004). Determinants of social presence in video-conferencing. In: *Environments for personalized information access: Proceedings of the AVI 2004 Workshop*. Presented at the Working conference on advanced visual interfaces, AVI 2004, Gallipoli, Italy, 1–9.

Bondareva, Y., Meesters, L. and Bouwhuis, D. (2006) Eye contact as a determinant of social presence in video communication. In: *Proc. 20th International Symposium on Human Factors in Telecommunication*. http://www.hft.org/HFT06/paper06/37_Bondareva.pdf

Bozkaya, M. (2008) The relationship between teacher immediacy behaviours and distant learners' social presence perceptions in videoconferencing applications. *The Turkish Online Journal of Distance Education*, **9**(1): 180–192.

Coleman, J. A., Hampel, R., Hauck, M. and Stickler, U. (2010) Collaboration and interaction: The keys to distance and computer-supported language learning. In: Levine G. and Phipps A. (eds.), *Critical and Intercultural Theory and Language Pedagogy*. Florence, KY, USA: Cengage Heinle, 161–180.

Creswell, J. W. (2007) *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*, 2nd ed. Thousand Oaks, London, New Delhi: Sage Publications.

de los Arcos, B. and Arnedillo Sánchez, I. (2006) Ears before Eyes: Expanding tutors' interaction skills beyond physical presence in audio-graphic collaborative virtual learning environments. *User-Centered Computer Aided Language Learning*. *Hershey: Information Science Publishing*, 74–93.

Develotte, C., Guichon, N. and Vincent, C. (2010) The use of the webcam for teaching a foreign language in a desktop videoconferencing environment. *ReCALL*, **22**(3): 293–312.

Gale, C. and Monk, A. F. (2000) Where am I looking? The accuracy of video-mediated gaze awareness. *Perception & Psychophysics*, **62**(3): 586–595.

Grayson, D. M. and Monk, A. F. (2003) Are you looking at me? Eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction*, **10**(3): 221–243.

Guichon, N. (2010) Preparatory study for the design of a desktop videoconferencing platform for synchronous language teaching. *Computer Assisted Language Learning*, **23**(2): 169.

Gumperz, J. J. (1982) *Discourse strategies*. Cambridge: Cambridge University Press.

Gumperz, J. J. (2003) Interactional sociolinguistics: a personal perspective. In: Schiffrin D., Tannen D. and Hamilton H. E. (eds.), *The handbook of discourse analysis*. Oxford: Blackwell, 215–228.

Hammersley, M. and Atkinson, P. (1995) *Ethnography: principles in practice*, 2nd ed. Oxon: Routledge.

Hampel, R. and Hauck, M. (2004) Towards an Effective Use of Audio Conferencing in Distance Language Courses. University of Hawaii, National Foreign Language Resource Center. http://llt.msu.edu/vol8num1/hampel/default.html

Hampel, R. and Stickler, U. (2012) The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL*, **24**(2): 116–137.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Herring, S. C. (2001) Computer-mediated discourse. In: Schiffrin D., Tannen D. and Hamilton H. E. (eds.), *The handbook of discourse analysis*. Oxford: Wiley-Blackwell, 612–634.

Jefferson, G. (1984) Transcription notation. In: Atkinson J. M. and Heritage J. (eds.), *Structures of social interaction*. New York: Cambridge University Press, ix–xvii.

Kehrwald, B. (2008) Understanding social presence in text-based online learning environments. *Distance Education*, **29**(1): 89.

Knapp, M. L. (1980) *Essentials of nonverbal communication*. New York: Holt, Rinehart and Winston.

Ko, C. J. (2012) A case study of language learners' social presence in synchronous CMC. *ReCALL*, **24**(1): 66–84.

Kress, G. R. and van Leeuwen, T. (2001) *Multimodal discourse: the modes and media of contemporary communication*. London: Arnold.

Lamy, M.-N. and Flewitt, R. (2011) Describing online conversations : insights from a multimodal approach. In: Develotte C., Kern R. and Lamy M.-N. (eds.), *Décrire la Conversation en Ligne : Le Face à Face Distanciel*. Lyon, France: ENS Éditions, 71–94.

Leontiev, A. A. (1981) *Psychology and the language learning process* (C. V. James, ed.). Oxford: Pergamon Press.

Lomicka, L. and Lord, G. (2007) Social presence in virtual communities of foreign language (FL) teachers. *System*, **35**(2): 208–228.

McAndrew, P., Foubister, S. P. and Mayes, T. (1996) Videoconferencing in a language learning application. *Interacting with Computers*, **8**(2): 207–217.

Mehrabian, A. (1969) Some referents and measures of nonverbal behavior. *Behavior Research Methods & Instrumentation*, **1**: 203–217.

Mondada, L. (2006) Participants' online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. *Discourse Studies*, **8**(1): 117–129.

Norris, S. (2004) *Analyzing multimodal interaction: a methodological framework*. New York: Routledge.

O'Dowd, R. (2000) Intercultural Learning Via Videoconferencing: A Pilot Exchange Project. *ReCALL*, **12**(1): 49–61.

Richards, K. (2003) *Qualitative inquiry in TESOL*. Palgrave Macmillan.

Richmond, V. P., McCroskey, J. C. and Payne, S. K. (1991) *Nonverbal behavior in interpersonal relations*, 2nd ed. New Jersey: Prentice Hall.

Rourke, L., Anderson, T., Garrison, D. R. and Archer, W. (1999) Assessing Social Presence In Asynchronous Text-based Computer Conferencing. *The Journal of Distance Education/ Revue de l'Éducation à Distance*, **14**(2): 50–71.

Satar, H. M. (2010). Social Presence in Online Multimodal Communication: A Framework to Analyse Online Interactions Between Language Learners. Unpublished PhD thesis. The Open University, Milton Keynes.

Scollon, R. and Scollon, S. (2003) *Discourses in Place: Language in the Material World*. London: Routledge.

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

Stake, R. E. (1995) *The art of case study research*. Thousand Oaks, London, New Delhi: SAGE.

Strauss, A. L. and Corbin, J. M. (1998) *Basics of qualitative research: techniques and procedures for developing grounded theory*, 2nd ed. London, Thousand Oaks, New Delhi: SAGE.

Vygotsky, L. S. (1978) *Mind in society: the development of higher psychological processes*. Harvard, USA: Harvard University Press.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press

Wang, Y. (2004a) Supporting Synchronous Distance Language Learning with Desktop Videoconferencing. *Language Learning & Technology*, **8**(3): 90–121.

Wang, Y. (2004b) Distance Language Learning: Interactivity and Fourth-Generation Internet-Based Videoconferencing. *CALICO Journal*, **21**(2): 373–395.

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.

Wang, Y. (2008) *Distance Language Learning and Desktop Videoconferencing*. Saarbrücken, Germany: VDM Verlag Dr. Müller.

Yin, R. K. (2003) *Case study research: design and methods*, 3rd ed. Thousand Oaks: Sage Publications.

https://doi.org/10.1017/S0958344012000286 Published online by Cambridge University Press