The Impact Of The Webcam On An Online L2 Interaction

Nicolas Guichon, Cathy Cohen

Guichon, N. & Cohen, C. (2014) The Impact Of The Webcam On An Online L2 Interaction. *Canadian Modern Language Journal*. vol. 70, n°3, pp. 331–354. (Authors' manuscript)

The Impact of the Webcam on an Online L2 Interaction

Nicolas Guichon and Cathy Cohen

Abstract: It is intuitively felt that visual cues should enhance online communication, and this experimental study aims to test this prediction by exploring the value provided by a webcam in an online L2 pedagogical teacher-to learner interaction. A total of 40 French undergraduate students with a B2 level in English were asked to describe in English four previously unseen photographs to a native English-speaking teacher of EFL via Skype, a free web-based videoconferencing tool, during a 10-minute interaction. Twenty students were assigned to the videoconferencing condition and 20 to the audioconferencing condition. All 40 interactions were recorded using dynamic screen capture software and were analyzed with ELAN, a sound and video annotation tool. Participants' perceptions of the online interaction are first compared with regard to the issues of social presence and their understanding and appreciation of the online interaction, using data gathered from a post-task questionnaire. The study then explores whether seeing the interlocutor's image impacts on the patterns of these synchronous exchanges and on the word search episodes. Results indicated that the impact of the webcam on the online pedagogical interaction was not as critical as had been predicted.

Keywords: audioconferencing, online interaction, social presence, videoconferencing, word search

Introduction

Since its inception, online language teaching has gradually incorporated functionalities that allow for more multimodal and synchronous means of communication. Yet, if synchronicity is generally seen as bringing real value to online pedagogical interactions (Hrastinski, 2008; Levy & Stockwell, 2006), research investigating the potential of a broader array of channels has been much less frequent. We contend that the crucial decision that distance learning institutions have to make concerning whether to include web-based audioconferencing or videoconferencing tasks in an online language learning program (see Hopkins, 2010) needs to be grounded in empirical evidence, especially since the provision of several channels in an online interaction is intuitively felt as beneficial to language learning.

Because it provides access to the interlocutor's image, two questions can be raised concerning the benefits of the webcam for the interaction: (a) Does the webcam image facilitate communication between distant partners by enhancing learners' perceptions of the online interaction? and (b) How does having access to one's partner's image change the patterns of the interaction?

In this experimental study, 40 French university students with a B2 level in English (according to the Common European Framework of Reference for Languages) were asked to describe four previously unseen photographs in English to a native English-speaking teacher via a videoconferencing facility during a 10-minute interaction. Twenty subjects did the activity with the webcam turned on (videoconferencing condition) and 20 with it switched off (audioconferencing condition). All 40 interactions were recorded using dynamic screen capture software and analyzed with ELAN, a sound and video annotation tool.

Participants' perceptions will first be investigated with regard to their feelings about the psychological and physical presence of the teacher and their understanding and appreciation of the online interaction, using data gathered from a post-task questionnaire. Here our aim is to compare the reactions of the two groups of participants to seeing or not seeing their teacher during the pedagogical interaction. We then compare the impact of each condition on the patterns of the synchronous interactions (silences, overlaps, number and duration of turns) and on the word search episodes, as specific samples of nonnative speaker (NNS) interactional practice (Brouwer, 2003).

The study of webcam-based online interaction

Comparative studies in the field of language learning

A small number of studies comparing audioconferencing and videoconferencing have been carried out in language learning situations to assess the potential of seeing one's partner. Yamada and Akahori (2009) carried out a well-designed experimental study whereby 40 university students in learner-to-learner dyads had to perform an explanation task in one of four conditions: (a) videoconferencing with both the learner's own and the partner's image; (b) videoconferencing with only the partner's image; (c) videoconferencing with only the learner's image; and (d) videoconferencing without images (i.e., audioconferencing). Two main conclusions can be drawn from this experiment: first, communication was facilitated when participants could see their partner's image but students' perceptions were more negative when they could not see their partners, because it allegedly augmented stress. Second, videoconferencing had a positive impact on participants' metacognition and comprehensibility in communication.

Yanguas (2010) conducted a qualitative study to examine how language learners negotiated meaning during task-based interaction. Learner-to-learner dyads were randomly assigned to (a) an audioconferencing group, (b) a videoconferencing group, or (c) a face-to-face (FTF) control group to complete a jigsaw task that included unknown lexical items. Yanguas (2010) concluded that in the audioconferencing condition, learners had to use linguistic resources they would not have used in the two conditions that provided visual cues (FTF and videoconferencing).

Rosell-Aguilar (2007) examined how 12 Open University tutors of a beginners' Spanish course compared online audiographic and FTF learning environments. His results suggest that, although tutors identified numerous similarities, they reported that the lack of paralinguistic cues in the audiographic condition made it difficult to perceive and interpret students' reactions, and that because there were no visual cues for the learner, the tutors had to provide warmth through the way they communicated and managed the learning environment. These comparative studies raise methodological and epistemological questions about the validity of comparing situations and drawing conclusions from experiments. As underlined by Yanguas (2010, p. 74), researchers "have taken very different perspectives and have used entirely different research designs making it very difficult to draw any general conclusions." Indeed, many variables need to be taken into account when carrying out comparative studies: the number of participants involved in the study, their digital literacy,

their level in L2, their attitudes toward language learning, whether they knew each other before the experiment, the type and length of the task they have to carry out, the type of data collected (self-reports, dynamic screen capture), the nature of the study (ecological and longitudinal, experimental and synchronic). Obviously, the methodological choices made by the researchers have an impact on the quality of the data and the validity of the results.
To the best of our knowledge, no study has yet compared an online L2 pedagogical interaction in audioconferencing and videoconferencing conditions when one of the partners is a language teacher. Thus the originality of the present research is its focus on an online teacher student interaction that aims to further our understanding of the value of two different synchronous conferencing tools for second language teaching (Dejean-Thircuir, Guichon, & Nicolaev, 2010; Develotte, Guichon, & Vincent, 2010).

Psychological factors in video-mediated L2 interaction

In this section, we will investigate what seeing one's interlocutor in an online situation brings to the quality of interpersonal communication at a psychological level and for language learning. Yamada and Goda (2012) have argued that significant communication in online settings requires creating a sense of presence facilitative of "smooth and meaningful interactions." Smoothness of mediated communication and psychological dimensions are usually united through the concept of "social presence," which is central in studies assessing the potential of the webcam in an interaction (Satar, 2013; Yamada & Akahori, 2009). In their seminal work, Short, Williams, and Christie (1976, p. 65) defined social presence as the "degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationship." Social presence then refers to the subjective activity attached to specific means of communication or their combination. Certainly this concept of social presence has been helpful when new means of communication have become available, and it has raised questions about the richness of the different media and their potential in distant exchanges. Almost contemporary to Short et al.'s (1976) paper, Korzenny's (1978, p. 3) theory of electronic propinquity explored "the psychological feeling of nearness" and offered "a general theory of mediated communication (…) given the role of electronic mechanisms in (…) interposing human contact." This theory "is conceptualized as a continuum of the subjective perception an individual holds that he or she is functionally, if not physically, close to someone else" (Walther & Bazarova, 2008, p. 624). Satisfaction, communication effectiveness, and task accomplishment are three elements that can be chosen to assess this feeling of proximity. One especially interesting aspect in Korzenny's (1978) theory is that of

the bandwidth of the communication medium: low bandwidth refers to few choices given to distant partners for communication (one channel for a telephone conversation, for example), whereas high bandwidth refers to an array of communication channels (several channels such as writing, speaking, and seeing one's partner in the case of videoconferencing).

For the cognitive psychologist Michinov (2008), social presence encompasses three dimensions: immediacy, intimacy, and sociability. Immediacy refers to the feeling of being physically close to one's partner. As McAndrew, Foubister, and Mayes (1996) have pointed out, physical barriers are partly removed thanks to videoconferencing, which seems to foster learner motivation to speak in the second language. It also facilitates the process of contextualization through the use of verbal and non-verbal signs by interlocutors, which link what is said at a particular moment and in a given place to their knowledge of the world, thus facilitating their engagement in the conversation. Besides, as indicated by Chamberlin Quinlisk (2008, p. 31), "gesturing, eye contact, smiling, head nodding" contribute to provide what she terms "immediacy" cues that help lessen "psychological distance between people."

Intimacy pertains to the perception of the quality of the interaction and the feeling of being understood. In a study of Japanese learners of English carrying out a task in audioconferencing and videoconferencing conditions, Yamada and Akahori (2009, p. 20) come to the conclusion that "the partner's image played an important role in helping learners comprehend each other's intended meaning" because having access to social cues, such as laughing or nodding, helps learners to assess the degree of comprehension of their partners. In addition, this study seems to indicate that the presence of one's own and one's partner's images in the videoconferencing condition "allows learners to perceive a similarity to a face-to-face situation" (Yamada and Akahori, 2009, p. 20), and the authors add that "this perception seems to lead to enhanced learning performance" (p. 20).

Finally, sociability refers to the psychological climate of the interaction between partners and the possibility to exchange informally and to develop affinities. This aspect has been brought to the fore by Garrison, Anderson, and Archer (1999), who have defined social presence as the capacity of the participants belonging to a learning community to project themselves socially and emotionally in all the dimensions of their personality through the communication media they use. Affective behaviours are, for these authors, one element to take into account to assess the degree of social presence in distant learning environments. Develotte et al. (2010) have studied the use of the webcam for teaching a foreign language using screen capture recordings of teacher-student interactions and semi-directive interviews. They concluded that "webcamming creates presence at a distance, installs an obvious connection

between the participants and, furthermore, develops the quality of the pedagogical relationship" (2010, p. 309).

Adapting this concept of social presence to situations in which a teacher and one or more learners are brought together through different web-mediated means for learning purposes, we propose to use the term "online teacher presence." By this we mean the feeling learners experience when they interact synchronously with their teacher online. This feeling depends on (a) the degree of immediacy felt toward one's teacher, (b) the extent to which one feels s/he understands or is understood by his/her teacher, and (c) the degree of social and emotional projection, that is, how students feel about the quality, naturalness, and enjoyment of the online interaction.

We hypothesize that "online teacher presence" is a subjective perception that varies across individual learners and is experienced differently depending on the modes used and the social and semiotic resources deployed by online teachers. Because videoconferencing provides rich clues about the teacher, it may thus enhance the perception of online teacher presence and may be conducive to interaction and mutual comprehension, two key features in language learning.

Patterns of web-mediated synchronous interactions and word search episodes

A computer-mediated pedagogical interaction can be studied from a great array of perspectives (Chun, 2008). Because of the synchronous nature of the interactions mediated by audioconferencing or videoconferencing, we choose to focus on the patterns of the interaction and closely examine teacher and learner management of turns and their mobilization of resources and strategies to keep the exchange as smooth and effective as possible. The notion of rhythm is thus crucial to studying the patterns of the exchange. As has been pointed by researchers from the field of mediated discourse analysis, rhythm not only pertains to the speed of accomplishment of certain processes (Lemke, 2005) but is also at the heart of some interactional decisions to slow down or accelerate certain processes and make visible relevant temporal behaviours to one's interlocutor (de Saint-Georges & Duc, 2007).

The study of silences and overlaps (segments of utterances that are pronounced simultaneously by at least two speakers) provides valid information about the management of the interaction and the teacher's capacity to listen and let the learners take the floor (Dejean-Thircuir et al., 2010). As Ricci Bitti and Garotti (2011, p. 89) have remarked about videoconference teaching, it "not only requires more commitment and attention from pupils, but also forces teachers to adapt their performance to a decidedly new interactive context

(familiarization with the technology, optimization of timing and rhythm of the interaction, keeping pupils' attention, etc.)."

Pauses – another crucial element of interactional practices – can be made up of silences signalling that the speaker does not have the means to continue the exchange right after her partner's turn and needs to think about her utterance before pursuing the interaction. In native speaker–non-native speaker (NS-NNS) interactions, the lack of resources sometimes pertains to syntax, but more generally signals a lack of lexis (Long, 1996) and leads to word search episodes, "where a speaker in interaction displays trouble with the production of an item in an ongoing turn at talk" (Brouwer, 2003, p. 535). Because word search episodes have a distinct impact on the patterns of a NS-NNS interaction, we have chosen to examine them more closely to see if they differ under the two experimental conditions.

A look at our data shows that such episodes usually occur as follows:

(a) Pauses intervene when a student encounters a problem in getting her message across, which leads to certain behaviours such as slower rhythm or a thinking face, signalling to the teacher that the learner will need some kind of support. (b) Then the learner engages in a word search or gets the teacher to understand the message despite the lexical problem. The decision rests with the teacher to interpret the different verbal and non-verbal cues so as to provide feedback more or less readily and lead the episode toward its resolution. (c) The episode usually ends when the student manages to get her message across or when the teacher provides the right term, enabling the interaction to proceed.

We take a quantitative approach to word search episodes here, comparing their number and duration across the two web-mediated conditions. Obviously, even if using quantitative methods to study an interaction is bound to yield results that will lack validity from a conversation analysis perspective, In view of the above discussion, we posit that (a) having access to the interlocutor's image may modify the patterns of the interaction, (b) the study of overlaps, silences, and turn duration will provide insights into the patterns of the synchronous exchange, and (c) in the videoconferencing condition the visual channel will give participants access to non-linguistic cues and non-verbal feedback, which may contribute to a more rapid resolution of word searches, compared to the audioconferencing condition in which these searches rely solely on the verbal channel.

Research questions and hypotheses

The present study was designed to answer two main research questions:

1. Is there a difference between learners' perceptions in the audioconferencing and videoconferencing conditions with regard to online teacher presence as defined above? In view of the findings reported above, it was hypothesized that non-native students in the videoconferencing condition would

- have a greater feeling of the psychological and physical presence of the online teacher than those in the audioconferencing condition;
- feel that they had a better understanding of the teacher and would be better understood by her than those in the audioconferencing condition; and
- have a better perception of the quality, naturalness, and enjoyment of the online interaction with their teacher than those in the audioconferencing condition.

2. Is there a difference in the patterns of the synchronous interactions and the word search episodes across the two experimental conditions? To investigate the patterns of the interaction, the following dependent measures are compared between the groups in the videoconferencing and audioconferencing conditions:

- number of student silences;
- number of overlaps;
- number and average duration of student and teacher turns;
- total number of words produced by the teacher.

Then the number and duration of word searches are compared to determine whether one condition is more facilitative of communication. For each of these dependent variables, the null hypothesis tested states that there will be no difference between the videoconferencing and audioconferencing conditions.

Methodology

Participants

Participants were second-year undergraduate students at a French university, who were all studying English for two hours a week as a compulsory non-specialist subject as part of their degree course. Seventy-four students in four different English classes were approached by their teacher and one of the researchers during their weekly English class. They were given very general information about the research project and encouraged to volunteer. All the students present in the four classes were then given Part 1 of the paper version of the Oxford University Press and University of Cambridge Local Examinations Syndicate's Quick

Placement Test (QPT; 2001) two weeks before the experiment began. To ensure that participants in the study had a similar level in English, only students scoring above 24 out of 40 were invited to participate. No incentives were offered to students to encourage their participation. Forty volunteer students with a mean age of 20 years and 2 months (SD = 1.32; range = 17.4–24.6) took part in the study. They were divided randomly between the two experimental conditions to create two groups of equal size, either audioconferencing or videoconferencing, and matched according to QPT score, age, and sex, as shown in Table 1. Mann–Whitney tests indicated that there were no differences across the two experimental groups for QPT score (U = 209, z = .25, p = .806, r = .04) or for age (U = 201, z = .03, p = .978, r = .004).

Table 1: Summary statistics for QPT score, age and sex

	QPT score mean rank	Age mean rank	Sex	
			Male	Female
Videoconferencing ($N = 20$)	20.05	20.45	8	12
Audioconferencing ($N = 20$)	20.95	20.55	8	12

The teacher had several years of experience teaching non-specialist university students in a classroom setting and was a regular user of online tools, such as Skype, for personal communication. However, this was the first time that she had taken part in an online pedagogical interaction. She was not informed of the study's purpose or hypotheses before the experiment. We considered it important to have only one teacher across both experimental conditions to ensure that the teacher effects were consistent across all the participants.

Procedure and data collection

Students were individually taken out of their weekly English class for around half an hour by an assistant. She accompanied them to a quiet office in the same building in which a laptop computer with an integrated webcam had been set up. Before beginning the online interaction with the as yet unknown native English-speaking teacher, students were asked by the assistant to read and sign a consent form, which explained how anonymity and confidentiality would be ensured and asked participants to agree to being recorded and filmed and to accept that the data may be used for research and educational purposes. Next, participants were asked to complete a pre-task questionnaire in French in which they gave their name, sex, and date of birth and answered two multiple-choice questions that aimed to assess their familiarity with online tools such as Skype as well as their perceptions about speaking English. The English translation of these questions is presented in Figure 1. Answers were assigned values (a = 1, b

= 2, c = 3, and d = 4), and these were used as a Likert-type scale to analyze the responses. The mean ranks are presented in Table 2.

On the following scale, how frequently would you say you use the computer to communicate orally online with applications like Skype or MSN [http://www.msn.com/]? (*Circle the answer below best corresponding to your situation.)

(a) I use these tools to speak online very frequently.
(b) I use these tools to speak online fairly frequently.
(c) I rarely use tools to speak online.
(d) I never use these tools to speak online.

On the following scale, how do you estimate your feelings toward speaking English? (*Circle the answer below best corresponding to your situation.)

(a) I really like speaking English and take every opportunity to use it/to practise.
(b) I really like speaking English but I'm afraid of making mistakes and that holds me back.
(c) I feel stressed when I speak English because I'm afraid of not being understood.
(d) I avoid speaking English and make up for it on my written work.

Figure 1 : Pre-task questionnaire

Table 2: Mean ranks for familiarity with online communication tools and perceptions toward speaking English

Variable	Videoconferencing ($N = 20$)	Audioconferencing ($N = 20$)
Familiarity with online communication tools	19.95	21.05
Perceptions toward speaking English	18.88	22.12

Mann-Whitney tests indicated that there were no statistically significant differences in the median scores for either of these questions between the two groups: for familiarity with online communication tools, U = 211, z = .31, p = .757, r = .05; for perceptions about speaking English, U = 232.5, z = .98, p = .329, r = .15. Thus we may conclude that the composition of the video and audioconferencing groups was very similar in terms of students' familiarity with online communication tools and their perceptions about speaking English.

Having completed the pre-task questionnaire, the students were given a short information sheet in French explaining that they were going to interact with an English teacher using Skype to determine how the teacher managed distant interactions with unknown students. The document explained that they would be speaking English for about 10 minutes during which they would first describe and interpret four previously unseen photographs and then try to invent a story linking the four images. Finally, the instructions encouraged them to avoid speaking French during the interaction and to request the teacher's help if they encountered linguistic difficulties. Full details of the task will be provided in the "Task" section.

Skype (http://www.skype.com) had been installed on the laptop computers used by the students and the teacher. In the videoconferencing condition, the laptop users saw the image of their interlocutor in full-screen and saw their own images in a much smaller screen. No images were present for either the students or the teacher in the audioconferencing condition. Skype was selected as the online platform because it was free and easy to use, had fairly good two-way audio and video quality, and was likely to be familiar to many of the participants.

The sessions on both computers were recorded using Camtasia, a recording tool that captures on-screen activity, including sound, and converts it into a video. This program was chosen as it was freely available, was simple to use, and had reliable audio and video quality.

Once students had read the information sheet, the assistant ensured that they were familiar with Skype and understood the experimental procedure. She then gave them the single page with the four photographs (the same four photographs were given to all participants). They were given two minutes to acquaint themselves with these and were instructed not to take any notes, as it was felt that working from notes might interfere with the rhythm and spontaneity of the interaction. Next, the assistant informed the students that she would be leaving the room for the duration of the Skype interaction but would be in the office next door should she be needed. It was hoped that having the assistant on hand to deal with any potential technical difficulties would alleviate some student stress and anxiety. She then activated and verified the Skype connection to the teacher and switched on the screen recording software. At the same time, the teacher activated the screen recording software on her computer.

Once the online interaction began, each student was alone in the office in front of the laptop computer. The teacher was in her own apartment, a decision taken to increase the ecology of the interaction. Indeed, for those students in the videoconferencing condition, we did not wish the teacher to appear in a setting that was familiar, such as a language classroom or a university office, because it may have added to their anxiety and made them more reticent to interact freely, perhaps equating the situation with some sort of assessment. The teacher was instructed to complete each interaction in around 10 minutes, a time chosen after a small-scale pilot study conducted before the main experiment: six other students having the same profile finished the task in approximately this time. The teacher was advised by the researchers to behave as spontaneously as possible with the students even though she had a prepared script that she was expected to follow. This lay on her desk next to her laptop computer but out of the students' view when the webcam was switched on in the videoconferencing condition.

The reason for imposing an approximate time limit and having a set script for each interaction was to facilitate the comparison of the performances of participants and the teacher in the two experimental conditions and also to make the data processing more manageable, given the number of participants in the study. A Mann–Whitney test revealed no between-group differences in the median duration in seconds of the online interaction in the videoconferencing condition and the audioconferencing condition, U = 197, z = -.08, p = .935, r = .01.The research sessions extended over a five-week period. To remain fresh and spontaneous, the teacher had only four or five online interactions in each session. Furthermore, to reduce the possibility of her developing fixed routines in either of the experimental conditions, she alternated between video- and audioconferencing interactions within each research session.

(*Circle the answer below best corresponding to your situation.) Indicate where along the scale						
1. you felt your partner was						
Very absent	1	2	3	4	5	Very present (psychologically)
Very far	1	2	3	4	5	Very close (physically)
2. the understanding of your partner was						
Very difficult	1	2	3	4	5	Very good
3. your partner's understanding of you was						
Very difficult	1	2	3	4	5	Very good
4. the communication with your partner was						
Very bad	1	2	3	4	5	Very good
Not natural at all	1	2	3	4	5	Extremely natural
Not pleasant at all	1	2	3	4	5	Extremely pleasant

Figure 2: Post-task questionnaire on perceptions

Once the Skype session was open but before beginning the task itself, the teacher tried to put participants at ease by greeting them warmly, telling them she was in her apartment, asking them their names, and then introducing herself using her first name only. After this short warm-up, she indicated to participants that she had not seen the pictures they were going to describe (although she was in fact familiar with them). Indeed, it was considered that if the students knew she had already seen the photographs, all the word search episodes would seem very artificial to them, because she would already know what information they were attempting to communicate and which lexical items were causing them difficulty. Finally, she instructed the students in the videoconferencing condition not to show her the pictures they were describing. When the interaction was complete, the assistant returned to the office and switched off Skype.

Participants were then invited to complete a post-test questionnaire in French, in which some of the items were based on previous questionnaires such as Yamada and Akahori's (2009).

Participants were asked to rate their perceptions of the interaction on seven items on a 5-point Likert scale, as displayed in Figure 2. The aim of this questionnaire was to examine if there was a difference in their perceptions of the interaction between the videoconferencing and audioconferencing conditions (see the section "Psychological factors in video-mediated L2 interaction" above).

Task

In the experimental task, participants had to describe in detail four previously unseen and unrelated photographs. Such a referential communication task was chosen because, as pointed out by White and Ranta (2002, p. 264), learners have to be "very precise in both vocabulary and structure, thus making demands on the learner's ability to quickly access specific linguistic knowledge." These photographs showed individuals in simple situations (a group of young people at an outdoor concert; an old lady in a hospital; an intimate funeral procession; a sad child holding a teddy bear). Because lexical items carry a heavy communicative load, it is essential to negotiate the meaning of these items if they are unknown to learners, to avoid communication breakdowns that would prevent the conversation from moving forward (Blake & Zysik, 2003). The four photographs were selected because each one contained what were considered to be problematic lexical items (e.g., loudspeakers, earring, wheelchair) likely to trigger word search episodes. If students did not provide sufficient details or justify what they were saying in their descriptions with precise references to elements in the photographs, or if their descriptions were considered to be unclear or not specific enough, the teacher was instructed to incite participants to elaborate, asking further open questions such as "How can you tell?" and "Why do you say that?" The aim of these probing questions was to provoke word search episodes. When the interaction came to a halt because students clearly lacked a particular key lexical item, they were encouraged to reformulate or describe the item in question. If they provided a word in French, the teacher feigned a lack of understanding, prompting students to find another way of communicating their idea.

Once all the data collection sessions were complete, the teacher was interviewed by the researchers. The aim was to assess how she had experienced the two experimental conditions, to examine her perceptions of the differences between them in terms of both her own and the students' behaviour and engagement, and to have her impressions of the impact of the webcam on the online interaction. We will refer back to the teacher retrospective interview in the discussion section, in order to to triangulate her comments and reflections with the results

from the quantitative analyses and thus increase the "interpretative validity" of our research as is deemed crucial by Dörnyei, (2007, p. 58).

Data analysis

The videos of the computer screen activity recorded with Camtasia were analyzed using ELAN (http://tla.mpi.nl/tools/tla-tools/elan/). This software allows researchers to access the image of the participants (labelled 1 in Figure 3), their verbal outputs (2), the transcription of the turns (3), and indications of duration for each turn (4).

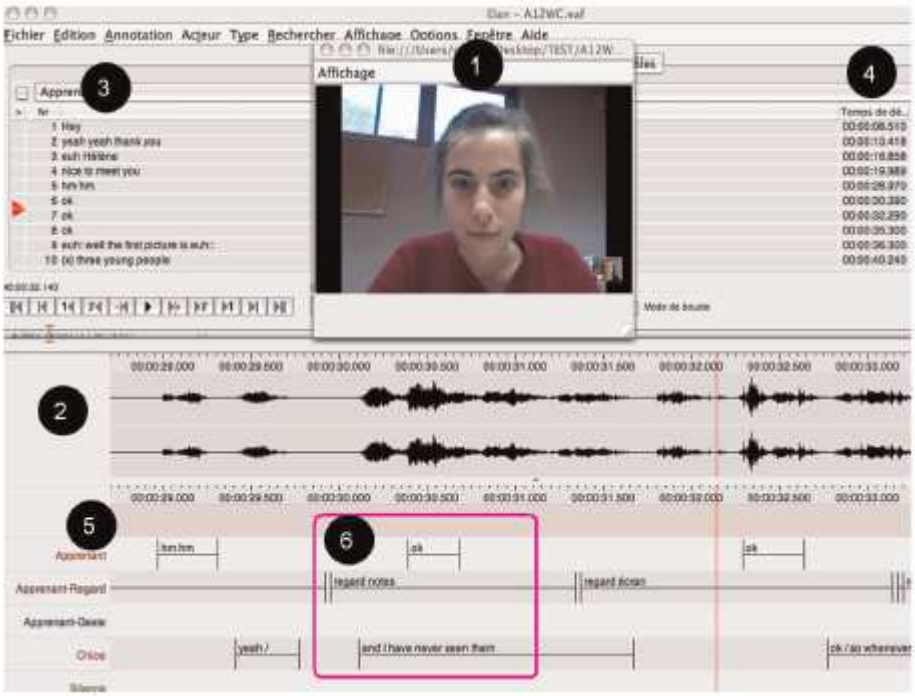

Figure 3: Screen capture of ELAN

It enables them to annotate the recordings on multiple parallel tiers, with a new tier for each feature (5) and the ability to identify different phenomena such as silences, overlaps, or word search episodes (6) (see Figure 3). The duration of each word search episode was calculated from the moment a student encountered a problem with a lexical item and then engaged in finding the proper word or getting the teacher to understand the message despite the lexical problem. It ended when the student succeeded in getting her message across or when the teacher provided the appropriate word, enabling the interaction to move on. ELAN was selected because it allows the transcription to be synchronized with the video and audio recordings and it provides precise temporal data. The numerical data from ELAN were then imported to the SPSS computer-based statistical package (version 20) for analysis using 2-tailed Mann–Whitney tests, allowing statistical comparisons to be made between participants in the videoconferencing and audioconferencing conditions. The Mann–Whitney test was

chosen in view of the small size of the sample (20 participants in each condition). The effect size (r) was calculated following Rosenthal (1991). The alpha level was set at .05.

Results

The results of this study examining the impact of the webcam on an online interaction will be reported in two parts. We begin by comparing students' perceptions of the online interaction across the two experimental conditions with regard to online teacher presence (research question 1). Then we compare the patterns of the interaction across the two experimental conditions, by addressing first the students' performance, then the teacher's performance; finally, the word search episodes are considered (research question 2).

Perceptions

To determine whether the students' perceptions of online teacher presence differed in the videoconferencing and audioconferencing conditions, a Mann–Whitney test was carried out on each of the dependent measures. Mean ranks are presented in Table 3.

Table 3: Mean ranks for perception measures

Variable	Videoconferencing ($N = 20$)	Audioconferencing ($N = 20$)
Psychological presence	20.30	20.70
Physical presence	21.32	19.68
Understanding of native partner	18.00	23.00
Understanding by native partner	20.08	20.92
Quality of communication	20.35	20.65
Naturalness of communication	20.88	20.12
Enjoyment of communication	20.30	20.70

The results indicated that there were no significant differences between the two groups on any of the perception variables: for psychological presence U = 182.5, z = -.68, p = .496, r = .12; for physical presence U = 183.5, z = -.47, p = .636, r = .07; for understanding of the native partner U = 250, z = 1.87, p = .062, r = .30; for understanding by the native partner U = 208.5, z = .28, p = .781, r = .04; for quality of communication U = 203.0, z = .10, p = .920, r = .02; for naturalness of communication U = 192.5, z = -.23, p = .820, r = .04; for enjoyment of communication U = 204.0, z = .16, p = .877, r = .02. For the variable understanding of the native partner, the result is tending toward significance (p = .062) with a medium effect size (r = .30), indicating a pattern of potential interest. Contrary to our hypothesis, this result suggests that students in the audioconferencing condition had a better understanding of the native English-speaking teacher than the students in the videoconferencing condition. To sum up, students in the videoconferencing condition did not have more positive perceptions of online teacher presence than those in the audioconferencing condition.

Table 4: Mean ranks for student performance

Variable	Videoconferencing ($N = 20$)	Audioconferencing ($N = 20$)
Number of student silences (> .1 sec)	16.28	24.72
Number of overlaps	23.72	17.28
Number of student turns	18.80	22.20
Average duration of a student turn (in secs)	22.52	18.48

Patterns of the interaction and word search episodes

The patterns of the synchronous interactions across the two experimental conditions are addressed here by comparing the dependent measures related first to the students' performance, then to the teacher's. The dependent measures investigated for the students are (a) number of student silences of > .1 seconds, (b) number of overlaps, and (c) number and duration of student turns. Mean ranks are presented in Table 4. Mann–Whitney tests indicated that there was a significant difference between the groups for one of the four dependent measures investigated. Indeed, student silences were significantly longer in the audioconferencing condition than the videoconferencing condition, $U = 284.5$, $z = -2.29$, $p = .022$, $r = .36$. For the remaining variables, the results were as follows: overlaps of student and teacher turns, $U = 135.5$, $z = -1.75$, $p = .081$, $r = .28$; number of student turns, $U = 234$, $z = .92$, $p = .358$, $r = .15$; average student turn duration, $U = 159.5$, $z = -1.10$, $p = .273$, $r = .17$. Although the reading for overlaps of student and teacher turns failed to reach statistical significance, the p-value was fairly low, with an effect size approaching the medium level, a result worthy of further investigation.

Table 5: Mean ranks for teacher performance

Variable	Videoconferencing ($N = 20$)	Audioconferencing ($N = 20$)
Number of teacher turns	20.12	20.88
Average duration of a teacher turn (secs)	18.52	22.48
Teacher's total number of words	17.22	23.78

The teacher's performance in the two experimental conditions is examined on the following dependent measures: (a) number and duration of teacher turns, (b) teacher's total number of words. Mean ranks for teacher performance are presented in Table 5.

Table 6: Mean ranks for word search episodes

Variable	Videoconferencing ($N = 20$)	Audioconferencing ($N = 20$)
Number of word search episodes	20.22	20.78
Duration of word search episodes (secs)	18.98	22.02

Mann–Whitney tests revealed no significant differences between the videoconferencing and audioconferencing groups on the three teacher variables: for number of teacher turns, U = 207.5, z = .20, p = .839, r = .03; for duration of teacher turns, U = 239.5, z = 1.07, p = .285, r = .17; for teacher's total number of words, U = 265.5, z = 1.77, p = .076, r = .28. The last result, total number of words used by the teacher, is tending toward significance, with an effect size falling just below the .30 criterion for a medium effect size. This indicates a pattern of potential interest that will be discussed further below. Mean ranks for the number and duration of word search episodes in the videoconferencing and audioconferencing conditions are given in Table 6. No differences were statistically established between the two experimental conditions for the two variables relating to word search episodes: for the number of episodes, U = 205.5, z = .15, p = .881, r = .02; for duration of episodes, U = 230.5, z = .83, p = .409, r = .13.

Discussion and conclusion

The results of this study show that there are fewer differences than we had anticipated on the dependent variables compared across the videoconferencing and audioconferencing conditions.

With regard to the question of students' psychological perceptions of online teacher presence, it was predicted that the videoconferencing condition would be the preferred medium. Our findings show that there were no significant differences between the two groups on any of the measures. Rosell-Aguilar's (2007) study comparing FTF tutorials to audiographic tutorials highlighted the importance in the latter condition of the tutor communicating warmth and humanity through the verbal channel to compensate for the lack of visual clues. The teacher in our study apparently succeeded in transmitting her warmth and presence to the students in both experimental conditions, which may account for the fact that there were no differences between the students' feelings of online teacher presence in the two conditions. On the one measure that approached statistical significance with a medium effect size, it was the students in the audioconferencing condition whose understanding of the native English-speaking teacher was better than that of the students in the videoconferencing condition, a result that runs counter to our hypothesis. Indeed, earlier research findings from the field of language teaching and learning had argued that being able to see the interlocutor's image should facilitate understanding, thanks to social cues such as laughing or nodding, and should make participants feel physically and psychologically closer to one another (e.g., Chamberlin Quinlisk, 2008; Develotte et al., 2010; Yamada & Akahori, 2009). Yet this did not seem to

happen in the present study. Indeed, during the retrospective interview the teacher remarked that she felt that "the lack of images helped students to focus on the words and their meaning, so maybe this obliged them to concentrate more." If this were indeed the case, we might hypothesize that, in a language learning and teaching online interaction, being able to see the image of the interlocutor and oneself during an videoconferencing interaction may in fact be distracting for some learners who, as a consequence, will be less focused on the verbal components of the teacher's message, thus hindering understanding to some extent.

Concerning the patterns of the interaction, our results revealed that there were significantly more student silences in the audioconferencing condition. Our results for overlaps were tending toward significance with more overlaps in the videoconferencing condition. An inability to see one's partner in the audioconferencing condition offers no paralinguistic cues as to when to take the floor, which could explain that participants here are more reticent to interrupt their interlocutor. This lack of paralinguistic and other social cues could also explain why there are more student silences in the audioconferencing condition, as learners wait for the teacher to respond or to react verbally before taking the floor or continuing their turn. Indeed, comments made by the teacher in the retrospective interview offer some support to this. She observed that "the flow of the conversation was interrupted more often in the audioconferencing condition when students had linguistic difficulties, like not knowing a particular word. So they had to stop and reflect or wait for me to intervene." She also commented that "with the video it was easier for me to pick up on when the students needed help through their non-verbal cues. This helped me intervene faster, helping the interaction to proceed more smoothly."

The difference between the results for the teacher's total number of words in the videoconferencing and audioconferencing conditions approached significance, with an effect size tending toward the medium level, with the teacher using more words in the audioconferencing condition. The teacher seemed to be aware of this, as she noted in the retrospective interview: "I felt like I had to provide the students with more information or assistance to move the interaction forward because of a lack of visual cues." Although the teacher had been asked to behave in a similar manner in the two conditions, our results suggest that the constraints of the situation may have led her to speak more abundantly when there was no webcam. Indeed, a possible explanation for this, brought to the fore in a similar online teaching situation (Dejean-Thircuir et al., 2010), might be that she felt the need to fill the silences more pressingly in the audioconferencing condition. In the videoconferencing

condition, on the other hand, the learners and the teacher used the webcam to move the conversation forward more rapidly and seamlessly. Indeed, the webcam enabled learners to provide non-linguistic cues to the teacher, to make gestures or to mimic, and to check that the teacher was ratifying crucial elements of their message. On the teacher's side, the webcam was used to give the learner indications about the degree of success of word searches, to give descriptions or explanations (smiles, nods, gestures), and to obtain clues about the learner's progress and engagement in the task. Neither the number nor the duration of the word search episodes differed significantly across the two conditions. Although the current study revealed no real quantitative differences between the two conditions, a further study of the data is currently being undertaken using a qualitative analysis to compare how learners negotiate the meaning of the problematic lexical items in the two conditions of this online teacher-learner pedagogical interaction and to explore the extent to which negotiation episodes lead to full or partial understanding of the target item.

Although it was carefully planned and yielded highly comparable data for both conditions, this study presents two main limitations. The first concerns the task chosen for the experiment, which might have been more genuine in terms of output if the teacher had not seen the photographs beforehand and did not have to pretend that she was unfamiliar with them. An alternative task could have required the learners to describe the layout of a room to the teacher while she produced a drawing according to their instructions. The second limitation concerns the duration of the experiment (around 10 minutes) and its punctual character (neither the learners nor the teacher were attuned to such a learning situation). Indeed, a second instance of the interactions might have shown that some of the variables were related to the novelty of the medium and/or the task (e.g., the duration of the exchanges and the number of turns) and that the familiarity with the medium and/or the task might account for these. A longer interaction or one repeated across several weeks with the same participants, all of whom experienced both conditions over several sessions, might have yielded different data and produced potentially different perceptions and behaviours. This limitation is inherent in the experimental conditions, which for reasons of comparability and cost are bound to lack ecological validity.

Overall, under the conditions of this study, the impact of the webcam on an online L2 interaction does not seem so straightforward, and the results obtained run counter to our original intuitions that more communication cues are better for pedagogical interactions in L2. As Walther (2011, p. 19) has underlined, "the inherent value of visual cues in online

communication may be as scientifically unfounded as it is intuitive." Indeed, in some of the studies he examined (e.g., Galegher and Kraut, 1994), he reported a discrepancy between perception and behaviour regarding the value of video, but our study produced results that show no such discrepancy. The participant's voice seems to be sufficient to make the rhythm of the interaction fluent enough without requiring visual support. As Walther and Bazarova (2008, p. 626) have remarked, "although users are inclined to prefer higher-bandwidth media for informationally complex conversations, when the choice of high-bandwidth alternatives is restricted, effective and satisfying communication takes place using lower-bandwidth media nevertheless."

Our findings seem to indicate that the way learners perceive online teacher presence, regardless of whether they are in an audioconferencing or videoconferencing condition, will depend on a complex interplay of factors relating to the pedagogical actions that can be deployed verbally and non-verbally by the teacher. These actions – including knowing when and when not to intervene, interrupt, prompt, step back, nod, laugh, smile or show warmth – will vary according to the teacher's perception of the learner's needs and to the situation itself. Such pedagogical actions, and the semiotic decisions they incur, require from online teachers what Van Lier (2004, pp. 148–149) has termed as "'just right' and 'just-in-time' responses and interventions [that] must be seen as among the most complex and demanding decisions experienced teachers make."

It should be emphasized that the present study is one of the first studies that does not examine peer-to-peer interactions, but rather teacher-to-learner interactions. In other studies (e.g., Yamada & Akahori, 2009), anxiety and unease seemed to be higher when the webcam was not on, but in our experiment, a teacher is in charge of leading the interaction, which might provide some comfort to the learners and augment feelings of online teacher presence. Yet this does not mean that the webcam image is not facilitative or does not change the quality of a mediated interaction. Following Levy and Stockwell (2006, p. 105), we believe that "material dimensions (or affordances) constrain, enable, and shape the communication that occurs, although these effects are subtle and difficult to pin down with precision." A qualitative study of the same data is thus needed to examine how the webcam image was used strategically by participants. If a teacher has a crucial role to play in an online exchange, only a minute study of non-verbal behaviour in the videoconferencing condition will allow us to identify when and how interaction is facilitated by the appropriate use of the webcam. Such a study should enable us to provide directions as to the selection of a medium by a teacher in

line with the work of Daft, Lengel, and Trevino (1987), the presence of the webcam image being detrimental for one task or one moment in an online course while it can be beneficial for some others.

References

Blake, R., & Zyzik, E. (2003). Who's helping whom?: Learner/heritage-speakers' networked discussions in Spanish. *Applied Linguistics, 24*(4), 519–544.

Brouwer, C. E. (2003). Word Searches in NNS–NS Interaction: Opportunities for Language Learning? *The Modern Language Journal*, 87, 534–545.

Chamberlin Quinlisk, C. (2008). Nonverbal communication, gesture and second language classrooms: a review. In S. G. McCafferty & G. Stam. (2008). (Eds), *Gesture. SLA and Classroom Research* (25-44). New York: Routledge.

Chun, D. (2008). Computer-mediated discourse in instructed environments. In S. Sieloff Magnan, (Ed.), *Mediating discourse online* (15–45). Amsterdam: John Benjamins.

Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly, 11*(3), 355–366.

Dejean-Thircuir, C., Guichon, N., & Nicolaev, V. (2010). Compétences interactionnelles des tuteurs dans des échanges vidéographiques synchrones. *Distances et savoirs, 8*(3), 377–393.

de Saint-Georges, I., & Duc, B. (2007). Order, duration and rhythm: tuning to complex temporal arrangements in workplace learning. In G. Morello (Ed.), *Actes du 5th Palermo International Conference on Social Time. Retroscapes and Futurescapes: Temporal tensions in organizations.*

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

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Galagher, J., & Kraut, R. E. (1994). Computer-mediated communication for intellectual teamwork: an experiment in group writing. *Information Systems Research, 5*, 110–138.

Garrison, D.R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: computer conferencing in higher education. *The Internet and Higher Education, 2*(2-3), 87–105.

Hopkins, J. E. (2010). Distance language learners' perceptions of assessed, student-led speaking tasks via a synchronous audiographic conferencing tool. *Innovation in Language Learning and Teaching, 4*(3), 235–258.

Hrastinski, S. (2008). Asynchronous and Synchronous E-Learning. *EDUCAUSE Quarterly, 31*(4), 51–55.

Korzenny, F. (1978). A theory of electronic propinquity: Mediated communications in organizations. *Communication Research, 5*(1), 3–24.

Lemke, J. (2005). Place, pace and meaning: multimedia chronotopes. In S. J. Norris & H. Rodney (Eds.), *Discourse in action: Introducing mediated discourse analysis* (pp. 110-122). London & New York: Routledge.

Levy, M., & Stockwell, G. (2006). *CALL dimensions. Options and issues in Computer-Assisted Language learning*. New Jersey: Lawrence Erlbaum Associates.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W.C. & T.K. Bahtia, T. K. (Eds.), *Handbook of second language acquisition* (pp. 413-68). New York: AcademicPress.

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

Michinov, E. (2008). La distance physique et ses effets dans les équipes de travail distribuées : une analyse psychosociale. *Le travail humain, 71*(1), 1–21.

Ricci Bitti, P. E., & Garotti, P. L. (2011). Non-verbal communication and cultural differences: issues for face-to-face communication over the Internet. In A. Kappas & N. C. Krämer (Eds.), *Face-to-face communication over the Internet* (pp. 81–99). Cambridge: Cambridge University Press.

Rosell-Aguilar, F. (2007). Changing tutor roles in online tutorial support for open distance learning through audio-graphic SCMC. *The JALT CALL Journal, 3*(1-2), 81–94.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd edition). Newbury Park, CA: Sage.

Satar, M. (2013). Multimodal Language Learner Interactions via desktop videoconferencing within a Framework of Social Presence: Gaze. *ReCALL, 25*(1) 122–142.

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

Van Lier, L. (2004). *The ecology and semiotics of language learning*. Dordrecht: Kluwer Academic.

Walther, J. B. (2011). Visual cues in computer-mediated communication: sometimes less is more. In A. Kappas & N. C. Krämer (Eds.), *Face-to-face communication over the Internet* (pp. 17–38). Cambridge: Cambridge University Press.

Walther, J. B., & Bazarova, N. N. (2008). Validation and Application of Electronic Propinquity Theory to Computer-Mediated Communication in Groups. *Communication Research, 35*(5), 622–645.

White, J., & Ranta, L. (2002). Examining the Interface between Metalinguistic Task Performance and Oral Production in a Second Language. *Language Awareness, 11*(4), 259–290.

Yamada, M., & Akahori, K. (2009). Awareness and performance through self- and partner's image in videoconferencing. *CALICO Journal, 27*(1), 1–25.

Yamada, M., & Goda, Y. (2012). Application of Social Presence Principles to CSCL Design for Quality Interactions, Jiyou Jia (Ed.). *Educational Stages and Interactive Learning: From Kindergarden to Workplace Training*, IGI Global, 31-48.

Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: it's about time! *Language Learning & Technology, 14*(3), 72–93.

Quick Placement Test (2001) Oxford University Press and University of Cambridge Local Examinations Syndicate

Software

CAMTASIA: http://camtasia-for-mac.en.softonic.com

ELAN: http://tla.mpi.nl/tools/tla-tools/elan/

MSN: http://www.msn.com/

SKYPE: http://www.skype.com

SPSS: http://www-01.ibm.com/software/fr/analytics/spss/