A semiotic perspective on webconferencing-supported language teaching.

Nicolas Guichon, Ciara R. Wigham

HAL Id: halshs-01236195

https://halshs.archives-ouvertes.fr/halshs-01236195

Submitted on 9 Dec 2015

A semiotic perspective on webconferencing-supported language teaching

Nicolas Guichon[1] & Ciara R. Wigham[2]

[1]Université Lumière Lyon 2 – Laboratoire ICAR - ENS de Lyon, France

(email: nicolas.guichon@univ-lyon2.fr)

[2]Université Lumière Lyon 2 - Laboratoire ICAR - ENS de Lyon, France

(email: ciara.wigham@univ-lyon2.fr)

Abstract

In webconferencing-supported teaching, the webcam mediates and organizes the pedagogical interaction. Previous research has provided a mixed picture of the use of the webcam: while it is seen as a useful medium to contribute to the personalization of the interlocutors' relationship, help regulate interaction and facilitate learner comprehension and involvement, the limited access to visual cues provided by the webcam is felt as useless or even disruptive.

This study examines the meaning-making potential of the webcam in pedagogical interactions from a semiotic perspective by exploring how trainee teachers use the affordances of the webcam to produce non-verbal cues that may be useful for mutual comprehension. The research context is a telecollaborative project where trainee teachers of French as a foreign language met for online sessions in French with undergraduate Business students at an Irish university. Using multimodal transcriptions of the interaction data from these sessions, screen shot data, and students' post-course interviews, it was found, firstly, that whilst a head and shoulders framing shot was favoured by the trainee teachers, there does not appear to be an optimal framing choice for desktop videoconferencing among the three framing types identified. Secondly, there was a loss between the number of gestures performed by the trainee teachers and those that were visible for the students. Thirdly, when trainee teachers were able to coordinate the audio and kinesic modalities, communicative gestures that were framed, and held long enough to be perceived by the learners, were more likely to be valuable for mutual comprehension.

The study highlights the need for trainee teachers to develop critical semiotic awareness to gain a better perception of the image they project of themselves in order to actualise the potential of the webcam and add more relief to their online teacher presence.

Key words: multimodality, webconferencing-based interaction, framing, gestures, teacher-training

1 Introduction

In webconferencing-supported teaching, the computer screen constitutes the interface, both technological and semiotic (see Souchier, Jeanneret & Le Marec, 2003: 35), between the protagonists of the interaction. The screen gathers an array of information that is made accessible through different modes (writing, aural and visual). Among the various tools that are available in this situation (text chat, whiteboard, etc.), the webcam provides participants with a restricted and imperfect access to everyone's image and voice (Zähner, Fauverge & Wong, 2000).

The image provided by the webcam is composed of a sequence-shot, that is to say a single, uninterrupted, ephemeral and unedited filmic element. Although the quality of desktop

videoconferencing has constantly been improving, bandwidth and processing limitations continue to entail imperfect synchronization between sound and image, micro-cuts and even breakdowns (Parkinson & Lea, 2011). Besides, the delivery of the webcam image is below the 24-image per second filmic norm, which can make its reception somewhat jerky. Yet, despite its many imperfections, the webcam image, when used in pedagogical interactions, seems to focus the interlocutors' attention and push the other semiotic resources into the background (Guichon & Cohen, 2014).

Given that the webcam image not only mediates but also organizes the pedagogical interaction, this article aims to study its meaning-making potential from a semiotic perspective. We follow van Lier's (2004: 62) proposition to envisage any learning context as "an activity space" which includes affordances that may become "meaning-making material" (*ibid.*) when they are used appropriately. Van Lier defines affordances as "possibilities for action that yield opportunities for engagement and participation that can stimulate intersubjectivity, joint attention, and various kinds of linguistic commentary" (2004: 81). Kress (2009) and Jewitt (2009, 2011) have shown that the study of language learning and teaching should include an examination of the variety of modes that make up a pedagogical situation and therefore all the semiotic resources that are available as well as the ways in which they are orchestrated. This theoretical standpoint requires a "multimodal analysis" (Jewitt, 2009) in order to explore how teachers (and learners) "make choices among various semiotic options in discursive practices" (Pinnow, 2011: 384) and assess their "meaning potential, based on their past uses, and affordances based on their possible uses" (Jewitt, 2011: 185).

Drawing from this semiotic perspective, the present study sets out to determine how the webcam is an artefact that organizes the pedagogical interaction and provides a certain number of affordances. In line with similar studies (Codreanu & Celik, 2013; Develotte, Guichon & Vincent 2010; Kern 2014), we will use data from a telecollaboration project using videoconferencing whereby trainee teachers mediate online sessions designed to develop interactions skills of their distant Irish learners of French. The specific aim of this study is to determine how trainee teachers learn to harness the potential of the webcam for language teaching by paying attention to framing.

2 Theoretical framework

2.1 Examining a pedagogical interaction from different perspectives: champ, contre-champ and hors-champ

In Figure 1, we identify the different elements that constitute a webconferencing-based interaction and use notions from the field of film semiotics to organize our study. These notions are those of *champ*, *contre-champ* and *hors-champ*,[1] which we will explain with reference to Deleuze's (1983) contribution to film analysis. We organize our presentation of these elements from the teachers' perspective, starting with the *champ*.

---

[1] We have chosen to keep the French terms because what is meant by *champ* and *contre-champ* is lost by the usual translation of "shot" and "reverse shot". Champ conveys an idea of depth better than shot does.
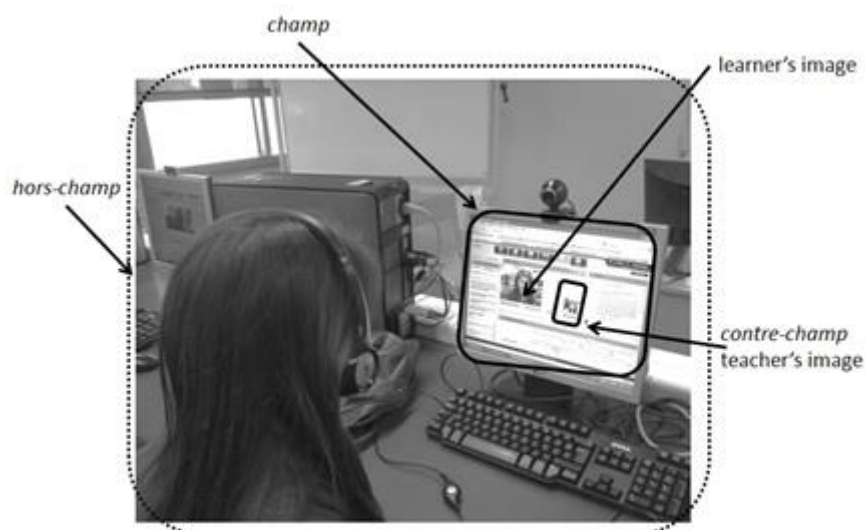
Figure 1. An online pedagogical interaction from different perspectives

The *champ* is restricted to what is visible within the frame of the screen. The information provided by the webcam during a videoconferencing interaction is rather poor in content. From the teacher's perspective, its main elements are:

- the learner gestures that appear in the frame;
- non-verbal micro-events that come across the learners' faces (smiles, frowns, mimics);
- lip movements when the learner speaks (usually with a slight time lag between image and sound);
- some information about the learner's contexts (classroom setting, decor, etc.) and personality (clothes, etc.) but, because of the close-up framing and the fixedness of the shot, contextual elements usually remain fairly unremarkable.

The visual cues produced by the webcam are scarce since the learners' faces saturate the *champ* by a close-up and there are very few changes in what is made visible to the interlocutors. Yet, being able to view the partner's image via the webcam seems to facilitate mutual comprehension and help decode an interlocutor's intended meaning (Guichon & Cohen, 2014; Yamada & Akahori, 2009).

The ***contre-champ*** is composed of the teacher's own image, either inserted in or next to the learner's image. This is a unique affordance of videoconferencing: contrary to a face-to-face teaching situation, the *contre-champ* image allows online teachers to see themselves. Yamada and Akahori (2009) have indicated that having access to one's self-image seemed to enhance a participant's awareness of his actions. If teachers choose to pay attention to this counter-image, they can regulate/modify their own image by getting an almost instant view of what they are endeavouring to make visible (the explicitness of a facial expression, the theatricality of a gesture). It can be hypothesized that the possibility to see oneself while engaged in aninteraction with a interlocutor can reinforce the empathic dimension of webconferencing-supported communication. Contrary to other online learning environments, which offer cultural artefacts (newspaper articles, TV reports, radio programmes), a teacher involved in a webconference *is* the actual document. By allowing his image and voice to become the main learning resource, the teacher's face gains a metonymic value: it becomes, by extension, the face of French

language and culture, a perception that is reinforced by the fact that trainees are actually in France – even if France is barely visible in the frame.

The **hors-champ** corresponds to the environment of the interaction, that is to say all the elements that remain out of the frame but are nevertheless part of the pedagogical interaction. Jones (2004) cautions researchers working in the computer-mediated communication (CMC) field not "to stop at the screen's edge" (p. 24) so as to gain a fuller understanding of context and not to separate what is going on online from what is going on around the participants.

For instance, Figure 2 shows an extract from our data depicting a trainee teacher shot from the side, which provides a glimpse of her environment.
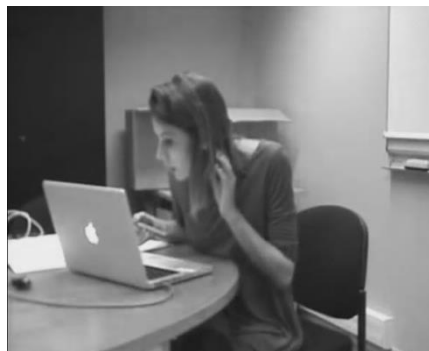

Figure 2. Hors-champ

As can be seen, the *hors-champ* is comprised of **physical** elements of the context (the pencil she is holding, the sheets of paper that are placed beside the computer, the space around her, the way she is sitting on her seat, the lighting, etc.). Although they are invisible, the heat and the noise made by the other trainees also form part of the context. Such a view helps remind us that below the faces and the shoulders that occupy most of the frame in desktop webconferencing-supported teaching, there are bodies that may be tired, relaxed or tense, in sum, bodies that have a life and remain invisible to the interlocutor.

The hors-champ also comprises **symbolical** elements: values, attitudes, teaching philosophy but also perceptions vis-à-vis the learners, their culture, needs and academic objectives. Despite its invisibility, the symbolical hors-champ is omnipresent: it depends on the construction of a common context through a "grounding" process (Clark & Brennan, 1991) that facilitates coordination and collaboration. It is negotiated by the *interactants* and involves the comprehension of cultural elements such as values, socio-economic aspects and language registers.

2.2 Attention to framing

In webconferencing-supported teaching, framing is a matter of choice determined by the participants: they can choose to focus on their face or can provide a view of their shoulders or their torso, thus providing some information about their clothes. There are two ways to attend to the content of the frame, either by adjusting the camera (see Figure 3) or by sitting more or less close to the webcam. It is important to underline that the *contre-champ* provides cues as to the exact content of the frame and the image that is projected.

Figure 3. Attention to framing

Sindoni (2013) insists that the way people position themselves in front of the camera is never neutral but involves "an *intentional* act on the part of each participant" and thus becomes "an integral part of the interaction" (Sindoni, 2013: 57). Such semiotic choices, Jewitt (2009) insists, realize social functions.

The work of Goffman (1974) that focused on non-mediated interactions is of paramount importance to feed the reflection on framing and enrich the cinematic reference. Goffman envisages framing as timely interactional procedures that provide interlocutors about one another's current action and projected identity. These procedures contribute to creating "an environment of mutual monitoring possibilities" (Goffman, 1974: 134) whereby each interlocutor both makes himself accessible to others while they simultaneously become accessible to him. In line with Goffman's approach, Jones (2004) highlights that CMC technologies offer unique semiotic possibilities allowing users "to be present to one another and to be aware of the people's presence" (Jones, 2004: 23). Guichon & Cohen (2014) have proposed the term "online teacher presence" by which they refer to the ways teachers make themselves present to their students during online synchronous interaction. Online teacher presence is the subjective perception experienced by students towards their online teacher and it depends on three elements: the degree of proximity felt towards their teacher (immediacy), the extent to which they feel they understand or are understood by him or her (intimacy), and the degree of social and emotional projection (sociability), that is, how they feel about the quality, naturalness, and enjoyment of the online interaction.

Framing in other - non-pedagogical - conversational contexts may not be that important. For instance, Sindoni (2013) has studied people interacting in a videochat (*Camfrog*, Camshare Inc. 2014) and has observed that, although they left their webcam on, participants were performing other tasks (e.g., checking emails or webpages). She concluded that they seemed "to demonstrate a parallel high level of reciprocal tolerance with regard to the low levels of the other participant's attention" (Sindoni, 2013: 77).

Yet, visual cues may prove important for pedagogical interaction since they contribute to "impression formation, rapport, and acquaintanceship development" (Manstead, Lea & Goh, 2011: 147) and help regulate the interaction, for instance by helping define turn-taking and reducing overlaps (Guichon & Cohen, 2014; Ricci Bitti & Garotti, 2011). Furthermore, the space that is shared by the teacher and the learner thanks to framing constitute a site of display that is defined by Jones (2009: 114) as "social occasions in which particular configurations of modes and media converge in a particular time and space in order to make particular social actions possible". Thus, it can be thought that attention to framing constitutes a crucial element of online teaching since what the online teacher chooses to show of herself (facial expressions, gestures, clothes, smile, nods, etc.)

contributes to the personalization of the relationship and facilitates learner comprehension and involvement.

2.3 Research questions

To conduct an online interaction, we suggest that language teachers need to develop a critical semiotic awareness that allows them to make the most of the semiotic resources and the tools that are at their disposal and to adapt their performance, in order to optimize the learning potential and maintain the learners' attention (Ricci Bitti & Garotti, 2011: 89). We see this critical semiotic awareness as an intricate part of what it is to teach a language online. It has a semiotic content because teaching is seen as the production of an array of signs that are verbal and non-verbal. It needs to become critical since teachers progressively have to (1) develop a high level of consciousness regarding all the information they are conveying when they interact with their learners and (2) become able to assess the impression with which they provide their learners.

The present qualitative study aims to analyse how framing is used by trainee teachers with regards to the pedagogical interaction. We choose to explore the following questions:

1. How do trainees position themselves in front of the webcam?
2. What are the communicative functions of gestures that are visible / invisible in the frame?
3. To what extent were the trainees aware of their framing choices? How were the framing choices perceived by the learners?

3 Methodology

3.1 Pedagogical context and participants

The context for this study was a telecollaborative project that brought together French as a foreign language (FFL) trainee teachers (henceforth 'trainees') and undergraduate Business students in their third semester of learning French at Dublin City University.

The trainees were second-year students on the Master of Arts in Teaching French as a Foreign Language programme at Université Lyon 2 (UL2). The telecollaboration project formed part of an optional module entitled 'Online teaching'. The primary objectives were for students to:

- develop professional skills to teach FFL online (activity preparation, mediation, corrective feedback);
- analyse their teaching practice and develop reflective analysis around this;
- understand issues linked to pedagogical engineering.

For the language students from *Dublin City University* (DCU) the project formed part of a 12-week blended French for Business module. This module, worth 5 ECTS credits, had CEFR level B1.2 (Council of Europe 2001) as its minimum exit level. The project was aligned with various module learning outcomes including being able to:

- understand standard written and spoken language (live or broadcast), on topics related to professions, work placements and job applications;
- apply for a job or a work placement through French;
- demonstrate a sufficient range of written and spoken language to talk about competencies, past academic or professional experiences, and personal goals;
- make effective use of digital resources and tools for communication purposes.

The participants met for six 40-minute online sessions in autumn 2013 via the webconferencing platform Visu (Bétrancourt, Guichon & Prié 2011). Visu was specifically designed for synchronous language teaching and was the outcome of a research and development project involving computer scientists and specialists of language education and cognitive psychology (see below for the description of functionalities). Following an introductory session, each online session was thematic and focused on Business French (professional experiences, preparing for an internship, project management, pitching a project, interviews, labour law). The pedagogical scenario for each of the six sessions was designed by two of the trainees and formed part of the circular learning design of their online teaching module (Figure 4).
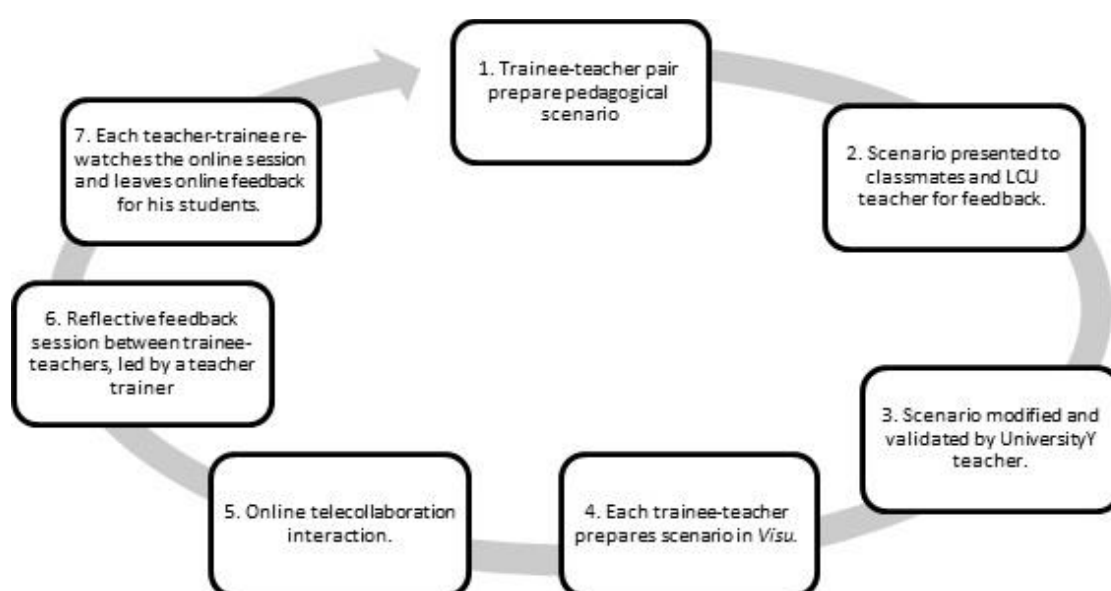

Figure 4. Circular learning design of the UL2 module 'Online teaching'

The webconferencing platform was used for steps four, five and seven of the process modelled in Figure 4. The platform offers three spaces: an area for preparing educational resources (step 4), a multimodal webconferencing room for synchronous interactions between participants (step 5) and a retrospection room where interactions can be replayed for training purposes (step 7). The interaction data focussed upon in this study comes primarily from the webconferencing-based interactions of step five.

A research protocol was designed by the researchers involved in the telecollaboration project and was approved by Dublin City University's research ethics committee. Participation in the study was voluntary: not all students gave permission to use their data: all trainees (ten females, two males) and 12 of the 18 students (eight females, four males) gave permission. To preserve anonymity, all participant names have been changed. Groups were formed randomly by the two course lecturers with no consideration of gender. Because of the uneven number on both sides, this resulted in five of the seven trainees working with two learners whilst the other two trainees had one-to-one sessions. Each

participant had their own computer and used headsets in order to ensure the best communication. Trainee teachers and students were in language labs during the exchanges. Students who were working in the same triad were, wherever possible, physically separated in the language lab

3.2 Data collection

To study the framing choices at play when teacher trainees interact with their distant learners and the impact they had on the interactions, three complementary data sets were used:

- **(1) the *hors-champ* and webcam films of three trainees** (Pam, Victor and Samia , see for example Figure 2) who were engaged in an interaction. Our data sample is limited to three sessions for technical and practical reasons: it was unreasonable to imagine that for each session we could position a video camera on the hors-champ environment of each participant as the size of the classroom did not allow us to do this and we did not have sufficient access to video cameras. We also felt that doing so would disturb the interaction: the research aspect of the project would become obtrusive. The three trainee-teachers for which a *hors-champ* recording was made were therefore volunteers. All webcam recordings were recorded and stored automatically by the *Visu* software.

- **(2) screen shot images**: These were taken from each of the twelve trainees' webcam videos to examine whether transformations exist in the trainees' webcam positioning over the six weeks of interaction. Around minute 17, a screen shot of the dominant framing choice was taken. Minute 17 was chosen because the sessions were underway and any initial technical issues, occurring most frequently during the opening phases of sessions, had been resolved.

- **(3) extracts of trainee feedback and learner interviews and** in an attempt "to "zoom in" on fine-grained detail and pan out to gain a broader, socially and culturally, situated perspective" as advocated by Flewitt, Hampel, Hauck & Lancaster, 2009: 44). With regards to trainee feedback, transcriptions of the reflective feedback sessions (step 7, Figure 4) were utilized in order to examine the importance the trainees grant to framing and gestures. Learner interviews comprised recordings of post-course interviews conducted face to face with the DCU students. The following questions pertaining to this study were asked:

- To what extent / when were you looking at your tutor's webcam image?
- What were you looking at more specifically?
- What did the webcam image bring to the conversation with your tutor?
- Did seeing your tutor sometimes bother you?
- During the sessions, were you aware of the image that you communicated of yourself through the webcam?

3.3 Data transcription and annotation

To address our first research question concerning how trainees positioned themselves in front of the webcam, we classified each screen shot image taken of the 12 trainees' framing at minute 17. Table 1 shows the four different categories used to classify the trainees' framing choices.

Framing category	Explanation
extreme close-up shot	The trainee's face saturates the frame but is only partially visible. At least one crucial element is missing (eyes, mouth or eyebrows are not visible).
close-up shot	The webcam is focused on the participant's face.
head and shoulders shot	The trainee's face and shoulders can be seen.
head and torso shot	The trainee's face, shoulders and torso can be seen.

Table 1. Classification of screen shots

With regards to our second research question concerning the functions of gesture that are visible or invisible in the frame, we devised a methodological framework to transcribe the webcam recordings and the videos of the hors-champ environments using ELAN (Sloetjes & Wittenburg, 2008). Building upon our previous work to understand multimodal communication structures (Develotte, Guichon & Vincent 2010; Wigham & Chanier 2013), we transcribed acts in both the verbal, co-verbal and non-verbal modes (see Table 2). To annotate communicative gestures in the kinesics modality, we used McNeill's (1992) schema that categorises gestures as iconic (representations of an action or object), metaphoric (illustrating an abstract concept), deictic (pointing gestures at concrete or abstract spaces) and beats (movements to accompany the rhythm of the discourse). To this schema we add the category of emblems (Kendon, 1982) referring to culturally specific gestures. We also annotated actions: communicative actions that impact on the communication, for example, writing something down or typing, and extra-communicative actions, for example scratching forehead, 'playing' with pen, as well as movements of the webcam or the computer screen. The latter allowed us to see when trainees explicitly changed their framing choices. Although text chat is not the focus of this article, it was included in transcriptions to facilitate future analyses on the interactions between different modalities in the verbal mode.

Communication mode	Communication modality	Act type	Explanation
Verbal	Audio	Audio act	Verbal act in the full duplex audio channel
		Silence	Interval between two audio acts greater than three seconds
	Text chat	Text chat act	Message entered into the text chat window
Co-verbal	Kinesics	Communicative gesture (*iconic, metaphoric, deictic, beat, emblem, communicative action*)	Gestures seen in the webcam or in the *hors-champ* recording.
Non-verbal	Proxemics	Movement	Participant moves webcam or computer as seen in the *hors-champ* recordings.
	Kinesics	extra-communicative gesture (*non-communicative action*)	For example, scratching forehead, pushing hair behind ear, 'playing' with pen.

Table 2. Classification of communication acts for transcription

The ELAN software was chosen as it is open-access and because it was specifically designed for the analysis of language, including sign language and, thus, is adapted to the transcription of non-verbal acts. Each horizontal line in ELAN is used to transcribe one modality that is used by a specific participant. Thus, ELAN gives researchers access to all modalities occurring at a given time and being used by different participants. The software allowed us to align the webcam and *hors-champ* videos of each session studied (see Figure 5).

Figure 5. Alignment of trainee *hors-champ* video with trainee and student webcam videos in ELAN.

For the trainees' feedback and the learners' interview data, we first conducted an exploration of all the data and then identified remarks and comments pertinent to our third research question that focuses on the extent to which trainees were aware of their framing choices and the perception of these by the learners.

3.4 Data coverage and annotation reliability

The screen capture data used for this study is comprised of the online interactions of involving three trainees. This data set totals 88 minutes and 40 seconds of interaction in which 304 audio acts (participants' verbal output) as well as 550 co-verbal and non-verbal acts were identified.

Ten minutes (11.4 per cent) of the total interaction data were annotated separately by two researchers with regards to gesture act types (*n*=101 gesture acts). The inter-rater reliability rate was 0.73 for the categorization of different gesture types. To calculate this rate, gesture type annotations agreed upon by both researchers were attributed the value of 1 and annotations that differed were attributed the value of 0. The researchers agreed on the annotations of 74 of the total 101 gesture acts. 74/101=0.73. This is high considering that there were seven possible annotation categories (see Table 2). Following initial coding of a sample of the data, the researchers met to reconcile any differences in categories before the other gesture acts were annotated.

4 Analyses and discussion

4.1 Framing choices

In relation to the first research question, we look at how the trainees positioned themselves in front of the webcam, using the screen shot images taken at minute 17 of each of the sessions. Since our data were captured in natural settings, absences and technical problems had an impact on data collection and out of 84 potential screen shots, only 66 could be retrieved. This moment was chosen because it is situated more or less in the middle of the session when all participants are fully engaged in the pedagogical activity and technical matters have been taken care of. The results that are obtained are organized along a continuum (see Figure 6) from extreme close-up shot, to close-up, to head and shoulder shot, to head and torso shot. The percentages that are provided are to be considered as simple indicators of frequency for each category but no statistical tests have been run.

Figure 6. Continuum of framing choices at minute 17 of the interaction

Extreme close-up shots are quite rare in the data and were mostly produced by the same trainee across all six sessions. Probably due to a lack of attention to the positioning of the webcam (and to her *contre-champ* image), it zooms in on the upper part of the trainee's head leaving out her eyes, mouth and gestures. All the expressive and affective cues were thus invisible and might have hindered the interaction with the learner as an image was produced that was devoid of empathy.

Close-up shots make up approximately one third of the whole. They give close access to the trainees' eyes and mouths and thus to their facial expressions. It can be hypothesized that such shots reduce the feeling of distance between interlocutors and facilitate the learner's attention to their teacher's voice. Two out of the four learners who were interviewed noted the fact they concentrated on their teacher's lips, which were perceived as crucial to understand certain characteristics of French pronunciation. Thus Aiden said: "when she uses words with /r/ I look to see if she's doing something special with her mouth". Since close-up shots saturate the frame and focus on the faces, they leave out many of trainee teachers' gestures but make their lip movements more visible to students.

Head and shoulder shots are favoured by a lot of the trainees (44% of the whole): thanks to this framing, the learner can see the teacher's entire face but can also glimpse certain expressive gestures when they are made visible by the teacher.

Head and torso shots can only be produced when the teachers sit back in their chairs, keeping the webcam at a certain distance and thus allowing the production of gestures to be quite visible. The communication space is thus larger and allows teachers to be more than 'talking heads' by revealing the way they are dressed and some of the context in which the interaction is taking place.

A close examination of the data reveals that a certain type of framing seems to be chosen by each trainee and that this choice evolves only marginally during the six-week telecollaboration project. This may indicate a difficulty to adjust one's posture to new teaching circumstances involving dealing with many other elements simultaneously. If we exclude extreme close-ups that do not seem to be semiotically adapted to this type of communication, framing seems to provide a trade-off between the visibility of the face and the visibility of the gestures, the first being maximized by close-up shots and probably facilitating the teacher's endeavour to show empathy while the second requires head and torso shots and foster the teacher's bodily expressivity (maybe to the expense of producing a certain distance with the learner). It is difficult to say which of the framing choices is most appropriate for webcam-mediated language teaching. Indeed, it can be hypothesized that framing may be usefully adapted by the teacher, interaction after interaction or moment by moment, according to the task at hand (fine-grained comprehension and focus on pronunciation might better be transmitted by close-up shots while some collaboration tasks might require head and torso shots), the familiarity of the interlocutors (getting literally closer with time but needing space at first), the learners' culture (proxemics being culturally significant as Hall showed back in 1966) or pedagogical intentions. This leads us to advance that trainees have to become critically aware of the semiotic effect each type of framing can have on the pedagogical interaction so that they make informed choices to monitor the image they transmit to their distant learners according to an array of professional preoccupations that are of affective, cognitive, and expressive orders.

4.2 Relationship between framing choices and gestural space

The framing choices of the trainees influenced whether a shared gestural space was established, or not, between the trainee and his/her students. In head and shoulders shots and head and torso shots (see Figure 5), the gestural space becomes shared and both the trainees' gestures, in addition to their facial mimics, can be seen by students, thus providing a common "site of display" (see Jones, 2009:115). The students attributed importance to the possibility of sharing this space, describing that it made them feel more at ease during the pedagogical interaction, helped to clarify potential misunderstandings and also increased their level of concentration, as Examples 1 & 2 demonstrate.

> **Example 1** (Catriona): "it felt more comfortable to see who you were talking to … we could see what was happening if she was laughing"
>
> **Example 2** (Catriona) "she tried to use her hands a lot to explain things …it's more attractive… you listen more to what she says"

The trainees' framing choice also had pragmatic implications. In Example 3, one student describes that she had said something in Spanish, rather than in the target language French, and that if she had not been able to see the pouting mimic of her teacher she would not have been able to interpret that she was joking about being angry at the use of another language.

> **Example 3** (Ana) "I said something in Spanish she pouted. It was a joke. If I hadn't seen her I would have thought she was angry"

In this example, the trainee exploits the characteristics of head and shoulders' framing to shape the interaction in a positive way by developing the social relationship between herself and her student through her co-verbal behaviour.

Examples 1-3 demonstrate that the content of the webcam image has an impact on students' appreciation and comprehension of the online situation: this reinforces the importance of making online teachers semiotically aware of the image they project of themselves to their students.

4.3 Visibility and communicational functions of gestures

The framing choices of the trainees also influenced whether gestures used in the interaction were visible to their students.

In Figure 7, image A, the emblem gesture to encourage the students is clearly directed at the webcam for the students' benefit and therefore is visible in both the hors-champ and the webcam views. In image B, however, although the circular gesture, associating the speaker and the interlocutor that accompanies the phrase 'our theme today' is performed by Pam, it is barely visible in the webcam view for the students who indeed would not be able to interpret its meaning from the webcam view.

A.
Emblem - Both hands, thumbs up.

"nickel"
(*great*)

B.
Metaphoric - Both hands, circular rotating movement.

"notre thème aujourd'hui"
(*our theme today*)

Hors-champ view   Webcam view   Transcription

Figure 7. Visibility of gestures in *hors-champ* and webcam views

Figure 8 shows, for the three sessions examined, the total number of gestures performed by the three trainees (seen in the *hors-champ* view on the left of Figure 8). When we examined whether each gesture was visible, barely visible or not at all visible in the webcam view, our data show a significant loss of information between the gestures performed (shown in cross-hatching) and those that were clearly visible to the students (shown in grey).
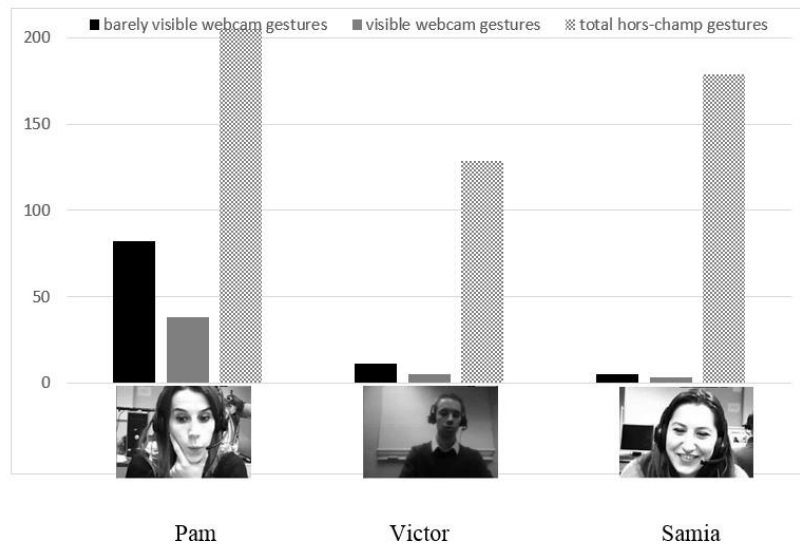

Figure 8. Gestures in and out of the frame for the three trainees

Indeed, for all three trainee-tutors examined there is a distinct loss between the numbers of communicative gestures performed (emblems, metaphorics, iconics, deictics and beats) and those visible for their interlocutors in the webcam view. For Pam, 18 of her communicative gestures (*n*=129) are visible in the webcam view, for Victor three of his communicative gestures (*n*=11) are seen in the webcam and for Samia only four of her total communicative gesture acts (*n*=74) are visible to her students. It should be noted that the participant Samia had technical difficulties during the session and that for short intervals during the session her webcam was not connected. This could partially explain the difference between the number of gestures produced and those that were visible. In contrast with Pam and Samia, Victor uses significantly fewer communicative gestures (*n*=11) than extra-communicative actions (*n*=61).

To illustrate this, in the *hors-champ* shot in Figure 9, Samia's gesture to accompany the phrase 'the last two weeks' clearly duplicates what is being said in the audio modality and therefore may have been performed to aid students' comprehension (see Tellier's (2006) "pedagogical gesture"). However, in the webcam shot this information is not conveyed maybe because it was a self-regulatory gesture made by the trainee to help her with her own speech. There has been a missed opportunity for the trainee to coordinate her actions in the two semiotic modalities, the audio and the kinesics, in order to aid comprehension for her students.
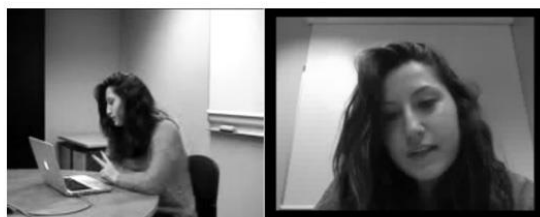
Left hand index and middle finger extended.

"ces <u>deux dernières semaines</u> t'as fait quoi ?"

(*what did you do the <u>past two weeks?</u>*)

Figure 9. Missed opportunity for communicative gestures

In contrast, when trainees coordinate the two semiotic modalities, their communicative gestures may become valuable for mutual comprehension. Figure 10 illustrates this. Pam is explaining the difference between the vocabulary items *salarié (*employee) and *bénévole* (volunteer). The trainee looks directly at the screen and communicates in the kinesic modality to her students with the first deictic gesture 'you' which is visible in the webcam view. Although both trainees in these examples make the same framing choice of a head and shoulders shot, Pam manages to better frame her gestures to aid the interaction. She shows she is aware, through her framing choices and use of a shared gestural space, that she is communicating using both the kinesic and audio modalities. In this gesture sequence, Pam's gestures are held long enough within the framed image, so that these particular gestures can be perceived by the students. Her attention to both framing and purposeful gesturing testify that she is developing her critical semiotic awareness.

Following her first deictic gesture, Pam then illustrates the concept of earning money. Using an emblem that is culturally specific, she rubs her thumb and her index upward as the French do to mean "make money". Here, we can see that the choice to frame gestures has implications from a cultural perspective even if the meaning of this gesture might be lost on the Irish learners would do not share the same emblem. The teacher then uses a self-deictic 'I'm a volunteer' before insisting on the difference with a gesture that accompanies 'I' but that is not seen by the students. Mutual understanding is then aimed for with a final gesture that accompanies 'I don't earn money'. This abstract deictic gesture moves backwards pointing back to a shared communication space between the teacher and her students earlier in the explanation. In this example, we clearly see that the teacher is aware of the modalities in which she is communicating and uses them to try to achieve mutual understanding.

Figure 10. Coordinated acts in the audio and kinesic modalities

Our transcriptions of the reflective feedback sessions illustrate that the extent to which gestures are visible may be influenced by framing choices but also, considering Pam and Samia both of whom used a head and shoulders framing shot, by the extent to which the trainees are semiotically aware and feel able to devise appropriate communication strategies (Examples 4 and 5).

> **Example 4** (Pam) "c'est vrai que je trouve qu'il y a vraiment du langage qui passe par le corps moi quand je m'exprime j'ai souvent j'ai besoin de bouger faire des choses avec mes mains... je me regardais quand même, et je me rendais compte que je disais dis-le et je faisais ça ça et des trucs comme ça et je disais ah oui... je me rendais compte qu'on était là vraiment pour regarder ça et du coup beh oui je le faisais car effectivement c'est de se regarder"
>
> (It's true that I find that I really use body language when I'm talking I often need to move to do things with my hands...I looked at myself all the same and I realised that I say 'say it' and I did this and this and things like this and I kept saying 'oh yes'...I realised that we were there really to look at ourselves and so I did because indeed it's about looking at yourself) (Our translation)
>
> **Example 5** (Samia) "j'ai l'impression qu'il y a un gros mur là face à moi et je suis limite obligée de passer fin pour me faire comprendre et c'est vraiment c'est c'est ce que je ressens...d'avoir un gros mur face à moi"
>
> (I was under the impression that there was a huge wall in front of me and that I more or less had to get over it to be understood and it's really it's what I felt...like I had a huge wall in front of me) (Our translation)

Opening up the shared gesture space, through portrait framing choices, means that some extra-communicative actions come into view (see Figure 11). Extra-communicative actions have little impact on the communication for they do not possess any semiotic meaning. These are actions that are for the trainees themselves (self-centred) rather than for their students.

To illustrate this, our minute 17 screen shot data showed two predominant types of extra-communicative actions. Firstly, the trainees used body-focused adaptors such as scratching a part of their head, touching their face and biting their lips (See Figure 11). These actions allow the trainees to maintain a sense of coherence for themselves, helping them in their thinking processes (Codreanu & Combe Celik, 2013). Whilst these actions are performed for the trainees themselves, they do not go unseen during the interaction as Example 6 illustrates.
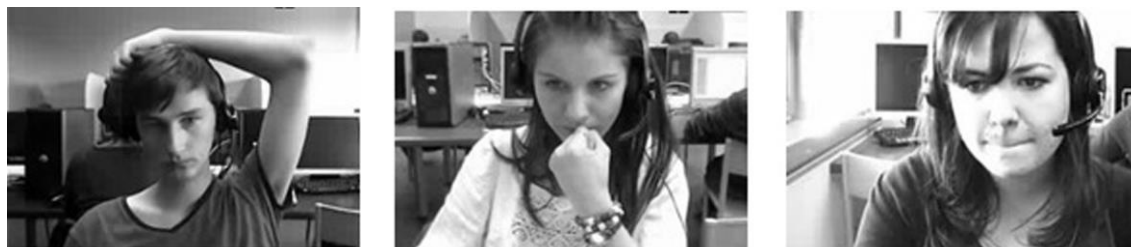

Figure 11. 'Thinking' extra-communicative actions

Example 6 (Siobhan, interview data) "it was really easy to see when she *[the trainee]* was in difficulty, for example when she asked a question and *[my classmate in Dublin]* didn't reply and I was thinking you could see in her face, ARGH what do I do now, what do I say, when she made her 'I-don't-know-what-to-do face' I tried to think more quickly".

Secondly, our data illustrate examples of potentially distracting extra-communicative actions (See Figure 12). These gestures are, like the adaptors, self-centred. They could cause interference as they are potentially distracting and detract from the message. We can wonder whether the trainees either forget about the presence of the webcam and thus these gestures yield some kind of naturalness to the interaction, or whether the opposite is true and the participants are too focused on their own images causing them to readjust their hair etc.

Figure 12. Potentially distracting extra-communicative actions in both webcam and *hors-champ* views

5 Conclusion

The analysis of our data has helped us uncover what critical semiotic awareness means when teachers are faced with the use of a tool that reveals only partially the communication space they are involved in. The *champ* gives access to a limited site of display and it is up to the online teacher to enrich it with cultural and socio-affective cues so as to increase its potential for communication in the target language. Not only is framing essential to display online teachers' gaze, mimics and mouth to render the interaction smoother and more enjoyable, but it requires that they make sure they produce, ostentatiously, communicative gestures that are within the frame and held long enough to be perceived by the learners. Indeed, student interviews revealed that there was an impact of the visibility of mimics and gestures on the perceived quality of the pedagogical interaction, even if our data do not make it possible to determine if the impact was on a socio-affective level only or also for language learning.

As long as the webcam gives access to the teacher's face, there does not seem to be an optimal framing choice for desktop videoconferencing among the three framing types that we identified. As seen in Section 4.3, two trainees (Pam and Samia) used the same framing choices but one used gestures to help move the interaction forward whilst the other did not use the site of display as richly. Yet, we saw that most gestures are invisible to the webcam but if a particular iconic or metaphoric gesture is to be used to help comprehension, online teachers should be careful to display it so that it is visible in the webcam field. This also highlights the importance of communicative facial expressions and mimics which really are visible in this type of interaction.

Thus, choosing either a close-up shot, a head-and-shoulder shot or a head-and-torso shot in order to frame oneself is a decision that can be made timely by the teacher depending on perceived learning needs, pedagogical intentions, task types, familiarity with the learner and intercultural considerations. If the image produced by the webcam is not to be flat, intrusive or devoid of meaning, the online teachers should develop appropriate strategies to actualise the potential of their projected image and thus give more relief to what Guichon and Cohen (2014), also referring to webconferencing mediated pedagogical situations, called "online teacher presence"(see supra). Besides, attention to framing demonstrates online teachers' attentiveness to their learners' cognitive and socio-affective needs and may help to contribute to what Kern (2014: 341) terms "relational pedagogy".

Critical semiotic awareness may be enhanced if teachers use their *contre-champ* image, as some kind of rear-view window to the interaction, so as to get a better perception of the image they project of themselves and, when needs be, adjust it by reframing themselves, getting closer or further from the webcam or checking that their communicative intentions are not lost by being too furtive or invisible. It can also be a way to reduce extra-communicative gestures that are not necessarily distracting but whose repetition could detract from the message by being not congruent with what is being said. It can be hypothesized that as online teachers become attuned to this situation, they will need only occasional glances at their *contre-champ* image at critical moments (openings, changes of tasks, cases of misunderstandings etc.).

Developing critical semiotic awareness is thus a matter of learning to adjust one's communication to the constraints of a technology not "to recover the information 'lost in translation' from off-line to online communication, but instead to augment and improve interaction" (Kappas & Krämer, 2011: 9). Mediated pedagogical interactions have long been evaluated as being impoverished and lacking in relief. Now that the technical aspects are

improving, we believe that teacher training, that is sensitive to the semiotic elements of online teaching as well as other important aspects of online pedagogy (see Guichon, 2009; Hampel & Stickler, 2012), is needed to help teachers make the most of original sites of display without feeling that they are losing something in the process.

Finally, let us not forget that the webcam frame is embedded in the bigger frame of the computer window, an interface which provides the learner with a complex semiotic ensemble. Two questions remain unanswered: (1) how does the content within the webcam frame mesh with the rest of the content included in the frame of the computer ; in other words, how does this specific semiotic resource connect with other dynamic (chat) and static (layout) semiotic resources to make a coherent whole?; (2) how is the activity in the webcam frame *perceived* by the learners on a cognitive rather than a symbolic level: what attentional resources are allocated to the webcam frame and to the rest of the computer window and with what changes in intensity? For these two questions, data collected in an ecological situation may be irrelevant, and punctual experiments using eye-tracking might provide insight to complete the present study.

References

Bétrancourt, M., Guichon, N. and Prié, Y. (2011) Assessing the use of a Trace-Based Synchronous Tool for distant language tutoring. *Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning*, Hong-Kong, July 2011. pp. 478-485

Camshare Inc. (2014). *Camfrog* [software] [http://www.camfrog.com].

Clark, H.H., and Brennan, S.E. (1991), Grounding in communication. In Resnick, L. B.; Levine, J. M.; Teasley, J. S. D. (Eds.). *Perspectives on socially shared cognition*, Washington DC: American Psychological Association. pp. 127-149.

Codreanu, T. and Combe Celik, C. (2013). Effects of webcams on multimodal interactive learning. *ReCALL*, 25, pp 30-47.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Deleuze, G. (1983). *L'image-mouvement*. Paris: Editions de Minuit.

Develotte, C., Guichon, N., and Vincent, C. (2010) The use of the webcam for teaching a foreign language in a desktop videoconferencing environment, *ReCALL* (23 (3), pp. 293-312.

Flewitt, R., Hampel, R., Hauck, M., and Lancaster L. (2009). What are multimodal transcription and data? In Jewitt, C. (Ed.). *The Routledge handbook of multimodal analysis*. London: Routledge. pp. 40-53.

Goffman, E. (1974) *Frame Analysis: An essay on the organization of experience*. New York: Harper and Row.

Guichon, N. (2009). Training future language teachers to develop online tutors' competence through reflective analysis. *ReCALL*, 21(2), 30–49.

Guichon, N. and Cohen, C. (2014). The impact of the webcam on an online L2 interaction. *Canadian Modern Language Review,* 70(3), pp. 331–354.

Hall, E. T. (1966). *The Hidden Dimension*. Garden City, N.Y.: Doubleday.

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

Jewitt, C. (ed.) (2009). *The Routledge handbook of multimodal analysis*. London: Routledge.

Jewitt, C. (2011). The changing pedagogic landscape of subject English in UK classrooms. In K.L. O'Halloran & B.A. Smith (Eds.), *Multimodal studies: Exploring issues and domains*. New York: Routledge. pp. 184-201.

Jones, R. H. (2004). The problem of context in computer-mediated communication. In Levine, P. & Scollon, R. (Eds.). *Discourse and technology – Multimodal discourse analysis*. Washington DC: Georgetown University Press. pp. 20-33.

Jones, R. H. (2009). Technology and sites of display. In C. Jewitt (ed.). *The Routledge handbook of multimodal analysis*. London: Routledge. pp. 114-126.

Kappas, A. and Krämer, N.C. (Eds.). (2011). *Face-to-face communication over the Internet*. Cambridge: Cambridge University Press.

Kendon, A. (1982). The study of gesture: some observations on its history, *Recherches Semiotique/Semiotic Inquiry*, 2(1). pp. 25-62.

Kern, R. (2014). Technology as Pharmakon: The Promise and Perils of the Internet for Foreign Language Education. *The Modern Language Journal*, 98, 1, pp. 340–357.

Kress, G. (2009). What is mode? In C. Jewitt (ed.). *The Routledge handbook of multimodal analysis*. London: Routledge. pp. 54-67.

Manstead, A. S. R., Lea, M., and Goh, J. (2011). Facing the future: Emotion communication and the presence of others in the age of video-mediated communication. In A. Kappa, & N. C. Krämer (Eds.), *Face-to-face communication over the Internet: Emotions in a web of culture, language, and technology*. Cambridge: Cambridge University. pp.17-38.

McNeill, D. (1992). *Hand and Mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.

Parkinson, B. and Lea, M. (2011). Video-linking emotions. In Kappas, A. & Krämer, (Eds.). *Face-to-face communication over the Internet*. Cambridge: Cambridge University press. pp. 100-126.

Pinnow, R. J. (2011). "I've got an idea": A social semiotic perspective on agency in the second language classroom. *Linguistics and Education* 22, pp. 383–392.

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

Sindoni, M. G. (2013). *Spoken and written discourse in online interactions*. New York: Routledge.

Sloetjes, H. & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*.

Souchier, E., Jeanneret, Y., & J. Le Marec (Eds.), (2003). *Lire, écrire, récrire*, Paris: Bibliothèque Centre Pompidou.

Tellier, M. (2006). L'impact du geste pédagogique sur l'enseignement/apprentissage des langues étrangères : Etude sur des enfants de 5 ans. PhD Thesis. [http://tel.archives-ouvertes.fr/tel-00371041]

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

Wigham, C.R. & Chanier, T. (2013). A study of verbal and non-verbal communication in Second Life - the ARCHI21 experience. *ReCALL* 25(1), pp. 63-84.

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

Zähner, C., A. Fauverge, & J. Wong (2000). Task-based Language Learning via Audiovisual Networks. In M. Warschauer & R. Kern (eds), *Network-based Language Teaching: Concepts and Practice*. Cambridge: Cambridge University Press. pp. 186-203.