AUDIOCONFERENCING (AC) AND VIDEOCONFERENCING (VC) FOR LANGUAGE TEACHING AND LEARNING*

Nayara Nunes Salbego
Universidade Federal de Santa Catarina (UFSC)
Instituto Federal de Santa Catarina (IFSC)

Abstract: This article addresses arguments in relation to using audio or videoconferencing for online synchronous language classes. Audioconferencing (AC) differs from Videoconferencing (VC) because of the utilization of the webcam. In the former, communication takes place through audio channels, while the latter complies images of the participants through synchronous video chat. In regard to that, online synchronous interaction through AC and VC may affect learning in different ways. Through research results about this issue, this article presents possible pros and cons for the use of the webcam. Results point to benefits of AC, not eliminating, however, the importance of the webcam if used with awareness.

Key-words: audioconferencing, videoconferencing, language learning.

1. Introduction

Synchronous online multimodal communication through Audioconferencing (AC) and Videoconferencing (VC) have become popular ways of communication in everyday life through free resources such as *Skype*, *FaceTime* and *Google Hangouts*. Likewise, in learning contexts, especially in online language learning and teaching, learners can communicate with others orally via video and/or audio calls as a way of interacting in the target language. On these grounds, researchers have been demonstrating interest in understanding how synchronous online communication happens for the purpose of teaching and learning languages. Results have shown advantages and disadvantages regarding the use of AC and VC for language pedagogy. In spite of that, teachers and students are the only ones eligible to decide which modality of online interaction, AC or VC, works better for them. Hence, this article presents research results unveiling pros and cons of AC and VC for language teaching and learning to argue that we have not one answer, but many, regarding the use of AC and/or VC. It ends with the final argument and conclusion that the decision relies on learners and teachers to identify how they profit from each modality of communication.

2. Audioconferencing (AC) and Videoconferencing (VC) for Language Learning

Audioconferencing (AC) and Videoconferencing (VC) can be included in the concept of webconferencing, since the latter is a broader term. Cruz and Barcia (2000) conceptualize webconferencing as a digital resource that allows oral, visual and written communication for people who are physically distant. While AC limits its affordances to the oral channel, VC combines visual and oral features to the interaction (Cruz, 1997; Cruz & Moraes, 1997). Both AC and VC, many times, can make use of written communication resources through a chat box, besides graphics such as emoticons.

There has been a number of studies investigating whether AC and/or VC, used individually or in combination, could be beneficial for language learning and teaching (e.g. JEPSON, 2005; STICKLER, BATSTONE & DUENSING, 2007; HEINS, DUENSING, STICKLER & BATSTONE, 2007; YANGUAS, 2010; HAMPEL & STICKLER, 2012; CODREANU & CELIK, 2011; SATAR, 2015; GUICHON & COHEN, 2015; GUICHON, & WIGHAM, 2016). For the purpose of this essay, I will focus on the most recent research results that specifically contrast AC and VC (Yanguas, 2010; Guichon & Cohen, 2015) or focus on analyzing the use of the webcam for language teaching and learning (Guichon, & Wigham, 2016).

Yanguas (2010) conducted a qualitative study to examine how language learners negotiated meaning during task-based interaction via AC, VC and face-to-face (F2F). Learner-to-learner dyads were randomly assigned to (a) an AC group, (b) a VC group, or (c) a F2F control group to complete a jigsaw task that included unknown lexical items. Yanguas (2010) concluded that in the AC condition, learners had to use linguistic resources they would not have used in the two conditions that provided visual cues (F2F and videoconferencing). However, the linguistic resources used in the AC modality presented a higher percentage of partially understood lexical items. Yanguas (2010, p. 84) posits a question regarding this issue:

> The question is then what is more important for L2 acquisition, to linguistically elaborate on the item augmenting the possibility of non-communication to occur, or to fully and unequivocally understanding the concept being focused on as a source of non-communication? In an interactionist context such as ours, it seems that the most logical answer would be the former rather than the latter but more research is needed to properly address this question.

In this case, participants were mainly trying to unveil the meaning of unknown lexical items. In the VC and F2F groups, participants could mimic instead of speaking, while in the AC group, participants resorted to elaborating verbally on the lexical items, although sometimes they did not fully understood each other.

As the author points out along the discussion of the results, both AC and VC can affect language teaching and learning positively. While AC may push learners to production, VC can foster complete understanding due to the visual resources (YANGUAS, 2010). No significant differences were identified between the F2F and VC conditions.

In the same lines, Guichon and Cohen (2015) studied if visual cues should enhance online communication in relation to social presence[1]. Through an experimental study, they explored the value provided by a webcam in an online L2 pedagogical teacher-to-learner interaction. A total of 40 French undergraduate students were asked to describe in English four previously unseen pictures to a teacher via *Skype*, during a 10-minute interaction. Half of the students (n = 20) were assigned to a VC condition and the other half (n = 20) to AC. The author aimed to explore whether seeing the interlocutor on the webcam impacts on the patterns of interaction and on word search episodes.

Among the results, Guichon and Cohen (2015) found that there were no differences between the students' feelings of online teacher presence in the two conditions (AC and VC). Another interesting result was that students in the AC condition understood the teacher better, in comparison to the VC group, contrarily to Yanguas' (2010) findings. According to Guichon and Cohen (2015), this result goes against the idea that "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" (p. 349). In agreement with Yanguas (2010), though, students in the AC condition produced more target language, than the ones in the VC group. As stated by the participant teacher in Guichon and Cohen's (2015, p. 349) investigation, "the lack of images helped students to focus on the words and their meaning, so maybe this obliged them to concentrate more".

As the authors put it:

> [W]e 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. (GUICHON; COHEN, 2015, p. 349)

Guichon and Cohen's (2015) research results point to the idea that the use of the webcam is not essential for language learning and teaching. Other differences regarding AC versus VC were the number of silences and overlaps[2]. There were more silences in the AC, possibly explained by the lack of visual clues, and more overlaps in the VC.

The authors also identified that the teacher had to speak more abundantly in the AC condition. The teacher herself stated that she had to provide more information or assistance to students in order to move the interaction forward due to the lack of visual cues. On the other hand, in the VC group the conversation forwarded more rapidly and seamlessly

---

[1] In short, Social Presence (SP) can be defined as the 'degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationship' (WILLIAMS AND CHRISTIE, 1976, p. 65, as cited in SATAR, 2015, p. 481).

[2] Silences and overlaps are constructs better understood according to Conversation Analysis. Silences may refer to moments in which talk is expected, but does not happen, and overlaps may be considered evidence of an incoming speaker's failure to notice whether the current speaker has or has not finished (HUTCHBY & WOOFFITT, 2002).

because of the linguistic cues. In relation to the word search episodes, no significant differences were found between the groups.

In a very recent study carried out by Guichon and Wigham (2016), the results also emphasize that the use of the webcam in VC is not essential. The authors claim that the limited access to visual cues provided by the webcam is felt as useless or even disruptive. The authors depart from the point that students only visualize the teacher's face and, sometimes, shoulders and torso, and that may not help learning when compared to AC. The data for the study consisted of multimodal transcriptions of online interactions, screen shot data and post-course interviews.

Through an analysis of shots, Guichon and Wigham (2016) tried to identify what would be the best position for teachers to stand in front of the webcam when teaching. The types of shots can be seen in **Table 1**.

Table 1 *Classification of screen shots* (GUICHON; WIGHAM, 2016, p. 70).

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.

The results found by Guichon and Wigham (2016) showed that there is no optimal framing choice for VC. Furthermore, there was a loss between the number of gestures performed by the teachers and those visible for the students, meaning that the teachers were not careful in relation to making gestures that would show in the camera. Consequently, the study highlights the need for teachers to develop critical semiotic awareness in order to have a better perception of the image they project in VC for language teaching and learning.

Although the studies presented tend to emphasize that AC works for language learning and teaching, putting aside the necessity of having visual cues through webcams (VC), the overall discussions and results allow room for the interpretation that both conditions, AC and VC, are affordable of language development. While AC fosters language production, interaction forwards more fluently in VC due to smiles, nods and gestures.

Considering that, teachers and students are the only ones eligible to decide which modality of online interaction, AC or VC, works better for them. Despite the cons in relation to VC shown throughout research results, it is evident that some teachers and learners may profit from the visual cues, be they a few or numerous. According to Dörnyei and Skehan (2003), individual differences play a role when learners have the task of deciding in which type of interaction to engage in order to foster language learning.

In the same way, teachers when planning may pay attention whether or not it is necessary or beneficial to use the webcam. According to the activities planned for class, teachers can opt for AC or VC, even making use of the resources individually or in

combination in specific moments of each class. As Blake (2008) and Thomas, Reinders and Warschauer (2014) assure, technological resources do not change education, being that they are inherently neutral. However, the use of such resources by educators may have a positive impact on language pedagogy.

3. Conclusion

To conclude, this essay presented research results regarding AC and VC, emphasizing that there are no conclusive results in the sense of stating which modality is more effective for teaching and learning languages. Through the presentation of empirical studies, I addressed how AC and VC can be effectively used for L2 instruction. Besides that, the findings indicated important aspects to be considered when using AC and VC, individually or in combination.

Although most evidences refer to more negative aspects of using the webcam for synchronous online language teaching and learning (YANGUAS, 2010; GUICHON; COHEN, 2015; GUICHON; WIGHAM, 2016), I argue that teachers and students are the only ones eligible to decide which modality of online interaction, AC or VC, works better for them. Individual differences and careful planning allows students and teachers, respectively, to decide upon which mode to choose, bearing in mind that technology cannot change education alone, it simply provides potential affordances for well-aware pedagogical rationales.

References:

BLAKE, R. J. *Brave New Digital Classroom*: Technology and Foreign Language Learning. Washington, D.C.: Georgetown University Press, 2008.

CODREANU, T. & CELIK, C. C. An in-depth Exploration of the Effects of the Webcam on Multimodal Interactive Learning. The EUROCALL Review. *Proceedings of the EUROCALL 2011 Conference*. 2012, Vol. 20, pp. 45-48.

CRUZ, D. M.; BARCIA, R. M. Educação a Distância por Videoconferência. In Cruz, D. M., Barcia, R. M. (eds.). *Tecnologia Educacional*, ano XXVIII, 2000, n. 150/151, v. 29, julho/dezembro, p. 3-10.

CRUZ, D. M.; MORAES, M. *Manual de Videoconferência*. Florianópolis: Laboratório de Ensino a Distância : UFSC, 1997.

CRUZ, D. M. Videoconferência e Educação à Distância: a comunicação a serviço da interatividade. *INTERCOM* Revista Brasileira de Comunicação, São Paulo, Volume XX, nº. 2, páginas 139-142, jul./dez/ 1997.

DÖRNYEI, Z.; SKEHAN, P. Individual Differences in Second Language Learning. In C. J. Doughty, & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 589-630). Oxford: Blackwell, 2003.

GUICHON, N.; WIGHAM, C.R. A Semiotic Perspective on Webconferencing Supported Language Teaching. *ReCALL*, 28(1), 2016, pp.62- 82.

GUICHON, N.; COHEN, C. The Impact of the Webcam on an Online L2 Interaction. *The Canadian Modern Language Review*, 70, 3, August, 2015, pp. 331–354.

JEPSON, K. Conversations – and Negotiated Interaction – in Text and Voice Chat Rooms. *Language Learning and Technology*, 2005, 9(3), pp. 79–98.

HAMPEL, R.; STICKLER, U. The use of Videoconferencing to Support Multimodal Interaction in an Online Classroom. *ReCALL*, 2012, Vol. 24 (2), 116-137.

HEINS, B.; DUENSING, A.; STICKLER, U.; BATSTONE, C. Spoken Interaction in Online and f2f Language Tutorials. *CALL* (Computer Assisted Language Learning), 2007, 20(3), pp. 279–295.

HUTCHBY, I.; WOOFFITT, R. *Conversation Analysis*. Principles, Practice and Applications. Cambridge, UK: Polity Press, 2002.

SATAR, H. M. Sustaining Multimodal Language Learner Interactions Online. *CALICO Journal*, 2015, vol. 32 (3), pp. 480-507.

STICKLER, U.; BATSTONE, C.; DUENSING, A.; HEINS, B. Distant Classmates: speech and silence in online and telephone language tutorials. *EURODL* (European Journal of Open, Distance and E-Learning), 2007, (II).

THOMAS, M.; REINDERS, H.; WARSCHAUER, M. (eds.). *Contemporary Computer-Assisted Language Learning*. London: Bloomsbury, 2014.

YANGUAS, I. Oral Computer Mediated Interaction between L2 Learners: It's about time! *Language Learning & Technology,* October 2010, Volume 14, Number 3, pp. 72–93.