# The use of teaching gestures in an online multimodal environment: the case of incomprehension sequences

*Benjamin Holt[1], Marion Tellier[2], Nicolas Guichon[3]*

[1]Université Lumière Lyon 2, Labex ASLAN, Laboratoire ICAR and Aix-Marseille Université, CNRS, LPL UMR 7309, 13100, Aix-en-Provence, France
[2] Aix Marseille Université, CNRS, LPL UMR 7309, 13100, Aix-en-Provence, France
[3]Université Lumière Lyon 2, Laboratoire ICAR

## Abstract

This study aims to describe the gesture use and multimodal behavior of future French teachers in Lyon, France during videoconferenced pedagogical interaction with learners of French in Dublin, Ireland. These multimodal conversations were recorded, transcribed and annotated using an annotation scheme that was designed for this multimodal corpus. The use of gesture was measured and compared quantitatively during sequences of incomprehension and sequences in which there was no manifest incomprehension between participants to measure the extent to which gesture was used to repair incomprehension. Results show that contrary to expectations, teacher trainees did not use gestures to repair incomprehension.

**Index Terms**: exolingual communication, incomprehension, teaching gesture, gesture rate, computer-mediated communication, videoconferencing, multimodality

## 1. Introduction

A project called *français en première ligne* [1] started in 2002 that allowed distant learners to interact asynchronously with native French speakers. With increased bandwidth and the ubiquity of webcams, the project became synchronous in 2006 [2], which allowed interlocutors to see each other, including their hand gestures. It has already been shown that gestures play an important role in the teaching and learning of foreign languages, [3]–[6] but what happens when the pedagogical interaction takes place in a videoconferencing environment where the gesture space is drastically reduced [7]? This case study, devised to prepare a PhD, will focus on webcammed pedagogical interaction between master's students in France and undergraduate learners of French in Ireland. More specifically, the use of hand gestures during incomprehension sequences will be quantitatively and qualitatively measured.

## 2. Theoretical Framework

In foreign language acquisition, the understanding of input is of the utmost importance and is preliminary to production [8], [9]. Therefore we must take a look at moments that favor the modification of input. Such moments occur during *exolingual conversation*, which has been defined by Porquier [10] as interaction that takes place between interlocutors who do not share equal linguistic proficiency in the language that is being spoken, and who consciously make adaptations according to the disparity. Native speakers use *foreigner talk*, altering and simplifying the way they speak [9]. Porquier & Py [11] think that *exolingual competence* is probably learned and not innate, and Sarré [12] agrees that negotiation of meaning is an important interactional skill.

Input is modified most during negotiation of meaning and incomprehension sequences [11], [13] where the interlocutors strive to facilitate comprehension. As Sarré [12] points out, videoconferencing environments are propitious for negotiation of meaning, and for this study we will use the four-step model devised by Varonis & Gass [13]:

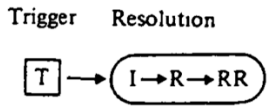



Figure 1: Incomprehension sequence model [13, pp. 74]

A *trigger* (T) is an utterance from the native speaker that is not understood by the non-native speaker, and is usually found at the beginning of the sequence [13]. The trigger can come from many things including lexical items, the context, the difficulty of the task, etc. [8], [14]. The trigger never exists out of context, and thus does not become a trigger unless it elicits a reaction from the non-native. Until this happens, it remains a "potential trigger" [13, pp. 78]. If the non-native gives feedback [15], this is considered to be the *indicator* (I), which "halts the horizontal progression of the conversation and begins the downward progression, having the effect of 'pushing down' the conversation rather than impelling it forward" [13, pp. 75]. If the native speaker does not ignore the indicator, s/he gives a *response* (R), which is the most important step because it reveals the negotiation strategies used by the teacher. Yanguas [14, pp. 82] explains that:

> "responses are perhaps the most vital element in the negotiation routines because, on the one hand, they include the feedback provided to the interlocutor to fix the communication problem and, on the other, they are pushed output on the part of the speaker."

Videoconferenced interactions have been shown to produce more negotiation sequences than other types of online conversation [12]. All of the reparation strategies enumerated by Long [9] are strictly verbal and prosodic, completely leaving out nonverbal means of repairing incomprehension, such as the use of coverbal gestures. Coverbal gestures, which have been defined as "the movements of the hands and arms that we see when people talk", [16, pp. 1] "offer themselves as a second channel of observation of the psychological activities that take place during speech production – the first channel being overt speech itself" [17, pp. 350]. Coverbal gestures are part of exolingual competence [11], [18] and aid in foreign language comprehension [5], [19], [20]. Furthermore, the act of teaching influences and changes nonverbal behavior [21]. In a pedagogical context, gestures and body language are used

more consciously to fulfill various functions, which incites us to consider this type of gesture as *teaching gestures* [3].

Teaching gestures, which are strongly related to the speech of the teacher, facilitate foreign language comprehension and memorization of lexis [4], [22]–[25]. Just as speech is modified during foreigner talk [9], [15], so too are gestures [26]. An experimental study by Tellier & Stam [6] examined the same future French teachers in two different situations: either sitting across from a native French speaker or from a non-native. As in a game of Taboo, these master's students had to explain a word to their partner without saying it. Tellier and Stam's analysis showed that when explaining words to non-native partners, future French teachers made gestures that lasted longer and were more iconic. Furthermore, the gesture rate increased and the gesture space grew larger when speaking to non-natives. Wagner *et al*. [27] mention that if the use of gesture is the best way to clarify ambiguity, an interlocutor will tend to use it. For these reasons, analysis of gesture during incomprehension sequences is indispensable.

There are certain particularities of webcammed interaction that can have an effect on communication and gesture production. In general, temporal delay, poor audiovisual quality and inadvertent disconnections are prone to making synchronous foreign language teaching more difficult, and teachers must learn to anticipate these technical problems [28]. Direct eye contact is impossible, because in order to give the impression of direct eye contact, one must look directly at the camera and off screen, therefore paradoxically not really looking at one's partner [7], [29]. Whether or not the webcam enhances or facilitates online communication is still up for debate. Wang [30] found that "video has been greatly appreciated by the distance participants and its pedagogical values are indispensable to language learning at a distance"; Develotte, *et al*. [31] found that visual clues provided learners with supplementary psycho-pedagogical, cultural and linguistic information; Develotte *et al*. [2] found that the video contributed to a more fluid interaction, as nonverbal clues helped to complete oral instructions. However, Develotte *et al*. [31, pp. 309] note that it can be distracting to see one's partner "either because it contradicts the oral message, because it makes no sense and adds nothing to it, or because it distracts the learner's attention," and that certain interlocutors prefer the audio channel, as "it appears that the use of a webcam image is more important in terms of its availability as a possible resource in case of need than as a favored type of communication." Guichon & Cohen [32] compared audio-synchronous and videoconferenced pedagogical interaction and found that the audiovisual condition enhanced neither the feeling of online presence, nor comprehension.

Most important, the webcam's field of view typically captures only the head and shoulders of each interlocutor, dramatically reducing the gesture space and leaving many gestures off screen and not visible [7]. As noted above, Tellier & Stam [6] found that when talking to non-natives, teacher trainees tended to make large gestures, thus widening the gesture space. In order for gestures to be visible onscreen, however, gestures must be made in a small space close to the face and shoulders, which is not natural, especially during exolingual conversation. Teacher trainees must adapt, as Develotte *et al*. [31] explain: "the video window can be compared to a theatre stage that the teacher trainees use to enact their role: they learn to adapt their gestures to the size of the stage." Because of this, we postulate that most gestures that are visible onscreen are allocentric [33], meaning that they were made in the center gesture space with a communicative purpose.

A teacher trainee is someone who is developing techniques to strategically monitor the online pedagogical interaction, and has three channels through which to deploy an array of strategies: the verbal channel, the gestural channel, and the textual channel. Verbal strategies used to facilitate comprehension may include use of synonyms, definitions, examples, translations, repetitions, reformulations, metalanguage, questions, and verifications of comprehension[1]. Foreign language teachers can consciously use gestures to inform, evaluate and animate [3]. The text chat is an invaluable tool for online language teachers because it can be used, among other things, to correct oral productions without interrupting the flow of the conversation [34], facilitate comprehension by repeating what is said orally [2] and allow learners to communicate in real time, modifying input and production, and responding to feedback all the while focusing on the form and structure of the language [14]. We shall see how teacher trainees divide their pedagogical moderation between each of these three channels during the interaction, especially during sequences in which there is manifest incomprehension.

## 3. Corpus

Our corpus is from the ISMAEL project[2] [35], which organizes webcammed multimodal interactions between future teachers of French at the University of Lyon 2 in France and business students who are learning French at Dublin City University in Ireland. Our corpus from the fall semester of 2013 is comprised of six weekly 45-minute interactions between groups of two (one teacher trainee, one learner) or three (one teacher trainee, two learners). For this study, we have chosen four interactions from the first week: two groups of two and two groups of three with durations of 41:19, 33:25, 43:53 and 42:14, respectively. The theme of this session was the French business world and its constituents (35-hour work weeks, paid vacation, coffee breaks, strikes, etc.). These interactions took place on an online language-learning platform called *Visu* (see figure 2) [34] and were saved using a screen recorder.

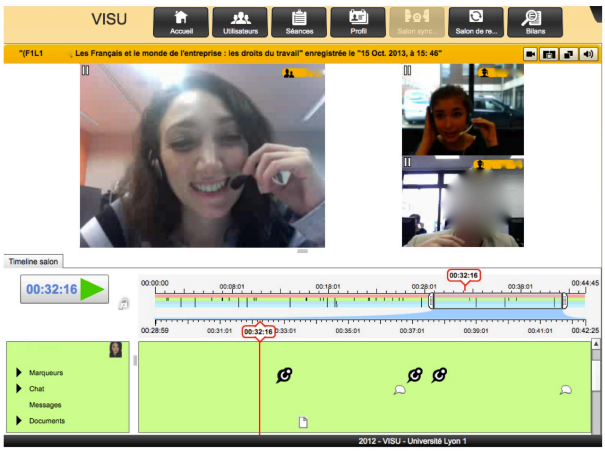



Figure 2: A multimodal interaction on *Visu*

[1] Strategies gathered from our corpus.

[2] InteractionS et Multimodalité dans l'Apprentissage et l'Enseignement d'une Langue (in English: Interactions and multimodality in the learning and teaching of a language).

## 4. Research questions and hypotheses

The aim of this study is to find out whether or not teacher trainees use their gestures to repair incomprehension and whether or not they make more gestures during incomprehension sequences. Our hypotheses are that teacher trainees will make more gestures during incomprehension sequences than during normal sequences (i.e. no manifest incomprehension) and that the text chat will be used more during incomprehension sequences.

## 5. Methodology

All four video recordings were transcribed and annotated using ELAN, which is open-source software designed for annotating multimodal interactions [36]. The spoken verbal messages were transcribed using the ICOR transcription convention [37], and the textual messages sent by the teacher were copied without alteration. We then annotated the incomprehension sequences, which begin with the teacher trainee's response (see figure 1) and end when the teacher trainee chooses to end the sequence.

For gesture annotation, we classified all hand gestures visible on the screen into six categories, including the four dimensions proposed by McNeill [16], [38]: iconics, metaphorics, deictics and beats, and added emblems and non-identifiable gestures. To compare use of gesture with other channels used, we annotated the oral verbal strategies listed above and counted the number of messages sent in the text chat window. Indeed, the chat window can be used for verbal strategies such as synonyms and translations, but since the purpose of this study is to measure the use of each channel, we left a single category for the text chat messages and kept verbal strategies as oral only. To remove some subjectivity and test the viability of our transcription guide, we calculated an inter-annotator agreement percentage for 21 gestures and for incomprehension sequences during 10 minutes of video, achieving ample agreement scores of 67% and 78%, respectively.

## 6. Analysis

The gesture, verbal and text chat strategies were counted for each interaction, and the percentage of each was compared during "normal" sequences (NS) and during incomprehension sequences (IS), with the total number listed under "entire interaction" (EI). Table 1 shows the duration of each type of sequence in seconds, and Table 2 shows the repartition of strategies counted among the three channels.

| | EI | NS | IS | %NS | %IS |
|---|---|---|---|---|---|
| Duration | 10066 | 9252 | 814 | 91.9% | 8.09% |

Table 1: Duration of sequences types (in seconds)

| | EI | NS | IS | %NS | %IS |
|---|---|---|---|---|---|
| Verbal | 611 | 511 | 100 | **69.0%** | **74.6%** |
| Text | 43 | 29 | 14 | **3.91%** | **10.4%** |
| Gesture | 221 | 201 | 20 | **27.1%** | **14.9%** |

Table 2: Number of strategies counted per modality per type of sequence

We see that during normal sequences, the audio channel was used for nearly 70% of all strategies observed, with text accounting for only 3.91% and gestures 27.1%. During incomprehension sequences, use of the audio channel rose slightly while use of the text chat nearly tripled, validating our second hypothesis. However, contrary to our predictions, the gesture channel was used less during sequences in which there was manifest incomprehension, dropping from 27% to 15%.

Even though the gesture channel does not seem to be a favored method of repairing incomprehension, it was used in some cases. In order to illustrate the usefulness of this channel, we will show two examples[1], in which teacher trainees used gesture to help repair incomprehension. In the following example, Victor (teacher trainee) uses a gesture[2] (see figure 3) to help explain the meaning of a 35-hour workweek to his learner during an incomprehension sequence:

VIC so **thirty-five hours is the maximum allowed by the law**

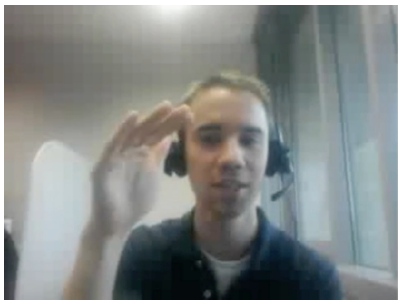


Figure 3

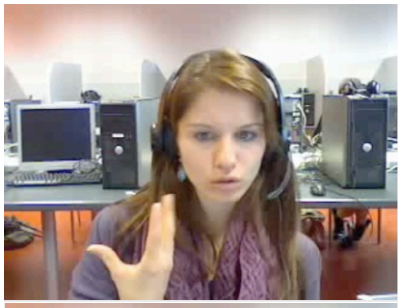
Victor's informational gesture [3], which is placed clearly within the webcam's field of view, complements his explanation of the 35-hour workweek law. The gestural representation of the concept of "maximum" may be easier for the learner to decode than the verbal explanation in a foreign language, and when used in conjunction with the word "maximum," which is the same in English as in French, helps to disambiguate its meaning.

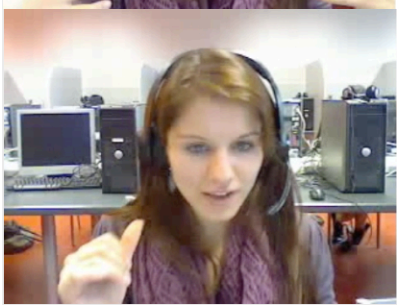
In the following example, Melissa (teacher trainee) uses gestures (see figures 4, 5 and 6) to verify her own comprehension of what Alejandra (learner) is trying to say:[3]

ALE hum: in hum spain I hum: (0.9) lots of things/ (.) (inaud.) hum hot/ (0.3) (inaud.) (1.0) [°(inaud.)°]
MEL [°hot° yes/ **hot drinks/]**
ALE (yes) (.) (0.5) yes and (0.4) hum hum hum (0.6) hum the beer/ [hum]
MEL [ah] (0.7) ok it's very hot in spain (1.3) **it's [hot in] spain/**
ALE [hum:] (0.9) yes because I live in hum the canary islands\ (1.4) [th-]
MEL [and] so you (0.2) **you drink beers/ (.) to [cool** off a bit\]
ALE [no n-] norma[lly] no but ((laughs))
MEL [°no/°] (1.1) o[k\]

[1] Translated from French to English by us.
[2] The part of the sentence corresponding temporally with the gesture is in boldface text. The gestures are pictured in order of occurrence.
[3] : = lengthened word; (#.#) = duration of pause greater than 200 milliseconds; (.) = micro pause shorter than 200ms; / = rising intonation; \ = falling intonation; <((description)) production> = description of production; °° = whispered or spoken very softly; [] = overlapping turns; UPPERCASE = perceptual salience; x = inaudible syllable; (inaud.) = inaudible series of syllables

ALE [the] hum (inaud.) in the: (0.6) the break/ (.) for coffee/ and ((laughs)) x (0.5) [and drink] a glass
MEL [°ok\°] (1.6) ok





Figures 4, 5 and 6

Melissa uses gestures to give feedback and to show what she understands, so that Alejandra can adjust her output accordingly. Feedback given through the gesture channel may be easier to decode than verbal feedback in a foreign language. First, Melissa simultaneously verifies her own comprehension of what Alejandra is talking about and proposes the words "hot drinks" accompanied by a gesture that represents a drink (see figure 4). During Melissa's following turn, she checks her own comprehension by asking if it's hot in Spain, and then repeats the question while using her hands to show that it's hot (see figure 5) to ensure comprehension. In figure 6, Melissa uses a French emblematic gesture of drinking to verify her own comprehension of drinking beer to cool off and to show Alejandra what was understood by her output in the foreign language. Alejandra's negative response shows that Melissa's question was understood, and due to Alejandra's low level of comprehension overall during this conversation, it is reasonable to believe that Melissa's gesture aided comprehension.

To further our understanding of gesture usage, we calculated the gesture rate (the number of gestures divided by the number of words spoken) for the entire interaction (EI), for normal sequences (NS), and for incomprehension sequences (IS) for each of the four teacher trainees (see table 5 on next page). Then, in table 3, we divided the gesture rate calculated for incomprehension sequences by the gesture rate calculated for normal sequences to see the difference. If the ratio is greater than one, then the gesture rate is higher during incomprehension sequences, and the inverse is true if the ratio is less than one. It is interesting to compare teacher trainees because each interlocutor has his or her own interactive and gestural profile which does not disappear during online interaction [7]. We found that teacher trainees 2 and 4 had a higher gesture rate during incomprehension sequences (ratio>1), whereas teacher trainees 1 and 3 had a lower gesture rate during incomprehension sequences (ratio<1), which supports [7]'s claim that each person uses gestures differently. The overall gesture rate for all four interactions (see table 4) shows that for our corpus, the gesture rate during incomprehension sequences was lower than the gesture rate during normal sequences.

| Interaction | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Incomp/norm | 0.56 | 1.30 | 0.64 | 1.7 |

Table 3: Ratio of incomprehension sequence gesture rate/normal sequence gesture rate

| | IS | EI | NS | IS/NS |
|---|---|---|---|---|
| # gestures | 20 | 221 | 201 | 0.76 |
| # words | 1265 | 10974 | 9709 | |
| Gesture rate | 0.016 | 0.020 | 0.021 | |

Table 4: Overall gesture rate

## 7. Discussion/conclusion

This was a preliminary study to examine how teacher trainees use multimodality in a videoconferencing-based teaching setting. The fact that these teacher trainees, overall, did not produce more gestures during incomprehension sequences than during normal sequences invalidates our first hypothesis. Perhaps the presence of a screen, keyboard and webcam alters the exolingual behavior of online teachers. Since the keyboard is widely used during technical problems [2], [39]–[41], teacher trainees might resort to using it during any type of problem, not distinguishing technical problems from linguistic incomprehension or miscommunication. Since it is impossible to make hand gestures and type at the same time, it seems in retrospect that our hypotheses may have been mutually exclusive. It would be interesting therefore to repeat this study in a videoconferencing environment without the possibility of sending text chat messages. The fact that these teacher trainees made fewer gestures during incomprehension sequences does not void the possibility of these gestures having specific, novel purposes worthy of study. One goal of our future research is therefore to define the functions of body language during videoconferenced pedagogical interaction.

It is difficult to draw generalizable conclusions based on these four teacher trainees' interactions. It is important to keep in mind that this was the first week of online interactions and that these trainees were not attuned to harnessing the affordances of the videoconferencing platform. Whereas the use of verbal strategies is known by most teachers and can be effectively transferred from face-to-face to online pedagogical interaction, making visible hand gestures in front of the webcam and using all three channels harmoniously is an entirely new skill that must be developed. To address this corpus it was necessary to identify strategies channel by channel, but it is clear that teacher trainees often use multiple channels simultaneously, two and sometimes three at a time, and thus the interaction between the channels must be studied to better understand the multimodal nature of these interactions. This exploratory research was necessary to

become familiar with a pedagogical situation that has been seldom studied, and for our future research we aim to see whether or not the teacher trainees progressively develop new strategies and ways of exploiting and combining the affordances offered by the videoconferencing platform over the course of a semester.

## 8. Acknowledgements

The authors are grateful to the ASLAN project (ANR-10-LABX-0081) of the University of Lyon for its financial support within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the National Research Agency (ANR). Special thanks is given to Erika Dussire and Jorane Soubesty for annotating sections of the corpus to obtain our inter-annotator agreement score.

| Interaction | 1 IS | 1 EI | 1 S | 2 IS | 2 EI | 2 NS | 3 IS | 3 EI | 3 NS | 4 IS | 4 EI | 4 NS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # gestures | 4 | 43 | 39 | 5 | 27 | 22 | 5 | 113 | 108 | 6 | 38 | 32 |
| # words | 340 | 2193 | 1853 | 413 | 2776 | 2363 | 186 | 2745 | 2559 | 326 | 3260 | 2934 |
| Gesture rate | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 0.03 | 0.04 | 0.04 | 0.02 | 0.01 | 0.01 |

Table 5: gesture rates by teacher trainee

## 9. References

[1] C. Develotte and F. Mangenot, "Discontinuités didactiques et langagières au sein d'un dispositif pédagogique en ligne," *Glottopol*, no. 10, pp. 127–144, 2007.

[2] C. Develotte, N. Guichon, and R. Kern, "'Allo Berkeley ? Ici Lyon... Vous nous voyez bien ?' Étude d'un dispositif de formation en ligne synchrone franco-américain à travers les discours de ses usagers," *Alsic*, vol. 11, no. 2, pp. 129–156, 2008.

[3] M. Tellier, "Dire avec des gestes," *Fr. Dans Monde Rech. Appl.*, no. 44, pp. 40–50, 2008.

[4] M. Tellier, "The effect of gestures on second language memorisation by young children," *Gesture*, vol. 8, no. 2, pp. 219–235, 2008.

[5] M. Tellier, "Usage pédagogique et perception de la multimodalité pour l'accès au sens en langue étrangère," in *La place des savoirs oraux dans le contexte scolaire d'aujourd'hui*, R. Bergeron, G. Plessis-Bélair, and L. Lafontaine, Eds. Montréal: Presses Universitaires du Québec, 2009, pp. 223–245.

[6] M. Tellier and G. Stam, "Stratégies verbales et gestuelles dans l'explication lexicale d'un verbe d'action," in *Spécificités et diversité des interactions didactiques*, V. Rivière, Ed. Paris: Riveneuve éditions, 2012, pp. 357–374.

[7] J. Cosnier and C. Develotte, "Le face à face en ligne, approche éthologique," in *Décrire la conversation en ligne, Le face à face distanciel*, C. Develotte, R. Kern, and M.-N. Lamy, Eds. Lyon: ENS Éditions, 2011, pp. 27–50.

[8] C. Cornaire and C. Germain, *La compréhension orale*. Baume-les-Dames: CLE Internationale, 1998.

[9] M. Long, "Native speaker/non-native speaker conversation and the negotiation of comprehensible input," *Appl. Linguist.*, vol. 4, no. 2, pp. 126–141, 1983.

[10] R. Porquier, "Communication exolingue et apprentissage des langues," in *Encrages : acquisition d'une langue étrangère*, 1984, pp. 17–47.

[11] R. Porquier and B. Py, *Apprentissage d'une langue étrangère : contextes et discours*. Le mesnil-sur-l'Estrée: Imprimerie Nouvelle Firmin Didot, 2008.

[12] C. Sarré, "Computer-mediated negotiated interactions: how is meaning negotiated in discussion boards, text chat and videoconferencing?," in *Second language teaching and learning with technology: views of emergent researchers*, S. Thouësny and L. Bradley, Eds. Dublin: Research-publishing.net, 2011, pp. 189–210.

[13] E. M. Varonis and S. Gass, "Non-native/Non-native Conversations: A Model for Negotiation of Meaning," *Appl. Linguist.*, vol. 6, no. 1, pp. 71–90, 1985.

[14] Í. Yanguas, "Oral computer-mediated interaction between L2 learners: it's about time!," *Lang. Learn. Technol.*, vol. 14, no. 3, pp. 72–93, 2010.

[15] M. Long, "Input, interaction, and second-language acquisition," *Native Lang. Foreign Lang. Acquis.*, vol. 379, pp. 259–278, 1981.

[16] D. McNeill, *Hand and Mind*. Chicago: The University of Chicago Press, 1992.

[17] D. McNeill, "So you think gestures are nonverbal?," *Psychol. Rev.*, vol. 92, no. 3, pp. 350–371, 1985.

[18] B. Azaoui, "Multimodalité des signes et enjeux énonciatifs en classe de FL1/FLS," in *Le corps et la voix de l'enseignant : théorie et pratique*, Paris: Ed. Maison des langues, 2014, pp. 115–126.

[19] D. Sime, "'Because of her gesture, it's easy to understand'. Learners' perception of teachers' gestures in the foreign language class.," in *Gesture: second language acquisition and classroom research*, Routledge., S. G. McCafferty and G. Stam, Eds. New York, 2009, pp. 259–280.

[20] S. Kellerman, "'I see what you mean': the role of kinesic behaviour in listening and implications for foreign and second language learning.," *Appl. Linguist.*, vol. 13, no. 3, pp. 239–258, 1992.

[21] L. Cadet and M. Tellier, "Le geste pédagogique dans la formation des enseignants de langue étrangère," *Cah. Théodile*, no. 7, pp. 67–80, 2007.

[22] M. Macedonia, K. Müller, and A. Friederici, "The impact of iconic gestures on foreign language word learning and its neural substrate.," *Hum. Brain Mapp.*, vol. 32, no. 6, pp. 982–998, 2011.

[23] W. C. So, C. S. Chen-Hui, and J. L. Wei-Shan, "Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall?," *Lang. Cogn. Process.*, vol. 27, no. 5, pp. 665–681, Jun. 2012.

[24] M. Macedonia and K. von Kriegstein, "Gestures enhance foreign language learning," *Biolinguistics*, vol. 6, no. 3–4, pp. 393–416, 2012.

[25] M. L. Rowe, R. D. Silverman, and B. E. Mullan, "The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language," *Contemp. Educ. Psychol.*, vol. 38, no. 2, pp. 109–117, Apr. 2013.

[26] T. W. Adams, "Gesture in Foreigner Talk," University of Pennsylvania, 1998.

[27] P. Wagner, Z. Malisz, and S. Kopp, "Gesture and speech in interaction: An overview," *Speech Commun.*, vol.

57, pp. 209–232, 2014.
[28] C. Dejean-Thircuir, N. Guichon, and V. Nicolaev, “Compétences interactionnelles des tuteurs dans des échanges vidéographiques synchrones,” *Distance Savoirs*, vol. 8, pp. 377–393, 2010.
[29] H. De Chanay, “La construction de l’éthos dans les conversations en ligne,” in *Décrire la conversation en ligne, Le face à face distanciel*, Lyon: ENS Éditions, 2011, pp. 145–172.
[30] Y. Wang, “Supporting synchronous distance language learning with desktop videoconferencing,” *Lang. Learn. Technol.*, vol. 8, no. 3, pp. 90–121, 2004.
[31] C. Develotte, N. Guichon, and C. Vincent, “The use of webcam for teaching a foreign language in a desktop videoconferencing environment,” *ReCALL*, vol. 22, no. 3, pp. 293–312, 2010.
[32] N. Guichon and C. Cohen, “The Impact of the Webcam on an Online L2 Interaction,” *Can. Mod. Lang. J.*, vol. 70, no. 3, pp. 331–354, 2014.
[33] J. Holler, M. Tutton, and K. Wilkin, “Co-speech gestures in the process of meaning coordination,” in *Proc. 2nd GESPIN—Gesture and Speech in Interaction Conf., Bielefeld, 5–7 September 2011*, 2011.
[34] N. Guichon, M. Bétrancourt, and Y. Prié, “Managing written and oral negative feedback in a synchronous online teaching situation,” *Comput.-Assist. Lang. Learn.*, vol. 25, no. 2, pp. 181–197, 2012.
[35] N. Guichon, F. Blin, C. Wigham, and S. Thouësny, “ISMAEL Learning and Teaching Corpus. Dublin, Ireland: Centre for Translation and Textual Studies & Lyon, France: Laboratoire Interactions, Corpus, Apprentissages & Représentations.” 2014.
[36] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, “Elan: a professional framework for multimodality research,” in *Proceedings of LREC*, 2006, vol. 2006, p. 5th.
[37] Groupe ICOR, *Convention ICOR*. Lyon: Université de Lyon. http://icar. univ-lyon2. fr/documents/ICAR_Conventions_ICOR_2007. doc, 2007.
[38] D. McNeill, *Gesture & Thought*. Chicago and London: The University of Chicago Press, 2005.
[39] C. Develotte, R. Kern, and M.-N. Lamy, Eds., *Décrire la conversation en ligne: la face à face distanciel*. Lyon: ENS, 2011.
[40] C. Kerbrat-Orecchioni, “Conversations en présentiel et conversations en ligne : bilan comparatif,” in *Décrire la conversation en ligne, Le face à face distanciel*, Lyon: ENS Éditions, 2011, pp. 173–195.
[41] A. Liddicoat, “Enacting participation : hybrid modalities in online video conversation,” in *Décrire la conversation en ligne, Le face à face distanciel*, Lyon: ENS Éditions, 2011, pp. 51–69.