Seeing the learner during a mediated tutoring session : influence of non-verbal cues on tutoring dialogue

Federico Tajariol, Jean-Michel Adam, Michel Dubois

Seeing the learner during a mediated tutoring session: influence of non-verbal cues on tutoring dialogue

Federico Tajariol, ARCADE, Laboratoire Clips (Imag), Université Joseph Fourier, Grenoble.
Jean-Michel Adam, ARCADE, Laboratoire Clips (Imag), Université Joseph Fourier, Grenoble.
Michel Dubois, Laboratoire de Psychologie Sociale, Université Pierre Mendès France, Grenoble.

ABSTRACT : *This paper deals with the effect of non-verbal cues on a video-mediated tutoring situation. We will first state communication theoretical underpinnings and then we will focus on the role of kinesic cues and ostensive-inferential cues in communication and mediated activities. After that we will describe the coding scheme we used to analyze the tutoring dialogue and we will present two experimental studies that evaluate the effects of these non-verbal cues on the following measures: the tutor's proactive behaviour, the tutor-student's mutual understanding, the efficiency of the dialogue and the intrinsic tutoring speech acts. We will finally conclude with some considerations for the design of interfaces to support video mediated tutoring dialogue.*

KEY WORDS
video-person, video-actions, tutoring dialogue, kinesic and ostensive-inferential cues

Introduction

Communication technologies are more and more used in mediated human activities, like in project meeting [1], in repairing equipment [2] and also in teaching activities [3]. Nowadays, audio-video technologies allow distant participants to see one another during the interaction [4] and even to show one another the actions they are doing in their working environment [5]. Nevertheless, several researches illustrate that the richness media framework [6] is not a completely adequate solution for the special needs in mediated activities [7]. From a pragmatic point of view, we do consider that affordances of communication tools are different according to the type of conveyed information [8] and the specific human activity.

In the next section we will define a background about communication from face-to-face interaction studies. Next, we will describe our experimental method and the two studies we conducted about the affordances of two non-verbal cues (kinesic and ostensive-inferential) on mediated tutoring dialogue. The outcomes show that the efficiency of tutoring dialogue is different according to the conveyed non-verbal cue and we suggest some principles to improve the design of interfaces supporting the mediated tutoring dialogue.

Background

Communication as a joint action

We communicate with another person to perform different activities, as planning, bargaining, creating, teaching, etc. Therefore we use language to act on our environment [9] and to make our addressee believe by saying words [10]. When we communicate with our partner, we cannot directly know his (her) thoughts, feelings and intentions. We are just able to infer them by interpreting our partner's utterances [11] and (her) his non-verbal behaviour [12]. Nevertheless, communication is not just a one-way human activity, where the speaker sends content and the addressee receives it. Communication is rather a joint action as playing a piano duet [13]. For instance, the tutoring dialogue is a cooperative action [14], where a tutor chooses the content of his (her) contribution according to the knowledge (s)he supposes student owns. Both tutor and student coordinate their turns to ground on a mutual understanding (dis-) confirming that they understand what has been said [15].

Communication as a multimodal process

When the speaker and the addressee are in a face-to-face setting, they enjoy the multimodal properties [16]. In fact, participants produce and interpret verbal and non-verbal cues. In the non-verbal cues class, we distinguish kinesic cues (e.g., facial expressions, postures, etc.) and ostensive-inferential cues (actions and deictic gestures). The former ensures the speaker-listener turns transitions and informs the participants about the other person's feelings and intentions [17]; the latter facilitates the verbal referring process, helps participants to coordinate their actions and to anticipate the other's needs [5]. For instance, in the tutoring dialogue, the tutor infers when and how to help students without disturbing by observing facial expressions [14]. So, we can suppose that in video-mediated communication setting it is sufficient for distant participants to have all non-verbal cues to act properly.

Non-verbal cues in mediated human communication

Nevertheless, in the literature related to the video-mediated communication, the effects of non-verbal cue are still discussed. On the one hand, several studies show that non-verbal cues do not always support participants to better perform their mediated activities [18, 19, 20, 21]. On the other hand, other studies show that kinesic and ostensive-inferential cues help distant participants to establish a mutual understanding [20, 22, 23] and that ostensive-inferential cues support participants to perform a mediated activity more quickly [24, 25]. We outline that these studies focus much more on the nature of non-verbal cues than on the socio-cognitive properties of a specific human activity. Following this point of view, we consider the synchronous tutoring activity [26, 14] and their socio-cognitive categories to analyze the effect of non-verbal cues on the efficiency and the contents of tutoring dialogue.

Research Problem

Our main research question is the following: which type of non-verbal cue improves the efficiency of tutoring dialogue? Would it be better if the tutor and the student see each other at a distance (kinesic cues)? Would it be better if the tutor observes the student's actions to improve (her) his learning (ostensive-inferential cues)?

To answer these questions, we set two experimental studies: the first one was dedicated to understand the affordance of kinesic cues; the second one was oriented to understand the affordance of ostensive-inferential cues and to compare their effects on tutoring dialogue. We will describe in the next section the experimental tasks and apparatus, as well as the measures we collected.

Method

In this section, we will describe the experimental task, the apparatus, the procedure and the dependant measures, which are common to the two studies.

Task

Tutors and students were involved in a sort of practical pedagogical work. One tutor had to help two students to learn basic commands of HTML language and to create an easy web page. In fact, they had to edit some sentences of a text in bold and italic, as well as building an internal link. The tutor could help and communicate with only one student at once, as in a dyadic tutoring situation: in fact, when the tutor was talking to a student, the other one could not hear their dialogue.

Apparatus

The tutor and the two students were in three separate rooms, each one was equipped with a personal computer (central unit and monitor) and another monitor was used to show the person. Anytime the tutor wanted to communicate with a student, (s)he chose the student by pressing a button. Both students could ask the tutor to help them anytime, sending him/her a standard message by means of synchronous chat software. Each computer was equipped with a web browser and a simple text editor.

Procedure

Each experimental session lasted nearly an hour.

In the preparing phase (15 minutes), the researchers introduced the tutor to the two students and the researchers explained the aims of the experimentation. Then, the tutor and each student settled down in three separate rooms. Each subject filled in a consent form and was briefed on the main functions of the

communicating technical apparatus. Next, each student answered a pre-test to evaluate his/her HTML knowledge.
Then the experimental session started and lasted 35 minutes. At the beginning, each student received a four-page HTML manual, which contained some HTML basic commands. Of course, the students needed their tutor's help to design the web page. The tutor had to help each student spontaneously or when a student asked for help.
In the last phase, each subject was asked to complete a questionnaire to evaluate the quality of the interaction with others, before he/she was debriefed and then dismissed.

Measures
We videotaped the 24 experimental sessions that we transcribed verbatim. Researches concerning the common grounding processes in communication [13], the affordances of video information in mediated communication [2, 27] and the tutoring dialogues [28] inspired us the following measures:

a) the proactive behaviour of the tutor : we distinguished the tutors' spontaneous interventions towards the students as proactive interventions from the tutors' reactive interventions when they had replied to a student's call;
b) the learning score : we measured the score of the web page that the students had realized[1];
c) the mutual understanding : we developed a coding scheme to categorize verbal markers the students and tutors had used in common ground process (Table 1);

Role	Category	Utterance/Example
Tutor	To accept student's utterance	T: "*yes*" "*ok, right*"
	To check student's understanding	T: "*is it clear now?*", "*is it ok?*"
Student	To accept tutor's utterance	S: "*yes*", "*ok*"
	To check tutor's utterance	S: "*could you repeat, please?*", "*what?*"

Table 1 - Mutual understanding coding scheme

d) the intrinsic speech acts concerning tutor-student dialogue : we classified the tutors' and students' speech acts which were related to the essential nature of the tutoring dialogue (Table 2).

Role	Category	Utterance/Example
Tutor	To find out student's ongoing task	T: "*Did you try moving it on the red icon?*"
	To help student	T: "*Close the*

[1] We mark 1 point for a right tag and 0.4 points for a right tag which was badly inserted in the html code.

		window and open the other file"
	To encourage student	T: "*That's good you've nearly finished*"
Student	To give tutor information about ongoing task	S : "*I still have to finish this part of the exercise*"
	To ask tutor's help	S: "*Is the I tag in the HEAD part of the text?*"

Table 2 - Intrinsic tutor's and student's tutoring speech acts

Three trained coders analysed the transcriptions of the dialogue. We calculated the reproducibility test [29] to evaluate the reliability of coding: we obtained an average value for the students' speech acts (K=.58) and a high value for the tutors' speech acts (K=.82).

Experimental Studies

In the next two sections we will describe two studies: first, to understand the effects of the video-person (kinesic cues) and second, the study concerning the effects of video-actions (ostensive-inferential cues) on tutoring dialogue.

Experiment 1: what are the effects of seeing our partner on tutoring dialogue?

Hypotheses

The tutor's behaviour was expected to be more proactive when a video-person link was available, because the tutor could observe the student's face and infer his/her difficulties and motivation during the practical work.

We expected the efficiency of the dialogue to be improved when a video-person link was available, because the subjects could better coordinate their contributions to the dialogue by seeing each other's face and producing fewer speech acts.

Furthermore, we also predicted the students' performance to improve because the tutors could better help and encourage the students while they performed the task.

Subjects

As tutors, we recruited twelve subjects, half men and half women (age M=30.9, S.D.=7.7). All of them were computer scientists owning good skills at HTML programming. Two of them were associate professors, whereas the others were Ph.D. students with almost one-year teaching experience in computer science. They were paid 25 euros each for their participation.

As students, we recruited forty-eight undergraduate students in psychology (thirty-six women and twelve men; age M=24.1, S.D.=6.2), all of them unskilled at HTML programming. For their participation, they received a credit for their social psychology course.

Experimental conditions

We set the following conditions:

c1) audio only: the tutor chose the student s/he wanted to talk to.

c2) audio & video-person: the tutor and each of the two students could see each other's face and upper torso on a personal monitor, and they communicated by means of the audio channel.

Each tutor performed in both conditions (within-participants experimental design), whereas half students were assigned to one of the two conditions (between- participants design). We controlled that the age of the students involved in the two conditions were equivalent (t(46)=1.318;n.s.) and so was the HTML knowledge (t(46)=0.09;n.s.). To guard against order effects, we counterbalanced the tutors' performing order setting.

Figure 1 – Experimental conditions of study 1: audio only (left) *vs*. audio & video-person (right).

Main outcomes

We conducted independent *t-test* and we calculated the value of the effect size *d*.

a) There were significantly more tutors' proactive interventions in the audio & video-person than in the audio only condition (M = 4.6 (2) vs. 1.8 (0.9), t (22) = 4.27; $p<.001$, $d = 1.8$).
b) Concerning the learning score, no significant differences were found between the two conditions (audio only M=4.2(1.4) *vs.* audio & video-person M = 4.6 (1.5), t(22)=0.83 ; n.s., $d =1.1$).
c) Concerning the mutual understanding, statistical analysis showed that tutor had significantly produced fewer verbal markers to check students' understanding in audio & video-person than in audio only condition (M=3.9 (3.6) *vs*. 10 (6.4), t(22)=2.85; $p<.01$, $d=1.17$). About other students' and tutors' verbal markers, we did not find any significant differences between the two conditions.
d) Concerning the speech acts referred to tutoring dialogue, we first presented outcome about tutors' speech acts and then about students' speech acts.

   i) Encouraging acts - Tutors produced significantly more speech acts to encourage students in audio & video-person than in audio only condition (M=3.2(1.3) *vs*. 1.2(1.2), t(22)=2.9; $p <.01$, $d=1.14$).

   * The number of tutors' speech acts oriented to know students' ongoing task was significantly higher in audio & video-person than in audio only condition (M=60.7 (33.2) *vs*. 35.3 (8.1), t(22)=4.27; $p<.02$, $d=1.05$).

   ii) Consequently, to inform their tutors about their ongoing task, the students significantly resorted to more speech acts in audio & video-person than in audio only condition (M=63.6 (28.4) *vs*. 43.8 (14.4), t(22)=2.15; $p <.05$, $d=0.88$).

No significant differences between the two conditions were found about other speech acts.

From these outcomes, it seems that communicative affordances of kinesic cues are the following: the tutor more often takes the floor spontaneously, (s)he produces fewer verbal markers to check students' understanding and (s)he more willingly encourages them. Furthermore, the tutors and students spend more time talking about the ongoing task.

Experiment 2: the effects of observing students' actions on tutoring dialogue.

The task, the procedure and the measures of this second experiment were the same than for the first one. In addition, we revised the coding scheme to include deictic utterances the tutors and students had used in the mutual understanding process. Deictic utterances were used to refer to the objects they had seen. We coded deictic expressions (i.e. T: "*you move the icon down*", S: "*here*?", T: "*yes*") and pronouns (i.e., T: "*take this and move it on your desktop*").

Hypotheses

We expected the tutor to be more proactive when (s)he could observe the students' actions, because (s)he was always aware of their difficulties and (s)he could take the floor before they asked for help.

Mutual understanding was expected to be easier when the tutor could observe the student's actions, because they shared a visual environment. Moreover, we expected the tutor and students to describe less explicitly their actions and all related objects, so they would produce more deictic utterances.

We also predicted the student's performance to improve when the tutor could monitor them during the practical work session, because the tutor could choose the most appropriate help according to their needs.

Subjects

We recruited the same tutors (N = 12) as for the first study, whom we paid 25 euros each for their participation.

We also recruited seventy-two undergraduate students in psychology (fifty women and twenty-two men; age M=23.8, S.D.=5.1), all of them unskilled at HTML programming. They received a credit for their social psychology course.

Experimental conditions

We set the three following conditions:

c1) audio & video-person: the tutor and the students could see each other's face and upper torso on one screen and they communicated by means of the audio channel;

c2) audio & video-actions: the tutor and the students could communicate by the audio channel, and the tutor could observe both students' computer screens (by means of VNC software). So, the tutor's computer screen was split in two windows showing the students' desktops and (s)he could observe their activity and make the decision to help them at will;

c3) audio & video-person & video-actions: the tutor could observe both students' face and upper torso and their computer screens, while the verbal communication between the tutor and each student was supported by the audio channel.

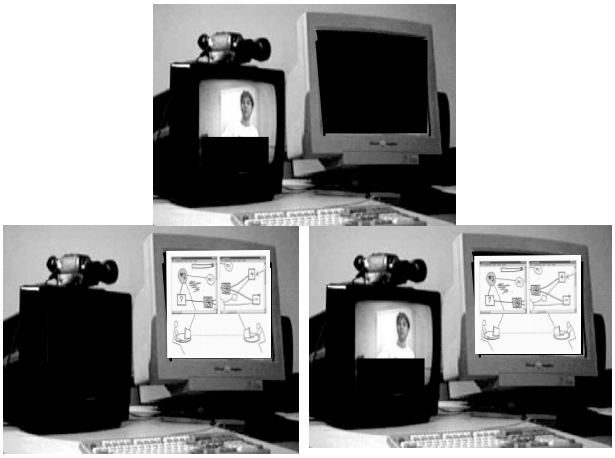

Figure 2 – Experimental conditions study 2:
audio & video-person (up) *vs*. audio & video-person (left)
***vs*. audio & video-person & video-actions (right).**

As for the first study, each tutor performed task in all conditions (within-participants design) and we counterbalanced the order of the tutors' performance to guard against order effects. About students, they were equally shared between the three experimental conditions. We controlled the age ($F(2, 69) = 1.73$; n.s.) and the HTML knowledge ($F(2, 69) = 0.16$; n.s.) of each group of students.

Main Outcomes

We conducted an analysis of variance (ANOVA) and *a priori* comparisons between conditions.

a) About the tutor's proactive behaviour, the number of tutor's spontaneous interventions is significantly different between conditions ($F(2, 33) = 4.671$; $p < .02$, $\eta^2 = 0.22$). In fact, the tutor is significantly more proactive in audio & video-person & video-actions than in audio & video-actions condition (M = 6.6 (1.9) *vs*. 4.6 (2), $t(33) = 2.366$; $p < .02$, $d = 1.02$) and than in audio & video-person condition (M = 6.6 (1.9) *vs*. 4.2 (2.2), $t(33) = 2.859$; $p < .007$, $d = 1.16$).

b) The students' web page score is significantly different between conditions ($F(2, 69) = 5.776$; $p < .005$, $\eta^2 = 0.14$). The students perform significantly better in audio & video-person & video-actions than in audio & video-person condition (M = 5.9 (1.6) *vs*. 4.5 (1.5), $t(69) = 3.27$; $p < .002$, $d = 0.88$) and better in audio & video-actions than in audio & video-person condition (M=5.5 (1.2) *vs*. 4.5 (1.5), $t(69) = 2.4$; $p < .02$, $d = 0.73$).

c) About mutual understanding, the number of deictic utterances is significantly different between conditions ($F(2, 33) = 5.453$; $p < .01$, $\eta^2 = 0.99$). A-priori comparisons show that the tutors and the students produce significantly more deictic utterances in audio & video-actions than in audio & video-person (M=125.2 (21.6) *vs*. 90.9 (29.3), $t(20.2) = 3.27$; $p < .005$, $d = 1.33$) and more in audio & video-person & video-actions than in audio & video-person (M=134.3 (46.2) *vs*. 90.9 (29.3), $t(18.6) = 2.75$; $p < .01$, $d = 1.12$). About the tutors' and students' verbal markers targeted to mutual understanding, outcomes do not show any significant differences between conditions.

d) About tutoring speech acts, we will first present outcomes for tutors (i) and then for students (ii).

   i) One-way ANOVA shows that the number of tutor's speech acts to find student's ongoing task is significantly different between conditions ($F(2,33) = 4.529$; $p < .02$, $\eta^2 = 0.21$). The Tutor produces fewer speech acts oriented to the student's task in audio & video-actions than in audio & video-person condition (M = 27.4 (7.6) *vs*. 35.3 (8.1), $t(22) = 2.8$; $p < .01$, $d = 1$) and fewer in audio & audio-person & audio-actions than in audio & video-person condition (M = 26.6 (6.8) *vs*. 35.3 (8.1), $t(22) = 2.6$; $p < .005$, $d = 1.16$).

   * the Tutor's speech acts to help the student. A one-way analysis of variance shows significant differences between conditions ($F(2, 33) = 15.248$; $p < .000$, $\eta^2 = 0.48$). There are fewer help speech acts from the tutor in audio & video-person than in audio & video-actions (M=14.8 (10.4) *vs*. 40.2 (12.9), $t(33) = 4.99$; $p < .000$, $d = 2.18$) and fewer in audio & video-person than in audio & video-person & video-actions (M=14.8 (10.4) *vs*. 37.9 (13.8), $t(33) = 4.55$; $p < .000$, $d = 1.91$).

* The Tutor's encouraging speech acts - A one-way analysis of variance shows significant differences between conditions ($F(2, 33) = 4.813$; $p < .01$, $\eta^2 = 0.22$). The Tutors more often encourage the students in audio & video-person than in audio & video-actions condition (M=3.2 (2.3) *vs*. 0.8 (1.1), t(33)=2.69; $p<.01$, $d=1.33$) and more often in audio & video-person & video-actions than in audio & video-actions (M=3.2 (2.8) *vs*. 0.8 (1.1), t(33)=2.69; $p < .01$, $d=1.12$).

ii) About the student's speech acts to give their tutor information about the ongoing task: the one-way analysis of variance shows significant differences between conditions ($F(2, 33) = 4.529$; $p < .02$, $\eta^2=0.21$). the Students less often give their tutors information about their ongoing task in audio & video-person & video-actions than in audio & video-person condition (M=26.5 (14.1) *vs*. 43.8 (14.4), t(22)=2.5; $p<.01$, $d=1.21$) and less often in audio & video-actions than in audio & video-person condition (M=27.7 (18.4) *vs*. 43.8 (14.4), t(22)=2.6; $p<.005$, $d=0.97$).

All other students' speech acts are substantially equivalent between conditions.

Measures	**Categories**		**(c1) audio & video-person**	**(c2) audio & video-actions**	**(c3) audio & video-person & video-actions**	**Comparisons**
Tutor's proactivity	Tutor's spontaneous interventions		M = 4.2 (2.2)	M = 4.6 (2)	M = 6.6 (1.9)	C3 > C1 ($p<.007$) C3 > C2 ($p<.02$)
Mutual understanding	Deictic terms		M = 90.9 (29.3)	M = 125.2 (21.6)	M = 134.3 (46.2)	C2 > C1 ($p<.005$) C3 > C1 ($p<.01$)
Learning score	Student's task performance		M = 4.5 (1.5)	M = 5.5 (1.2)	M = 5.9 (1.6)	C3 > C1 ($p<.002$) C2 > C1 ($p<.02$)
Intrinsic tutoring speech acts	Tutor	To find out student's ongoing task	M = 35.3 (8.1)	M = 27.4 (7.6)	M = 26.6 (6.8)	C1 > C2 ($p<.01$) C1 > C3 ($p<.005$)
		To help student	M = 14.8 (10.4)	M = 40.2 (12.9)	M = 37.9 (13.8)	C2 > C1 ($p<.000$) C3 > C1 ($p<.000$)
		To encourage student	M = 3.2 (2.3)	M = 0.8 (1.1)	M = 3.2 (2.8)	C1 > C2 ($p<.01$) C3 > C2 ($p<.01$)
	Students	To give tutor information about ongoing task	M = 43.8 (14.4)	M = 27.7 (18.4)	M = 26.5 (14.1)	C1 > C2 ($p<.005$) C1 > C3 ($p<.01$)

Table 3 - Main outcomes study 2

Conclusions

The aim of these experiments was to understand communicative affordances of kinesic cues (video-person) and ostensive-inferential cues (video-actions) on a same activity, the mediated tutoring dialogue. The outcomes of study 1 show that when the tutor and the student can see each other's face, the tutor more often takes the floor spontaneously, (s)he produces fewer verbal markers to check the student's understanding and (s)he more willingly encourages him/her.

Outcomes from study 2 partially corroborate study 1, while confirming some previous researches [20, 24]. For example, the tutor's communicative behaviour is more proactive when kinesic cues are available. On the contrary, it appears that ostensive-inferential cues let the tutor and the student focus on intrinsic learning contents and improve their mutual understanding. Moreover, the student's performance is better and the efficiency of the dialogue is higher.

For the design of the interface supporting mediated tutoring dialogue, these main outcomes imply that:

If we want the tutor to be proactive, it is better to let the tutor and the student see each other;

If we prefer the student to improve learning and the dialogue to be more efficient, it is better to let the tutor observe the student's ongoing actions.

However, these suggestions have to be taken according to the limits of our studies. The first concerns the content of tutoring interaction: we are planning to conduct another experimentation in which the content will be rather declarative (about the comprehension of a dialogue and text in a foreign language). The second limit is related to the number of students who were involved in each experimental session. We think that if a tutor had to assist more students at the same time, it would be necessary to investigate the form of ostensive-inferential cues. For instance, we are going to study a visualisation form by which the tutor can just observe the steps twelve students perform during a practical work.

References

[1] J.C. Tang & E. Isaacs, Why do users like video video? Studies of multi-media supported collaboration. *Computer supported cooperative work, 1*(3), 1993, 163-196.

[2] R. E. Kraut, S.R. Fussell & J. Siegel, Visual information as a conversational resource in collaborative physical tasks. *Human Computer Interaction, 18*, 2003, 13-49.

[3] B. Collis, *Tele-learning in a digital world: The future of distance learning.* (London: International Thomson Publications, 1996).

[4] C. Heath, P. Luff & A. Sellen. Reconfiguring media space : supporting collaborative work. In K.E. Finn, A.J. Sellen & S.B. Wilbur (Eds.) *Video-mediated communication* (NJ : Lawrence Erlbaum, 1997, 323-349).

[5] B. Nardi, A. Kuchinsky, S. Whittaker, R. Leichner & H. Schwarz. Video-as-Data : Technical and social aspects of a collaborative multimedia application. In K.E. Finn, A.J. Sellen & S.B. Wilbur (Eds.) *Video-mediated communication* (NJ : Lawrence Erlbaum, 1997, 487-519).

[6] R.L. Daft & R.H. Lengel, Organisational information requirements, media richness and structural design. *Management Science, 32*, 1986, 554-571.

[7] S. Whittaker, Theories and methods in mediated communication. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123-164). (Hillsdale, NJ: Lawrence Erlbaum, 2002)

[8] H.H. Clark & S.E. Brennan, Grounding in communication. In L. B. Resnick, J. Levine & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*. (pp. 127 - 149). (Washington, DC: American Psychological Association 1992)
[9] J.L. Austin, *How to do things with words*. (Oxford : Oxford University Press, 1962).
[10] J. Searle, *Speech Acts, An Essay in the Philosophy of Language*. (Cambridge:Cambrige University Press,1969).
[11] D. Sperber & D. Wilson, *Relevance: Communication and cognition*. (Oxford: Blackwell, 1987).
[12] M. Argyle & J. Graham. The Central Europe Experiment : looking at persons and looking at things. *Journal of Environmental Psychology and Nonverbal Behaviour, 1*, 1977, 6-16.
[13] H.H. Clark, *Using languages* (Cambridge : Cambridge University Press, 1996)
[14] B.A Fox, *The Human Tutorial Dialogue Project: Issues in the Design of Instructional Systems*. (Hillsdale NJ: Lawrence Erlbaum, 1993).
[15] H.H. Clark & D. Wilkes-Gibbs, Referring as a collaborative process. *Cognition, 22*, 1986, 1-39.
[16] H. McGurk & J. MacDonald, Hearing lips and seeing voices. *Nature, 264*, 1976, 746-748.
[17] G. Amado & A. Guittet, *Dynamique des communications dans les groupes*. (Paris : Colin, 2003).
[18] A.H. Anderson, A. Newlands, J. Mullin, A.M. Fleming, G. Doherty-Sneddon & J. Van Der Velden, Impact of video-mediated communication on simulated service encounters. *Interacting with computers, 8*, 1996, 193-206.
[19] O. Daly-Jones, A.F. Monk & L.A. Watts, Some advantages of video-conferencing over high-quality audio conferencing : fluency and awareness of attentional focus. *International Journal of Human Computer Studies, 49*, 1998, 21-58.
[20] G. Doherty-Sneddon, A. Anderson, C. O'Malley, S. Langton, S. Garrod & V. Bruce, Face-to-face interaction and video mediated communication : a comparison of dialogue structure and co-operative task performance. *Journal of Experimental Psychology : Applied, 3*, 1997, 105-125.
[21] C. O'Malley, S. Langton, A. Anderson, G. Doherty-Sneddon V. Bruce, Comparison of face-to-face and video-mediated communication. *Interacting with computers, 8, 2*, 1996, 177-192.
[22] E. Isaacs & J.C. Tang, What video can and cannot do for collaboration. *Multimedia Systems, 2*, 1994, 63-73.
[23] S. Whittaker, E. Geelhoed & E. Robinson, Shared workspaces: how they work and when are they useful? *International Journal of Man-Machine Studies, 39*, 1993, 813-842.
[24] S.R. Fussell, L. D. Setlock & E.M. Parker, Where do helpers look? Gaze targets during collaborative physical tasks. In *Proceedings of CHI 03 Extended Abstracts* (pp. 768-769). NY: ACM Press, 2003.
[25] D. Gergle, R.E. Kraut & S.R. Fussell, Language efficiency and visual technology: minimizing collaborative effort with visual information. *Journal of Language and Social Psychology, 23*, 2004, 491-517.
[26] M.T.H. Chi, N. De Leeuw, M.H. Chiu & C. Lavancher. Eliciting self-explanations improves understanding. *Cognitive Science. 18*, (3), 1994, 439-477.
[27] L. Karsenty, Cooperative work and shared visual context : an empirical study of comprehension problems in side-by-side and remote help dialogues. *Human Computer Interaction, 14*, 1999, 283-315.
[28] M.T.H. Chi, Quantifying qualitative analyses of verbal data : a practical guide. *The Journal of the Learning Sciences, 6*(3), 1997, 271-315.
[29] Carletta, J., Isard, A., Isard, S., Kowtko, J.C., Doherty-Sneddon, G., & Anderson, A.H., The reliability of a dialogue structure coding scheme. *Computational Linguistics, 23*(1), 1996, 13-31.