



HAL
open science

Developing a multimodal communication to enhance the writing competence in an audio-graphic conferencing environment

Maud Ciekanski, Thierry Chanier

► To cite this version:

Maud Ciekanski, Thierry Chanier. Developing a multimodal communication to enhance the writing competence in an audio-graphic conferencing environment. EUROCALL 2007, Mastering Multimedia: Teaching Languages Through Technology., Sep 2007, University of Ulster, Coleraine Campus., United Kingdom. edutice-00200851v1

HAL Id: edutice-00200851

<https://edutice.hal.science/edutice-00200851v1>

Submitted on 21 Dec 2007 (v1), last revised 21 Feb 2008 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Developing a multimodal communication to enhance the writing competence in an audio-graphic conferencing environment

Maud Ciekanski, Thierry Chanier

LASELDI, Université de Franche-Comté, 30, rue Mégevand, 25030 Besançon cedex, France

{maud.ciekanski, thierry.chanier}@univ-fcomte.fr

Abstract

Since the last decade, most studies in Computer Mediated Communication (CMC) have highlighted how online synchronous learning environments implement a new literacy related to multimodal communication. The environment used in our experiment is based on a synchronous audio-graphic conferencing tool. This study concerns false-beginners in an English For Specific Purposes (ESP) course, presenting a high degree of heterogeneity in their proficiency levels. An original coding scheme was developed in order to transcribe the video data into a set of users' actions, part of them being speech acts, which occurred into the different modalities of the system (aural, textchat, text editing, websites).

The paper intends to shed further light on and increase our understanding of multimodal communication structures through learner participation and learning practices. On the basis of evidence from an ongoing research investigation into online CALL literacy, it will seek to identify how learners use different modalities to produce collectively a writing task, and how the multimodal learning interaction affects the learner's focus and engagement within the learning process. The adopted methodology combines quantitative analysis of learner's participation involved in a writing task in regards to the use of multimodal tools, and qualitative analysis focusing on how the multimodal dimension of communication enhances language and learning strategies. A particular attention will be paid to the benefits of a group producing process in writing task (whether collaborative or cooperative), in terms of metacognitive strategies and social learning strategies, through self and co-evaluation practices. By looking at the relationship between how the learning tasks are designed by tutors and how they are implemented by learners, that is to say taking into account the whole perception of multimodal communication for language learning purposes, it attempts to provide a framework for evaluating the potential of such an environment for language learning.

Keywords

multimodality, CMC, collaborative learning, writing competence, audio-synchronous environment, learning strategies.

1. Introduction

In his last review concerning the evolution of the technology choice in the area of computer-assisted language learning (CALL), Stockwell (2007: 113) shows how technology has moved on from CALL to computer-mediated communication (CMC) and computer-supported collaborative learning (CSCL). This evolution concerns every language skills and areas. Writing then is no more perceived only in its personal dimension, but as an interactive process which may be mediated successfully by computers and group of learners. Previous experiments, such as (Dejean & Mangenot, 2000), have shown the successful collaboration between learners in front of the same computer to write collectively one text, and how the screen provided a convergent effect, facilitating the collaboration. These experiments have highlighted the importance of learning discussions about writing in order to help learners develop their writing learning awareness.

The recent development of synchronous online environments integrating a large range of modes has given the opportunity to set up pedagogic scenarios and communication scenarios designed to enhance collaboration supported by a combination of modes and modalities. Three types of multimodal synchronous online environment can be listed. These environments refer to different kinds of multimodality and communication:

- audio synchronous environment integrating audio and chat (verbal communication) (eg. Jepson, 2005),
- videoconferencing environment integrating audio, video and chat (verbal and non-verbal communication) (Andrews, 1993; McAndrew, 1996; Stevens & Altun, 2002; Wang, 2004),
- audio-graphic conferencing environment integrating audio, graphic and chat (verbal and non-verbal communication).

This paper concerns an Audio-Graphic Synchronous Environment (AGSE) which includes communication tools and shared editing tools. Previous recent studies analysed multimodal communication in AGSE. These studies concern a wide range of dimensions: modes affordances (Hampel & Baber, 2003), task design (Hampel, 2006), oral communication (Lamy, 2006; Chanier *et al.*, 2006), tutoring practices (Vetter, 2004, Hampel & Hauck, 2005), multimodal communication model (Chanier & Vetter, 2006), multimodal communication strategies (Jeannot *et al.*, 2006), etc.

Our study carries on the first observations made by Vetter and Chanier (Vetter & Chanier, 2006) concerning the correspondence between modes. Whereas the majority of previous investigations were based on oral communication, we focus here on the writing skill. This paper sets an original methodological approach which aims at contributing to a better understanding of the specificities of multimodal communication in AGSE and its relevance to the development of the writing competence, supported by a writing process perceived as a complex and procedural activity which does not require much instruction but action, and as a social event. It may be of interest to language teachers who want to design online collaborative writing activities intertwined with synchronous communication, as well as researchers who want to further investigate the relationships between writing and multimodal communication from a general perspective stressed by Lamy (2007:237):

[...] the aim is [...] to identify methods for the analysis of language learner conversations in such environments so as to better understand how to promote multimodal conversation as a legitimate learning activity of the electronically literate. We do not endorse the view that technologies are a mere support for conversational activity, the script of which is then decoded through traditional language-centered methodologies [...]. Instead we look upon technologies as mediating the social event that is the conversational process.

In section 2 we set the scene by presenting the learning situation, an English for Specific Purpose (ESP) course for students involved in a master degree in open distance teaching (ODT), and the characteristics of the AGSE environment. Then, in section 3, we give a definition to the term "multimodality" related to our framework and emphasize three main modalities which will support the writing process, i.e. textchat, audio, and actions in a shared word processor. We discuss methodological approaches for analysing conversations where notions such as participant's perspectives and context become prominent. These general considerations are backed on works coming from the Ethnography of Communication or Conversation Analysis and more recent works which developed analytic frameworks such as (Baldry & Thibault, 2005; Kress *et al.*, 2001).

In order to closely investigate the relationships between multimodality and language learning, more particularly the writing competence, after having presented our original coding schema used when transcribing video screenshots, we start section 4 with the analysis of two writing activities. We explain how participants play with multimodalities in order to accomplish their tasks, at an individual level or a group level where the expected collaboration appears. Relying on this comprehension of the writing and learning processes we subsequently, in

section 5, unfold constraints and patterns of use between modalities, successively considering internal (intramodal) and external (intermodal) relationships, thanks to new tools developed by our research team (Betbeder *et al.*, 2007b). Lastly, in section 6, we outline the interest of such environment in second language learning and compare it to other collaborative writing environments, which are asynchronous. Part of the conclusion draws the reader's attention to the current tendency in CALL to overlook the writing competence when working in synchronous environment.

2. The research experiment: population, and environment

2.1. The CoPéAs project

The ESP course was designed as part of the research project CoPéAs (*Communication Pédagogique et environnements orientés Audio-Synchrones*) ran between the Université of Franche-Comté (France) and the Open University (UK) in 2005, involving 16 French-speaking students divided into two groups of eight according to their level (false-beginner or advanced learners). Each group was tutored by an English-native tutor of the Open University, proficient in designing pedagogic materials for online distance learning. Tutors and learners met in the audio synchronous environment for eight sessions from one to one-and-a-half hour each during ten weeks. Learners worked at home with their own computers. In addition, they used an asynchronous learning management system in order to consult instructions and publish individual pieces of written work. The course aimed at developing vocational English and competences in ODT through spoken and written English (for further details on the CoPéAs project, see Chanier *et al.*, 2006). The study presented here will focus on the less proficient group where some learners had not practiced the target language for years (15 to 30 years).

The research protocol includes audio and video recording of the AGSE (screen captures with a software), saving learners' productions (individual and collaborative), pre-questionnaires and post-interviews of the tutors and learners. To give a glimpse of the corpus recorded, there are 37 videos corresponding to 27 hours, 512 files (productions, audiograms of the interviews, questionnaires, etc.) for 35 Go.

2.2. Specificities of the AGSE

The AGSE used in our experiment is Lyceum¹, developed and used within the Open University, and designed to facilitate distance tutorials. Its structure allows tutor and learners to meet each other synchronously. The different participants connected to the environment are able to communicate orally in real time, participate in the textchat, and read/modify simultaneously textual or graphic productions. The interest of Lyceum in language learning has already been stressed by various authors (see references in the previous section). We only define here our own view of the structure of the environment.

In Lyceum, every participant (tutor and learner) shares the same interface and the same rights. The interface is composed of three components, which we outlined in 3 frames on [figure 1](#):

- **Spatial component** (frame1): participants move from room to room or from document to document within one room. Participants can be located thanks to one grey rectangle in the spatial component; here the participant is in room 101. It is also possible to see who is in the lobby. The participants can only perceive each other (audio, graphic, chat, production) if they share the same room. They are then listed in the communication component (frame 2)
- **Communication component** (frame 2): it includes audio, vote and textchat tools. Each participant can, at any time, talk to the others by one click on the button “Talk” (eg. Tim and Sophie), raise the hand to ask for talking (eg. Lucas), to vote (tick) “Yes” (eg. Sophie) or “No” (eg. Laetitia) to answer to a question asked to the whole participants, or take collectively one decision. It is also possible to notify one’s way out (eg. Julie). The textchat is another tool in its communication cluster.
- **Shared editing tools component** (frame 3): three kinds of shared editing tools are provided: a whiteboard which allows learners to write, draw and import images or text, a concept map for writing and organizing information, and a word processor (mistakenly labelled "Document" in the interface) providing the opportunity to write by several hands a single text. Up to five documents, generated by these tools, can be opened at the same time. Every participant can only see and work with one document at a time, which means that participants who share the same room and communication tools may not visualize the same document. Icons at the top of the frame display participants' distribution among

¹ Lyceum : <http://kmi.open.ac.uk/projects/lyceum/>

documents. Everyone can add or suppress a document, save or download productions in a document, and act in every open document.

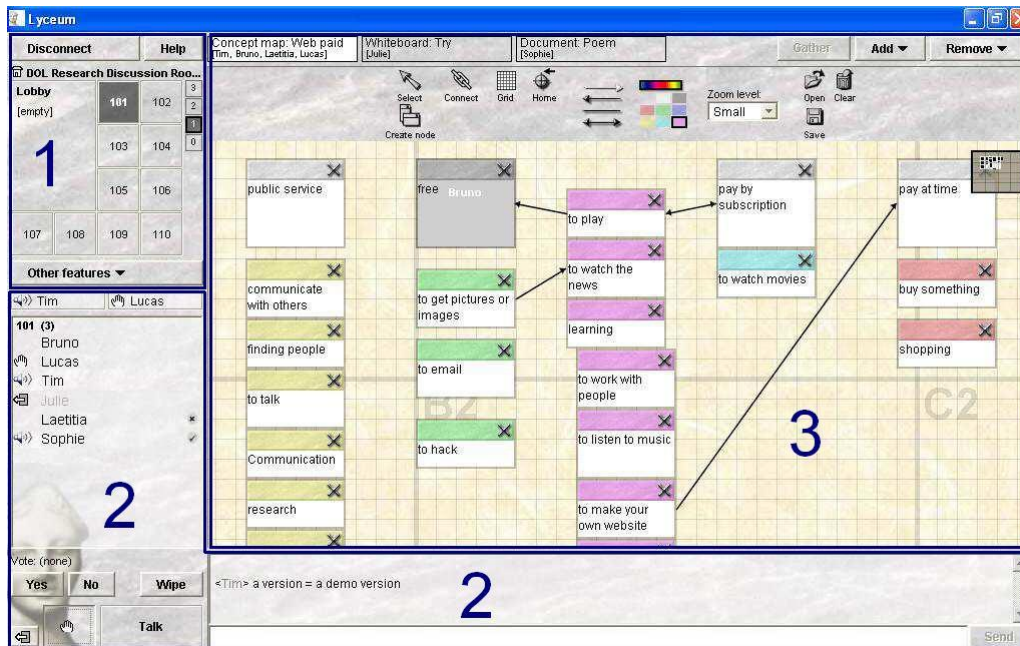


Fig. 1. The Lyceum interface and its 3 components

3. Multimodality in synchronous environments: definition and methodology of analysis

3.1. Defining multimodality in AGSE

The AGSE supports modes of communication, which are semiotic resources constructing discourse in interaction, such as textual, speech, graphic, iconic and spatial (which corresponds to the participants' localisation and movement in the different rooms and documents). Modalities are attached to each mode (see [table 1](#)). For example, the written linguistic mode is realised within the different modalities of textchat, word processor or whiteboard (on which textboxes may be created). A single mode may therefore be associated with several modalities, or with a single one (as for speech mode and the audio modality). In the following data, we are concerned with two modes (written language and spoken language) and three modalities (audio, chat, word processor, now WP).

Modes	Modalities
Textual	Chat, word processor, conceptual map, whiteboard
Speech	Audio
Graphic	Conceptual map, whiteboard
Iconic	Vote, in/out, away for a moment, raising hand, taking of the flow
Spatial	Movement (room + document)

Table 1: Correspondences between modes and modalities in Lyceum

Participants can communicate from a large semiolinguistic repertoire of particular interest in language learning. The richness of such a repertoire asks for an organisation of all these modes. Thus, Kress & Van Leeuwen (2001) defines multimodality as:

the use of several semiotic modes in the design of a semiotic product or event, together with the particular way in which these modes are combined – they may for instance reinforce each other [...], fulfil complementary roles [...] or be hierarchically ordered

3.2. Methods for the analysis of language learner conversations

Since the 90's, numbers of works, especially in Discourse Analysis, define the multimodality as a dynamic process of meaning-making (Kress *et al.*, 2001, Scollon & Levine, 2004). Multimodal communication is co-constructed through the interaction (between pieces of semiotic resources or participants) and cannot be studied block by block but as a whole, made of pieces of different natures like a patchwork. Interpreting multimodality means re-building the meaning given by the participants while communicating. As in any communication, the meaning given by the speaker may differ from the one perceived by the addressee. Thus, multimodality cannot be studied as a static composite production. From a macro-scale, different aspects have to be taken into account, in particular, the participant's perspective and the context.

The participant's perspective

A majority of research on multimodality back their approaches on the Ethnography of Communication or on Conversation Analysis, based on Malinowski's studies, who discussed the notion of context and culture. The notion of **action** (i.e. what is being done in the situation with various semiotic tools, called resources) is prevalent. The focus of the analysis is not the

“potential” meaning of a text², but the way people interpret it in different situations and activities. Participants, through their actions, show to each other how they understand the situation, (“*what is going on*”, citing Goffman’s motto), what is the focus of their attention (what they consider as relevant in a composite communication). Following Halliday’s social theory of communication (1978), we argue that in verbal interactions with others we have at our disposal a network of options (or sets of semiotic alternatives) which are realized through sets of options of the semantic system. The alternatives selected within the network of meanings can be considered as traces of decision made by sign-makers (participants) about what is the most appropriate and plausible signifier for the expression of the intended meaning in a given context (Kress, 1997; Kress *et al.*. 2001). Following this tradition, we chose here to analyse the learning process from the learners’ and tutor’s perspectives, from choices they made out of a set of available meaning-making resources, in a particular situation at a given moment.

The notion of context

The notion of context is of the utmost importance in the study of every interaction (Goodwin & Duranti, 1992; Goffman, 1974). And it is even more fundamental in distance interaction occurring in AGSE. As we said before, context may be plural thanks to possibilities which participants have to move from one room to another or from one document to another. Following Goodwin and Duranti (1992:3), “*the notion of context involves a fundamental juxtaposition of two entities: (1) the focal event; (2) a field of action within which that event is embedded*”. What is the focal event in AGSE? As Jones said (2004: 27) “*In the 'digital surround' created by new communication, communication is more polyfocal*”. Will participants be lost among the multiple possibilities offered by the learning environment? Our experiment rather shows that learners make consistent **individual** choices to participate to multimodal discourse. It is possible to discern “*focused engagements involving clear and discernable involvements*” (Goffman, 1983). They also make **collective** choices. AGSE can be depicted as an environment of mutual monitoring possibilities, characterised by “*the moment-by-moment shifts of alignment participants bring into interaction to signal 'what they are doing' and 'who they are being'*” (Goffman, 1964).

² We define “text” following Halliday (1989: 10), as “*any instance of living language that is playing some part in a context of situation [...]. It may be either spoken or written, or indeed in any other medium of expression that we like to think of*”.

Taking into account these two dimensions, our work develops an actional perspective (the nature of the action impacts on the choice of multimodal component). The "mode action", as named by Kress, *et al.* (2001) is fundamental in that specific communication situation.

Therefore, the diversity of actions which may be performed requires a heterogeneous code of transcription to symbolize verbal and non-verbal modes, oral and written forms, etc. The unit chosen to describe interactions is the “**act**”. As Baldry and Thibault remind us (2005: xvi) “*Transcription is a way of revealing both the co-deployment of semiotic resources and their dynamic unfolding in time along textually constrained and enabled pathways and trajectories*”.

Three dimensions are taken into account so as to “cut” the text into phases:

- the dynamicity of the text (the text is here considered in interaction)
- the historicity of the action (the text is studied in its historicity, from a longitudinal perspective)
- the meaning-making unit (shared understanding)

Each modality is not analysed for itself. Multimodality is seen as a cluster of modalities connected to each other. We do not consider action through one single modality (only the word processor, for example).

Text is always multimodal, making use of, and combining, the resources of diverse semiotic systems in ways that show both generic (standardised) and text-specific (individual, even innovative) aspects (Baldry & Thibault, 2005:19).

Our coding schema

We defined an original coding schema. Every act (whether verbal or non-verbal) has got a **time** and **duration**, i.e. beginning and end, and is attached to a **workspace**, defined here as a basic frame (space + time) in order to describe participants' actions within one collaborative tool (concept map, whiteboard or WP). The notion of workspace is important because actions occurring in one space at a given time may not be perceived by participants located in a separate space at that time. Each act is also placed in a **sequence**. Every Lyceum session is divided into sequences, linked to the pedagogical scenario such as: greetings, tutor's guidelines in one common room, group divided into sub-group which attend separate rooms, after sub-task have been completed the group meets again in the common room in order to share results. An act is defined by the preceding attributes and by an actor, a modality (audio, vote, chat, production), its value (what has been done or said) (For further details, see

Betbeder, *et al.*, 2007a). [Figure 2](#) displays a series of acts, extracted from the fifth session, from the sequence 3 (S5.3) which corresponds to one of the two writing activities discussed in this paper. Attributes such as end time and workspace have been removed in order to simplify the presentation. Silences occurring in the audio modality are not represented here.

In this extract, 3 learners are working in a sub-group. They fill the quiz (10 questions) the tutor gave them at the beginning of the session. The tutor (TutT) came in the room while AT1 was writing a definition of ESL in the word processor with some explanations where the word "mainly" was included. The extract starts when AT6 orally requests a clarification about this word. Green arrows follow up this conversation (numbered 1): AT1 replies orally (action 2), then the tutor answers 3 times in the textchat (3, 5 and 12), and the learners apparently paying no attention to this, leaves the room (action 14). A second conversation (orange arrows) starts with AT3's contest, who sees no reason for talking about ESL (action 4). It is closed in 10 with AT3's agreement. This conversation occurred only in the audio modality. A few moments later, AT3 starts a third conversation by proposing to switch to the third question of the quiz. This conversation (blue arrows) alternate between audio and word processor modalities.

User	T	Start. Time	Audio transcript	Textchat Log	Production (WP)	Spatial move
AT6	1	43:38	what's mean euh + mainly ++ please what's mean mainly			
AT1	2	43:46	mainly is euh + usually ++			
TutT	3	43:49		principally		
AT3	4	43:51	but euh what euh do you write the definition of ESL it's not the + it's not the question	①		
TutT	5	43:58		generally		
AT1	6	43:59	no because because the + the the interactive element + on the whiteboard + euh + name ESL games + ESL help + ESL quizz ++			
AT6	7	44:16	yes +			
AT3	8	44:18	yes but ++ but it's not important think +			
AT1	9	44:24	I euh +++ for me euh ++ it's important ++			
AT3	10	44:32	ok			
AT3	11	44:42	do you want to answer at the third ++ euh question			
TutT	12	44:45		③ ... are both equivalents of "mainly"		
AT1	13	44:52	euh + yes			
TutT	14	45:04				leave
AT1	15	45:23	I euh + for me it's a games ++ maybe + I don't know <laugh>			
AT3	16	45:35			create : 'perhaps games'	
[...]						
AT6	17	46:09	and and ++ and help is no + no good			
AT3	18	46:18			edit : 'perhaps games an also help'	

Fig. 2: Simplified versions of a tabular transcription of a screen video, extract from S5.3.

Actions occurring in the shared editing tools component are transcribed as series of production acts ("prod", for short). In the shared word processor (WP), different types of acts and a unit of action have been defined in relationships with the functionalities of the tool (Table 2).

Type of acts	Unit taken into account
create	Paragraph (write)
select	Paragraph
edit	Paragraph (modify form and/or content)
suppress	Paragraph
enter	Paragraph

Table 2: Acts and unit of analysis in the WP

After having defined our framework, we can now qualitatively study the writing process in L2 and understand (1) how the collaborative work can support the development of the writing competence, and (2) the affordances of AGSE on the writing competence.

4. Analysis of two writing activities

4.1. Aims of the ESP course

The overall objective of the course was to enable learners to develop skills for working with foreign partners in distance learning field. Students practiced their English by doing some complex tasks involving reading/listening and production activities in the AGSE. The very challenge of the course was to develop the participation of false-beginner learners for ESP. Learners were also asked to do writing tasks collectively. These tasks were seen by the tutors as an incitation to orally negotiate meaning and form.

The pedagogical aim of the writing tasks is to make learners improve their writing competence, by developing their awareness of the writing process. A “learning-by-doing” approach has been chosen. False beginners had the possibility to prepare their written production before the online meeting. They were asked to write the text in collaboration, which permitted learners to alleviate the cognitive load by focusing on a limited number of aspects at a time. As Dam *et al.* (1990) highlight it, while writing, learners are faced to three main difficulties: linguistic, lack of automatic strategies for producing a text, and sociocultural difficulties. The writing tasks proposed during the course are much more centred on the writing process (planning, writing, revising, correcting and publishing) than on the production in itself. The design of the tasks aims at developing writing strategies in L2 for false beginners.

We report here two types of writing activities built around the word processor: a questionnaire, based on the selection of pertinent items contained in the web-site “English-Club” (learners have to recognize forms, and contents and to reformulate them), and a guided production which evaluates the AGSE used during the training (criteria were elicited). The first activity corresponds to a sub-part of session 5, (S5, located approximately in the middle of the course). The false beginners group had been divided into 2 sub-groups gr1 (N=3), and gr2 (N=3), respectively in sequence 2 and 3. The tutor gave oral guidelines and moves from one room to another. The activity lasted 15 minutes. The second activity lasted approximately 30 minutes, was also designed for sub-groups and occurred at the end of the training, in

session 8 (S8, sequence 2 and 3 for every sub-group). Learners reported after the experiment that they considered this activity as much more difficult than the one in session 5.

4.2. When writing the "text" with modes and modalities

The analysis presented in this section concerns the word processor and its correspondences to other modes and modalities. In this section, we examine the way participants (principally learners) combine modalities with the word processor in order to collaboratively write a text, what were their individual strategies, and collaborative practices.

Table 3 offers an overview of the distribution of acts according to modalities in the two writing activities and in every sub-group. We will often refer to it.

Session.Sequence	Sub-group	Distribution of acts according to modalities
S5.2	gr1	79 acts : 31chat > 25 audio > 23 prod
S5.3	gr2	85 acts: 62 audio > 17 prod > 6 chat
S8.2	gr1	337 acts: 199 audio > 173 prod > 65 chat
S8.3	gr2	300 acts: 166 audio > 104 prod > 31 chat

Table 3. Distribution of acts according to modalities ("prod" refers to production acts in the word processor (WP)).

4.2.1 Considering writing acts only

Let us start by considering only one modality, the one corresponding to writing acts (the prod modality in the WP). As mentioned previously, we distinguish four types of acts (see [Table 2](#)): create, edit, select and suppress. [Figure 6](#) displays the proportion of acts performed in both groups in S5 and S8. We have to keep in mind that the total amount of writing acts is 7 times higher in S8 than in S5 (see [table 3](#)). Acts of creation correspond to a deeper involvement of the writer than acts of edition, and even more than acts of selection or deletion. In session 5, the proportions of creation and edition acts are the same in both groups and the total number of acts almost equivalent. They represent between 70 and 80% of the total amount.

Apparently both groups managed the writing of short sentences and clauses for the quiz in the same way, alternating between creation and edition.

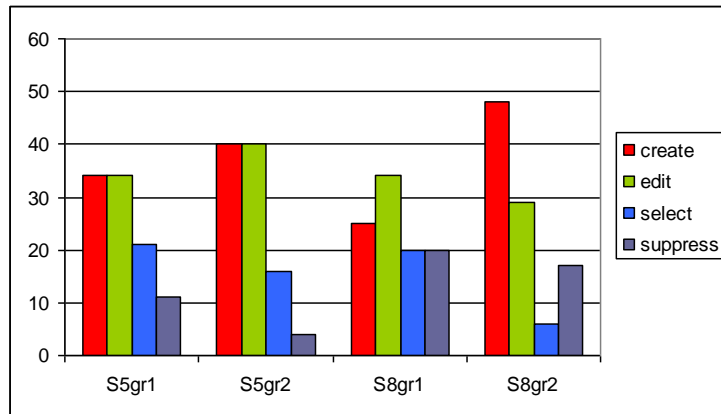


Fig6. Percentages of type of acts in the word processor per subgroup and activity

The situation is strikingly different in session 8. Firstly, creation and edition acts are no longer balanced in the two groups. Secondly, each sub-group performed the activity in a very different way: in gr1 the number of editions is greater than the number of creations whereas it is the opposite in gr2. A deeper involvement in the writing process seems to have occurred in gr2.

But examining one modality at a time is like scrutinizing a monocolour copy of a painting and trying to guess out of it how the painter composed it. We need, as we said, examine "the co-deployment of different semiotic resources and their dynamic unfolding" in order to understand how learners achieved the canvas.

4.2.2 A hierarchical organization

[Figure 3](#) displays the proportion of acts per modality (vote, audio, textchat and word processor), per subgroup performed during the two activities. It is surprising to note that, even when involved in a writing task, learners perform most of the activity orally (see, S5, gr2), by discussing what and how to write. When comparing sub-groups and activities, we can see (1) the variety of combinations performed by learners (chat dominated in S5 within the group who had the lower L2 competence; gr1 and gr2 had recourse to the vote in S8 in this more demanding activity), and (2) the prevalence of the speech mode.

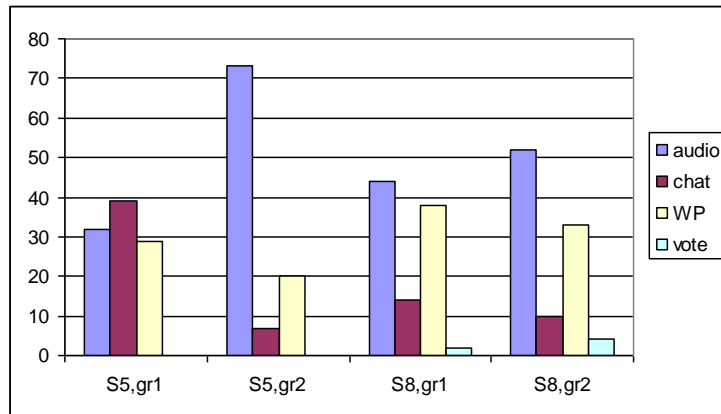


Fig. 3. Proportions of modalities used in sessions 5 and 8 per writing activity and per sub-group

These three verbal modalities (audio, textchat and WP) compose the repertoire for language production and play complementary roles to support the writing process. It may be explained by their very natures which differ when considering such attributes as synchronicity and duration.

Synchronicity: audio > textchat > WP
 Duration: WP > textchat > audio

Audio is more "synchronous" than textchat. Turn-taking happens more rapidly, there is no need to type, but also no possibility to rephrase one's production, once it is uttered.

Intervening in the word processor is a slower process. Only one participant at a time can edit a paragraph (but several paragraphs can be edited by different users). On the contrary, language has a longer lifecycle / duration in the WP than in the textchat, where the scrolling is permanent. In its turn, production in the textchat lasts longer (before it disappears from the current window; recalling past turns in the chat with scrolling device is rarely ever done) than oral utterances. These attributes may influence learners' choices while working: it is easier to quickly comment a piece of the text by oral than by textchat. These aspects are described in the next sections.

Comparing the use of the multimodal repertoire between S5 and S8, it appears that the use of the speech mode differs within the same group. Various parameters may impact on the use of the organization of the multimodality such as individual factors we will now consider.

4.2.3 Participant and group perspectives

Let us now pay attention to the way individual participants employ modalities they have at their disposal. During the quiz activity of session 5 in sub-group 1 (gr1), learners performed 79 acts evenly distributed among modalities (see [Table 3](#)). [Figure 4](#) displays individual

practices. Productions in the word processor are equally distributed among the 3 learners, and learner AT5 uses in equal proportion the three modalities. However learners AT4 and AT2 adopted opposite behaviour. AT4 rather participated in the WP and audio, whereas AT2 participated in textchat in complement to the WP. Their use of textchat and WP are proportionally inversed.

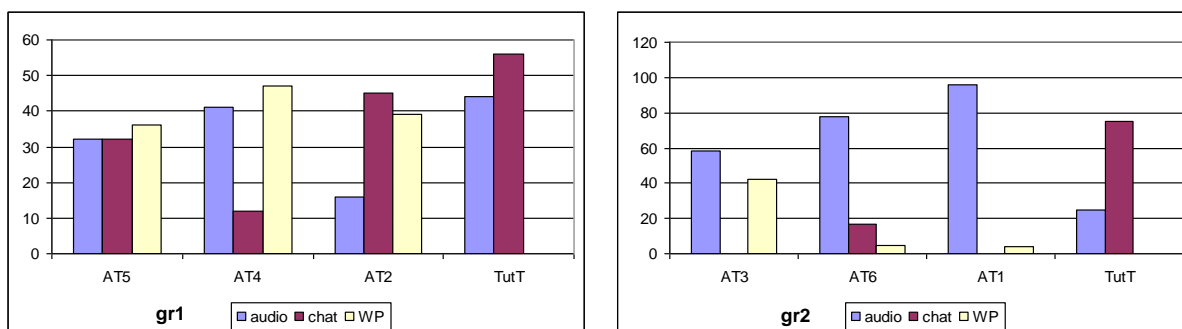


Fig.4. Percentage of modalities per participant in subgroups in the written activity of session 5

The other sub-group, gr2, performs during the same activity 85 acts distributed among the three modalities in a very unbalanced way (see [Table3](#) and [figure 4](#)). The speech mode is predominant; textchat has almost not been used, except by the tutor; one participant took over the word processor.

The detailed analysis of the individual practices also reveals a singular distribution of roles among the learners from group to another. Whereas learners share the same role in gr1 (they all write and correct the text), learners in gr2 distribute roles and functions.

The written production process in gr1 can be characterized as "cooperative" (each one contributes in a separate manner to the common performance). The "modal density" (Norris, 2004) for that space-time is concentrated on the WP with some punctual use of audio and textchat, in order to solve difficulties. The negotiation among learners comes *a posteriori*, during the correction phase of the writing process.

In gr2 only one learner (AT3) is in charge of the role of the writer. The other two participate orally to the construction of the text. The use of another mode and modality (here speech and audio) helps learners with difficulties participate at different levels and different degrees to the writing process. The interventions made by AT3 in the WP are the result of speech acts by the two other learners who are co-builders of the meaning and also correctors (sometimes by text, often by oral). [Figure 2](#) illustrates this phenomenon, which occurred during the

discussion tagged 3 (blue arrows). In gr2 negotiation takes place before writing in the WP, showing a collaborative way of achieving the text.

4.2.4 Contextual factors

The previous focalization on the modalities per participants gave an indication on the way individuals in a given context perceive the function of each mode and modality. Nevertheless, contextual factors may impact on the participants' choices. [Figure 5](#) compares subgroups during the second written task in session 8. As far as gr2 is concerned, the sub-group adopted the same collaborative model than in session 5 (AT3 is still the writer of the group). However the situation in gr1 varies widely from the one encountered in S5. Note that the intervention of the tutor in the WP corresponds only to the act of saving the document. He does not intervene in the text.

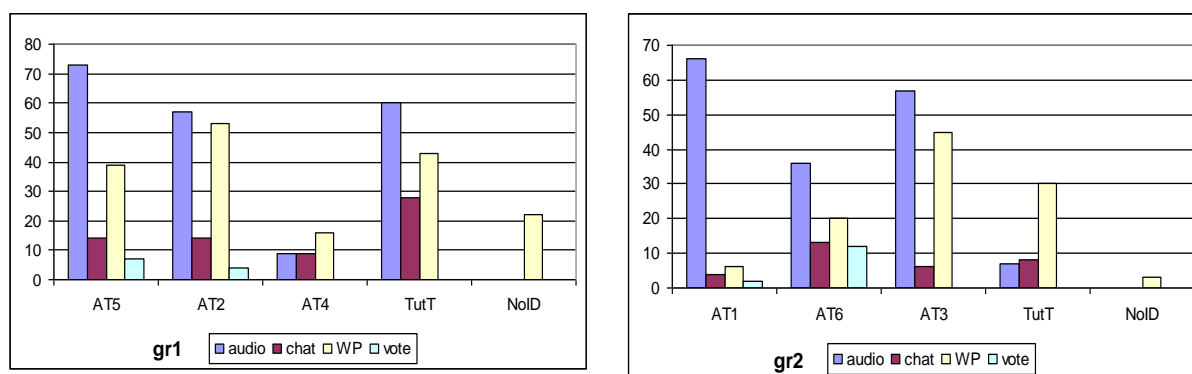


Fig.5. Percentages of modalities per participant in subgroups in the written activity of session 8. "NoID" refers to action in the word processor where the writer could not be identified.

Whereas in S5, learners only had to pick sentences or words in a referenced hypertext to fill a questionnaire, in S8 the activity requires to formulate full sentences and to express opinions. Accordingly, in the latter situation, learners changed their multimodal strategies in order to face several difficulties. In gr1, the dual strategy including speech mode and WP is used by AT5 and AT2 (AT4 encountered sound difficulties). The context challenges the individual preferences. For example, AT2 does not like much using audio (this fact came out of other analyses and of post-interviews). But here, the bi-modal strategy shared with AT5 is successful. She is even the one who proposed to compensate written difficulties by the cut of the writing process into different steps (conception of the meaning by oral, followed by a first formulation in the WP, then a correction by oral or in the WP).

The interesting aspect in the other sub-group, gr2, is the use of the textchat besides the WP in order to propose, enrich, and correct the text typed in the WP. The chat here functions as a “writing laboratory” where everything is allowed. On the contrary, gr2 seems to perceive the text in the WP as the final state of the text (they try to achieve to the better content and form before laying it down in the WP).

5. Unfolding intra and intermodal constraints in the multimodal discourse

In a discourse analysis perspective, the analysis of the multimodal organization in a given context highlights the way modes and modalities are intertwined to compose a discourse by which the collective action can be displayed. In the framework of our research project, CoPéAs, a pattern recognition tool has been developed (Betbeder *et al.*, 2007b) It is based on Manilla's Winnepi algorithm (*ibid*). The issue is to try to identify among recurrent patterns, systematic manners of accomplishing actions with modalities. Patterns can be represented as a kind of rule with a left hand and a right hand side and a connector "=>" meaning "followed by".

```
(1) [modality1, participantX] => [modality2, participantY]
(2) [modality1, participantX], [modality2, participantY] => [modality3, participantZ]
```

In (1) the algorithm considers two consecutive acts. It can be phrased as "the algorithm found a tendency (a recurring pattern) where modality1 used in an act by a participantX is followed by another act where modality2 is used by a participantY. In (2), considers three acts, i.e. it searches a recurring sequence where two acts are followed in a systematic way by another third.

5.1. Intramodal patterns

Firstly, let us consider patterns extracted by the algorithm where modalities on the left and the right hand sides are the same, i.e. intramodal patterns. In our corpus, for all sessions running from S1 to S8, the general trend is that 75.6% of patterns are monomodal. These patterns have a form like the ones in (3) and (4) where audio acts are followed by audio acts, or in (5) and (6), where acts accomplished with one share editing tools are followed by acts in the same modality.

```
(3) [audio, x] => [audio, y]
(4) [audio, x], [audio, y] => [audio, z]
```

- (5) [prod, x] => [prod, y]
 (6) [prod, x], [prod, y] => [prod, x]

Indeed, this intramodal organization automatically computed is typical of the corpus extracts analysed in our study. Examples (7) illustrate it within the audio modality and (8) within the word processor tool.

- (7)
 [AT3, audio]: do you want I write in the document
 [AT1, audio]: yes ok XXX right++ok I (euh++)

- (8)
 [AT2, (WP select)]: I like when you want my opinion on such or such thing
 [AT5, (WP edit)]: I like when you want my opinion on such or such thing
 that enables me to make efforts so that me am included/understood"
 [AT4, (WP enter)]:
 [AT2, (WP create)]: we like when we need to give your opinion

This “**modal density**” (Norris, 2004) shows that participants at some moments of the action focus on one single modality, which becomes prevalent in their interaction in order to perform the target action. The notion of polyfocalisation is defined here as successive monomodal phases. The difference between the two subgroups and between the two activities reveals that the multimodal organization is flexible according to what participants think to be the most efficient for them.

5.2. Intermodal patterns

Secondly, if we consider intermodal patterns (left and right hand sides display different modalities), they only represent a minor part of the overall number of patterns but some of them may occur in a significant number of times within a given activity, more particularly in sub-group's work. In sessions S5 and S8, we can find recurrent bi-modal patterns (audio & production / chat & production) such as (9) and (10), where there is a switch from audio to production (WP). In (9), one speech act accomplished by AT1 and followed by one production act by AT3 "generates" a speech act by AT6. in (10), the same actor (AT6) accomplishes one communicative goal through the successive use of two modalities, talk first, then writing in the word processor. Example (11) illustrates pattern (9).

- (9) [audio, AT1], [prod, AT3] =>[audio, AT6]
 (10) [audio, AT6] =>[prod, AT6]
 (11)
 [AT1, audio]: euh in help euh interactivity is euh ++ FAQ+
 [AT3 (WP suppress), (WP edit)]: help only
 [AT6, audio]: yes ++ just+choose help++ok

The notion of “modal density” takes here another signification and highlights the fact that the complexity of the action experimented by participants generates the use of more than one modality to accomplish the target action. The intermodality appears in our corpus as

scaffolding strategies to participate into and achieve the collective action (a distributed multimodality). The polyfocalisation is characterized by the integrated use of various modes and modalities which required that participants pay attention to several, oral and visual, elements of the environment. The addition of another mode has to be negotiated in order to be perceived by the participants. For example, in [figure 2](#), in the discussion numbered one (green arrows), the tutor gives several explanations in the textchat after a request of AT6. But learners are involved in longer term task where the chat is not prevalent. They ignore it and the tutor eventually leaves the room.

5.3. Summing up the multimodal layout

The prevalent multimodality organization is a succession of monomodal phases. Yet, the analysis of the work in sub-groups shows that participants use more often multimodal strategies. Therefore, the “modality-switching” is a strategy frequently used by learners involved in a collaborative writing task. Two bimodal couples have been identified and each mode has its own function:

audio (process-oriented)+ prod (product-oriented)
 textchat (product-oriented) + prod (product-oriented)

From a pedagogical standpoint, it seems that intermodal strategies enhance written participation (the writing task is performed in the textchat and in the WP). The textchat in complement to the WP is used for several purposes: proposal of draft versions, content focus on particular words of the text in order to discuss their meaning, or form focus where norms (correction, appropriateness) of the written text are argued.

In addition, intermodal strategies also support the writing process. The use of a combination of textchat, audio and WP occurs in the corpus as follow:

- Textchat is used to improve the text, especially the form of the message (cleaning procedure).
- Audio is rather used to comment and negotiate the content of the message (punctually, after a first written version).
- Audio is used to provide a first version of the written message (in L2 or even sometimes in L1), and to facilitate the written process (the deep procedure is realized through audio and WP. The focus is here on the content rather than on the form).

(12)
 [AT3, (WP, edit)]: to have comparaisn between web sites, to know more

criteria for a good site
[AT1, chat]: according to the different criteria ?

In (12), the textchat gives, in an interactive manner (note the question mark) another version of a sub-part of the written message which came out in the word processor. Each element of the whole text (according to Halliday's perspective) nourishes the global reflection of the participants, in a dynamic way.

6. Outcomes for language learning

After having depicted the organisation of multimodality related to writing tasks situated in an AGSE, we would like to discuss its impact on L2 language learning from two perspectives: language learning awareness, and language autonomy.

6.1. Enhancing language learning awareness

The multimodal learning environment because of its process-oriented and collaborative nature helps learners to pay more attention to the writing process than to the result of their writing. The integrated word processor gives a shared visible image of the writing process. WP could be compared to the Wiki, another collaborative writing tool of a more asynchronous nature. In a Wiki the history of the modifications and interventions are recorded for later examinations. With the WP, learners work simultaneously from the conception of the idea to its formulation and assessment. The synchronicity of the device, combined with the communication component of the AGSE put the emphasis on the negotiation of meaning and forms. On the one hand, learners are "spectators" of the writing process during sufficient time to develop a deep analyze of it. Their reflection on their own writing process may then be enriched by the opportunity to compare their practices to the others'. On the other hand, the collaborative nature of the task, backed on various modalities, gives every participant the possibility to have a different share in the common work, while enriching their language learning strategies. Thus, learners punctually develop "learning conversations" by which they focus on the learning process. These learning conversations are of three types: (1) comments on the writing process, (2) scaffolding of the learner's reflection, and (3) shared references from which assess their language learning, according to the previous typology developed by Harri-Augstein & Thomas (1981). This last point is particularly relevant in our corpus in terms of the assessing process. Learners compare their production to the others' and auto-correct what they've first written. They can also correct others' productions. It is interesting to note that usually they will ask by oral to the others if they don't mind being corrected, or they

formulate excuses orally or in the textchat (with smileys, for example) to save the face of their peers.

6.2. Enriching individual language repertoire and language autonomy

In terms of language production, the study of our corpus highlights two aspects: a rather high degree of participation, and occurrences of language autonomy strategies. As it has been shown in section 5.3., the use of multimodality by learners has sustained their participation, especially for learners with the lowest fluency. The learning environment, because of its multimodal nature, enriches the learners' language repertoire. Indeed, learners may choose the modality by which they prefer to convey their message (rather by textchat than by oral, for example), or may use combined modalities. Thus, multimodal strategies can be described as scaffolding strategies which enrich the learner's communication repertoire (following Gumperz' perspective). They may be described as "compensatory strategies" when they sustain the learner's communication. When the multimodal strategies are negotiated by participants who are used to work together in sub-group, they are also characteristics of a new kind of communication, typical of these environments.

Then, the multimodal and collaborative aspects of the task help learners develop language strategies which contribute to enhance their language autonomy. We notice several occurrences in our corpus: learners know how to answer to their own questions (consulting authentic websites, online dictionary), they are able to compare their productions and to correct them in terms of accuracy and appropriateness, and to reuse spontaneously, in their own production, what they notice as "correct" in others' productions. Thus, the multimodality aggregated to collaboration and synchronicity constitutes a positive framework to enhance language autonomy.

7. Conclusions

This paper sets an original methodological approach, which aims at contributing to a better understanding of the specificities of multimodal communication in AGSE and its relevance to the development of the writing competence, supported by a writing process perceived as a complex and procedural activity and as a social event. It may be of interest to language teachers who want to design online collaborative writing activities intertwined with synchronous communication, as well as researchers who want to further investigate the relationships between writing and multimodal communication.

Our paper focuses on the writing process in L2 in an interactive integrated word processor and explores the way learners use the variety of modes and modalities displayed by the AGSE to perform a collaborative text. The communication practices of two sub-groups of false-beginners have been compared in two writing tasks of different degree of difficulty. We backed our methodological approach for analysing multimodal conversations on works coming from the Ethnography of Communication or Conversation Analysis where notions such as participant's perspectives and context become prominent to explain the communication process. We chose to analyse the learning process from the learners' and tutor's perspectives, from choices they made out of a set of available meaning-making resources, in a particular situation at a given moment. This conception of communication thus involves an appropriated coding, which enables the understanding of various semiotic resources as a meaning-making unit. Interpreting multimodality means then re-building the meaning given by the participants while communicating. Thus, multimodality cannot be studied as a static composite production.

Focusing on the specificity of writing in such a communication context, several trends have been noticed. Learners used the word processor in combination with others modalities, which highlights the strategic use of certain modes to facilitate the writing process. They also make consistent individual choices to participate to multimodal discourse, and collective choices. Thus, AGSE can be depicted as an environment of mutual monitoring possibilities. The analysis of the work in sub-groups shows that participants use more often multimodal strategies. Therefore, the "modality-switching" is a strategy frequently used by learners involved in a collaborative writing task. Moreover, the interactive writing process enhances the writing competence. Two types of scaffolding are then listed: *collaborative* scaffolding (writing by several hands, the correction of somebody's production, or after observation of somebody's production, etc.) and *multimodal* scaffolding which both encourage metacognitive strategies and autonomous strategies.

This paper has showed the interest for learners to practice an interactive writing while using a set of modes and modalities. Though, the current tendency in CALL is to overlook the writing competence when working in synchronous environment. Moreover, a general trend nowadays is to focus on the use of video to enhance the oral communication. Indeed, the use of video will not allow learners to benefit of the richness of such a communication environment. Still further experiments and researches are needed to complete the understanding of the

multimodal communication in collaborative tasks and its potential in terms of language learning, in order to promote a real appropriated communication in AGSE.

8. References

- Andrews, T. (1993) Videoconferencing: an interactive communication tool for distance learners. *ONCALL*, **8**(1), <http://www.cltr.uq.edu.au/oncall/andrews81.html>
- Baldry A. and Thibault P. (2006) *Multimodal Transcription and Text Analysis, a multimedia toolkit and coursebook with associated on-line course*. London, Equinox.
- Betbeder M.-L., Ciekanski, M., Greffier, F., Reffay C., Chanier, T. (2007) Interactions multimodales synchrones issues de formations en ligne: problématiques, méthodologie et analyses. In : Lamy, M-N, Mangenot, F. and Niessen, E. (eds.), *EPAL 2007*, Grenoble, France, 7-9 juin 2007. <http://w3.u-grenoble3.fr/epal/pdf/betbeder.pdf>
- Betbeder M.-L., Tissot R, Reffay C. (2007) Recherche de patterns dans un corpus d'actions multimodales. *EIAH*, Lausanne, Suisse, 27-29 juin 2007. <http://edutice.archives-ouvertes.fr/edutice-00158881>
- Chanier, T., Vetter, A., Betbeder, M.-L., Reffay, C. (2006) Retrouver le chemin de la parole en environnement audio-graphique synchrone. In : Dejean-Thircuir, C. and Mangenot, F (eds.), *Les échanges en ligne dans l'apprentissage et la formation, Le Français dans le Monde, Recherche et applications*, **40** : 139-150. <http://edutice.archives-ouvertes.fr/edutice-00084388>
- Chanier T. and Vetter A. (2006) Multimodalité et expression en langue étrangère dans une plate-forme audio-synchrone, *Alsic*, **9** : 61-101. <http://edutice.archives-ouvertes.fr/edutice-00001436>
- Dam, L., Legenhausen, L. & Wolff, D. (1990) Text production in the foreign language classroom and the word processor. *System* **18**, (3): 325-334.
- Dejean, C. and Mangenot, F. (2000) Gestion à deux d'une tâche informatisée d'écriture à contraintes. *Actes du colloque "L'écriture à contraintes"*, In : Bilous, D. and Oriol-Boyer, C. (eds.), may 2000, Université Stendhal-Grenoble 3.
- Goffman, E. (1964) The neglected situation. *American Anthropologist* **66**:133-136.
- Goffman, E. (1974) *Frame analysis: an essay on the organization of experience*. New York: Harper & Row.
- Goffman, E. (1983) The interaction order. *American Sociological Review*, **48** (1): 1-17.
- Goodwin, C. and Duranti, A. (1992) Rethinking Context: An introduction. In: Goodwin, C. and Duranti, A (eds), *Rethinking Context: Language as an interactive phenomenon*, Cambridge, CUP: 191-227.
- Halliday, M.A.K. (1978). *Language as Social semiotic: The social interpretation of language and meaning*. London: Arnold.
- Halliday, M.A.K. (1989) Part A. In: Halliday M.A.K. and Hasan R. (eds). *Language, Context and Text: Aspects of language in a social-semiotic perspective*. Oxford. Oxford University Press, 55-79.

- Hampel R. (2006) Rethinking task design for the digital age : A framework for language teaching and learning in a synchronous online environment, *ReCALL*, **18** (1):105-121.
- Hampel, R. and Baber, E. (2003) Using internet-based audio-graphic and video conferencing for language teaching and learning. In: Felix, U. (ed.), *Language Learning Online: Towards Best practice*. Lisse: Swets & Zeitlinger, 171-191.
- Hampel, R. and Hauck, M. (2004) Towards an effective use of audio conferencing in distance language courses. *Language Learning and Technology*, **8** (1): 66-82.
- Harri-Augstein S. and Thomas L. (1981) Comment devenir un apprenant : apprendre à apprendre. In : Holec H. (ed). *Autonomie dans l'apprentissage et apprentissage de l'autonomie, Etudes Appliquées de linguistiques*, **41** : 86-101.
- Jeannot, L., Vetter, A. and Chanier, T. (2006) Repérage des stratégies des apprenants et du tuteur dans un environnement audio-graphique synchrone. In : Dejean-Thircuir, C. and Mangenot, F (eds.), *Les échanges en ligne dans l'apprentissage et la formation, Le Français dans le Monde, Recherche et applications*, **40** : 151-161. <http://edutice.archives-ouvertes.fr/edutice-00087741>
- Jones, R. (2004) The problem of Context in Computer-Mediated Communication. In: Levine P. & Scollon, R. (eds). *Discourse & Technology, multimodal discourse analysis*, Washington, Georgetown University Press, 20-33.
- Jepson, K. (2005) Conversations – and negotiated interaction- in text and voice chat rooms. *Language Learning & Technology*, **9** (3): 79-98.
- Kress, G. (1997) *Before Writing: rethinking the paths to literacy*. London: Routledge.
- Kress, G., Jewitt, C, Ogborn, J & Tsatsarelis Charalampos (2001) *Multimodal Teaching and Learning, the rhetorics of the science Classroom*. London: Continuum.
- Kress G. and Van Leeuwen T. (2001) *Multimodal Discourse: the modes and media of contemporary communication*. London: Arnold.
- Lamy, M-N. (2006) Conversations multimodales : l'enseignement-apprentissage à l'heure des écrans partagés. *Revue Le Français Dans le Monde, Recherche et Applications*, **40** :129-138.
- Lamy, M-N. (2007) Multimodality in Online Language Learning Environments: Looking for a methodology. In: Baldry, A. and Montagna E. (eds.) *Interdisciplinary Perspectives on Multimodality: Theory and practice*. Proceedings of the Third International Conference on Multimodality. Campobasso: Palladino : 237-254.
- McAndrew, P, Foubister, S.P., and Mayes, T. (1996) Videoconferencing in a language learning application. *Interacting with Computers*, **8**(2): 207-217
- Norris S. (2004) Multimodal Discourse Analysis: A Conceptual Framework. In: Levine P. and Scollon, R. (eds). *Discourse & Technology, multimodal discourse analysis*, Washington, Georgetown University Press, p. 101-115.
- Scollon, R. and Levine, P. (2004) Multimodal Discourse Analysis as the Confluence of Discourse and Technology. In Scollon, R. and Levine, P. (eds), *Discourse & technology, multimodal discourse analysis*. Washington D.C., GUP.
- Stevens, V. and Altun, A. (2002) The Webheads community of language learners online. In: Syed, Z. (ed.) *The process of language learning: An EFL perspective*. Abu Dhabi: The Military Language Institute, 285-318.

- Stockwell, G. (2007) A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL*, **19** (2): 105-120.
- Vetter A. (2004) Les spécificités du tutorat à distance à l'Open University, *ALSIC*, **7** : 107-129. http://alsic.u-strasbg.fr/v7/vetter/alsic_v07_06-pra2.htm
- Vetter A. and Chanier T. (2006) Supporting oral production for professional purpose, in synchronous communication with heterogeneous learners, *ReCALL*, **18** (1): 5-23. <http://edutice.archives-ouvertes.fr/edutice-00080316>
- Wang, Y. (2004) Internet-based desktop videoconferencing in supporting synchronous distance language learning. *Language Learning and Technology*, **8**(3): 90-121.