Evaluating evaluation as a communication process. What role for formative evaluation in ICT-based knowledge acquisition?

Philippe USEILLE
Laboratoire des Sciences de la Communication
Université de Valenciennes et du Hainaut-Cambrésis
Le Mont-Houy 59313 Valenciennes Cedex 9 France
philippe.useille@univ-valenciennes.fr

ABSTRACT

This article examines how formative evaluation as a communication process contributes to knowledge acquisition in using ICT (Information and Communication Technologies). Previous studies, especially in the field of education and training, have shown that formative evaluation plays a crucial part in the learning process because it contributes to learning to learn. Through formative evaluation, the learner becomes aware of errors and can adjust learning strategies to the situation. In addition, formative evaluation provides the teaching side with significant and useful information.

Consequently, ICT researches have developed a wide range of solutions for this specific purpose. It is however difficult to check the efficiency of these tools by considering the effects of ICT in the knowledge acquisition process. I suggest that formative evaluation includes also a communication system that has an effect on the learning process.

This study tackles the issue by proposing an alternative approach to formative evaluation that considers it as both a learning and a communication process. The study is based on SADT (Structure Analysis and Design Technique) that provides a suitable description for the whole complex communication process. It allows a rigorous understanding and identification of the variables of evaluation as a communication process in order to take care of an ICT frame.

Finally, this article outlines a multidisciplinary method to evaluate formative evaluation by focusing on the validity facets of the communication process.

Keywords: formative evaluation, communication process, validity criteria, ICT training context.

INTRODUCTION

Evaluation is clearly at the heart of the instruction process. In the field of education, the different purposes of evaluation have been considered; the criteria of a valid evaluation have been also explored. More precisely, formative evaluation has been studied as a way to improve teaching/learning performance.
Why did I study evaluation function implemented by ICT tools ? The last years has seen the increasing use of computer technology for the development of evaluation, more precisely in the field of language tests. For example, Web-based tests are popular because learners from anywhere in the word can log on them and receive quick instant feedback on their performance. Among them, DIALANG [1], a web-based system, is available in 14 languages and is intended for placement and diagnostic purposes. It tests different skills, and feedback is provided to learners in terms of their levels as defined by Council of Europe. The European multilingual context increases these kind of products. In addition, the potential for using computers instead of humans is considerable and could significantly reduce the costs [2]. Given this, I consider that I need to investigate evaluation, and more specifically formative evaluation, as a communication process that includes ICT dimensions.
The understanding of this is of primary importance not only for pedagogic studies but also in the ICT field, especially regarding the production of information about learner performance, the adaptation of the teaching role, and their interactions through a complex process
Although much research has been focused on different aspects of formative evaluation but not as a communication process. The present work addresses the issue and presents a method for analysis of formative evaluation, step by step, which considers its specific purposes. It takes into account not only ICT aspects but also human or material ones which are involved in the process. This method should serve as a basis for further investigations in ICT applications and should be useful to improve its efficiency. This paper is organized as follows. The first section contains a presentation of the characteristics of formative evaluation. Then, in the following section, the method of analysis is described. In the last section, the method is applied to a few examples and discussed to see further developments.

1. BOTH SIDES OF FORMATIVE EVALUATION

The word evaluation is derived from value, and in its most basic sense simply means the "process of assessing the worth or value of something" [3]. In this sense, it is possible to say that this covers all the procedures that specify whether the training program manages to be relevant to identified needs. All training programs require evaluation not only as part of a teaching/learning process but also as a criterion of quality of the process.
There is a wide range of possible activities that is involved in evaluation. However, one traditional way of studying pedagogical evaluation has been to classify it into three general types, notably:

- "summative evaluation" corresponds to what the learner has learned at the end of the program. This is past-focused. Often, quantitative results allow comparing the learner with programs standards. It is like a "rear-view mirror".

- "prospective evaluation" is future-focused. It takes place at the beginning of the training and gives data about the learner's level compared to the prerequisites of the training. In this way, the appropriate resources are suggested to the learner. This is a kind of "window" through which the best suited training is determined.

- "formative evaluation" is a continuing action which allows regulation of training. Focused on the present, this measures the acquisition and errors in order to inform and adapt teaching and learning to each other. This is aimed at helping learners achieve the aims of the training in their own individual way. It does not consist necessarily in grades. It is used for making a diagnostic. It is like a self-reflecting "mirror" of training.

evaluation types	focus	function
summative	past	Rear-view mirror
prospective	future	window
formative	present	Mirror

Figure 1 : different types of evaluation

Let us mention that there are other concepts defining various kind of evaluation such as "alternative assessment", "authentic assessment" or "performance-based assessment. These studies point out that evaluation may be mixed with the training process: it suggests what could be formative evaluation [4].

Formative evaluation has aroused considerable interest and attention in the education field because it is designed to feed the whole teaching/learning process. I would like to focus specially in the ICT area, where formative evaluation is crucial for learners. A survey I carried out showed that learners consider it as «very beneficial » [5]. It may prevent learners from dropping out when ICT are used as self-training tools. In this case, it appears that formative evaluation is a driving force of the whole teaching/learning process. This topic could generate questions such as:

- What makes a good error diagnostic?

- How do we drive the learner to improve performances?

- How do we design good technical supports to work out a relevant evaluation?

- What could be the automatic part and the non automatic part in the evaluation process?

- How do we isolate what is wrong when formative evaluation does not work as we planed?

It is to be noted that these questions are related to my personal professional experience: I teach French as a foreign language by using ICT. It provides important research questions that I would like to address by studying examples from this area. I have deliberately made these questions of very different types in order to show the range of questions that "formative evaluation" can convey specially in an ICT environment. This also shows that they are very problem-focused in their approach and very practical in their intended outcomes. This is, in my view, a case for also having available as a source for our reflection a systematic approach and technique which help me to answer these questions and take good decisions about the suitable tools I need. The whole of the present paper is essentially an exploration of one way in which a reflection about formative evaluation can be managed by using appropriate analysis tool of investigation.

Let come back to these questions; they cover different "problem areas":

- Problems of the functions formative evaluation performs: error diagnostic, improving learning;

- Problems of the principles formative evaluation refers to what makes a valid formative evaluation;

- Problems of the adapted means devised to perform it.

- Problems of the nature of data produced by formative evaluation mediated by ICT.

Some of these can be investigated in two ways:

The first approach concerns the way in which the pedagogic approach specifies the requirements of the contents of the training (the syllabus) and the teaching/learning options. The features of the contents are likely to inform the whole teaching/learning process (which includes evaluation); it is so significant that it has to be tackled. For instance, teaching/learning a language differs from mathematics. French has the characteristics of a language and the specific features of its history, the culture it represents, a different way to see the things. Teaching/learning French language can vary deeply according to the learner cultural profile. What is related to the contents can provide us with information of great importance about the functions and the validity criteria. This approach might, however, be partly beside the point to answer some of these questions.

The second approach examines this in a different way. The point is that formative evaluation is defined as a communication process. This separates and identifies information and activities which take part in the process. I also maintain that it is related to communication. It means that the functions are connected and each one informs the other. This involves transformations of data into information and circulation of them through a complex system. What do data and information mean? Data refers to all things selected or generated by tools for social purposes as formative evaluation. It becomes information when they are given meaning. These fundamental concepts would need deeper exploration as does Marcia BATES [6].

This method enables us to go into all the details of process which may contain a wide range of machine generated tools or not, media, software: they may have an influence on formative evaluation as a communication and information process for the learner. In this second approach, I could examine another of the validity of the whole process of communication in order to know if the evaluation is formative or not. I wonder if the learner is able to give meaning to the delivered data. .

Both, I argue, are useful and I thus combine different kind of validity. In so doing, as well as focusing my attention on the communication aspects and taking teaching/learning features into account. This allows one to understand what is on line in an evaluation process and get the point in relevant way

A functional analysis like IDEF0 (Integration Definition language 0) method, in my view, provides a suitable description of the phenomenon of formative evaluation.

2. UNDERSTANDING FORMATIVE EVALUATION WITH IDEF0 METHOD

I decided to face with formative evaluation seen both as a teaching/learning and communication and information process. Therefore, I need a method which is designed to explore complex systems like the evaluation process. IDEF0 is likely to describe such a process (cf. http://www.idef.com/idef0.html). But let us briefly mention the essential features of this approach. It provides a structured representation of the function, activities or processes within a

modeled system. It performs systems analysis and design at all levels, for systems composed of people, machines, material, computers and data of all varieties so that it takes into account all the aspects of evaluation process, particularly in an ICT environment. This an appropriate technique for separating the various aspects of the complex evaluation process. Based on SADT (Structured Analysis and Design Technique) developed by Douglas T. Ross and SofTech, Inc. this technique consists in a graphical modeling language (syntax and semantics). In this paper, IDEF0 is used to facilitate the analysis of the functions the evaluation system performs by modelizing it.

The model consists of hierarchical series of diagram, text, and glossary: the functions are represented on the diagram by boxes and the data and objects that inter-relate those functions are represented by arrows. The boxes provide a description of what happens in a function. A box has a name which is an active verb or verb phrases. Each side of the function has a meaning in terms of box/ arrow relationships. Arrows entering the left side of the box represent inputs: they are used by the function to produce outputs. Arrows link to the top of the box are controls: they specify the conditions required for the function to produce good outputs. Arrows leaving a box on the right side are outputs: they are data produced by the function. Arrows connected to the bottom side of the box represent mechanisms. They are necessary to support the execution of the function. The diagram looks like the following example:
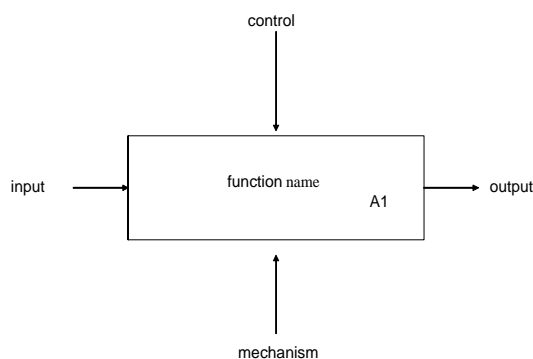

Figure 2: Representation of box-and-arrows.

In the following section, I will apply the IDEF0 Model to formative evaluation, especially in an ICT context.

3. IDEF0 MODEL OF FORMATIVE EVALUATION IN AN ICT CONTEXT

The whole graph (figure 3) represents the model of formative evaluation as a communication process. It needs to be pointed out that I cannot draw up the exhaustive IDEF0 graph in this paper. If this were the case, it should be made up of various levels connected together. Moreover, I point out that same arrows may be linked to different boxes… This means that IDEF0 graph is relevant for outlining very complex objects. To simplify it, I have excluded the inter-connections and have insisted on the leading features.

IDEF0 illustrates a certain point of view about the subject. At first, I have isolated the phenomenon within the following boundaries: from the production of an evaluation task to the prescription of an action to correct the teaching/learning process. I do not mention the elements that are outside even if

they can be significant. Secondly, our viewpoint has determined what can be seen within the model: I intend to draw attention to communication aspects of the process such as information using, teaching/learning requirements. All the examples I give come from the area of teaching/learning French as a foreign language as found on the web.

I propose a graph that stands for the whole formative evaluation in ICT (fig. 3). This graph can be broken into five main parts related to the different functions represented by the boxes. They represent the different phases of the process. Each box located on the right side provides data to the left one.

This graph can be broken into five main parts related to the different functions represented by the boxes. They represent the different phases of the process. Each box located on the right side provides data to the left one.

The first step of the process consists in the production of evaluation performance. It has to sum up the most important skills and abilities the syllabus plans. It is controlled by teaching/learning principles which define the validity of the task. It has to produce what it is supposed to evaluate. A valid test of writing comprehension will test writing comprehension and not reading skills.
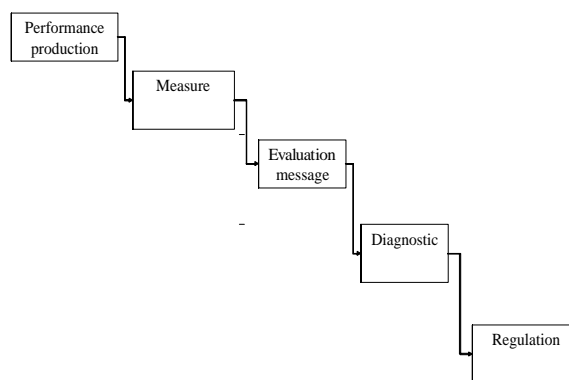

Figure 3: Representation of the whole process

I drew a basic and simplified graph which does not respect all the IDEF0 syntax but it shows the essential steps of the process. It is a "top-level diagram" providing the most general or abstract description of the subject. This diagram may be followed by a series of "child diagrams" providing more details about subject. I have developed them in our thesis.

The first step (figure 4) of the process consists in the production of evaluation performance: it has to be related to the most important skills and abilities the training plans or others linguistic learning aspects. It is controlled by teaching/learning principles that define the validity of the task. It has to produce what is supposed to be evaluated. A valid test of writing comprehension will, for instance, only test writing comprehension and not reading skills by summing up the essential acquisitions. The evaluation task may also reconstruct the original ones to simulate a real situation.

teaching/learning
options

↓

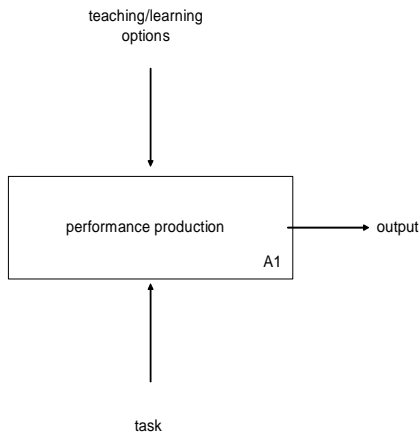| performance production |
| --- |
| A1 |

→ output

↑

task

Figure 4: performance production

This first box provides the measure function with input coming from learner performance (fig. 5). They may be assessed in a quantitative or qualitative way. This operation is controlled by the validity of measurement tool and the cultural conventions of what is possible to measure or not. The measure tool which supports this function is affected by the features of the task. Very complex performances such as an essay can not be measured but only appreciated in qualitative terms. I notice that designing technical tools for quantitative measurements is easier than for qualitative ones, particularly in ICT environments. Consequently, this would exclude the more complex activities from ICT area.

Teaching/learning
validity

Technical validity

↓

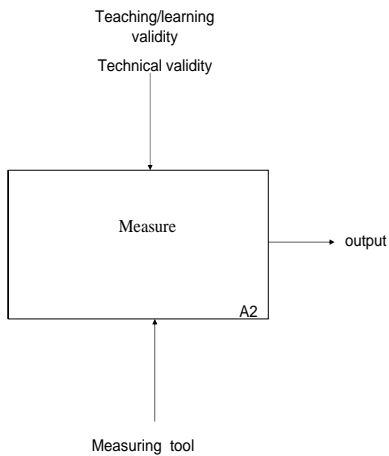| Measure |
| --- |
| A2 |

→ output

↑

Measuring tool

figure 5: measure

Afterwards, the third function consists in evaluation message production. I can make the best of IDEF0 modelling by showing a part of the process that is often mixed with the former function: providing the learner with evaluation information. IDEF0 enables us to specify the characteristics of the message: it may be a text, a graph or every kind of admitted codes (the mechanism). In ICT environments, these codes are not stable. There are, I mean, lot of varieties of available codes that are in use. They have to be perfected according to semiotic and ergonomic standards. This function is controlled by the logical and pragmatic rules of assertion (a performance is, for example, told to be "right" or not). It means that the message must be pertinent and relevant. Admittedly, the more cognitively efficient is a message, the more it is pertinent [7]. I make much of it in an ICT context.

teaching/learning
validity

message validity

↓

evaluation message
A3

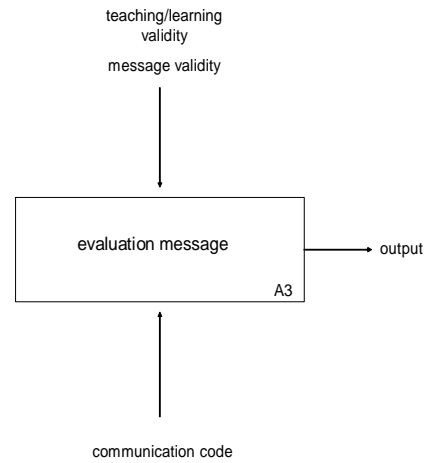→ output

↑

communication code

Figure 6: evaluation message

The fourth step is one of the specific parts of formative evaluation: the diagnostic. It consists in proposing hypothesis to explain the performance of the learner. According to teaching and learning principles, the performance is related to the actual learner's skills: this relation is acceptable as long as it can be justified by scientific or empiric hypothesis in the teaching/learning area. ICT environment does not allow large scale of errors automatic analysis so it is necessary to think about human support to do this. The diagnostic message can also be expressed in various conventional codes and have to be pertinent to be efficient. Its means that the message put down learner's errors to intelligible causes so that he can get the point.

teaching/learning
validity

errors analysis validity

↓

diagnostic
A4

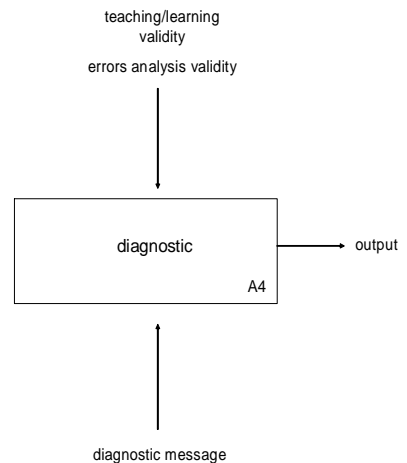→ output

↑

diagnostic message

Figure 7: diagnostic

The last step I identify in this paper consists in the production of instructions to lead the learner or the teacher to correct the process. This regulation can be machine generated or performed by any efficient means. This has to respect as well pragmatic rules (as pertinent ones). The mechanism is a message using a certain code.
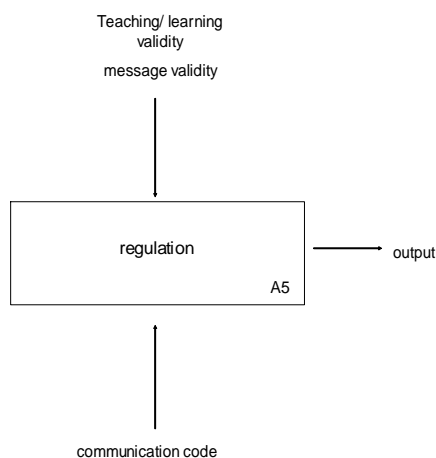
Figure 8: regulation

The whole diagram obtained by using IDEF0 shows that formative evaluation is a complex communication process. This model reveals different implicit hidden variables which support this process:
- Teaching/learning validity
- Pragmatic pertinence or message validity
- Tool dependence
- Potential conflicts between these three variables

Thus, the IFEF0 model suggests studing the different parts of the process, the integration of these functions and the validity of the whole process.

4. ICT APPLICATIONS OF FORMATIVE EVALUATION

Let us apply IFEF0 model to few examples coming from the foreign language the teaching/learning area (especially French as foreign language). In this section, I have collected data and classified them in a table.

− Example 1: http://www.laits.utexas.edu/fi/tv/index.php
The useful web site of the Austin University proposes evaluation activities for students who learn French. The method is very simple: "In each exercise, you will answer fill-in-the-blank questions. When you have answered all the questions on a page, click the Submit button. For answers submitted, you will be presented with feedback indicating suggested answers. You may print results, if you wish". I took the example of the verb to be.

function	*example*	*Ccontrol*	*mechanism*
performance	writing the correct form of the verb to be	french grammar	Drill exercice
measure	none	none	none
evaluation message	none or maybe implicitly	none	none
diagnostic	none	none	none
regulation	none	none	none

Figure 8 : table of data collected in example n°1

What kind of evaluation is it? This web site section dealt with evaluation (which is called "testez-vous") does not refer to the lesson section. It gives no mention about the type of evaluation. That is why the cases which are located in the right side of the table are empty. The evaluation process is here partly performed. Even the measure function is not explicit: the learner is supposed to compare the performance with the correct answers. In some cases, we have observed learners repeating the evaluation task until they carry out to do it (they did not in fact) [3]. It would have required an adapted understanding of the errors. It means that kind of evaluation requires an expert learner or some extra-help to achieve the process.

− Example 2: http://fog.ccsf.cc.ca.us/~creitan/grammar.htm
The web site of the San Francisco City College Language Centre allows learner to practice grammatical skills in the following way:

function	*example*	*control*	*mechanism*
performance	MCQ	french grammar	MCQ tool
measure	none	none	none
evaluation message	symbols	no specify	software
diagnostic	yes	related to the contens	software
regulation	yes	Pragmatic rules	text

Fig. 9: table of data collected in example n°2

It is not yet formative evaluation through lack of diagnostic and regulations steps. In comparison with the former example, the measure is followed by an evaluation message: it is only quantitative because of the characteristics of the MCQ tool that fits in with this approach. This shows complex interrelations between, on the one hand, functions, and, on the other hand, control and mechanism: the first can be transformed by the others. The machine generated correction provides right answers without any justification. The score does not give any accurate data about the learner level. I may say that these very quantitative data are not really significant: they have to be interpreted.

− Example 3: http://www.longman.co.uk/testbuilder2/template.asp?testid=40

function	*example*	*control*	*mechanism*
performance	MCQ	french grammar	MCQ tool
measure	machine generated and quantitative	mathematic validity	software
evaluation message	mark and percentage of success	mathematic validity	numeric code

function	example	control	mechanism
diagnostic	none	none	none
regulation	none	none	none

Fig. 10: table of data collected in example n°3

The Longman British editor proposes these two characteristic features of formative evaluation: diagnostic and regulation. It manages to do it by both identifying the learner mistakes and providing a relevant regulation. Let us notice that errors are only related to teaching/learning contents and not to other aspects as the ICT environment. I have postpone that the mechanism is not the function and, so, it can not manage completely to perform it. Moreover, as this regulation is automatic, it does not take care of the suitable moment the regulation has to be done. It could be too early or too late in relation to the teaching/learning process [8]. Nevertheless, there is no doubt that this tool attributes some meaning to learner answers. Let us notice that this site resorts to specific software that has to be able to analyse the possible learner's errors, according to a certain teaching/ learning point of view.

Finally, the following conclusions can be drawn from the IDEF0 modelling of evaluation, particularly formative evaluation in ICT environment:

- Separating the functions of formative evaluation help us to specify eventually the learner's errors diagnostic: it does not only concern teaching/learning contents but may be due to the mechanism and the control which are involved;

- The quality of formative evaluation using ICT is affected by various kind of validities such as :

  The teaching/learning validity controls the evaluation task, the measure function, the diagnostic function, the regulation function;

  I can add the tool validity that concerns mainly the mechanisms (task, measure, diagnostic for example) which can influence the whole process, especially in an ICT context.

  The global communicational validity in ICT context which mainly consists of semiotic codes and pragmatic pertinence and the interactions between various kind of validity ;

These aspects belong to different analysis levels. Each one may be significant and does not necessarily match. There is not any obvious correlation between these levels. In consequence, I can either attempt to make it harmonious, or I can acknowledge a certain diversity of the variables which are part of a complex construction which is related to the user's situation. Does it make sense? By this I mean the do learners elaborate their own construction of information that is given by the formative evaluation context process. It is up to the learner to answer. We are carrying out further studies to investigate learner's responses to this process and how he/she reacts to.

5. CONCLUSION

As a conclusion, I think that defining formative evaluation as also a communication process is relevant in the case of an ICT context. This approach leads us to apply an appropriate model such as IDEF0: which can outline the process step by step. Moreover, it enables us to go into the specific relations between different variables and to collect significant data in a systematic way. This method can help us to evaluate evaluation in an ICT context. However, the model has to be confronted with the learner's point of view. This needs to be investigated in further studies.

6. REFERENCES

[1] DIALANG : http://www.dialang.org/english/index.htm
[2] G. FULCHER, "Machines get clever at testing", **The GUARDIAN**, Thursday May 17, 2001. Dr FULCHER is a researcher in the second language testing field
[3] M.J. Wallace, **Action Research for Language Teachers**, Cambridge University Press, 1998.
[4] M.B. Kane, N. Khattri, A. L. Reeve & R. J. Adamson, **Assessment of Student Performance: Studies of Education Reform**, U.S. Department Education, Office of Educational Research and Improvement,1997.
[5] P.Useille, L. Verclytte, M. Labour, **Créer du sens par l'évaluation de la lecture de documents hypermédias**, Actes de H2PTM'03, Hermes-Science, Lavoisier, Paris, 2003
[6] M.J. BATES**,** Information and knowledge : an evolutionnary framework for information science, **Information Research**, Vol. 10 No.4, July 2005.P
[7] D. Sperber, D. Wilson, **Relevance Communication and Cognition**, Backwell,1986.
[8] L. Verclytte, **Le sur-mesure au moment opportun : Vers l'élaboration de documents d'accompagnement en situation d'acquisition de connaissances**, Unpublished doctorat thesis, 2002.

web sites consulted in may 2006 :
IFEF0 method : http://www.idef.com/idef0.html
the Austin University web site:
http://www.laits.utexas.edu/fi/tv/index.php
the San Francisco City College Language Centre web site:
http://fog.ccsf.cc.ca.us/~creitan/grammar.htm
web sites consulted in may 2004 :
The Longman Editor web site:
http://www.longman.co.uk/testbuilder2/template.asp?testid=40