A Genre-Based Perspective for the Development of Communicative Computerized Adaptive Tests

Javier López-Cuadrado
Department of Computer Languages and Systems, University of the Basque Country (UPV-EHU). Spain // javilo@ehu.es

Ana J. Armendariz
Department of Computer Languages and Systems, University of the Basque Country (UPV-EHU), Spain // jiparlea@si.ehu.es

Marion Latapy
Laboratoire d'Informatique de l'Université de Pau et Pays de l'Adour, France // Marion.Latapy@iutbayonne.univ-pau.fr

Philippe Lopistéguy
Laboratoire d'Informatique de l'Université de Pau et Pays de l'Adour, France // Philippe.Lopisteguy@iutbayonne.univ-pau.fr

ABSTRACT
This paper discusses the communicative potentials of Computerized Adaptive Testing. The study is based on a model that offers a set of independent communicative concepts to describe the Genre of an interactive application. This model will be the starting point to analyze the stages of the interaction cycle that are typically inherent to every generator of computerized adaptive tests. In each of these stages, the authors present the most appropriate communication structures that should support the intentions of the developers of this kind of software. The paper finishes giving a synthesis of the interactive potential of these communication stages from a Genre perspective, and discussing the compatibility between the proposed approach and some model-based methods that are currently used during the design of interactive applications.

Keywords
Genre, Communication structures, Interactive applications, Software development, Computerized adaptive testing

Introduction
The systems that use Computerized Adaptive Tests (CATs) aim to measure the maturity or knowledge level of a student in a specific domain by emulating the intelligent behavior of human evaluators. In other words, these tests dynamically apply the questions that are more informative for that purpose (Wainer, 2000). The authors have some experience in this area within the language learning domain, particularly in the Basque language field (López-Cuadrado, Armendariz, & Pérez, 2005; López-Cuadrado, Pérez, Arruabarrena, Vadillo, & Gutiérrez, 2002).

The issues concerning the administration of items at run-time, which depends on previous responses, are well documented in literature. For instance, topics based on experimental protocols and psychometric measures such as stabilization of domain criteria, identification of existing dependences among them, classification of items before they feed the item bank, test validity, precision and security, item selection rules, and scoring methods are first and foremost treated in research and development projects (Ólea & Ponsoda, 2003; van der Linden & Glas, 2000; Wainer, 2000). However, there is little work centered on the communicative dimension of CATs, in other words, how the successive stages of the assessment activity communicate with the receiver, that is, the examinee.

This paper works on a new a research line to propose guidelines to make easier the explicit description of the communicative dimension. This component, which belongs to the human-computer interaction, distinguishes between vertical interaction, i.e. studies of the dynamic offered to the final user, and horizontal interaction, i.e. studies about the presentation and also about the manipulation of the user-interface (Nogier, 2005).
2003), which provide a scenario specification for interactions; and the latter is derived from graphic charts and ergonomic patterns (Lynch & Horton, 2005). The resulting interactive system is generally stabilized by means of participative and iterative approaches combined with prototyping techniques.

In the particular case of CAT generators, the vertical interaction is accepted and established by consensus (Thissen & Mislevy, 2000), as will be shown in section 3 by an algorithmic representation of their behavior. Actually, the horizontal interaction is the one that gathers less results and specific recommendations. Precisely, the contribution of this paper lies in this field.

While specifications of a horizontal interaction are basically centered on graphic charts and ergonomic patterns, this paper proposes the Genre model as perspective for complementary studies. Genre theory states explicitly that the development of a software product within a specific Genre – for instance, educational – makes the author to use communication structures – such as an exercise or course – that the receiver cognitively associates to the idea he or she has about the specific Genre (Latapy, Lopisteguy, & Dagorret, 2004). Schematically, this theory says that communication between transmitter and recipient is built by conventionally admitted communication structures, so when those structures are chosen adequately (that is, when they cover the intentions of the sender), this approach improves the acceptance of the product by the final receiver.

This paper is divided into seven parts. Next section introduces the concept of Genre, and proposes a generic model that allows the Genre specification for interactive applications. This model is built based on communication concepts that are derived from the Genre theory. The third section provides the background on computerized adaptive testing and presents the main features of CAT systems. In the following section a CAT generator, called GenTAI, is presented as a particular case of this kind of interactive application. Using GenTAI as basis, the communication stages that feature the vertical interaction of any CAT system are identified. Sections fifth and sixth are dedicated to analyze the communicative potentials of every of these stages by means of the most appropriate communication structures. Finally, the last section presents the conclusions and describes how this original approach to the concept of interaction can be included within a CAT system production process.

After giving a brief historic overview of the Genre, from its appearance to its recent consideration in interactive software production, next section will introduce a Genre model that integrates the most relevant concepts to analyze the features of a given application from a Genre perspective.

The fundamentals of the Genre concept: a model for Genre specification

The concept of Genre is a dimension culturally inherent to any communication activity. A good example of this is the fact that in 384 B.C. Aristotle based his theory on the role of the language in communication and creation. According to the movement of the Formalists, in the literary field, Genres differ thank to their dominant feature, which is related to the functions fulfilled by the created work (Canvat, 1998). This position leads to consider that a Genre is determined by a communication intention (i.e. a communication goal that guides the transmission of information). Moreover, Canvat (1998) considers Literature as a coherent system and defines the concept of Genre as a structure of social communication. This structure takes into account the situation of communication between the author and his or her audience. A situation of communication includes elements related to the message, such as the position and visibility of the transmitter inside the message, the implication marks destined to the recipient, the contents chosen, the channels used, codes, implemented languages, and so on. Therefore, one can say that Genre leads to the establishment of communication rules that control the communication exchanges between interactors, that is, transmitters and recipients.

As a result of that, and after applying the Communication theory to different media (other than written text), researchers in the Genre domain state that productions belonging to one Genre make possible to answer the expectations of any kind of recipient: readers, spectators, and users (Canvat, 1998; Chandler, 2000). Additionally, users enjoy while observing the way in which an interactive application handles, mixes and makes the communication conventions evolve, even reaching non-expected situations for the concrete Genre in which the application is supposed to stand.
The concept of Genre also changed towards the area of Human Computer Interaction. More particularly, the authors point out research on the role of Genre in understanding interaction (Finlay, 2004) and the study of Genre as a structuring concept for interaction design (Pemberton, 2000). So, the concept of Genre evolved from Aristotle to the Computer Science era, and it remains characterized by a dominant feature: its communication intention. Nowadays Genre is not simply a set of rules, but a framework that conceptualizes reality by defining and including communication conventions, expectations, structures and interactors, which are the fundamental molds in which communication is cast (Bakhtin, 1986).

The present work, which participates to these purposes, is based upon the three following points:
- Genre is a concept inherent to any communication activity;
- Genre defines structures of communication;
- Mixing Genres conveys creativity. Indeed, innovating works often result from a mixture of communication structures that belong to distinct Genres. Ludo-pedagogical applications, that combine pedagogical structures and games, are examples of such mixtures.

Therefore, the concept of Genre provides all interactors with well known and accepted communication structures, and facilitates for transmitters the implementation of their communication intentions. The proposed model (Figure 1) provides designers of interactive applications with a language for Genre description. It results from a survey, briefly presented above, at three levels of abstraction: a theoretical level (Genre theory), a specific Genres level and a products level (Latapy, 2005).

![Figure 1 – Model of Genre](image)

An application deals with a field which can be described either formally (e.g. ontologies) or informally. The Genre specificity of an application is defined by a communication intention, that can be considered as expressing the justification of the system usefulness, and associated communication structures.

The interactors (actors of the interaction) communicate within the exclusive framework of activities. They alternatively play the role of transmitter and receiver in an interaction. For example, the author of an interactive program (sleeping partner) plays the role of the sender via the application, as does the user when he or she handles it; and the user who receives information through the application is the receiver, just like the application when it gets information from the user.

The context represents the situations in which the application is used. It is defined in terms of place, moment, necessary resources and rules (social, cultural, historical...).
The communication structures concern rhetoric, and they participate in the implementation of the intention. On one hand, among these structures, one can stress rhythm, analogy figure and utterance that result from traditional rhetoric. They facilitate some aspects of the communication form. On the other hand, the exchange structures are dynamic communication structures that describe the actions allowed to interactors.

The exchange structures partly result from the verbal interactions theory. Table 1 suggests some of the communication structures and their objectives, though it does not make up an exhaustive list.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>selection</td>
<td>lead an interactor to select relevant information</td>
</tr>
<tr>
<td>resolution</td>
<td>lead interactors to identify, analyze and clarify a given problem</td>
</tr>
<tr>
<td>interrogation</td>
<td>lead the instigator of the exchange to ask a question and the receiver to formulate a response</td>
</tr>
<tr>
<td>discussion</td>
<td>lead interactors to introduce, justify and accept or deny arguments about a particular subject</td>
</tr>
<tr>
<td>notification</td>
<td>lead interactors to inform each other</td>
</tr>
<tr>
<td>coordination</td>
<td>lead the interactors to coordinate their activities, respecting a mode which can be direct supervision, mutual adjustment or standard</td>
</tr>
</tbody>
</table>

Rhythm allows to express an idea with a certain insistence, a certain periodicity, thanks to a suitable use of the communication elements to be transmitted by various media. Rhythm can be defined as continuous stress organization of discourse. Binary, ternary and enumerated are particular rhythms (Table 2).

<table>
<thead>
<tr>
<th>Rhythm</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>binary</td>
<td>double association of communication elements, which have the same function</td>
</tr>
<tr>
<td>ternary</td>
<td>triple association of communication elements</td>
</tr>
<tr>
<td>enumerated</td>
<td>accumulation of same level communication elements, which are not necessarily coordinated</td>
</tr>
</tbody>
</table>

Analogy figures describe forms which make closer or identify two realities belonging to separate fields. Metaphor, comparison and personification are particular analogy figures that are described in Table 3.

<table>
<thead>
<tr>
<th>Analogy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>metaphor</td>
<td>it brings implicitly together two communication elements with at least one common characteristic</td>
</tr>
<tr>
<td>comparison</td>
<td>it brings explicitly together two elements of communication with at least one comparable characteristic</td>
</tr>
<tr>
<td>personification</td>
<td>it assigns person capabilities to an inanimate element, such as an object or event</td>
</tr>
</tbody>
</table>

Table 4 – Essential enunciative modes.

<table>
<thead>
<tr>
<th>Enunciative mode</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>narrative</td>
<td>describe events</td>
</tr>
<tr>
<td>explanatory</td>
<td>analyze a phenomenon or an idea, to be well understood. It gives explanations, details. Generally speaking, it includes technical elements.</td>
</tr>
<tr>
<td>descriptive</td>
<td>produce an image the receiver does not see but can imagine. The descriptive communication makes possible the receiver to imagine a concept, a place or a person aspect. It informs by providing qualifiers, indications of place...</td>
</tr>
<tr>
<td>injunctive</td>
<td>the sender enforces the receiver to do something. The injunctive communication proposes an action. It is characterized by the presence of orders, advice or prohibitions.</td>
</tr>
<tr>
<td>argumentative</td>
<td>convince, persuade or prove something. It defends a thesis, by giving arguments or examples.</td>
</tr>
</tbody>
</table>
**Utterance** describes a form of discourse in the communication (textual, visual or aural). The utterance is made up of an *enunciative authority* and an *enunciative mode*. The enunciative authority defines the author of the statement. The enunciative mode defines the form of the statement, expressed by various media. The *narrative* one, the *explanatory* one, the *description* one, the *injunctive* one and the *argumentative* one are particular enunciative modes. Their purposes are detailed in Table 4.

During activities, documents dealing with the field are handled. Actually, in this context a *document* is any set of information recorded on a support that allows its use.

An *interaction scenario* describes activities, which can be simple or recursively constructed, to be carried out by *interactors* in a *context*. An interaction scenario satisfies an *intention*, and uses and organizes the *communication structures* to fulfill this intention through *activities*. In fact, the activities support the tasks performed by the application.

The aim of this paper is to support the integration of the communicative dimension into CAT designing and development processes, thanks to the explicit consideration of the recently presented concepts. Next section gives an overview of the main concepts regarding computerized adaptive testing, while the following one will present GenTAI, a specific CAT generator, and will give a deeper view of the Genre model in terms of this software.

**The fundamentals of Computer Adaptive Testing**

Computerized Adaptive Tests (CATs) emulate the intelligent behavior of human evaluators (Wainer, 2000). In fact, they dynamically select and administer the most appropriate items or questions depending on the previous responses given by the examinees (i.e. those that really provide useful information about student’s ability). One can find different theories that support the generation of CATs, as the Sequential Probability Ratio Test (Reckase, 1983), the combination of granularity hierarchies and Bayesian nets (Collins, Greer, & Huang, 1996) or the Measurement Decision Theory (Rudner, 2002). However, the most used framework for computerized adaptive testing is the *Item Response Theory* (IRT – Lord, 1980), an item-oriented background that offers models that associate the ability of the examinee with the probability of a correct response. To choose the proper item from the bank, the CAT algorithm needs to know the values of some psychometric characteristics (called *parameters*) that feature the items. In terms of the IRT this means that the item bank must be calibrated according to some model, which can be one-dimensional or multidimensional, depending on the number of factors or traits that are measured by the test, dichotomous or polytomous, depending on the considered number of possible response categories, and logistic or normal, depending on the function used to formulate the relationship between examinees and correct responses. The most used IRT models are the 1-parameter model (Rasch, 1960), which characterizes an assessment item by its difficulty, an the 3-parameter model (Birnbaum, 1968), which manages items’ discriminative power and guessing probability too. Both models are one-dimensional, since they take just one trait into account, concretely the ability the item bank measures; they are dichotomous because they consider that an item response can only be correct or incorrect; and both models use the logistic function for practical reasons.

Most of the existing CAT generators follow the steps described by the following algorithm (Thissen & Mislevy, 2000) even if they are based on different IRT models. In fact, Figure 2 schematizes the minimal functionalities that a generic CAT system must provide, showing how it works: after presenting the instructions to be followed during the assessment (1), the system takes an initial ability estimate ($\hat{\theta}$) and selects the first item to be shown (2); then, the examinee answers the item (3) and the ability estimate is updated by some maximum likelihood or Bayesian method (4); if the stopping criterion is satisfied (5), then the CAT finishes and the final score is computed; otherwise, the algorithm selects a new item (6) (providing both maximum information for the provisional $\hat{\theta}$ and fulfilling any defined constraint), and presents it to the examinee (3); after collecting the answer, the ability estimate is updated (4), and so on. The cycle continues until the stopping criterion is satisfied, something that, depending on the particular implementation of the algorithm, can happen, for instance, when a predefined number of items has been administered, if the error of the new estimate is smaller than a certain value, or when a time limit has been reached.
Previous figure has shown how the generic CAT algorithm works, but there are alternative approaches, such as assisted self-adapted tests (Wise, Ponsoda, & Olea, 2002), testlet-based models (Wainer, Bradlow, & Du, 2000) and tests with content balancing (Kingsbury & Zara, 1989), which are very interesting from a communicative perspective, as will be discussed in next sections. An assisted self-adapted test will select (step 6) but will not necessarily administer (step 3) the best item in each iteration of the CAT algorithm. It will ask the examinee instead, giving him or her the chance to choose among the recommended item or another one (for instance, a significantly easier one after a wrong answer). On the other hand, the Testlet Response Theory is a generalization of the IRT where items are substituted in the administration algorithm by sets of questions that are applied together and usually share a common environment or supporting document. Lastly, the majority of the existing CAT systems impose restrictions during the next item selection step, in most cases to avoid the item overexposure, that is, to prevent the most informative items to be constantly selected and administered to every examinee that performs a CAT. Some of them also implement content balancing procedures to establish limits on the number of items of every learning area to be applied in a test, so every generated CAT will include similar proportions of items for every considered area.

GenTAI and the Genre model

GenTAI (López-Cuadrado & Sánchez Fernández, 2005) is an interactive application that allows the user to generate TRI-based CATs. Although it can perform tests from any set of items, the system is intended to use an item bank for the assessment of knowledge on the Basque language (López-Cuadrado, Armendariz, & Pérez, 2005).

In this section, the concepts of the Genre model (Figure 1) are specifically informed for GenTAI purpose (Figure 3). The CAT generator was developed independently of the previously presented theory, but, as one could expect, the implemented system matches the model of Genre. This section, as well as the following two, will describe how the communicative features of GenTAI fit the model from a theoretical perspective. Its evaluation in terms of the Genre model at a practical level, together with the development of enhancements in the application (such as those regarding the expectations of the final user), is left for future research.

GenTAI follows the CAT-based evaluation Genre, which describes a set of software applications with some common characteristics. First of all, in this kind of application the interactors are the examinee (receiver) and the CAT generator (transmitter). The intention of these interactive programs is to obtain an ability estimate for the examinee. This goal is reached thanks to the CAT administration algorithm interaction scenario, which is implemented within the CAT generator module. The recently presented CAT administration algorithm selects the items the examinee should answer. Actually, the activities consist in answering to these items.

Since GenTAI is classified within the CAT-based evaluation interactive applications, it makes use of the communication structures linked to this Genre. It manages some documents: text-based multiple-choice items stored in the item bank about the Basque language. Those items are characterized by the 3-parameter logistic IRT model, so their difficulty, discrimination and guessing values must be known to successfully perform the tasks regarding item selection, ability update and, in some cases, verification of stopping criterion. The items are also classified by

0
domain category (syntax, verbs, declension, vocabulary, suffix, and connectors). GenTAI is expected to be used in a context where the examinee is alone behind the computer, that is, without help, such as other people or books, while finding the correct answer to the items he or she has to solve. In most cases the user is expected to be in an euskaltegi – a centre where adult students learn the Basque language – while interacting with GenTAI, since the application has been developed to determine new learners’ ability with the final purpose of deciding their starting level in Hezinet, the adaptive hypermedia system for the Basque language learning currently in use in many centers (Pérez, 2000; López-Cuadrado, Pérez, Vadillo, & Arruabarrena, 2002).

The CAT administration algorithm scenario allows to satisfy GenTAI’s intention, in other words, to obtain an estimate for the examinee’s ability, by means of managing different communication structures. To achieve this purpose the CAT administration algorithm arranges the stages needed to present items and to show instructions and final reports to the user. This scenario also establishes the constraints of the CAT administration algorithm, such as those related to the item selection procedure, the ability estimation method, stopping criteria or other restrictions (like those imposed by the techniques that control item overexposure and underexposure), or decisions (such as allowing response omissions, test review, or notification of correct or incorrect answers).

The CAT administration algorithm (Figure 2) is the starting point to identify the following underlying communication stages, which are contemplated in each of the iterations supported by any system that generates CATs:

- Request for examinee’s data.
- Presentation of instructions.
- CAT administration, which includes the “Administration of a single item” communication stage and may also contain the “Notification of item response’s result” stage.
- Notification of CAT’s final result.

Once the communication stages of a CAT generator have been identified, the next step is to associate the most appropriate communication structures with them. It is necessary to discuss the suitability of alternative solutions with interactive potential, depending on the intentions of the developers of the CAT system. When designing and implementing the interaction system of a CAT generator like GenTAI, it is recommended to consider first the communication structures that are potentially shared by all of these communication stages, and next, the communication specificities of each stage. This is why next section is dedicated to present the common structures, while the following one will discuss the particularities of every stage.

**Common communication structures**

Figure 4 shows the communication structures that most interactive applications employ: exchange, rhythm, analogy and utterance. The communication stages identified in GenTAI, which have been just mentioned, clearly make use of them.
**Figure 4 –** Most commonly used communication structures

*Exchange* is a very common structure within the communication stages of a CAT generator, particularly in the appearance of selection, interrogation, and notification.

- The *selection* shows the user several options to choose among them. It can be used through every of the communication stages. For example, during the request for examinee’s data the system can list the considered educational levels (such as, no studies, primary/higher education, graduate, postgraduate); during the presentation of instructions the application can ask whether advices and terms have been comprehended or not; during the administration of a multiple-choice item the user has to select the correct option; if the implemented CAT is an *assisted self-adapted test*, during the notification of item response’s result the system must ask the user if the next item to be administered should be easier or harder than the last one; finally, during the notification of CAT’s final result, GenTAI lets the user choose among different formats to visualize the result.

- The *interrogation* is employed by applications like GenTAI to ask the user different questions, some of which, as just seen, can be implemented by a selection structure. However, interrogations are usually more complex than a simple choice among several given options, and it becomes clear, for example, during the request for examinee’s data, when the CAT generator can make inquiries about user’s birthday, address, comments or suggestions.

- The *notification* informs the user about some topics. GenTAI makes use of this communication structure during the request for examinee’s data, concretely when explaining what these data are going to be used for, and when exposing which fields are compulsory and which are optional; during the presentation of instructions, the examinee is informed about how a CAT works; even during the administration of a particular multiple-choice item, the statement (*stem*) itself alerts the examinee about the correct answer; if the notification of a response’s result stage was implemented, the system should let the user know if recently answered items had been succeeded or not; lastly, during the notification of CAT’s final result, GenTAI shows the score obtained in the test.

The *rhythm* is very flexible. One of its variants concerns the language in which the texts of GenTAI are displayed. Thus, it is desirable to use a binary or even a ternary rhythm by combining different languages, since the application evaluates the Basque language level of people that, in most cases, only understand Spanish. This is why GenTAI shows every message in a binary pattern (Basque/ Spanish), and, as soon as the system is ready to be used in the French Basque Country, it will also include a French version of the contents.

*Analogy figures* are usual, not only in applications that generate CATs, but also in other interactive programs. The most significant figures are metaphors, comparisons and personifications.

- The most important *metaphor* in GenTAI is the use of radio buttons, which emulate the mechanism of old radios by switching off the selected channel as soon as any other is selected. Radio buttons are very commonly employed to let the user select one and only one of the options that are available, for instance, within a selection.

- *Comparisons* can be implemented by the use of animations or screen video-captures, where the system shows the user how certain tasks should be performed. It might be implemented during the presentation of instructions...
to explain what kind of interaction is expected, for instance, during the request for examinee’s data or the administration of a single item.

- **Personifications** can be used all through the execution of the CAT generator, especially if the application is intended to assess the knowledge of children or elderly people. The final score of the CAT can be personified, for example, by the use of an avatar that has a better or worse appearance depending on the result. Thus, if a smile is designed to characterize the score, the better its value is, the happier the smile will be. One can also conceive personified metaphors, for instance, when designing an animated assistant, like the used by the Office suite (Microsoft, 2004), to help examinees that are not used to computerized testing during the different stages of the assessment process.

Finally, the utterance is shared by all the stages. The enunciative authority in this case is the CAT generator itself, which makes use of different enunciative modes, depending on the communication enunciation. The most significant ones are narrative, argumentative, descriptive, explanatory, and injunctive.

- The narrative mode is used whenever the system has to inform the user, in particular, when presenting the instructions and reporting a response’s result – if the CAT generator applies assisted self-adapted tests – or the score of the test.
- If the system displays a notification together with some kind of reasoning – for example, if the final score is not only shown but also discussed – then argumentative mode is employed.
- The descriptive mode appears often, for instance, during the request for examinee’s data, the system can explain the user what every field in the questionnaire means; while presenting the instructions, most CAT generators describe how the test administration is going to happen; and in notification stages, the system can give some explanations about presented data, such as the scale in which the final score is given.
- The explanatory mode can be used as extension of the descriptive one, mainly when the system offers the user a more technical description, including, for example, terms related to the IRT.
- The injunctive mode appears whenever the system needs the user to do something, as performing a selection, so it can be used in each and every communication stage.

**Particular structures for every stage**

Last section has presented the communication structures that are common to all the communication stages implemented by a CAT generator. This one will discuss the particularities of every of these stages. For instance, three types of rhythm will be distinguished: rhythm by combining several languages, binary rhythm and enumerated rhythm. In all of them it is clear that the sender wants the receiver not to miss the message, even if offering different languages in an interactive application may well seem not to be a rhythm case. However, in this context the authors are supposing that GenTAI users are almost bilingual, that is, they manage their native language (Spanish, French, English...) but they are not necessarily Basque speakers. So, instructions and notification texts should be displayed at least in one language other than Basque, if one wants the user to understand the message. The idea is that, as soon as the user that reads the Basque version of these texts has a doubt, for instance, about the meaning of a concrete expression, he or she will have the chance to read the same words in his or her native language to compare both versions and understand the message.

**Request for examinee’s data**

The first time the user interacts with a CAT generator, the system uses to ask for some personal data, as the name or age of the examinee. Sometimes more specific information may be also asked, for instance, if it is projected to carry out statistical analysis or obtain empirical conclusions about how the application is working.

In the case of language assessment systems like GenTAI, where the aim is to determine the starting level for new students that are going to interact with the learning system Hezinet, it is recommended to display every message by implementing a binary or even ternary rhythm by combining several languages. The fact is that some examinees are not too well up on the language the CAT measures, as happens with GenTAI, most of whose users are expected to be mainly Spanish or French speakers.
Presentation of instructions

This stage is crucial for the examinees. Beginning a test knowing the goals, how a CAT works, and which are the existing restrictions should prevent the user from confusions, distractions, frustrations and other negative effects derived from the fact that a CAT differs from a traditional test. Actually, details like applying the items one by one or not allowing omissions and revisions can falsify the results if the examinee is not prepared for such circumstances.

The rhythm should be carefully designed for this stage. To use a binary rhythm during the presentation of instructions is useful if sender’s intention is to insist in some point, but ternary or enumerated rhythms border on repetitiveness, and in some cases can exasperate the recipient. For this reason, the authors recommend to use them only in concrete situations, i.e. in those whose importance the transmitter explicitly wants to emphasize. The designers of a CAT generator can apply different implementations for the rhythm. As mentioned above, the system can display the information in several languages, so the recipient can choose the preferred one; it is also possible to use different formats to explain the instructions, for instance, by combining plain text with explanatory audio or video clips; finally, a more interesting alternative from the psychometrical perspective consists in allowing the examinee to answer a couple of sample items so he or she will become familiar with the user interface.

CAT administration

When talking about communication structures, the administration of a CAT is an element that can be fractally divided into smaller ones, since it includes several instances of the administration of a single item stage, and sometimes it also implements the notification of item response’s result.

The designers of a CAT system must bear in mind all the communication structures for this stage. Among the exchanges, selection and notification are the most important ones, and should be carefully implemented. Items are iteratively presented one by one, as the CAT administration algorithm determines, and this enumerated rhythm is broken only when the stopping criterion is satisfied. Some CAT generators implement content balancing procedures, so they establish limits on the number of items of every learning area – such as declension, verbs, syntax and vocabulary in the case of GenTAI – to be administered in a CAT. This situation, which can also be seen as a rhythm rupture, becomes very useful to avoid keeping on applying items that cover the same learning contents. Analogy figures can be used to enhance the capabilities of the CAT, particularly if the recipients are children. Designers can for example conceive the test administration as an adventure, or present it using some kind of metaphor like a book that has to be read over, as Hezinet does with its learning sessions (Pérez, 2000). However, from a psychometric point of view, this kind of frills might change the final score, so it is recommended to consider their effects on the results before designing an analogy figure for a CAT administration. Finally, utterance in this context is characterized mainly by an injunctive mode, since the examinee has to continuously answer items.

Administration of a single item

The application of an item is a communication stage that the CAT administration algorithm repeats until the stopping criterion is fulfilled.

The rhythm is not very relevant in this stage, unless the CAT generator administers items with some audio or video, for instance, as basis to build questions. Depending on how many times the examinee is allowed to listen or watch every multimedia document, and also on the existence of its text version, the rhythm can be binary, ternary, or enumerated. Due to the difficulties of calibrating this kind of item bank, most of the operative CAT generators do not contain multimedia in their items. However, the importance of complex, i.e. polytomic and multidimensional, IRT models is growing every day (van der Linden & Hambleton, 1997), as happens with Testlet Response Theory, so enumerated rhythm gets importance if CAT designers want to implement computerized adaptive testlet-based testing.

As mentioned before, analogy figures can be used during the item administration. For example, an animated speaking question mark can be the one that presents the examinee the items. However, although personifications like
this can be interesting from a pedagogical point of view, mainly if CATs are going to be applied to children, from a psychometrical perspective they might produce negative effects on the results.

Finally, the enunciative mode used by the utterance communication structure in the item administration stage is mainly injunctive, since the CAT generator expects the examinee to select the correct answer among the offered options. However, depending on how the question is made, injunctive mode can appear combined with others, such as interrogative (who invented the kaleidoscope?), narrative (A train leaves San Sebastián for Bayonne at 7.00 a.m. traveling at 90 km per hour; 30 minutes later another train leaves Bayonne for San Sebastian at 110 km per hour...), explanatory (A number is said palindromic if it is read the same both ways. How many of the following numbers are palindromic?...), or descriptive (In the following video you will be able to see the distorted image of an object and listen to a voice. The voice will describe the object by giving some hints. Your work is to guess which object it is.).

**Notification of item response’s result**

Although it is not usual to inform the examinee about the result of every administered item, sometimes, particularly when implementing assisted self-adapted tests, this communication structure is employed to notify the user if the item has been answered correctly or not. In this situation, the system usually suggests, depending on the result, either an easier or more difficult item to be displayed next, but also gives the examinee the chance to decide the difficulty for the next item.

The enunciative modes that the notification of each response’s result can use are essentially the same as those implemented for the notification of the test’s final result. The only difference is that the descriptive mode is useless to inform about a single item’s response, but it might be practical when explaining the CAT’s final score.

**Notification of CAT’s final result**

After administering the last item of the CAT, the system will report the obtained final result. The rhythm in this communication stage is linked to the idea that the score can take different formats. For instance, GenTAI lets the examinee know the results simultaneously by the numeric IRT ability, the same value scaled to a more comprehensible scale, a percentage of correctly answered items, an evolution graph and a table of administrations or log.

This communication stage is ideal to explore the possibilities that analogy figures offer, since their employment during the notification of CAT’s score will surely have no effect on the results. CAT designers can use, for instance, metaphors as a thermometer bar to represent where the examinee’s ability has been measured, comparisons by using the centile scale to contrast CAT’s score with the results obtained by other people, and personifications like those mentioned before to explain the results to children.

**Conclusions**

The functional aspects of CAT generators are very well documented in literature. They cover a wide range of proposals about criteria, models, protocols, and metrics to be taken into account during the design of a CAT system. In fact, these principles are the basis for the implementation of concrete CAT generators. As happens with any other interactive system, during the development of the interactive part of a CAT generator, the designers can also consider proposals such as ergonomic recommendations (Scapin & Bastien, 1997) or style guides (Lynch & Horton, 2005), whose purposes are general for any kind of interactive system.

In this paper, the authors have focused on the communicative specificity of CAT generating systems with the aim of providing a chart for the design of the interaction of this kind of applications. To achieve this, the communication stages supported by any CAT generator have been specified, and afterwards the communicative potential of each of them has been studied. This potential is derived from the analysis of every stage within a Genre model that eases the specification of the Genre of any interactive application. Actually, this model proposes a set of interrelated concepts
whose instantiation, for a given application, takes part during the definition of the Genre for the desired interaction between sender and receiver.

For a particular CAT, the communicative capabilities of its stages can be stated and stabilized depending on the concrete framework or background pursued by the designers. Concretely, Table 5 offers a reference that might help the processes of decision and line of argument while designing the communication stages of an interactive CAT system. The Genre model can be useful to study not only CAT systems, but also other kinds of interactive e-learning software, such as intelligent tutoring systems or adaptive hypermedia systems. The designers can improve the communicative capabilities of their applications in two steps. First, the developers should define the communicative stages to be implemented, not only from a classical rhetoric perspective but also from the interactive point of view, and then, they should find and study different appropriate communication structures for these stages. This paper has presented the theoretical results for GenTAI, a particular CAT generator, including some examples of structures that could improve the communicative practical capabilities of the system.

Table 5 – Communicative consideration of the CAT’s communication stages

<table>
<thead>
<tr>
<th>COMMUNICATION STAGES</th>
<th>Request for examinee’s data</th>
<th>Presentation of instructions</th>
<th>CAT administration</th>
<th>Notification of CAT’s final result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Request for examinee’s data</td>
<td>Presentation of instructions</td>
<td>CAT administration</td>
<td>Notification of CAT’s final result</td>
</tr>
<tr>
<td></td>
<td>Administration of a single item</td>
<td>Notification of a response’s result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhetoric of the interaction</td>
<td>Selection</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Interrogation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Notification</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Classic rhetoric</td>
<td>Rhythm</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Metaphor</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Comparison</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Personification</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Narrative</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Argumentative</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Descriptive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Explanatory</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Injunctive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Interrogative</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

From a methodological point of view, the presented method enhances the existent approaches for the design of interactive systems, regardless of they are based on data as OOHDM (Schwabe & Rossi, 1998), RMM (Isakowitz, Stohr, & Balasubramanian, 1995) and WebML (Ceri, Fraternali, & Bongio, 2000), based on tasks as CTT & SWC (Winckler & van der Donckt, 2005) or based on interactions as UMLi (Pinheiro da Silva, 2002). All of them are helpful to design an interactive application such as a CAT generator, but they can offer better results if they are combined with the two suggested steps in this paper. In fact, the designer can use the identified communication stages, perhaps with some specific changes in the naming and focusing for each approach, as the conceptual bricks that will form the final design. For instance, RMM proposes slices and access directives to navigate through interactive nodes, OOHMD requires the establishment of navigational contexts and the definition of the interface objects that will be responsible for the interaction, ConcurTaskTrees & StateWebCharts use containers to organize the interactive contents, as also does UMLi, and WebML proposes the arrangement of pages with units and their interactive capabilities, and with optional selectors. The complementary contribution of the approach introduced in this paper to these existing alternatives is founded on the systematized and organized assessment of the communicative capability the designer desires for every communication stage. The already mentioned methods propose models and schemas that formalize the dynamic and the contents of the interactions, but they do not necessarily consider the communicative intention. The approach presented by the authors helps the designer to formally define communicative capabilities, and, following the Genre model and its concepts, offers a reference for the construction and establishment of reasoned communicative intentions.
Finally, from a perspective based on the conception of processes, nowadays software production procedures try to consider the final user in a significant way. This tendency is typically implemented during software engineering by means of the inclusion of prototyping and participative approaches into the life cycle. These tendencies pursue both to increase the probabilities of acceptance of the final product by its target users and to minimize the re-working risks, in other words, to avoid repeating a production task because of an incomplete consideration of the vision that the different actors of the production process have about the application.

In the future the authors intend to prepare surveys to work on the opinion of GenTAI users. Actually, a more ambitious analysis could include, besides the intentions of the designers, also the expectations of the users. This could require to form a dynamic adaptation of the communicative behavior according to the reactions of the user and some reference model. Furthermore, since CAT systems do not use more information about the examinee than an ability estimate and the answers given to the set of administered items, it is necessary to study the practicability of taking additional personal information into account while performing the implementation of a CAT generator. Such an analysis should determine how the inclusion of complementary data, such as historical learning achievements, ages, genders, learning styles or attitudes, could affect not only the evaluation outcomes but also the communicative potentials of the resulting CAT system.

Acknowledgements

The authors want to thank Tomás A. Pérez and Pantxika Dagorret for their help and feedback during the writing of this article, and they are grateful to the reviewers for their careful reading and comments.

This work was supported by the projects Decuple e-learning (nº2005/11/07/1320) and Desarrollar una Competencia Universitaria Transfronteriza sobre el Proceso de e-Learning (BOPV nº2004/12/03/232), both of them financed by the Common Euskadi - Aquitania funds.

References


Canvat, K. (1998). Enseigner la littérature par les genres; pour une approche théorique et didactique de la notion de genre littéraire, Bruxelles (Belgique), De Boeck Université.


