INTELLIGENT SYSTEM TO ASSESS AND TREAT DEVELOPMENTAL DYSLEXIA IN SPANISH LANGUAGE

Moreno L **González C**

Sánchez J. **Muñoz V** **Estévez J**

Aguilar R. **Sigut J.**

Departamento de Física Fundamental y Experimental, Electrónica y Sistemas. Edificio de Físicas y Matemáticas
Universidad La Laguna. Islas Canarias. Spain
Email: jsanrosa@ull.es

ABSTRACT

This paper describes the architecture of the already implemented package where three important elements interact: a multimedia interface, an inference module and a database. This architecture provides the system with the flexibility to support a large variety of tasks, dynamic presentations, and complex teaching strategies. The application of speech recognition technology is also an important part of the dyslexi c children evaluation and treatment. A research effort has been done to develop a software environment to help tutors of dyslexic children with the diagnostic and treatment tasks. Nowadays it is being used in several Spanish Schools as part of its validation process.

INTRODUCTION

Our research team had designed a system on the base of numerous studies about learning and reading difficulty in a transparent language as it is the Spanish. In this paper, we will aboard the design and implementation of an adaptive multimedia system to help tutors to diagnose and treat developmental dyslexia in the Spanish language (Jiménez, 2000). Reading problems are originated, in the majority of cases, due to difficulties in accessing to lexis. Developmental dyslexia term is applied to children that present this difficulty. Dual model of access to lexis describes with certain detail how the lexis process level works and permit us to make some predictions. There are two subtypes of developmental dyslexia in this model: phonologic dyslexics and superficial dyslexics.

NOMENCLATURE

Intelligent multimedia, Adaptive hypermedia systems, Developmental dyslexia, Speech recognition

1. DYSLEXIA BASIC CONCEPTS

Frequently, teachers are concerned about the great amount of children that cannot acquire the reading ability. Approximately, the 3% s shows a hard and persistent difficulty. This situation limits the student progress due to reading is a fundamental tool to acquire knowledge, and that it has social and personal consequences for the children. For this reason, a fast identification and an appropriate treatment of this problem are necessary to help with this learning difficulty.

Researchers generally agree that difficulties in word recognition produce reading problems (Olson, Kliegel, Davidson y Foltz, 1985; Stanovich, 1988; Van Den Bos y Spelberg, 1994), and also that good readers shows high automatism out of context word recognition (Perfetti, 1989). Term “developmental dyslexia” means “ retarded readers with high IQ and low progress in reading whose deficits are not due to sensorial, physical, motor, emotional problems or lack of learning opportunities” (DSM IV, 1994).

According to dual model to lexis access (Coltheart, 1978), there are different ways or procedures to identify words. A way can be to identify the word as a draw. That is, identify the word as a whole, in the same way that a visual stimuli. This way is called “lexical route or by visual route”. The unity of recognition is the word. There are empirical data that support the visual route existence. So, dyslexics delay more reading words than pseudo-words (lexis effect) and considering words, they read faster frequent words (familiarity effect). Also, there are differences in reading errors: bigger in the pseudo-words and unfamiliar words than in familiar words (Valle, 1989). Therefore, recognition times are lower when words are presented in their usual "box" or shape. Thus, "elefante" is read faster than "EleFaNte" (McClelland, 1979).

The possibility of reading unfamiliar words and pseudo-words suggest that must be other ways to read. In such ways, recognition units must be pieces smaller than words as letters, syllables, etc. After these units have been identified conversion rules of grapheme-phoneme are applied, assembling all. This reading mode is called "phonologic route".

Many researchers (v.g.,Castles y Coltheart;1993; Manis, Seidenberg, Doi, McBridge-Chang y Petterson, 1996; Stanovich, Siegel y Gottardo, 1997) had concluded that there are as many dyslexia subtypes as the dual model proposed plus one. Developmental dyslexia can be phonologic, superficial or mixed (with characteristics of both previous subtypes).

Phonologic dyslexic has problems in the phonologic route, for this reason familiars words will be easily read but great difficulties are shown with unfamiliar words and pseudo-words. Due to phonologic dyslexics are supported in the visual route to read, their errors will appear in functional words (articles, prepositions, etc.) and lexis (Patterson, Marshall y Coltheart, 1985). On the contrary, if the problem is in the visual route, regular words and pseudo-words can be read without difficulties, but irregular words (there are not in Spanish) and differentiating homophones words ("ola"/"hola") are difficult tasks. Errors will be produced by the attempt to regularize words and reading could loose fluency because they are supported by phonologic route (Patterson, Marshall y Coltheart, 1985).

It is important to say that some authors argue that developmental dyslexia classification in subtypes have worked in English language. They claim that the essential problem of dyslexic is the weakness in phonologic process (v. G., Felton y Wood, 1992). But in a transparent orthography as Spanish language, it is demonstrated that phonologic problem, even being the main difficulty, is not the only one reason. (Jiménez y Ramírez, 2001).

How to differentiate a bad reader and a dyslexic pupil on one hand and the analysis of different difficulties presented in this problem on the other hand are the main problems of human tutors at the head of a reading class. So, after pupils had been identified as dyslexics, another problem is to identify which developmental dyslexia subtype they belong to. Independently of identified subtype, it is necessary a functional evaluation of linguistic and cognitive abilities to find where students have deficits. After that, treatment will be applied only in the presented deficits.

New technologies can be helpful for people with learning difficulties (v.g.,Cisero, Royer, Marchant y Jackson, 1997) and, in this sense, we have designed and implemented a system as a support tool (Gonzalez 2001) for the professionals who work with dyslexic children.

We have organized this paper as follow: firstly, how we have modelled and implemented the system with adaptive characteristics, secondly how we have particularized the application in the case of developmental dyslexia, and finally we describe our actual research on the implementation of Automatic Speech Recognition in the system.

2. MODELS AND METHODOLOGIES USED

Each time the student interacts with the tutorial the results are stored in a database and this information is taken into account in subsequent actions. In this way, the achievements and progresses of the user can be analysed, and according to the results of this analysis, the "task thread" and the characteristics of the presentation are dynamically changed. The decisions about dynamic changes are supported by the knowledge based system (KBS).

A KBS is the part of the system where the concepts, facts and relations between the components and the adaptation rules to afford contextual changes are represented. Hence, it is clear that the correct design of the KBS is a primary step. For this reason we have used well known methodologies to represent the expert knowledge, in particular KADS (Schreiber G., Wielinga B., Breuker J., 1993) and IMMPS (Bordegoni M., Faconti G., Rist T., 1996; Karagiannidis C., Koumpis A., Stephanidis C., 1995).

Adaptability is considered the better characteristic of IMMPS to facilitate interactive systems access to people from different categories and requirements. Interactive processes are characterized by several attributes, which are not clearly separated:

- Determinants: are factors that guide the adaptation process (i.e., adaptation based on the users characteristics, kinds of task or interaction, etc.);

- Constituents: are interaction aspects that are adapted (i.e. content, presentation primitives as interaction techniques, media or modalities);

- Objectives: are particular goals that guided the adaptation process (i.e. to minimize errors numbers, to optimise the efficiency and efficiency, etc.);

- Rules: Guides the process of adaptation constituent's instantiation according to the state of adaptation determinants.

The adaptation process is carried out trough the selection of constituents based on the determinants using the adaptation rules.

2.1 DETERMINANTS

To design the adaptation in our system we have categorized the determinants in two spaces:

a) Interaction Context Space (IC)

b) Information Space (I)

More specifically, the IC space is built with the follows determinants:

- Student Model: formed by a perfil (personal characteristic of student) and records (variables about the learning evolution).

- Teaching strategies: hold the knowledge about the plan of didactic sequence, and are based on a set of rules that allows system adaptation to the student according to the learning goals.

On I space, determinant is the convenience level about media use. We have designed different evaluation and treatment modules according to the problems that are present in each dyslexia type. The next modules were defined: syntactic, audible discrimination, levels of phonologic awareness, knowledge of letters and lexemes and suffixes. The nature of each task depend of module and stimulus presented into the module. Text can be presented as visual or audible way, with diverse styles (script, bold, ...), in a graphic way, with sounds, etc. On the other hand, graphics and animations can be presented isolated or joined. For instance, in the case of audible discrimination module, we must use the audible media that have more incidences in the task (words, couple of words, syllables, etc.).

2.2 CONSTITUENTS

Adaptation constituents had been defined as follows:

- Information presentation: the information is presented according to the kind of dyslexia and the level of affected routes.

- Task simplification: is the process where complementary object are deleted and/or options are reduced.

- Visual components of presentation: phrases, words, images and animations are presented in accordance with their incidence level in the kind of task that the pupil has to do and/or the type of errors committed.

- Audible components of presentation: as the visual components, phrases, words, feedbacks and explanations they are presented in relation to their incidence level in the kind of tasks that the pupil has to do and/or the type of errors achieved.

2.3 OBJECTIVES

Our goals are:

-In technical terms:

- Easy scalability and system reuse

- Decreasing computing resources cost

For these reason, we have designed the system with a scalable architecture. So, with this architecture we can represent dynamically the media and the system can increase easily in media and stimulus with a simple addition of their references and relations in the database.

-In pedagogical terms:

- Increasing student attention and motivation.

- Encouraging the significant learning through a right combination of media and modalities

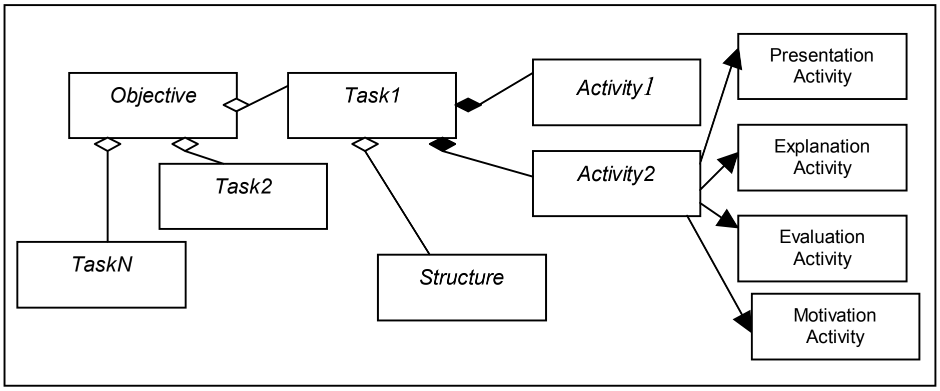


Figure 1 . Abstract object model of the system.

Consequently, to achieve the goals, we have defined the instruction system with a set of components. Components are several kinds of tasks and events for encourage the attention, motivation and learning process during the program execution.

2.4 RULES

Rules define the allocation of constituents to determinants taking into account the satisfaction level of the adaptation goals.

Some of the rules that we have defined to our system are:

- If the user has phonologic dyslexia, then show pseudo words or familiar words in the evaluation.

- If user has phonologic dyslexia, then show irregular words and introduce pseudo words or familiar words gradually into the treatment.

- If user has superficial dyslexia, then show irregular words in the evaluation.

- If user has superficial dyslexia, then show pseudo words or familiar words and introduce gradually into the treatment irregular words

- If user has superficial dyslexia then enlarge gradually words long.

- If user has superficial dyslexia then enlarge gradually words complexity.

- Be IO = { io1, io2,….., iox} the set of images associated to the object o, and be o={word, letter, sentence}, if the student has made an error related to the object o, it is presented with its associated image io in light up modality.

- If user has committed an error during the treatment, run the feedback using the agent, with the explanation related to error type produced

- If user makes error n times, decrease complexity task.

- If user makes error n maximum times, review the requisites previous and go to another task related of less complexity.

- If user makes a task satisfactorily, run the positive feedback using pedagogical agent and go to next task in complexity order.

- If is the first presentation of a task, run the demonstration, merging in the time the sequence of audible and visual media with the agent execution.

The interface which allows the communication between the student and system has been designed taking into account cognitive theories about the learning with multimedia (Mayer, 1997). The exercises are presented by pedagogic agents in a verbal and gesture way further learning and student motivation (Moreno y Mayer, 1999). Therefore, we have four formats to the task: a) motivation (simple tasks); c) explanation (to show, to describe); d) evaluation (to ask/ to relate) and d) reinforcement (do again changing presentation) (Moreno et.all,2000).

Our system has been designed on the base numerous researches about the learning of reading and his difficulties in a transparent orthography system as the Spanish (Jiménez y Ortiz 2000; Ortiz y Guzmán, 2001; Rodrigo y Jiménez, 1999; Rodrigo y Jiménez, 2000; Jiménez y Hernández-Valle, 2000). All these studies show the relevance of lexis units in the recognition of word.

In the following section, we will describe how the system has been designed and developed.

3. FUNCTIONAL DESIGN, INFERENCE MODEL AND IMPLEMENTATION ISSUES.

3.1 FUNCTIONAL DESIGN

The main functional subsystems are the database, the inference module and the interface. The database keeps crucial information about different aspects of the package. On one hand, it stores the available media components and the relation between these components and the concepts involved in the tasks. On the other hand, it stores the profile and history composing our student model.

This system has an inference module with the responsibility of selecting the next task to be proposed taking into consideration the information stored in the database. Furthermore, the media components to be used in the selected task are also chosen examining both the relation between all the components stored in the database, the specifications of the task and the profile and history of this user. All the inference process is not concentrated in the inference module, as a part of the inference is carried out by the interface.

The interface is a piece of software where the interaction with the user is controlled. The interface has to determine the basic parameters related to the media components to produce coherent responses to the user actions (for example, which media to show when the answer is correct or wrong). To get this, once again the information stored in the database has to be analyzed.

3.2 INFERENCE MODEL

The previously explained ideas about how to carry out all the inferences needed to build a presentation has been implemented using an inference model composed of four basic classes of "objects": the class "Objective", the class "Task", the class "Activity" and the class "Structure".

1.- An instantiated object of the class "Objective" works as a container for a group of instantiations of the class "Task" (objects classified as Tasks). Tasks objects are added to a container of the class "Objective" through an "aggregation" relation. In this way, if the container is destroyed the instantiated task is not eliminated. The containers of the class "Objective" are related between them, and therefore the information stored in a container of the class "Objective" allows the system for finding out which tasks should be solved by the user in order to progress through a search structure (for instance, a decision tree).

2.- Each instantiation of the class "Task" is an object composed of other objects of the class "Activity" through a relation of the type "composition" (if the task object is destroyed all the activities composing the task are also destroyed). However, an instantiation of the class "Task" includes the guidelines of how the information related to the task to perform is going to be presented to the user (which and how the media components should be used in the interface). For this reason, aggregation relations also exists between a Task object and one or several structure objects described below.

3.- Each activity object can inherit attributes from different subclasses of activities. These subclasses are Presentation, Explanation, Evaluation and Motivation. Each subclass has a particular purpose and behavior.

4.- The structure object is a template with specific directives about how the media components should be used to obtain the desired presentation.

3.3 IMPLEMENTATION ISSUES

Building a tool with the mentioned features requires the collaboration of a multidisciplinary team. In this sense, from the implementation point of view there are two important issues: the technical design (software engineering) and the artistic development of the media components.

This last point requires the collaboration of different groups. In our case, a multimedia specialist together with a professional artist under the supervision of the psychologists developed and produced most of the media included. The suitability of the artistic work was validated with a number of children asking them about the style and the contents of the drawings. One of the key points very related to the user motivation is the variability of the scenarios and media components: at the moment there are about 200 original drawings produced and there are still another 600 to produce in the final system. There are also 2400 sonorous elements (explanations, feedback, sounds related to the scenarios ...) already produced and we are estimating that the final system will contain around 4000. This high number of elements together with the capability of the system to compose dynamically coherent scenarios with continuous variations avoids the routine in each session, increasing the motivation and interest of the students. These numbers also explain the important role that the inference module plays in the coherent production of the multimedia presentation.

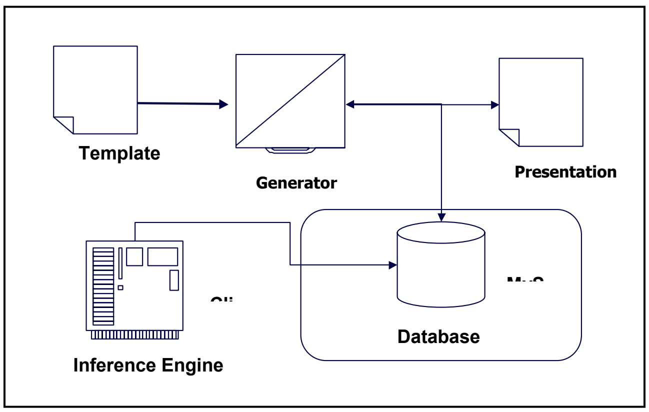


Figure 2. Interrelation between the main software components of SICOLE.

The other keystone of the package is the design of the software. The system is based in a number of available tools to cover the different requirements. At the present moment, the database is supported by the MySQL environment. This decision was made after evaluating the robustness and fastness to maintain high amounts of data and to exchange pieces of information with other Windows applications of several products. The core of the multimedia interface is based in commercial applications as Macromedia Authorware 6.0 and Flash 5.0. The media production was carried out with Macromedia Fireworks, Corel Draw and Corel Photo Paint. Finally, the inference module is based in the CLIPS environment and an interface between CLIPS and the database developed for this project with Microsoft Visual C++ 6.0. Additionally, an agent is used in the presentation to guide and motivate the student: it is based on the Microsoft Agent Technology and programmed as an Active X object.

The software architecture is designed to allow the system to adapt each presentation to the individual characteristics of each student through a software module. This module generates activities dynamically and builds the presentation following the learning style of the student. In this framework the structure element is very important. The structure of a particular task is stored in the database. We can see this structure as a template that the inference module has to complete with data extracted from the database. The type of data included in a template is diverse: interaction types, definition of text containers, graphics, sounds and animation, input and output variables,... To understand better this concept, an example (see Example 1) corresponding to a simple structure related to the "Lexemes and Suffixes" module is now described. Once the structure is filled it represents a task that makes the user must select the correct picture between two, representing the meaning of the word that appears in the screen.

The slot "id" is a number representing the module related to this task. It has a default value of 6, which means that the default module is the "Lexemes and Suffixes" module. The slots "picturex" are the identifiers of the pictures that are going to be used. One of the pictures will be the correct one, while the other is a distracter. The slot word is the identifier of the word to show in the screen. The slot ID_ist is the identifier of the instructional message that describes to the student how to carry out the task. The slot ID_exp is the identifier of the explanation of this task, it could be an example of how a particular task is solved. The slots f_correct and f_mistake are identifiers for the feedback that the user has to receive when the task was solved correctly or incorrectly. Finally, the slot relation establishes which of the two pictures is correct.

```
(defclass  EVAL_ESTRUCTURA_A  (is-a  EVAL)  (role  concrete)  (pattern-match
reactive)
(slot id       (create-accessor read-write)
(type INTEGER)
(default 6))
(slot picture1        (create-accessor read-write)
(type INTEGER))
(slot picture2        (create-accessor read-write)
(type INTEGER))
(slot word     (create-accessor read-write)
(type INTEGER))
(slot id_inst  (create-accessor read-write)
(type INTEGER)
(default 0))
(slot id_exp   (create-accessor read-write)
(type INTEGER)
(default 0))
(slot f_correct (create-accessor read-write)
(type INTEGER))
(slot f_mistake       (create-accessor read-write)
(type INTEGER))
(slot relation        (create-accessor read-write)
(type INTEGER)))
```

Example 1. Simple structure related to the "Lexemes and Suffixes"

The role of these structure elements in the whole system can be viewed in figure 2 where the interrelations between the basic software tools are shown. As it is represented, the communication between the interface (Authorware presentation) and the inference module is always performed through the database. The CLIPS system is checking continuously a flag in a control data structure in the database. When the interface determines that a new inference is needed to change the state of the presentation, it changes this flag and waits until the CLIPS system produces all the required inferences. The result is to fill a number of required templates that are read by the Authorware module to change the media components and consequently the state of the presentation.

4. AN APPROACH FOR THE DIAGNOSTIC AND TREATMENT OF THE DYSLEXIA.

The elements described above could be applied in general for any Intelligent Tutorial System with adaptive requirements. In this part of the paper, the particularities about the specific task of dyslexia diagnose and treatments in children together with the solutions proposed and implemented are discussed.

4.1 EVALUATION MODULES.

According to the double route model of Colheart, the first problem to solve is to determine if the child can be categorized as affected by phonologic dyslexia, surface dyslexia, a mixture of both or the child is not affected by dyslexia. The distinction between the two subtypes of dyslexia is important because deficits in the reading activity of the children in each category are of different nature. The phonologic dyslexic has a selective deficit reading pseudo words or unfamiliar words, while the surface dyslexic has a selective deficit reading irregular words despite the fact that he reads words and pseudo words correctly.

The SICOLE environment classifies the student in one of the categories mentioned above through a computer monitored naming exercise. In the naming task, the student is asked to pronounce a sequence of words and pseudo words. The words are shown in the screen one after the other repeating the following operations: a blank screen is shown (200 ms.), a sound is emitted alerting the student about the immediate occurrence of the following word to read, the word is shown in the middle of the screen surrounded by a rectangle (there is a soft color in the background to show a comfortable appearance in contrast with the word color) and finally the computer acquires and stores the word pronounced by the child, registering the reaction time (RT), i.e. the time from the moment the word is shown until the moment the child starts the reading activity. The time between the stimuli varies since the student must press a key to indicate that he has pronounced the word. That avoids a word would not be completely recorded as it will occur if recording time is fixed.

It must be remarked that once the sound is recorded, it is also analyzed using speech processing techniques for voiced and unvoiced segmentation. In order to identify an accurate RT, it is necessary to use signal processing techniques and calculate parameters as the Short Time Energy and Zero Crossing Rate values. Although the program determines other important features as the beginning and ending of each syllable, the most useful parameter in the actual implementation, is the reaction time (RT). Children with higher RT in pseudo words but under the upper limit of the confidence interval defined for normal readers reading familiar words are identified by the system as phonological dyslexic (Seymor, 1984). On the other hand, those with higher RT reading familiar words but under the upper limit of the confidence interval defined for normal readers reading pseudo words are identified as surface dyslexic.

Once the system identifies the type of dyslexia affecting the student, it starts a functional evaluation to determine the processing deficits that cause of low reading performance. This part of the diagnostic subsystem has three modules, one for each processing level: lexical processing, perceptual processing and syntactic-semantic processing. Each one of these modules is also composed of different submodules where specific cognitive components affecting the different processing levels are evaluated.

Module Description	Submodules to detect specific reading deficits
Syntactic	Gender
	Number
	Order
	Functional Words
	Assignation of Syntactic Roles
	Punctuation
Auditory discrimination	Direct Syllabus
	Double Syllabus
	Rhyme
	Auditory discrimination in words
Syllabic Conscience	Final Syllabic Omission
	Final Trisyllabic Omission
	Initial Bisyllabic Omission
	Initial Trisyllabic Omission
Phonemic Conscience	Isolate
	Omit
	Synthesise
Knowledge of letters	Letters
Lexemes and Suffixes	Lexemes and Suffixes
Homophones Words	Homophones Words

Table 1. Classification of evaluation modules and submodules

The diagnostic modules can be viewed in table 1. The reading capabilities intensive evaluation student results are stored in the database and are used in the treatment modules. Taking into consideration the deficits discovered in each student, different treatments are applied.

4.2 TREATMENT MODULES.

As an example of a treatment module, an overview of the Syllabic Conscience Treatment Module (SCTM) is now given. This module has three submodules. There is a global set T of different tasks related to the SCTM. Each submodule is defined by a subset $Si < T$. Inside each submodule there are three different difficulty levels. A difficulty level is defined by the way in that a set of activities are combined in a presentation. For instance, a presentation of the most difficult level is a sequence of activities where all the possible deficits in the SCTM framework are combined, while a presentation in the easiest level has activities related to only one deficit. With this strategy, the training of the children with more problems is improved by focusing his attention in solving exercises with only a class of difficulties.

The first submodule is the most difficult one, and it is accessed by the student after the evaluation stage. It has three difficulty levels and the student will access to the adequate level depending on the evaluation stage results. The measure that determines which difficulty level is the correct for a student comes from the percentage of correctly performed activities in a presentation (sequence of activities). Our system considers that the SCTM is necessary for the children if their result in the Syllabic Conscience Evaluation Module (SCEM) was lower than 70%.

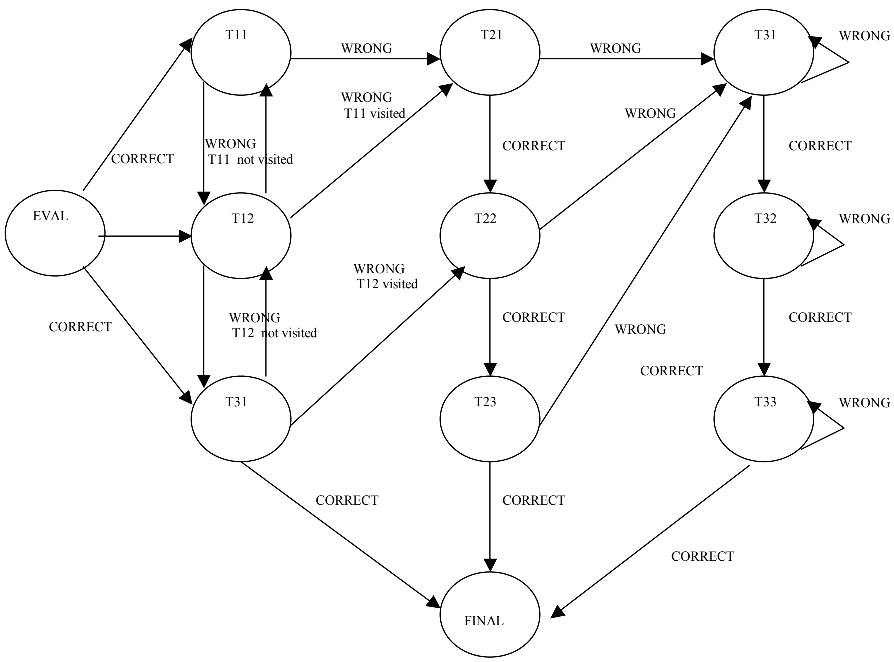


Figure 3. Graph showing the transitions between submodules and difficulty levels in the Syllabic Conscience Training Module.

Once the level i of the first submodule has been assigned to the student a sequence of randomly generated multimedia activities are started. The activities are chosen in a random manner, but under the requirement that at least one activity belongs to each task in the set S1. In general the system drives the student through difficulty levels and submodules. Figure 3 shows a diagram where the transitions are represented. Tij means submodule i, difficulty level j. For example, if the student has been assigned to T11 and performed the presentation of this place under the minimum percentage of correct activities he is driven to T21 (second submodule, first level), however if he passed correctly the presentation of T11 he is driven to the second difficulty level of the submodule 1, T12. Notice, that if the T12 presentation is not correctly performed by the student, the system distinguishes between two cases: if the student was previously in T11, he is driven to the second submodule (T21), while if he has not been in T11, the systems drives him to the T11 place. From this last point, it is clear the importance for the inference process of the history data stored in the database of the system.

5 RESEARCH ON THE IMPLEMENTATION OF AUTOMATIC SPEECH RECOGNITION IN THE TUTORIAL SYSTEM.

Nowadays the team is exploring how to take advantages of the current state of the art in the automatic speech recognition (ASR) technology. The main motivation of including ASR in the system is to extract the largest amount of diagnostic and treatment tasks relevant information from the waveform emitted by the child. A part of this technology has been already included in the prototype as was described above (naming task). Obviously, the goal of detecting automatically each phoneme pronounced by the child and the time parameters related to them (starting time and end time of each phoneme) is more ambitious and it will require a deeper research and even an improvement of the state of the art in the ASR technology. However, we consider valuable all the research efforts to include the ASR technology in the tutorial system for two main reasons:

-Improvement of the evaluation stage. Using the reaction time together with other parameters (for instance, time length of each phoneme, time separation between relevant phonemes in the word or the rhythm of the pronunciation) could be very valuable to get better results in classifying different types of children according to their reading capabilities.

-Improvement of the training stage. A better evaluation of the progress in the reading capabilities obtained by the system users would made the training stage more efficient. In this sense, the previously mentioned parameters are also useful.

It must be remarked that we are not trying to produce activities where all the interaction with the user is based in the ASR (for instance, pronouncing a word and automatically detecting the mistakes made by the user). According to our evaluation of different commercial and research ASR environments, the ASR technology under the established working conditions for SICOLE is not enough robust to make the system completely dependent on it. It must be considered the working conditions mentioned in the introduction of this paper, what makes necessary several features for the ASR:

-Isolated word recognition. Evaluation and training is most of the time based on reading correctly a sequence of isolated words and/or pseudowords. This implies that ASRs that make use of contextual information (sentence where the word is included) to reduce the mistakes are not valid for this application. The error rate of ASRs without contextual information is sensibly higher.

User independent recognition. Many ASR environments reduce the error rate through a previous supervised training with samples of the user utterance. Taking into consideration again the working conditions and the fact that the users are not good readers, a supervised training stage could be inapplicable.

-Phonemic level recognition. The evaluation and training stages require information of the utterance at the phonemic level. This discards those ASRs that are specifically designed to recognize at the word level.

With these ideas in mind, a tool to study the applicability of the ASR to the evaluation of dyslexic children has been developed. The chosen ASR environment has been the CSLU toolkit. Some of the benefits evaluated for this election have been:

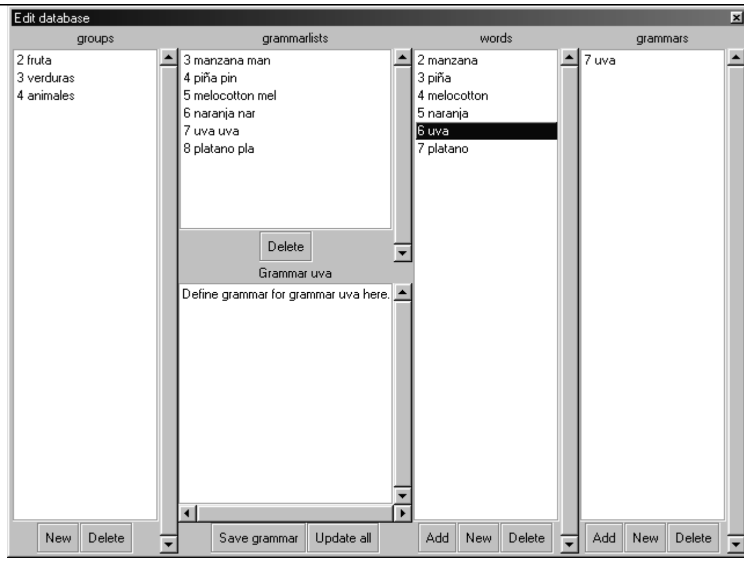
- Fulfillment of the three previous requirements together with high accuracy and good technical support. Furthermore, this package is used by big organizations to develop their own ASR products.

- Open environment, not only from the application development point of view, but also for building new corpora. It also supports the Spanish language.

- Different levels of application development: visual environment for rapid application development (RAD), scripting language (TCL/TK) and finally C language.

The main objective is to present to the user a naming task. The words are read by the user and the utterances are analyzed by the prototype showing the most likely phonemes recognized, their time parameters and the confidence of the recognition. Each word in the naming list could have several related grammars. A grammar contains symbolic information that describes the expected utterance at the phonemic level. Grammars are flexible enough to consider many natural variations of the pronounced word. Using several grammars for the same word is a requirement because in our case (dyslexic readers) is not rare that the same word will be pronounced in many different ways. All the information about the users and the recognition results are stored in a database for further analysis.

Figure 4 (upper part) shows one of the control screens of the prototype where an administrator can build lists of words with different grammars for each word. In the same figure (lower part, left) an example grammar for the Spanish word “plastilina” is also given. Regular expressions are used to specify the utterance. Notice how the expected utterance can be obtained from basic building blocks, which can be combined allowing repetitions or omissions and including other “special sounds” as “garbage” or “silence”. In the same figure (low part, centre) it is shown the output information: first column is the phonetic sound description, the following columns are the start and end times, and the confidence level. In the right side part, this information is graphically represented: each point represents a phoneme, the Y axis is the confidence of the recognition for each phoneme and the X axis is the time for each phoneme.



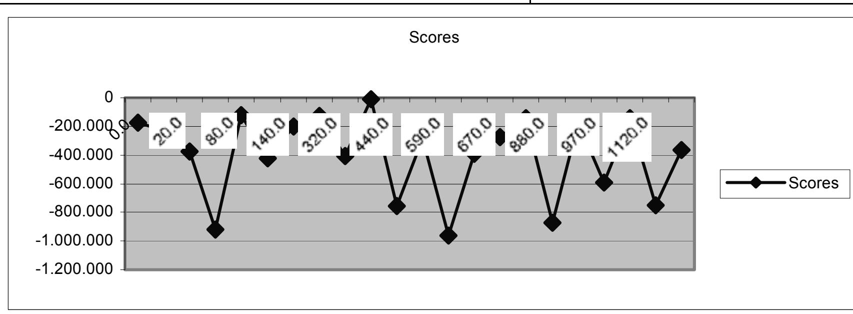
Control Screen for tasks and grammar design.

plastilina {$s1= {plas | ples | plis | plos | pas};
$s2={ta | te | ti | to | tu | tel};
$s3= {ca | co};
$s4={la | le | tili | lo | lu};
$s5={na | ne | ni | no | nu};
$sep = *sil%% [*any%%] *sil%%;
$grammar = [$sep] <[$s1]> [$s2] [$s3] [$s4] [$s5] [$sep];
}

Example of a grammar defined for the word “plastilina”

Phonem	Begin	End	Score
<.pau>	0.0	10.0	-173.562
<.pau>	10.0	20.0	-225.333
<.pau>	20.0	40.0	-373.909
$m_l8<m	40.0	80.0	-918.376
m>$m_r2	80.0	120.0	-122.279
$a_l11<a	120.0	140.0	-424.705
<a>	140.0	270.0	-198.377
a>$a_r6	270.0	320.0	-128.869
$n_l0<n	320.0	370.0	-406.841
n>$n_r0	370.0	440.0	-12.483
<.pau>	440.0	470.0	-753.486

Numerical results for an utterance.



Graphical representation of an utterance measured with the prototype

Figure 4 . Control screen, grammar example and measurements carried out with the prototype for researching ASR applicability.

6. CONCLUSIONS AND STATE OF THE PROJECT.

This is a research about how a multimedia system can help in the diagnosis and treatment of the developmental dyslexia. We have developed the basic infrastructure to implement simple and complex pedagogical strategies. The software architecture designed is now supporting several modules that are able to establish an automatic diagnostic about the dyslexic subtypes and also to perform training routines to improve several reading deficits.

We have learned from this research several important points in the technical aspects:

- It is a primary condition for any system of this type to be flexible enough to support a wide range of different solutions to the variety of individual dyslexic cases with different features. This has leaded us to select the architecture described above to implement SICOLE.

- One of the main benefits of this architecture is that it can incorporate new strategies and modifications to the implemented modules allowing the psychologist for experimenting different approaches for the expected problems.

- The inference module based on a KBS has been found very useful to provide the system with the capability of composing dynamic presentations: tasks and media components are chosen according to the inference results based on both static and variable information stored in the database.

- In this sense, the general strategy of using two dependent templates, one to be filled by the inference module and the other to be filled by the interface software has given flexibility in two different scales. The template filled by the KBS serves as the ground to complete the second template at the moment of generating the presentation.

- The speech recognition technology has been also applied in several levels: basic parameters of the sound waveform (reaction time) are used to classify the dyslexic subtypes, while more ambitious applications as detecting mistakes at the phonemic level in the user utterance are still researched through an implemented prototype where the using of multiple grammars and interpretations has demonstrated to be a practical strategy.

Although the system has already been verified through the critic examination of a group of expert psychologists, at this moment the system is being used in some Schools with the purpose of validating and measuring the real benefits of SICOLE. We are using a reading level match design with pretest-posttest control groups. The effect of SICOLE in the reading proficiency will be analyzed in a sample of students with learning difficulties in the reading activity divided as follows:

a) Experimental group composed of 30 retarded readers with ages between 8 and 9 years old coming from the second cycle of the Primary School.

b) Control group for the chronological ages composed of 30 students with the same features as the ones in the experimental group.

c) Control group for the reading level composed of 30 normal reader students with ages between 6 and 7 years old coming from the first year of the first cycle in the Primary School.

The experimental group is being trained with the SICOLE system, while the control groups are trained with other multimedia software in activities that are not related to the reading activity. The dependent variables are:

a) Reaction time in lexical decision tasks.

b) Latency time in word reading tasks.

c) Mistakes in the graph-phonemic decoding of the words.

d) Measurements of the reading level (i.e. lexical, syntactic and semantic processes).

ACKNOWLEDGMENTS

This research has been founded by Fondos Europeos para el Desarrollo Regional (FEDER), 1FD97-1140, and the Dirección General de Investigación Científica y Técnica (DIGICYT), Spain

We are specially grateful to the group of expert psychologist of our research team: Juan E. Jiménez, Alicia Díaz Megolla, Eduardo García Miranda, Adelina Estévez Monzó, María del Rosario Ortiz González, Mercedes Rodrigo López, Remedios Guzmán Rosquete and Isabel Hernández-Valle.

REFERENCES

Bordegoni M., Faconti G., Rist T., Ruggieri S., Trahanias P., MD Wilson (1996), "Intelligent Multimedia Presentation Systems: A Proposal for a Reference Model Multimedia Modelling: Towards the Information SuperHighway p.3-20 JP Coutaz, M Diaz, P Senac (Eds), World Scientific Publishing Co., Singapore.

Castles, A. y Coltheart, M. (1993). Varieties of developmental dyslexia. Cognition, 47, 149-180.

Cisero, C.A., Roger, J.M., Marchant, H.G.y Jackson, S.J .(1997). Can the computer-based academic assessment system (CAAS) be used to diagnose reading disability in college students?. Journal of Educational Psychology, 4,599-620.

Coltheart, M. (1978). Lexical access in simple reading tasks. En G. Underwood (ed.), Strategies of information processing. London: Academic Press.

Cuetos, F., Valle-Arroyo, F., y Suárez, M.P. (1996). A case of phonological dyslexia in spanish. Cognitive Neuropsychology, 13 (1), 1-24.

Gonzalez C.S. (2001), Sistema Tutorial Inteligente para niños con discapacidades intelectuales y cognitivas. Tesis doctoral. Universidad de La Laguna. Spain.

Jiménez, J. E. y Ramírez, G. (2001). Identifying subtypes of reading disability in a transparent orthography. Remitido para su publicación.

Jiménez, J.E. y Hernández-Valle, I. (2000). Word identification and reading disorders in the Spanish language. Journal of Learning Disabilities, 33, 44-60.

Karagiannidis C., Koumpis A., Stephanidis C. (1995), Media/Modalities Allocation in Intelligent Multimedia User Interfaces: Towards a Theory of Media and Modalities, in Proc First International Workshop on Intelligence and Multimodality in Multimedia Interfaces (IMMI-1).

Karagiannidis C., Koumpis A., Stephanidis C. (1997), Decision Making in Intelligent User Interfaces, in: Third Int. Conf. on Intelligent User Interfaces.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32, 1-19.

Moreno L., González C.S., Aguilar R.M., Estévez J.I., Sánchez J.L., Barroso C. (2000), "Adaptive Multimedia Interface for Users with Intellectual and Cognitive Handicaps". Lecture Notes in Computer Science 1839. Gilles Gautier, Calude Frasson, Kurt VanLehn (Eds.) of Intelligent Tutorial Systems. p.363-372. 5th International Conference. ITS 2000. Montréal, Canada.

Schreiber G., Wielinga B., Breuker J. (1993). KADS. A Principle Approach to Knowledge- Based System Development. Ed. Academic Press Limited.

Seymor, P.H. K. y McGregor, C.(1984). Developmental dyslexia: A cognitive experimental analysis of phonological, morphemic and visual impairments. Cognitive Neuropsychology, 1, 43-82.

Stanovich K.E. y Stanovich, P. J. (1996). Rethinking the concept of learning disabilities: The demise of aptitude/achievement discrepancy. En D. R. Olson y N. Torrance (eds.), The handbook of educational and human development (pp.117-147). Oxford: Blackwell.

The CSLU Toolkit is a product of the “Center for Spoken Language Understanding” at the OGI School of Science and Technology (Oregon Health and Science University). http://cslu.cse.ogi.edu.