Implications of a Formative Evaluation for Improving the Educational Effectiveness of an Educational Game for Spanish Orthography

Cristina Carmona, Eva Millán

Departamento de Lenguajes y Ciencias de la Computación

Universidad de Málaga

{cristina, eva}@lcc.uma.es

Abstract

An educational game is a recreational activity designed to teach people, typically children, about a certain subject or help them learn a skill as they play. These games are usually successful in engagement, but sometimes fail in triggering learning. This paper shows the results obtained when evaluating MITO, an educational game to teach Spanish orthography. This experiment let us find the main deficiencies of MITO and opens a new line of work that tries to fix them.

1. Introduction

Electronic educational games are learning and recreational environments that try to increase the learner's motivation by embedding pedagogical activities in highly enjoyable interactions. In [1], a former review of literature on the effectiveness of games versus traditional classroom instruction is made. By analysing 68 studies from 1963-1991 in social science, math, language arts, logic, physic and biology, it was determined that games were more effective in the domains of language arts and mathematics. The reason is that, in these fields, greater specificity of content and more effective use of computers created a clear advantage for the exercises over traditional teaching methods. Furthermore, greater retention was shown in games, and students reported greater interest in game activities than in more conventional classroom instruction. However, there is little empirical evidence that electronic educational games can promote learning unless the interaction is led by teachers and integrated with other instructional activities [2]. A possible reason for this limitation is that learning how to play does not imply learning the domain [3].

Another important point that must be taken into account is that the educational game should adapt the proposed activities to the learner's cognitive stage. Probably, one of the most accurate and commonly used

definitions of children cognitive stages is the defined by Piaget's theory of cognitive development [4][5]. The cognitive abilities that Piaget studied are important when teaching young students, as they help determine how much and in what way students will understand the topic being taught.

Examples of educational systems designed for children and based on Piagetian stages of cognitive development are not easily found in research. An interesting work in this field is presented in [6]. Later on, same authors presented in [7] an independent, adaptive, and easy-to-integrate web-based component to evaluate a student's cognitive development that can be used as the pre-test of any ITS This component was constructed by including existing test's items [6] into the SIETTE [8] web-based adaptive testing system.

2. MITO

MITO¹ (which stands for Multimedia Intelligent Tutor of Orthography) is a stand-alone application focused to help children between 8 and 12 years old learn Spanish orthography [9].

According to the American Heritage dictionary, the definition of the word orthography is:

- The art or study of correct spelling according to established usage.
- The aspect of language study concerned with letters and their sequences in words.
- A method of representing a language or the sounds of language by written symbols; spelling.

But in the case of Spanish, the word "orthography" has a different meaning. The definition of the "Real Academia de la Lengua Española" dictionary could be translated as follows:

• Set of rules that determine the writing of a language.

¹ MITO is a tool designed with educational and non-commercial purposes. Installation files and instructions are accessible at http://www.lcc.uma.es/~cristina/mito.

The most significant difference between English and Spanish orthography, is that, in the case of Spanish, there are rules that determines how words are spelt. However, from the nearly 600 rules that exist, there are only a few really useful (in the sense that they contemplate many words and few exceptions, for example, rules for accents). In [10], authors did an study with a basic vocabulary consisting of 674 words to determine how many of them were covered by a rule, and found that, if the more general rules were not considered, there were only 48 words out of the 674 covered, i.e., the probability that a randomly chosen word is covered by a rule is only 0.07. On the other side, it is known that there is a very close relation between orthography and visual memory. The study presented in [11] states that, in orthography, the 83% is learnt visually, the 11% is learnt by audition, and only 6% is learnt using other senses.

So our view of how orthography should be taught is not based in learning the rules, but shares the directions presented in [12], which we briefly explain next: to our purposes, orthography will be used as a synonymous of "correctly spelling the words". Therefore our main goal will be to teach the strategy that persons with good orthography apply. Solving exercises is not effective if the mental strategy is wrong, so the fundamental thing is to teach a mental process, which can be described as follows: When in doubt, good writers mentally search for the image of the word and visualize it in their minds. Writing becomes then a recall of the word previously stored in their minds.

This visual recall of the word can be good enough to be confident on its correctness, but in some cases good writers clarify doubts by writing the word in different ways. Consequently, our initial goal when developing MITO is to help in learning how to use visual memory for orthography (by showing correctly and incorrectly written words). But, given that Spanish orthography does have rules, MITO uses them. Words to be shown in MITO exercises have been chosen so that there is always a rule that determines its correct spelling.

As also stated in [12], in order to learn Spanish orthography, the student must satisfy some conditions:

- Be advanced in the concrete operational stage (about eight years),
- Being able to write and read at a reasonable speed,
- Being aware of the existence of rules and exceptions,
- Being motivated to improve (though our system is designed to improve motivation).

The contents in MITO are divided in four modules that aim to teach words corresponding to different sets of rules: a) Written accent rules; b) H, G and J rules; c) B, V, C, Z, D and Q rules; and d) M, N, Y, LL, R and RR rules. As explained before, rules are not learning goals on themselves, but a way to provide support to being able to correctly spell the words. In this sense, rules are used in our system to group the exercises to be posed to the student (Fig. 1).



Fig. 1 Screenshot of the main menu

For each learner, MITO creates and maintains a simple student model that keeps record of the number of correctly solved exercises of each module. In order to achieve the learning goals fixed, MITO presents a series of features that have been designed to take into account good practices in educational games Such good practices were identified in a literature review (for details see [9]), and are represented here by keywords. Below is the list of such features, together with the keywords that describe the good practices they intend to promote.

- The user can choose a different module at any moment (NAVIGATIONAL FREEDOM, INTERACT IVITY).
- Each exercise is explained briefly, but a more complete explanation can be shown by user's demand (SIMPLICITY, ACCESSIBILITY).
- The user's answers are checked immediately (IMMEDIACY).
- Each group is divided in several different exercises (VARIETY) that progressively get more complicated (i.e., apply more or more difficult rules) as the user correctly solves them CHALLENGING). In order to trigger learning, the user must show knowledge about between 90% and 100% of a module's content (either by visual memory or by direct application of its rules) to be able to finish it (LEARNING EFFECTIVENESS).
- The system has a help mechanism that explains the rules in an accurate and clear way. This help is context depending and can be shown by user's demand (SIMPLICITY, SUPPORT).

- Feedback messages are provided immediately (for both the cases of correct and incorrect answer) and are clear and brief (orthography rules are used as a basis for explanations) (SUPPORT, IMMEDIACY).
- There are many multimedia elements (childish sounds, nice pictures, etc.) that keep the user motivated.(MOTIVATION)
- As the user goes on with the game, his/her knowledge about the topic increases and the system adaptively selects more difficult exercises. (ADAPTATION, CHALLENGING, WELL-STRUCTURED)
- The system keeps a simple student model that is updated in every exercise. (OUTCOMES)
- At any moment, the user can see his/her student model. (OUTCOMES, ORIENTATION)
- Each time the user correctly solves an exercise, some points are granted (REINFORCEMENT).
- MITO keeps record of the score reached by its registered users, which is available at any time (ranking). In this way, competition triggers motivation and keep the children engaged. (MOTIVATION)

Fig. 2 shows an exercise in MITO. The relevant elements are (from left to right and from up to down): user name, exercise stem, points reached, motivational feedback, window of the exercise, motivational goal (in this case, Minnie must reach Mickey and each correct answer will move her one position forward), number of correct and incorrect answers and four buttons: main menu, show results, help, and how to play (that provides an extended game description). When the student selects a word in the exercise, the background is changed to green if it is correct and to red if incorrect and a childish voice congratulates or scolds the action while appropriate feedback is shown.



Fig. 2 Screenshot of an exercise

3. Evaluation

A first formative evaluation of the MITO system has been performed. This evaluation had two main goals: a) to determine the degree of acceptance of the game among the users it was focused to, identifying relevant aspects that can improve learner's motivation and b) to study the effectiveness of the game in helping children to learn orthography, identifying possible ways to improve the design and behavior of the system from an educational point of view.

Sixteen children from the school "CEIP Ricardo León" in Málaga took part in the experiment. The children belong to 5th grade (10 years old) and used the game during 15 minutes. Before the game, they had to take a pre-test with 40 items that are grouped according to the rules applied: 5 items for the B, V rule; 5 items for de C, Z, D, Qu rule; 5 items for the M, N rule; 5 items for the R, RR rule; 5 items for the written accent rules; 5 items for the H rule; and 5 items for the G, J rule. Each item is a word partially written, and the child has to complete the missing letter. After the game, the children answered a post-test with the same 40 items plus four questions (free text) about the game: 1) Did you like the game?; 2) Did you ask for help?; 3) What did you like most?; 4) What did you like least?

While the children were playing, a human tutor observed their behaviour and assisted them when needed (children had different computer literacy degrees, so some of them needed more support than others, though in general we found that the system was very easy to use and children only needed assistance at the very beginning).

Regarding the motivational aspects, first results have been very encouraging. Children did really become very engaged with the game. The features of the system that they liked better were the multimedia aspects, i.e., Disney©² characters and childish sounds (which fully fulfilled their goal to capture children's attention). The fact that they could help their favourite character to reach a goal was also very motivating, as also was the possibility to see their position in the ranking. When asked for features they did not like, only one student pointed out that the explanation of the exercise was not clear, seven students explicitly said "I liked everything" and the rest left the question without answer. However, results in the use of feedback/help messages to trigger learning were not so good. More than half of the children did not want to pay any attention to feedback messages, even when pointed out that these messages could give a clue for the next exercise. It was somehow surprising for us that children did not want to use the "help" facility either, even when it could help them to get the points they wanted so much to beat their partners. When reaching an impasse,

² Disney characters were taken from http://clipart.disneysites.com and are used in our application with no commercial purposes.

most children asked their tutor for the correct answer, but very few of them tried to use help (only 3 out of the 16 involved in the experiment). In summary, children only wanted to quickly get the correct answer from their tutor and keep playing. In our opinion, the reason is that, even when help messages are short and clear, children were too involved with the game and did not want to waste their time in reading. However, during the experiment we observed that children were willing to accept help instead of the correct answer from their tutor, if it was presented as a "trick" that could provide the answer to not only the current exercise but probably to the subsequent ones. This suggested that the use of a personal helping agent (as in [13]) could improve the effectiveness of the system from an educational point of view. The helping agent could be a nice character, assigned to the child at the beginning of the interaction, and presented as a friend that can be called when needed to provide help by means of "tricks" (simplified orthographic rules).

Regarding the effectiveness of the game, Table 1 shows the number of mistakes made by the students in the pre and post-tests. As can be seen, after playing with the system, 8 students make it better, 4 make it worse and 4 stay the same, being the mean rate of improvement of 5.3% and of impairment of 5.6%.

Learner ID	Pre-test	Post-test
1	13	14
2	7	10
3	10	10
4	13	11
5	6	5
6	8	10
7	11	11
8	11	11
9	15	12
10	7	5
11	7	6
12	10	13
13	2	2
14	18	14
15	7	6
16	14	12

Table 1 No of mistakes in the pre-test and post-test

In a second analysis of the same data, we have given punctuation to students according to their performance, so we can deduce how much they learn:

• $-1 \rightarrow$ if a word is correct in the pre-test and incorrect in the post-test (UNLEARN)

- $0 \rightarrow$ if a word is correct (incorrect) in the pre-test and correct (incorrect) in the pos-test (NO CHANGE)
- 1 \rightarrow if a word is incorrect in the pre-test and correct in the post-test (LEARN)

So, if the punctuation is less than -2, the student unlearn a lot; if it is between -2 and -1, they unlearn a bit; if it is 0, they do not learn; if it is between 1 and 2, they learn a bit; and if it is more than 2, they learn a lot. Table 2 shows the percentage of students in each level. These results show that the 50% of the students learn and the rest either no learn (25%) or unlearn (25%).

% Learners	Status
12.5 %	Learn a lot
37.5 %	Learn a bit
25 %	No change
12.5 %	Unlearn a bit
12.5 %	Unlearn a lot

Table 2 Percentage of students in each status

Before the experiment, it seemed clear to us that a fifteen minutes session was not enough to change student's knowledge significantly. But these results seem to show that, all together, students learn as much as they unlearn. In our opinion, the most plausible explanation for this is that students did not take much interest when answering either the pre-tests or the post-tests, which confirms our subjective impression during the experiment. On the other hand, we have also identified a potential problem here: MITO's exercises show both correct and incorrect words (for the student to detect errors), which might be inconvenient for novice users because it interferes with the correct building of a visual representation of the word. This problem might also explain why some students seem to "unlearn" while using the game. Finally, we are convinced that the scarce use of feedback and help facilities has contributed against learning effectiveness.

4. Conclusion and Future Work

Our main goal while developing MITO was to build an application for children that could help them in the somewhat tedious but very important task of learning Spanish orthography. To this end, motivational aspects have had a great importance in the design and development of the system, and a first evaluation shows encouraging results regarding motivation. But more importantly, this evaluation also showed some weaknesses of the system that will definitely influence the design of its subsequent versions:

- The current version of the system shows incorrectly written words to all students. We have learned that exercises based on identifying orthographical mistakes must be proposed only to advanced students, because it can be inadequate for students that still do not have correct visual representations of the words in their minds.
- The design of the system does not sufficiently encourage students to use feedback and help, which are designed to trigger learning. We have learned that this facilities need to be redesigned so that the interference with game playing is minimized. In this sense, we plan to use helping agents. In the evaluation, the human tutor played this role with satisfactory results, which encourages us to pursue this direction.
- The architecture and the user modeling techniques used are too simple and therefore the adaptive capabilities leave room for improvement. Currently the system only stores the student score (number of correct words) in each game. The solution is clearly to implement a better user model. Based on this user model, students will be classified into categories and the adaptive capabilities will be improved, so for example exercises for identifying incorrectly spelt words will only be shown to advanced students.
- Finally, another important issue is that the evaluation plan needs to be redesigned so we can obtain significant conclusions in the summative evaluation. In this sense, future versions of the system (probably web-based and designed to store the logs of each session) will be distributed in several schools so students can use it for a whole semester or academic year. In this way, we will have many more users and each of them would have use the system for a longer time. In order to obtain more reliable measures of the learning gain, pre and post test will be integrated as part of the game and include motivational aspects (each correct answer will give the kid points for the ranking, as we identified the ranking as one of the main sources of motivation).

Other improvements of the system we are planning include the development of an interface for tutors/parents that allows the inclusion of new words, to provide a greater variety of the exercises posed to the student. However, we must make clear that we are using the Spanish Orthography domain only as a test-bed, because our ultimate goal is to identify the kind of interactions and features of an educational game that promote significant learning, so that we can make an abstraction of such features to be included in a more general architecture that allows the

implementation of effective educational games for different domains.

5. References

- [1] J.M. Randel, B.A. Morris, C.D. Wetzel, and B.V. Whitehill, "The Effectiveness of Games for Educational Purposes: A Review of Recent Research", *Simulation & Gaming 23*, 1992, pp. 261-276.
- [2] M. Klawe, "When does the use of Computer Games and other Interactive Multimedia Software help students learn Mathematics?", NCTM Standards 2000 Technology Conference, Arlington, VA, USA, 1998.
- [3] C. Conati and J.F. Lehman, "EFH-Soar: Modeling Education in Highly Interactive Microworlds", *Advances in Artificial Intelligence, Proceedings of the third Congress of the Italian Association for Artificial Intelligence*, AI-IA'93, 1993, pp. 47-58.
- [4] J. Piaget, "How children form mathematical concepts", *Scientific American 189* (5), 1953, pp. 74-79.
- [5] J. Piaget, "Piaget's Theory", *Carmichael's Manual of Child Psychology*, third edition, volume 1. New York: Wiley, 1970.
- [6] I. Arroyo, J.E. Beck, K. Schultz and B.P. Woolf, "Piagetian Psychology in Intelligent Tutoring Systems", *Proceedings of the Ninth Artificial Intelligence in Education*, AIED'99. Le Mans, France, 1999, pp. 600-602.
- [7] I. Arroyo, R. Conejo, E. Guzmán and B.P. Wolf, "An adaptive web-based component for cognitive ability estimation", *Proceedings of the Tenth International Conference on Artificial Intelligence in Education*, AIED 2001. San Antonio, TX., 2001, pp. 456-466.
- [8] R. Conejo, E. Guzmán, E. Millán, M. Trella, J.L. Pérez-De-La-Cruz and A. Ríos, "SIETTE: A Web-Based Tool for Adaptive Testing", *International Journal of Artificial Intelligence in Education, 14*, 2004, pp. 29-61.
- [9] E. Millán, C. Carmona, R. Sánchez and J.L. Pérez-De-La-Cruz, "MITO: an Educational System for Learning Spanish Orthography", *Workshop on Educational Games as Intelligent Learning Environments*, AIED'05, Amsterdam, The Netherlands, 2005, pp. 39-47.
- [10] M.J. Esteve and J.M. Jiménez, "La disortografía en el aula", *Editorial Disgrafos*, Alicante, 1988.
- [11] J. Mesanza, "Didáctica actualizada de la Ortografía", *Editorial Santillana*, 1987.
- [12] D. Gabarró and C. Puigarnau, "Nuevas estrategias para la enseñanza de la Ortografía en el marco de la PNL", *Editorial Aljibe*, Archidona-Málaga, 1996.
- [13] C. Conati and X. Zhao, "Building and evaluating an intelligent pedagogical agent to improve the effectiveness of an educational game", *Intelligent User Interfaces*, 2004, pp. 6-13.