A Serious Game for Second Language Acquisition in a Virtual Environment

Marilisa Amoia, Treveur Brétaudière, Alexandre Denis, Claire Gardent, Laura Perez-Beltrachini

A Serious Game for Second Language Acquisition in a Virtual Environment

Marilisa AMOIA, Treveur BRETAUDIERE, Alexandre DENIS, Claire GARDENT, Laura PEREZ-BELTRACHINI
INRIA Nancy Grand Est, CNRS,
54500 Vandoeuvre-les-Nancy, France

ABSTRACT

In this paper, we present I-FLEG, a 3D language game designed for interactively learning French as a second language. I-FLEG differs from previous computer-aided language learning (CALL) approaches in that it combines a situated, language learning environment with advanced artificial intelligence and natural language generation techniques which support user adaptivity and the automatic, context-aware generation of learning material. In addition, because it is integrated in a 3D virtual reality environment, I-FLEG naturally supports e-learning and facilitates the collection of test data.

Keywords: Virtual Learning Environments, Computer Assisted Learning, Intelligent Tutoring Systems, Immersive Learning.

1. INTRODUCTION

As argued by e.g [15] immersive experiences play a fundamental role in learning a foreign language. This is because learning a foreign language is a complex process which requires not only the acquisition of the grammar rules of the target language (L2) but also the knowledge of the societal and cultural habits of the native speakers of that language. That is, a successful interaction in the target language L2 is one that is aware of the cultural differences and fulfills both the semantic and the pragmatic constraints of L2.

As the technology for 3D simulation and virtual reality is now mature and routinely exploited in games and serious games, the question therefore naturally arises of whether and how 3D world and games could be used to support situated, immersive language learning. Moreover, 3D technology exhibits several features which are attractive for e-learning. It can be combined with web technologies thereby providing large, universal access to learning resources. It allows for synchronous and asynchronous learning thereby presenting an interesting alternative to classical teaching/learning environments. It has many of the characteristics of real-life experiences without their shortcomings (such as risk of injury). And thanks to its immersive character, it might help maintain learner motivation and engagement.

In this paper, we present I-FLEG, a 3D game designed for exploring the potential usefulness of 3D technology for language learning. I-FLEG differs from other existing 3D approaches to computer aided language learning (e.g., [24], [1] and [11]) in that it combines 3D technologies with sophisticated, natural language generation techniques for automatically producing learning material that is both varied and adapted to the current context. In addition, the game includes some basic scoring, error detection and error correction mechanisms which lay the basis for personalized feedback. And because it is web-based and relies on a database for permanent storage, it makes it easy to collect and analyze learner/game interactions.

The paper is structured as follows. Section 2 situates our approach with respect to existing work on Computer Aided Language Learning. Section 3 presents the architecture of the I-FLEG language game. Section 4 describes the learning strategies implemented in the game and Section 5 explains how learning material is dynamically generated from a description of the 3D world in the context of the game. Section 6 concludes and gives pointers for further research.

2. STATE OF THE ART TECHNOLOGY FOR SECOND LANGUAGE ACQUISITION

In the last few years, there has been an increasing interest for exploiting web, gaming and simulation technologies to teach foreign languages. Indeed, these technologies offer some interesting and innovative possibilities for teaching and learning foreign languages. By allowing long distance learning (e.g. through chats, blogs, emails exchange, etc.), they facilitate information exchange between learners and teachers thus highlighting the communicative aspect of language learning. By providing an immersion into reality-like virtual worlds, they give the learning experience a fun factor which is lacking from traditional face-to-face or book-based methods and might help maintain learner motivation and engagement. A closer look at current systems for second language acquisition shows three main research trends: systems providing contextualized learning; systems providing error detection and personalized feedback; and systems supporting the automatic authoring of learning content. We now review each of these trends and situate the I-FLEG approach.

Contextualized Learning

Interactive or situated learning has been acknowledged as one of the main pedagogical principles in second language acquisition. Thus [15] has shown that adult learners acquire a second language more easily and that their knowledge is better anchored if they are exposed during learning to real life like situations like those that children experience when acquiring their first language. And [21] and [22] have demonstrated that raising the awareness by the learner of phenomena specific to the target language e.g., by highlighting them, fosters learning.
The first experiments in situated computer-aided language learning (CALL) go back to the 1990's when the first CALL software were produced on CD-ROM. *Escape From Planet Arizona*[1] and *Who is Oscar Lake*[2] for instance, are first attempts to integrate traditional learning content (i.e. vocabulary and grammar exercises) in a situational context (an adventure story) thus producing in the learner the impression of an immersive experience.

More recently, some research on computer aided second language acquisition has focused on using online 3D virtual reality environments and video game technology for teaching languages. In particular, *Thethis* (see [23] and [24]) implements a web application for the language training of hotel receptionists. The learner is exposed to situations simulating the reception desk of an hotel. He can interact by means of preset dialogues with virtual agents simulating e.g., telephone calls and hotel guests arriving or by chatting with a human tutor or with other learners. In this way, the users can train oral comprehension, reading and classical grammar skills in a situated setting. The main innovative aspect of *Thethis* is the social, communicative aspect of the learning platform which allows learners to share their learning experience with fellow students, with human teachers and with virtual conversational agents. 3D video game technology has also been used in so called culturally-aware tutorial systems such as BiLAT [13] to train social cultural skills. In these approaches, the learner acts as an avatar in the simulated environment and must learn behaviors and relation building strategies which are conform with cultural norms. That is, the virtual world is used mainly as a mean to immerse the learner in a simulation of the societal and cultural world. The Tactical Language and Cultural Training Systems (TLCTS, [11]) goes one step further and provides a game based learning platform for acquiring functional skills in foreign language and culture. Finally, the Croquelandia [25] adventure game allows learners to interact with other students in Spanish whereby the conversational options selected by the learner determine the game flow. While TLCTS includes cultural interactions and grammatical exercises, Croquelandia focuses on pragmatics, teaching how to make request and to apologise in Spanish.

Error Detection and Personalized Feedback
Providing personalized feedback in a system for interactive second language learning is a difficult task which involves interpreting the, often ill-formed, learner input, detecting errors and proposing a correction. Consequently, most of the existing systems strongly constrain the user input so as to facilitate analysis, error detection and error correction. Among the systems performing error analysis, both *Robo Sensei* [19] for learning Japanese and *Spanish for Business Professionals* [9] have been used in classrooms. These systems provide a collection of exercises on grammar, phonetics etc. and make use of photos and computer animations to contextualize grammar training activities. These systems also have the ability to categorize learners errors and to provide user feedback with grammar explanation. *E-Tutor* [10], a system for learning German, improves on these two systems in that it can provide a differentiated feedback for each level of language proficiency; and *TAGARELLA [1]*, an intelligent web-based workbook for learning Portuguese, provides advanced adaptive error feedback for the following types of activities: reading comprehension, listening comprehension, picture description, rephrasing, fill-in-the-blanks, and vocabulary. In *TAGARELLA*, the generated feedback takes into account both the learner proficiency and the language content exemplified by the exercise.

Automatic Authoring of Training Activities
The authoring of training activities is labour intensive and requires expertise. In order to reduce these costs efforts, there has been an increasing interest in developing systems for the automatic authoring of training activities. Work in that direction includes *VISL* [2], a visual interactive syntax learning tool accessible though the internet for learning the syntax of different languages. [18] describe *WERTI*, a prototype system for the automatic generation of exercises based on arbitrary web content selected by the learner. The system can generate training activities such as reorder scrambled words or phrases and fill-in-the-blank exercises for some grammar category. The generation is based on more or less deep NLP analysis such as tokenization, part-of-speech tagging and shallow parsing of the arbitrary input. Another important trend is the automatic generation of multiple choice vocabulary cloze items from the web, from very large corpus or using thesauri (see [4] and [16]) or by applying word sense disambiguation-based methods to texts extracted from web documents (cf. [17]).

In sum, the research on computer-aided second language acquisition has been carried out into three main directions. First, there are approaches that focus on situated language learning and provide learning content in contexts recreating reality-like situations. Second, there is work on providing the learner with an adaptive, personalized feedback. Third, there are attempts to automatize the production of learning material.

The system we present in this paper integrates features from each of these three trends. It uses an immersive 3D environment to embed learning in a simulated reality. It monitors user/system interactions and maintains a learner model to provide the learner with a feedback that takes into account her language proficiency. And it exploits Artificial Intelligence (AI) and Natural Language Generation (NLG) techniques to automatically generate learning activities that varies depending on the learner level, on the current teaching goal being pursued and on the learner position in the virtual world.

3. I-FLEG: A SERIOUS GAME FOR SECOND LANGUAGE ACQUISITION

---

1 Escape from Planet Arizona: An EF Multimedia Language Game. [Software]. (1995). Stockholm, Sweden: EF Education.

2 Who is Oscar Lake: A Language Learning Adventure Game. (1995). Jersey COW Software.

I-FLEG is a serious game designed for exploring the potential contribution 3D games and NLP technology can make to CALL. In what follows, we start by describing the game scenario used for language teaching. We then describe the features, the goals and the coverage of this system, the game aspects and the system architecture. Finally, we show how I-FLEG supports a learning process that is *adaptive* (the system takes into account a learner model to vary its output), *situated* (the exercises vary depending on the current context), *free* (the learner can take the initiative) and *supervised* (the learner receives personalized feedback).

Game Scenario

The I-FLEG game can be played by connecting to the Allegro Island in Second Life[3]. This island, shown in Figure 1, is populated with houses of different colors whose content is identical, each of which representing a game unit. In this way, the game can be deployed in a classroom setting with several students playing simultaneously. The I-FLEG house is currently the main game unit (although it can easily be extended to other settings such as a restaurant, a cinema, etc.). It contains different rooms each corresponding to a particular lexical field to be acquired by the learner.

A first person perspective is used whereby the learner is an avatar that can freely move in the game world. In order to start the game, she enters one of the houses where she can then interact with the 3D world by clicking on objects and entering text in a chat box.

Figure 1. The I-FLEG Island.

Figure 2. The interior of an I-FLEG house.

The game responds through messages displayed in the Head-Up Display (HUD, Figure 3) which acts as a virtual tutor guiding the player throughout the game. The upper part of the HUD shows information about the current state of the game, the score of the player and the elapsed time. The main window in the center of the HUD is used for guiding the navigation of the learner through the game world, for communicating learning exercises to the learner, and for conveying feedback information. The menu on the right lists the grammar topics (also called teaching goals) that can be exercised in a given session. The content of this menu varies with the language proficiency of the user. Finally, on the bottom right corner of the main window, the "start evaluation" button can be used at any time to switch to evaluation mode.

[3] http://secondlife.com/

We differentiate between the navigational instructions generated by the system to guide the user through the different phases of the game and the system feedback on learner input. The system uses English as a language for guiding the user though the game and for explaining training activities but French, the target language, for feedback messages and error correction.

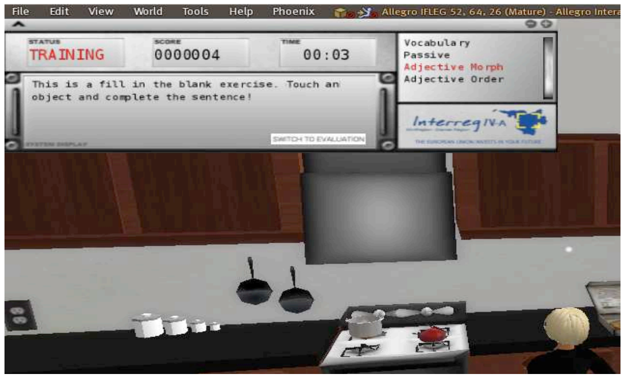

Figure 3. I-FLEG Game Hud.

Finally, I-FLEG can be used in four different modes: initialization, tutorial, training and evaluation. In the initialization mode, the learner either registers or identifies herself. Accordingly, the learner information is either stored in the database for later retrieval or displayed to the learner in the form of a greeting.
The tutorial mode permits revising vocabulary by clicking on objects and getting their name. This step can be parameterized to fit the learner level. For instance, instead of simply naming the objects, the system might produce a more complex noun phrase including e.g., an adjective or a relative clause once the learner has the appropriate L2 knowledge.
In training and in evaluation mode, clicking on an object triggers an exercise session with the system which consists of a system query, the learner answer and the system feedback. In training mode, the system provides the target answer if the learner makes a mistake. Moreover, the learner can practice the same exercise as many times as she likes. In evaluation mode on the other hand, the only feedback provided by the system is the update of the score. In addition, the learner can play each exercise only once (just as a student can only hand in one copy after an exam).

Goals and Coverage
The educational goal of I-FLEG is to facilitate the acquisition of French as a second language. Currently, it can be viewed as a "situated textbook" providing the learner with lexical and grammatical drills i.e., with exercises which can be used to practice lexical and grammatical skills. In addition, a tutorial mode permits revising vocabulary. Although in the current version of the game, interpersonal communication skills cannot be trained as the system only monitors player interactions with objects, we are currently working on integrating a dialog system into the game[4] to support learning by situated conversational interaction.

The current implementation targets learners of French at A1-A2 levels as defined in the FLE (French as a foreign language) standard i.e., beginners and intermediate level learners. In particular, the linguistic knowledge conveyed to the learners includes: lexical fields centered on some specific domains, e.g. *household items*, *food*, *clothes*; morpho-syntactic features such as verb governed prepositions or adjectival/verbal morphology; and basic syntactic constructions such as active, passive, intransitive, transitive and ditransitive verbs. As explained in Section 5 however, the modular architecture of the system and the computational grammar used to generate the exercises make it easy to rapidly extend the linguistic coverage of the system and even, to switch to another language. We could in particular modify the system to teach English by plugging in an already existing grammar of English and extending the lexical database with the appropriate lexical entries.

Game aspects
It has often been noted (cf. [5]) that such games features as the immersive, reality-like environment of the game, the autonomy given to the player in making game relevant decisions, and the fun aspect promoted by the game logic might make games particularly appropriate as environments for language learning. Accordingly, I-FLEG exploits the following game techniques to promote user-system interaction, to elicit the learner use of the L2 language in context and to maintain her attention:

- free exploration of the environment
- items gathering
- point-and-click interaction with objects
- scoring mechanism

---

4 Interaction with other avatars is possible but is not taken into account by the game.

The game techniques applied in I-FLEG include scoring and the storyline behind the game scenario. Although this aspect is not yet fully developed, I-FLEG is designed as an adventure game in which the player has to master a challenge. More specifically, the player is in the house of a crazy linguist who invented a mysterious machine capable of understanding all languages of the world. The linguist has disappeared but the parts of the machine are still hidden in the house. The task of the player then is to find out where they are and to collect them. Indeed, in the house, there are strange objects hidden, that are not part of the furniture. When touched by the learner, these objects trigger learning activities. The player has a limited amount of time to find out where they are, touch them and solve the learning activities linked to them. By solving exercises the player earns credits points (the score). This game structure permits modeling different learner level whereby each game level represents a language proficiency level in the second language (e.g. A1) and includes a set of training activities appropriate for that level in terms of lexical, morphosyntactic and syntactic coverage. A level is complete when the learner has totalized the score that defines the accomplishment of that language level. While the exploration, items gathering and point-and-click interaction with objects makes I-FLEG closer to the adventure game genre, the scoring mechanism and the fact that the scores of all players are displayed in the game area, makes it closer to the arcade game genre.

Game Architecture

The I-FLEG game has been implemented within the virtual world of SecondLife.

The choice of SecondLife has not been motivated by social interaction between players but rather by the technical convenience that SL provides. Indeed, SL provides altogether: a graphic library, a network library, a scripting engine, and a user management service. The scripting engine, being powerful enough to enable HTTP access in and out SL has been of considerable importance in the choice of SL. Moreover, SL eases a lot the content creation process by providing both internal tools to build objects and environments and a strong market to buy existing objects. These two facilities were worth to build the house since half of the objects were built on request to a builder player, while the other half has been purchased or freely obtained (aka *freebies*). Thanks to the inherently multiplayer framework that is SL, it is possible for several players to play the game at the same time in different houses. It thus allows classroom based teaching where all pupils could be trained in their own 3D house.

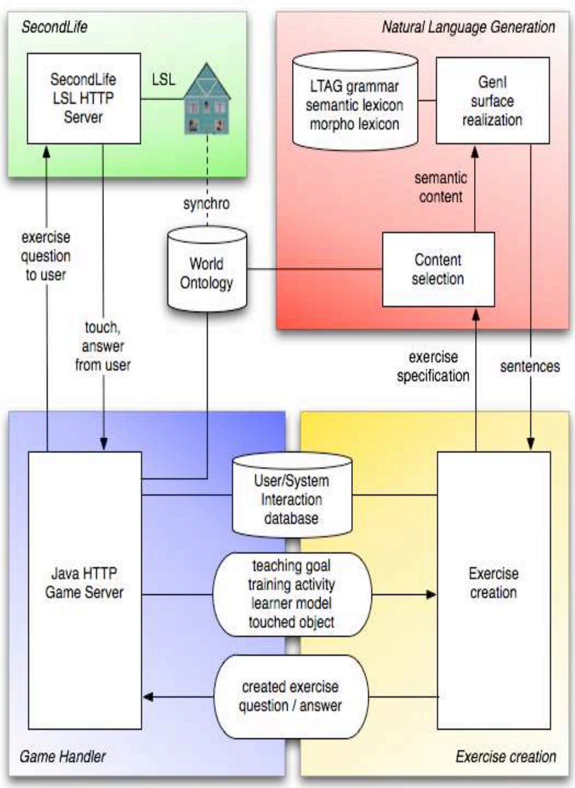

Figure 4. The Game Architecture.

Figure 4 shows the architecture of I-FLEG. The game includes four main modules: a Second Life LSL-HTTP server, the Java game server, a natural language generation module and a component for the generation of training activities. Second Life (SL) is used as a graphical interface providing the 3D virtual environment for the game. An HTTP-service manages the messages exchanged between the user, the SL platform and the Java game server.

The game server (Game Handler) implements the game logics and is responsible for the management of user/system interactions. It keeps track of the world state through an ontology synchronized with the game world visualized in the SL platform. The game server also

monitors the user and his interactions with the virtual world and stores in a database both a user model and the logs of the user/system interactions.
The Natural Language Generation and Exercise Creation modules are responsible for the automatic authoring of training activities. The Exercise Creation module is responsible for building the exercise instances relying on the parameters sent by the Game Handler module, i.e. the teaching goal, the training activity type, the learner model and the object touched by the learner. It further extracts from the game ontology the facts true of that object and sends all this information to the natural language generation module. This NLG module combines a content selection component with a surface realizer [8] to produce, from the set of facts input by the Exercise Creation Module, one or more target sentence(s) together with detailed information about their derivational history. The information output by the NLG module constitutes the basis for defining the question and answer pairs that are returned to the Game Handler and eventually displayed to the learner as the text of the exercise and/or as error correction feedback of the system.

4. I-FLEG AND LEARNING

As already mentioned, I-FLEG supports a situated, interactive learning process in which the flow of learner actions is minimally constrained and where the system provides feedback in terms of scoring and target sentences. We now discuss each of these features in more detail.

Free learning
Constructivist and cognitivist teaching theories ([22], [3], [26]) emphasize both the fundamental role of learner autonomy in enhancing learning performance and the importance for the learner to be aware of the acquired language skills during the learning process. In line with these approaches, the I-FLEG game does not follow a prescriptive teaching strategy but rather supports a free learning flow while highlighting the specific construct being learned through the selection, by the learner, of a given teaching goal. The system lets the learner free to explore the various learning contents and to organize the learning process meeting his own individual needs. She is free to navigate between tutorial, training and evaluation mode in any order and any number of times. Similarly, in tutorial and in training mode, she can choose the lexical field (room in the house) and the specific objects she wishes to work on. Finally, she can switch between teaching goals at will, for instance, learning first passive, then adjective order, then passive again.

Situated, interactive Learning
It has been often argued that learning in 3D virtual worlds both provides high level of interactivity and immersion ([6], [12]). [27] describe the relationship between virtual environments and learning on three dimensions, representational fidelity, immediacy of control and presence. 3D environments are usually higher than 2D environments on the three dimensions, as it is the case for I-FLEG game. Indeed, I-FLEG offers representational fidelity given that the house environment the player can evolve in corresponds closely to a real house, with usual rooms and with objects that are already familiar to the learner. The immediacy of control is provided by the point-and-click mechanism which allows the learner to select the objects of her choice to support the learning. The presence is reinforced by the avatar and its freedom of movement in the house.The learning content is directly influenced by the presence of the avatar and its control over the objects.

Learning in I-FLEG is situated and interactive in that the learner learns by interacting with the virtual world and the learning content depends on the current context of the learner avatar. More generally and as explained in detail in the next section, the generation of training activities in I-FLEG is a context-aware process, and so are the texts output in the tutorial and training phases. That is, the system generates training items on the basis of both absolute (e.g. object color, function, form) and relative context properties, such as the player position and/or the state of the touched object.
Thus, in tutorial mode, when the learner enters a room, the system makes use of a relative property to generate a training item, namely the position of the player in the virtual world, and outputs the name of the place she is in, as shown in (1); while, when the learner touches objects, the system displays their name, as shown in (2). Similarly, in training mode, the exercise produced by the system will systematically mention the name of the object being clicked on as illustrated in (3) or the state of the object if it was affected by the interaction with the user as in (4).

(1) System: *Vous êtes dans une salle à manger*.
You are in the dining room.

(2) System: *C'est une machine à café.*
This is a coffee machine.

(3) System: "Make a sentence with the following words":
machine à café / plein / être
coffee machine / full / be

(4) System: *Le frigo est ouvert/fermé.*
The fridge is open/closed.

That is, depending on where the user is and which objects she touches, the teaching content as well as the learning flow will be different.

Learner model, System feedback and Adaptive learning

I-FLEG provides the architecture for a personalized learning process in which the learning content is automatically tuned to the learner level of proficiency. This tuning process relies on a learner model which is constantly updated as the learner interacts with the system and on an expert model relating grammar, lexicon and morphosyntax to learner levels. A basic error detection strategy combined with a scoring mechanism provides the learner with feedback and helps maintain her motivation. We now detail each of these components.

The expert model: the *expert or domain model* implemented in I-FLEG encodes the linguistic knowledge required for each language level as recommended by the FLE standard for French. For instance, Table 1 shows some of the linguistic knowledge implemented in the current version of I-FLEG for the language proficiency levels A1 and A2. This information is used to decide on the grammar exploited by the NLG module to generate exercises thus defining the grammatical coverage of the system. In the A1 level, the grammar used is restricted to cover the most basic syntactic constructions. This coverage then increases with the level of the learner.

Figure 5. Example of tutorial interaction.

A1	nouns, definite determiners, adjectives, indefinite determiners, Adj Small Clause (copula, raising), NP Small Clause (copula, raising), ...
A2	A1, transitive verbs, demonstrative pronouns, ...

Table. 1 Example of FLE level specifications and required language knowledge.

The learner model: *The learner model* is structured into two components:

1. the *user profile*, a component containing static information about the user i.e., information that does not change during a game session such as age, gender and level of L2 proficiency,
2. the *interaction profile*, a component storing information about the user/system interactions.

The static information stored in the user profile is shown in Table 2, we store the age and the gender of the learner, his native language or languages (in the case of a polyglot). We also store information about other foreign languages known by the user and his exposure to the target second language, i.e. whether the learner has studied French at school or at University or lived in a french speaking country, and her language proficiency level. This information is obtained by querying the user at her first game interaction.

Currently, this static information is used to initialize the system to A2 level. Once the system coverage is more developed however, it will be used to set the parameters determining the lexical, morpho-syntactic and syntactic coverage of the system or equivalently, the learner level of L2 proficiency. For instance, depending on whether the user is a child or an adult, the system might choose different ontologies tailored to fit the interests of each age group. Similarly, the system might choose to use a more or less large coverage grammar depending on the level of the learner: a more advanced learner will trigger the selection of a larger coverage grammar and as a result, more complex exercises will be produced. For instance, after the user has touched a table in tutorial mode, the output might be (5a) if the user is a beginner but (5b) or (5c) if she is more advanced and has already acquired adjectives and personal pronouns respectively.

(5)

a. *C'est une table*.
This is a table.

b. *C'est une petite table.*
This is a small table.

c. *C'est une petite table. Elle est blanche*.
This is a small table. It is white.

User ID	167	53
Age	23	25
Gender	F	M
L1 Background	English	Spanish
L1 Second Background	_	_
L2 Background	_	English
L_Target Exposure	live in France for 2 years	Learned French at school
L_Target Level	Basic	Intermediate

Table 2. User Profile. L1 Background and L1 Second Background classify the user depending on her native language and/or native languages (if the user is a polyglot). L2 Background stores information about the learner knowledge of foreign/second languages. L_Target_Exposure and L_Target Level represent the time the user was exposed to the target language and the proficiency level she has achieved, respectively.

Similarly, the interaction profile is built by the system monitoring the user during the whole game session and storing in a database detailed information about user/system interactions. Thus, for each exercise instance, information such as query/answer pairs, the time needed to solve an exercise, the game score, etc. are stored in the database. In future work, we plan to exploit this information to determine, at the beginning of each game, the complexity and the type of interactions allowed for, depending on the results obtained by the learner in previous games. That is, the training activities proposed to the user will be tailored to her level of L2 proficiency.

Error detection, Scoring and Feedback*:* I-FLEG integrates a basic error detection mechanism which compares the learner input against one or more expected answers (targets) allowing for minor deviations such as differences in capitalization, determiner definiteness and punctuation. Although this is obviously limited, the training activities currently proposed are sufficiently constrained for the error detection mechanism to be adequate. Using this mechanism, the system provides the learner with two types of feedback namely, a verbal and a numerical feedback. The verbal feedback indicates whether the answer given by the learner is correct or not. In case it is not, the target answer is displayed thereby giving the learner the correct solution.

Figure 6. Example of user/system interaction during the training phase.

The numerical feedback is used to quickly give the player an idea of her performance and to enhance the game feeling as well as to evaluate the learning progress. It is based on a structured datatype mapping learning performance to the credit points the learner scores while playing. Recall that the game can be played both in training and in evaluation mode. While in training mode, the score is used only to motivate the learner and give her an idea of her level; in evaluation mode, the score is exploited to update the learner profile and when appropriate, to upgrade her to a new level. In both cases, the score is systematically displayed to the user both during playing and at the end of the (training or evaluation) session. As shown in Figure 7, the system shows the learner the achieved credit points as an indication of her performance for each of the teaching goals that have been evaluated or trained.

Depending on the learner level of L2 proficiency and on the type of training activity, the system could be amended to adopt different strategies for analyzing the learner input and for providing her with a feedback. For instance, if the user is a beginner, the system could allow for some imprecisions in the input while if the learner is intermediate or expert, scoring can be made more stringent in that all deviations from the target are penalized not just those bearing on the current teaching goal.
Table 3 illustrates this by showing the different feedbacks which could be generated by the system in the case of a beginner and of an intermediate user for a training activity focusing on adjective order. If the learner is a beginner, the system accepts an imprecise input containing the right word order but with morphological imprecisions and outputs a score upgrade of half a point together with a feedback message containing the correct answer, i.e. *"Jean est un petit garçon" (*Jean is a little boy*).*
The same input by an intermediate learner will score no point however as will an input by a beginner which fails on the current teaching goal such as here, "*Jean est un garcon petit*" where the "adjective order" exercise is incorrectly handled by the learner.

Figure 7. The final score

Training Activity Type: Adjective Order

Please put the following words in the write order.	
System Query: *petit est Jean un garçon*	System Feedback
a. User Level = Beginner User Input: *Jean est un petite garçon.* Jean is a little boy	System Output: Oui, presque*!* *"Jean est un petit garçon"*. Score = Score + 0.5
b. User Level = Beginner User Input: *Jean est un garçon petit.* Jean is a little boy	System Output: *Eh non!* *"Jean est un petit garçon"*. Score = Score + 0
c. User Level = Intermed User Input: *Jean est un petite garçon.* Jean is a little boy	System Output: *Eh non!* *"Jean est un petit garçon"* Score = Score + 0

Table 3. Example of System Feedback in case of a beginner (a and b) and intermediate (c) learner level.

5. AUTOMATIC GENERATION OF TRAINING ACTIVITIES

A distinguishing feature of I-FLEG is that it exploits Natural Language Generation techniques to produce test items. As mentioned in Sections 3 and 4, exercises are triggered by the learner interactions with the world. Whenever the learner touches an object, an exercise query is displayed and the learner answer is checked against the expected answer(s) referred to as "the target(s)". In I-FLEG, these exercise queries and targets are dynamically generated by the GenI surface realizer [8] from a Description Logic (DL) knowledge base describing the objects of the 3D world. In what follows, we first describe the knowledge base content and creation. We then sketch the process by which queries and targets are produced. Finally, we discuss three features of the NLG approach to the automated generation of language learning exercises: (i) it permits adapting the exercises to the context (object, teaching goal, learner level); (ii) it allows for versatile and varied output (given the same context, several, distinct training items can be produced); (iii) it is generic and therefore easy to port either to new grammatical constructs or to a new lexicon (specialized terminology, new lexical field, etc.).

Knowledge Base Content and Creation

As mentioned above, exercises and targets i.e., solutions to the exercises, are generated from a knowledge base. This knowledge base contains facts about the 3D world objects and more specifically, attributes and relations. For instance, the knowledge base associated with a table named *t* in the virtual world might contain the following information where *t* and *c* denote individuals, *Table*, *White*, *Large* and *TableCloth* are attributes and *With* is a binary relation.

t:Table, t:White, t:Large, (t,c):With c:TableCloth

The knowledge base associated with each object is encoded in OWL Web Ontology Language and can be edited in two ways : either through a graphical interface developed for the I-FLEG game or through the well-known Protégé editor for Knowledge Bases. Both tools (the 3D GUI and the Protégé graphical interface) support the authoring, by non experts, of the content necessary to generate exercises. However while Protégé requires some understanding of the underlying OWL language, the 3D GUI, because it is tailored to the needs of our application (describing objects), is much simpler to use. The GUI supports the quick and efficient authoring of a relatively small amount of content supporting the generation of multiple, distinct exercises. In addition, because the patterns of facts necessary to support the generation of exercices can be caracterised in terms of syntax and of concepts (for instance, to support the generation of an adjective exercise, the knowledge base must contain at least one concept of type *Adjective*), it should in fact be possible to automate the generation of the knowledge base associated with each object provided the appropriate attribute and relations can somehow be selected from some existing resource such as for instance, WordNet or CYC.

Target and Queries

I-FLEG currently integrates four types of exercise and four "teaching goals" i.e., four constructs to be acquired by the learner[5]. These four teaching goals are: Vocabulary (learning the name of an object and its gender), Passive (learning the passive transformation), Adjective agreement (correctly inflecting an adjective in context) and Adjective Order (learning whether a given adjectives comes before or after the noun it modifies). The four activity types are Fill in the blank (FIB) with lemmas, Name an object, Shuffle and Transpose.

5 As explained in the following paragraphs, the algorithm and resources used are generic thus facilitating coverage expansion particularly, lexical and grammatical coverage.

A “Fill in the blank (FIB) with lemmas” item is an item as shown below where the learner must fill a blank given a lemma. FIB exercises permit in particular exercising morphological inflections and agreement, a difficult topic for learners of French. The query contains a blank and a lemma, the target is the query where the blank has been replaced with the inflected form of the lemma to be substituted in.

Q: *C'est une table ... (blanc).* T: C'est une table blanche

This is a table (white) This is a white table

The "Name an object" query is always “Qu’est ce que c’est?” (What is this?) and the target is an NP of the form “Determiner Noun” where the noun should appropriately name the object clicked on by the learner and the determiner must be in the appropriate gender (e.g., “la table”, not “le table”). Although simple, this exercise supports the acquisition of both vocabulary and gender information.

Q: *Qu’est ce que c’est?* T: Une table

What is it? A table

A “Shuffle” query is a set of lemmas which the learner must use to make a sentence e.g.,

Q: *regarder/Jean/blanc/nappe/table/avec*
T: Jean regarde la table avec la nappe blanche[6]
Jean looks at the table with the white tablecloth

Finally, a “Transpose” item consists of a sentence which the learner must transpose to a different grammatical form (e.g., transforming an active voice sentence to a passive voice one):

Q: Transpose to passive the following active voice sentence:
Jean nettoie la table
Jean cleans the table
T: La table est nettoyée par Jean
The table is cleaned by Jean

How are training items generated?
The parameters input to the generation of a test item are the following:

- Language Content: the syntactic form the output should have, e.g. a sentence, a noun phrase, etc.
- Semantic Content: the object the test item refers to i.e., the object clicked on by the learner and the set of facts encoded in the ontology about that object
- Teaching Goal: the linguistic construct which the test item should exercise e.g. adjective morphology or passive.
- Activity Type : the type of test item that should be generated, e.g. name an object, fill in the blank, shuffle or transpose.

Starting from this input and from a computational grammar and lexicon for French, the generator produces test items in three main steps. First, a set of facts about the object is selected (*content selection*). Second, the selected facts are turned into one or more targets (*surface realization*). Third, the query is derived from the target(s) (*exercise generation*).
Content Selection (CS) selects from the knowledge base associated with the target object, a set of facts that satisfy the constraints imposed by the input parameters. For instance, if the teaching goal is Passive and the activity type is Transpose, the selected content must be verbalizable by a sentence containing a transitive verb. Thus, a binary relation must be selected together with content describing the two arguments of that relation. CS ensures that this is the case by proceeding breadth first through the search space, checking the imposed constraints against the ontology and halting when all constraints are satisfied.

Surface Realization (SR) transforms the facts selected by content selection into a phrase verbalizing those facts. For instance, given the facts (*t:Table, t:White),* SR will produce either the sentence “La table est blanche” or the NP “La table blanche”.[7] Like Content Selection, SR is non deterministic and can produce several solutions for the same input. For instance, given the facts (*j:John(j), e:Clean, (e,j):agent, (e,t):patient, t:Table, t:White),* SR will non deterministically produce *“Jean nettoie la table blanche, C’est la table blanche que nettoie Jean, La table blanche est nettoyée par Jean, C’est par Jean que la table blanche est nettoyée, etc.[8]”* Importantly however, SR can also be constrained to produce only those phrases satisfying some constraints. For instance, imposing the constraint that the output phrase must be in the active voice will eliminate from the output all sentences in the passive voice.

---

6 Note that the query does not specify whether it is the table or the tablecloth that is white. Because the exercises are generated from semantic representations, the output target will however be constrained to reflect only one of these possibilities. To constrain the learner choice so that it maps the target, we use either the world (by entering facts that matches the actual state of the world) or bracketing (by bracketing together the noun and the adjective that go together).

7 For details of how this works, we refer the reader to (Gardent and Kow 2007).

8 Jean cleans the white table, It is the white table that Jean cleans, The white table is cleaned by Jean, It is by Jean that the white table is cleaned, etc.

Finally, queries are derived from targets based on the information output for each generated phrase by the surface realizer namely, meaning (i.e., the selected facts), syntax (the phrase structure tree of the output phrase) and meta-information describing the grammatical constructs illustrated by the output phrase (e.g., Canonical Object vs. Cleft Object). In specific, FIB queries are produced by searching the yield of the syntactic tree for the lemma to substitute in and removing it from the output query. Similarly, shuffles are produced by extracting from this yield, the lemma of the content words it contains and ordering them randomly. Finally, Transpose items are created by inputing the surface realizer with the same set of facts for the query and for the target but with different constraints. For instance, to produce a "Transpose to Passive" item, we generate the query (e.g., "*Jean nettoie la table*" / "Jean cleans the table") using the Active voice constraint and the target (e.g., "*La table est nettoyée par Jean*" / "The table is cleaned by Jean") using the Passive voice constraint.

Versatility

I-FLEG training item generation module is versatile in that given the same input parameters (object, teaching goal, activity type), it can produce several outputs. As explained in the previous section, the reason for this is that both the content selection and the surface realization module are non deterministic. Thus for instance, given the following input parameters, I-FLEG will indifferently produce any of the queries listed below. Note that because we use a grammar rather than templates, all grammatical and lexical constraints are fulfilled so that for instance depending on the adjective the blanks will occur in the correct position ("blanc" is always post-nominal, "grand" pre-nominal and both can be used in predicative and in attributive position).

Input Parameters

> Facts: (*t:Table, t:White, t:Large, (t,c):With c:TableCloth)*
> Teaching Goal: Adjective order
> Activity Type: FIB
> Object: *t* (the table)

I-FLEG Output Queries:

C'est une table … (blanc), C'est une … table (grand), C'est une … table … (grand,blanc), La table est … (blanc), La table est … (grand), La … table est … (grand,blanc), La table … est … (grand,blanc), Une … table (grand), Une table …. (blanc), Une … table … (grand,blanc), Une table … avec une nappe (blanc), Une … table avec une nappe (grand), Une … table … avec une nappe (grand,blanc), La table qui est … (blanc), La table qui est … (grand), La … table qui est … (grand, blanc), etc.

Adaptivity

I-FLEG exercises are adaptive in that among the many possibilities made available by content selection and surface realization, context (in terms of input parameters) can restrict the output. Trivially, each query is related to the object the learner clicks on. More interestingly, the queries output by I-FLEG are constrained by mapping the input parameters such as the teaching goal, the activity type and the learner proficiency level, to constraints regulating content selection and/or surface realization. Thus for instance, if the learner has not yet learned prepositional phrases, a constraint can be used to block the generation of phrases containing a preposition thereby preventing the generation, in the above example, of the queries containing "avec" (e.g., "Une table … avec une nappe" (blanc)).

Genericity

The algorithm used by I-FLEG for generating language learning exercises is generic and can easily be ported to a new lexical field or a new grammatical construct. To extend the lexical coverage, it suffices to appropriately modify the lexicon and the ontology (cf. Section 3). The I-FLEG algorithm for generating exercises will then produce arbitrarily many training items covering the extended lexicon. To extend grammatical coverage and provide additional Teaching Goals (e.g., relative pronouns, clitics, tense morphology, etc.), the computational grammar [7] exploited by the generator can be drawn upon. This grammar of French covers some 35 basic verbal subcategorization frames and for each of these frames, the set of argument redistributions (active, passive, middle, neuter, reflexive, impersonal, passive impersonal) and of argument realizations (cliticization, extraction, omission, permutations, etc.) possible for this frame. Predicative (adjectival, nominal and prepositional) and light verb constructions are also covered as well as subcategorizing nouns and adjectives. Basic descriptions are provided for the remaining constructions i.e., adverbs, determiners and prepositions. We are currently investigating how these phenomena relate to the teaching goals of French teaching courses and to what extent, these can be covered by the I-FLEG approach.

6. CONCLUDING REMARKS

We have introduced I-FLEG, an interactive language game specifically designed to explore the contribution 3D worlds and NLG technology could make to CALL. This game brings together 3D graphics, virtual reality and NLG technologies thereby supporting immersive, situated language learning, learner adaptivity and the context-driven automatic generation of learning activities.

We conducted a preliminary evaluation of the game prototype during a demonstration session at our institution open house day. The game was set to train English and was presented to a general audience mostly composed of French speakers. The test participants played short game sessions including a tutorial, a vocabulary and a syntax test unit. Table 4 shows the game results achieved by five test subjects. The total score represents the number of exercise items successfully solved by each player and might be used together with the total playing time as a metric for ascertaining the language proficiency of the player. I-FLEG further outputs partial scores (e.g. vocabulary and syntax score) that provide metrics for comparing the performance of the player on different teaching goals. In the next version of the game, we plan to generate more accurate scoring by abandoning absolute scores in favor of relative scores, e.i. the score will measure the number of right solved exercises out of the total number of exercises done for each teaching goal.

Player	Vocabulary Score	Syntax Score	Total Score	Start Time	End Time
AvatarT 2	9	4	13	11:59:22 AM	12:12:11 PM
AvatarT 4	3	2	5	11:40:35 AM	11:53:45 AM
AvatarT 5	0	0	0	10:47:37 AM	10:49:34 AM
AvatarT 3	6	1	7	10:49:48 AM	11:12:45 AM
AvatarT 1	14	8	22	11:12:01 AM	11:36:01 AM

Table 4. Game scores examples from the demo session.

Despite the limited number of test participants (about 10 people) we believe the feedback we received is very encouraging. We asked them to express their subjective opinion about the game. They found the learning platform fun and natural to interact with. Further, most of them judged playing the game a very engaging experience.

We are currently working on designing an experiment for assessing the impact of the technology embodied by the I-FLEG system on computer aided language learning. In essence, the aim is to compare the effectiveness of learning with I-FLEG with that of learning in a traditional classroom setting. In particular, we envisage a comparison of the efficiency of different teaching methodologies, such as for instance free vs. prescriptive learning flow using teaching strategies induced using reinforcement learning. Currently, the learner is let free to switch between training and evaluation mode at will. Using reinforcement learning, it is possible to learn a teaching strategy from the logs of the web interaction between learner and I-FLEG. We aim to evaluate the impact of both teaching strategies (free vs imposed by the system) on learning.

As mentioned in Section 4, we are furthermore exploring various feedback strategies and their impact on learning with the aim to define more efficient ways to present feedback to learners.

Finally, we are working on extending the linguistic and pedagogical content of the system with the aim of extending not only its lexical and grammatical coverage but also of making available, additional types of training activities which would make better use of the situated setting provided by the virtual world. In particular, we are exploring how constrained dialogs could be integrated in the game thereby permitting the exercising of lexical and grammatical constructs within a conversational setting.

7. ACKNOWLEDGMENTS

The research presented in this paper was partially supported by the European Fund for Regional Development within the framework of the INTERREG IVA Allegro project.

8. REFERENCES

[1] Amaral, L. and D. Meurers. 2007. Conceptualizing Student Models for Intelligent Computer-Assisted Language Learning. Cristina Conati and Kathleen F. McCoy (Eds.): **User Modeling 2007: Proceedings of the Eleventh International Conference**, Lecture Notes in Computer Science. Wien, New York, Berlin: Springer.

[2] Bick, E. 2005. Grammar for Fun: IT-based Grammar Learning with VISL. In P. J. Henriksen (ed), **CALL for the Nordic Languages**, Copenhagen Studies in Language, pp. 49- 64.

[3] Buell, J. 1997. Constructing education: Computers and the transformation of learning, ***CALL Journal***, Vol. 7, 1997.

[4] Coniam, D. (1997). A preliminary inquiry into using corpus word frequency data in the automatic generation of English language cloze tests. **Computer Assisted Language Instruction Consortium** 16(2–4), pp. 15–33.

[5] Csikszentmihalyi, M. 1990. **Flow: The Psychology of Optimal Experience**, Harper & Row, New York, 1990.

[6] Dalgarno, B., Hedberg, J. and Harper, B. (2002). The Contribution of 3D Environments to Conceptual Understanding. In A.Williamson, C. Gunn, A. Young and T. Clear (Eds) **Winds of change in the sea of learning: Charting the course of digital**

education, proceedings of the 19th annual conference of the Australasian Society for Computers in Learning in Tertiary Education (pp. 149-158). Auckland, NZ: UNITEC Institute of Technology.
[7] Gardent, C. *2008*. Integrating a unification-based semantics in a large scale Lexicalised Tree Adjoininig Grammar for French. **Proceedings of the 22nd International Conference on Computational Linguistics (COLING'08)**, pp. 249-256, Manchester, UK, 2008.
[8] Gardent, C. and E. Kow (2007). A symbolic approach to Near-Deterministic Surface Realisation using Tree Adjoining Grammar. In **Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL 2007)**, pp. 328-335, Prague, Czech Republic.
[9] Hagen, L. K. (1999). Spanish for Business Professionals. Project Web Page. http:/ /www. uhd.edu/academic/research/sbp/.
[10] Heift, T. (2001). Intelligent Language Tutoring Systems for Grammar Practice. **Zeitschrift für Interkulturellen Fremdsprachenunterricht** (online), 6(2), pp.1-15.
[11] Johnson, W. L. and A. Valente. 2009. Tactical Language and Culture Training Systems: Using AI to Teach Foreign Languages and Cultures. **AI Magazine**, 30(2), pp. 72– 83.
[12] Kapp, K. M., & O'Driscoll, T. (2010). **Learning in 3D, Adding a New Dimension to Enterprise Learning and Collaboration**. San Francisco, CA: Pfeiffer.
[13] Kim, J., R. Hill, P. Durlach, H. C. Lane, E. Forbell, M. Core, S. Marsella, D. Pynadath and J. Hart. 2009. BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. **Journal of Artificial Intelligence in Education**, 19(2), pp.289–308.
[14] Krashen, S. 1982. **Principles and practice in second language acquisition**, Pergamon.
[15] Krashen, S. 1997. **The monitor model for adult second language performance**. In Viewpoints on English as a Second Language, New York: Regents. Burt, Dulay, Finocchiaro (eds), pp. 152-161.
[16] Leacock, C., M. Chodorow, M. Gamon and J. Tetreault (2010). **Automated Grammatical Error Detection for Language Learners**. Morgan & Claypool (eds).
[17] Liu, Chao-Lin, Chun-Hung Wang and Zhao-Ming Gao (2005). Using lexical constraints for enhancing computer-generated multiple-choice cloze items. **International Journal of Computational Linguistics and Chinese Language Processing**, Vol. 10.
[18] Metcalf, V. and D. Meurers, D. 2006. Generating Web-based English Preposition Exercises from Real-World Texts. In **Proceedings of EUROCALL 2006**, Granada Spain.
[19] Nagata, N. (2009). Robo-Sensei's NLP-Based Error Detection and Feed- back Generation. ***CALICO Journal*** 26(3), pp. 562–579.
[20] Raybourn, E., E. Deagle, K. Mendini and J. Heneghan. 2005. Adaptive Thinking & Leadership simulation game training for Special Forces Officers. In **Proceedings of the Interservice/ Industry Training, Simulation and Education Conference (I/ITSEC 2005)**, Orlando, Florida, USA.
[21] Rutherford, W. 1987. **Second Language Grammar: Learning and Teaching**. Longman.
[22] Schmidt, R. (1995). **Consciousness and foreign language: A tutorial on the role of attention and awareness in learning**. In R. Schmidt (ed.), Attention and awareness in foreign language learning, Honolulu: University of Hawaii Press, pp. 1–63.
[23] Segond, F. and T. Parmentier. 2004. NLP serving the cause of language learning. **Proceedings of the Workshop on eLearning for Computational Linguistics and Computational Linguistics for eLearning**, Morristown, NJ, USA, 2004, pp. 11– 17, Association for Computational Linguistics.
[24] Segond, F., T. Parmentier, R. Stock, R. Rosner, and M. U. Muela. 2005. Situational language training for hotel receptionists. **Proceedings of the 2nd ACL-Workshop on Building Educational Applications Using NLP**, Ann Arbor, Michigan, 2005, pp. 85–92.
[25] Sykes, J., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. **CALICO Journal** 25, 528-546.
[26] Uhl Chamot, A. and J.M. O'Malley. 2009. **The Calla Handbook: Implementing the Cognitive Academic Language Learning Approach**, Addison Wesley, 2009.
[27] Whitelock, D. Brna, P. & Holland, S. (1996) What is the value of Virtual Reality for Conceptual learning? Towards a theoretical Framework. in P. Brna, A. Paiva and J.A. Self (eds.) **Proceedings of the European Conference on Artificial Intelligence in Education**, Lisbon pp. 136-141.