

# Immersive 3D Environments and Multilinguality: Some Non-Intrusive and Dynamic e-learning-oriented Scenarios based on Textual Information

Samuel Cruz-Lara, Nadia Bellalem, Lotfi Bellalem, Tarik Osswald

## ► To cite this version:

Samuel Cruz-Lara, Nadia Bellalem, Lotfi Bellalem, Tarik Osswald. Immersive 3D Environments and Multilinguality: Some Non-Intrusive and Dynamic e-learning-oriented Scenarios based on Textual Information. Journal of virtual worlds research, Journal of virtual worlds research, 2009, 2 (3). inria-00426562

**HAL Id: inria-00426562**

**<https://hal.inria.fr/inria-00426562>**

Submitted on 27 Oct 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Immersive 3D Environments and Multilinguality:  
Some Non-Intrusive and Dynamic e-learning-oriented  
Scenarios based on Textual Information

Samuel Cruz-Lara<sup>\*</sup>, Nadia Bellalem, Lotfi Bellalem and Tarik Osswald  
LORIA / INRIA Nancy Grand Est – Nancy-Université

---

## Abstract

Virtual Worlds may become primary tools for learning many aspects of history, for acquiring new skills, for job assessment, and for many of our most cost-effective and productive forms of collaboration (Metaverse Roadmap Report, 2007).

We will present some non-intrusive and dynamic e-learning based scenarios related to multilingual textual information within an immersive 3D environment. We refer to these scenarios as non-intrusive because they do not interrupt the user's activities within the immersive 3D environment. Rather, they enrich his/her individual experience. Obviously, these scenarios need to be dynamic, because user interaction occurs mostly in real-time. In addition, these non-intrusive and dynamic e-learning-oriented scenarios exemplify how a standardized framework for textual multilingual support associated with an immersive 3D environment, may considerably change the way people usually deal with multilingual information and with language learning on the Internet. We would like to illustrate that, in the context of immersive 3D environments, dealing with multilinguality is more than just localization and real-time automatic translations. Finally, the analysis of these non-intrusive and dynamic e-learning based scenarios lead us to propose a general architecture allowing immersive 3D environments to deal with multilinguality in the most general and dynamic way possible.

**Keywords:** Multilinguality, Virtual Worlds, Standardization, Non-Intrusive and Dynamic Scenarios, Language Learning

---

Linguistic information plays an essential role in the management of information, as it bears most of the descriptive content associated with more visual information. Depending on the context, it may be seen as the primary content (text illustrated by pictures or videos), as documentary content for multimedia information, or as one among several possible information components

---

in specific contexts such as interactive multimedia applications. Linguistic information can also appear in various formats: spoken data in an audio or video sequence, implicit data appearing on an image (caption, tags, etc.) or textual information that may be further presented to the user graphically or via a text to speech processor.

In this context dealing with multilinguality is crucial to being able to adapt the content to specific users' targets. It has to take into account situations in which the linguistic information contained in a multimedia sequence is already conceived in such a way that it can be adapted on the fly to the linguistic needs of the user. It also has to take into account situations in which the content should be adapted by additional processes before it is presented to the user.

The extremely fast evolution of the technological development in the sector of Communication and Information Technologies, and in particular, in the field of natural language processing, makes the question of standardization particularly acute. The issues related to this standardization are of industrial, economic and cultural nature.

The scope of activities in localization and translation memory (TM), as well as any type of online multilingual customization (subtitling, iTV, karaoke) is very large, and numerous independent groups are working on these aspects, such as LISA, OASIS, W3C, ISO, etc. Under the guidance of the above-mentioned groups, many formats have been developed. Some of the major formats of specific interest for localization and translation memories are TMX (LISA/OSCAR), XLIFF and OAXAL (OASIS), ITS (W3C). There are many identical requirements for all the formats, irrespective of the differences in final output.

Second Life, in common with other virtual world applications, has opened up the potential for users and learners, teachers and trainers, policy makers and decision makers to collaborate easily in immersive 3D environments. Through their presence as an avatar in the immersive space, the user can readily feel a sense of control within the environments and more easily engage with the experiences as they unfold.

In this paper, we would like to illustrate that within immersive 3D environments dealing with multilinguality involves more than just localization and translation issues.

The paper is structured as follows: firstly, we define some important concepts such as Multilinguality, Globalization, Localization, Internationalization, and Automatic Translation; secondly, we explain why normalization and standardization are key-issues in the framework of multilinguality, and we introduce the Multi Lingual Information Framework - MLIF [ISO CD 24616]; thirdly, we introduce several non-intrusive and dynamic e-learning scenarios that allow us to illustrate how multilinguality may be approached within immersive 3D environments. Finally, from the analysis of these scenarios, we propose a general architecture allowing immersive 3D environments to deal with multilinguality in the most general and dynamic way.

## **Multilinguality**

This section provides some important definitions in the frame of multilinguality, such as localization, internationalization and automatic translation. It should be noted that the scenarios described in Section 4 mainly focus on automatic translation.

## Globalization: Localization and Internationalization

As the LISA association explains, globalization can best be thought of as a cycle rather than a single process, as shown in Figure 1.

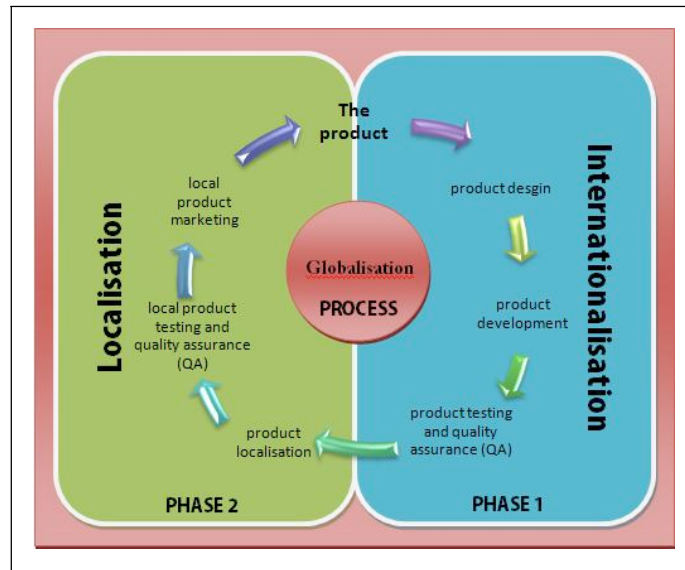


Figure 1: The Globalization process

In this view, the two primary technical processes that comprise globalization - internationalization and localization - are seen as part of a global whole:

- Internationalization encompasses the planning and preparation stages for a product in which support for global market is built in by design. This process means that all cultural assumptions are removed and any country- or language-specific content is stored externally to the product so that it can be easily adapted.
- Localization refers to the actual adaptation of the product for a specific market. It includes translation, adaptation of graphics, adoption of local currencies, use of proper forms for dates, addresses, and phone numbers, and many other details, including physical structures of products in some cases. If these details were not anticipated in the internationalization phase, they must be fixed during localization, adding time and expense to the project. In extreme cases, products that were not internationalized may not even be localizable.

### Automatic translation

Automatic or Machine translation (MT) is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitutions of words in one natural language for words in another. Using corpus techniques, more complex translations may be attempted, allowing for better handling of differences in linguistic typology, phrase recognition, and translation of idioms, as well as the isolation of anomalies.

Current MT software often allows for customization by domain or profession (such as weather reports) - improving output by limiting the scope of allowable substitutions. This technique is particularly effective in domains where formal or formulaic language is used. It follows then that machine translation of government and legal documents more readily produces usable output than conversation or less standardized text<sup>2</sup>.

# Standardization

## Introduction

Nowadays, an increasing number of standards are frequently being used within most scientific and technical domains. Translation and localization activities simply cannot remain isolated from this important and novel situation. The advantages of normalization are currently fully recognized by most professional translators. Using standards means working with a high level of quality, performance, and reliability within a very important market that is becoming more and more global and thus more and more challenging. Indeed, the standards combine simplicity and economy by reducing the planning and production costs, and by unifying several kinds of terminology (i.e. validated vocabulary) and several kinds of products. At the national and international levels, standards stimulate cooperation between different communities of trade while ensuring interoperability within information exchanges and reliability of all generated results by using standardized methods and procedures; this is why normative information has become fundamental.

The scope of research and development within the localization and translation memory process development is very large. Within this context, several industrial standards have been developed, including TMX, XLIFF and OLIF. However, when we closely examine these different multilingual textual information representation standards or formats by subject field, we find that they have many overlapping features. All the formats aim at being user-friendly, easy-to-learn, and at reusing existing data or knowledge bases. All these formats work well in the specific field they are designed for, but they lack the synergy that would make them interoperable when using one type of information in a slightly different context. Modelization corresponds to the need to describe and compare existing interchange formats in terms of their informational coverage and the conditions of interoperability between these formats. One of the issues here is to explain how a uniform way of documenting such data takes into account the heterogeneity of both the formats and the descriptors.

We also seek to answer the demand for more flexibility in the definition of interchange formats so that any new project may define its own data organization without losing interoperability with existing standards or practices. Such an attempt should lead to more general principles and methods for analyzing existing multilingual databases and mapping them onto any chosen multilingual interchange format.

## Normalization: a key issue for translation

The translator is the most important element of the translation process: thanks to their experience, their knowledge and the automatic translation services they may use, they ensure that the translated document is accurate with respect to the original one [3]. A good translation does not only need linguistic awareness but it also needs a good knowledge of the technical or scientific field of the documents that have to be translated. Most texts are addressed to specialists and ignorance of the specialized expressions can justifiably cause rejection by the reader. In the same way, English technical terms - whose equivalents nevertheless exist in the

---

<sup>2</sup> Definition from <http://www.wikipedia.org>

target language - are often used untranslated (as is the case in French, for example). Obviously, technical translations are not limited to data processing or computer science translations. Technical translators must have, in addition to their knowledge, a high-quality set of documents. Even the more knowledgeable must continuously seek advice from technical documents (i.e. journals, specialized dictionaries, magazines, data bases, etc.). These technical documents constitute a set of essential tools that allow a translator or a translation service to analyze the information located on the covered subject. So they must evolve constantly by acquiring new information and new experiences, **because this way** they will obtain additional linguistic and non-linguistic knowledge related to their **domains** [4].

Given that they are high-level models for technical specifications (i.e. symbols, definitions, codes, figures, methodology, etc), standards constitute a fundamental tool for translation, because they will provide abundant, exact, and above all, interoperable and reliable information. Unfortunately, several fields - and especially translation - have numerous and often non-compatible standards. This requires a parallel activity of normalization of these standards in order to ensure, at least, a minimal degree of compatibility. This activity constitutes the main issue of the "Documentation on standards" or the "Structured set of standards" [5]. As the texts, objects of translation, have a great diversity of subjects, the documentary activities applied to the standards can also guide and direct the translator in the search for standards relative to a given field. The production of standards is sometimes so prolific that it makes it quite difficult to understand exactly what method or procedure has to be used.

The documentary techniques bring essential assistance. The services of information and dissemination of the national and international organizations of standardization give easy access to all this information. The standards issued by the organizations of standardization (i.e. ISO, W3C and LISA) are becoming more and more accepted by the translation services and the translators to achieve a high-level quality in their services and their products. Standardization thus becomes a synonym of quality for the customers who desire the best possible results. In addition, standards also represent an essential tool for translators and translation services, as they aim at creating normalized terminological and methodological linguistic resources, in order to improve the national and international exchanges as well as cooperation within all possible fields. It is also necessary to point out the important efforts of standardization of ISO and W3C in the field of information technologies, especially those related to the computerized processing of multilingual data and the Internet.

Standardization is present during the whole process of translation. So, the access to the normative information **represents a stage impossible to circumvent** within the activity of translators and translation services. Within this task (i.e. standardization) human translators are assisted not only by resource centers in charge of disseminating the information of the standards worked out by the national and international organizations, but also by other specialized private agencies (**i.e. PRODOC<sup>3</sup>**) whose main objective is to advise the customers with regards to standards. In the same way, the Internet allows access to thousands of web sites (i.e. terminology trade, databases, etc.) that provide important information related to using standards, as well as access to several interesting research projects in progress whose objectives are the development of standards and recommendations in the field of the "industry of language".

---

<sup>3</sup> <http://www.prodoc.de>

## **MLIF: The Multi Lingual Information Framework (MLIF)**

MLIF provides a generic platform for modeling and managing multilingual information in various domains: localization, translation, multimedia, document management, digital library, and information or business modeling applications. MLIF provides a metamodel and a set of generic data categories for various application domains. MLIF also provides strategies for the interoperability and/or linking of models including (but not limited to): XLIFF, TMX, OAXAL, SMILText and ITS.

### **What is a metamodel?**

A metamodel does not describe one specific format, but acts as a high level mechanism based on the following elementary notions: structure, information and methodology. A metamodel can be defined as a generic structure shared by all other formats and which breaks out the organization of a specific standard into basic components. A metamodel should be a generic mechanism for representing content within a specific context. Actually, a metamodel summarizes the organization of data. The structuring elements of the metamodel are called "components" and they may be "decorated" with information units. A metamodel should also comprise a flexible specification platform for elementary units. This specification platform should be coupled to a reference set of descriptors that should be used to parameterize specific applications dealing with content.

### **What is a data category?**

A metamodel contains several information units related to a given format, which we refer to as "Data Categories". A selection of data categories can be derived as a subset of a Data Category Registry (DCR) [ISO 12620]. The DCR defines a set of data categories accepted by an ISO committee. The overall goal of the DCR is not to impose a specific set of data categories, but rather to ensure that the semantics of these data categories is well defined and understood. A data category is the generic term that references a concept. There is one and only one identifier for a data category in a DCR. All data categories are represented by a unique set of descriptors. For example, the data category /languageIdentifier/ indicates the name of a language which is described by 2 [ISO 639-1] or 3 [ISO 639-2] digits. A Data category Selection (DCS) is needed in order to define, in combination with a metamodel, the various constraints that apply to a given domain-specific information structure or interchange format. A DCS and a metamodel can represent the organization of an individual application, and the organization of a specific domain.

## **Specifying the Multi Lingual Information Framework**

Linguistic structures exist in a wide variety of formats ranging from highly organized data (e.g. translation memory) to loosely structured information. The representation of multilingual data is based on the expression of multiple views representing various levels of linguistic information, usually pointing to primary data (e.g. part of speech tagging) and sometimes to one another (e.g. references or annotations). The following model identifies a class of document structures that could be used to cover a wide range of multilingual formats, and provides a framework that can be applied using XML.

All multilingual standards have a rather similar hierarchical structure but they have, for example, different terms and methods of storing metadata relevant to them. MLIF is being designed in order to provide a generic structure that can establish a basic foundation for all these standards. From this high-level representation we are able to generate, for example, any specific

XML-based format. We can thus ensure the interoperability between several standards and their committed applications.

Figure 2 describes the MLIF metamodel.

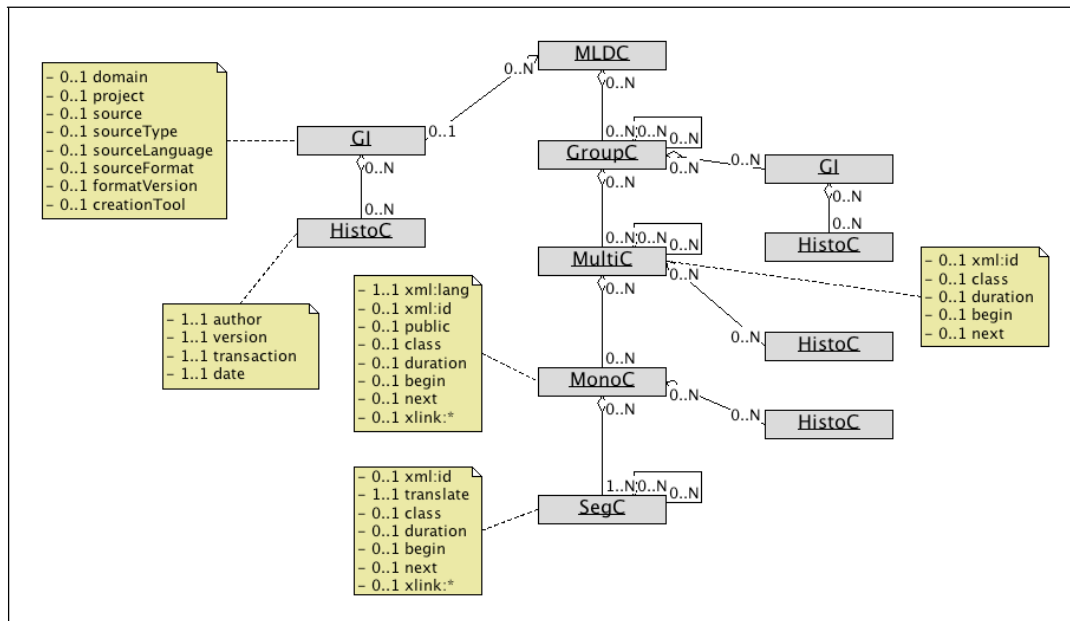


Figure 2: MLIF metamodel with related Data Categories

It is made of the following components:

- **MLDC (MultiLingual Data Collection)** Represents a collection of data containing global information and several multilingual units.
- **GI (Global Information)** Represents technical and administrative information applying to the entire multilingual data collection.
- **GroupC (Grouping Component)** Represents a sub-collection of multilingual data having a common origin or purpose within a given project.
- **MultiC (Multilingual Component)** Groups together all variants of a given textual content.
- **MonoC (Monolingual Component)** Part of a multilingual component (**MultiC**) containing information related to one language
- **HistoC (History Component)** This generic component allows tracing of modifications in the component it is anchored to (i.e. versioning).
- **SegC (Segmentation Component)** A recursive component allowing any level of segmentation for textual information.

Any format compliant with this standard may use the MLIF metamodel in two possible ways:

- By fully implementing the MLIF metamodel starting at the level of the **<MLDC>** component;
- By specifically embedding MLIF compliant information within another model, by implementing one of the lower level MLIF component, namely **<GroupC>**, **<MultiC>** or **<MonoC>**.



## **Relation with other standards**

As with TMF used for terminology, MLIF introduces a metamodel in combination with chosen data categories as a means of ensuring interoperability between several multilingual applications and corpora.

MLIF deals with multilingual corpora, multilingual fragments, and the translation relations between them. In each domain where MLIF can be used, we may consider a specific granularity of segmentation and description, built on MAF, SynAF and TMF, for morphological description, syntactical annotation or terminological description respectively. Supporting the construction and the interoperability of localization and "Translation Memories" TM resources, MLIF also deals with the description of a metamodel for multilingual content. MLIF does not propose a closed list of description features. Rather, it provides a list of data categories, which is much easier to update and extend. This list represents a point of reference for multilingual information in the context of various application scenarios.

However, MLIF will not only describes elementary linguistic segments (i.e. sentence, syntactical component, word, part of speech, etc.), but it may also be used to represent document structure (i.e. title, abstract, paragraph, section, etc.). In addition, MLIF allows using external and internal links (i.e. annotations and references).

The MLIF is being designed with the objective of providing a common platform for all the existing tools developed by the groups listed in the introduction section. It promotes the use of a common framework for the future development of several different formats: TMX, XLIFF, etc. MLIF can be considered as a parent for all these formats, since all of them deal with multilingual data expressed in the form of segments or text units. They can all be stored, manipulated and translated in a similar manner.

## **E-learning scenarios: the Language Academy**

We would like to present now some non-intrusive and dynamic e-learning based scenarios on multilingual textual information within an immersive 3D environment. We refer to these scenarios as non-intrusive because they do not interrupt the user activities within the immersive 3D environment. Rather, they enrich his/her individual experience. Obviously, these scenarios need to be dynamic, because user interaction occurs mostly in real-time situations. We would like to illustrate that, in the context of immersive 3D environments, dealing with multilinguality is more than just localization and real-time automatic translations.

### **The General Scenario**

Pierre is a French student, who speaks English fluently but recently moved to the Netherlands. Therefore, he decides to learn Dutch through various ways. The following scenarios represent Pierre's evolution phases for virtual world usage in order to learn Dutch.

#### **Pierre alone, at home**

This scenario represents the already existing solutions in the language learning domain.

At the beginning, Pierre learned the grammatical ground knowledge using book learning methods or through group or private Dutch lessons given by a teacher. After acquiring strong ground knowledge, Pierre decides to go further and to learn this new tongue by watching

Dutch movies with French or English subtitles.

His media player with multilingual support enables him to pause anytime in order to obtain the Dutch translation of the displayed subtitles. He can also have the audio part corresponding to one subtitle played back, or he can ask for some definitions and synonyms for a single word – using specific web **services**<sup>4</sup> – by searching in his web browser. If the Dutch subtitles are not available, a translation web service may generate them from the French or English ones.

It would be interesting to be able to click directly on the subtitles (in order to get synonyms, definitions or translations) without needing to use a web browser and in order to sensibly improve the interactivity. This functionality is not currently available but is conditionally conceivable.

Pierre can also get information about the Dutch culture, thanks to on-line resources like Wikipedia, for example. This possibility is particularly useful for understanding some subtle scenes.

Nevertheless, this learning method has several weaknesses:

- Pierre only gets information without having any opportunity to use his knowledge for speaking or writing. Therefore, it is a one-way learning.
- The resources (movies and subtitles) may be limited, on the one hand by their costs, and on the other hand, by Pierre's cinematographic tastes.

That is why Pierre will take the plunge and look into virtual worlds, in order to discover more interactive and more immersive learning methods.

### **Pierre alone, in a virtual world**

Pierre now discovers the virtual worlds. He hears about the very accurate modeling of the historic museum of Amsterdam and decides to teleport there in order to take a virtual tour.

At his arrival, Pierre is offered a flying carpet to take him through the museum. The flying carpet is actually a robot which communicates through the chat interface in the virtual world (see Figure 3). It can talk and understand what Pierre says. Pierre will now virtually visit the museum, riding and directing the flying carpet.

At first, Pierre asks the carpet in Dutch to have a brief guided tour. The sentence written by Pierre is analyzed and then the carpet flies through some museum key-places. At each place, the carpet gives Pierre a description of the place where he is or a description of some works.

Pierre can also ask the carpet to fly to a given place. Then, the carpet computes the shortest way to go there. It is a means of conveyance which is more efficient than walking and which enables Pierre to use his Dutch knowledge in order to express himself.

If he wants to, Pierre can ask, in Dutch, for more accurate information about the works he discovers. The carpet then extracts the date, the author or the story of the work from the available data stored on a web server.

---

<sup>4</sup> Online databases like WordNet (<http://wordnet.princeton.edu/>) or ConceptNet (<http://conceptnet.media.mit.edu/>)



Figure 3: The flying carpet in Second Life

While visiting the museum, Pierre can also click on any work and get its complete description. All the textual information which Pierre gets is clickable. So, he can get translations of any word group, or get synonyms and definitions of a word.

During Pierre's visit, the server memorized the works about which Pierre asked for information. At the end, Pierre is offered a quiz which has been established according to the concerned works, which is written in Dutch too.

Through the museum visit, Pierre can start expressing himself in Dutch and learn a lot about Dutch culture. These two elements are essential for language learning.

### **Pierre participating in a group to learn Dutch, in a virtual world**

Pierre is now on an island in the virtual world, where conferences are held in Dutch. In particular, there are many documentary movies about the Netherlands, directly presented in the virtual world by their respective directors. Many people from several countries who wish to learn Dutch meet on this island in order to improve their skills.

In fact, the directors are native Dutch speakers and they can talk in the virtual world using VoIP technology. On the server, a software application analyzes the voice information and transcribes it into text format for the avatars who are attending the conference. Doing so, the avatars may at once read the speech and listen to it, so that they can easily understand and assimilate the Dutch language.

Moreover, once the speech is transcribed into text, it may be translated by an online tool into each avatar's tongue. As a consequence, one speech may be listened to in Dutch and read in various languages. As before, it is possible to click on one word to get synonyms, definitions or translations.

The attendees at one conference may also ask questions to the speaker if they virtually put up their hands. They may do it in different ways: through the chat-based communication system or directly with VoIP. In order to practice, they often prefer expressing themselves in Dutch.

Moreover, the same transcription and translation system as before is available to the speaker (the movie director), so that he can understand questions asked in other languages than Dutch. Figure 4<sup>5</sup> shows how such a conference may be seen by Pierre.

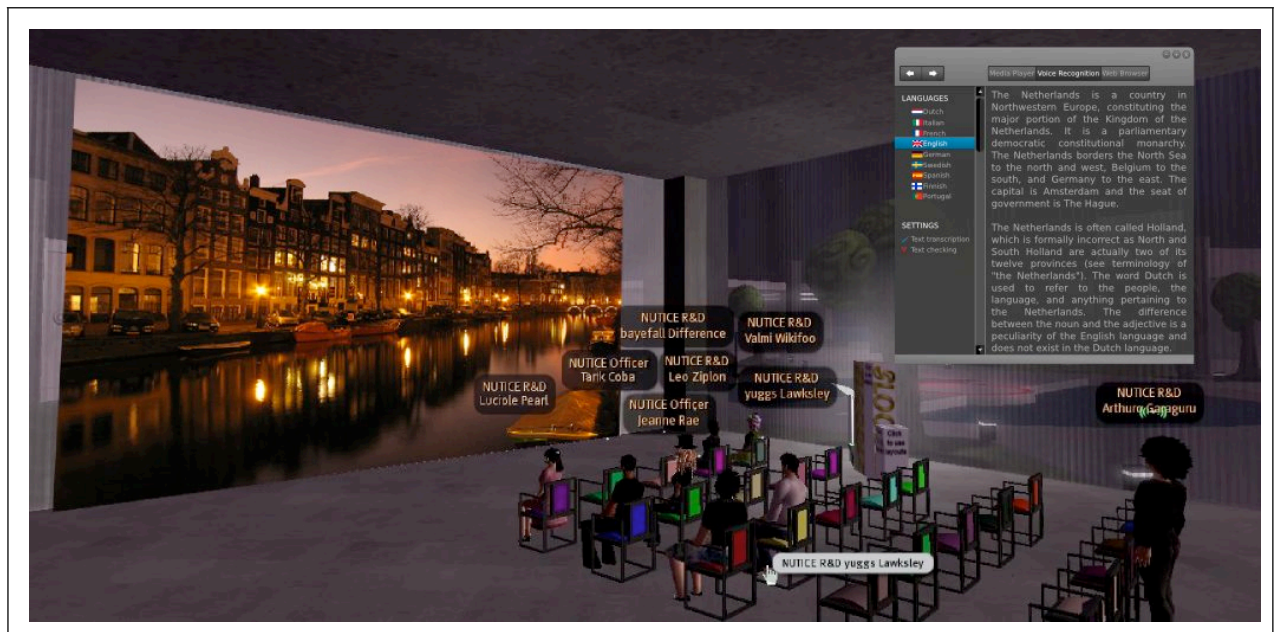


Figure 4: Virtual conference with voice transcription and translation

Finally, once again in order to facilitate language learning, the Dutch text may be analyzed by a grammar parser. Doing so, a user may for example configure their display so that all the verbs are displayed in a red font color.

### **Pierre with Dutch people, in a virtual world**

Pierre has now improved his Dutch level a lot. He knows the Dutch culture much better and has heard many native Dutch people talking. Nevertheless, his skills are not perfect yet and he now wishes to talk in real time with Dutch people. That is why he teleports to one of the most popular virtual islands in the Netherlands.

He can now talk through the chat interface with native Dutch people. The virtual world makes the situation less alarming than in the real world so that Pierre feels more comfortable during the talks.

He does not feel as nervous as he usually is in real life, particularly when talking in a foreign language. The virtual world also makes the chat-based talk more immersive than the usual chat interfaces, as it offers the users a common environmental context.

Pierre is going to meet Dutch people, and suggests that they watch a movie together in a virtual cinema. Pierre may display on his screen the subtitles in a selected language whereas his virtual friends do not need them. He can also talk about his impressions concerning the movie or about the cultural differences. In that case, the displayed subtitles are clickable.

---

<sup>5</sup> Image from Wikimedia Commons (<http://commons.wikimedia.org/>), text from Wikipedia (<http://www.wikipedia.org>)

But Pierre's Dutch knowledge is still not good enough to understand everything he hears or reads and he still needs to click on some words in order to get synonyms, translations or definitions in real time. Sometimes, he also enables the syntax coloration, for example in order to highlight the verbs displayed in the chat window of his local user interface (see Figure 5).



Figure 5: Red font colour for the verbs displayed in the chat interface

If he wants to, Pierre can also use VoIP technologies to talk with the avatars in place, using the same transcription technology as in the scenario presented in Section 4.4. Although he is a bit shy, Pierre talks easily and understands the talks better thanks to the textual transcription.

Every time Pierre clicks on one word, the system memorizes the searched information. Doing so, Pierre can delay dealing with the words and expressions which were a bit difficult to him. This is particularly useful if Pierre has not had time to consult the information right away. This functionality leaves Pierre more time to get focused on the talk in the virtual world. The sentences containing the searched words or expressions are also kept in memory, so that Pierre can retrace his searches back to the original context and have a usage example for each word.

### Required system components

In this part, we are going to explain the different technologies required for each scenario presented in Section 4. Some components are transversal to the described scenarios and implement their own mechanisms, and some other components have already been defined before. To make things easier to follow, we will first make a list of the required system components and then associate them with each scenario in Section 6.

Figure 6 shows the relationship between all the system components that we are going to explain further.

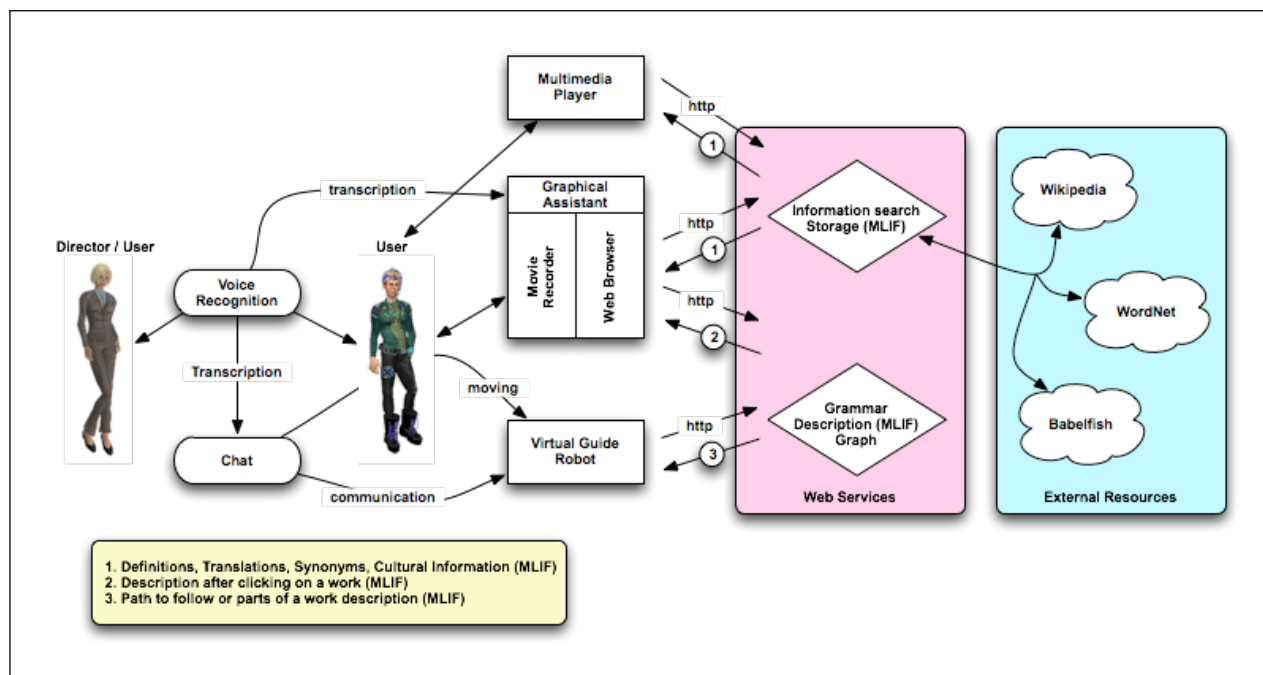


Figure 6: Relationship between the system components

### A multimedia player

The needed multimedia player should support video files with synchronized textual subtitles, in various languages. As each word or sentence in the subtitles should be clickable, the corresponding lexical units have to be representable in the text files containing the subtitles.

The player links the subtitles with the corresponding audio part (using their corresponding time signature), so that the user can ask to hear the audio part corresponding to a given subtitle. This player can display and control subtitle streams in order to be able to change the subtitle language while the video is running. That is why the subtitles should be linked together in order to be able to switch from a language to another. This functionality is enabled by MLIF.

As a consequence, it is necessary to have a data structure which represents both time information and language equivalences.

### A web service linked to external resources

This service uses external resources which already exist on the web and allows one to work easily with up-to-date information.

It relies on services like WordNet (definitions, synonyms), Reverso (word, sentence or paragraph translations) or Wikipedia (cultural information). The way those services are structured should make the needed information easy to extract. The web service should also be able to provide information which would not only be indexed to nouns, but also to cultural singularities.

### A virtual guide robot

We are now talking about a robot having the appearance of a flying carpet, which moves in a marked network according to the visitor's orders. The markers (nodes of the graph

representing the network) need to have names in various languages, including Dutch in our example. The markers have to be linked together in order to represent the graph edges.

The robot program communicates with a web service which includes a dynamic and multilingual grammar parser (depending on the marker names), in order to make accurate syntax and semantic analysis.

The implemented grammar manages several kinds of orders: moves and information inquiries.

The least-cost path computing, the management of the links between markers, the marker names in various languages, the information display and the syntax and semantic analysis are relocated on a PHP/MySQL server working as a web service in order to make the processing faster. That means that the virtual world should also be able to send HTTP requests.

As for the information inquiry (date, type, author), the text data describing the works located in the virtual world should be structured in such a way that the required information may be extracted easily. To do so, the virtual world textual data has to be formalized. The resources to fill in the descriptions might also be taken from online encyclopedias if the relevant information is standardized too.

It should also be possible to obtain information about one work by clicking on it. This information may be displayed directly in a web browser but also in the graphical virtual assistant (see Section 5.5). The information elements may be sorted according to the inquiries of the visitor and that is why they should also be standardized.

The data describing the works is stored on a remote server which is accessible through the web service.

### **A voice recognition application**

This module allows the transcription of talks into text, which is then displayed to the listeners through the chat interface window or the graphical assistant. Matching voice and text allows several kinds of learning:

- Phonetics: to be able to identify the sounds of a tongue (perception and discrimination);
- Lexical: to learn the word spellings;
- Morphosyntax: agreement of adjectives, nouns, past participles (depending on the language);
- Semantic: understanding words according to their context.

In order to get a better voice recognition rate, the conferences will be thematic and tools like Viavoice, Dragon Naturally Speaking and Freespeech will be used depending on each one's speciality.

### **A graphical virtual assistant**

This graphical assistant, dedicated to language learning, is an interface which is integrated to the virtual world. This interface is available for each user who would like to use the functionalities of the Language Academy. It encompasses the following modules:

- A movie recorder (in order to watch the virtual conferences again);
- A web browser or alike, to display the information relative to the works in a museum, the results of a search when clicking on one word and the search history (storage with MLIF format in the database);
- A voice recognition module –which may be an external stand-alone component– in order to transcribe into text the talks of the speakers (scenario 4.4) and of the Dutch people Pierre talks to (scenario 4.5).

### **An interactive chat interface**

The chat interface of the virtual world has to be flexible enough in order for the user to be able to configure syntax coloring and to click one word if he wants to. Syntax coloring can be made, as the user wishes, in the chat window and/or in the graphical assistant.

All of the obtained text data is clickable, in order to get information about the selected words or groups of words. Besides, a clicked word may be put back into its context (the sentence). Therefore, the granularity of the textual information (paragraphs, sentences, words, etc.) must be included in the standard.

Since the chat-based talks have a time signature, it is also important to be able to match time information to each text element.

## **Discussion**

### **Relationship between system components and scenarios**

Figure 7 shows the system components involved in each scenario. A checked box means that the corresponding scenario involves the corresponding system component.

		Scenarios			
		4.2	4.3	4.4	4.5
System components	Multimedia player	✓	-	-	-
	Web service	✓	✓	-	✓
	Virtual guide robot	-	✓	-	-
	Voice recognition application	-	-	✓	✓
	Graphical virtual assistant	-	✓	✓	✓
	Interactive chat interface	-	✓	✓	✓

Figure 7: Relationship between system components and scenarios

### **Theoretical and technical limits**

The functionalities of scenarios 4.4 and 4.5 mainly rely on a very good voice recognition quality. The current systems are mostly based on the use of linguistic corpora in order to improve the



recognition quality. It means that, within scenario 4.4, every conference should be associated to one available corpus. But this is not so easy to do in scenario 4.5, as the talk themes can't be defined in advance.

Moreover, the translation of the transcribed talk into another language mainly depends on the voice recognition quality. The loss of accuracy caused by text automatic translation should also be taken into account. Therefore, depending on the vocabulary and expressions used, the results of the transcription and translation functionality may be more or less accurate.

However, those limitations are not theoretical limitations. The research for this kind of application is advancing and that is why we think we can get better results in the coming years.

Finally, all the scenarios are based on a certain representation of multilingual textual information. Therefore, the emergence of new multilingual textual information standards is necessary. The MLIF standard proposal has been developed in a similar vein in order to standardize the multilingual textual information representation and fully satisfies the needs of the scenarios presented in Section 4.

## Conclusion

We have proposed some non-intrusive and dynamic e-learning based scenarios related to multilingual textual information within an immersive 3D environment. We have shown that in the context of immersive 3D environments, dealing with multilinguality is not only localization and real-time automatic translations. Immersive 3D environments may really change the perception that users may have of multilingual issues, in particular, in the field of language learning.

The analysis of these scenarios leads us to propose a general architecture allowing immersive 3D environments that are able to deal with multilinguality in the most general and dynamic possible way. It should be noted that several scenarios have already been implemented (at least partially) in Second Life, and in Open Sim. All this work is being performed in the framework of the ITEA2 Project METAVERSE1 (ITEA2 07016).

These scenarios have also led us to propose a list of minimal multilingual requirements to ISO's MPEG-V working group. The requirements that must be fulfilled are:

- **Multilinguality:** textual information must be managed, accessed, and represented by taking into account multilingual aspects. Here, some key technologies and standards are: Unicode, XML, etc.
- **Multilingual links:** it must be possible to set up links between documents using different languages. This will allow us to represent localization and translation issues. Here, some key technologies and standards are TMX, XLIFF, MLIF, OAXAL, and ITS.
- **Linguistic granularity:** it must be possible to represent textual information at several levels: paragraphs, sentences, words, syllables, part-of-speech, etc. Here, some key technologies and standards are those of ISO's TC37/SC4 "Linguistic Resources Management".
- **Time issues:** it must be possible to associate time to multilingual textual information. A classical example in this domain is illustrated by captions or sub-titles related to video frames. Here, some key technologies and standards are W3C's SMILText and Timed-Text, SRT, etc.

- External links: it must be possible to set up links between multilingual textual information and some external knowledge databases such as Wikipedia, WordNet, ConceptNet, etc. Obviously, external links may also be used to translate textual information on a real-time basis. Here, some key technologies are represented by on-line translation tools such as Yahoo's Babelfish, Google's and Translator.

## Glossary

ISO: International Organization for Standardization.

<http://www.iso.org>

ITS: Internationalization Tag Set. ITS is designed to be used with schemata to support the internationalization and localization of schemas and documents.

<http://www.w3.org/TR/its/>

LISA: Localization Industry Standards Association.

<http://www.lisa.org>

MAF: Morpho-syntactic Annotation Framework. MAF provides a reference format for the representation of morpho-syntactic annotations. ISO DIS 24611.

[http://lirics.loria.fr/doc\\_pub/maf.pdf](http://lirics.loria.fr/doc_pub/maf.pdf)

MLIF: MultiLingual Information Framework. MLIF is a high-level model for describing multilingual data, which can be used in a wide range of translation or localization applications. ISO CD 24616.

<http://mlif.loria.fr/>

OASIS: Advancing Open Standards for the Information Society.

<http://www.oasis-open.org/home/index.php>

OAXAL: Open Architecture for XML Authoring and Localization.

[http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=oaxal](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=oaxal)

SMILText: Synchronised Multimedia Integration Language Text. The purpose of including text content functionality into SMIL 3.0 is to allow authors to define small amounts of lightly-formatted text containing embedded temporal markup within the context of a SMIL presentation.

<http://www.w3.org/TR/2008/REC-SMIL3-20081201/smil-text.htm>

SynAF: Syntactic Annotation Framework. The Syntactic Annotation Framework (SynAF), is a high level model for representing the syntactic annotation of textual documents. ISO DIS 24615

[http://lirics.loria.fr/doc\\_pub/SynAF\\_LREC2006.pdf](http://lirics.loria.fr/doc_pub/SynAF_LREC2006.pdf)

TMF: Terminological Mark-up Framework. TMF provides a framework for the XML terminological database representation. ISO 16642:2003.

<http://www.loria.fr/projets/TMF/>

TMX: Translation Memory eXchange. TMX is an XML format for the exchange of translation memory data created by computer aided translation and localization tools.

<http://www.lisa.org/Translation-Memory-e.34.0.html>

W3C: World Wide Web Consortium.

<http://www.w3.org>

XLIFF: XML Localization Interchange File Format. XLIFF is intended to give any software provider a single interchange file format that can be understood by any localization provider.

<http://docs.oasis-open.org/xliff/xliff-core/xliff-core.html>

## References

Samuel Cruz-Lara, Nadia Bellalem, Julien Ducret, and Isabelle Krammer. "Standardising the Management and the Representation of Multilingual Data : the Multi Lingual Information Framework". Topics in Language Resources for Translation and Localisation. Editor Elia Yuste. John Benjamins Publishers, 2008. ISBN: 9027216886.

Sara de Freitas. "Serious Virtual Worlds: A Scoping Study". The Serious Games Insitute. November 2008.

Gómez, Carmen et María Pinto (2001): La normalisation au service du traducteur, in META, XLVI, no 3, p. 564.

Hurtado Albir, A. (ed.) (1996), La enseñanza de la traducción, Castellón : Universidad Jaume I.

Pinto, M. (1993). Análisis documental. Fundamentos y procedimientos, 2<sup>a</sup> ed. rev. y auhm., Madrid, Eudema.