Lantern
Part 1/2

Volume 7, Number 1
Lantern part 1/2
January 2014

Managing Editor

Yesha Sivan,
Metaverse-Labs Ltd.
Tel Aviv-Yaffo Academic College, Israel

Issue Editors

Yesha Sivan,
Metaverse-Labs Ltd
Tel Aviv-Yaffo Academic College, Israel

Abhishek Kathuria,
The University of Hong Kong

David Gefen,
Drexel University, Philadelphia, PA, USA

Maged Kamel Boulos,
University of Plymouth, Devon, UK

Coordinating Editor

Tzafnat Shpak

Linguistic and Multilingual Issues in Virtual Worlds and Serious Games: a General Review

Samuel Cruz-Lara, Alexandre Denis, and Nadia Bellalem
LORIA (UMR 7503) / University of Lorraine, France

Abstract

Within a globalized world, the need for linguistic support is increasing every day. Linguistic information, and in particular multilingual textual information, plays a significant role for describing digital content: information describing pictures or video sequences, general information presented to the user graphically or via a text-to-speech processor, menus in interactive multimedia or TV, subtitles, dialogue prompts, or implicit data appearing on an image such as captions, or tags. It is obviously crucial to associate digital content to linguistic information in a non-intrusive way: the user must decide, whether or not, he wants to display the linguistic information related to the digital content he is dealing with in any particular language.

In this paper we will present a general review on linguistic and multilingual issues related to virtual worlds and serious games. The expression "linguistic and multilingual issues" will consider not only any kind of linguistic support (such as syntactic and semantic analysis) based on textual information, but also any kind of multilingual and monolingual topics (such as localization or automatic translation), and their association to virtual worlds and serious games. We will focus on our ongoing research activities, particularly in the framework of sentiment analysis and emotion detection. Note that we will also dedicate special attention to standardization issues because they grant interoperability, stability, and durability.

The review will essentially be based on our own experience but some interesting international research projects and applications will be also mentioned, in particular, research projects and applications related to sentiment analysis and emotion detection.

1. Introduction

Virtual worlds and serious games are fascinating good examples of fields of development where applications offering linguistic and, in particular, multilingual support have become an absolute necessity. One may obviously include localization (i.e., adapting the natural language used by the graphical user interface), and automatic translation (i.e., using some specific software to translate text or speech from one natural language to another). But beyond that "classical multilingual support", one may also include interactive and non-intrusive multilingual tools like a linguistic enhanced chat (Cruz-Lara, S. et al., 2009), as well as natural language processing-based activities, such as e-learning - in particular language learning - (Amoia M., 2011; Amoia M. et al., 2011), conversation support in serious games (Bretaudière, T. et al., 2011), and even sentiment analysis and emotion detection based on textual information.

For example: today, talking to people via a textual chat interface has become very usual. Many web applications have an embedded chat interface, with a varying array of features, so that the users can communicate from within these applications. But all applications have their own peculiarities, and their chats also serve various requirements. A distinctive feature of virtual worlds is that people are more likely to converse interactively with other people who cannot speak their native language. In such cases, the need for some kind of "innovative" multilingual support becomes obvious. As we will explain in the following sections, a linguistic enhanced chat interface with multilingual features can meet these requirements. Moreover, such an interface can be turned into an advantage for people who want to improve their foreign language skills in virtual life situations.

As most of our ongoing research activities are devoted to sentiment analysis and emotion detection from textual information, these two fields will be described in a much more detailed way. We will also dedicate special attention to standardization issues, in particular, in the framework of sentiment analysis and emotion detection. An extensive presentation of the "MultiLingual Information Framework" standard (MLIF) [ISO 24616:2012] that we have developed and that has become in 2012 an ISO international standard may be found in (Cruz-Lara S. et al., 2010; Cruz-Lara, S., 2011). In this paper we will show how other, a priori, non-linguistics related standards such as EmotionML (i.e., a W3C proposed recommendation for the annotation and representation of emotions) that could be successfully associated to linguistic and multilingual issues in virtual worlds and serious games.

2. Classical multilingual issues

2.1 Globalization: Localization and Internationalization

As the former LISA association (Localization Industry Standards Association; http://www.gala-global.org/lisa-oscar-standards) explained, **Globalization** can best be thought of as a cycle rather than a single process, as shown in Figure 1.
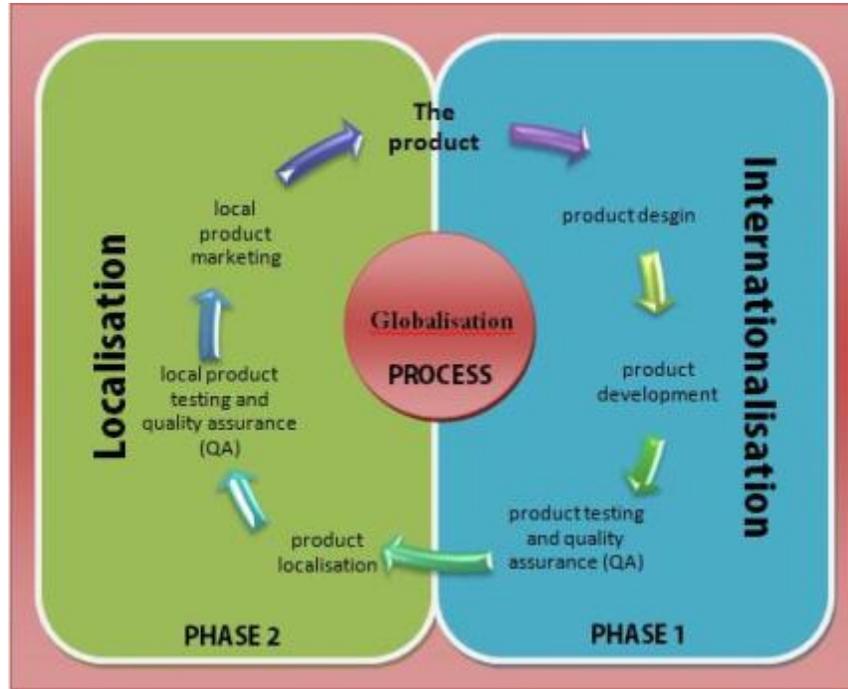
**Figure 1: The globalization process (http://commons.wikimedia.org/).**

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.

2.2   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 substitution 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. (Wikipedia. http://en.wikipedia.org/wiki/Machine_translation)

3. Innovative Multilingual and Linguistic Issues

In this section we will present several innovative multilingual and linguistic issues that may be successfully associated to virtual worlds and serious games.

3.1   Interactive and Linguistic Chat-based Communication

Here, the general idea is to improve foreign languages skills (Cruz-Lara, S. et al., 2009; Cruz-Lara S., 2011). Let's say that Pierre is a French student who has a good knowledge of English. However, his skills are not perfect yet and he wishes to talk in real time with native English-speaking people. That is why he connects to a virtual world, such as Second Life, and teleports to one of the most popular virtual islands in the United Kingdom. He can now talk through the chat interface with native English-speaking 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, 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 nice 3D environmental context.

Pierre is going to meet English 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.

But Pierre's English 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 2).

Figure 2: Different font colors for nouns and verbs displayed in the chat interface.

Every time Pierre clicks on a word, the system memorizes the searched information. Doing so, Pierre can afterwards deal with the words and expressions that were a bit difficult to him. In addition to that, he may not always have time to consult the information right away, and 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 replace his searches in the context and have a usage example for each word.

3.2   E-language Learning

The INTERREG-III ALLEGRO project (Amoia M., 2011; Amoia M. et al., 2011) aims at making language learning more challenging and entertaining. It attempts to differentiate itself from the regular serious gaming market by relaying extensively on Natural Language Processing (NLP) systems. The introduction of these technologies allows for very sophisticated language learning exercises, contrasting with previously developed products, based on keywords, template based generation, or a choice of fixed answers.

Thanks to language generation, the player can be confronted with many different ways to perform the very same exercise, increasing the players interest, lowering its ability to remember previous exercises answers, and adapting the difficulty to the players level; in short: it raises the challenge to another level.

This versatility enabled the use of dynamic environments, and this is why virtual universes were naturally chosen as a base for the game environment, allowing for a broader audience, easy maintenance, and cut on development costs in comparison with the costs of a dedicated 3D application (Bretaudière T. et al., 2011).

3.3   Conversation Support in Serious Games

Serious games engage players towards the acquisition of new skills rather than merely entertain them. Typically, they are designed for education, training or general counselling and assistance. Virtual Characters (VCs) or Embodied Conversational Agents (ECAs) can enrich users' engagement through interactive conversations (Cassell J. et al., 1999). For instance, VCs can help players to achieve some goals inside the game such as orienting players to find some place in a 3D environment, or providing relevant information at each level of the game.

Cognitive models have been proposed for representing conversational agent's beliefs, desires and intentions. The levels of thrust, credibility and cooperativeness have been subject of study in the last years, in which VCs can be collaborative or not (Roque, A., and Traum D.R., 2007). Our aim in the EUROSTAR EMO-SPEECH project (Bretaudière T. et al., 2012) is to provide virtual characters in serious games with conversation capabilities in collaborative environments.

In our first case study (see Figure 3), we integrated this dialogue framework with the serious game "Mission Plastechnologie" (http://www.mission-plastechnologie.com) developed by Artefacto (http://www.artefacto.fr). In this game, the player has to manufacture a joystick in the Plastech Factory. The employees of the factory are very collaborative and help the player to manufacture it. They explain the production and recycling processes, the different jobs and the security measures. Several characters are engaged in twelve different dialogues during the game (the game and its dialogues are in French). Particularly, the roles of the player and the system are interchanged during the game depending on the game level. Three possible characters can represent the player: Julie, Lucas, and Ben (see Figure 3, right-hand side).

Figure 3: Scenes of the game "Mission Plastechnologie".

4. Sentiment Analysis and Emotion Detection

4.1   Foreword

In this section we will introduce in a rather detailed way, two fields that represent our main ongoing research activities: sentiment analysis and emotion detection on textual information.

First, we consider that it is extremely important to define as rigorously as possible these two fields. That is the reason why a detailed presentation is given in this section. Second, once that sentiment analysis and emotion detection have been properly defined, we explain how they can be related to linguistic and multilingual issues, and then associated to virtual worlds and serious games.

4.2   Introduction

The experience of the user may be conveyed by the textual content of his utterances. The earliest aspect that is investigated in literature is the sentiment expressed by a user through an opinion. That is, how he feels, or more specifically, how he feels about a certain topic. The most general task is to find whether a text conveys a positive, negative or neutral sentiment. This task is at hand when there is a need to determine the opinion conveyed by the texts produced by a user or a community of users: detecting hostile messages and flame wars (Spertus 1997), determining whether movies are appreciated (Turney, P., 2002; Pang, B. et al. 2002) and thereby trying to predict movie success (Mishne, G. and Glance, N., 2006; Sadikov, E. et al. 2009), opinions about products (Hu, M. and Liu, B., 2004) including predicting sales (Liu, Y. et al. 2007), assessing citizens satisfaction of public policies (Kim, S. and Hovy, E., 2006) and predicting election results from Twitter (Tumasjan, A. et al. 2010). Note that predicting an election result from Twitter is still a controversial topic (Metaxas, P.T. et al. 2011).

4.3   Sentiment Analysis

The term sentiment in the expression "sentiment analysis" is a bit misleading. It refers to polarized elements in a text and not to the general meaning of sentiment as an emotion, which is closer to the topic of emotion detection. Moreover, in the literature opinion and sentiment are often used interchangeably, an opinion being a sentiment about something. Sentiment analysis is related to subjectivity analysis (that is whether a text expresses a subjective or objective point of view), however a sentiment is not necessarily expressed through a subjective assessment, and not all subjective assessments express a sentiment. Table 1 shows examples of objective and subjective sentences that can convey a sentiment or not. The sentence "This car is red" is a purely objective sentence without polarized information, while

"This car is broken" is also an objective sentence but which contains a negative sentiment about the car. The sentence "I think that this car is red" is a subjective sentence expressing a belief but without polarization, while "I love this car" is also a subjective sentence but which conveys a positive sentiment.

The most complete definition of the sentiment analysis task is probably provided by Liu, B. (2012) in which an opinion is defined as a quintuple (*e, a, s, h, t*) where *e* is the name of an entity, *a* an aspect of this entity, *s* is a sentiment about *a*, *h* is the opinion holder, and *t* is the time when an opinion is expressed by *h*. For instance in "Linux configurability is loved by geeks", the entity *e* would be Linux, the aspect *a* would be the configurability of Linux, the sentiment *s* would be positive, the holder of the opinion *h* would be geeks, and the time *t* would be the time of utterance, that is (Linux, configurability, positive, geeks, now). The sentiment analysis task is then to output for a given text or sentence the set of opinions that this text conveys.

Table 1: Examples of subjectivity and sentimental sentences

	without sentiment	*with sentiment*
objective	This car is red	This car is broken
subjective	I think that this car is red	I love this car

4.4   Multilingual Sentiment Analysis

All the research works reviewed so far cover the sentiment analysis task for English, machine learning algorithms are trained on English corpora and lexicons are in English only. Most work in multilingual or cross-lingual sentiment analysis focus on subjectivity analysis that consists in determining whether a document or sentence is subjective or objective, but we assume that these approaches may be applied to polarity analysis as well. When performing sentiment analysis in another language several strategies are possible (Banea, C. et al. 2011), sorted here from the best to the worst cases. The ideal case is when an annotated corpus is available in the target language, or if it is possible to build such corpus automatically, then classical supervised learning methods are possible. If not, automated translation can be used: from an annotated English corpus to the target language and then use the translated corpus to train a classifier (Banea, C. et al. 2008), or vice versa by translating the target corpus into English and running an existing classifier for English (Kim and Hovy, 2006; Denecke, 2008), this operation can also be performed with manually aligned corpora (Mihalcea et al. 2007). Another approach is to build directly the lexicon for the target language, by lexicon-crawling and bootstrapping techniques. The last strategy is to translate instead the English lexicon to the target language and use the translated lexicon (Kim and Hovy, 2006; Mihalcea et al. 2007), however this method does not seem to yield the best accuracy. We refer the interested reader to more resources (Banea et al. 2010; Banea et al. 2011; Liu 2012) for a good overview of the topic.

4.5   Emotion Detection

The sentiment defined as a polarity is a worthwhile metrics of the user satisfaction, however in some domains we might want to explore more fine-grained metrics. The topic of "emotion detection" goes deeper in the description of the user feelings. It consists in finding the emotions that are conveyed by a sentence or a document. It can be considered as a more specific and harder task than sentiment analysis: sentiment analysis is restricted to two or three classes (positive, negative, neutral) while

emotion detection consists in finding emotions from a larger set of emotions and those are more complex than polarity. Typically, the set of emotions that is the most used is the Ekman's model (Ekman, P., 1972) that contains seven emotions (joy, sadness, fear, disgust, anger, surprise, interest), sometimes restricted to six, five or four emotions (joy, sadness, anger, fear). For instance the sentence "Trolley Square shooting leaves 6 dead, including gunman" could be said to convey a mixture of sadness, fear and anger.

It is interesting to compare the domain of application of emotion detection with regard to sentiment analysis. Emotion detection is mostly at hand in domains that really carry strong emotional content such as fairy tales (Alm, C. et al., 2005), dream diaries (Frantova, E. and Bergler, S., 2009), mood in blog posts (Mishne, G., 2005), news headlines (Strapparava, C. and Mihalcea, R., 2007), interaction with virtual agents (Ochs, M., et al. 2012), suicide notes (Pestian, J.P. et al., 2012), etc. Actually, in any domain where emotion detection is relevant, it is possible to carry a sentiment analysis task. However emotion detection is much more subtle and cannot be applied in all domains where sentiment analysis is relevant: the set of emotions needs to be chosen according to the domain. For instance, in the tutoring domain, Ekman's emotions do not seem to address correctly the experience of the student, it is unlikely that the student will feel anger or sadness per se. Kort, B. et al. (2001) propose instead to consider cognitive-affective states described by the valence of the experience and its relationship to the learning effects (constructive learning or un-learning). A similar claim is made by Calvo et al. (Calvo, R.A. and D'Mello, S., 2010; Baker, R. et al., 2010) in which cognitive-affective states like boredom or frustration are proposed. Interestingly these cognitive-affective states can be expressed in terms of valence/arousal, and then in this domain a sentiment analysis task would also be perfectly relevant.

To summarize, both sentiment analysis and emotion detection are interesting measures of the user experience. However, emotion detection may not be relevant in all domains and it is moreover widely acknowledged to be a more complex task than sentiment analysis. On the contrary, sentiment analysis is a relatively easier task, which can be applied in most domains but only provides binary polarity information about the user experience.

5. Implementing Some Prototypes for Sentiment Analysis and Emotion Detection

In the framework of the ITEA2 Empathic Products project (ITEA2 11005; http://www.empathic.eu) we are currently developing several prototypes for sentiment analysis and emotion detection (see Figure 4).
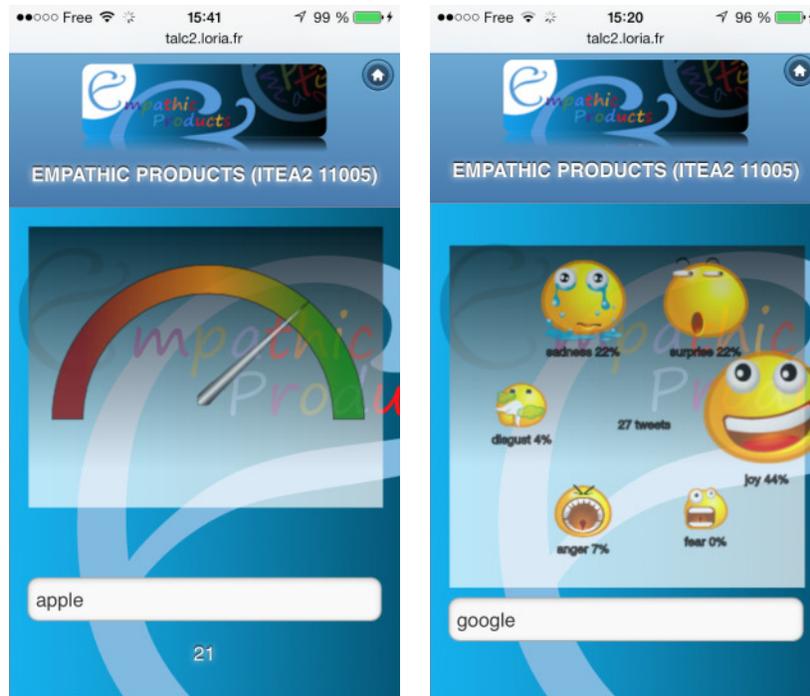
Figure 4: Sentiment analysis and emotion detection on a mobile device

On the left-hand side, a gauge indicates a positive (i.e. green), negative (i.e. red), or neutral (i.e., orange) sentiment detected on a keyword (i.e. apple). The sentiment analysis is performed in real-time on Twitter™. The number of tweets that have been analyzed is shown below the keyword (i.e. 21). On the right-hand side a similar prototype implements emotion detection. Once again, the detection is performed on a keyword (i.e. google), in real-time on Twitter™. The "emotion display" shows by using icons and text, the emotions that have been detected and the percentage related to each of the detected emotions (i.e., joy, fear, anger, disgust, sadness, and surprise). The number of tweets that have been analyzed is shown in the middle of the display (i.e., 27).

Both prototypes (i.e., sentiment analysis and emotion detection) are based on natural language processing tools. Roughly, for emotion detection, the approach is similar to Chaumartin, F.R.(2007), that is, we use an emotion lexicon, namely WordNet-Affect (http://wndomains.fbk.eu/wnaffect.html). WordNet-Affect is an extension of WordNet Domains (http://wndomains.fbk.eu/) including a subset of synsets (i.e., set of synonyms) suitable to represent affective concepts correlated with affective words (Strapparava, C. and Mihalcea, R., 2007). The sentence is analyzed with respect to several ordered filters (smileys, semantic rules, keywords) with a naive treatment of negation that inverts the found emotion. For sentiment analysis, we are currently exploring two opposed approaches, one symbolic and one with machine learning. The symbolic approach attempts to both tackle some linguistic difficulties (i.e., a word can have different polarities in different contexts) and domain dependence by using some general-purpose lexicons. The second approach relies on machine learning by training a classifier, namely a Random Forest classifier evaluated on the Semeval-13 dataset (SemEval, 2013).

6. Associating Sentiment Analysis and Emotion Detection to Virtual Worlds

In the framework of the ITEA2 METAVERSE1 (ITEA2 07016) project we have associated emotion detection from textual information (i.e. chat-based dialogues) to the Solipsis (http://en.wikipedia.org/wiki/Solipsis) and Second Life™ virtual worlds.

The general scenario is the following: two avatars meet somewhere in Solipsis (or SecondLife™) and start a discussion by using the chat tool. Roughly, we have used here the same approach as that of figure 4 (right-hand side). That is, each sentence in the chat is analyzed in real-time with respect to several ordered filters (smileys, semantic rules, keywords). Emotion detection is performed by using WordNet-Domains, WordNet-Affect, and a corpus containing several manually annotated phrases.

The Emotion detection architecture that has been implemented is shown in Figure 5. Elements in orange color are internal and thus dependent of a particular virtual world (Solipsis or Second Life™ in the ITEA2 METAVERSE1 project); elements in blue color are independent from any virtual world and they are related to our emotion detection prototype; finally, elements in green color represent external tools that we use associated to the emotion detection prototype (WordNet-Domains, WordNet-Affect, and part-of-speech taggers).
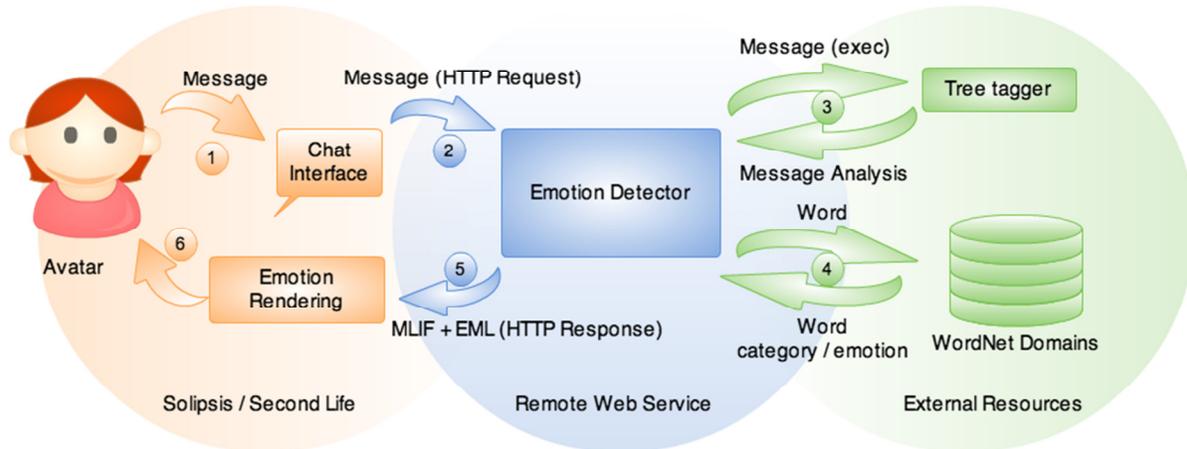
Figure 5: The emotion detection architecture in the ITEA2 METAVERSE1 project.

There are six steps represented in Figure 5:

1. A user, represented by an avatar, uses the chat tool in order to write a message;

2. The message is sent to the emotion detection prototype by using an http request;

3. The message is parsed and analyzed by an external tool: the "tree tagger" (a tool for annotating text with part-of-speech and lemma information (http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/). Roughly, this tool identifies the syntactic categories of the elements in the sentence: verbs, nouns, adjectives, adverbs, etc.

4. A search is then performed in WordNet-Domains and WordNet-Affect in order to try to find an association between domains, synsets and emotions;

5. If an association is found, the detected emotion is encoded (see the Standardization section) by using MLIF (multilingual information) and EmotionML (emotions) standards (see Figure 6).
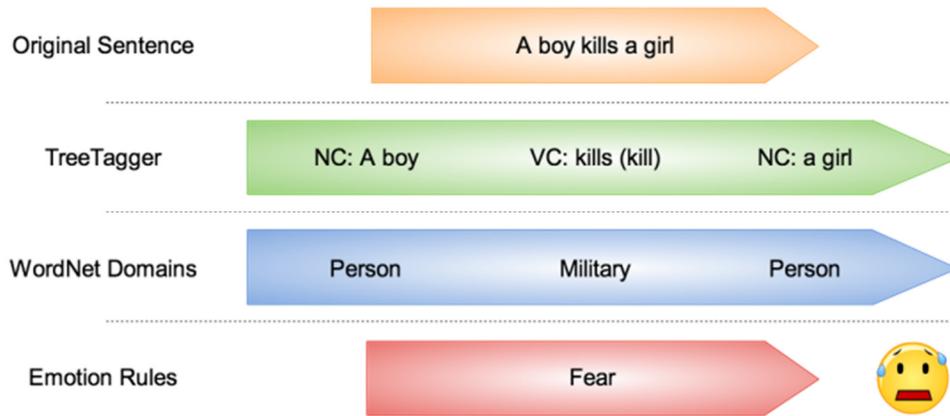
Figure 6: Detecting emotions in the ITEA2 METAVERSE1 project

6. Finally, the emotion rendering tool changes the face of the avatar according to the detected emotion. Figure 7 illustrates emotion detection in Solipsis.

Figure 7: Emotion detection in Solipsis.

Our emotion detection tool can also be directly accessed on the web (see Figure 8): http://talc.loria.fr/~metaverse/web/emotions/filterDetection/corpusCreation.php

This was the best summer I have ever experienced.

Send   Previous   Next

Sentence: This was the best summer I have ever experienced.

Detected emotion: (joy)

Filter	Smileys	WordNet-Domains	WordNet-Affect	Annotation Phrases
Detected Emotion			joy	joy
Key Elements			good	best summer, ever, experienced
Duration	0.0001	0.0024	0.0023	0.0013

Group types	NC	VC	NC	NC	VC
Words	This	was	the best summer	I	have ever experienced
Fields		be ○ Factotum	good ○ Factotum ○ Quality ○ Food ○ Medicine ○ Psychological_Features ○ Metrology Quality ○ Sociology ○ **joy** summer ○ Time_Period	I ○ Person	have ○ Factotum ○ Sexuality ○ Animals Biology ○ Sport ○ Gastronomy ever ○ Factotum experience ○ Factotum ○ Psychological_Features ○ Psychology

Figure 8: Figure 8. Our web-based emotion detection tool

6.1   E-learning Virtual world Emotion Tagging

E-learning in a virtual world requires some level of investment by the participants and is eased by their collaboration. It has been shown that emotional information can enhance the collaboration. For instance (Kamada, M. et al., 2005) show that participants interact more if provided with emotional clues about their partner's current state. We propose to integrate our sentiment analysis tool to the Umniverse collaborative virtual world (Reis, F., and Malheiro, R., 2011): the virtual world works as a situated forum in which participants can move around and submit posts (see Figure 9). When participants submit

posts they can annotate by hand the emotion that their post carries (with Ekman's emotions). Our tool can then be used to pre-annotate each post by proposing automatically an emotion.
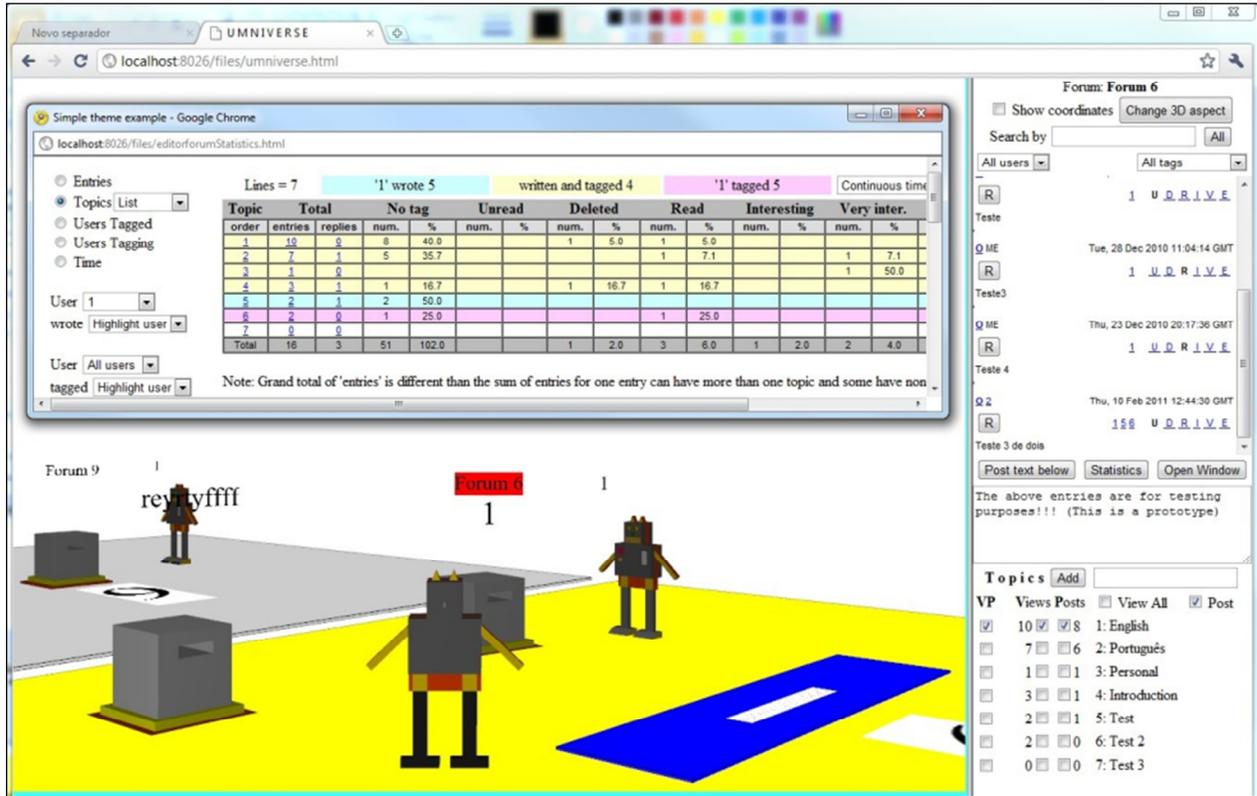

Figure 9: The Umniverse collaborative virtual world

After posts have been annotated and published to the forum, it is possible to filter the existing posts by their annotated emotion and as such find all the posts that carry sadness for example. Similarly to (Kamada, M. et al., 2005) the assumption is that the participants will collaborate more or better if they can express their emotions and react to their peers' emotions as well.

7. Standardization

In order to enhance interoperability between virtual worlds, applications, and multilingual corpora, using standards is mandatory. That is the reason why we have spent lot of efforts in the specification of the "MultiLingual Information Framework" (MLIF) [ISO 24616:2012] (Cruz-Lara S. et al., 2010; Cruz-Lara, S., 2011).

In this section, we will briefly present the EmotionML standard, and we will explain how we are currently using it within the ITEA2 Empathic Products project (ITEA2 11005).

7.1   The EmotionML Standard (from W3C's Website)

The Emotion Markup Language (EmotionML) is a W3C candidate recommendation aiming to strike a balance between practical applicability and scientific well-foundedness. The language is conceived as a "plug-in" language suitable for use in three different areas: manual annotation of data, automatic recognition of emotion-related states from user behavior, and generation of emotion-related system behavior.

So, roughly, use cases for EmotionML can be grouped into three broad types: First, **manual annotation of material involving emotionality**, such as annotation of videos, of speech recordings, of faces, of texts, etc; second, **automatic recognition of emotions from sensors**, including physiological sensors, speech recordings, facial expressions, etc., as well as from multi-modal combinations of sensors; and third, **generation of emotion-related system responses**, which may involve reasoning about the emotional implications of events, emotional prosody in synthetic speech, facial expressions and gestures of embodied agents or robots, the choice of music and colors of lighting in a room, etc.

Concrete examples of existing technology that could apply EmotionML include: **opinion mining / sentiment analysis in Web 2.0**, to automatically track customer's attitude regarding a product across blogs; **affective monitoring**, such as ambient assisted living applications for the elderly, fear detection for surveillance purposes, or using wearable sensors to test customer satisfaction; **character design and control for games and virtual worlds; social robots**, such as guide robots engaging with visitors; **expressive speech synthesis**, generating synthetic speech with different emotions, such as happy or sad, friendly or apologetic; expressive synthetic speech would for example make more information available to blind and partially sighted people, and enrich their experience of the content; **emotion recognition** (e.g., for spotting angry customers in speech dialog systems); **support for people with disabilities**, such as educational programs for people with autism. EmotionML can be used to make the emotional intent of content explicit. This would enable people with learning disabilities (such as Asperger's Syndrome) to realize the emotional context of the content; EmotionML can be used for media transcripts and captions. Where emotions are marked up to help deaf or hearing-impaired people who cannot hear the soundtrack, more information is made available to enrich their experience of the content.

In April 2011, an XML Schema and a MIME-type for EmotionML were defined. The following example shows how automatically annotated data from three affective sensor devices might be stored or communicated:

```xml
<emotionml xmlns="http://www.w3.org/2009/10/emotionml"
    category-set="http://www.w3.org/TR/emotion-voc/xml#everyday-categories">
 ...
<emotion start="1006526160000" expressed-through="face">
  <!--the first modality detects excitement.
      It is a camera observing the face. A URI to the database
      is provided to access the video stream.-->
  <category name="excited"/>
  <reference uri="http://www.example.com/facedb#t=26,98"/>
</emotion>

<emotion start="1006526160000" expressed-through="facial-skin-color">
  <!--the second modality detects anger. It is an IR camera
      observing the face. A URI to the database
      is provided to access the video stream.-->
  <category name="angry"/>
  <reference uri="http://www.example.com/skindb#t=23,108"/>
</emotion>

<emotion start="1006526160000" expressed-through="physiology">
  <!--the third modality detects excitement again. It is a
      wearable device monitoring physiological changes in the
      body. A URI to the database
      is provided to access the data stream.-->
  <category name="excited"/>
  <reference uri="http://www.example.com/physiodb#t=19,101"/>
</emotion>

<emotion start="1006526520000" expressed-through="physiology">
```

```
  <category name="angry"/>
  <reference uri="http://www.example.com/physiodb2#t=2,6"/>
</emotion>
 ...
</emotionml>
```

Example 1. Automatically annotated data from three affective sensor devices

7.2   Using EmotionML in the ITEA2 Empathic Products Project (ITEA2 11005)

We are currently using EmotionML for sentiment analysis and emotion recognition. The general idea is to perform sentiment analysis and emotion recognition from the syntactic and semantic analysis of sentences from chat-based tools or social networks such as Twitter™.

In order to be able to use different kinds of resources, we defined a unified XML schema for representing both lexicons and corpora. This format allows representing linguistic information in relation to EmotionML. The lexical aspects of the schema model are represented by lexical items (i.e., lemmas and their part-of-speech). The corpora aspects of the schema represent a corpus as a collection of texts, which can be decomposed into lists of sentences. The sentences themselves can be annotated as sequences of inflected forms/lemmas pairs, including syntactic information such as the dependencies. Both sentences and texts can be annotated with one or several EmotionML emotions. The schema has been implemented in XSD, and is used in our Java API.

The following example (SemEval-07, http://nlp.cs.swarthmore.edu/semeval/) shows how annotated lexical data might be associated to Ekman's basic emotions:

```
<sentence id="1414">
      <content>Trolley Square shooting leaves 6 dead, including gunman</content>
      <tagged-content>
         <tagged-word pos="NN" lemma="trolley" morph="trolley"/>
         <tagged-word pos="NN" lemma="square" morph="Square"/>
         <tagged-word pos="NN" lemma="shooting" morph="shooting"/>
         <tagged-word pos="VBZ" lemma="leave" morph="leaves"/>
         <tagged-word pos="CD" lemma="6" morph="6"/>
         <tagged-word pos="JJ" lemma="dead" morph="dead"/>
         <tagged-word pos="," lemma="," morph=","/>
         <tagged-word pos="VBG" lemma="include" morph="including"/>
         <tagged-word pos="NN" lemma="gunman" morph="gunman"/>
      </tagged-content>
      <emotion>
         <info>
             <origin source="semeval-ekman" xml:lang="en" />
         </info>
         <category confidence="1.0" value="0.87" name="sadness"/>
         <category confidence="1.0" value="0.18" name="surprise"/>
         <category confidence="1.0" value="0.0" name="joy"/>
         <category confidence="1.0" value="0.12" name="anger"/>
         <category confidence="1.0" value="0.0" name="disgust"/>
         <category confidence="1.0" value="0.57" name="fear"/>
      </emotion>
      <emotion>
         <info>
             <origin source="semeval" xml:lang="en" />
         </info>
         <category confidence="1.0" value="0.0" name="valence"/>
      </emotion>
</sentence>
```

Example 2. Associating basic emotions to annotated lexical data

Remarks and comments about Exemple 2:

<sentence> represents the sentence that has been analyzed;

<content> represents the sentence as a string;

<tagged-content> contains the elements in the sentence. The meaning of attributes is the following:

pos: part-of-speech morphological tags. Morphological tags represent different kinds of words such as: verbs, nouns, adjectives, etc.

lemma: the canonical form of a word;

morph: a word.

<emotion> an EmotionML element allowing to associate emotions to sentences. This element is fully based on EmotionML.

In the framework of sentiment analysis, we will also use Emotion-ML. The following example shows how annotated lexical data might be associated to positive, negative, or neutral sentiments:

```
<sentence id="266290644207681536">
       <content>@bryony_q Omg really totes loving this.. :) xxx</content>
       <emotion>
           <info>
               <origin source="twitter" xml:lang="en" />
           </info>
           <category confidence="1.0" value="1.0" name="valence"/>
       </emotion>
</sentence>
<sentence id="266460172858310656">
       <content>Major migrane going on now :( q</content>
       <emotion>
           <info>
               <origin source="twitter" xml:lang="en" />
           </info>
           <category confidence="1.0" value="0.0" name="valence"/>
       </emotion>
</sentence>
```

Example 3. Associating positive, negative, or neutral sentiments to annotated lexical data

Remarks and comments about Example 3:

<sentence> represents the sentence that has been analyzed;

<content> represents the sentence as a string;

<emotion> an EmotionML element allowing to associate sentiments to sentences;

Within the <category> element, the "value" attribute means: 0.0 is a fully negative sentiment, 1.0 is a fully positive sentiment, and 0.5 is a neutral sentiment.

7.3   Emotion Annotation

We first defined a Java interface EmotionAnnotator for annotating emotions for texts or sentences. In order to do several experiments on sentiment analysis, we implemented this interface in several ways. We first integrated the annotator that we have developed for the ITEA2 METAVERSE1 project. We then tried several other approaches, first at the lexical level (using only lexicon words valence) and at the syntactic level (using also valence shifting rules). The lexical level requires the use of a tagger, a tool that produces the lemmas from the inflected forms found in the text. As for the syntactic level, it requires the use of a parser to produce the syntactic dependencies between lemmas. For both purposes we based our work on the Stanford Core NLP (http://nlp.stanford.edu/software/corenlp.shtml) a high quality library that contains both taggers and parsers. Next, we tried statistical sentiment analysis. Thanks to the WEKA library (data mining software in Java, http://www.cs.waikato.ac.nz/ml/weka/), after importing our data within the WEKA format, we experimented with several supervised machine-learning algorithms. However, the two kinds of approach have their own advantages and deficiencies. While the lexical/syntactic annotators, including the annotator of the ITEA2 METAVERSE1 project, are relatively domain independent and perform well for simple sentences not too complex, they fail at capturing the diversity of valence phenomena and thus obtain low accuracy on our evaluation corpus (the Sanders corpus around 3300 annotated tweets with valence; http://www.sananalytics.com/lab/twitter-sentiment/). On the contrary, the statistical annotators manage to obtain fair accuracy on the Sanders corpus, but fail to be domain independent and thus tend to have wrong results on small generic sentences. The next objective is to see if it is possible to combine both kinds of annotators, either in a sequential way, or by learning domain independent valence shifting rules from our corpora.

7.4   The SATI API

The SATI API (http://talc2.loria.fr/empathic) enables to perform Sentiment Analysis from Textual Information. It is accessed through HTTP GET requests. The format of requests is: http://talc2.loria.fr/empathic?QUERY, where QUERY is formatted as parameter/value pairs separated by the "&" character.

For instance http://talc2.loria.fr/empathic?text=I+love+apples&engine=metaverse will perform the analysis of the sentence "I love apples" using the emotion detection prototype of the ITEA2 METAVERSE1 project, and return the value "joy". Figure 10 shows the available parameters of the SATI API while Figure 11 illustrates its multilingual part.

Parameters

text

name	text
description	the text parameter specifies the input text that is to be analyzed. The text value needs to be formatted using the URL encoding. See also RFC 3986 . The text value is preferably in English. We provide analysis in other languages than English, the input language being automatically detected by default. Note that if the text parameter is missing, the engine returns a default emotion.
type	string

lang

name	lang
description	the lang parameter specifies the input text language as standardized by the ISO 639-1. It accepts the default value "auto" to automatically detect the input language. The success of automatic language detection depends on the length of the input text. For short input texts, it is advised to specify the lang parameter.
type	string
default	auto

engine

name	engine
description	the engine parameter specifies which sentiment analysis engine will be used for sentiment analysis. Different engines may have different kinds of outputs. See the EmotionML vocabulary document for getting the types of vocabulary used byeach engine.
type	string
values	syntactic_liu_urban,syntactic_mpqa,metaverse
default	syntactic_liu_urban

output

name	output
description	the output parameter specifies how the output will be formatted. The "value" output is the simplest format: for engines that return valence information it returns a real value between 0 and 1, 1 being positive, 0 negative and 0.5 neutral. For engines that return emotion information it returns a value defined in a set of values that depends on the engine, for instance the metaverse engine returns a value in {joy, sadness, fear, disgust, ...}. The "xml" output corresponds to a XML representation of the emotion in the EmotionML format. The "json" output corresponds to the same content than the XML representation but formatted using JSON.
type	string
values	value,xml,json
default	value

query

name	query
description	the query parameter specifies the function of the api that is to be called. The "analysis" query corresponds to the main function that analyzes text. The "info" query returns information about the api formatted in XML. The "vocabularies" query returns the vocabularies used by the engines in EmotionML format.
type	string
values	analysis,info,vocabularies
default	analysis

twitter

name	twitter
description	the twitter parameter enables to perform the analysis of tweets that contain the given text. If true, it returns the sentiment analysis of the first found tweet containing the text as provided with the text parameter. Since this query waits for a tweet to be published in real time, it is better to call it asychronously with AJAX requests.
type	boolean
default	false

**Figure 10: Available parameters of the SATI API**
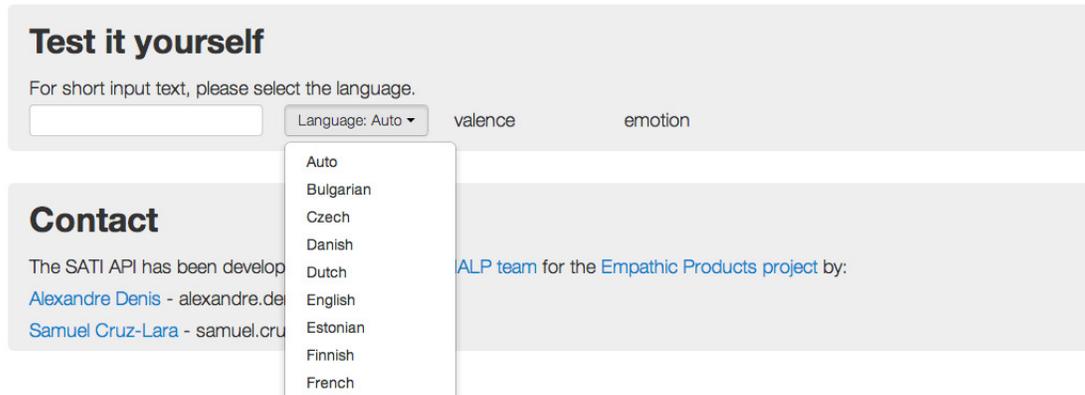
Figure 11: Multilinguality in the SATI API

It must be noted that multilingual sentiment analysis is actually performed in English (cf. multilingual sentiment analysis). All required translations, as well as language detection, are accomplished by using the Google Translate API.

8. Conclusion

We have presented a general review related to the association of linguistic and multilingual issues to virtual worlds and serious games. Beyond the "classical multilingual support" such as localization and automatic translation, we shown that several non-intrusive and interactive tools such as a linguistic enhanced chat, as well as several natural language processing-based activities, such as e-learning – and in particular language learning -, conversation support, sentiment analysis and emotion detection, may be successfully associated to virtual worlds and serious games.

Sentiment analysis and emotion detection have been extensively explored, and we think that we have clearly exposed how they can be associated to linguistic and multilingual issues in order to improve the user experience, both in virtual worlds and serious games.

Our research work, particularly in the framework of representing and managing multilingual textual information has contributed to the development of an ISO international standard the "Multi-Lingual Information Framework" (MLIF) [ISO 24616:2012] (Cruz-Lara S. et al., 2010; Cruz-Lara, S., 2011), but globally, our research work mainly contributes to the successful and innovative association of standardized linguistic-related tools to virtual worlds and serious games. As we have also shown in this paper, other a priori non linguistic-related standards such as EmotionML, can also be successfully used in this interesting association. Using standards is mandatory because they grant interoperability, stability, and durability.

It is also important to note that beyond virtual worlds and serious games, there is a third paradigm where linguistic and multilingual issues may be particularly interesting for improving the user experience: the augmented reality paradigm. This will certainly be our next field of research.

Augmented reality (Wikipedia. http://en.wikipedia.org/wiki/Augmented_reality) is a live, direct or indirect, view of a physical, real-world environment whose elements are augmented (or supplemented) by computer-generated sensory input such as sound, video, graphics or GPS data. Augmented reality lets people use familiar, everyday objects in ordinary ways. The difference is that these objects also provide a link into a computer network. Rather than immersing people in an artificially created virtual world, the

goal is to augment objects in the physical world by enhancing them with a wealth of digital information and communication capabilities.

For us, it is important to consider that virtual worlds and augmented reality are not two totally separated paradigms or, even less so, two opposite paradigms: augmented reality must be perceived as a way for linking real and virtual worlds (Mackay W.E. 1998).

Note that some interesting international research groups such as LangAR (http://labs.pearson.com/prototypes/langar-augmented-reality-talking-phrasebook/) are already working in this innovative field. Figure 12 illustrates how multilingual textual information may be added to the screen of a mobile device (i.e., augmented reality) when recognizing some particular images such as croissants and chocolate tarts. Amazing!

Figure 12: Multilinguality in the augmented reality LangAR project

References

Alm, C., Roth, D., Sproat, R. 2005. Emotions from text: machine learning for text-based emotion prediction. *Proceedings of Human Language Technologies Conference / Conference on Empirical Methods in Natural Language Processing*. Vancouver, Canada.

Amoia, M. (2011). I-FLEG: A 3D-Game for Learning French. *Proceedings of the 9th International Conference on Education and Information Systems, Technologies and Applications*. Orlando, Florida, USA.

Amoia, M., Gardent, C., and Perez-Beltrachini, L. (2011). A serious game for second language acquisition. *Proceedings of the Third International Conference on Computer Aided Education*. Noordwijkerout, The Netherlands.

Baker, R., D'Mello, S., Rodrigo, M.M., and Graesser, A.C. (April, 2010). Better to be frustrated than bored: the incidence, persistence, and impact of leaerner's cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4)*, 223-241.

Banea, C., Mihalcea, R., Wiebe, J., and Hassan, S. (2008). Multilingual Subjectivity Analysis Using Machine Translation. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Waikiki, Honolulu, Hawaii.

Banea, C., Mihalcea, R., and Wiebe, J. (2010). Multilingual Subjectivity: Are More Languages Better? *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, China.

Banea, C., Mihalcea, R., and Wiebe, J. (2011). Multilingual Sentiment and Subjectivity Analysis. In I. Zitouni $ D. Bikel (Eds.), *Multilingual Natural Language Processing*, NJ, USA: Prentice Hall.

Bretaudière, T., Cruz-Lara, S., and Rojas Barahona L. (2011). Associating Automatic Natural Language Processing to Serious Games and Virtual Worlds. *The Journal of Virtual Worlds Research, 4*(3).

Calvo, R.A., and D'Mello, S. (2010). Affect Detection : An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing, 1*(1)*, 18-37.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., Yan, H. (1999). Embodiment in Conversational Interfaces: Rea. *ACM CHI 99 Conference Proceedings*, Pittsburgh, PA.

Chaumartin, F.R. (2007). UPAR7: A Knowledge-Based System for Headline Sentiment Tagging. *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval 2007),* 422-425.

Cruz-Lara, S. (2011). Mundos virtuales: Una infraestructura global para facilitar las interacciones sociales multilingües y el aprendizaje de idiomas. *Proceedings of the VII International Conference on Engineering and Computer Education. Workshop "New trends in Engineering Education", Guimarães, Portugal.*

Cruz-Lara, S., Bellalem, N., Bellalem, L., and Osswald, T. (2009). Immersive 3D Environments and Multilinguality: Some Non-Intrusive and Dynamic e-learning-oriented Scenarios based on Textual Information. *The Journal of Virtual Worlds Research* 2(3).

Cruz-Lara, S., Francopoulo, G, Romary, L., and Semar, N. (2010). MLIF: A Metamodel to Represent and Exchange Multilingual Textual Information. *Proceedings of the Language Resources and Evaluation Conference, Valletta, Malta.*

Cruz-Lara, S., Osswald, T., Guinaud, J., Bellalem, N., and Bellalem, L. (2011). A Chat Interface Using Standards for Communication and E-learning in Virtual Worlds. *Lecture Notes in Business Information Processing 73,* Part 6, 541-554.

Denecke, K. (2008). Using SentiWordNet for Multilingual Sentiment Analysis. *In Proceedings of the IEEE 24th International Conference on Data Engineering Workshop (ICDEW 2008)*.

Ekman, P. (1972). Universals and Cultural Differences In Facial Expressions Of Emotions. *In J. Cole (ed.), Nebraska Symposium on Motivation, 1971* (pp. 207-283). Lincoln, Neb.: University of Nebraska Press.

Frantova, E., and Bergler, S. (2009). Automatic Emotion Annotation of Dream Diaries. *Proceedings of the Analyzing Social Media to Represent Collective Knowledge Workshop at K-CAP 2009, The Fifth International Conference on Knowledge Capture.*

Hu, M. and Liu, B. (2004). Mining opinion features in customer reviews. *Proceedings of AAAI 2004,* (pp. 755-760).

Kamada, M., Ambe, M., Hata, K., Yamada, E., and Fujimura, Y. (2005). The Effect of the Emotion-related Channel in 3D Virtual Communication Environments. *Psychology Journal, 3(*3), 312-327.

Kim, S. and Hovy, E. (2006). Identifying and Analyzing Judgment Opinions. *Proceedings of the Human Language Technology / North American Association of Computational Linguistics conference*. New York, NY.

Kort, B., Reilly, R., and Picard, R.W. (2001, August). An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. *Proceedings of the IEEE International Conference on Advanced Learning Technologies,* Madison, USA.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Liu, Y., Huang, X., An, A., and Yu, X. (2007). ARSA : a sentiment-aware model for predicting sales performance using blogs. *Proceedings of ACM SIGIR.*

Mackay, W.E. (1998). Augmented Reality: Linking real and virtual worlds: A new paradigm for interacting with computers. *Proceedings of ACM AVI '98, Conference on Advanced Visual Interfaces* (pp. 1-9). *L'Aquila, Italy: ACM. Keynote address.*

Mihalcea, R., Banea, C., and Wiebe, J. (2007, June). Learning Multilingual Subjective Language via Cross-Lingual Projections. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic*.

Mishne, G. (2005). Experiments with Mood Classification in Blog Posts. *Proceedings of the first Workshop on Stylistic Analysis Of Text For Information Access. 2005.*

Mishne, G. and Glance, N. (2006). Predicting Movie Sales from Blogger Sentiment. *Proceedings of AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*.

Ochs, M., Sadek, D., and Pelachaud, C. (2012). A Formal Model of Emotions for an Empathic Rational Dialog Agent. *Autonomous Agents and Multi-Agent Systems, 24*(3), 410-440.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of EMNLP-02,* (pp. 79-86).

Pestian, J.P., Matykiewicz, P., Linn-Gust, M., South, B., Uzuner, O., Wiebe, J., Cohen, K.B., Hurdle, J., and Brew, C. (2012). Sentiment Analysis of Suicide Notes: A Shared Task. *Biomedical Informatics Insights*, *(5)* Suppl. 1. Auckland, New Zealand: Libertas Academica.

Reis, F., and Malheiro, R. (2011, May). *Umniversity virtual world platform for massive open online courses University Platform. IE Challenges 2011 VII International Conference on ICT in Education, Braga, Portugal.*

Roque, A., and Traum, D.R. (2007). A model of compliance and emotion for potentially adversarial dialogue agents. *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*.

Sadikov, E., Parameswaran, A., and Venetis, P. (2009, May). Blogs as Predictors of Movie Success. *Proceedings of the Third International AAAI Conference on Weblogs and Social Media (Poster Papers).* San Jose Mc Enery Convention Center, USA.

SemEval (2013, June). SemEval-2013: Semantic Evaluation Exercises. *International Workshop on Semantic Evaluation*. Atlanta, USA.

Spertus, E. (1997). Smokey: Automatic recognition of hostile messages. *Proceedings of the Conference on Innovative Applications of Artificial Intelligence* (pp.1058-1065). Menlo Park, CA: AAAI Press.

Strapparava, C., and Valitutti, A. (2004, May). WordNet-Affect: an affective extension of WordNet. *Proceedings of the Language Resources and Evaluation Conference* (pp. 1083-1086), Lisbon, Portugal.

Tumasjan, A., Sprenger, T.O., Sandner, Ph.G., and Welpe, I.M. (2010, May). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.* Washington D.C., USA.

Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Philadelphia, Pennsylvania, USA, 417-424.