Automatic Translation System to Spanish Sign Language with a Virtual Interpreter

Sandra Baldassarri, Eva Cerezo, and Francisco Royo-Santas

Advanced Computer Graphics Group (GIGA), Computer Science Department,
Aragon Insittute for Engineering Research (I3A), University of Zaragoza, Spain
{sandra,ecerezo,497328}@unizar.es

Abstract. In this paper, an automatic translation system from Spanish language into Spanish Sign Language (LSE) performed by a virtual interpreter is presented. The translator is based on rules from Spanish grammar considering the syntactical and morphological characteristics of words and the semantics of their meaning. The system has been incorporated to an animation engine in which a virtual character acts as an interpreter that communicates using LSE. The mood of the interpreter is considered, so that the translation and the signs are modified depending on whether the interpreter is happy, angry, etc. The system has been tested with satisfactory results in speed and quality.

Keywords: Sign language translation, virtual characters, animation, emotion.

1 Introduction

In the last few years, the design of computer application interfaces has evolved in order to guarantee the accessibility of applications to everyone. Regarding the deaf community, a considerable amount of work has been done in the automatic translation into sign languages (SL). These languages, unfortunately, are not universal and each country has its own variety. In fact, most of the work done [1] is based on English grammar. This is the case of the works derived from ViSiCAST [2] and eSIGN [3] projects. Regarding Spanish Sign Language (LSE), San-Segundo et al. in [4] have developed a translator based on Spanish grammar that uses VGuido, an eSIGN avatar, but their application domain is very restricted (sentences spoken by an official when assisting people who are applying for their Identity Card).

None of the previous works includes mood. As in face-to-face communication, mood, emotions and facial expressions are an integral part of sign languages [5]. Words can considerable change their meaning depending on the mood/emotion of the speaker. Moreover, communicating in sign language without facial expressions would be like speaking in a monotonic voice: more boring, less expressive and, in some cases, ambiguous. The system presented in this paper is based on Spanish grammar, takes mood into account and has been integrated in an existing animation engine capable of managing expressive virtual characters to perform the signs.

2   Automatic Translation System to LSE

An automatic translation system from phrases in Spanish into LSE was developed as an independent module in C++ language. The system considers the syntactical and morphological characteristics of words and also the semantics of their meaning. The translation of a sentence or phrase is carried out by four modules (see Fig 1).
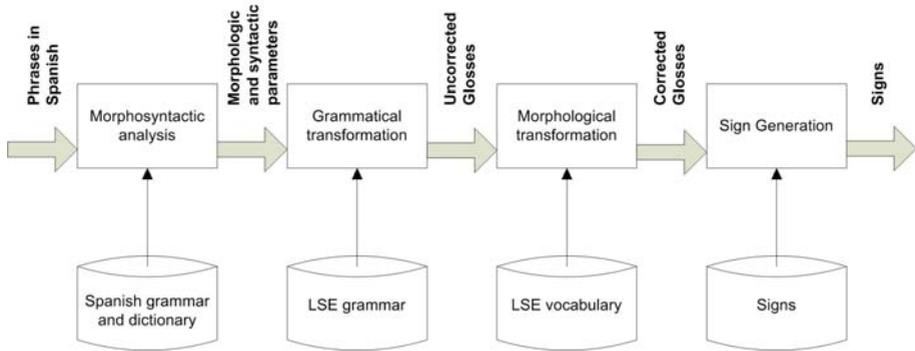

Fig. 1. Translation process

Morphosyntactic analysis: A phrase in Spanish is used as input. A series of parameters containing all the morphological information of the words as well as the relations and syntactical dependencies among them are drawn from it. This module uses the FreeLing analyzer [6], which was migrated to the Windows system.

Grammatical transformation: On the basis of the syntactic information gathered during the previous step, and through the application of grammatical rules, this module generates a series of glosses. The grammatical transformation module considers **mood** when it generates the translation, since mood influences the meaning and leads to the repetition of certain words, such as the nucleus, or to the appearance of new ones (similar to question tags or pet expressions). It can also be the case that certain blocks alter the order of the words to emphasize some of them.

Morphological transformation: Some of the glosses resulting from the previous step could be incorrect. This may occur when the original word in Spanish has no direct correlation to a term in LSE. Sometimes a synonym of that term will correlate directly to a term in LSE, but it can also occur that several signs are required in LSE to render a single word in Spanish. Or sometimes in Spanish several words can be used to express an idea that LSE expresses in a single sign. So, in this step, all the changes in the words are implemented, resulting in grammatically correct glosses.

Sign generation: Once the appropriate glosses have been produced (those which correspond directly to signs), in this step they are translated into a representation format that allows to generate the appropriated animations. **Mood** also influences the way in which specific words are signed. For example, the word "*no*" can be accompanied by different gestures. When the person signing is happy, he or she will move their finger, but if the person is angry, he or she usually resorts to dactylology and signs "N-O". In order to take these cases into account, the dictionary used for final

translation of the glosses into signs has been modified, allowing a word to be translated differently depending on the mood parameter.

3   Managing the Virtual Interpreter

The translation system has been incorporated in a scripted animation engine called Maxine [7]. In Maxine, virtual actors are provided with full expressive body, facial animation and with an emotional state which can modify character's answers, expressions and behaviours.
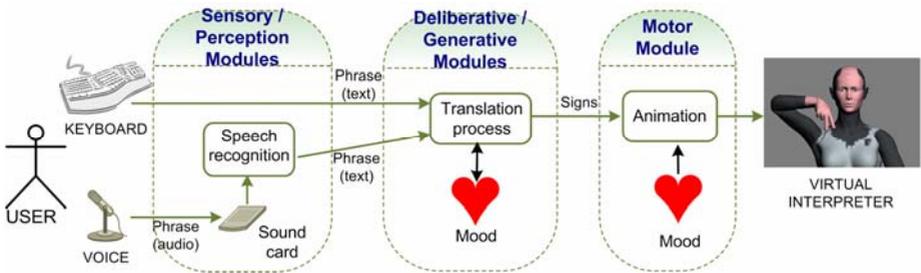

Fig. 2. Integration of the translation tool in the animation engine

The inputs of Maxine can be either written or spoken Spanish, generating phrases (see Fig 2). These phrases are given to the translation system, which returns the signs that must be animated. Also, the virtual character mood can be captured by a webcam [8] and it is supplied to the translator. The output consists in a virtual interpreter playing the signs corresponding to the translation, in Spanish Sign Language.

4   Assessment

Assessment was done by two teachers of a school of interpreters considering the accuracy of two aspects: the translation and the synthesis of the signs by the virtual interpreter.

The quality of translations was tested by means of a selection of phrases usually used during conversations.92 sentences with 561 words were selected. The phrases used in the measurement tests varied in length, ranging from simple phrases such as "How are you?" to more complex ones containing up to 25 words and several verbs along with subordination structures. The results obtained are: 539 well translated words (96%); 526 well translated and in the correct place inside the sentence (93.7%). 15 words were added without affecting the understanding of the sentences (2.7%), and 3 words were wrongly added affecting the meaning (0.6%). The results of the tests performed to check the translation speed were satisfactory and allowed to work in real time. The maximum time was 55 ms and the minimum 15 ms being the average for each sentence 19.59 ms.

Regarding sign synthesis, the results were validated video-recording a real interpreter signing different sentences and comparing with the same sentences performed

Fig. 3. Virtual and Real interpreters, signing the words "Sign Language" (in Spanish)

by the virtual interpreter in order to verify the visual quality of the sign animations (see Fig. 3).

5   Conclusions

The system presented performs automatic translation from Spanish to LSE and has been integrated in an animation platform that allows the signs to be performed by a virtual interpreter. This interpreter has a mood which may influence translations and signs. The quality of the results has been validated with sign language teachers, but in the near future a more formal evaluation with deaf people must be carried out

References

1. Ong, S., Ranganath, S.: Automatic sign language analysis: A survey and the future beyond lexical meaning. IEEE Trans Pattern Analysis and Machine Intelligence 27(6), 873–891 (2005)
2. http://www.visicast.sys.uea.ac.uk/
3. http://www.sign-lang.uni-hamburg.de/esign/
4. San-Segundo, R., Barra, R., Córdoba, R., D'Haro, L.F., Fernández, F., Ferreiros, J., Lucas, J.M., Macías-Guirasa, J., Monero, J.M., Pardo, J.M.: Speech to sign language translation system for Spanish. Speech Communication 50, 1009–1020 (2008)
5. Olivrin, G.: Composing Facial expressions with signing avatars. In: Speech and Face-to-Face Communication Workshop, Grenoble (2008)
6. FreeLing, http://garraf.epsevg.upc.es/freeling
7. Baldassarri, S., Cerezo, E., Seron, F.: Maxine: A platform for embodied animated agents. Computer & Graphics 32, 430–437 (2008)
8. Cerezo, E., Hupont, I., Manresa, C., Varona, J., Baldassarri, S., Perales, F., Seron, F.: Real-time Facial Expression Recognition for Natural Interaction. In: Martí, J., Benedí, J.M., Mendonça, A.M., Serrat, J. (eds.) IbPRIA 2007. LNCS, vol. 4478, pp. 40–47. Springer, Heidelberg (2007)