Providing Accessibility to Hearing-disabled by a Basque to Sign Language Translation System

María del Puy Carretero, Miren Urteaga, Aitor Ardanza, Mikel Eizagirre,
Sara García and David Oyarzun
Vicomtech-IK4 Research Center, P. Mikeletegi, 57, 20009 San Sebastián, Spain

Keywords: Virtual Character, Automatic Translation, Sign Language, LSE, Natural Language Processing.

Abstract: Translation between spoken languages and Sign Languages is especially weak regarding minority languages; hence, audiovisual material in these languages is usually out of reach for people with a hearing impairment. This paper presents a domain-specific Basque text to Spanish Sign Language (LSE) translation system. It has a modular architecture with (1) a text-to-Sign Language translation module using a Rule-Based translation approach, (2) a gesture capture system combining two motion capture system to create an internal (3) sign dictionary, (4) an animation engine and a (5) rendering module. The result of the translation is performed by a virtual interpreter that executes the concatenation of the signs according to the grammatical rules in LSE; for a better LSE interpretation, its face and body expressions change according to the emotion to be expressed. A first prototype has been tested by LSE experts with preliminary satisfactory results.

1 INTRODUCTION

This paper presents a modular platform to translate Basque into LSE. The modularity of the platform allows the input to be audio or text depending on the needs and the available technology of each application case. The main objective of translating spoken language into Sign Language is to allow people with hearing loss to access the same information as people without disabilities. However, beside social reasons, this research project has also been motivated by legal and linguistic reasons

1.1 The Importance of Sign Languages

Sign Languages are the natural languages that deaf people use to communicate with others, especially among them. The following facts on Sign Languages and deaf people highlight the need of doing research on Sign Languages:

- *Not all deaf people can read*. A deaf person can read well if deafness has come at adult life. However, in most deafness cases at prenatal or infant age the ability to interpret written natural text does not develop, due to the difficulty to acquire grammatical and void word concepts. Therefore, these types of deaf people are unable to read texts and/or communicate with others by writing.
- *Lipreading.* Some deaf people can read lips, but it is not a general ability. Furthermore, lipreading alone cannot sufficiently support speech development because visual information is far more ambiguous than auditory information. Sign Language involves both hands and body. Hand gestures should always be accompanied by facial and corporal expressiveness. The meaning of a sign can change greatly depending on the face and body expression, to the extent that it can disambiguate between two concepts with the same hand-gesture.
- *Sign Language is not Universal.* Each country has its own Sign Language, and more than one may also co-exist. For example, in Spain both Spanish Sign Language and Catalan Sign Language are used. A deaf person considers his/her mother tongue the sign language used in his/her country.

1.2 Legal Framework

According to statistics collected by the National Confederation of Deaf People in Spain (CNSE),

Carretero M., Urteaga M., Ardanza A., Eizagirre M., García S. and Oyarzun D..
Providing Accessibility to Hearing-disabled by a Basque to Sign Language Translation System.
DOI: 10.5220/0004748502560263
In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence* (ICAART-2014), pages 256-263
ISBN: 978-989-758-015-4

there are approximately 120.000 deaf people in Spain who use Sign Language as their first language. Backing up deaf people's rights, the Spanish General Law of Audiovisual Communications (CESyA, 2010) lays down that private and public channels have to broadcast a certain amount of hours per week with accessible contents for disabled people. Table 1 summarizes the amount of hours per week that broadcasters must provide with accessible content for deaf people by the end 2013:

Table 1: Accessibility requirements of the Spanish General Law of Audiovisual Communication.

	Private service	Public service
Subtitles	75%	90%
Sign language	2 hours	10 hours
Audio description	2 hours	10 hours

In order to adapt the contents to Sign Language an interpreter is necessary. Adapting contents to Sign Language is expensive and requires several hours of work for off-line contents and more than one interpreter for live contents.

In this paper we present a platform to adapt not only audiovisual contents but also other kind of information to Sign Language by using virtual characters whose role is to interpret the meaning of the content.

1.3 Inaccessibility to Basque Contents

We have started to develop a translator from Basque to LSE because of two main reasons:

On the one hand, in the Basque Country there are two main spoken languages: Spanish and Basque. Not all deaf people have the ability to read and for those who are able, learning two languages is very challenging. Hence, those who learn a spoken language tend to choose Spanish because of its majority-language status (Arana et al., 2007). Due to the hegemony of the Spanish language, all the audiovisual content and information in Basque is out of reach for the deaf community. Therefore, there is a need to provide tools to make all these Basque contents accessible to the deaf community.

On the other hand, while research on translation from Spanish to LSE is being already studied and developed by other research groups, no research has been made on Basque translation into LSE. This project takes this challenge to make research on Basque translation and analysis.

2 RELATED WORK

At an international level, the following research works are the most relevant ones at trying to create a translation platform from spoken language to Sign Language using virtual characters:

The European project ViSiCAST (Verlinden et. al., 2001) and its continuation eSIGN (Zwiterslood et. al., 2004) are among the first and most significant projects related to Sign Language translation. It consists of a translation module and a sign interpretation module linked to a finite sign data-base. The system translates texts of very delimited domains of spoken language into several Sign Languages from the Netherlands, Germany and the United Kingdom. To build the sign data-base, a motion capture system was used and some captions were edited manually when needed. The result was good regarding the execution of the signs and the technology of that period. However, facial and body expression were not taken into account, which is one of the most fundamental aspects to understand Sign Language correctly.

Dicta-Sign is a more recent European project (Efthimiou et. al., 2010). Its goal was to develop technologies to allow interaction in Sign Language in a Web 2.0 environment. The system would work in two ways: users would sign to a webcam using a dictation style; the computer would interpret the signed phrases, and an animated avatar would sign the answer back to the user. The project is being developed for the Sign Languages used in England, France, Germany and Greece. In addition authors affirm that their internal representation allows them to develop a translator among different sign languages.

There have been other smaller projects designed to resolve local problems for deaf people in different countries, each project working with the Sign Languages from: United States (Huenerfauth et. al., 2008), Greece (Efthimiou et. al., 2004), Ireland (Smith et. al., 2010), Germany (Kipp et. al., 2011) Poland (Francik and Fabian, 2002) or Italy (Lombardo et. al., 2011). All of these projects try to solve the accessibility problems that deaf people have to access media information, communicate with others, etc.

Regarding LSE, there have been several attempts to build an automatic translator. Unfortunately, none of them seem to have a huge repercussion on the Spanish deaf community. The Speech Technology Group at Universidad Politécnica de Madrid have been working on a both way translation-system from written Spanish into LSE (San-Segundo et. al.,

2006) as well as translation from LSE into written Spanish (Lopez et al., 2010). The system is designed for a very delimited usage domain: the renewal of Identity Document and Driver's license. In a more challenging project Baldassarri et. al. (2009) worked on an automatic translation system from Spanish language to LSE performed by a virtual interpreter. The system considers the mood of the interpreter modifying the signs depending whether the interpreter is happy, angry, etc. A more recent research translates speech and text from Spanish into LSE where a set of real-time animations representing the signs are used (Vera et. al., 2013). The main goal is to solve some of the problems that deaf people find in training courses. Finally, the project textoSign (www.textosign.es) is a more oriented product whose goal is to provide a translation service working in real time which may be integrated into websites, digital signage scenarios, virtual assistants, etc.

The objective of our first prototype is to provide a domain specific spoken language-LSE translation platform. Once its performance is validated, it could be easily adapted to other domains in further developments. The prototype has been built and tested on the weather domain. This domain was chosen for (1) using a relatively small and predictable vocabulary, (2) having just one speaker and (3) showing graphic help such as weather maps as a cue for potential mistranslation cases. The first prototype introduces two novelties: on the one hand, the avatar will process the hand-gesture and the bodily expression separately according the required emotion; on the other hand, it will translate from Basque, a co-official language in Spain and until the moment of writing of this article, inaccessible for the deaf community.

3 BASQUE TO LSE PROTOTYPE

The main goal of the prototype is to make audiovisual content accessible to deaf people in a domain specific content, so the global architecture is validated before further development.

The system translates the input message into LSE with the help of a virtual avatar. Figure 1 shows the architecture of the platform and its functionality. The input of the platform can be of diverse origin: television, web pages, health-care services, airports, conferences, etc.

For the first prototype, the system only allows text and written transcriptions of spoken language as input; however, a speech transcription module is

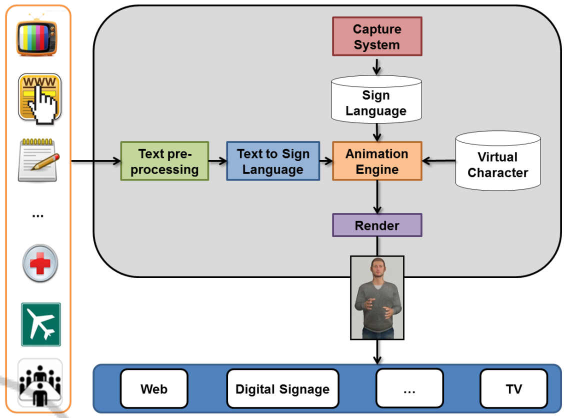

Figure 1: System workflow.

needed for spoken inputs such as conferences, live television shows or talks. Even if the development of a Basque speech recognizer is under development, right now we apply voice alignment technologies to align subtitles with the audio. The system has a modular architecture to allow integrating further speech and voice technologies in the future. Right now the prototype works with .srt files that are prepared before the shows based on the autocue scripts. Voice alignment technologies are used to match the speech to the script and synchronize the translation.

The current system consists of five different modules: (1) a text-to-Sign Language translation module, (2) a gesture capture system to create an internal (3) sign dictionary, (4) an animation engine and a (5) rendering module. The output of the platform can also be very diverse depending on the ultimate context and target audience: television, web pages, digital signage systems, mobile devices, etc.

3.1 Text-to-Sign Language Translation Module

The first prototype to validate the project is domain-specific: it automatically translates a TV weather forecast program in Basque into LSE.

Due to the lack of annotated data and the fact that it is impossible to gather parallel corpora in Basque and LSE, statistic machine translation approaches were discarded, and a Rule-Based Machine Translation approach was chosen. The rules were designed taking into account a corpus from the application domain of the prototype. To do so, a code-system has been created in order to represent LSE signs in written strings. Each concept that in LSE has a fixed sign has its corresponding tag in our written representation of LSE.

3.1.1 Linguistic Analysis of the Application Domain

Before building the rules from scratch, the language used in the application domain was linguistically analysed. To do so, a domain-specific corpus was compiled *ad hoc*. This corpus is composed by transcriptions of 10 weather shows from the Basque-speaking TV channel ETB1. It contains over 8000 words (2176 distinct words and 472 lemmas), corresponding to approximately 55 minutes of manually transcribed contents. In order to make a thorough linguistic analysis, the corpus was tagged semi-automatically indicating the lemma, morphemes and the morphosyntactic information of each token. The morphemic information, only reachable by a deep linguistic analysis, is of special relevance taking into account that Basque is an agglutinative language. This is one of the major challenges encountered in Basque automatic translation. The linguistic information extracted was used to spot linguistic patterns and to build robust translation rules.

3.1.2 Text Pre-processing

The current system runs with subtitle text as input (.srt files). These subtitles are produced before the show and are also used as autocue help for the TV presenter.

```
00:00:30,410 --> 00:00:33,060
Atzo guk iragarritakoak baino ekaitz
gutxiago izan ziren baina gaur...
(Yesterday there were fewer storms
than we had foreseen, but...)
```

First, the input is pre-processed to make it suitable for automatic translation. Subtitles as such, without text processing, follow readability standards that hinder automatic translation: there are no more than 32 characters per line, not more than two lines per screenshot, etc. Hence, these sentences need to be reconstructed before translation. The text pre-processing module joins and splits different sentences taking into account the information obtained from the tokenization and the capital letters. Besides, the whole text is lemmatized and tagged using an inner dictionary.

```
[atzo|den][gu][iragarri|ad][ekaitz][gutxi][baina|0][gaur|den][hemen][ipar][haizea][sartu|ad][kosta]
([yesterday|time][we][foresee|v][storm][few][to_be|v][to_be|aux][but|0][today|time][north][wind][enter|v][to_be|aux][coast])
```

This tagging includes grammatical remarks (word-class, time-related word or morpheme, etc.) and other kind of linguistic information that should be taken into account when signing in LSE. Within this process, sentence splitting marks are also inserted following LSE standards. Not all Basque sentences match with LSE sentences; subordinate clauses, for example, must be expressed in separate sentences in LSE. The strict word order and little abstraction of Sign Languages makes that sentences force these languages to build shorter and linguistically simpler sentences. The LSE sentence splitting process is done by rules taking into account linguistic cues.

3.1.3 Sentence Translation

The objective of this module is to give as output a sequence of signs that strictly follows the LSE grammar. The translation module takes as input the processed subtitles previously obtained. The output is a string with tags that correspond to specific signs in LSE:

```
[atzo|den][gu][ekaitz][gutxi][iragarri|ad][baina|0][gaur|den][hemen][ipar][haizea][kosta][sartu|ad]
([yesterday|time][we][storm][few][foresee|v][but|0][today|time][north][wind][coast][enter|v])
```

Taking into account the segmentation tags and other morphosyntactic information, it translates the sequence that follows Basque syntax into a string of codes that follow the LSE linguistic rules though pattern identification. There are three types of translation rules according to the action they imply:

- *Explicitation:* in LSE grammatical information is expressed with explicit gestures. However, in Basque it is very common to have grammatical information in elliptical form (e.g.: subjects, tense, objects, etc.). This information is usually expressed by inflected word-forms and suffixes. The explicitation rules combine all the grammatical information contained in the linguistic tags and the semantic information contained in the lemmas. Thus, the output number of tokens, each one of them containing one grammatical or semantic piece of information. Example:

```
[[.*].*[.*|v][.*|aux_past][.*].*]
                        ↓
[[.*].*[.*|v][.*|aux_past][yesteday|
time][.*].*]
```

- *Selection of concepts:* LSE does not use many grammatical and semantically void words that verbal languages usually have (e.g.: articles, grammatical words, etc.). The selection rules select the tokens that have to be expressed in LSE and leave out the ones that do not make sense in this language. Example (erasing auxiliary verbs):

```
[[.*].*[.*|v][.*|aux_past][yesteday|
time][.*].*]
                        ↓
[[.*].*[.*|v][yesteday|time][.*].*].
```

- *Reorganization or syntax rules*: these rules change the order of the tokens to adjust it to the syntactic rules of LSE. The output of this module consists of a set of tokens that can be translated directly into a sequence of gestures. The reorganization rules involve splitting of sentences according to LSE rules. Example (SOV pattern with time cue at the beginning)

```
[[.*].*[.*|v][yesteday|time][.*].*]
                        ↓
[[yesteday|time][.*].*[.*|v]]
```

3.2 Capture System Module

In order to translate Basque into LSE, we had to compile a LSE data-base. To do so we have developed a capture system combining two different motion capture systems. It uses non-invasive capturing methods and allows entering more sign-entries quite easily. The system can be used by any person but only one person can use the system in each capture session.

3.2.1 Caption of Wrist and Hand Movements

The majority of the full-body motion capture systems are not accurate enough when capturing hand movements, especially movements involving wrist and finger movements. These movements usually require greater precision. Given their importance in LSE, two motion capture CyberGlove II gloves were used, one for each hand. These gloves allow tracking precise movements of both hand and fingers. They connect to the server via Bluetooth, which allows more comfortable and free movements when signing.

3.2.2 Caption of Body Movements

As mentioned before, body movements have also a great significance in LSE. In order to capture the movements of the whole body, the Organic Motion system was used. This system uses several 2D cameras to track movements. The images are processed to obtain control points that are triangulated to track the position of the person that is using the system. Thanks to this system, the person signing does not have to wear any kind of sensors, allowing total freedom of movement. The captured movements result more natural and realistic.

It is important that the signs are made by Sing Language experts for a better accuracy and understanding. The person signing has to wear the gloves while standing inside the Organic Motion System at the same time.

3.2.3 Merging of Hand and Body Captions

In order to join the animations captured with both systems it is necessary to join and process the captions before saving them as whole signs. Autodesk Motion Builder is used for that purpose. This software is useful to capture 3D models in real time and it allows creating, editing and reproducing complex animations.

Both CyberGlove II and Organic Motion provide plug-ins to use with Motion Builder that allows tracking and synchronizing the movements in the same 3D scenario in real time. For that purpose three skeletons are needed: one for each hand and one for the whole body. These skeletons are joined in one unique skeleton to simulate the real movements of the person who uses the capture system. Motion Builder allows making it in a semi-automatic way reducing pre-processing.

Although both CyberGlove II and Organic Motion obtain realistic movements, it is possible that in certain cases the capture may contain errors due to different reasons, such as calibration or interferences. In these cases, manual post edition of the capture is needed. This manual edition consists of comparing the obtained animations with the real movements and adjusting the capture using Motion Builder. Capture edition is always done by an expert graphic designer.

Once the realistic animations are obtained, they are stored in a database to feed the platform with vocabulary in Sign Language.

3.3 Sign Dictionary

The sign or gesture dictionary contains the words used in the code given to each concept linked to the actual gesture that the avatar has to interpret. The gesture dictionary is composed by a finite number of lemmatized concepts gathered from the domain-specific corpus. Furthermore, all synonyms are gathered within the same entry. The sign dictionary can contain three types of entries:

- *One-to-One Concepts*: concepts that match a word-token in Basque and that are expressed in one sign in LSE. Synonyms are listed under the same LSE sign.
- *Grammatical or void Words*: these entries are listed in the dictionary as evidence of processing, but are linked to an empty concept. They do not trigger any kind of movement because in LSE they do not exist.
- *Multi-word Concepts*: some concepts may map to more than one word-token in Basque. These concepts are registered as one entry in the gesture dictionary and they map to just one concept in LSE.

The current sign dictionary contains 472 lemmas. These entries have proved to be enough to translate the domain-specific corpus used to extract the translation rules. All these concepts or lemmas need to be captured with Capture System Module so they are added to the sign dictionary and can be interpreted by the virtual interpreter

3.4 Animation Engine

As explained in the second section, there are several research projects involving a signing virtual character. Our Animation Engine is developed with the aim of providing natural transitions between signs as well as modifying the execution of the signs depending on the emotion of the virtual interpreter. Emotion is essential in LSE. Each sign should be represented using not only the hands and the face, but at least, also the upper body of the interpreter. It is based on executing the corresponding sign and changing the speed of the animation depending on the emotion that the virtual interpreter has to reproduce according to the real input at that moment.

The Animation Engine runs as follows: the appearance of the virtual interpreter is loaded from the Virtual Character database. While the virtual interpreter does not receive any input it has a natural behaviour, involving blinking, looking sideways, changing the weight of the body between both feet, crossing arms, etc. When the Text to Sign Language module sends the translation to the Avatar Engine module, it stops the natural behaviour (except blinking) and starts the sequence of signs. If any emotion or mood cue is registered as input, the speed of the animation changes accordingly; for example, it slows down if sad, speeds up if angry. Additionally, the virtual interpreter’s expression is also modified using morphing techniques. For the first prototype, the emotion is indicated manually choosing one of the six universal emotions defined by Ekman (1993), since the automatic emotion recognition module is at a very early stage of development.

The Animation Engine module is developed using Open Scene Graph. It applies any sign animations stored in Sign Language database captured with the Capture System Module previously. In order to concatenate several animations and to obtain realistic movements, a short transition between the original signs is introduced. The implemented algorithm takes into account the final position of a sign and the initial position of the next sign; it is implemented using the technique explained by Dam et., al. (1998).

Thus, the final result is the virtual interpreter signing with very realistic movements. The execution of the signs, as well as facial expression, is modified depending on the emotion. Figure 2 shows the same sign executed with different emotions.

Figure 2: The same sign with different emotions. From left to right: neutral, angry and sad.

3.5 Rendering

The objective of this module is to visualize the virtual interpreter synchronized with other possible multimedia contents. For the current prototype, this module inserts the avatar in the broadcasted TV show. For synchronizing the virtual interpreter with the visual content the system takes into account the time stamps in the .srt subtitles.

4 EVALUATION

The evaluation of the system was carried out at two levels: the translation module and animation module were evaluated separately by LSE experts. This allows spotting more easily the source of the mistakes. The feedback of the evaluation is being used as material for improvement.

4.1 Translation Evaluation

The text-to-LSE module was evaluated by one real interpreter. In order to do the evaluation we used the subtitles of four different TV weather program as input of the module and we obtained the corresponding string of signs (translated into Spanish in request of the evaluator). Then the result was sent to the real LSE interpreter. The sentences to be evaluated showed the original sentence and its "translation" to LSE with written codes in brackets referring to one sign each.

```
   Kaixo, arrasti on guztioi
ikusentzuleok.
   [Hola][tardes][buenas][todos]
   [Hello][afternoon][good][everybody]
```

PThe real interpreter sent back the result correcting the sentences that contained mistakes.

```
   Kaixo, arrasti on guztioi
ikusentzuleok.
   [Hola][tardes][buenas][todos]
   HOLA BUENAS TARDES TODOS
   [Hello][good][afternoon][everybody]
```

There were translated 368 sentences with the text to LSE module. The real interpreter returned 34 of these sentences corrected. Hence, according to this evaluation, the 90.7% of the translated sentences follow strictly the LSE rules. Most of the mistakes deal with the position of numbers or adjectives; hence, according to the evaluator, the general meaning of the sentences was well transmitted despite these mistakes.

4.2 Animation Engine Evaluation

For the evaluation of the LSE animation engine a hearing disabled person, an LSE interpreter and a Basque/LSE bilingual person collaborated with us. After watching the avatar execute some sentences, they made several comments that are summarized as follows:

Movements and their transition should come up smoothly, especially when spelling.

The emotions could be recognized, but the expression of the virtual interpreter ought to be more exaggerated, using eyebrow, gaze and shoulder movements. The waiting state of the avatar comes up very natural, but should move less not to attract unnecessary attention.

Some hand-modelled signs were not natural. It was agreed that the movement caption system was preferable.

The avatar should always wear dark clothes in order to give more contrast between hands and clothes.

In general, the evaluation of experts in LSE was positive; they appreciated the efforts made to make media services and other kind of information more accessible. They gave a very positive feedback to the first prototype. We used their assessment to improve the animation engine in a second round. The animation engine and spelling were improved, in order to result more natural; the avatar's clothes have been changed to a darker colour; and both expressions and emotions have also been exaggerated.

In addition we are adding rules to interpret questions and exclamations showing them on the character's face.

5 CONCLUSIONS

This prototype system is developed to translate Basque text into LSE. This first version it is domain-specific and its dictionary and grammar are limited. In spite of these limitations, the evaluators high lightened the impact this platform may have for the deaf community: it makes Basque contents accessible for the first time. Beside its social impact, this project has proven the usability of Rule-Based translation approaches in language combinations where no parallel corpora is available. The Capture System used has also been another of the keys for the success of the project. Being a non-invasive system with Bluetooth technology has allowed recording more natural movements. The work done on the animation engine has also given its results on the smooth concatenation of signs and the execution of the virtual interpreter's expressions. The expression of the avatar is another of the aspects that was positively remarked by the experts, since the system modifies it depending on the mood or emotion that has to be expressed.

As future work, apart from integrating

technology to adapt other kind of inputs (e.g. audio input), we are planning to add another motion capture system to track the face of the person signing. This will allow the signs to be even more realistic, without the need of editing face expressions manually. Furthermore, we are planning to extend the Text to Sign Language module to other domains. The system could also integrate other languages such as Catalan and Catalan Sign Language in further developments. Finally, we plan to add an emotion recognition module in order to recognize emotion from the voice of the speakers and automatically modify the virtual interpreter accordingly. Thanks to the modular architecture of the project, it is relatively easy to integrate external modules and/or reuse some modules in other projects, multiplying the usability and impact of the work done.

All these steps are planned to be taken under LSE experts' supervision and constant feedback.

ACKNOWLEDGEMENTS

This work have been evaluated by AransGi (http://www.aransgi.org)

REFERENCES

Arana, E., Azpillaga, P., Narbaiza, B., 2007. Linguistic Normalisation and Local Television in the Basque Country. In *Minority Laguage Media: Concepts, Critiques and Case Studies. Multilingual Matters. Noth York, 2007.*

Baldassarri, S., Cerezo, E., Royo-Santas, F., 2009 Automatic Translation System to Spanish Sign Language with a Virtual Interpreter. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I, Uppsala, Sweden: Springer-Verlag, 2009, pp. 196-199.*

CESyA. 2010. Ley General Audiovisual. Available at: http://www.cesya.es/es/normativa/legislacion/ley_general_audiovisual (accessed September 11, 2013).

Dam, E., Koch, M., Lillholm, M., 1998. *Quaternions, Interpolation and Animation*, Technical Report DIKU-TR-98/5, Department of Computer Science, Univ. Copenhagen, Denmark.

Ekman, P., 1993. Facial expression and emotion. In *American Psychologist, Vol. 48(4) pp 384–392.*

Efthimiou, E., Sapountzaki, G., Karpouzis, K. Fotinea, S-E., 2004. Developing an e-Learning platform for the Greek Sign Language. In *Lecture Notes in Computer Science (LNCS) , in K. Miesenberger, J. Klaus, W. Zagler, D. Burger, Springer (eds), Vol. 3118, 2004, 1107-1113.*

Efthimiou E., Fotinea S., Hanke T., Glauert J., Bowden R., Braffort A., Collet C., Maragos P., Goudenove F. 2010. DICTA-SIGN: Sign Language Recognition, Generation and Modelling with application in Deaf Communication. In *4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 80-84.*

Francik J. and Fabian, P., 2002. Animating Sign Language in the Real Time. In *20th IASTED International Multi-Conference on Applied Informatics AI 2002, Innsbruck, Austria, pp. 276-281.*

Huenerfauth, M., Zhou, L, Gu, E., Allbeck, J., 2008. Evaluation of American Sign Language generation by native ASL signers. In *ACM Trans. Access. Comput. 1, 1, Article 3.*

Kipp, M., Haleoir, A. and Nguyen, Q., 2011. Sign Language Avatars: Animation and Comprehensibility. In *Proceedings of 11th International Conference on Intelligent Virtual Agents (IVA)*. Springer.

Lombardo, V., Battaglino, C., Damiano, R. and Nunnari, F., 2011. An avatar–based interface for the Italian Sign Language. In *International Conference on Complex, Intelligent, and Software Intensive Systems.*

Lopez, V., San-Segundo, R., Martin, R., Lucas, J.M., Echeverry, J.D., 2010 Spanish generation from Spanish Sign Language using a phrase-based translation system. In *VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop, Vigo, España.*

San-Segundo, R., Barra, R., D'Haro, L.F., Montero, J.M., Córdoba, R., Ferreiros. J., 2006. A Spanish Speech to Sign Language Translation System. In *Proceedings of Interspeech 2006 Pittsburgh, PA.*

Smith, R., Morrissey, S., and Somers, H., 2010. HCI for the Deaf community: Developing human-like avatars for sign language synthesis. In *Irish Human Computer Interaction Conference'10, September 2–3, Dublin, Ireland.*

Vera, L., Coma, I., Campos, J., Martínez, B., Fernández, M. Virtual Avatars Signing in Real Time for Deaf Students. In *GRAPP 2013 – The 8th International Conference on Computer Graphics Theory and Applications.*

Verlinden, M., Tijsseling, C., Frowein, H., 2001. A Signing Avatar on the WWW. In *Proceedings of the 18th International Symposium on Human Factors in Telecommunication, Bergen, Norway.* K. Nordby, (ed.).

Zwiterslood, I., Verlinden, M., Ros, J., van der Schoot. S., 2004. Synthetic signing for the deaf: Esign. In *Proceedings of the Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment, Granada, Spain, July 2004.*