Language Resources for Spanish - Spanish Sign Language (LSE) translation

Rubén San-Segundo[1], Verónica López[1], Raquel Martín[1], David Sánchez[2], Adolfo García[2]
[1]Grupo de Tecnología del Habla-Universidad Politécnica de Madrid
[2]Fundación CNSE

Abstract

This paper describes the development of a Spanish-Spanish Sign Language (LSE) translation system. Firstly, it describes the first Spanish-Spanish Sign Language (LSE) parallel corpus focused on two specific domains: the renewal of the Identity Document and Driver's License. This corpus includes more than 4,000 Spanish sentences (in these domains), their LSE translation and a video for each LSE sentence with the sign language representation. This corpus also contains more than 700 sign descriptions in several sign-writing specifications. The translation system developed with this corpus consists of two modules: a Spanish into LSE translation module that is composed of a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (for converting a word sequence into a sequence of signs) and a 3D avatar animation module (for playing back the signs). The second module is a Spanish generator from LSE made up of a visual interface (for specifying a sequence of signs in sign-writing), a language translator (for generating the sequence of words in Spanish) and a text to speech converter. For each language translation, the system uses three technologies: an example-based strategy, a rule-based translation method and a statistical translator.

1. Introduction

In Spain, 92% of the Deaf have a lot of difficulties in understanding and expressing themselves in written Spanish and around 47% of the Deaf, older than 10, do not have basic level studies (information from INE –Spanish Statistics Institute- and MEC –Ministry of Education-). The main problems are related to verb conjugations, gender/number concordances and abstract concepts.

In 2007, Spanish Sign Language was accepted by the Spanish Government as one of the official languages in Spain, and it was defined a plan to invest in new resources in this language. One important problem is that LSE is not disseminated enough among people who can hear. This is why there are communication barriers between deaf and hearing people. These barriers are even more problematic when they appear between a deaf person and a government employee who is providing a personal service, since they can cause the Deaf to have fewer opportunities or rights. This happens, for example, when people want to renew the Identity Document or the Driver's License (DL). Generally, a lot of government employees do not know LSE, so a deaf person needs an interpreter for accessing to these services. Thanks to associations like the Fundación CNSE, LSE is becoming not only the natural language for the Deaf to communicate with, but also a powerful instrument when communicating to people who can hear, or accessing information.

2. State of the Art

The research into sign language has been possible thanks to corpora generated by several groups. Some examples are: a corpus composed of more than 300 hours from 100 speakers in Australian Sign Language (Johnston T., 2008). The RWTH-BOSTON-400 Database that contains 843 sentences with about 400 different signs from 5 speakers in American Sign Language with English annotations (Dreuw et al., 2008a). The British Sign Language Corpus Project tries to create a machine-readable digital corpus of spontaneous and elicited British Sign Language (BSL) collected from deaf native signers and early learners across the United Kingdom (Schembri, 2008). And a corpus developed at Institute for Language and Speech Processing (ILSP) and that contains parts of free signing narration, as well as a considerable amount of grouped signed phrases and sentence level utterances (Efthimiou E., and Fotinea, E., 2008).

In recent years, several groups have developed prototypes for translating Spoken language into Sign Languages: example-based (Morrisey and Way, 2005), rule-based (San-Segundo et al 2008), full sentence (Cox et al, 2002) or statistical approaches (Bungeroth and Ney, 2004; Morrissey et al, 2007; SiSi system) approaches. About speech generation from sign language, in the Computer Science department of the RWTH, Aachen University, P. Dreuw supervised by H. Ney is making a significant effort into recognizing continuous sign language from video processing (Dreuw et al, 2008b; Dreuw, 2009). The results obtained are very promising.

This paper describes the parallel corpus obtained for developing a Spanish-Spanish Sign Language (LSE) translation system in two specific application domains: the renewal of the Identity Document and Driver's License.

3. Spanish-LSE parallel corpus

The corpus developed in this project has been obtained with the collaboration of Local Government Offices where the mentioned services (the renewal of the Identity Document (ID) and Driver's License (DL)) are provided. The most frequent explanations (from government employees) and the most frequent questions (from the user) were taken down over a period of three weeks and more than 5,000 sentences were noted and analysed.

Not all the sentences refer to ID or DL renewal (Government Offices provide more services), so sentences had to be selected manually. Finally, 1360 sentences were collected: 1,023 pronounced by government employees and 337 by users. These sentences were translated into LSE, both in text (sequence of glosses) and in video, and compiled in an excel file. This corpus was increased to 4,080 by incorporating different variants for Spanish sentences, maintaining the LSE translation. The main features of the corpus are summarised in Table 1. These features are divided depending on the domain (ID or DL renewal) and whether the sentence was spoken by the government employee or the user.

	ID		DL	
Government employee	**Spanish**	**LSE**	**Spanish**	**LSE**
Sentence pairs	1,425		1,641	
Different sentences	1,236	389	1,413	199
Running words	8,490	6,282	17,113	12,741
Vocabulary	652	364	527	237
User	**Spanish**	**LSE**	**Spanish**	**LSE**
Sentence pairs	531		483	
Different sentences	458	139	389	93
Running words	2,768	1,950	3,130	2,283
Vocabulary	422	165	294	133

Table 1: Main statistics of the corpus

All signs were written in the parallel corpus using glosses (capitalised words with a semantic relationship to sign language). In order to consider other sign-writing notations, a database with 715 signs (including all signs in the parallel corpus) was generated. This database includes sign descriptions in glosses, SEA (Sistema de Escritura Alfabética) (Herrero, 2004), HamNoSys (Prillwitz et al, 1989), and SIGML (Zwiterslood et al, 2004). Also, the database includes signs for all of the letters (necessary for word spelling), numbers from 0 to 100, numbers for hour specification, months, week days (Figure 1).

GLOSA	SEA	HAMNOSYS	SIGML
ABAJO	olemuawu	[illegible]	SIGML\ABAJO.txt
ACOMPAÑAR-A_MI	saca íájwe-ye	[illegible]	SIGML\ACOMPAÑAR-A_MI.txt
ACTUAL	s omèawud	[illegible]	SIGML\ACTUAL.txt
ADIÓS	omaudahb	[illegible]	SIGML\ADIÓS.txt
ADJUNTAR	sòaméha òamèug	[illegible]	SIGML\ADJUNTAR.txt
AHÍ	elewe	[illegible]	SIGML\AHÍ.txt

Figure 1: Example of sign database

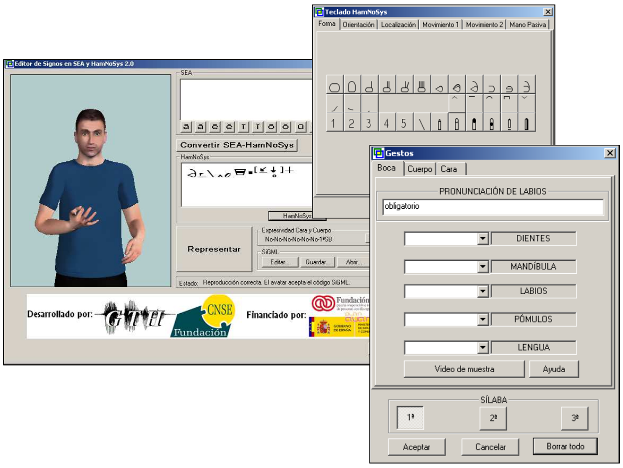

Figure 2: New version of the visual sign editor

The sign database has been generated using a new version of the eSIGN Editor. This new editor has three windows. In the main window, the eSign avatar shows the sign that is being designed at this moment (using a SEA or a HamNoSys specification). The second window allows HamNoSys characters to be introduced, and the last one permits non-manual gestures to be added. The SEA characters can be introduced using the PC keyboard together with auxiliary buttons (Figure 2).

This new version incorporates a Spanish graph to phoneme tool and a SEA-HamNoSys converter. On the one hand, the Spanish graph to phoneme tool is a rule-based converter that generates a sequence of SAMPA (Speech Assessment Method Phonetic Alphabet) phonemes given a Spanish sentence. This sequence is necessary to make the avatar move the lips according to this pronunciation. On the other hand, SEA (Sistema de Escritura Alfabética) (Herrero, 2004) is an alphabet (based on ASCII characters) for sign-language. Like HamNoSys, SEA allows a sequence of characters to be specified that describe aspects such as hand-shape, hand-position, location and movements. The reason for developing this converter is that the first normative dictionary for LSE (http://www.fundacioncnse.org/tesorolse developed at Fundación CNSE:) has SEA descriptions for more than 4,000 signs, but the eSign avatar needs HamNoSys descriptions for sign representation (previously converted into SiGML files).

The SEA-HamNoSys converter has been implemented in three steps: SEA characteristic detection, SEA-HamNoSys conversion for individual characteristics, and the generation of HamNoSys sign descriptions. These steps are repeated for all syllables that make up the sign, if there is more than one. The SEA-HamNoSys converter has been evaluated with 100 signs selected for including all the main SEA characteristics and the HamNoSys structure generated is useful and syntactically correct. These results are due to the fact that these two sign-writing notations have different specification levels. SEA presents a higher level because it has been designed to be easy to learn. On the other hand, HamNoSys allows a very detailed level of sign design. Because of this, when converting from SEA to HamNoSys, it is sometimes necessary to incorporate additional information by making some assumptions that are not always correct.

For designing a sign, it is necessary to specify hand movements (manual part) and other gestures including face, head and body movements (non-manual part). For designing the manual part, two processes have been followed: if the sign was included in the normative dictionary from Fundación CNSE, its SEA description has been automatically converted into HamNoSys (and lightly modified if necessary). On the other hand, if the sign was not in the dictionary, the HamNoSys sign specification has to be generated from scratch, using the videos recorded by the Fundación CNSE as the reference. Most of the signs (around 70%) were included in the dictionary, so the SEA-HamNoSys conversion tool has been very useful: the design time was reduced significantly, by approximately

50%. For the non-manual part of the sign, the design was always made from scratch, using the tools provided in the Visual Editor.

4. Spanish into LSE translation

The Spanish into LSE translation module is composed of three modules (Figure 3).

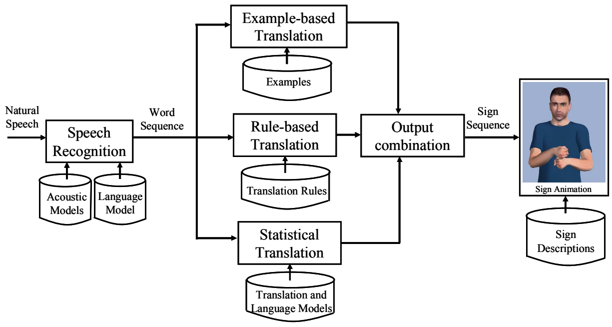

Figure 3: Spanish into LSE translation module

The first one is a speech recognition module that converts natural speech into a sequence of words (text). The second one is a natural language translation module that converts a word sequence into a sign sequence. For each translation, three different strategies are combined at the output step. The first one consists of an example-based strategy: the translation process is carried out based on the similarity between the sentence to be translated and the items of a parallel corpus with translated examples. Secondly, a rule-based translation strategy, where a set of translation rules (defined by an expert) guides the translation process. The last one is based on a statistical translation approach where parallel corpora are used for training language and translation models. We have considered two statistical alternatives: phrase-based one and Finite State Transducers (FST). Table 2, summarizes the results for rule-based and statistical approaches in laboratory tests: SER (Sign Error Rate), PER (Position Independent SER) and BLEU (BiLingual Evaluation Understudy).

		SER	PER	BLEU
Statistical approach	**Phrase-based**	39,01	37.05	0.5612
	FST-based	34.46	33.29	0.6433
Rule-based approach		21.45	17.24	0.6823

Table 1. Result summary for rule-based and statistical approaches

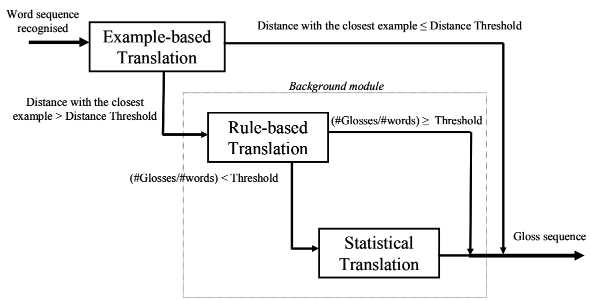

Figure 4: Diagram of natural language translation module combining three different translation strategies

The translation module has a hierarchical structure (Figure 4). Firstly, an example-based strategy is used to translate the word sequence. If the distance with the closest example is lower than a threshold, the translation output is the same than the example. But if the distance is higher, a background module translates the word sequence, using a combination of rule-based and statistical translators. The last module represents the signs with VGuido (the eSIGN 3D avatar). It is important to remark that this system translate Spanish into LSE, not into Signed Spanish.

5. Spanish generator from LSE

The spoken Spanish generation system is composed of three modules (Figure 5).

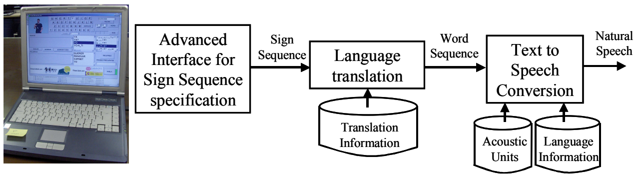

Figure 5: Diagram of Spanish generation system

The first module is an interface for specifying a sign sequence. This interface includes several tools for sign specification: avatar for sign representation (to verify that sign corresponds to the gloss), prediction mechanisms, calendar and clock for date or time definitions, etc. With this visual interface the Deaf can build a sign sentence that will be translated into Spanish and spoken to a hearing person. The sign sequence is specified in glosses but signs can be searched by using specific sign characteristics in HamNoSys notation. (Figure 6)

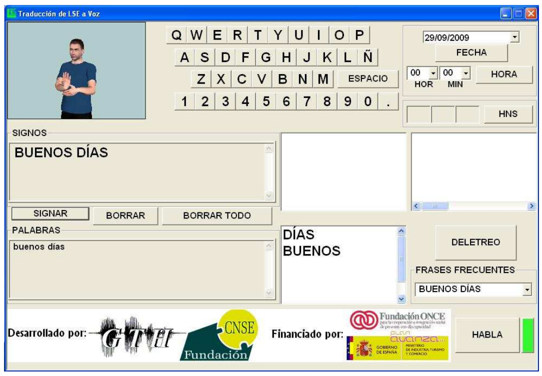

Figure 6: Visual interface for sign sequence specification

The second module converts a sign sequence into a word sequence with three different strategies combined: an example-based, a rule-based and a statistical translation strategy. The procedure is the same as in the Spanish into LSE translation system. The last module converts the word sequence into spoken Spanish by using a commercial Text to Speech converter. In this project the Loquendo system has been used (http://www.loquendo.com/en/).

6. Evaluation

An evaluation has been performed for testing the speech into LSE translator and the spoken Spanish generator for Driver's License renewal. The speech-LSE system translates the government employee's explanations and the spoken Spanish generator helps Deaf to ask questions.

The evaluation was carried out over two days. On the first day, a one-hour talk, about the project and the evaluation, was given to government employees (2 people) and users (10 people) involved in the evaluation. Six different scenarios were defined in order to specify real situations. The sequence of scenarios was randomly selected for each user. Ten deaf users interacted with two government employees at the Toledo Traffic Office using the developed system. These ten users (six males and four females) tested the system in almost all the scenarios described previously: 48 dialogues were recorded (12 dialogues were missing because several users had to leave the evaluation session before finishing all the scenarios). The user ages ranged between 22 and 55 years (40.9 average). For both systems the translation accuracy was very high (> 90%) but the users reported several problems related to avatar naturalness and LSE normalization.

7. Conclusion

This paper has described the first Spanish-LSE parallel corpus for language processing research focusing on specific domains: the renewal of the Identity Document and Driver's License. This corpus includes 4,080 Spanish sentences translated into LSE. This corpus also contains a sign database including all sign descriptions in several sign-writing specifications: Glosses, HamNoSys and SEA: Sistema de Escritura Alfabética. This sign database includes all signs in the parallel corpus and signs for all the letters (necessary for word spelling), numbers from 0 to 100, numbers for time specification, months and week days. The sign database has been generated using a new version of the eSign Editor.

This paper also has described the design and development of a Spanish into Spanish Sign Language (LSE: Lengua de Signos Española) translation system. This system is made up of a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (for converting a word sequence into a sequence of signs belonging to the sign language), and a 3D avatar animation module (for playing back the signs). For the natural language translator, three technological proposals have been evaluated and combined in a hierarchical structure: an example-based strategy, a rule-based translation method and a statistical translator.

Finally, this paper has presented a spoken Spanish generator from sign-writing of Spanish Sign Language (LSE: Lengua de Signos Española). This system consists of an advanced visual interface where a deaf person can specify a sequence of signs in sign-language, a language translator (for generating the sequence of words in Spanish), and finally, a text to speech converter.

8. Acknowledgements

The authors want to thank the eSIGN (Essential Sign Language Information on Government Networks) consortium for permitting the use of the eSIGN Editor and the 3D avatar in this research work. This work has been supported by Plan Avanza Exp Nº: PAV-070000-2007-567, ROBONAUTA (MEC ref: DPI2007-66846-c02-02) and SD-TEAM (MEC ref: TIN2008-06856-C05-03) projects.

9. References

Bungeroth J., Ney, H.,: Statistical Sign Language Translation. In Workshop on Representation and Processing of Sign Languages, LREC 2004, 105-108.

Cox, S.J., Lincoln M., Tryggvason J., Nakisa M., Wells M., Mand Tutt, and Abbott, S., 2002 "TESSA, a system to aid communication with deaf people". In ASSETS 2002, pages 205-212, Edinburgh, Scotland, 2002

Dreuw P., Neidle C., Athitsos V., Sclaroff S., and Ney H. 2008a. "Benchmark Databases for Video-Based Automatic Sign Language Recognition". In International Conference on Language Resources and Evaluation (LREC), Marrakech, Morocco, May 2008.

Dreuw, P., D. Stein, T. Deselaers, D. Rybach, M. Zahedi, J. Bungeroth, and H. Ney. 2008b "Spoken Language Processing Techniques for Sign Language Recognition and Translation. Journal Technology and Dissability. Volume 20 Pages 121-133. ISSN 1055-4181.

Dreuw, P., Stein D., and Ney H. 2009. "Enhancing a Sign Language Translation System with Vision-Based Features". LNAI, number 5085, pages 108-113, Lisbon, Portugal, January 2009.

Efthimiou E., and Fotinea, E., 2008 "GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI" LREC 2008.

Johnston T., 2008. "Corpus linguistics and signed languages: no lemmata, no corpus". 3rd Workshop on the Representation and Processing of Sign Languages, June 1. 2008.

Herrero, Ángel. 2004 "Escritura alfabética de la Lengua de Signos Española" Universidad de Alicante.

Morrissey S., and Way A., 2005. "An example-based approach to translating sign language". In Workshop Example-Based Machine Translation (MT X–05), pages109–116, Phuket, Thailand, September.

Morrissey S., Way A., Stein D., Bungeroth J., and Ney H., 2007 "Towards a Hybrid Data-Driven MT System for Sign Languages. Machine Translation Summit (MT Summit)", pages 329-335, Copenhagen, Denmark, September 2007.

Prillwitz, S., R. Leven, H. Zienert, T. Hanke, J. Henning, et-al. 1989. "Hamburg Notation System for Sign Languages – An introductory Guide". International Studies on Sign Language and the Communication of the Deaf, Volume 5. University of Hamburg, 1989.

San-Segundo R., Barra R., Córdoba R., D'Haro L.F., Fernández F., Ferreiros J., Lucas J.M., Macías-Guarasa J., Montero J.M., Pardo J.M, 2008. "Speech to Sign Language translation system for Spanish". Speech Communication, Vol 50. 1009-1020. 2008.

Schembri. A.,2008 "British Sign Language Corpus Project: Open Access Archives and the Observer's Paradox". Deafness Cognition and Language Research Centre, University College London. LREC 2008.

Zwiterslood, I., Verlinden, M., Ros, J., van der Schoot, S., 2004. "Synthetic Signing for the Deaf: eSIGN". Workshop on Assistive Technologies for Vision and Hearing Impairment, CVHI 2004, 29 June-2 July 2004, Granada, Spain.