Source Language Categorization for improving a Speech into Sign Language Translation System

V. López-Ludeña, R. San-Segundo, S. Lutfi, J.M. Lucas-Cuesta, J.D. Echevarry, B. Martínez-González

Grupo de Tecnología del Habla
Universidad Politécnica de Madrid
{veronicalopez|lapiz|syaheerah|juanmak|jdec|beatrizmartinez}@die.upm.es

Abstract

This paper describes a categorization module for improving the performance of a Spanish into Spanish Sign Language (LSE) translation system. This categorization module replaces Spanish words with associated tags. When implementing this module, several alternatives for dealing with non-relevant words have been studied. Non-relevant words are Spanish words not relevant in the translation process. The categorization module has been incorporated into a phrase-based system and a Statistical Finite State Transducer (SFST). The evaluation results reveal that the BLEU has increased from 69.11% to 78.79% for the phrase-based system and from 69.84% to 75.59% for the SFST.

Keywords: Source language categorization, Speech into Sign Language translation. Lengua de Signos Española (LSE).

1   Introduction

In the world, there are around 70 million people with hearing deficiencies (information from World Federation of the Deaf http://www.wfdeaf.org/). Deafness brings about significant communication problems: most deaf people are unable to use written languages, having serious problems when expressing themselves in these languages or understanding written texts. They have problems with verb tenses, concordances of gender and number, etc., and they have difficulties when creating a mental image of abstract concepts. This fact can cause deaf people to have problems when accessing information, education, job, social relationship, culture, etc. According to information from INE (Statistic Spanish Institute), in Spain, there are 1,064,000 deaf people. 47% of deaf population do not have basic studies or are illiterate, and only between 1% and 3% have finished their studies (as opposed to 21% of Spanish hearing people). Another examples are the the figures from the National Deaf Children's Society (NDCS), Cymru, reveal for the first time a shocking attainment gap between deaf and hearing pupils in Wales. In 2008, deaf pupils were 30% less likely than hearing pupils to gain five A*-C grades at General Certificate of Secondary Education (GCSE), while at key stage 3 only 42% of deaf pupils achieved the core subject indicators, compared to 71% of their hearing counterparts. Another example is a study carried out in Ireland in 2006, of 330 respondents "38% said they did not feel confidence to read a newspaper and more than half were not fully confidence in writing a letter or filling out a form" (Conroy, 2006).

Deaf people use a sign language (their mother tongue) for communicating and there are not enough sign-language interpreters and communication systems. In Spain, there is the Spanish Sign Language (Lengua de Signos Española LSE) as the official sign language. In the USA, there are 650,000 Deaf people (who use a sign language), although there are more people with hearing deficiencies, but only 7,000 sign-language interpreters, i.e. a ratio of 93 deaf people to 1 interpreter. In Finland we can find the best ratio, 6 to 1, and in Slovakia the worst with 3,000 users to 1 interpreter (Wheatley and Pabsch, 2010). In Spain this ratio is 221 to 1. This information shows the need to develop automatic translation systems with new technologies for helping hearing and Deaf people to communicate between themselves.
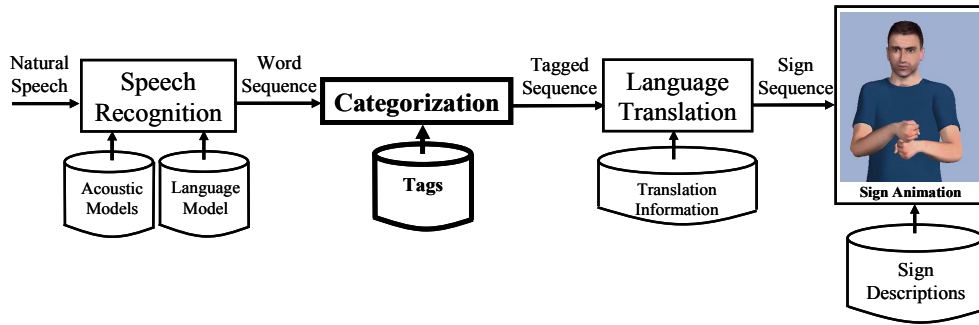
Figure 1. Spanish into LSE translation system.

It is necessary to make a difference between "deaf" and "Deaf": the first one refers to non-hearing people, and the second one refers to hearing and non-hearing people who use a sign language to communicate between them, being part of the "Deaf community". Each country has a different sign language, but there may even be different sign languages in different regions in the.

This paper describes a categorization module for improving the performance of a Speech into Sign Language Translation System. This system helps Deaf people to communicate with government employees in a restricted domain: the renewal of Identity Documents and Driver's License (San-Segundo et al., 2008). This system has been designed to translate the government employee's explanations into LSE when government employees provide these face-to-face services. The system is made up of a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (a phrase-based system 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) (Figure 1). This paper proposes to include a fourth module named "categorization" between the speech recognition and language translation modules (Figure 1). This categorization module replaces Spanish words with associated tags as it will be shown further.

For the natural language translation module, two different statistical strategies have been analyzed: a phrase-based system (Moses) and a Statistical Finite State Transducer (SFST). The proposed categorization module has been incorporated into and evaluated with both translation strategies.

This paper is organized as follows: section 2 describes the state of the art. Section 3 describes the parallel corpus used in these experiments. The main characteristics of the LSE are presented in section 4. Section 5 details the two main translation strategies considered. The categorization module is described in section 6. Section 7 includes the main experiments and the obtained results, and finally, sections 8 and 9 include the main conclusions and the future work.

2 State of the art

In recent years, several groups have developed prototypes for translating Spoken language into Sign Language: example-based (Morrissey, 2008), rule-based (Marshall and Sáfár, 2005; San-Segundo et al. 2008), full sentence (Cox et al., 2002) or statistical approaches (Stein et al., 2006; Morrissey et al., 2007; Vendrame et al., 2010) approaches.

Given the sparseness of data for researching in Sign Languages, in the last five years, several projects have started to generate more resources: in American Sign Language (Dreuw et al.., 2008), British Sign Language (Schembri, 2008), Greek Sign Language (Efthimiou and Fotinea, 2008), in Irish Sign Language (Morrissey et al., 2010), NGS (German Sign Language) (Hanke et al., 2010), and Italian Sign Language (Geraci et al., 2010). For LSE, the biggest database was generated two years ago in a Plan Avanza project (www.traduccionvozlse.es) (San-Segundo et al., 2010) and it is has been used in this work. Not only the data but also new practice (Forster et al., 2010) and new uses of traditional annotation tools (Crasborn et al., 2010) have been developed.

The work presented in this paper describes experiments with a relevant database, despite the small amount of data available for research into sign languages. The system presented in this paper demonstrates a very good performance compared to similar systems previously developed. The presented results are also the best results for translating Spanish into LSE using the biggest database that includes these languages.

In Europe, the two main research projects involving sign languages are DICTA-SIGN (Hanke et al., 2010; Efthimiou et al., 2010) and SIGN-SPEAK (Dreuw et al., 2010a and 2010b), both financed by The European Commission within the Seventh Frame Program. DICTA-SIGN (http://www.dictasign.eu/) aims to develop the technologies necessary to make Web 2.0 interactions in sign language possible: users sign to a webcam using a dictation style. The computer recognizes the signed phrases, converts them into an internal representation of sign language, and then it has an animated avatar that signs them back to the users. In SIGN-SPEAK (http://www.signspeak.eu/), the overall goal is to develop a new vision-based technology for recognizing and translating continuous sign language into text.

3 Parallel corpus

This section describes the first Spanish-LSE parallel corpus developed for language processing in two specific domains: the renewal of the Identity Document (ID) and Driver's License (DL). This corpus has been obtained with the collaboration of Local Government Offices where these services are provided. Over several weeks, the most frequent explanations (from the government employees) and the most frequent questions (from the user) were taken down. In this period, more than 5,000 sentences were noted and analyzed.

Not all the sentences refer to ID or DL renewal (Government Offices provide more services), so sentences had to be selected manually. This was possible because every sentence was tagged with the information about the service being provided when it was collected. Finally, 1360 sentences were collected: 1,023 pronounced by government employees and 337 by users. These sentences have been translated into LSE, both in text (sequence of glosses) and in video (containing replayed sentences by native LSE signers), and compiled in an excel file. Videos are not used in this study but they were collected for generating a complete parallel corpus.

This corpus was increased to 4,080 by incorporating different variants for Spanish sentences (maintaining the LSE translation) (San-Segundo et al. 2010). Table 1 summarizes the main features of this database.

	Spanish	LSE
Sentence pairs	4,080	
Different sentences	3,342	1,289
Words/signs per sentence	7.7	5.7
Running words	31,501	23,256
Vocabulary	1,232	636

Table 1. Main statistics of the corpus

For the experiments presented in this paper, this database has been divided randomly into three sets: training (75%), development (12.5%) and test (12.5%). The training set was used for tuning the speech recognizer (vocabulary and language model) and training the translation models. The development set was used for tuning the translation systems and finally, the test set was used for evaluating the categorization module.

4 Spanish Sign Language (LSE)

Spanish Sign Language (LSE), just like other sign languages, has a visual-gestural channel, but it also has grammatical characteristics similar to spoken languages. Sign languages have complex grammars and professional linguists have found all of the necessarily linguistic characteristics for classifying sign languages as "true languages". In linguistic terms, sign languages are as complex as spoken languages, despite the common misconception that they are a "simplification" of spoken languages. For example, The United Kingdom and USA share the same language. However, British Sign Language is completely different from American Sign Language. W. Stokoe supports the idea that sign languages have four dimensions (three space dimensions plus time), and spoken languages have only one dimension, time, so it cannot say that sign languages are a simplification of any other language.

One important difference between spoken languages and sign languages is sequentially. Phonemes in spoken languages are produced in a sequence. On the other hand, sign languages have a large non-sequential component, because fingers, hands and face movements can be involved in a sign simultaneously, even two hands moving in different directions. These features give a complexity to sign languages that traditional spoken languages do not have. This fact makes it very difficult to write sign languages. Traditionally, signs have been written using words (in capital letters) in Spanish (or English in the case of BSL, British Sign Language) with a similar meaning to the sign meaning. They are called glosses (i.e. 'CAR' for the sign 'car').

In the last 20 years, several alternatives, based on specific characteristics of the signs, have appeared in the international community: HamNoSys (Prillwitz et al, 1989), SEA (Sistema de Escritura Alfabética) (Herrero, A., 2004) and SignWriting (http://www.signwriting.org/). HamNoSys and SignWriting require defining a specific picture font to be used by computers. SignWriting includes face features in the notation system but HamNoSys and SEA do not include them. All of these alternatives are flexible enough for dealing with different sign languages including LSE. However, in this work, glosses have been considered for writing signs because it is the most familiar and extended alternative according to the Spanish Deaf Association. These glosses include non-speech indicators (i.e. PAY or PAY? if the sign is localized at the end of an interrogative sentence) and finger spelling indicators (i.e. DL-PETER that must be represented letter by letter P-E-T-E-R).

LSE has some characteristics that differ from Spanish. One important difference is the order of arguments in sentences: LSE has a **SOV** (subject-object-verb) order in contrast to **SVO** (subject-verb-object) Spanish order. An example that illustrates this behaviour is shown below:

Spanish: Juan ha comprado las entradas (Juan has bought the tickets)
LSE: JUAN ENTRADAS COMPRAR (JUAN TICKETS TO-BUY)

There are other typological differences that are not related to predication order:

- Gender is not usually specified in LSE, in contrast to Spanish.
- In LSE, there can be concordances between verbs and subject, receiver or object and even subject and receiver, but in Spanish there can be only concordance between verb and subject:
  - Spanish: Te explica (*he explains to you*)
  - LSE: EXPLICAR-él-a-ti (*EXPLAIN-HIM-TO-YOU*)
- The use of classifiers is common in LSE, but they are not in Spanish.
  - Spanish: debe acercarse a la cámara (*you must approach the camera*)
  - LSE: FOTO CLD_GRANDE_NO CLL_ACERCARSE DEBER (*PHOTO CLD_BIG_NO CLL_APPROACH MUST*)
- Articles are used in Spanish, but not in LSE.
- Plural can be descriptive in LSE, but not in Spanish.
- In Spanish, there is a copula in non-verbal predications (the verb 'to be', *ser* and *estar* in Spanish), but there is not in LSE.
- There are Spanish impersonal sentences, but not in LSE.
- LSE is more lexically flexible than Spanish, and it is perfect for generating periphrasis through its descriptive nature and because of this, LSE has fewer nouns than Spanish. (i.e. mud is translated into SAND+WATER)
- To finish, LSE has less glosses per sentence (5.7 in our database) than Spanish (7.7 in our database).
- LSE has smaller vocabulary variability. LSE has a vocabulary of around 10,000 signs while Spanish has several millions of different words. Good examples are the different verb conjugations.

5 Statistical translation strategies

In this paper, two different statistical strategies have been considered: a phrase-based system and a Statistical Finite State Transducer. The proposed automatic categorization has been evaluated with both translation strategies. This section describes the architectures used for the experiments.

5.1 Phrase-based translation system

The Phrase-based translation system is based on the software released at the 2009 NAACL

Workshop on Statistical Machine Translation (http://www.statmt.org/wmt09/)(Figure 2).

Target Language: Sign Language
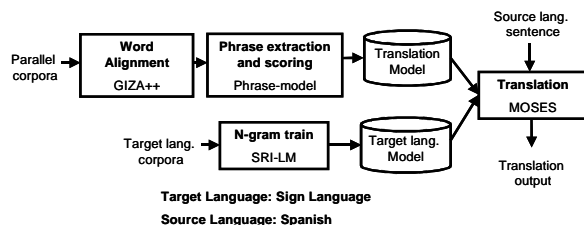**Source Language: Spanish**

Figure 2. Phrase-based translation architecture.

In this study, a phrase consists of a subsequence of words (in a sentence) that intends to have a meaning. Every sentence is split in several phrases automatically so this segmentation can have errors. But, the main target, when training a phrase-based model, is to split the sentence in several phrases and to find their corresponding translations in the target language.

The phrase model has been trained starting from a word alignment computed using GIZA++ (Och and Ney, 2003). GIZA++ is a statistical machine translation toolkit that is used to train IBM Models 1-5 and an HMM word alignment model. In this step, the alignments between words and signs in both directions (Spanish-LSE and LSE-Spanish) are calculated. The "alignment" parameter has been fixed to "target-source" as the best option (based on experiments over the development set): only this target-source alignment was considered (LSE-Spanish). In this configuration, alignment is guided by signs: this means that in every sentence pair alignment, each word can be aligned to one or several signs (but not the opposite), and also, it is possible that some words were not aligned to any sign. When combining the alignment points from all sentences pairs in the training set, it is possible to have all possible alignments: several words aligned to several signs.

After the word alignment, the system performs a phrase extraction process (Koehn et al. 2003) where all phrase pairs that are consistent with the word alignment (target-source alignment in our case) are collected. The phrase extraction

The maximum phrase length has been fixed at 7 consecutive words, based on development experiments over the development set (see previous section).

Finally, the last step is phrase scoring. In this step, the translation probabilities are computed for all phrase pairs. Both translation probabilities are calculated: forward and backward.

For the translation process, the Moses decoder has been used (Koehn, 2010). This program is a beam search decoder for phrase-based statistical machine translation models. In order to obtain a 3-gram language model, the SRI language modeling toolkit has been used (Stolcke, 2002).

5.2 Phrase-based translation system

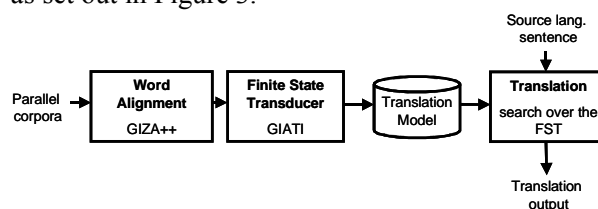The translation based on SFST is carried out as set out in Figure 3.

Figure 3. Diagram of the FST-based translation module.

The translation model consists of an SFST made up of aggregations: subsequences of aligned source and target words. The SFST is inferred from the word alignment (obtained with GIZA++) using the GIATI (Grammatical Inference and Alignments for Transducer Inference) algorithm (Casacuberta and Vidal, 2004). The SFST probabilities are also trained from aligned corpora. The software used in this paper has been downloaded from http://prhlt.iti.es/content.php?page=software.php.

6 Categorization module

As it was presented in Figure 1, the categorization module proposed in this paper analyzes the source language sentence (sentence in Spanish) and replaces Spanish words with their associated tags. This module uses a list of 1014 Spanish words (the vocabulary in this restricted domain) and the corresponding tags. For every word, only one syntactic-semantic tag is associated. In the case of homonyms, the most frequent meaning has been considered for defining the syntactic-semantic tag. Figure 4 shows an extract of the word-tag list. This list is composed of Spanish words and their corresponding tags, including the English translation in parenthesis.

```
word TAG (word and tag in English)
…
cerrado CERRAR-YA (closed CLOSE-ALREADY )
cerramos CERRAR (we close CLOSE )
cerrar CERRAR (to close CLOSE)
cobradas COBRAR-YA (charged CHARGE-ALREADY)
cobro COBRAR (I charge CHARGE)
coge COGER (you get GET)
cogido COGER-YA (gotten GET-ALREADY)
coja COGER (you get GET)
…
```

Figure 4. Extract of the word-tag list.

The categorization module executes a simple procedure: for all words in a Spanish sentence, the categorization module looks for this word in the list and replaces it with the associated tag. It is important to comment two main aspects. The first one is that there is a tag named "non-relevant" associated to those words that are not useful for translating the sentence. The second one is that if the Spanish word is not in the list (it is an Out Of Vocabulary word: OOV), this word is not replaced with any tag: this word is kept as it is.

In order to train the statistical translation modules when using the categorization module, it is necessary to retrain the translation models considering the tagged source language, not the original word sentences, and using the training set. This way, the translation models learn the relationships between tags and signs.

The main issue for implementing the categorization module is to generate the list of the Spanish words with the associated tags. In this work, the categorization module considers the categories used in the rule-based translation system previously developed for this application domain (San-Segundo et al., 2008). These categories were generated manually during one week, approximately. In this case, the natural language translation module was implemented using a rule-based technique considering a bottom-up strategy. The translation process is carried out in two steps. In the first one, every word is mapped into one syntactic-pragmatic tag. After that, the translation module applies different rules that convert the tagged words into signs by means of grouping concepts or signs and defining new signs. These rules can define short and large scope relationships between the concepts or signs.

When implementing the categorization module, several strategies for dealing with the "non-relevant" words have been proposed:

- In the first alternative, all the words are replaced by their tags with the exception of those words that they do not appear in the list (OOV words). As, it was commented before, they are kept as they are. In the word-tag list, there is a "non-relevant" tag mapped to words that are not relevant for the translation process (named "basura" (non-relevant)). This alternative will be referred in the experiments like "**Base categorization**". For example:
  - Source sentence: debes pagar las tasas en la caja (*you must pay the taxes in the cash desk*)
  - Categorizated source sentence: DEBER PAGAR basura DINERO basura basura DINERO-CAJA (*MUST PAY non-relevant MONEY non-relevant non-relevant CASH-DESK*)
  - Target sentence: VENTANILLA ESPECÍFICO CAJA TU PAGAR (*WINDOW SPECIFIC CASH-DESK YOU PAY*)

- The second proposed alternative was not to tag any word in the source language but removing non-relevant words from the source lexicon (associated to the "non-relevant" tag). This alternative will be referred in the experiments like "**Non-relevant word deletion**". For example:
  - Source sentence: debes pagar las tasas en la caja (*you must pay the taxes in the cash desk*)
  - Categorizated source sentence: debes pagar tasas caja
  - Target sentence: VENTANILLA ESPECÍFICO CAJA TU PAGAR (*WINDOW SPECIFIC CASH-DESK YOU PAY*)

- Finally, the third alternative proposes to replace words with tags (with the exception of OOVs) and to remove "non-relevant" tags. This alternative will be referred in the experiments like "**Categorization and non-relevant word deletion**". For example:
  - Source sentence: debes pagar las tasas en la caja (*you must pay the taxes in the cash desk*)
  - Categorizated source sentence: debes|DEBER pagar|PAGAR tasas|DINERO caja|DINERO-CAJA
  - Target sentence: VENTANILLA ESPECÍFICO CAJA TU PAGAR (WINDOW SPECIFIC CASH-DESK YOU PAY)

In the next section, all the alternatives will be evaluated and discussed.

7   Experiments and discussion

For the experiments, the corpus (described in section 3) was divided randomly into three sets: training (75%), development (12.5%) and test (12.5%). Results are compared with a baseline. This baseline consists of training models with original source and target corpus without any type of factorization, i.e, sentences contain words and signs from the original database. For example: this sentence "debes pagar las tasas en la caja" (*you must pay the taxes in the cash desk*) is translated into "VENTANILLA ESPECÍFICO CAJA TU PAGAR" (*WINDOW SPECIFIC CASH-DESK YOU PAY*).

For evaluating the performance of the translation systems, the BLEU (BiLingual Evaluation Understudy) metric (Papineni et al., 2002) has been used. BLEU is one of the most well-known metric for evaluating automatic translation systems because this metric presents a good correlation with human evaluations. This metric has been also adopted to evaluate speech into sign language translation systems (Stein et al., 2006; Morrissey et al., 2007; Vendrame et al., 2010, San-Segundo et al. 2008). In order to analyze the significance of the differences between several systems, for every BLEU result, the confidence interval (at 95%) is also presented. This interval is calculated using the following formula:

$$\pm \Delta = 1{,}96 \sqrt{\frac{BLEU(100 - BLEU)}{n}} \qquad (1)$$

n is the number of signs used in evaluation, in this case n=2,906. An improvement between two systems is statistically significant when there is no overlap between the confidence intervals of both systems.

Related to the speech recognizer, it is important to comment that the Word Error Rate (WER) obtained in these experiments has been 4,7%.

Table 2 compares the baseline system and the system with the categorization module for translating the references (Reference) and the

speech recognizer outputs (ASR output) using the phrase-based translation system.

Phrase-based translation System		BLEU	±Δ
Baseline	Reference	73.66	1.60
	ASR output	**69.11**	**1.68**
Base categorization	Reference	81.91	1.40
	ASR output	**74.55**	**1.58**
Non-relevant words deletion	Reference	80.02	1.45
	ASR output	**73.89**	**1.60**
Categorization and non-relevant word deletion	Reference	84.37	1.32
	ASR output	**78.79**	**1.49**

Table 2. Evaluation results for the phrase-based translation system.

Table 3 compares the baseline system and the system with the categorization module for translating the references (Reference) and the speech recognizer outputs (ASR output) using the SFST-based translation system.

SFST		BLEU	±Δ
Baseline	Reference	71.17	1.65
	ASR output	**69.84**	**1.67**
Base categorization	Reference	71.86	1.63
	ASR output	**68.73**	**1.69**
Non-relevant words deletion	Reference	76.71	1.54
	ASR output	**72.77**	**1.62**
Categorization and non-relevant word deletion	Reference	81.48	1.41
	ASR output	**75.59**	**1.56**

Table 3. Evaluation results for the SFST-based translation system.

Comparing the three alternatives for dealing with the non-relevant words, it is shown that adding tags to the words and removing "non-relevant" words are complementary actions that allow reaching the best results.

In order to better understand the main causes of this improvement, an error analysis has been carried out, establishing a relationship between

these errors and the main differences between Spanish and LSE.

The most important type of error (35% of the cases) is related to the fact that in Spanish there are more words than signs in LSE (7.7 for Spanish and 5.7 for LSE in this corpus). This circumstance provokes different types of errors: generation of many phrases in the same output, producing a high number of insertions. When dealing with long sentences there is the risk that the translation model cannot deal properly with the big distortion. This produces important changes in order and sometimes the sentence is truncated producing several deletions.

The second most important source of errors (25% of the cases) is related to the fact that when translating Spanish into LSE, there is a relevant number of words in the testing set that they do not appear in the training set due to the higher variability presented in Spanish. These words are named Out Of Vocabulary words. For example, in Spanish there are many verb conjugations that are translated into the same sign sequence. So, when a new conjugation appears in the evaluation set, it is an OOV that provokes a translation error.

Other important source of errors corresponds to ordering errors provoked by the different order in predication: LSE has a SOV (Subject-Object-Verb) while Spanish SVO (Subject-Verb-Object). In this case, the frequency is close to 20%

Finally, there are others causes of errors like the wrong generation of the different classifiers needed in LSE and not presented in Spanish (11%) and the existence of some deletions when translating very specific names, even when they are in the training set. Some of these names (i.e. 'mud' is translated into SAND + WATER) need some periphrasis in LSE that not always are properly generated.

Based on this error analysis, the main causes of the translation errors are related to the different variability in the vocabulary for Spanish and LSE (much higher in Spanish), the different number for words or signs in the sentences (higher in Spanish) and the different predication order.

The categorization module allows reducing the variability in the source language (for example, several verb conjugations are tagged with the same tag) and also the number of tokens composing the input sentence (when removing non-relevant words). Also, reducing the source language

variability and the number of tokens provoke an important reduction on the number of source-target alignments the system has to train. When having a small corpus, as it is the case of many sign languages, this reduction of alignment points permits to obtain better training models with less data, improving the results. These aspects allow increasing the system performance. Presumably, if there were a very large corpus of Spanish-to-Spanish-Sign-Language available, the system could learn better translation models and the improvement reached with this categorization module would be lower.

The evaluation results reveal that the BLEU has increased from 69.11% to 78.79% for the phrase-based system and from 69.84% to 75.59% for the SFST.

8 Conclusions

This paper describes a categorization module for improving a Spanish into Spanish Sign Language Translation System. This module allows incorporating syntactic-semantic information during the translation process reducing the source language variability and the number of words composing the input sentence. These two aspects reduce the translation error rate considering two statistical translation systems: phrase-based and SFST-based translation systems. This system is used to translate government employee's explanations into LSE when providing a personal service for renewing the Identity Document and Driver's License.

9 Future work

The main issue for implementing the categorization module is to generate the list of the Spanish words with the associated tags. Generating this list manually is a subjective, slow and difficult task. Because of this, in the near future, authors will work on the possibility to define a procedure for calculating this list automatically.

References

Casacuberta F., E. Vidal. 2004. "Machine Translation with Inferred Stochastic Finite-State Transducers". Computational Linguistics, Vol. 30, No. 2, pp. 205-225, 2004.

Conroy, P. 2006. Signing in and Signing Out: The Education and Employment Experiences of Deaf Adults in Ireland. Research Report, Irish Deaf Society, Dublin. 2006.

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, Edinburgh, Scotland, pp 205-212, 2002.

Crasborn O., Sloetjes H. 2010. "Using ELAN for annotating sign language corpora in a team setting". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, 2010. pp 61-65

Dreuw P., Neidle C., Athitsos V., Sclaroff S., and Ney H. 2008. "Benchmark Databases for Video-Based Automatic Sign Language Recognition". In International Conference on Language Resources and Evaluation (LREC), Marrakech, Morocco, May 2008. pp 1115-1121.

Dreuw P., Ney H., Martinez G., Crasborn O., Piater J., Miguel Moya J., and Wheatley M., 2010 "The SignSpeak Project - Bridging the Gap Between Signers and Speakers". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, 2010a. pp 73-80.

Dreuw P., Forster J., Gweth Y., Stein D., Ney H., Martinez G., Verges Llahi J., Crasborn O., Ormel E., Du W., Hoyoux T., Piater J., Moya Lazaro JM, and Wheatley M. 2010 "SignSpeak - Understanding, Recognition, and Translation of Sign Languages". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010b. pp 65-73

Efthimiou E., and Fotinea, E., 2008 "GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI" LREC 2008. pp 1-10.

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.

Forster J., Stein D., Ormel E., Crasborn O., Ney H. 2010. "Best Practice for Sign Language Data Collections Regarding the Needs of Data-Driven Recognition and Translation". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 92-98.

Geraci C., Bayley R., Branchini C., Cardinaletti A., Cecchetto C., Donati C., Giudice S., Mereghetti E., Poletti F., Santoro M., Zucchi S. 2010. "Building a corpus for Italian Sign Language. Methodological issues and some preliminary results". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 98-102.

Hanke T., König L., Wagner S., Matthes S. 2010. "DGS Corpus & Dicta-Sign: The Hamburg Studio Setup". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 106-110.

Herrero, A., 2004 "Escritura alfabética de la Lengua de Signos Española" Universidad de Alicante. Servicio de Publicaciones.

Koehn P., F.J. Och D. Marcu. 2003. "Statistical Phrase-based translation". Human Language Technology Conference 2003 (HLT-NAACL 2003), Edmonton, Canada, 2003. pp. 127-133.

Koehn, Philipp. 2010. "Statistical Machine Translation". phD. Cambridge University Press.

Marshall, I., Sáfăr, E. (2005) "Grammar Development for Sign Language Avatar-Based Synthesis", In Proceedings HCII 2005, 11th International Conference on Human Computer Interaction (CD-ROM), Las Vegas, USA, July 2005. pp 1-10.

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)", Copenhagen, Denmark, 2007. pp 329-335.

Morrissey, S. 2008. "Data-Driven Machine Translation for Sign Languages". Thesis. Dublin City University, Dublin, Ireland.

Morrissey S., Somers H., Smith R., Gilchrist S., Dandapat S., 2010 "Building Sign Language Corpora for Use in Machine Translation". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 172-178.

Och J., Ney. H., 2003. "A systematic comparison of various alignment models". Computational Linguistics, Vol. 29, No. 1, 2003. pp. 19-51.

Papineni K., S. Roukos, T. Ward, W.J. Zhu. 2002 "BLEU: a method for automatic evaluation of machine translation". 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, PA. 2002. pp. 311-318.

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. Institute of German Sign Language and Communication of the Deaf, 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. 2008. pp. 1009-1020.

San-Segundo, R., Pardo, J.M., Ferreiros, F., Sama, V., Barra-Chicote, R., Lucas, JM., Sánchez, D., García. A., "Spoken Spanish Generation from Sign Language" Interacting with Computers, Vol. 22, No 2, 2010. pp. 123-139.

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. pp 1-5.

Stein, D., Bungeroth, J. and Ney, H.: 2006 "Morpho-Syntax Based Statistical Methods for Sign Language Translation". 11th Annual conference of the European Association for Machine Translation, Oslo, Norway, June 2006. pp 223-231.

Stolcke A., 2002. "SRILM – An Extensible Language Modelling Toolkit". Proc. Intl. Conf. on Spoken Language Processing, vol. 2, Denver USA, 2002. pp. 901-904,

Stokoe W., Sign Language structure: an outline of the visual communication systems of the American deaf, Studies in Linguistics, Buffalo University Paper 8, 1960.

Vendrame M., Tiotto G., 2010. ATLAS Project: Forecast in Italian Sign Language and Annotation of Corpora. In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT 2010), Valletta, Malta, May 2010. pp 239-243.

Wheatley, M., Annika Pabsch, 2010. "Sign Language in Europe". In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies. LREC. Malta 2010. pp 251-255.