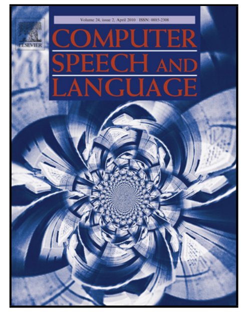
Computer Speech and Language 28.3 (2014): 788-811

DOI: http://dx.doi.org/10.1016/j.csl.2013.10.003

Accepted Manuscript



Title: A rule-based translation from written Spanish to Spanish Sign Language glosses

Author: Jordi Porta Fernando López-Colino Javier Tejedor José Colás

PII: S0885-2308(13)00086-7
DOI: http://dx.doi.org/doi:10.1016/j.csl.2013.10.003
Reference: YCSLA 626

To appear in:

Received date: 1-2-2012
Revised date: 26-8-2013
Accepted date: 11-10-2013

Please cite this article as: Porta, J., López-Colino, F.,A rule-based translation from written Spanish to Spanish Sign Language glosses, *Computer Speech & Language* (2013), http://dx.doi.org/10.1016/j.csl.2013.10.003

A rule-based translation from written Spanish to Spanish Sign Language glosses

Jordi Porta[1], Fernando López-Colino, Javier Tejedor, José Colás

Human Computer Technology Laboratory
Escuela Politécnica Superior – Universidad Autónoma de Madrid
Av. Francisco Tomás y Valiente, 11. 28049 Madrid, Spain

Abstract

One of the aims of Assistive Technologies is to help people with disabilities to communicate with others and to provide means of access to information. As an aid to Deaf people, we present in this work a production-quality rule-based machine system for translating from Spanish to Spanish Sign Language (LSE) glosses, which is a necessary precursor to building a full machine translation system that eventually produces animation output. The system implements a transfer-based architecture from the syntactic functions of dependency analyses. A sketch of LSE is also presented. Several topics regarding translation to sign languages are addressed: the lexical gap, the bootstrapping of a bilingual lexicon, the generation of word order for topic-oriented languages, and the treatment of classifier predicates and classifier names. The system has been evaluated with an open-domain testbed, reporting a 0.30 BLEU (BiLingual Evaluation Understudy) and 42% TER (Translation Error Rate). These results show consistent improvements over a statistical machine translation baseline, and some improvements over the same system preserving the word order in the source sentence. Finally, the linguistic analysis of errors has identified some differences due to a certain degree of structural variation in LSE.

Keywords: Machine Translation, Spanish, Spanish Sign Language, LSE, Deaf People Communication

1. Introduction

Translation helps people to communicate across linguistic and cultural barriers. However, according to Isabelle and Foster (2005), translation is too expensive, and its cost is unlikely to fall substantially enough, to constitute it as a practical solution to the everyday needs of ordinary people. Although it remains to be seen if machines will ultimately compete seriously with humans in translation, machine translation can help break linguistic barriers and can make translation affordable to many people. This situation is especially important for Deaf people, since translation helps Deaf and Hearing communities to communicate with each other, and provides Deaf people with the same opportunities to access information as everyone else.

1.1. Sign Languages

Sign Languages (SLs) exploit a different physical medium from the oral-aural system of spoken languages. SLs are gestural-visual languages, and this difference in modality causes SLs to constitute another branch within the typology of languages. However, there are still many myths around SLs. One of the most common and enduring myths is that sign language is universal; however, every Deaf community has its own SL, even within the same country. For example, in Spain, apart from Spanish Sign Language (LSE), there exists another recognised SL, known as Catalan Sign Language (LSC) and used within the Catalan Deaf community. Another common myth is that there is a correlation between spoken and SL families. American Sign Language (ASL) and British Sign Language (BSL), however, despite the fact that both are SLs used in English-speaking countries, are mutually unintelligible. Sign

Email addresses: jordi.porta@uam.es (Jordi Porta), fj.lopez@uam.es (Fernando López-Colino), javier.tejedor@uam.es (Javier Tejedor), jose.colas@uam.es (José Colás)
[1]Corresponding Author. Tel.: +34 91 497 33 64 / Fax: +34 91 497 22 35

languages do not derive from spoken languages, but, as any other languages, can be influenced by contact with other languages. As with spoken languages, when the use of an SL is extended, dialects and varieties are developed. In other words, SLs are natural languages that arise spontaneously in Deaf communities to fulfil the function of communication.

Phonocentrism is a view in which speech is considered to be superior to, or more natural than, written language. This attitude, which is still dominant in western culture, has negatively affected the consideration of sign languages, adding to their status of minority languages the status of minorised languages, i.e., languages whose value is not recognized on the interactional scene by speakers of a sociolinguistically dominant language. This also encourages the assumption that speakers of the minorised language conform to the usage and interactive norms set by their interlocutors. Nevertheless, scientific claims regarding the status of SLs as "real" human languages have been made since the work of W. Stokoe on the ASL (Stokoe, 1960). SLs in developed countries, especially SLs in Europe and North America, dominated research during the first decades of study. Currently, language typologists still have some difficulties accessing research on a number of regions like Central and South America, Africa, or Asia because most publications are written in national languages not accessible to a wider international audience. According to Zesman (2007), the state of knowledge regarding SLs has developed like a mosaic with many untiled gaps, but this is increasingly giving typologists a clearer picture of the range of diversity in SLs. Some cross-linguistic and typological studies of SLs, such as those by Sandler and Lillo-Martin (2006) or Brentari (2010), have shed light on both the universals and the diversity of SLs, contributing to the understanding of human languages in general.

The lack of a writing system is characteristic of SLs, and one shared with two-thirds of the spoken languages of the world. Strictly speaking, the only way of representing SLs is to use motion pictures. However, several notational systems exist. The most important today are SignWriting (Sutton, 1974) and HamNoSys (Prillwitz et al., 1989). SignWriting was conceived primarily as a writing system, and has its roots in DanceWriting (Sutton, 1973), a notation for reading and writing dance movements. HamNoSys was conceived as a phonological transcription system for SLs, with the same objective as the International Phonetic Alphabet (IPA) for spoken languages. There is another alphabetic writing system, designed specifically for LSE, and called SEA (*Sistema de escritura alfabética*) (Herrero et al., 2001); this uses the Latin alphabet, and has LSE's phonology as its basis. As this paper focuses primarily on syntax, glosses will be used here instead of a phonological notation. Glossing is a commonly used system for explaining or representing the meaning of signs and the grammatical structure of signed phrases and sentences in a text written in another language. However, glossing is not a writing system that could be understood by sign language users. A machine translation (MT) system needs to produce an animation to be considered a complete and useful system.

Most contemporary works on SLs have adopted language theories created for spoken language instead of developing new theories. This adoption leads naturally not only to the study of the phonology, morphology, and morphosyntax of SLs, but also to the study of all other descriptive levels found in spoken languages. However, from the point of view of natural language processing, SLs are still under-resourced or low-density languages – that is to say, little or no specific technology is available for these languages, and computerised linguistic resources, such as corpora or lexicons, are very scarce. This situation, of course, is not exclusive to SLs, since it in fact applies to most of the languages of the world.

1.2. Motivation and organisation of this paper

According to INE[2] and MEC[3], there are 1,064,000 Deaf people in Spain. Half of that population is more than 65 years old. Among the Spanish Deaf population, 47% of people do not have basic-level education, and in many cases are illiterate. Only one to three percent of Spanish Deaf people have graduated from basic school (INE, 1999; MEC, 2000). Moreover, about 92% of the Spanish Deaf population has significant difficulties in using written Spanish (INE, 2008). The unemployment rate for Deaf people is almost 20%, and that of Deaf women is as high as 30% (INE, 2003).

LSE is the name of the most widespread gestural-visual language used by the Spanish Deaf community and by persons who live and interact with them. The estimated number of users of LSE is 100,000. For about 20 to 30% of those users, LSE is their second language. LSE was finally recognised as an official language in Spain in 2007, but

[2]Instituto Nacional de Estadística (Spanish National Statistics Institute)
[3]Ministerio de Educación y Cultura (Spanish Ministry of Education)

the ratio of interpreters to LSE users in Spain is 1/221, below the European average. From a linguistic point of view, LSE refers to a variety of the Sign Language used in a large central area of Spain, with Madrid being its cultural and linguistic epicentre. There are other local varieties in some areas of Asturias, Aragón, Murcia, Western Andalucía (Seville), etc. The mutual intelligibility among these varieties is high due to their lexical similarities.

Technology can help with the integration of the Deaf into general society. This paper describes the machine translation module used by the TARTESO translation service (López-Colino et al., 2011), which translates Spanish into LSE. This service is composed of three main modules (see Figure 1): a speech recognition module, which transcribes speech into text; the machine translation module, which uses speech transcriptions or regular text and translates them into a gloss notation; and the SL synthesis module (López-Colino and Colás, 2012), which generates animations using an avatar technology.

The rest of this paper is organised as follows: Section 2 presents a sketch of LSE describing phenomena common to many other SLs. Section 3 presents a review of machine translation systems for SLs and SL corpora that have been built to date. Section 4 presents our rule-based Spanish-LSE machine translation system. The corpus used for experimentation, the evaluation procedure, and the results are presented in Section 5, together with some discussion. Finally, the work is concluded in Section 6.

2. A sketch of LSE

The most important documentation for a language is a reference grammar, which documents the principles governing the construction of words and all kinds of grammatical structures found in a language. Fortunately, the first LSE grammar has recently been published as *Gramática didáctica de la lengua de signos española* (GDLSE) (Herrero Blanco, 2009). It is a complete grammar, describing the linguistic structures of LSE with a pedagogical orientation. Since the grammar is addressed to students and teachers of LSE, it includes exercises and multimedia material. Furthermore, the *Diccionario normativo de la LSE* (DILSE) (Fundación CNSE, 2008) represents another milestone for LSE. Despite being descriptive works, both works have a normative-oriented guidance, so it is in their aim to represent the contemporary standard for LSE.

2.1. Phonology and non-manual components

The articulators in SLs are the hands, the arms, the face (including eyes, eyebrows, mouth, cheeks and tongue), the head, the neck, the shoulders, and the body. During signing, the visual field of one signer covers all the articulatory elements, hands and body, of the other signer. However, in a conversation, attention is focused on the gaze of the other signer, leaving the rest in the peripheral vision. The gaze, as in spoken communication, plays a major role in SL communication. In LSE and in other SLs, hands are the main articulator. The dominant hand is the hand involved in the articulation of one-handed signs. In two-handed signs, each hand can act either symmetrically or antisymmetrically, or there can be a dominant hand acting as a coordinator and a so-called non-dominant hand, or passive hand, used as the place of articulation. In right-handed signers, the dominant hand is the right hand and the left hand is the non-dominant hand. The role of the hands is reversed in the case of left-handed signers. This difference, however, generates no confusion in understanding. Signs are formed by a set of minimal, meaningless elements or parameters (formerly called cheremes), whose variation expresses the appearance of a new sign. There are four main kinds of these articulatory parameters or phonemes: the handshape, the hand orientation, the place of articulation, and the movement. Moreover, the contact of the hand with the body or the other hand, and the type of symmetry in two-handed signs, are also considered phonologically relevant in LSE. Some studies also consider the non-manual components (NMCs) of a sign as parameters of that sign, but the number of signs whose meaning depends on NMCs in LSE is so few that this is usually ignored in phonology.

NMCs have several functions in morphology, and can be performed alone or in conjunction with the manual component. In LSE, the most significant functions are the expression of intensification or restriction, and the expression of the verbal modality. Intensification/restriction is usually applied to adjectives, and can be alternatively expressed in LSE by means of adverbs as ‘MUCHO (much)’ or ‘ALGO (some)’. In the case of the non-manual expression of intensification, clenched teeth, semi-closed eyes, and tense articulation denote the intensification of a positive property, as for ‘ALTÍSIMO (very tall)’, and intensive negative properties are denoted by inflated cheeks and blowing, as for ‘MUY-CERCA (very close)’. These NMCs can be also applied to classifiers representing adjectives. Some verbal aspects have no manual realisation. This is the case with inchoative aspect value, for which the performance of the sign is suspended, with open mouth and raised eyebrows. Deontic modality is marked with raised eyebrows and pursed lips, whereas epistemic modality uses frowns and a u-like mouth shape, along with an affirmative movement of the head. The main functions of NMCs in syntax are the distribution of information and agreement. Information as topic is marked with pauses and eyebrows. Agreement can be marked by the gaze and directing the body towards the position of the space where the object or the subject of a verb has been previously placed. The interrogative construction in LSE is realised by raising the eyebrows and bending the torso forward while signing the verb of the sentence. The negative constructions may be realised by including the sign 'NO (not)' after signing the verb. However, it is also common simply to negate with the head while signing the verb. NMCs also have other roles in discourse and in conversation (in particular in turn-taking), but these fall outside the scope of this work.

2.2. The use of space

Signs are produced in the signing space, i.e., the physical space around the signer. The space used for the articulation of signs is a lexical space and has a phonological value. This space comprises some body areas and the central and lateral planes of symmetry. However, space has additional and complex uses. Space can be used topographically, i.e., as a map, when visible objects used in communication are signalled, or analogically, when objects are placed in analogy to the real world, this use being iconic or metaphoric. Some extensions of the frontal lexical space have grammatical purposes. They are used for deixis, anaphora, and agreement. Space is used syntactically, assigning arbitrary places (the locus) for pronominal reference. These locations have no relation to the physical locations of the objects situated there, since these objects can be abstract. In subsequent references, grammatical relationships to objects can be established through locations, which can act as morphemes.

2.3. Parts of speech

Hengeveld et al. (2004) distinguish among specialised, flexible, and rigid languages, depending on how grammatical functions and lexemes relate to each other. In specialised or rigid languages, each function has a particular lexeme, whereas flexible languages display classes of lexemes that can be used in more than one function without requiring lexical or syntactic derivation. Flexibility generates homography. Spanish is considered a specialised language but, as Herrero Blanco (2009) points out, LSE displays some flexibility in its part-of-speech (PoS) system. Sometimes the noun, the adjective, and the verb are represented by the same sign, as is the case for '*dolor* (pain)', '*doloroso* (painful)' and '*doler* (to hurt)', which are all represented by the single sign DOLOR. The topic of LSE flexibility will be further explored in Section 4.2.1.

2.4. Classifiers

According to Aikhenvald (2000), almost all languages have some grammatical means for the linguistic categorisation of nouns and nominals. Classification is a categorisation device found in languages for marking humanness, animacy, shape, size, and other inherent and salient properties of objects. In SLs, the phenomenon of classification is pervasive, and has received different analyses in the SL literature (Schembri, 2003). Supalla (1986) defined classifiers as morphemes incorporated into verbs, but Sutton-Spence and Woll (1999) imposed three conditions on classifiers: to reference a group of elements sharing some common features; to be a proform; and to be found in verbs of movement and location. Proforms have an anaphoric primary function (Engberg-Pedersen, 1989), and can be defined as anything that refers to, and stands in the place of, something previously presented in the discourse. Consequently, pronouns are a subclass of proforms. Proforms are realised as a handshape with a specific orientation, which represents the prototypical orientation of the object. Both parameters can be considered morphological components of the proform. For example, a proform can represent a person, using a particular handshape with a horizontal orientation to indicate that the person is lying down.

Broadly speaking, we will refer to classifiers as handshapes that substitute for other signs and have a morphological value. There are three groups of classifiers in LSE: entity classifiers, handle classifiers, and extension classifiers. Entity classifiers are descriptive classifiers that represent objects by their dominant dimension or shape. Therefore, an object can be classified according to its dominant dimension as one-dimensional (pen, spoon, etc.), two-dimensional (book, sheet, table, etc.), circular (dish, coin, etc.), three-dimensional (ball, building, fruit, etc.), cylindrical (bottle, column, etc.), or fluid (water, smoke, crowds, etc.). Four-wheeled and two-wheeled vehicles have different classifiers, as do objects having frameworks (swimming pool, picture, etc.). Other classifiers can represent the thickness of an object or stacked objects. Persons, animals, and parts of the body also have classifiers. Handle or instrument classifiers describe the handling of an entity rather than representing the entity itself. Extension classifiers depict the perimeters of objects or make reference to their surfaces. Classifiers, as well as path movements, are among the most iconic elements of SLs.

2.5. Morphology

Grammatical values can be expressed in the form of independent single words, called free morphemes, or as additions or alterations to the phonology of words, in which case they are called bound morphemes. Not all languages have the same grammatical values or express them in a similar way. For example, Spanish has gender morphemes, but LSE expresses gender values by other means, using free morphemes. Languages can be classified according to the degree of presence of the two kinds of morphemes. At one end of the spectrum there are isolating languages, like Chinese, with a low morpheme-per-word ratio, and at the opposite end there are synthetic languages, like Inuit languages, with lengthy words. However, in general, the vast majority of languages, including SLs, contain both kinds of morphemes. In addition, SLs have a widely used morphological resource for the phenomenon of classification noted in Section 2.4.

SLs contain not only inflection but also derivation, compounding, and classification (see Section 2.6). Inflectional morphology is the study of the changes undergone by words in expressing grammatical values. LSE, and virtually all SLs studied, have a rich inflection system. Bound morphemes found in LSE belong to one of two classes, flective or introflective. Flective morphemes are concatenated to the phonological form, like a movement or a repetition[4] of the original movement, sometimes in a direction. They are mainly used for expressing several forms of plural, some values of the verbal aspect, or as a mechanism for morphological derivation. Introflective morphemes cause internal changes to the phonological form of the sign, and hence they are non-concatenative in nature. This kind of morpheme is also found in Semitic languages. Introflection is used to agree with the subject, the object, the receiver or the place. The classifier of the subject and/or the object are also introflected in verbs. The number is introflected in pronouns, etc. LSE has, in some cases, irregularities such as suppletive base forms, i.e., different unrelated base forms for the same sign, indicating different grammatical values.

LSE does not mark gender in signs, and hence gender is expressed by adding 'HOMBRE (man)' / 'MUJER (woman)' after the sign. Nouns and adjectives can inflect, changing some of their phonological components to express number, to be classified, to express place or agreement with a subject, etc. Nouns admit adjectival inflection of size and shape. There are several ways of forming plurals in LSE, depending partly on the phonological structure of signs. Personal pronouns inflect in number and, like some temporal nouns, they can incorporate numerals, e.g., 'VOSOTROS-3 (you three)' or 'SEMANA-2 (week two)'. Place deictics have values for first person 'AQUÍ (here)' / 'ESTE (this)', second person 'AHÍ (there)' / 'ESE (that)', and third person 'ALLÍ (there)' / 'AQUEL (that)'.

Additionally, in LSE, the third person deictics and the personal pronoun 'ÉL (he)' are split into two values – proximal or present, and distant or absent. It is worth noting that this distinction does not exist in ASL. Adverbs can be modified to express intensification or restriction by means of the execution speed and accompanying facial and other non-manual components. Verbs can inflect for both subject and object agreements, and can express a variety of aspectual values such as habitual, continuous, etc. Verbal agreement in LSE is not as systematic as in Spanish, where all verbs agree in number. Padden (1990) proposed a classification for verbs of several SLs on the basis of which affixes may be added. She proposed three basic classes: plain, agreement, and spatial verbs. Plain verbs neither inflect for person and number nor take locative affixes. Agreement verbs use the syntactic space and can inflect in person and number, but cannot incorporate spatial information. Finally, spatial verbs use the topographical space and incorporate spatial information, but cannot incorporate person and number. Aspect marking is widely used by SLs. In LSE, reduplicating a sign rapidly can mark habitual aspect. An abrupt stop before the completion of the sign marks an unrealised inceptive aspect. If the sign is repeated with a slower circular movement, its aspect value is continuous. Some nouns and adjectives can undergo aspect marking, for instance with habitual values, e.g., 'ENFERMO (sick)' > 'ENFERMIZO (sickly)', and continuous (monotonous) values. All the classes of verbs can be slightly modified to incorporate aspectual values.

[4]Repetition can be seen as reduplication, another non-concatenative operation found in some languages as Tagalog.

2.6. Classifier nouns

A classifier noun is a common noun acting as a free morpheme prepended to another noun to complement its lexical meaning, or to other lexical categories to produce lexical derivatives. The use of classifier nouns is common to all SLs. For instance, LSE resorts to classifier nouns for tree names, so that the Spanish noun '*naranjo* (orange tree)' is translated as a construction headed with the sign 'ÁRBOL (tree)' and followed by the sign for the fruit of the tree 'NARANJA (orange)' forming 'ÁRBOL NARANJA (lit. tree orange)'. The deverbal noun '*trabajador* (worker)' is translated as 'PERSONA TRABAJAR (lit. person to-work)' and the adjective '*envidioso* (envious)' as 'PERSONA ENVIDIA (lit. person envy)'. For LSE, the most widely used classifier nouns are 'AGUA (water)', 'ÁRBOL (tree)', 'CASA (house)', 'DINERO (money)', 'TRABAJO (work)', 'PERSONA (person)' and 'TIENDA (shop)'. They present a different degree of grammaticalisation so that some of them can be suppressed, as is the case with 'PERSONA (person)' but not with 'TIENDA (shop)'. However, in general, classifier nouns do not need to be realised when they have been used before or when the context contains enough information to infer them. More examples of Spanish words translated to classifier noun constructions are '*cuartel* (barracks)' → 'CASA MILITAR (lit. house military)', demonyms (the name given to a person from a particular region or country), like '*europeo* (european)' → 'PERSONA EUROPA (lit. person Europe)', '*billete* (bill)' → 'DINERO TARJETA (lit. money card)', and '*apicultura* (beekeeping)' → 'TRABAJO ABEJA (lit. work bee)'.

2.7. Classifier predicates

Classifier predicates (CPs) pose a challenge to the definition of what constitutes a linguistic expression. Many times, CPs are spatial metaphors, scene visualisations, or even pantomimic descriptions. A CP is a complex sign typically preceded by a nominal phrase. A CP is created by selecting a classifier. The signer performs a three-dimensional movement with the hand that communicates an outline, a position in the space around the signer, a movement through the dimensional space, a physical or abstract dimension, or any other property of the object that needs to be communicated. This type of predicate is ideal for describing scenes, manipulation tools, movements, size, and other information of visual or spatial scenes or processes.

2.8. Syntax

Traditionally, languages have been classified according to the prevailing order of syntactic functions (Greenberg, 1963). Thus, both Spanish and English have a Subject-Verb-Object (SVO) basic order, although the Spanish word order is freer than that of English due to the richer morphology of Spanish. Early research in ASL word order during the 60's was influenced by the functional-oriented typology of that decade. In the 80's, the description of other SLs led other researchers to consider other motivations, such as semantics or pragmatics, to explain the word order of different SLs. Other typological studies have highlighted the fact that some languages do not organise the structure of the basic components of the sentence from the syntactic functions of subject and object, but from the grammaticalisation of the functions of topic and comment, as do Mandarin, Mayan languages, and the languages of the Philippines. However, still other studies (Li and Thompson, 1976) have shown that some languages use both types of organisation, resulting in four types of languages: subject-prominent languages, topic-prominent languages, languages prominent in either subject or topic, and languages organised according to other parameters. LSE is considered a topic-prominent language. Topicalisation, i.e., movement of the topic element at the beginning of the sentence, is widespread in SLs, but to varying degrees. The topic is usually marked by NMCs and followed by an intonation break. In general, topicalisation occurs only in the main clause. SLs have a considerable variation of basic word order of a sort very similar to that found in the so-called discourse-oriented languages, including Chinese and Russian. SLs productively employ devices for presenting information in a different order determined by discourse factors, such as presenting the old information (topic) before the new (comment). The presence of specific non-manual markers is noted when the order is other than basic.

3. Review of machine translation systems to Sign Languages and Sign Language corpora

As was reported by Huenerfauth (2003), all MT systems for SLs published up to 2003 were just works in progress or simple demonstrators. However, some systems were particularly distinguished, including the ZARDOZ system (Veale and Conway, 1994), the ViSiCAST Translator (Bangham et al., 2000), the ASL Workbench (Speers, 2001) and the TEAM project (Zhao et al., 2000). All these systems were rule-based, following the predominant MT model

Corpus	Languages pair	Sents.	Tokens		Types		Reference
			Words	Signs	Words	Signs	
-	Italian-Italian SL	585	15,000	6,000	1,442	300	(Bertoldi et al., 2010)
RWTH-Phoenix	German-German SL	2,468	16,500	10,500	1,302	1,895	(Bungeroth et al., 2006)
-	Chinese-Thaiwanese SL	1,983	-	11,501	-	2,159	(Chiu et al., 2007)
-	Catalan-Catalan SL	199	2,416	4,305	446	648	(Massó and Badia, 2010)
ATIS	English-Irish SL	595	4,436	4,333	600	544	(Bungeroth et al., 2008)
ATIS	German-German SL	595	4,903	4,291	627	498	”
ATIS	English-South African SL	595	4,436	2,525	600	422	”
ID	Spanish-Spanish SL	2,000	20,000	15,000	800	400	(San-Segundo et al., 2010)
DL	Spanish-Spanish SL	2,000	11,000	8,000	1,000	500	”

Table 1: Counts for sign language parallel corpora for different language pairs.

of that time, and made use of either transfer-based or interlingua-based approaches. The only approach dealing with classifier predicates was that of Huenerfauth (2006), who presented a multi-path approach combining interlingua, transfer, and direct approaches as a whole.

Today's prevalent research approaches for MT are data-driven. Example-based machine translation (EBMT) and statistical machine translation (SMT) have replaced the earlier rule-based machine translation (RBMT) approaches. Results obtained using these EBMT and SMT approaches have reached such a high level of quality that they make machine translation genuinely useful, especially when the target language is English. However, data-driven approaches estimate their parameters from an aligned bilingual corpus, and their accuracy depends heavily on the quality and size of this corpus. Unfortunately, corpora for SLs are still far from reaching the state of the art of those for spoken languages. The problem of modality and the lack of a standardised writing system make data acquisition for SLs a time-consuming and expensive task compared to the acquisition of spoken or written data. SL data are usually acquired first by video-recording the signer's performance, and then aligning the video sequence with multi-channel annotations that describe in a gloss format the signs and the non-manual components. The use of annotation tools such as ANVIL (Kipp, 2001), SignStream (Neidle, 2002), iLex (Hanke, 2002), ELAN (Wittenburg et al., 2002), or the annotation tool presented in (Braffort et al., 2004) makes easy the construction and annotation of new corpora for SLs. On the whole, if corpora are essential for language research, parallel corpora are essential for machine translation. To construct machine translation corpora, SL data annotations are extracted and aligned with the equivalent expressions in other languages. The amount of parallel data currently available for SLs is typically a few hundred sentences or even just isolated words in restricted domains and on very few topics. However, some parallel corpora are being used for research in machine translation; these are summarised in Table 1.

Despite the lack of parallel corpora, the success of data-driven approaches to machine translation between spoken languages has led to the application of the same techniques to SLs. However, according to Morrissey (2011), most research in SL MT, with a few exceptions, has emanated from sporadic and short-term projects as opposed to long-term research investment. Some works are still worth mentioning: The Thai-to-Thai SL machine translation system (Dangsaart et al., 2008) presents a direct translation system with reordering rules. The system for Thai reaches an F-score of about 97% for a set of 297 test sentences. Morrissey (2008) presented exhaustive experiments on the MaTrEx, a hybrid approach combining EBMT and SMT (Stroppa and Way, 2006). Results of MaTrEx on the ATIS corpus were about 0.39 BLEU for English-to-Irish Sign Language translation, and about 50% for German-to-German Sign Language (DGS) translation.

For Spanish to LSE, Baldassarri and Royo-Santas (2009) described a rule-based demonstrator. Spanish is analysed using FreeLing (Atserias et al., 2006) dependency analysis, in which the relations between words are interpreted as the relations between blocks, such as subject, predicate, verb, etc. Through the application of a series of grammatical rules, the dependency analysis is transformed into a series of glosses. These grammatical transformations are carried out in reverse order. During transformation, some information flows in a bottom-up direction: partial translation and semantic information. Semantic information includes meaning, temporal and numeral information, and information regarding the type of block. A morphological transformation is applied to glosses before generating signs. In this stage, some glosses are replaced by their synonyms when they do not correspond to signs, or by their hyponyms when they have neither equivalent sign nor synonyms. In addition, some individual glosses can be replaced by sequences

[FIGURE 1]

Figure 1: Architecture of the Spanish-to-LSE Machine Translation System. 'ASR' stands for Automatic Speech Recognition.

of signs. All these replacements are based on a dictionary of equivalences. The system was tested with 92 sentences containing a total of 561 words. Sentences varied in length, ranging from simple phrases to more complex ones containing up to 25 words and several verbs along with subordination structures. Appropriate dictionary entries were created for the evaluation, with very satisfactory results: 96% of the words were correctly translated, and 93.7% of them were in correct order. Also worth mentioning is the existence of another rule-based Spanish-to-LSE MT system based in Apertium, a free/open-source platform (Forcada et al., 2011). The system uses shallow transfer rules and delivers an SEA representation of LSE. To our knowledge, there are no published results on this system but it is available on the Internet.[5]

The most complete work in Spanish-to-LSE translation to date is summarized in (San-Segundo et al., 2011). Their system focuses on helping Deaf people in a highly specific domain, the renewal of driving licenses. The final version of the system combines three approaches in the following order of decreasing priority: an EBMT strategy, a rule-based translation method, and an SMT approach. All three approaches compute a confidence value, and thresholds determine when to use the next approach. EBMT uses a set of translated sentences and a heuristic distance function based on the Levenshtein distance (Levenshtein, 1966). In the rule-based approach, every word is first mapped into syntactic-pragmatic categories. After that, handcrafted rules are applied bottom-up to convert tagged words into signs. Rules define short- and long-scope relationships between concepts or signs. Categorisation, as described in (López-Ludeña et al., 2011b), consists of tagging a word as non-relevant or giving it a list of manually defined tags such as lemmas or word categories. For the SMT approach, two methods were evaluated: the Moses system and a Stochastic Finite-State Transducer using the GIATI algorithm (Casacuberta and Vidal, 2004). Both methods use the same alignments between words and signs. However, to improve their Spanish-to-LSE SMT system, Factored Translation Models are combined with the phrase-based SMT model (López-Ludeña et al., 2011a). Factors are semantic categories and linguistic information such as part of speech or morphological features. The best BLEU obtained is above 0.70.

Examining the brief history of MT systems for SL, it seems that research and development of such systems generally followed contemporary MT trends. Within rule-based MT, emphasis has been placed on the modelling of particular phenomena, using the computational linguistic theories and formalisms available at the time of any given study. Conversely, the data-driven paradigm exploits statistical regularities found in available parallel data, which are scarce for SLs. Scarceness has led to an attempt to reduce parameters by using semantic classes and incorporating linguistic information as factors. Doing so has resulted in some level of success, but only in very limited domains. However, parallel data for SLs are expected to remain scarce in the foreseeable future, unless research efforts are focused on their acquisition.

4. A rule-based machine translation system for Spanish-to-LSE translation

The rule-based approach to translation is mainly motivated by the fact that available resources for data-driven approaches, especially corpora as seen in Section 3, are costly to acquire, and it is difficult for data-driven systems to estimate parameters from small corpora and still have reasonable coverage in wider domains. The source of knowledge for rule-based machine translation is contrastive linguistic studies. Contrastive linguistics focuses on similarities and differences in the structure and functioning of two or more languages. For the case of Spanish and LSE, contrastive studies have been carried out by Minguet (2000) and by *Gramática contrastiva español / LSE*, published as an electronic resource at the *Biblioteca Virtual Miguel de Cervantes*[6]. Additionally, some other, more specific contrastive studies, such as that of Villameriel (2008), can be found for discourse markers and other linguistic phenomena occurring in discourse or in conversations, such as role shifting or turn-taking.

The strategy chosen in this paper for translating Spanish to LSE is the transfer at the level of syntactic functions from dependency analyses. Figure 1 shows the architecture of the system presented here and its connection to other

[5] <http://aplica.prompsit.com/en/text_es_ssp> (Accessed: September 2012)
[6] <http://bib.cervantesvirtual.com/seccion/signos> (Accessed: January 2012)

[FIGURE 2a] [FIGURE 2b]

(a) Constituency tree (b) Dependency tree

Figure 2: Two kinds of analyses for '*Cuando está solo no se aburre* (When he is alone he does not get bored)'. Syntactic functions are labelled in the constituency tree branches of (a) and on dependency arcs of (b), where arrows go from heads to dependents.

external modules. A dependency tree from the analysis module abstracts the constituency structure and the surface word order in the source sentence. Only functional relations between words are represented in a dependency tree. The transfer module makes use of the bilingual lexicon and of the knowledge in the language pair-specific rule database to map dependency structures and to transfer a Spanish dependency tree into the corresponding dependency tree in LSE. Word ordering and morphological rules are applied to the transferred dependency tree so that the output of the generation stage is a sequence of glosses with morphological indications. It is important to note here that glossing is a transcribing system that abstracts away the phonological representation of signs. That is why, properly speaking, the system presented in this paper does not produce a true translation. However, glosses should be considered an intermediate result from which signs, after being translated to phonological forms, will be synthesised by some avatar technology. The different stages of the machine translation strategy are detailed next.

4.1. Analysis

The analysis stage is carried out by MaltParser (Nivre et al., 2007), a data-driven dependency parser. The parser makes use of the tokenisation and part-of-speech tagging delivered by FreeLing (Atserias et al., 2006). In addition, FreeLing provides the transfer phase with named-entity classification and recognition and word sense disambiguation using EuroWordNet (Vossen, 1998). In dependency parsing, the syntactic structure of a sentence is described in terms of words and binary semantic or syntactic relations between words; constituency or phrase structure does not play any role. Figure 2 contains the two kinds of analysis for a sentence. Given a phrase-structure analysis annotated with grammatical functions, it is not difficult to see how it can be transformed into a dependency analysis. MaltParser has been trained with the IULA Spanish LSP Treebank (Marimon et al., 2012), a corpus consisting of almost 590,000 tokens in 42,000 sentences from different technical domains. The resulting Spanish model has been evaluated with a 20% held-out set of the LSP Treebank, which reports a Labelled Attachment Score (LAS) of 93.14% and a Labelled Complete Match (LCM) of 47.60%. LAS is the accuracy obtained in assigning both a head and a function label; LCM is the proportion of completely correct analysed sentences. Results from a newspaper corpus, the Tibidabo Treebank (Marimon, 2010), report an LAS of 88.95% and an LCM of 36.20%. These results represent the state of the art in dependency parsing for Spanish.[7]

4.2. Transfer

The transfer stage involves two different sub-stages: lexical transfer and structural transfer. The lexical transfer employs a bilingual LSE-Spanish lexicon to convert the Spanish dependency tree generated from the analysis stage into an intermediate representation between the Spanish and LSE of this dependency tree. The structural transfer converts this intermediate dependency tree into an LSE dependency tree that is next moved to the generation stage. The bilingual lexicon is expanded by the incorporation of morpho-lexical and lexical-semantic relationships.

4.2.1. Lexical transfer

The *Diccionario normativo de la LSE* (DILSE) is a truly bilingual LSE-Spanish electronic dictionary created by the CNSE[8]. The DILSE is the first LSE normative dictionary containing equivalences in Spanish, definitions, and SEA transcripts, as well as other linguistic information, for about 4,100 combinations of signs and senses corresponding to 2,500 Spanish words. In addition to the signs of the dictionary, all the signs defined in the GDLSE are also considered normative. All these signs, together with their bilingual equivalences, constitute the basis of our transfer lexicon.

[7] <http://www.iula.upf.edu/recurs01_mpars_ul.htm> (Accessed: May 2013)

[8] Spanish Confederation of Deaf People

The lexical transfer module contains mapping between Spanish expressions and LSE expressions. A sequence '*papá y mamá* (dad and mom)' or the plural word '*padres* (ancestors)' are both translated to the single sign 'PADRE-MADRE (lit. father-mother)'. Sometimes the choice of a sign depends on the word sense assigned by FreeLing, as in the case of 'CAPITAL (capital)', which is translated with the gloss CAPITAL_1 when the sense corresponds to *Administration/Geography* or with CAPITAL_2 for the *Finance* sense.

The coverage of the lexicon is rather low if we consider that the number of Spanish lemmas in the DILSE is 2,500, that a Spanish dictionary such as the DRAE (Real Academia Española, 2001) contains more than 100,000 lemmas, and that our Spanish analysis lexicon contains 550,000 word-forms corresponding to 76,000 lemma-PoS combinations. To bridge the gap between vocabularies, several mechanisms for expanding the bilingual lexicon with new Spanish-LSE entries will be proposed in the following sections. On the one hand, given that the part-of-speech system of LSE is more flexible than that of Spanish, a Spanish derivative morphology has been implemented to relate Spanish lemmas. On the other hand, lexical ontologies provide semantic lexical relations like synonymy and hypernymy that will be used for lexical substitution.

Morpho-lexical relationships. The morpho-syntactic properties of a sign determine if the sign is a noun, an adjective or a verb. In Spanish, lexemes or stems are the bearers of meaning, while morphemes express grammatical values and determine the part of speech of the word. Spanish derivational processes are carried out mainly by means of affixation, so that morpho-lexical relationships can be established by means of a computational morphology. These lexical relations in Spanish have been extensively studied in (Santana-Suárez et al., 2004, 2005, 2006). However, in LSE, a sign is flexible enough as to be used as a noun, as an adjective, or as a verb without changing its form. In general, in LSE, it is the semantic class of a sign that determines its flexibility. Abstract nouns, for example, display such flexibility in LSE. When a noun names a property or an entity, its form is identical to the adjectival form. In the other cases, when the noun denotes a state, a process or an action, or a mental phenomenon, its form is shared with adjective and verb. Flexibility of verbs in LSE is not as general as for nouns. To give some examples, the mechanisms for verbal derivation sometimes add a repetition of the complete sign, as in 'VESTIR-SE (dressing oneself)', which derives from 'VESTIR (to dress)'. Other times, the derived sign becomes two-handed, as in 'ADELGAZAR (to slim)', which comes from the adjective 'DELGADO (slim)'. However, most verbs are flexible with nouns and adjectives.

In order to relate signs to their corresponding Spanish lemma derivatives, a finite-state derivational morphology for Spanish has been implemented. Nouns for names of quality, state, and condition are also derived, like '*humanidad* (humanity)' < '*humano* (human)', or denominal adjectives like '*poderoso* (powerful)' < '*poder* (power)', among others. The derivational morphology is also used to relate Spanish lemmas like '$\textit{nacimiento}_N$ (birth)' with the sign 'NACER (to be born)' through the pre-existing Spanish-LSE bilingual entry $\textit{nacer}_N$ (to be born) ↔ NACER (to be born). Using derivational morphology, the bilingual lexicon is increased by over 3,600 new entries relating new signs with previously known signs.

Lexico-semantic relationships. Another extension of the bilingual lexicon is achieved by the use of lexical semantic relations from the EuroWordNet (EWN) lexical ontology (Vossen, 1998). Apart from the obvious inferences from synonymy relations, by which the bilingual lexicon is increased with approximately 4,200 entries, other inferences can be drawn from hypernymy-hyponymy relations by allowing for substitution of an unknown word by its hypernym – for example, '*coche* (car)' ↔ '*vehículo* (vehicle)'. However, EWN can be exploited further; for example, as Rodríguez González (1992) noted, there exist lexical gaps between Spanish and LSE. Some Spanish word for which there is no equivalent sign can be validly expressed in LSE by enumerating some of its hyponyms, like '*metal* (metal)' ↔ 'ORO PLATA ... (gold silver ...)' or '*muebles* (furniture)' ↔ 'MESA SILLA ... (table chair ...)'. Over 200 bilingual new entries relating new signs with their hyponyms are obtained.

Classifier nouns. The detection of translatable cases using classifier nouns seen in Section 2.6 is implemented through Spanish EWN semantic relationships. For example, for the nouns naming trees, translatable cases are deduced with the hypernym (IS-A) and holonym-part (HAS-A) relationships, which can be formulated in predicate calculus as in (1a), in which the translation of '*naranjo* (orange tree)', which is 'ÁRBOL NARANJA (lit. tree orange)', is deduced from the fact that '*naranjo*' is a tree having a part which is the fruit '*naranja* (orange)'. For the translation of the names of the inhabitants of a certain place, like '*americano* (american)', the corresponding country name is deduced with (1b), so that it can be translated as 'PERSONA AMÉRICA (lit. person America)'. Since classifier nouns in LSE have different

[FIGURE 3a]

(a) Spanish dependency analysis

[FIGURE 3b]

(b) Intermediate dependency analysis

Figure 3: Spanish dependency analysis for '*El objetivo final es dominar el mundo y crear imperios* (The final objective is world domination and empire creation)' and its transferred analysis to LSE where some nodes are disconnected and some new nodes appear. The transferred sentence has been converted into a pseudo-interrogative sentence, which is a proper way to translate a copulative sentence.

degrees of grammaticalisation, a set of eight rules covers the six principal cases of classifier nouns: 'ÁRBOL (tree)', 'CASA (house)', 'DINERO (money)', 'TRABAJO (work)', 'PERSONA (person)' and 'TIENDA (shop)'.

(1) a. $translation(x, \text{ÁRBOL } y) \leftarrow hypernym(x, tree) \wedge holonym\text{-}part(x, y) \wedge hypernym(y, fruit)$

b. $translation(x, \text{PERSONA } y) \leftarrow human(x) \wedge holonym\text{-}member(x, y) \wedge derived\text{-}from(x, y) \wedge place(y)$

Fingerspelling. Fingerspelling, which represents alphabets or syllables of spoken languages, is used in different languages for different purposes. Fingerspelling or reduced fingerspelling, using only the first letters of the words, may be used to represent words from the source language with no equivalent sign in the lexicon, like proper nouns. In so doing, a proper noun like 'Juan (John)' will be glossed 'DL-JUAN ' and signed as 'J-U-A-N' or 'J-U-A'.

4.2.2. Structural transfer

Intermediate dependency trees resulting from the lexical transfer are transferred to LSE. In this process, some nodes representing words are removed, as in the case of definite articles, which have no equivalence in LSE. Other information, however, can appear in the new structure, like the non-overt subjects in Spanish represented in LSE as personal pronouns whose person and number are inferred from the morphology of the Spanish verb. However, all these transferences are not mandatory, since they are controlled by parameters indicating if a phenomenon has to be transferred. Figure 3 contains an example of transference of a copulative sentence into a pseudo-interrogative one in which an interrogative pronoun ('QUÉ (what)') is introduced. Note that definite articles (*el* (the)) disappear in the transferred tree, since they have no translation as signs in LSE, and that despite the fact that some words are not transferred to signs, their nodes are maintained because their function should be preserved.

4.3. Generation

The generation stage, from an LSE dependency tree, and as a previous step before the synthesis, generates the sequence of glosses. It is divided into two different stages. First, a word order generation takes the LSE dependency tree produced by the transfer stage and produces a sequence of signs representing it; next, a morphological generation takes this sequence of signs and adds the corresponding morphological information.

4.3.1. Word order generation

Another consequence of the Spanish inflectional system is that its constituent order is freer than is the case in other languages with impoverished morphologies, like English. LSE order, as has been seen in Section 2.8, is determined by the verb type and the topicalisation of complements. It is said that word order in LSE, from a functional point of view, is freer than in Spanish, but at the same time, informative order is stricter. Expression of place and time at discourse level differs in both languages, being less redundant in LSE than in Spanish. LSE places this frame information at the beginning of the discourse and does not repeat it in subsequent sentences. On the other hand, Spanish carries time information in each verb. In the case of locatives, LSE situates this information in the linguistic space, and uses it for agreement in subsequent references. LSE can also use the passive hand as a location for the object signed with the active hand. According to the GDLSE, the linear precedence for constituents in LSE predicative sentences is:

(2) Unmarked word order in LSE sentences
$\text{Tc} \prec \text{To} \prec \text{S} \prec \text{IO} \prec \text{DO} \prec \text{V} \prec \text{T/L} \prec \text{Asp} \prec \text{Mod}$

where Tc is the 'contextual theme', consisting of generic temporals and locatives, To is the topicalised argument, S is the subject, IO is the indirect object, DO is the direct object, V is the verb, T/L are 'specific' temporals and locatives, and Asp and Mod are aspect and mood auxiliaries. The difference between contextual and specific temporals/locatives depends on their roles in the argument structure of the verb. If they are complements, they usually appear after the verb, otherwise they constitute the frame of the predication. LSE has no copulative verbs (linking verbs), and hence copula is expressed as a pause between the subject and the attribute, accompanying the attribute with an assertive movement of the head, and frequently doubling the subject with a pronoun, as in 'JUAN MÉDICO ÉL (lit. Juan doctor he)'.

Broadly speaking, in a predication, the position of its head, normally a verb, is taken as a reference for ordering the rest of the constituents according to their function in the predication. Non-argumental temporal and locative adjuncts appearing before the verb are fronted to occupy the contextual thematic positions. Topicalised or dislocated arguments are heuristically detected as those appearing in front of the verb. They occupy the second position in the LSE sentence, after the contextual theme and in front of the subject and the objects of the verb. These arguments are advanced from positions after the verb to prior positions. All other prepositional complements and adjuncts after the verb are left behind the verb. Modal and aspectual markers are moved to their appropriate positions.

Dealing with complex and compound sentences requires further inspection of form and function. For example, sentential subordination in LSE has fewer lexical markers than in Spanish, and relative constructions are expressed in LSE by means of repetition, pronominalisation, and agreement, as in (3). Completive complementisers in Spanish are expressed with interrogative pronouns, as in (4), etc.

(3) *el hombre que vino ayer es un ladrón* (the man who came yesterday is a thief)
HOMBRE VENIR AYER. HOMBRE LADRÓN ÉL (lit. man to-come yesterday. man thief he)

(4) *él dice que no puede venir* (he says he cannot come)
ÉL DICE QUÉ-?. NO PODER VENIR (lit. he says what cannot to-come).

Noun phrases constitute another of the major structures in LSE, and their constituents have the following linear precedence:

(5) Unmarked word order in LSE noun phrases (NPs)
NCL $\prec$ N $\prec$ Compl $\prec$ Dem $\prec$ Pos $\prec$ Num $\prec$ Loc $\prec$ CL-compl $\prec$ Indef

where NCL are classifier nouns, N is the head of the NP, Compl are the nominal complements and adjuncts of N, Dem are demonstratives, Pos are possessives, Num are numerals, Loc are locatives, CL-compl are complements with classifiers, and Indef are indefinites. However, in the case of marked structures, the basic order differs from the previous established order. As an example, consider a Spanish noun phrase with the structure 'N_1 de N_2 (N_1 of N_2)'. When translating this expression to LSE, sometimes N_1 precedes N_2 ($N_1 \prec N_2$), as in '*hermano de Luis* (Luis's brother)', which is translated as 'LUIS HERMANO (lit. Luis brother)' because '*hermano* (brother)' denotes a permanent relationship. In other cases, the order is 'N_2 N_1', as in '*mesa de estudio* (study table)', because in this case '*estudio* (study)' is the function or purpose of '*mesa* (table)', and it is thus translated as 'MESA ESTUDIO (lit. table study)'. Meronymy relations, i.e., part-whole relations, usually expressed as 'N_1 de N_2', are translated as 'N_2 N_1', as in '*el volante del coche* (car's steering wheel)', translated as 'COCHE VOLANTE (lit. car steering wheel)'.

The ordering of signs is formulated as a constraint satisfaction problem. A precedence graph is built from the information about the relative order between a head and its dependent or between pairs of dependents sharing a head. In (6), an example rule is shown stating that the prepositional modifier of a noun precedes it if the head noun denotes a permanent relationship.

(6)
$$\boxed{1}\left[\text{CAT: } \textit{noun}\right] \xrightarrow{\text{MOD}} \boxed{2}\begin{bmatrix}\text{CAT: } \textit{prep} \\ \text{PFORM: } \textit{de} \\ \text{POBJ: } \left[\text{CAT: } \textit{noun}\right]\end{bmatrix} \wedge \textit{permanent\_relationship}(\boxed{1}) \implies \boxed{2} \prec \boxed{1}$$

For the generation of constraints on word order from the dependency tree, the algorithm applies the following steps:

1. For each word H, we create two convenience elements '$(_H$' and '$_H)$', which represent the boundaries of the phrase headed by H, which maintains the relation '$(_H \prec H \prec {}_H)$'.

[FIGURE 4a] [FIGURE 4b]

(a) Dependency analysis (b) Precedence graph

Figure 4: Dependency analysis of ‘*Tiene tres pelotas de colores y muñecos de goma* (He has three coloured balls and rubber dolls)’ and precedence graph from its output from the transfer stage. The circuit outlined with solid lines represents a topological ordering of the graph that will result in the gloss sequence ‘PELOTA DE COLOR TRES Y MUÑECO ESPECÍFICO GOMA HAY (lit. ball of colour three and dolls specific rubber there-is)’.

2. For every non-topicalised dependent D of a head H, if D precedes H then ‘$(_H \prec (_D$’ and ‘$_D) \prec H$’, and if H precedes D, then ‘$H \prec (_D$’ and ‘$_D) \prec {}_H)$’, otherwise ‘$(_H \prec (_D$’ and ‘$_D) \prec {}_H)$’.
3. For every pair of non-topicalised dependents D_1 and D_2 of a head H, which has been placed on the same side in the previous step, if D_1 precedes D_2, then ‘$_{D_1}) \prec (_{D_2}$’ and if D_2 precedes D_1, then ‘$_{D_2}) \prec (_{D_1}$’. Otherwise, they maintain the same relative order as in the source sentence.
4. Apply topological sort to the graph of precedence, breaking any detected cycles, to obtain the word ordering.

Since step 4 of the previous algorithm can only be applied on an acyclic graph, errors during development on inconsistencies are easy detected as cycles and reported to the grammar writer. Figure 4 contains an example of the application of the algorithm to a dependency analysis. As it can be seen in Figure 4a, the dependency structure for the transferred tree is identical to the original tree but not to the transferred lexical items. The gloss sequence is obtained by applying a topological ordering to the precedence graph in Figure 4b, generated with the previous algorithm. It is worth noting that conditions in step 2 of the algorithm ensure non-projectivity in the constituency of the realised sentence. However, evidence exists in ASL of constructions with a topicalised constituent of an embedded clause fronted into the main clause. This phenomenon, which is clearly non-projective, is neither common nor frequent, and it is not documented in GDLSE.

4.3.2. Morphological generation

Some glosses not realised as signs are simply removed, but some other glosses are annotated with morphological features that modify their phonological representation, like an absent third person personal pronoun located to the left of the signer, a transitive verb agreeing with its object, a plural realised by repeating the noun, or a verb with a continuous aspect. The generation of phonological representations from morphologically annotated glosses, suitable for their synthesis by an avatar technology, is outside the scope of this paper, but some details can be found in (Porta et al., 2012).

Special cases in generation are classifier predicates. Some LSE postpositions are mainly nominal complements with locative or temporal meanings. In Spanish, these complements correspond to nominal phrases with prepositional phrase modifiers introduced by locative prepositions and locutions as ‘*sobre* (on)’, ‘*bajo* (under)’ or ‘*tras* (behind)’ and by temporal prepositions as ‘*entre* (between)’ or ‘*desde* (from)’. In these cases, the locative construction can be generated as a two-handed sign in which the passive hand represents the head of the phrase and the active hand represents the term of the preposition, both with their entity classifier. The only approach to the automatic generation of classifier predicates has been developed by Huenerfauth (2006) for ASL. His approach makes use of a 3D scene generator from natural language descriptions. These 3D scenes, together with a set of predefined templates, were used to define the position of the hands and perform the classifier predicate.

5. Evaluation

Human evaluation is fundamental and remains crucial to proper assessment of the quality of machine translation systems. When an SL animation from the output of an MT system is evaluated, however, the whole process is taken into account – not only the translation into a symbolic representation, but also its synthesis as an animation, how realistic and fluid the performance of the avatar is, etc. The fact remains, however, that the lack of a generally accepted writing system for SLs poses problems to MT evaluation. In spite of the existence of transcription systems, much vital information for the production of signed utterances is missing in any text-based representation, and automatic evaluation using distances between a reference and candidate translations as mere strings is inadequate for measuring how much of the essence of the reference translation is preserved (Morrissey and Way, 2006). In addition, the choice of transcription representation influences the evaluation result. SiGML, glosses, and sign identifiers were used in (Morrissey, 2011) to explore this effect. Although the scores achieved by SiGML were found to be better than those achieved by glosses, and despite the fact that glosses are not considered an adequate representation of signs by many, automatic evaluation of text-based representations remains useful to assess the internal progress of MT systems. In addition, the results obtained could be comparable with those found in the research literature and can lead to a correct SL output with a synthesis module using a lexicon relating glosses and animated forms.

5.1. LSE corpus description

A parallel Spanish-LSE corpus has been created from the material used in the psycholinguistic study published in (de los Reyes Rodríguez Ortiz, 2005). The study seeks to assess the effectiveness of LSE as a language of instruction compared with an oral language and the effect of late learning of LSE on comprehension. An experienced human interpreter was in charge of translating the texts originally written in Spanish into LSE. The performance was first video-recorded and then transcribed into a gloss notation. A bimodal bilingual second interpreter (daughter of two deaf parents) translated the video recordings in LSE back into Spanish. The resulting Spanish speech was then manually transcribed. We will call the original Spanish texts corpus A; the LSE glosses resulting from the transcribed video performance corpus B; and the transcription of the speech translation of the video-recorded performance corpus C. According to the author of the study, the meaning of the back-translated sentences in C agrees quite well with their correspondent sentences in A, confirming both the skill of the first interpreter in translating into LSE and the naturalness and fluency of the Spanish in the back-translation performed by the second interpreter.

Experiments are conducted using six different texts with different topics and with different degrees of difficulty of understanding. The topics covered by texts one to six are as follows: student accommodation; history and war; a day in a dog's life; population and demographic problems; domestic violence; and a day in a little girl's life. The difficulty of a text is determined by the background knowledge required to understand it, the topic variation, its thematic progression, the structure of each text, the ease of extraction of the main ideas, the domain vocabulary, the lexical richness, and the degree of abstraction. Texts one and two are of low difficulty, texts three and four are of average difficulty, and texts five and six are the most difficult to understand.

All the texts of the corpus come in three versions: (A) the Spanish source version; (B) an LSE translation of A; and (C) a Spanish translation of B. The three versions of the corpus have been manually segmented into sentences and aligned appropriately. As an example, (7) contains three aligned sentences extracted from the corpus. The sentence in (7a) is the source sentence in A, the sentence in (7b) corresponds to the LSE translation of (7a) in B, and that, in turn, is translated back into Spanish in (7c) in C.

(7) a. *La guerra es tan vieja como la humanidad* (War is as old as humanity)
    b. GUERRA EDAD IGUAL HUMANIDAD IGUAL+ (lit. war age equal humanity equal-ASP.CONT)
    c. *La guerra tiene la misma edad que la humanidad* (War has the same age as humanity)

The number of sentences and lexical items is shown in Table 2. In Spanish counts, for versions A and C, word-forms, multiwords, numbers, quantities, punctuation, dates, proper names, etc., are included. LSE counts for version B include signs, classifier predicates as a unit, and punctuation. The average number of tokens per sentence in Spanish is fourteen to fifteen, while the average number of tokens in the LSE sentences is approximately ten.

The LSE part of the corpus has been annotated following the guidelines published in (Alonso Baixeras et al., 1998). Glosses are accompanied by several symbols describing prosodic, syntactic and morphological values. Classifier predicates are also tagged and described. Unfortunately, marking has not been applied uniformly and consistently. Not all the instances of the same phenomenon have been marked, and not all the marked instances have always been marked in the same way. The symbol '(?)' marks pseudo-interrogatives, which are realised by means of NMCs. The symbol '(cond)' is a mark used in conditionals, also performed by NMCs. The single or multiple repetitions of a sign are marked using the symbols '+', '++' and '+++'. However, they are not used consistently, even within the same text. The most frequent use of the symbol '+' corresponds to the plural marking in nouns, like 'PERSONA+ (people)', to some kind of adverbial intensification in verbs like 'PREOCUPAR+ (to worry-INTENS)', or to the aspectual marking in verbs, like 'BUSCAR+ (to search-ASP.FREQ)'. The symbol '++' is also used for quantifying in either nouns like 'PERSONA++ (people)', adjectives like 'TRAVIESO++ (naughty-PL)' or verbs like 'JUGAR++ (to play-ASP.CONT)'. The

	Language	Sentences	Words			Punctuation	
			Tokens	Types	Lemmas	Tokens	Types
Corpus							
A	Spanish	229	3,033	1,082	782	373	9
C	Spanish	229	2,946	912	642	446	8
B	LSE	229	1,992	611	-	348	2
Testbed							
A	Spanish	195	2,538	957	700	307	7
C	Spanish	195	2,519	813	580	394	8
B	LSE	195	1,709	532	-	-	-

Table 2: Counts of sentences, tokens (or running items), types (or vocabulary), lemmas and punctuation signs (excluding full stops) for versions A, B and C of the texts in (de los Reyes Rodríguez Ortiz, 2005).

symbol '+++' seems to be used to mark descriptive plural, i.e., a kind of plural realised by repeating the signs in different places, sometimes in locations analogous to the locations of these objects in the real world. The symbol '(mofl)' stands for an NMC realised by inflating the cheeks with a value of intensification. Finally, the symbol '(2m)' states that one-handed signs have been made two-handed, as a realisation of plurals, aspectual values, or reciprocity, among other possibilities. There are different marks for some kinds of classifier predicates: descriptive constructions (CLD), locative constructions (CLL), pronominal constructions (CLS), body-as-subject related actions (CLC), body-part related actions (CLCP), and instrumental constructions (CLI). The corpus contains thirty-four classifier constructions, but no one contains the phonetic information for its articulation.

5.2. Experiments and discussion

To evaluate the system output quality, a testbed has been created using the corpus described in the previous section. Because Spanish sentences in C are obtained from LSE sentences in B, and B sentences are obtained from Spanish sentences in A, C and A can be considered paraphrases, since C is a back-translation of A. However, from lexical and structural points of view, the differences between A and B and the differences between B and C are not the same. It is worth noting that some kind of interpretation has been applied in the translation from A to B. The existence of classifier predicates in the LSE translation supports that hypothesis. Both versions A and C convey the same meaning, however, the translation from B to C is a more 'literal' translation in lexical and structural terms and B seems *a priori* easier to obtain from C than from A. In order not to introduce artificial differences, morphological mark-up is not taken into account, since there is not enough consistency in its use. Sentences containing classifier predicates have not been incorporated into the testbed, although the possibility of generation of these classifier predicates will be discussed later. Each testbed used for experimentation consists of 195 sentences. Counts for words and signs can be found in Table 2.

In automatic machine translation evaluation, the most widespread metric used is BLEU (Papinemi et al., 2002). BLEU is calculated by combining individual n-gram precision of up to four with a brevity penalty. Translation error rate (TER) is another common metric used for MT evaluation. TER attempts to measure the minimum amount of editing that a human would have to perform to convert the system output into the reference translation (Snover et al., 2006). The metric is computed from the number of insertions, deletions, and substitutions of words, and from the number of phrasal shifts, i.e., the movement of word sequences to other locations of the output translation. Note that while BLEU is better if higher, TER is better if lower.

In order to evaluate and compare the RBMT approach presented in this paper with other approaches, several experiments have been conducted with the phrase-based Moses translation model (Koehn et al., 2007) providing a baseline with a data-driven approach. For Moses, initial experiments with development sets led to poor results because of the scarcity of the data of the testbed and the sparseness of the problem, which cause overfitting. These poor results agree with those reported by Morrissey (2011) as having been obtained in similar experiments. Results for Moses are estimated using cross validation.

In the experiment called Experiment I in Table 3, standard ten-fold cross-validation is used, splitting data randomly into training and test. Both the language model and the phrase translation table are obtained from the training

Experiment	Testbed	TER	BLEU	Modified N-gram Precision				Methods
				1-gram	2-gram	3-gram	4-gram	
I	A	140.31	0.000	0.152	0.012	0.002	0.000	Moses + GIZA++ + crossval
II	A	144.14	0.003	0.142	0.009	0.002	0.001	Moses + GIZA++ + crossdom
III	A	82.45	0.018	0.394	0.082	0.021	0.004	Moses + Bilingual Resources + crossval
IV	A	82.85	0.033	0.384	0.086	0.022	0.004	Moses + Bilingual Resources + crossdom
V	A	72.56	0.088	0.492	0.157	0.058	0.019	RBMT without parsing + Spanish word order
VI	A	72.21	0.089	0.492	0.164	0.058	0.019	RBMT + MaltParser

Experiment	Testbed	TER	BLEU	Modified N-gram Precision				Methods
				1-gram	2-gram	3-gram	4-gram	
I	C	116.06	0.000	0.260	0.026	0.003	0.000	Moses + GIZA++ + crossval
II	C	121.03	0.005	0.237	0.017	0.003	0.002	Moses + GIZA++ + crossdom
III	C	46.47	0.217	0.765	0.332	0.157	0.082	Moses + Bilingual Resources + crossval
IV	C	46.90	0.216	0.762	0.325	0.154	0.082	Moses + Bilingual Resources + crossdom
V	C	41.61	0.296	0.801	0.405	0.244	0.154	RBMT without parsing + Spanish word order
VI	C	42.73	0.302	0.800	0.410	0.252	0.159	RBMT + MaltParser
VII	C	41.96	0.306	0.801	0.417	0.256	0.162	RBMT + MaltParser + Manual Corrections

Table 3: Results of the experiments: BLEU and precision values are given as indices but the TER metric is expressed as percentage. A TER higher than 100% indicates that there are more words in the output translation than in the corresponding reference translation. Note that while BLEU value is better if higher, TER is better if lower.

set of each fold with GIZA++ (Och and Ney, 2003). As it can be seen from the results in Table 3 BLEU is zero. The reason for this low result is that three- and four-gram precisions are zero for almost all the test folds. An according high average TER indicates that much postediting work on the output will be required to match the reference translation. Moreover, a TER higher than 100% indicates that there are more words in the output translation than in the corresponding reference translation. When inspecting the phrase table obtained, it seems clear that, due to the experiment's small size, data are incorrectly aligned, causing most of the errors. This result applies to both testbeds.

In Experiment II, six-fold cross-validation is used but in each fold one of the six texts is left out the training set and used as test set. The second experiment aims to evaluate how well Moses works when training and test domains differ when translating Spanish into LSE. As can be seen in Table 3 for both testbeds, modified n-gram precisions are slightly lower and TER is slightly higher than in Experiment I. This result might indicate some domain dependency but a reliable conclusion should be based on more data.

For the third and fourth experiments using Moses (Experiments III and IV in Table 3), different language models are obtained from the training data of each fold, as in the previous experiments. However, in an attempt to improve the results obtained with GIZA++ in Experiments I and II, Moses is provided with a phrase table obtained from the bilingual lexicon of the RBMT system, i.e., the bilingual lexicon. In that table, the same probability has been given to every possible translation of a word. The phrase table is the same for each of the folds. From the results, the good average adequacy obtained from the cross-validation tests in light of the precision figures for shorter n-grams should be noted. The use of good and accurate bilingual resources, like bilingual lexicons or manually aligned data, is clearly helpful to SMT when parallel text data are scarce. Similar results are obtained within both testbeds.

Experiment V uses the RBMT system of this paper without the generation module described in Section 4.3.1. In this case, the generation algorithm has been replaced with an algorithm preserving the original order of the corresponding words in the Spanish source sentence. The resulting order looks like Signed Exact Spanish (SES), which is not a natural language but a manually coded language, i.e., a representation of an oral language in a gestural-visual form. However, the output differs from SES mainly in that, because of the lexical transfer, not all Spanish words are expressed in LSE, and the correspondence between words and signs is not one-to-one. It is worth noting that similar experiments can be found in the literature. In (Kanis et al., 2007), using a training set of 12,616 sentences, translation from Czech to Signed Czech, a manually coded language, reached a BLEU of 0.81, a WER of 13.14% and a PER of 11.64%. In (Stein et al., 2010), two experiments applied to both German and DGS involving simple lowercasing and four-letter stemming have been conducted. The BLEU/PER obtained in each experiment was 0.021/85.7% and 0.026/81.1% respectively. However, the reported baseline with Moses was 0.181 BLEU and a 71.0% TER with a training set of 2,565 sentences and a test set of 512 sentences. Combining several systems, they finally reached a BLEU of 0.234 and a TER of 65.5%. Note that the disparity between these results is because Czech and Signed Czech have the same surface order, but German and DGS do not. Here, figures for the RBMT system are given in Table 3. It is worth noting that, as explained in Section 5.1, source sentences in the testbed C have been obtained as back-translations from LSE sentences, and a slight bias towards LSE surface order can be detected, which is perfectly accommodated in a relatively free word-order language like Spanish. This could explain the relatively good BLEU obtained in this experiment.

The complete RBMT system described in this paper has been evaluated as Experiment VI in Table 3. This system has been provided with several parameters to deal with LSE variation, which are used in the experiment to approximate linguistically the system output to its corresponding reference translation in the testbed. The most important phenomenon revealed by the parameters is the realisation of pronominal subjects, since, contrary to the norm in standard LSE, they are realised in the majority of the LSE sentences of the testbed. In view of the results, there is a general gain in terms of BLEU and TER with respect to the figures obtained in the rest of the experiments. However, the most important thing to note here is that, in order to get better results than in Experiments III and IV, many implemented word order rules were deactivated. These rules were implemented according to the LSE grammar but the order of signs of the testbed does not seem to follow the current norm in LSE.

In order to quantify the effect of the parser quality on results in terms of BLEU and TER, the dependency trees delivered by the parser have been manually corrected. The number of corrections applied to testbed C totals 399 distributed as follows: 220 head assignments and 179 function labels. Corrections affected 263 tokens (a 7.7% of the tokens). The results of this experiment can be seen in Experiment VII. The improvement in BLEU and TER with respect to the parsers uncorrected is very small. This result supports the hypothesis that the word order found in LSE for the testbed B is based on the word order of the source sentences and not on the grammatical functions, since the correction of parsers does not have a dramatic improvement in translation results.

Results confirm that data scarcity and domain sparseness lead the data-based approach to perform worse than the rule-based system. Providing bilingual lexical resources have a positive effect in data-based approaches but differences using cross validation or leaving one text out are small. We think that this result should not be interpreted as domain independence. Instead, we consider that data are not still enough to measure the out-of-domain effect.

On the rule-based translation side, the most important conclusion that can be drawn from the experiments is that the order of signs is similar to the order of words in the Spanish fragments. We think that this result should not mean that LSE and Spanish have similar word orders or that the order generated by the system is not valid. We consider that LSE order admits some degree of freedom and that the order of signs in the testbed is also valid for the purpose of communication. At this point, deeper and more extensive experiments measuring human understanding should be performed to draw reliable conclusions.

5.3. Analysis of errors

Because data-driven approaches base much of their success on data, an analysis of errors is rarely found in evaluations. However, a proper classification of errors, which should take into account the linguistic phenomena involved in the error, can be helpful in choosing more appropriate system architecture and in the development of new rules for RBMT systems and resources for translation. As an exception, López-Ludeña et al. (2012) contains an analysis of errors for the Spanish-to-LSE SMT system described in Section 3 applied to a highly specific domain. In that work, the main sources of errors are reported in order of decreasing importance as follows: (a) the differences in the number of words and signs for parallel sentences, due to the absence in Spanish of pronominal subjects, and the absence in LSE of definite articles, prepositions, and copula, and the different realisations of plural, and others; (b) the differences in word order (SVO versus SOV); (c) the incorrect generation of classifier predicates; (d) out of vocabulary words; and (e) specific names in LSE periphrastic expressions.

Apart from the expected differences caused by the lexical gap, a detailed analysis of errors between Experiment VI and VII reveals that some differences are caused by the incorrect assignment of functions in the dependency analysis. However, many other differences come from the choice of alternative structures to convey meaning. Some examples are reported below, showing the source sentence ‘a.’, the target or reference translation ‘b.’ and the system’s output 'c.'. Sometimes the differences arise from an alternative form of realisation, as in (8), where the intensification of a property as in '*muy feliz* (very happy)', which plays the role of a predicative complement in '*lo vemos muy feliz* (we see him/her very happy)', is translated by doubling the predicate, resulting in 'VER FELIZ, VER FELIZ (lit. see happy, see happy)', which is a more iconic expression of quantity.

(8) a. *Cuando abrimos lo vemos muy feliz.* (When we open, we see him/her very happy)
b. TOCA ABRIR, VER FELIZ, VER FELIZ. (lit. turn to-open, see happy, see happy)
c. TOCA ABRIR FELIZ MUY VER. (lit. turn to-open happy very see)

Other examples in the testbed also display variation in the choice of referring expressions. For example, the Spanish female reference '*Ana* (Ann)' is realised either as the finger-spelled sign 'DL-ANA (A-n-n)', or as the deictic 'IX-"él" MUJER (she woman)', a pronoun with a free morpheme carrying the information about the female sex of the referent, or as the generic 'MUJER (woman)'. All these three forms of realisation can be seen respectively in examples (9)–(11).

(9) a. *Allí Ana juega con sus amigos.* (Ann plays there with her friends)
b. ALLÍ PARQUE AMIGO DL-ANA JUGAR. (lit. there park friend A-n-n to-play)
c. ALLÍ ANA JUGAR AMIGO SU. (lit. there A-n-n to-play friend her)

(10) a. *A Ana le encanta la bata.* (Ann likes her dressing gown)
b. IX-"él" MUJER ENCANTAR BATA. (lit. he woman to-like dressing gown)
c. BATA ANA ENCANTAR. (lit. dressing gown Ann to-like)

(11) a. *Ana no hace nada.* (Ana does nothing)
b. MUJER HACER NADA. (lit. woman to-do nothing)
c. ANA NADA NO-HACER. (lit. Ann nothing not-to-do)

It is worth noting that example (11) contains a double negative, often required in Spanish, whose translation in (11c) does not match the reference in (11b). However, as in other languages, the double negative could be understood as a form for adding emphasis to the negative interpretation. In any case, the translation of the general negation is not fully implemented, because of the complex interactions of negation with other phenomena, which require further study.

Complex sentences, i.e., those with an independent clause and one or more dependent clauses, are stable in Spanish. However, according to GDLSE, LSE presents alternative forms of expression containing non-standard markers, which are still in the process of grammaticalisation, and hence have an uncertain future. This issue causes differences as seen in (12), where the clause structure is marked in square brackets. The first dependent clause is introduced by the complementiser '*que* (that)', which plays the role of direct object of the verb '*querer* (to want)'. The second clause, a causal clause, is introduced by '*porque* (because)'. The output translation in (12c) reflects the original structure in (12a), but differs significantly from the structure of the reference translation in (12b), in which both the causal marker and the verb are doubled.

(12) a. [*Mamá no quiere* [*que los niños vayamos a la calle*] [*porque está oscuro*]].
(Mom does not want children to go to the street because it is dark)
b. [MOTIVO MAMÁ QUERER-NO [NIÑOS IR CALLE] QUERER-NO] [MOTIVO OSCURO].
(lit. [reason mom not-to-want [children to-go street] not-to-want] [reason dark])
c. MAMÁ [NIÑOS IR CALLE] QUERER-NO [MOTIVO OSCURO]
(lit. [mom [children to-go street] not-to-want [reason dark])

Other differences arise from the interaction of several phenomena. A very common construction found in SLs is rhetorical questions or pseudo-clefts. Structurally, they consist of a question-answer pair and are used to bring a particular constituent into focus. An example of this construction can be found in (13b). In the example, several markers are introduced. First, '-?' is used to represent the CNMs expressing a question, and then ',' represents an intonation break. Note that, in this particular example, because of the interaction of a coordinating conjunction, the structure appears twice.

(13) a. *El objetivo final es dominar el mundo y crear imperios.*
(The final objective is to dominate the world and to create empires)

b. OBJETIVO FINAL QUÉ-?, MUNDO DOMINAR, CREAR QUÉ-?, IMPERIO.
(lit. objective final what, world to-dominate, to-create what empire)

c. OBJETIVO FINAL MUNDO DOMINAR IMPERIO CREAR.
(lit. objective final world to-dominate empire to-create)

There are several possible reasons for some of these differences. First, a single non-native signer has done the translation and, therefore, several aspects could have influenced the native language of the signer, such as word order. Second, data were acquired in Seville, in southern Spain, away from the influence centres of the standard, which are thought to be central and eastern communities. Third, when the corpus was created in 1995, no standard existed for LSE. Therefore, the language used surely corresponds to a variant with notable differences from the current LSE standard. And fourth, some degree of variation exists within LSE, as in all languages, which leads to the existence of different alternative forms of realisation of the same linguistic phenomenon. Being those different forms equally valid, some parameters could be introduced to the system in order to select one form of realisation or another.

5.4. Analysis of classifier predicates

Classifier predicates deserve special comment. Only two cases of the original thirty-four classifier predicates are identified as postpositions, which are tractable cases (see Section 4.3.2). With the rest of the classifier predicates, the ordinary translation of their corresponding Spanish expressions into sequences of signs should be valid for the purpose of communication. Almost all other classifier predicates found in the corpus could be generated by identifying some linguistic structure in the source language, e.g., '*el perro corre detrás de la pelota* (the dog runs after the ball)'. In this case, the classifier construction uses the spherical classifier handshape, representing the ball, and the quadruped classifier handshape, representing the dog. Both hands are moved: the 'dog's hand' after the 'ball's hand'. Another, much easier example is '*el perro está bajo la mesa* (the dog is under the table)', translated by placing the dog's hand under the table's hand. As can be noted, both examples present spatial movements or relations between objects. However, consider the example in (14) containing the expression '*no pueden dormir* (they cannot sleep)' which has been translated by a classifier construction with the meaning of "person tossing and turning into the bed".

(14) a. *Muchas veces por la noche no pueden dormir preocupados por ese problema*
(Many times they cannot sleep because they are worried about that problem)

b. MUCHAS-VECES POR-LA-NOCHE CLS: 2 "persona dando vueltas en la cama" PREOCUPAR+ PROBLEMA IX-"ese"
(lit. many.times at.night a.person.tossing.and.turning.into.the.bed to.worry-ASP.CONT problem that)

This construction could be performed iconically, using the passive hand to represent the bed, with the bed's classifier handshape, while the active hand represents the person, using the 'V' handshape and performing an alternative rotation over the passive one. It is important to note that world knowledge is required for the generation of this expression: sleeping takes place in a bed; when a person cannot sleep, s/he usually tosses and turns in the bed; this action is performed on a particular axis of rotation, etc. However, an alternative valid translation using signs could be accomplished with 'ELLOS DORMIR NO (lit. they to-sleep no)'. Other interesting translation examples can be found, as for 'CANIBALISMO (cannibalism)', in which the signer has created a two-handed classifier predicate simulating the grasping of different parts of his/her own body and the eating of them. The construction is marked with 'CLC'. However, the construction is similar to the sequence of signs: 'BRAZO COMER, CUERPO COMER, PIERNA COMER (arm to-eat, body to-eat, leg to-eat)'. In order to generate classifier constructions, a machine translation system should first identify them and then be able to generate a correct description. Due to the difficulty involved in the generation of such constructions, very few works on this subject are found in the literature – only Huenerfauth (2006) for ASL and López-Colino and Colás (2011) for LSE.

6. Conclusions and future work

The choice of a particular type of technology to process a language is greatly influenced by the density of the language, i.e., the availability of digital stored resources. Commercial research and development have concentrated on high-density languages. According to Varga et al. (2007), for medium-density languages, spoken by over half of humanity, parallel corpora can be compiled using digital resources like literary and religious texts, international law, movie captioning, software documentation, bilingual magazines, corporate home pages, and annual reports. Today, LSE, like any other sign language, is a low-density or under-resourced language. Because of modality, acquisition of sign language data is a time-consuming and expensive task compared to the acquisition of spoken or written data. For LSE, no parallel corpus currently exists of sufficient size to enable data-driven approaches to machine translation in non-restricted domains. However, LSE is being standardised, and the language's first grammar, as well as a normative dictionary, have recently been published. All these factors make the rule-based approach the only viable approach today for the development of a production-quality Spanish-to-LSE machine translation system for non-restricted domains.

This paper has presented a transfer-based approach for Spanish-to-LSE translation by delivering LSE glosses. A wide-coverage Spanish parser was used to obtain a dependency analysis. The PoS flexibility of LSE enables the augmentation of the bilingual lexicon by bootstrapping the initial lexicon using morpho-lexical relationships. Lexical-semantic relationships were used to bridge the lexical gap. Classifier names were also deduced using lexical-semantic relationships. A word order generation algorithm was presented to deal with the topic-oriented surface order of LSE. The algorithm made use of linear precedence rules operating at the level of syntactic functions and dealing with topicalisation. Glosses were annotated with number, aspect, or other morphological information. Finally, classifier predicates, which are one of the cornerstones of sign languages, had a translatable subclass corresponding to Spanish prepositional expressions with locative and temporal meanings. Moreover, some of the solutions presented in this paper to the problems found for machine translation between Spanish and LSE could be equally valid, with minor adaptations, for the translation between other spoken languages and other sign languages.

A parallel Spanish-LSE corpus was created from the data used in a psycholinguistic study. Although the corpus is comparable in size to other corpora used in data-driven approaches, it is not domain-specific. A parallel testbed was used to evaluate the system. Experiments reported a BLEU about 0.30 and a TER at about 42%. On the one hand, these results confirm the intuition that when data are scarce, or in sparse domains, induction is difficult or impossible, and hence bilingual data and handcrafted grammars are good alternatives. On the other hand, experiments show that LSE word order generation has played some little role in this corpus. This does not mean that word order in LSE is the same than in Spanish. In fact, they have been reported to be very different. However, order in LSE is free enough to accommodate the word order of the source sentences. Finally, a linguistic-oriented error analysis has shown that many differences between the system output and the reference translations arise from variations in the realisation of the linguistic structures, and that classifier predicates are the most difficult expressions to generate.

In spite of the relatively good results obtained with the approach presented in this paper, experiments should be extended in several directions in order to assess properly the performance of the system. The corpus should incorporate more linguistic phenomena in order to evaluate the coverage of the system and its components. In addition, the corpus should conform to standard LSE, since, from a sociolinguistic point of view, the standard is the variant used in addressing a linguistic community, in this case the LSE community. This paper is also incomplete, since it addresses only the translation of written Spanish-to-LSE glosses. Glosses are used as an intermediate symbolic representation for guiding the development and assessing the performance of the system in sign choices and in generating sign sequences. However, glosses are merely an underspecified symbolic representation of a signed message. Glosses are not even a writing system for Sign Languages that could be understood by Deaf people. A complete MT system for Sign Languages should produce animations, and a genuine and proper evaluation should involve Deaf people and should measure comprehension. Many other important aspects have been left unaddressed for the moment, and there is still a great deal of work to do, even in the field of animation synthesis, in order to make MT systems operational and useful.

References

Aikhenvald, A. Y., 2000. Classifiers: A typology of noun categorization devices. Oxford Studies in Typology and Linguistic Theory. Oxford University Press, New York, USA.

Alonso Baixeras, P., Martínez Sánchez, F., Nogueira Fos, R., Pinto Muñoz, J., Serna Serna, M., Vázquez Aznar, A., 1998. Signar: Enseñanza de la lengua de signos española. CNSE.

Atserias, J., Casas, B., Comelles, E., González, M., Padró, L., Padró, M., May 2006. FreeLing 1.3: Syntactic and semantic services in an open-source NLP library. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC-2006). Genoa, Italy, pp. 48–55.

Baldassarri, S., Royo-Santas, F., 2009. An automatic rule-based translation system to Spanish Sign Language. New Trends on Human-Computer Interaction, 1–11.

Bangham, J. A., Cox, S. J., Elliott, R., Glauert, J. R. W., Marshall, I., Rankov, S., Wells, M., April 2000. Virtual signing: capture, animation, storage and transmission - an overview of the ViSiCAST project. In: IEE Seminar on Speech and Language Processing for Disabled and Elderly People. London, England, UK, pp. 1–7.

Bertoldi, N., Tiotto, G., Prinetto, P., Piccolo, E., Nunnari, F., Lombardo, V., Mazzei, A., Damiano, R., Lesmo, L., Del Principe, A., May 2010. On the creation and the annotation of a large-scale Italian-LIS parallel corpus. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Languages Technologies. Valletta, Malta, pp. 19–22.

Braffort, A., Choisier, A., Collet, C., Dalle, P., Gianni, F., Lenseigne, B., Segouat, J., May 2004. Toward an annotation software for video of sign language, including image processing tools and signing space modelling. In: Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-2004). Lisbon, Portugal, pp. 201–203.

Brentari, D., 2010. Sign Languages. Cambridge University Press, New York, USA.

Bungeroth, J., Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way, A., van Zijl, L., May 2008. The ATIS sign language corpus. In: Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC-2008). Marrakech, Morocco, pp. 2943–2946.

Bungeroth, J., Stein, D., Zahedi, M., Ney, H., May 2006. A German Sign Language corpus of the domain weather report. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC-2006). Genoa, Italy, pp. 2000–2003.

Casacuberta, F., Vidal, E., June 2004. Machine translation with inferred stochastic finite-state transducers. Computational Linguistics 30 (2), 205–225.

Chiu, Y.-H., Wu, C.-H., Su, H.-Y., Cheng, C.-J., January 2007. Joint optimization of word alignment and epenthesis generation for Chinese to Taiwanese sign synthesis. IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (1), 28–39.

Dangsaart, S., Naruedomkul, K., Cercone, N., Sirinaovakul, B., November 2008. Intelligent Thai text - Thai sign translation for language learning. Computers and Education 51 (3), 1125–1141.

de los Reyes Rodríguez Ortiz, I., April 2005. Comunicar a través del silencio: Las posibilidades de la lengua de signos española. Universidad de Sevilla.

Engberg-Pedersen, E., 1989. Proformes en morphologie, syntaxe et discourse. In: Quertinmont, S. (Ed.), Études Européenes en Langue de Signes. Edirsa, Brussels, pp. 35–52.

Forcada, M. L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz-Rojas, S., Pérez-Ortiz, J. A., Sánchez-Martínez, F., Ramírez-Sánchez, G., Tyers, F. M., 2011. Apertium: a free/open-source platform for rule-based machine translation. Machine Translation 25 (2), 127–144.

Fundación CNSE, 2008. Diccionario normativo de la lengua de signos española. Fundación CNSE.

Greenberg, J. H., 1963. Universals of Languages. MIT Press, Cambrigde, Massachusets.

Hanke, T., May 2002. iLex - A tool for sign language lexicography and corpus analysis. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002). Las Palmas, Spain, pp. 923–926.

Hengeveld, K., Rijkhoff, J., Siewierska, A., November 2004. Parts of speech systems and word order. Journal of Linguistics 40, 527–570.

Herrero, Á., Alfaro, J. J., Cascales, I., 2001. Lecciones de escritura alfabética de la LSE. Universidad de Alicante.

Herrero Blanco, Ángel., 2009. Gramática didáctica de la lengua de signos española (LSE), 1st Edition. Ediciones SM.

Huenerfauth, M., September 2003. A survey and critique of American sign language natural language generation and machine translation systems. Tech. rep., University of Pennsylvania.

Huenerfauth, M., January 2006. Generating American Sign Language classifier predicates for English-to-ASL machine translation. Ph.D. thesis, University of Pennsylvania.

INE, 1999. Encuesta sobre discapacidades, deficiencias y estado de salud. INE.

INE, 2003. Las personas con discapacidad y su relación con el empleo. Encuesta de población activa del segundo trimestre de 2002. INE.

INE, 2008. Encuesta sobre discapacidades, autonomía personal y situaciones de dependencia. INE.

Isabelle, P., Foster, G., 2005. Machine Translation Part I - Overview. Encyclopedia of Language and Linguistics 15, 404–422.

Kanis, J., Zahradil, J., Jourcicek, F., Müller, L., 2007. Czech-Sign speech corpus for semantic based machine translation. Lecture Notes in Computer Science 4188, 613–620.

Kipp, M., September 2001. ANVIL - A generic annotation tool for multimodal dialogue. In: Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech-2001). Aalborg, Denmark, pp. 1367–1370.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E., June 2007. Moses: Open source toolkit for statistical machine translation. In: Proceedings of Demonstration and Poster Sessions at the 45th Annual Meeting of the Association for Computational Linguistics (ACL-2007). Prague, Czech Republic, pp. 177–180.

Levenshtein, V. I., 1966. Binary codes capable of correcting deletions, insertions and reversals. Soviet Physics Doklady 10, 707–710.

Li, C. N., Thompson, S. A., 1976. Subject and topic: A new typology of language. In: Li, C. N. (Ed.), Subject and Topic. Academic Press, London/New York, pp. 457–461.

López-Colino, F., Colás, J., February 2011. The synthesis of LSE classifiers: From representation to evaluation. Journal of Universal Computer Science 17 (3), 399–425.

López-Colino, F., Colás, J., 2012. Spanish Sign Language synthesis system. International Journal of Visual Languages & Computing 23, 121–136.

López-Colino, F., Tejedor, J., Porta, J., Colás, J., 2011. Integration of a Spanish to LSE machine translation system into an e-learning platform. Lecture Notes in Computer Science 6768, 567–576.

López-Ludeña, V., San-Segundo, R., de Córdoba, R., Ferreiros, J., Montero, J. M., Pardo, J. M., September 2011a. Factored translation models for improving a speech into sign language translation system. In: Proceedings of the 12th Annual Conference of the International Speech Communication (Interspeech-2011). Makuhari, Chiba, Japan, pp. 1605–1608.

López-Ludeña, V., San-Segundo, R., Lutfi, S., Lucas-Cuesta, J., Echevarry, J., Martínez-González, B., July 2011b. Source language categorization for improving a speech into sign language translation system. In: Proceedings of the Workshop on Speech and Language Processing for Assistive Technologies (SLPAT-2011). Edinburgh, Scotland, UK, pp. 84–93.

López-Ludeña, V., San-Segundo, R., Montero, J. M., de Córdoba, R., Ferreiros, J., Pardo, J. M., 2012. Automatic categorization for improving Spanish into Spanish Sign Language machine translation. Computer Speech & Language 26 (3), 149–167.

Marimon, M., 2010. The Tibidabo Treebank. Procesamiento del Lenguaje Natural 45, 113–119.

Marimon, M., Fisas, B., Bel, N., Arias, B., Vázquez, S., Vivaldi, J., Torner, S., Villegas, M., Lorente, M., 2012. The IULA Treebank. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC-2012). Istanbul, Turkey, pp. 1920–1926.

Massó, G., Badia, T., May 2010. Dealing with sign language morphemes in statistical machine translation. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Languages Technologies. Valletta, Malta, pp. 154–157.

MEC, 2000. Serie informes - CNICE. MEC.

Minguet, E. (Ed.), 2000. Signolingüística. Itroducción a la lingüística de la LSE. Fundación FESORD, Valencia, Spain.

Morrissey, S., August 2008. Data-driven machine translation for sign languages. Ph.D. thesis, University of Dublin.

Morrissey, S., 2011. Assessing three representation methods for sign language machine translation and evaluation. In: Proceedings of the 15th Annual Meeting of the European Association for Machine Translation (EAMT-2011). Leuven, Belgium, pp. 137–144.

Morrissey, S., Way, A., May 2006. Lost in translation: the problems of using mainstream MT evaluation metrics for sign language translation. In: Proceedings of the 5th SALTMIL Workshop on Minority Languages at LREC-06. Genoa, Italy, pp. 91–98.

Neidle, C., 2002. Signstream: A database tool for research on visual-gestural language. Sign Language and Linguistics 4 (1/2), 203–214.

Nivre, J., Hall, J., Nilsson, J., Chanev, A., Eryigit, G., Kübler, S., Marinov, S., Marsi, E., 5 2007. MaltParser: A language-independent system for data-driven dependency parsing. Natural Language Engineering 13, 95–135.

Och, F. J., Ney, H., March 2003. A systematic comparison of various statistical alignment models. Computational Linguistics 29 (1), 19–51.

Padden, C., 1990. Sign language research: Theoretical issues. Gallaudet University Press, Washington DC, USA.

Papinemi, K., Roukos, S., Ward, T., Zhu, W.-J., July 2002. BLEU: A method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-2002). Philadelphia, Pennsylvania, USA, pp. 311–318.

Porta, J., López-Colino, F., Colás, J., November 2012. A computational morphology for Spanish Sign Language nominal inflection. In: Proceedings of IberSPEECH-2012. Madrid, Spain.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Herming, J., 1989. HamNoSys. Version 2.0; Hamburg notation system for sign languages. An introductory guide. Signum-Verlag.

Real Academia Española, 2001. Diccionario de la lengua española, 22nd Edition. Espasa, Madrid.

Rodríguez González, M., 1992. Lenguaje de signos. Fundación CNSE, Barcelona.

San-Segundo, R., López-Ludeña, V., Martín, R., Sánchez, D., García, A., May 2010. Language resources for Spanish - Spanish sign language (LSE) translation. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Languages Technologies. Valletta, Malta, pp. 208–211.

San-Segundo, R., Montero, J. M., de Córdoba, R., Sama, V., Fernández, F., D'Haro, L., López-Ludeña, V., Sánchez, D., García, A., 2011. Design, development and field evaluation of a Spanish into sign language translation system. Pattern Analysis & Applications, 1–22.

Sandler, W., Lillo-Martin, D., 2006. Sign Language and Linguistic Universals. Cambridge University Press, New York, USA.

Santana-Suárez, O., Carreras-Riudavets, F., Pérez-Aguiar, J., 2004. Relaciones morfoléxicas sufijales para el procesamiento del lenguaje natural. Editorial MILETO, Madrid.

Santana-Suárez, O., Carreras-Riudavets, F., Pérez-Aguiar, J., 2005. Relaciones morfoléxicas prefijales para el procesamiento del lenguaje natural. Editorial MILETO, Madrid.

Santana-Suárez, O., Carreras-Riudavets, F., Pérez-Aguiar, J., 2006. Relaciones morfoléxicas parasintéticas para el procesamiento del lenguaje natural. Editorial MILETO, Madrid.

Schembri, A., 2003. Perspectives on classifier constructions in sign languages. Psychology Press, Ch. 1: Rethinking 'classifiers' in Signed Languages, pp. 3–34.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., Makhoul, J., 2006. A study of translation edit rate with targeted human annotation. In: Proceedings of Association for Machine Translation in the Americas. pp. 223–231.

Speers, D. L., December 2001. Representation of American Sign Language for machine translation. Ph.D. thesis, Georgetown University.

Stein, D., Schmidt, C., Ney, H., Dec. 2010. Sign Language Machine Translation Overkill. In: Federico, M., Lane, I., Paul, M., Yvon, F. (Eds.), International Workshop on Spoken Language Translation. Paris, France, pp. 337–344.

Stokoe, W. C., 1960. Sign language structure: An outline of the visual communication systems of the American deaf. Studies in Linguistics: Occasional Papers 8, 1–78.

Stroppa, N., Way, A., November 2006. MATREX: DCU machine translation system for IWSLT 2006. In: Proceedings of the International Workshop on Spoken Language Translation (IWSLT-2006). Kyoto, Japan, pp. 31–36.

Supalla, T., 1986. The classifier system in American Sign Language. In: Craig, C. (Ed.), Noun Classes and Categorization. No. 7 in Typological Studies in Language. John Benjamins, pp. 181–214.

Sutton, V., December 1973. Sutton movement shorthand, the classical ballet key, key one. Sutton Movement Writing Press.

Sutton, V., 1974. SignWriting.

Sutton-Spence, R., Woll, B., 1999. The linguistics of British Sign Language: An introduction. Cambridge University Press, Cambridge, UK.

Varga, D., Halácsy, P., Kornai, A., Nagy, V., Németh, L., Trón, V., 2007. Parallel corpora for medium density languages. In: N. Nicolov, K. Bontcheva, G. A., Mitkov, R. (Eds.), Recent Advances in Natural Language Processing IV. Selected papers from RANLP-05. John Benjamins, pp. 247–258.

Veale, T., Conway, A., June 1994. Cross modal comprehension in ZARDOZ an English to sign-language translation system. In: Proceedings of the 7th International Workshop on Natural Language Generation (INLG-1994). Kennebunkport, Maine, USA, pp. 249–252.

Villameriel, S., June 2008. Marcadores del discurso en la lengua de signos española y en el español oral: Un estudio comparativo. Ph.D. thesis, Universidad de León.

Vossen, P. (Ed.), 1998. EuroWordNet. A multilingual database with lexical semantic networks. Kluwer Academic Publishers, Dordrecht, Holland.

Wittenburg, P., Levinson, S., Kita, S., Brugman, H., May 2002. Multimodal annotations in gesture and sign language studies. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002). Las Palmas, Spain, pp. 176–182.

Zesman, U., 2007. Sign languages: Spinning and unraveling the past, present and future. Petrópolis: Editora Arara Azul., Ch. Roots, leaves and branches – The typology of sign languages, pp. 671–695.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N. I., Palmer, M., October 2000. A machine translation system from English to American Sign Language. In: Proceedings of the 4th Conference of the Association for Machine Translation in the Americas on Envisioning Machine Translation in the Information Future (AMTA-2000). Cuernavaca, México, pp. 191–193.

7. Vitae

Jordi Porta obtained his B.Sc. in Computer Science from the Universitat Politècnica de Catalunya in 1994. From 1994 to 1996, he worked for the Grup de Investigació en Lingüística Computacional (GilcUB) of the Fundació Bosch i Gimpera–Universitat de Barcelona participating in several national and European projects. In 1996, he joined the Real Academia Española, where he works as a staff engineer within the area of Computational Linguistics. Since 1999 he has been teaching part-time in the area of Computational Science and Artificial Intelligence at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid. He joined the Human Computer Technology Laboratory (HCTLab) in 2009 to carry out research in the fields of Sign Language Processing and Machine Translation.

Fernando López-Colino received his Bachelor degree in Computer Science from the Universidad Autónoma de Madrid in 2005, and a Ph.D. in Computer Science and Telecommunication Engineering from the same university in 2009. In 2005 he joined the Human Computer Technology Laboratory (HCTLab), focusing his research on Sign Language processing and synthesis as applied to mobile devices. In 2011, his Ph.D. dissertation received the *IMSERSO Infanta Cristina* award for its relevance to the Deaf community.

Javier Tejedor received a B.Sc. in Computer Engineering, M.Sc. in Computer and Telecommunication Engineering, and Ph.D. in Computer and Telecommunication Engineering in 2002, 2005, and 2009 respectively from the Universidad Autónoma de Madrid. He has been with the Human Computer Technology Laboratory (HCTLab) since 2001 in the School of Computer Engineering and Telecommunication at that same university, and is an assistant professor in the School of Computer Engineering and Telecommunication there. His main interests are speech indexing and retrieval, spoken term detection, and large vocabulary continuous speech recognition.

José Colás has been a professor in the area of Computer Architecture and Technology since 2002. He received his Bachelor degree in Telecommunication Engineering from the Universidad Politécnica de Madrid in 1990, and his Ph.D. in Telecommunications from the same university in 1999. In 1993, his group received the *Reina Sofia* award for a research trajectory focused on technologies for the disabled. In 2001 he founded the Human Computer Technology Laboratory (HCTLab) at the Universidad Autónoma de Madrid. In 2003, this group received the *Infanta Cristina* award for research related to new technologies for the disabled, with a focus on mobile devices. He is the head of the "Multimodal Interaction Oriented to Disabled People" research unit at the HCTLab.

Figure 1

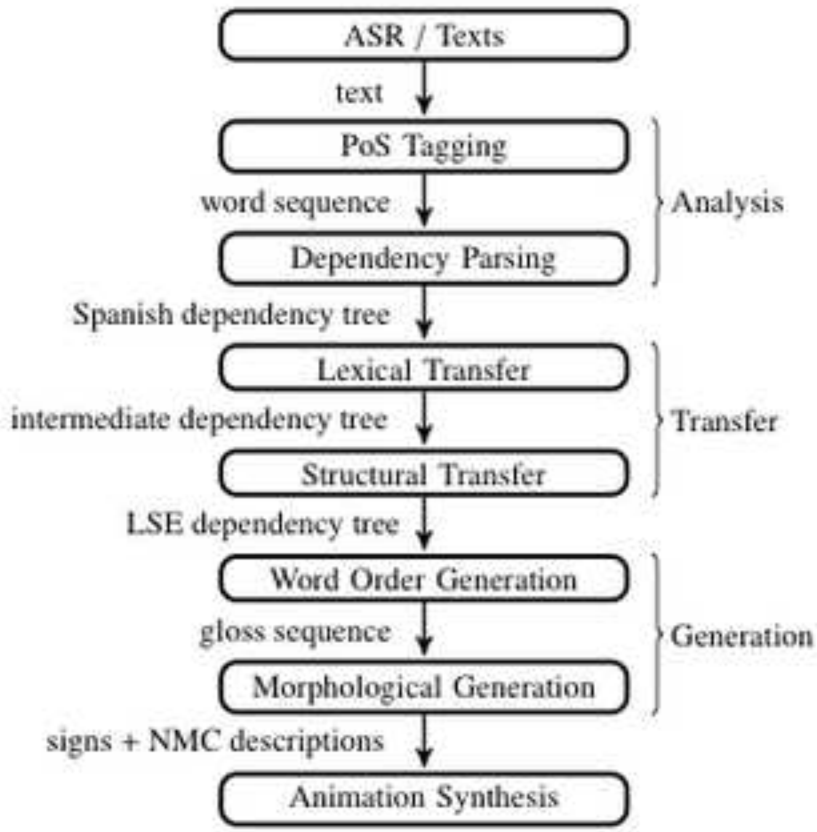
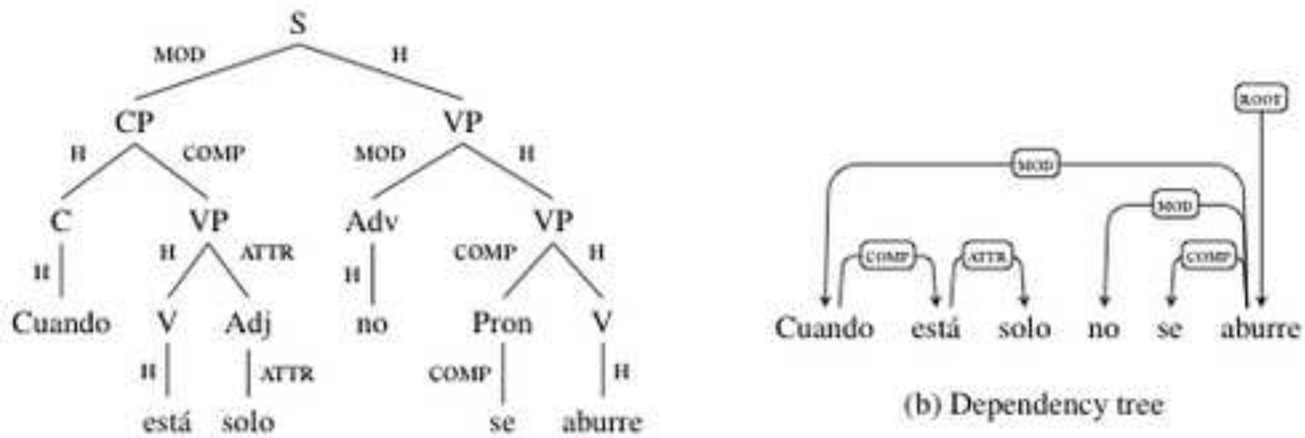


Figure 1: Architecture of the Spanish-to-LSE Machine Translation System. 'ASR' stands for Automatic Speech Recognition.

Figure 2

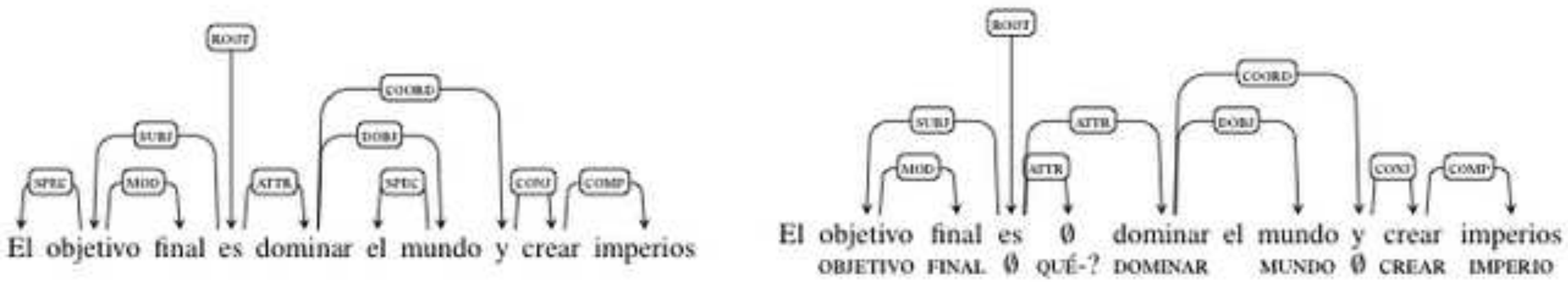


(a) Constituency tree

(b) Dependency tree

Figure 2: Two kinds of analyses for '*Cuando está solo no se aburre* (When he is alone he does not get bored)'. Syntactic functions are labelled in the constituency tree branches of (a) and on dependency arcs of (b), where arrows go from heads to dependents.

Figure 3

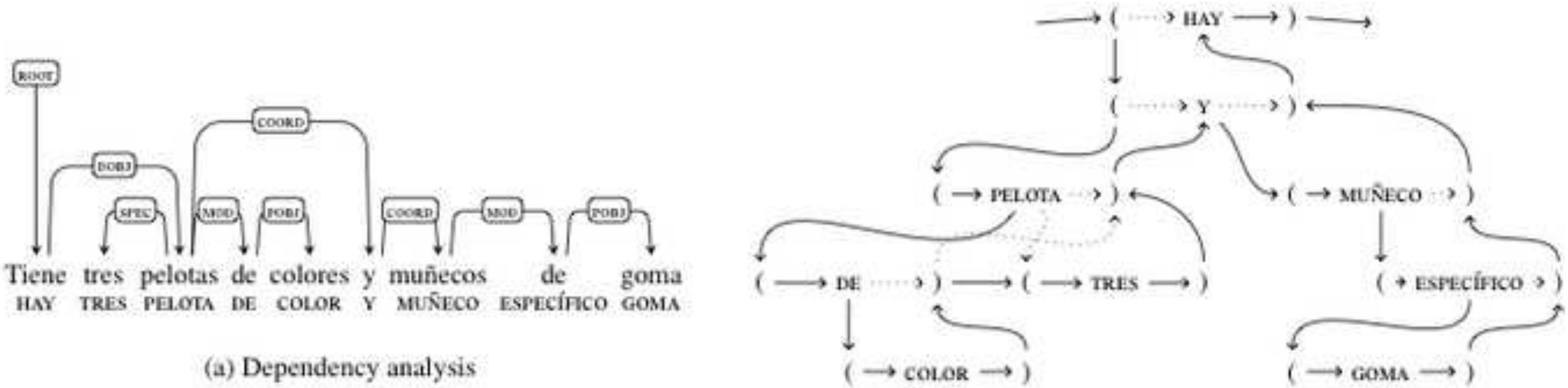


(a) Spanish dependency analysis

(b) Intermediate dependency analysis

Figure 3: Spanish dependency analysis for '*El objetivo final es dominar el mundo y crear imperios* (The final objective is world domination and empire creation)' and its transferred analysis to LSE where some nodes are disconnected and some nodes appear. The transferred sentence has been converted into a pseudo-interrogative sentence, which is a proper way to translate copulative sentence.

Figure 4



(a) Dependency analysis

(b) Precedence graph

Figure 4: Dependency analysis of '*Tiene tres pelotas de colores y muñecos de goma* (He has three coloured balls and rubber dolls)' and precedence graph from its output from the transfer stage. The circuit outlined with solid lines represents a topological ordering of the graph that will result in the gloss sequence 'PELOTA DE COLOR TRES Y MUÑECO ESPECÍFICO GOMA HAY (lit. ball of colour three and dolls specific rubber there-is)'.