Journal of Information Technology Research
Volume 12 • Issue 2 • April-June 2019

Designing High Accuracy Statistical Machine Translation for Sign Language Using Parallel Corpus:
Case Study English and American Sign Language

Achraf Othman, Research Lab. LaTICE, University of Tunis, Tunis, Tunisia
Mohamed Jemni, Research Lab. LaTICE, University of Tunis, Tunis, Tunisia

ABSTRACT

In this article, the authors deal with the machine translation of written English text to sign language. They study the existing systems and issues in order to propose an implantation of a statistical machine translation from written English text to American Sign Language (English/ASL) taking care of several features of sign language. The work proposes a novel approach to build artificial corpus using grammatical dependencies rules owing to the lack of resources for sign language. The parallel corpus was the input of the statistical machine translation, which was used for creating statistical memory translation based on IBM alignment algorithms. These algorithms were enhanced and optimized by integrating the Jaro–Winkler distances in order to decrease training process. Subsequently, based on the constructed translation memory, a decoder was implemented for translating English text to the ASL using a novel proposed transcription system based on gloss annotation. The results were evaluated using the BLEU evaluation metric.

KEYWORDS

Annotation System, Communication, Deaf, Machine Translation, Natural Language Processing, Parallel Corpus, Sign Language Processing, Statistical Machine Translation

INTRODUCTION

We can easily exchange our ideas, collaborate, and build strategies together, if we could speak all languages. Alternatively, we should have a communication tool that allows such innovations. Communication is the essence of human interaction. One of the most effective approaches of communication between human beings is through languages, which were born from human interaction. Communication and language exhibit a relation of interdependence. However, this natural method of communication can establish an obstacle in the cases where languages are postponed. Nobody can ignore the obvious barrier of communication between languages of different modalities, worth knowing the vocal and sign languages.

DOI: 10.4018/JITR.2019040108

In fact, Sign Languages (SLs), used by the deaf communities, which do not hear or hear badly, are visual-gesture languages. The message is transmitted by gestures and movements and received by the visual channel. The vocal languages, as for them, exhibit an audio-phonatory character. Their message is emitted through a phonatory canal and received thanks to the auditory canal. Canals used by the SLs are thus different from those used typically, and the SLs distinguish them from other languages. The linguistic knowledge acquired after several years of research and studies on diverse vocal languages are with difficulty transposable in the SL. Therefore, SL becomes a novel object of linguistic study, and it started from the 1960s. This research field is thus more recent, and it explains that the knowledge of the linguistic research on SL is constantly evolving.

This paper concerns NLP and is interested, in particular, in the case of ASL. Exotic by their implementation of gestures/movements and not sounds, these leave the traditional phonological frame and do not have a standard phonetic script similar to the vocal languages to transcribe their realizations. Moreover, exotic by their multi-linearity, they utilize their common canal to convey several elements of information simultaneously, whereas the voice device allows the production of only a sound at a time. The language models raise natural questions that are different from those suggested by the vocal languages.

Transcription is the operation that substitutes a grapheme or a group of graphemes of a writing system for every phoneme or for every sound. It thus depends on the target language, a unique phoneme that can correspond to various graphemes following the considered language. In short, it is the writing of words or pronounced sentences in a given system. The transcription also aims at being without loss, so that it should be ideally possible to reconstitute the original pronunciation from this one by knowing the rules of transcription.

From all cited perspectives, this paper concerns the transcription of SL and the implementation of a Statistical Machine Translation and concentrates more particularly on the machine translation of a text in English to ASL and conversely. The study produced articulates around four axes:

- Presenting an overview about SL and a state of the art about Sign Language Processing;
- Proposing a novel transcription system to write SL called XML-Gloss Annotation System;
- Building an artificial parallel corpus between English and ASL using XML-Gloss Annotation System;
- Implementing a Statistical Machine Translation between English and ASL based on the Artificial Parallel Corpus.

WebSign Project

This work is a part of the WebSign Project (Chabeb et al., 2008) which is a project developed within Research Laboratory LaTICE of the University of Tunis. The predominant objective of this project is to establish a communication system for the enhancement of the accessibility of deaf people to information. This tool allows to increase the autonomy of the deaf and does not require non-deaf to acquire special skills to communicate with them. This project is a multi-community, and it offers the possibility of registering a sign in several languages. In fact, it incorporates an interactive interface that allows the creation of dictionaries. The synthesis is based on a virtual avatar (Jemni et al., 2008). WebSign integrates a system of transcription Sign Modeling Language (SML). The synthesis kernel enabled them to develop several applications (Othman et al., 2010).

WHAT IS SIGN LANGUAGE?

SL are visual-gesture languages. They are considered by the community as the standard language for the deaf communities, where the message is conveyed by gestures and received by the visual channel. Vocal languages are audio–vocal languages. Their message is transmitted via the vocal canal and received through the ear canal. SL provides all the functions of the other vocal natural languages. SL is a language in itself, compared to the English or any language, which comprises all the features of natural languages and can be analyzed similar to any other language. Therefore, SL has a system of communication and representation. Thus, it exhibits the flexibility to create any novel vocabulary and new necessary grammatical structure and full capacity of abstraction and expression. It is transmitted from one generation to the next generation; there is not only a universal sign language, but also several SL, each being developed in a unique context. The main parameters in SL are:

- Gestures (Stokoe, 2005)
- Hand configuration and Orientation
- Localization and Movement
- Non-manual gestures
- Classes (Standard Signs, Shape and Size Specifiers, and Classifiers)

RESEARCH ON SIGN LANGUAGE

Research on SL covers many areas such as history, sociology, and anthropology. Linguistics studies have developed significantly. The dissemination of information and the technological advances in SL require, as for any other language, to develop software tools that are dedicated. Specific seminars and workshop are organized for presenting trends of research on SLs. This section presents the advances of existing tools and projects toward enhancing the communication between deaf communities or between non-deaf persons.

SL Resources

SL resources (called Corpora or Corpus) are a collection of data that are used for processing SL in machine translation, data mining, etc. Generally, the collection of resources is performed through national and international projects by specifying the characteristics of the resulting corpus. The predominant objectives are as follows:

- Data rate in words or sentences or hours;
- Policy of accessibility to resources;
- Number of participants (deaf, interpreters, or other);
- Transcription system (Gloss, HamNoSys, etc.);
- Lemmatization;
- Annotation tools;
- Liddel Model (Valli et al., 2000);
- Quality of the recorded videos.

Despite the initiatives for the collection of resources, still there is no large corpus that is useful for automatic processing and more particularly machine learning. This limit is due to the high cost of the production of signs by signer (Su et al., 2009). In the literature, there are many corpora for each community and they are not limited:

- American Sign Language (Neidle et al., 1997);
- British Sign Language (Schembri et al., 2008);
- Chinese and Taiwanese Sign Language (Chiu et al., 2007);
- German Sign Language (Hanke et al., 2004);
- Irish Sign Language (Bungeroth et al., 2008);
- Spanish Sign Language (San-Segundo et al., 2009);
- Unified-Arabic Sign Language (Abdel-Fattah et al., 2005).

Machine Translation for Sign Language

After citing different corpora for different sign languages, authors notice that most of the studies are on linguistic structures that are conducted in the first place to the generation of transcription systems based on linear decomposition of gestural signs in various characteristic parameters (hand shape, orientation, location, movement, eye gaze, and facial expressions). The SL interpretation is a complex process, which involves many cognitive tasks in parallel (watching, understanding, searching for equivalents, reformulation, control, etc.). Interpretation is intended to be an accessibility tool for communication with deaf communities.

The interpretation of the SL is invariantly limited by the presence of an expert interpreter for SL. A synthesis system can solve this problem. The synthesis of SL is a system based on 2D or 3D animation. Two approaches are being explored for the generation of SL animation: one is based on the video and the other on 3D synthesis through conversational agent or avatar. Several systems were developed for SL. Some conversational agents are valid for any SL (multi-community avatars), and others are specific to one SL in order to highlight one or more grammatical features. Furthermore, the goals of synthesis are multiple. Some conversational agents are dedicated for learning and others for machine translation, etc. In the literature, there are several projects such as the following:

- VisiCast (Elliott et al., 2004)
- eSIGN (Hanke et al., 2003)
- TESSA (Cox et al., 2002)
- Works of Matt Huenerfauth (Huenerfauth et al., 2005)
- Vcom3D (DeWitt et al., 2003)

All these synthesis tools for SL requires a translation stage before the interpretation since the user provides a sentence as an input. Moreover, this entry should be processed and translated to an annotation system toward synthesizing it using an avatar. Moreover, it is noted that the quality of interpretation or synthesis depends on the dictionary, which invariantly uses the motion sensors in order to have the positions of the articulations in the space of the conversational agent. Moreover, different transcription systems were used in order to construct the translation for the avatar or for a textual display. There are several machine translations to or from SL or between SLs:

- Albuquerque Weather (Grieve-Smith et al., 2008)
- TEAM (Zhao et al., 2000)
- VisiCast Machine Translation (Bangham et al., 2000)
- ASL Workbench (Liddel et al., 2003)
- SASL (Olivrin et al., 2008)
- Spanish SL (San Segundo et al., 2008)
- MaTrEx (Morrisey et al., 2013)
- Studies of Hung-Yu Su and Chung-Hsien Wu (Su et al., 2009)

Transcription Systems for Sign Languages

In 1960, Stokoe's study (Stokoe et al., 2005) showed that SLs are real and independent languages and have argued that SLs are the mother tongue of deaf people. Before its function, the SLs were not considered as real languages. They were observed as a set of meaningless gestures. Then, an SL could not be used in the education of deaf children. Afterwards, many SL notations were proposed. Some of them have become widely used by SL researchers, and others are used as teaching tools. SL can be written using gloss or notations. For example, this sentence "What is your name?" will be "name you what" in SL since the deaf person or the interpreter will not sign the word-for-word sentence, but rather reconstruct it according to the subjects, the objects, and the action generated by the verb. With regard to notation, the notation uses special symbols to describe the physical parameters of SL. Several studies bearing the writing of SLs are provided in the following: Gloss (Klima et al., 1979), HamNoSys (Hanke et al., 2004), Jouison (Jouison et al., 1990), Laban (Laban et al., 1975), Movement-Hold by Liddell et Johnson (Liddel et al., 1989), Newkirk (Newkirk et al., 1989), SiGML (Elliott et al., 2004), SignWriting (Stuart et al., 2011), SML (Jemni, El Ghoul et al., 2008) and Stokoe (Stokoe et al., 2005).

The representation of sign language in the written form is necessary for the generation when conversational agents are used. Although this problem can be solved using one of the SL notations discussed earlier, it should be noted that none of these annotations were accepted as a standard written form of sign language. In addition, each annotation is performed for a particular purpose and study.

Video Synthesis for Sign Language

Video synthesis is typically used in the SL dictionaries. Furthermore, many researchers investigate on SL video synthesis and uses the transcription techniques in websites to guide users and allow them to easily navigate. It is easier to record a large number of gestures in the SL by simply matching the video to a sentence. For a searched word in a dictionary, the tool displays the video of the corresponding entry, or for a sentence, it is sufficient to position the videos of elementary signs in sequences in order to create a phrase in the target language of the signs. Hence, responsible signatories, deaf or performers, produce natural and realistic sequences containing manual and non-manual functionalities that are the predominant advantage of this approach. In addition, the videos can be annotated using the tools mentioned previously, which increases the understanding of the sequences.

Two challenges emerge owing to the difficulties of concatenating these videos. The first challenge is to ensure that the transitions between the signs in each video are flexible and that the facial expressions and parameterization of the magnitude or speed of certain signs should be appropriately integrated with the corresponding signs. The second challenge is to manage the changes in hand configurations according to the context called “predicate classifiers” and their trajectories that are not known in advance.

3D Synthesis for Sign Language

A conversational agent or avatar is a virtual character in 3D. Avatar-based viewing systems use avatars rather than human signatories. They devote a more complex procedure to translate the vocal languages in sign languages. To perform this, they use a text annotation for the animation as an input. This visualization has become the fastest solution for evaluating the results of machine translation as well to solve the problems posed by visualization using videos. First, avatar-based visualization systems are completely controllable in their movements and speed of execution through simple commands. Second, the movements can be completely described in a considerably compact language consisting of control commands. Two major challenges are posed when using avatars. The first is to model a sufficiently articulated avatar to perform SL gestures. This should be as natural as possible. It should also be able to perform such subtle gestures as changes in facial expressions, such as smiling, raising eyebrows, or frowning. The second challenge is to animate the avatar. The avatar movements should be physically plausible and realistic in order to be interpreted and understood by humans.

MOTIVATION AND CONTRIBUTIONS

Studies on SLs are recent and innovative and at the same time are numerous, ranging from studies on the linguistic, cognitive, and grammatical aspects to the creation of corpus, machine translation, and synthesis in real time. This paper is a contribution for developing linguistic resources in SL in order to process them for machine translation, from an English text to ASL or reverse. This mission is not easy, because the problems mentioned earlier are numerous or several studies have reached satisfactory results. In fact, for processing SL, one should invariantly study a written form for representation and modeling. Despite the existence of scoring and annotation tools, each has disadvantages. However, for the text annotation in gloss, which exhibits more advantages for automatic processing, there is no detailed XML representation of this system for storage and animation via signing avatars. Hence, researchers created a new annotation system based on gloss, in connection with our particular requirements for machine translation to and from SL.

XML-Gloss Annotation System

In this framework, a novel annotation for the ASL is presented based on the annotation system in gloss. A gloss is defined as a linguistic comment added in the body of a text or a book, or in its margin, explaining a foreign or dialectal word, a rare term. XML is a general format of text-oriented documents. Owing to its simplicity, flexibility, and expansion possibilities, it is possible to adapt it to multiple domains. It will be used in our transcription system which is defined in detail. Also, a complete example is provided followed by a validation schema and a style sheet for rendering on browsers using related languages to XML.

The conventions described here reflect the annotation conventions we used for transcribing data. We refer to all the conventions described in the Liddell conventions (Liddel et al., 1989). Similarly, we provide a set of fields and values that are particularly designed for the ASL data annotation. Each gloss is represented by an English word, which is written in uppercase letters. This notation, although simplified, does not reflect the morphological richness of the ASL signs alone, but also the considerably important grammatical function of the facial expressions. The non-manual signs transmitted by the face can occur simultaneously with the manual components. They are represented by a line above the signs. The authors annotate all the important characteristics of SL grammar. A single English word in uppercase identifies a single ASL sign, for example, the sign of a cat will be "CAT." Capitalized words in English separated by hyphens also represent a single sign, that is, when more than one English word is required for translating a single sign. Let us take the case of an ASL transcription of the phrase "His aunt lived in Turkey. There had been no contact with the aunt. She died and left something to him in her will" which is shown in Figure 1 (Speers et al., 2002). At first glance, the transcript has no relation to the English language; however, it is not true. In fact, observing the signs through images or a video sequence alone is not sufficient, because the annotation in gloss becomes indispensable. The annotation in gloss is shown in Figure 1.

The glosses here are not simple words in English. Hyphens between uppercase letters indicate a sequence of signs of alphabetic characters used to spell a word letter by letter. For example, if the interpreter wishes to spell the sign "JOHN," he only has to represent it in the form "J-O-H-N". A gloss that begins with the symbol "#" represents an ASL lexical sign that started as a sequence of alphabetic character signs, particularly when the word is unknown, such as "#KATE." Our XML-Gloss Annotation focuses on the grammar of the ASL and the procedure to annotate each component in glosses. This study allows us to investigate all the morphological aspects in order to design a model in XML, which is used for processing SL. From this annotation convention, authors can then proceed to study the semantics between subjects and objects and their locations in the space of the trainer. The whole description and conventions of the XML-Gloss Annotation System was published in (Othman et al., 2013).

Figure 1. Gloss transcription of the sentence "His aunt lived in Turkey. There had been no contact with the aunt. She died and left something to him in her will" in ASL.

$$\overline{\text{POSS}^{\to[\text{MAN}]}\ \text{AUNT}}^{\,t}\ \text{LIVE THERE}^{\to[\text{T-U-R-K-E-Y}]}\ \text{T-U-R-K-E-Y}\quad \text{PRO}^{\to[\text{AUNT}]}\ \text{CONTACT ZERO}$$

$$\text{DIE}^{\to[\text{AUNT}]}\ \text{PRO}^{\to[\text{AUNT}]}\ \text{GIVE-OWNERSHIP}^{\to[\text{AUNT}]\to[\text{MAN}]}\ \text{W-I-L-L GIVE-OWNERSHIP}^{\to[\text{MAN}]}$$

Building Artificial Parallel Corpus Using Dependency Grammatical Rules

In this section, a novel approach to the generation of discourse in the ASL is presented in order to translate English text and at the same time build a parallel corpus between these two languages. The proposed approach was featured in the VisiCast project. However, it is limited to the syntactic level. In this study, authors incorporate more than 52 relations of grammatical dependencies during the generation. This allowed us to generate the non-manual components. Research on the analysis of lexical, syntactic, morphological, and semantic of an English text is considerably advanced, and authors find several tools that provide results with rates near 100% accuracy and a recall (recall) close to 90%. The approach is divided into two predominant parts: the first is the automatic processing of the provided input text and represent it as a semantic graph. The second step is the generation of the XML-Gloss transcript using the APIs.

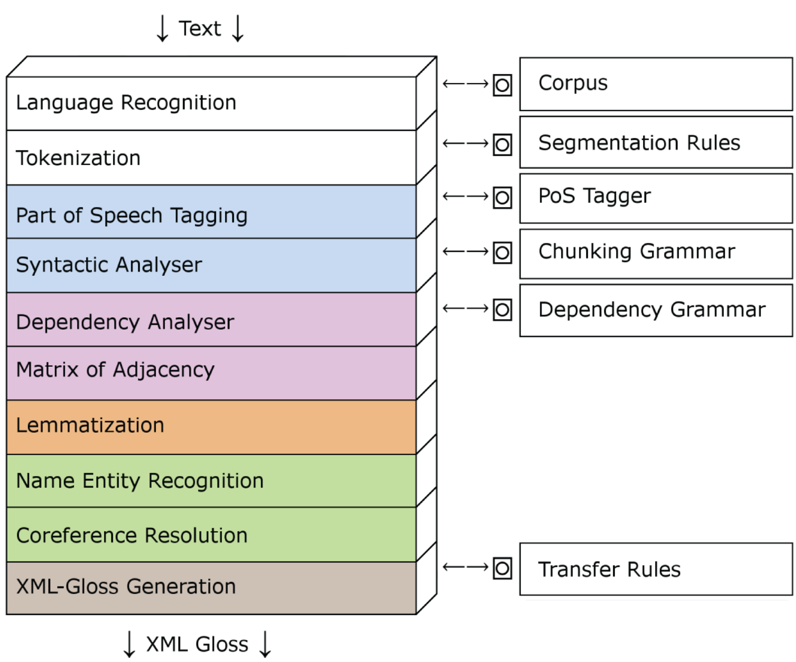
The approach uses a syntactic analysis coupled with a representation of the meaning of the words and a comparison of the words emerging from the first method. The approach implements the various notions of a thesaurus of ideas as vector space to recover the ideas of the meanings of the words provided in a semantic lexicon. The ideas of each word, manipulated in the form of vectors, are developed from the leaves of the syntactic structure to the top points to obtain the representational ideas of the sentences and texts in processing. The vectors associated with the text then represent the context of the text.

Architecture of Proposed System

The organization of the levels of linguistic treatment of our system is similar to those of the levels of the triangle of Vauquois. Thus, as shown in Figure 2, the system for analysis is organized according to three predominant levels: lexical, syntactic, and semantic, the software blocks of which constitute a chain of linguistic treatments.

The assembly is structured around several modules supervised by a control module. These modules contain the different data according to which the analysis and the generation of message are produced.

Figure 2. Overview of proposed system

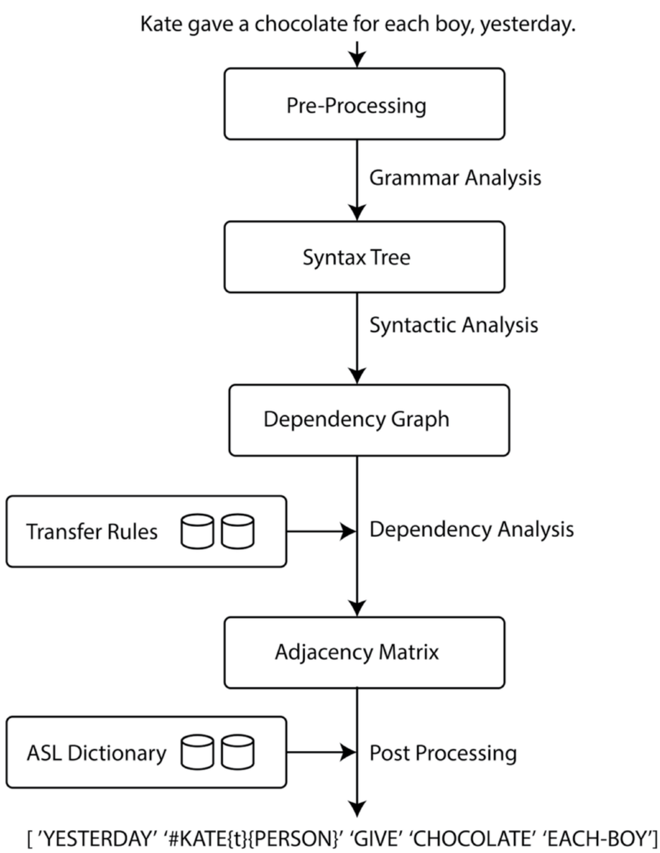


During the implementation process, lexicons, grammar, stroke translation, and segmentation data are compiled before they are used by the system. As far as its functioning is concerned, the proposed system takes a text as input, segments it into paragraphs, sentences, and words and generates several levels of analysis: Minimal analysis, Chunking, and, Semantic analysis.

Overview

Let us return to the translation chain of the proposed system by simulating the processing of an example. The analysis that is performed in a classical manner following the diagram presented in the previous figure, of which we transformed the output, which is the statement obtained in the target language for the other languages. Therefore, after a segmentation step, the morphological, syntactic, and semantic information of each word of the source utterance is searched in a lexicon. A dependency grammar is then used for constructing the tree representing the syntactic structure of the utterance. The tree obtained for the analysis of a sentence similar to "Kate gave chocolate for each boy, yesterday" is provided in Figure 3. In fact, the structure of the target statement in the ASL is "time → subject → verb → object → object complement." The first step is a pre-processing applied directly to the input sentence followed by a dependency analysis between the words that will be described in the following sections. Then, authors analyze the semantic structures between the words. Following this analysis, authors generate a syntactic structure specific to ASL, which from this structure is linearized and formatted according to the XML-Gloss transcription system.

Figure 3. Overview of the functioning of the proposed translation system



Implementation

From the previous example, authors describe the implementation phase step by step. The entry phrase is "Kate gave chocolate for each boy, yesterday."

Preprocessing and Lexical Analysis

Preprocessing is a necessary and indispensable phase for converting the raw data into a format suitable for automatic processing in our ASL statement generation system. They invariantly begin by segmenting the text into sentences, and they perform the same pre-treatment for each sentence. The first pre-processing operation is called "tokenization," the objective of which is precisely to transform the input string into "tokens". This operation applies to the source texts and considers the spaces to separate the words, the numbers, and the punctuation. In our example, the tokens are:

1. ["Kate" ; "gave" ; "chocolate" ; "for" ; "each" ; "boy" ; "," ; "yesterday"]

Then, all characters are converted to lowercase. We obtain the following:

2. ["kate" ; "gave" ; "chocolate" ; "for" ; "each" ; "boy" ; "," ; "yesterday"]

Grammar Analysis

From the lexical analysis, authors proceed to the grammatical analysis of each token in order to construct the syntactic tree that allows us for semantic analysis. Syntactic analysis is the association of a grammatical category (noun, verb, adjective, adverb, proper noun, etc.) for each word of the input sentence. For our input sentence, we obtain the following:

3. "kate" → NNP
"gave" → VBD
"chocolate" → NN
"for" → IN
"each" → DT
"boy" → NN
"," → SYM
"yesterday" → NN
"." → SYM

For the labeling of grammatical categories, the "Stanford Log-linear Part-Of-Speech Tagger" tool (Toutanova et al., 2000) is used, which is based on the Maximum Entropy Model. The model assigns a probability for each grammatical category t of the set of grammatical categories T or a provided word w in its context h. The overall accuracy rate of the labeling is 96.72% for words known during learning and 84% for unknown words. For example, the precision rate of the grammatical category "IN" is 97.3%. The precision rate for the "VB" is 94% and that for the "NNPS" is 41.1%.

Syntactic Analysis

After determining the grammatical categories of each word of the sentence provided as input, authors now pass to the syntactic analysis in order to construct the syntactic tree. The syntactic tree (Abney et al., 1989) contains a node for each word. The role of the parser is to establish a connection between two nodes to create a novel node. The task of a dependency analyzer between nodes is to take a series of words and impose links on it. There are three strategies: brute force, exhaustive left-to-right search, and enforcing uniqueness. From dependency search algorithms, the dependency links between the words are established on the basis of the parent–child principle. This grammar is defined by lexical rules and internal rules. The lexical rules are extracted from the grammatical analysis of the previous step. For the internal rules of grammar, authors use the model of Collins (Collins et al., 2003). This probabilistic model is based on statistics extracted from a corpus where each sentence of the corpus has a syntactic tree according to a lexical grammar provided by a linguist. For the above example, the lexical rules are as follows:

4. NNP("kate", NNP) → kate
VBD("gave", VBD) → gave
NN("chocolate", NN) → chocolate
IN("for", IN) → for
DT("each", DT) → each
NN("boy", NN) → boy
SYM(",", SYM) →,
NN("yesterday", NN) → yesterday
SYM(".", SYM) → .

And internal rules are:

5. S → NP VP SYM
NP → NNP
NP → NN
NP → DT NN
PP → IN NP
VP → VBD NP PP SYM NP

From the lexical rules and the internal rules, the syntactic tree of the sentence "Kate gave chocolate for each boy, yesterday." is built. The tree is illustrated in Figure 4.

Dependency Analysis

The analysis of dependencies allows to label the grammatical relations provided by the syntactic analysis. The dependency relation between two words is binary: one part is called the "agent" and the other part is the "dependent." In order to extract relations, one relies on the study of Marneffe et al. (De Marneffe et al., 2006). Their representation contains 53 dependency relations. For each relation, authors append an index to identify the name of the relation. For example, tmod dependency relation is indexed 50. From relationships, they construct the dependency graph starting with the "root" relation, knowing that the relationships do not cover all the words, but rather the predominant words. Figure 5 illustrates the dependency graph of our example sentence.

Adjacency Matrix

From the dependency relations, authors define a finite graph G with *n* vertices, such that n is the number of words of the input sentence. The edge joining two vertices *i* and *j* is denoted $\left(V_i, V_j\right)$. The set of edges is *V*. For our example, the dependency graph is illustrated in Figure 6.

Adjacency matrix A of size $n \times n$ is determined from G. An element of the matrix a_{ij} (an edge between two vertices i and j) is defined as follows:

Figure 4. Syntax tree of the sentence "Kate gave chocolate for each boy, yesterday."

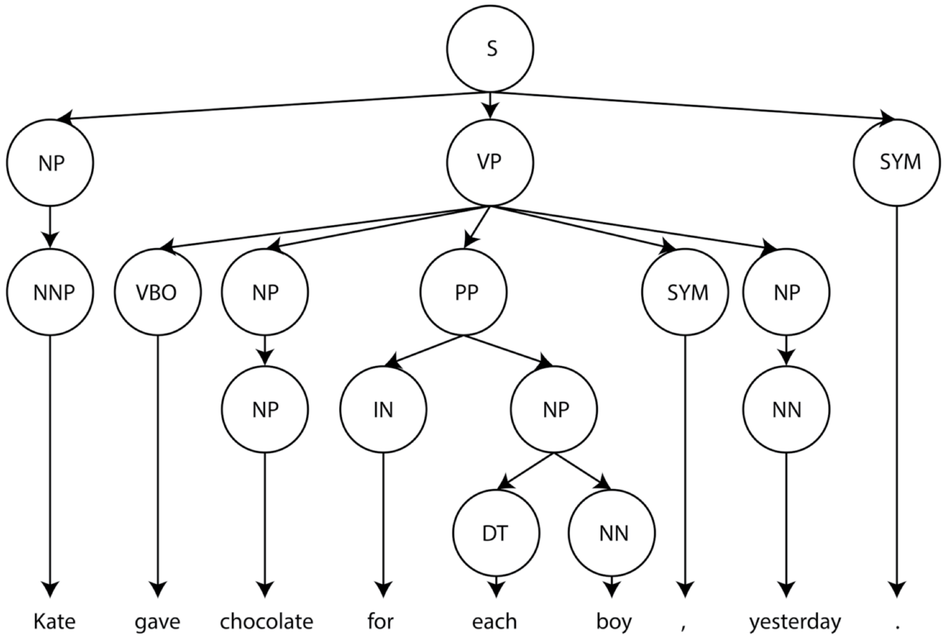
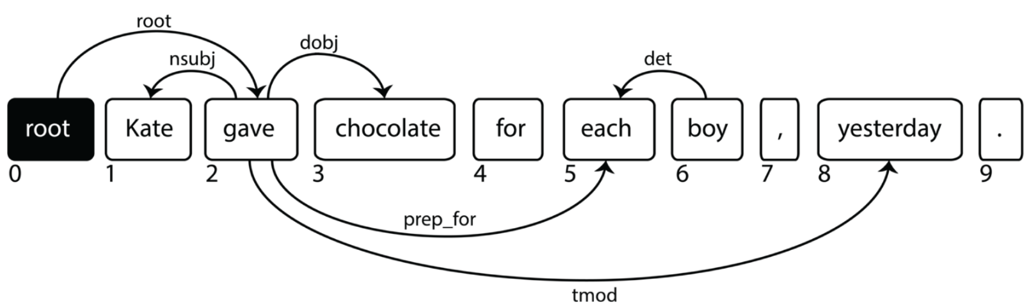


Figure 5. Dependency relations of the sentence "Kate gave chocolate for each boy, yesterday."



$$a_{ij} = \begin{cases} 0 & si\left(V_i, V_j\right) \notin E \\ > 0 & sinon \end{cases}$$

If the value of a_{ij} is strictly greater than zero, then its value corresponds to the index of a dependency relation. From this definition and from our example, authors construct the adjacency matrix of Figure7.

To generate the ASL statement, authors should define the output rule. In fact, the structure of the statement should respect the following rule:

6. " *tense* $\rightarrow$ *subject* $\rightarrow$ *verb* $\rightarrow$ *object* $\rightarrow$ *object compliment* ."

Figure 6. Sentence dependency graph “Kate gave chocolate for each boy, yesterday.”

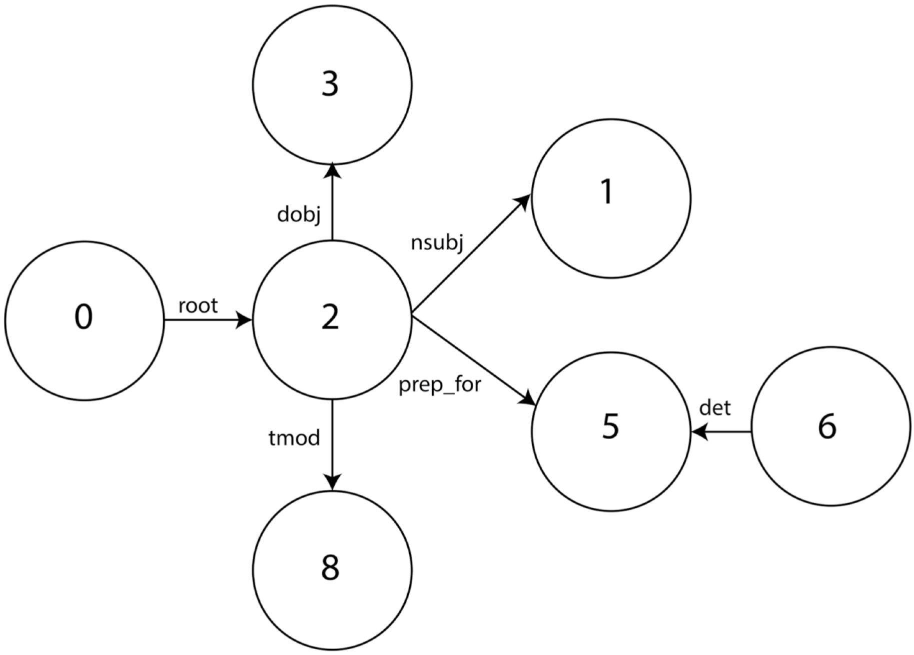


Figure 7. Adjacency matrix of the dependency graph for the sentence “Kate gave chocolate for each boy, yesterday.”

	00	01	02	03	04	05	06	07	08	09
00	0	0	0	0	0	0	0	0	0	0
01	0	0	28	0	0	0	0	0	0	0
02	49	0	0	0	0	0	0	0	0	0
03	0	0	19	0	0	0	0	0	0	0
04	0	0	0	0	0	0	0	0	0	0
05	0	0	40	0	0	0	18	0	0	0
06	0	0	0	0	0	0	0	0	0	0
07	0	0	0	0	0	0	0	0	0	0
08	0	0	50	0	0	0	0	0	0	0
09	0	0	0	0	0	0	0	0	0	0

From the adjacency matrix, authors extract the components that construct the utterance in order. For each component, all related words are retrieved from its column. For example, the components are extracted as follows:

7. "*tmod → nsubj → root → dobj → prep_for*"

The component "prep_for" also contains a recursive relation, which is the relation "det." Therefore, the order of extraction will be:

8. "*tmod → nsubj → root → dobj → prep_for+det*"

The result of the extraction algorithm is:

9. "*yesterday kate{t} gave chocolate each-boy*"

The suffix {t} is applied to the subjects of the sentence in order to mention the nonmanual component "topic."

The rules of transfer to the ASL from the relations of grammatical dependence are provided by the linguists. Furthermore, when generating the statement, it is necessary to specify in advance the type of the statement (SVO, SOV, etc.). This functionality, despite manual, provides the possibility for generating speech in sign languages in a more generalized manner.

Lemmatization and Formatting

Lemmatization associates a lemma with each word of the text. If a word cannot be lemmatised (number, foreign word, unknown word or its grammatical function is FW), then this word will be spelled ("fingerspelled") using the symbol #. Furthermore, all words will be converted to uppercase. For our example:

10. "kate" → Lemme: kate → KATE
"gave" → Lemme: give → GIVE
"chocolate" → Lemme: chocolate → CHOCOLATE
"each" → Lemme: each → EACH
"boy" → Lemme: boy → BOY
"yesterday" → Lemme: yesterday → YESTERDAY
The result is provided in the following:
11. "YESTERDAY KATE{t} GIVE CHOCOLATE EACH-BOY"

Name Entities Recognition NER

The task of NER is concerned with a certain number of particular lexical units, which are the names of persons, the names of organization, and the names of places, to which other phrases such as Dates, currency units, and percentages are frequently added. Its objective is twofold: on the one hand, to identify these units in a text, and, on the other hand, to categorize them according to the predefined semantic types. The result of these processes corresponds to the annotation of the entities, which most frequently materializes via tags enclosing the entity. This allows us to determine, for example, the names of the people in order to spell them during the transcription or identify a date or quantities (Nadeau et al., 2007). In our example, the phrase contains an entity named "person" that is the word "Kate." In this case, the transcription of this sign becomes "#KATE {t}," and the statement becomes:

12. "YESTERDAY #KATE{t} GIVE CHOCOLATE EACH-BOY"

Coreference Resolution

The coreference resolution is to find all the expressions that refer to the same entity in a text as observed in the previous section. Moreover, authors used the "Stanford Coreference Resolution" tool to determine the co-referential entities in a sentence (Raghunathan et al., 2010). The coreference resolution is used for introducing indices between the already extracted "root" entities: the subject and the objects. In this example, the entity "Nader" will be indexed by x and the verb "vote" will also be indexed by x and y, which refers to the first person in the singular. Then, the entity "he" will be replaced by a nonmanual component referring to x. Similarly, the entity "she" will be referred by y. The accuracy rate of the resolution of the coronations is 83.8% and the recall rate is 73%.

Generating XML-Gloss File

The generation of XML-Gloss transcription is a simple task. Simply route the entities resulting from the ASL construction algorithms and make a call to the API methods. Let us take the same example, the final statement is:

13. "YESTERDAY #KATE{t} GIVE CHOCOLATE EACH-BOY"

The different instructions for creating XML-Gloss file are:

(1)
```
g.create_sentence("en", "asl", "kate gave chocolate for each boy, yesterday.");
g.create_clause("none");
g.create_token("yesterday");
g.create_token("kate");
g. add_property("kate","t");
g.add_fingerspelling("kate");
g.create_token("give");
g.create_token("chocolate");
g.add_compound("each", "boy", "-");
save_file("sentence.xml");
```

The file will be as follows:

(2)
```
<?xml version= "1.0" encoding= "UTF-8"?>
<?xml-stylesheet type='text/xsl' href='form.xsl'?>
<bloc xmlns:xsi= "http://www.w3.org/2001/XMLSchema_instance" xmlns:noNamespaceSchema
    Location= "schemaXML.xsd">
<sentence src_lang= "en" lang= "asl" src_sentence= "kate gave chocolate for each boy, yesterday">
<clause type_clause= "none">
<token>YESTERDAY</token>
<token property= "t" fs= "yes">KATE</token>
<token>GIVE</token>
<token>CHOCOLATE</token>
<token compound= "-"><token>EACH</token><token>BOY</token></token>
</clause></sentence>
```

Moreover, the final output is shown in Figure 8.

Figure 8. Transcription in gloss of the sentence "Kate gave chocolate for each boy, yesterday."

YESTERDAY $\overline{\text{\#KATE}}^{\,t}$ GIVE CHOCOLATE EACH-BOY

Evaluation

The evaluation of the proposed approach is essential to ensure the quality of the transcription generated by our algorithms. To evaluate our system, authors manually compared each sentence and its transcription and generated transcription. However, this requires a long time, because the number of sentences exceeds 100k. The idea is to evaluate only the transfer rules between the two languages (English and ASL), and this considerably reduced the time and cost of the evaluation. In other words, let us consider a sentence e in English and its adjacency matrix M. Authors define a transfer rule R (i⇒j) by:

"tmod + nsubj + root + dobj + prep_for-det"
"T + S + V + O + CO"

At the evaluation level, only the transfer rule is checked for all sentences with an R[i] type structure. For our system, authors evaluated 820 transfer rules extracted from ASL Learning books with a precision rate equal to 82% for 6720 sentences calculated from the following formula:

$$T\left(precision\right) = \frac{count\left(valid\,sentence\right)}{count\left(sentence\right)} \times 100$$

Building Memory Translation for Statistical Machine Translation Between English/ASL

The approach of the ASL discourse generation from the rules of grammatical dependencies presents an interesting solution for the machine translation of a text into a transcription of the sign language. The overall architecture of this generation system was described with the different bricks constituting this system. This approach suffers from some limitations. The first is that the translation function is not bijective; in other words, only the generation of a text in English to the ASL is feasible. The second is that the evaluation is a complex task, because it is manual and no automatic approach is available.

Brown et al. (Brown et al., 1993) propose a probabilistic model according to which a sentence P in the source language has a possible translation T into the target language according to the probability p (T|P). This assertion can be interpreted as the probability that a human translator produces the target sentence T, knowing the source sentence P. Therefore, this model makes it possible to search for a sentence T, which is a possible translation of P by maximizing the probability p(T|P), according to observations performed in a parallel corpus. The SMT can therefore be defined according to the following equation:

$$\hat{T} = \max_{T} p\left(T \mid P\right) = \max_{T} p\left(P \mid T\right) \cdot P\left(T\right)$$

Building Parallel Corpus English-ASL

The problems encountered in creating the corpus in sign languages, because of their costs, are mentioned in the previous sections. The availability of parallel corpus thus poses problems, whether for the cover of the vocabulary or for the syntactic structures of the sentences to be translated. In addition, texts in specialized fields are generally more difficult to translate automatically, owing to the lack of corpus and the large number of non-vocabulary words. Automatic translation systems provide the best translations when there is sufficient size domain learning data. One of the methods to compensate for the lack of specialized data is the posteriori edition of translations from automatic systems. This revision stage conducted typically by humans ensures the quality of the translations and adapts them to the fields of specialization, if necessary (Krings et al., 2001). Because of this, the quality of automatic translations directly affects the post-publishing effort.

In this framework, authors used an approach to the generation of a parallel artificial corpus English-ASL. The idea emerges from generating, from a set of sentences in English, their respective transcriptions in the ASL annotated in XML-Gloss using the approach cited in the previous chapter. Then, a step of manual validation of the transfer rules by the users of the system, who are experts in the linguistic domain of the SL, is conducted. The English sentences were extracted from the Gutenberg corpus than contains more than 42k e-books and more than one hundred e-books collected from other partner sites. All these resources are free to access. The extraction of these resources was conducted using a tool that runs for each book and stores the sentences and the words in a database. This phase was conducted for over four stages.

Implementation and Experimentation

After constructing the English-ASL parallel corpus, authors implement our statistical automatic translator based on a sub-phrastic approach. This statistical approach currently yields prominent performances, and the measurements show the predominance of this system. We also adopt the methods proposed in the Moses machine translation toolkit in our study (Koehn et al., 2007). Moses proposes all the necessary tools for the construction of a translation model. A decoder also makes it possible to use these tools in order to produce hypotheses for translating an initial text. Several steps are necessary for constructing a translation system with Moses and use the various tools available. First, it is necessary to perform word alignment from the parallel corpus. Then, the set of alignments obtained forms a table that serves as a basis for the construction of the translation table or memory (Othman et al., 2013), which is described in the following section. The latter is elaborated according to the alignment scores obtained by the sub-phrastic segments present in the parallel corpus. Then, they constructed a re-scheduling model that contains the information of the positions in the sentences of the translated words with respect to the words translated previously. To estimate word alignment probabilities, Moses encapsulates the GIZA ++ tool (Och et al., 2003), which implements the algorithms of the IBM 1-5 models. Authors also use a parallelizable version of this program, MGIZA (Gao et al., 2008) in order to reduce the time during the alignment phase. Language models can be constructed using other tools, such as the one proposed by SRI-LM (Stolcke et al., 2002) or IRST-LM (Federico et al., 2008).

Building Lexical Translation Memory

In this section, the construction of an English-ASL translation memory from the ASL-PC12 corpus is described inspired by the IBM 1-5 models (Othman et al., 2012) (Berger al., 1994). The translation memory will be used in the decoding phase.

Lexical Translation

If we consider a source word in ASL "REVIEW," there are several possible translations in English: "review," "reviews," "reviewed," etc. Most words have several possible translations. The concept of SMT implies the use of statistics extracted from words and texts. What type of statistics is required to translate the word "REVIEW" into ASL or vice versa? If we assume that we have a large corpus or collection of data in the ASL coupled with a data collection in English, where for each sentence in English it corresponds to its ASL translation, we can count how many times it translates the word "REVIEW" into "review," "reviews," etc. Table 1 shows the possible results of the translation of the

Table 1. Probably translation of word "review"

Translation of "REVIEW"	Occurrence Count	Alignment Probability
reviewed	1	1.0000000
reviews	3	0.6666667
review	9	0.4000000
for	7216	0.0001840
$\varnothing$	-	0.0001261
of	10854	0.0001064
the	32608	0.0000326

word "REVIEW," knowing that the number of occurrences of the initial word in English is five times in the corpus ASL.

Here, authors note that the word "reviewed" was used only once and it was aligned with "REVIEW," which provides a translation probability equal to 1, in other words, in a novel translation request. For the others, when the number of occurrences increases, the probability of alignment decreases.

Probability Distribution Estimation

To calculate the probabilistic distribution of a word f, in our example the word "REVIEW," it is sufficient to determine the number of occurrences of this word f in the corpus in ASL and to calculate the occurrence of these possible translations in the corpus in English. Then, the ratio for each output e is calculated. They will therefore have:

$$P_f\left(e\right) = \begin{cases} 1.0000000 & \text{if e = 'reviewed'} \\ 0.6666667 & \text{if e ='reviews'} \\ 0.4000000 & \text{if e = 'review'} \\ 0.0001840 & \text{if e = 'for'} \\ 0.0001261 & \text{if e} = \varnothing \\ 0.0001064 & \text{if e = 'of'} \\ 0.0000326 & \text{if e = 'the'} \end{cases}$$

In this case, the choice of the best possible translation is the one with the greatest probability. The type of estimate used is called Maximum Likelihood Estimation, which maximizes the resemblance between the data.

Alignment

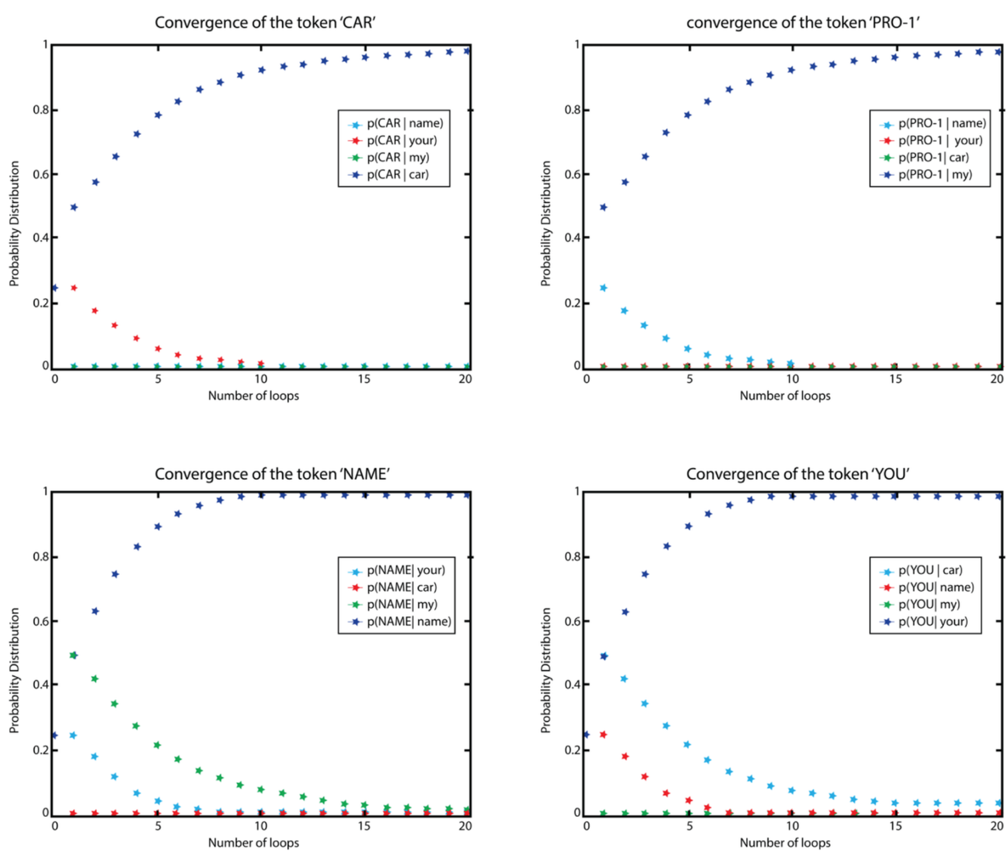
From the probabilistic distributions already computed previously for lexical translations, we can fix the first statistical automatic translator, which uses only the lexical translations. In previous table, we show the probabilistic distributions of three tokens in ASL and their alignments in English. We note the probability of translating a word f into a word in ASL e with a conditional probability function t (e | f). The resulting array is called T-tables. Provided an English-ASL parallel corpus, how can we make alignments between a source sentence and a target phrase to generate a dictionary? This can be performed by implementing the learning algorithms proposed by IBM (Brown et al., 1993). These alignment models are called “IBM Model” which are based on the probability of lexical translations using the alignment function a.

The probability of the translation is calculated from the product of all the lexical probabilities of each word e_j for j from 1 to l_e. At the level of the quotient authors added 1 to consider the token NULL. Therefore, they will have $(l_f + 1)^{l_e}$ possible alignment to map $(l_f + 1)$ words of the sentence f with all the words of target sentence e. μ is a constant of normalization of $P(e, a \mid f)$ that ensures $\sum_{e,a} P(e, a \mid f) = 1$. To conclude, authors apply the previous equation to the phrase “your blue car,” which is translated into “YOU BLUE CAR.”. The curves in Figure 9 show how the most likely translation converges to 1 and the others to zero for a given word.

Optimization Based on Similar String

In order to reduce the number of iterations and to accelerate the learning phase for the construction of an English-ASL lexical translation memory, authors have used the notion of similarity of strings. As the majority of English words are similar to their ASL translations and vice versa. For this, the distance of Jaro-Winkler was used (Jaro, 1989). The distance of Jaro–Winkler measures the similarity between two strings of characters. The distance is a value between 0 and 1. The Jaro–Winkler distance d_w of two strings S_1 and S_2 is defined as follows:

Figure 9. Convergence of the most probable lexical translations calculated from the IBM1 alignment model



$$d_w = d_j + \left(\ell p \left(1 - d_j\right)\right)$$

where ℓ is the length of the common prefix (maximum four characters), p is a coefficient that favors strings with a common prefix and d_j is the distance of Jaro between S_1 and S_2 defined as follows:

$$d_j = \frac{1}{3}\left(\frac{m}{\left|S_1\right|} + \frac{m}{\left|S_2\right|} + \frac{m-t}{m}\right)$$

where $\left|S_1\right|$ is the length of the string S_1, $\left|S_2\right|$ is the length of the string S_2, m is the number of corresponding characters, and t is the number of transpositions. This distance is integrated during the initialization of the probabilistic distributions before the learning phase. Initially, distribution $p\left(e|f\right)$ is initialized at $\frac{1}{l_e}$ where l_e is the number of words in target language e. In this case, the probability distribution $p\left(e|f\right)$ is defined as follows:

$$t\left(e|f\right) = \alpha \cdot \frac{1}{l_e} + \beta \cdot d_w\left(e, f\right)$$

Where α and β are two coefficients satisfying the following conditions: $\alpha, \beta \in \mathbb{R}$ and $\alpha + \beta = 1$. These two coefficients allow us to select the weight of the use of the distances of Jaro–Winkler during the initialization. For our experimentation, authors initialize α to 0.2 and β to 0.8 that favors the use of the similarity of the strings. The experiments were carried out on the same corpus as previously used containing three sentences They note that the distributions of the most probable translations converge rapidly toward 1. Figure 10 shows the evolution of convergence to lexical translations of four words (CAR, PRO-1st, NAME, YOU) learned from the corpus containing three sentences. It is noted that the distribution converges more rapidly than in the previous example. This shows the effectiveness of our approach. Let us return to the previous example of the pair p (YOU|you), Figure 10 (the blue curve) shows the probabilistic distribution using the Jaro–Winkler distance converges to 1 from 20 iterations.

Until now, authors have presented an implementation of the IBM1 learning model without and with the use of Jaro–Winkler distances. This allowed us to automatically build dictionaries of lexical translations of each word in the source language into a target language from a parallel corpus. It was shown that, at each iteration, the probabilistic distribution is refined up to the convergence toward the value 1. They also optimize the number of iterations, which reduces the learning time, by using the notion of similarity of strings.

Evaluation

Human evaluation is a method for determining the performance of a translation system. Human translations of machine translation are concerned with several aspects of translation, such as adequacy, fidelity, and mastery of translation. Human evaluation is widely discussed, as evidenced by the number of studies on the subject. The primary problem of human evaluation is the time it takes. It, therefore, corresponds more to a situation in which it is desired to evaluate a stable system. For our proposed system, the metric used is the BLUE score (Papineni et al., 2002). First, authors detail the results of the translation of a text in the ASL to a text in English. Table 2 shows the evolution of the

Figure 10. Convergence of the most likely lexical translations to 1 calculated from the IBM1 alignment model using Jaro–Winkler distance

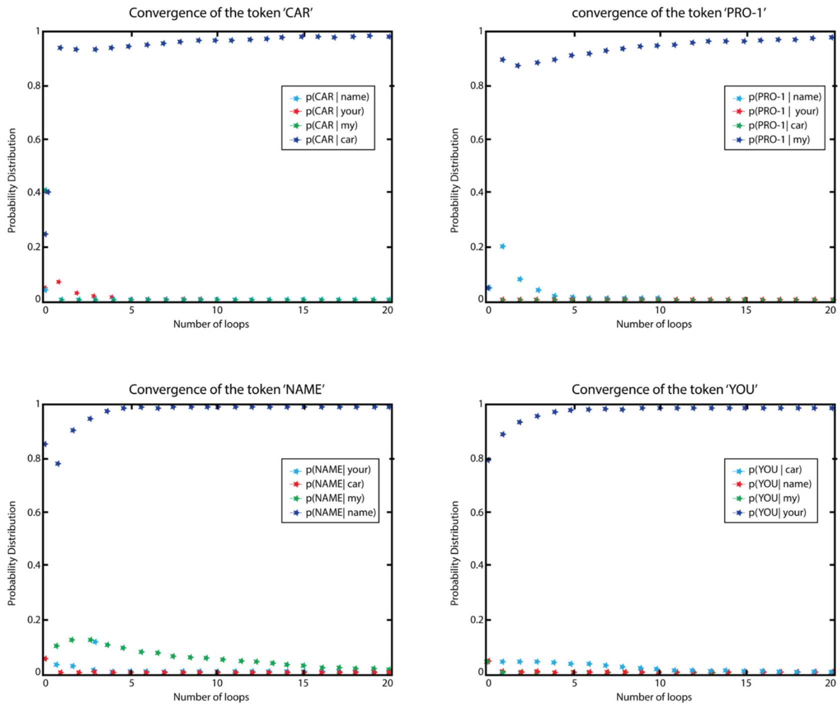


Table 2. Variation of BLUE score according to the size of evaluation corpus

#evaluation	BLUE	Precision			
		1-Gram	2-Gram	3-Gram	4-Gram
1000	41,19	81,3	52,3	37,8	28,8
2000	43,82	85,1	55,9	40,9	31,4
3000	45,16	85,6	57	42,3	32,6
4000	46,20	85,7	57,7	42,7	32,8
5000	47,57	87,8	59,5	44,6	34,7
6000	49,73	89,8	61,4	46,6	36,8
7000	48,66	88,9	60,2	45,3	35,6
8000	48,04	88,3	59,5	44,6	34,9
9000	46,61	86,6	57,9	43,2	33,4
10000	46,58	86,7	57,9	43,1	33,4

performance of the system as the size of the evaluation and learning corpus increases. In a second step, they repeated the same operation using Jaro–Winkler distance. This technique increased the BLUE score, which reflects the quality of the translation. The average BLUE score is 35.17 and 46.35 using the Jaro–Winkler distance.

CONCLUSION

The progress of research since half a century has enabled machine translation to affect our daily lives. Sign Language Translation is a recent theme of research because it combines two complex scientific problems: translation and the transcription of SL. However, it is not difficult to imagine its applications: online and real-time synthesis, access to information, indexing, cross-lingua of multimedia content, and assistant for communication and elearning. Many studies on SL are recent and innovative. The study includes the linguistic, cognitive, and grammar aspect until the creation of the corpus, machine translation, and real-time synthesis. As they perceive, Sign Languages are not universal. In general, the studies are focused on one community of deaf and do not share the same syntactic structures, phonological, lexical, morphological, and semantic aspects.

Despite existing tools for transcription and annotation, each presents drawbacks. However, for the textual annotation in gloss, authors proposed an XML representation exhibits more benefits for an automatic processing toward store and synthesis Sign Language via conversational agents. The proposed transcription system was based on Gloss Annotation, which were evaluated by comparison to manual transcription. Then, this paper focused on the machine translation and more particularly on the translation of English to ASL or vice versa. In accordance with the state of the art, the proposed machine translation is based on statistical models. These models comprise a large number of parameters of the order of several millions. Their training requires parallel texts: thousands or preference of million phrases translations of one another, through which the parameters of the models are estimated. These texts were generated automatically using rule-based approaches guided by the use of grammatical dependency graphs. Our experiences show that the model of alignment enhanced the performance using the Jaro–Winkler distances: between 8 and 11 blue points following the direction of translation.

REFERENCES

Abdel-Fattah, M. A. (2005). Arabic sign language: A perspective. *Journal of Deaf Studies and Deaf Education*, *10*(2), 212–221. doi:10.1093/deafed/eni007 PMID:15778217

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, *18*(1), 129–144. doi:10.1007/BF01069051

Arun, A., Haddow, B., & Koehn, P. (2010). A unified approach to minimum risk training and decoding. In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR* (pp. 365-374).

Bangham, J. A., Cox, S. J., Elliott, R., Glauert, J. R. W., Marshall, I., Rankov, S., & Wells, M. (2000). *Virtual signing: Capture, animation, storage and transmission-an overview of the visicast project*.

Berger, A. L., Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., Gillett, J. R., Lafferty, J. D., & Ureš, L. et al. (1994, March). The Candide system for machine translation. In *Proceedings of the workshop on Human Language Technology* (pp. 157-162). doi:10.3115/1075812.1075844

Brown, P. F., Cocke, J., Pietra, S. A. D., Pietra, V. J. D., Jelinek, F., Lafferty, J. D., & Roossin, P. S. (1990). A statistical approach to machine translation. *Computational Linguistics*, *16*(2), 79–85.

Brown, P. F., Pietra, V. J. D., Pietra, S. A. D., & Mercer, R. L. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, *19*(2), 263–311.

Bungeroth, J., Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way, A., & van Zijl, L. (2008). The ATIS sign language corpus.

Chabeb, Y., Elghoul, O., & Jemni, M. (2006). WebSign: A Web-Based Interpreter of Sign Language. In *International Conference on Informatics and its Applications, Oujda*.

Chiu, Y. H., Wu, C. H., Su, H. Y., & Cheng, C. J. (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. doi:10.1109/TPAMI.2007.250597 PMID:17108381

Collins, M. (2003). Head-driven statistical models for natural language parsing. *Computational Linguistics*, *29*(4), 589–637. doi:10.1162/089120103322753356

Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., & Abbott, S. (2002). Tessa, a system to aid communication with deaf people. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 205-212). doi:10.1145/638249.638287

De Marneffe, M. C., MacCartney, B., & Manning, C. D. (2006). Generating typed dependency parses from phrase structure parses. In *Proceedings of LREC* (pp. 449-454).

DeWitt, L. C., Lam, T. T., Silverglate, D. S., Sims, E. M., & Wideman, C. J. (2003). U.S. Patent No. 6,535,215. Washington, DC: U.S. Patent and Trademark Office.

Elliott, R., Glauert, J. R. W., Jennings, V., & Kennaway, J. R. (2004). An overview of the SiGML notation and SiGMLSigning software system. In *Sign Language Processing Satellite Workshop of the Fourth International Conference on Language Resources and Evaluation, LREC 2004* (pp. 98-104).

Federico, M., Bertoldi, N., & Cettolo, M. (2008). IRSTLM: an open source toolkit for handling large scale language models. In Ninth Annual Conference of the International Speech Communication Association (pp. 1618-1621).

Gao, Q., & Vogel, S. (2008). Parallel implementations of word alignment tool. In *Software Engineering* (pp. 49–57). Testing, and Quality Assurance for Natural Language Processing.

Grieve-Smith, A. B. (1999). English to American Sign Language machine translation of weather reports. In *Proceedings of the Second High Desert Student Conference in Linguistics (HDSL2)*, Albuquerque, NM (pp. 23-30).

Hanke, T. (2004). HamNoSys-representing sign language data in language resources and language processing contexts. In LREC (Vol. 4).

Hanke, T., Popescu, H., & Schmaling, C. (2003). eSIGN–HPSG-assisted Sign Language Composition. In *Gesture Workshop*.

Hoang, H., & Koehn, P. (2008). Design of the moses decoder for statistical machine translation. In Software Engineering, Testing, and Quality Assurance for Natural Language Processing (pp. 58-65). Association for Computational Linguistics. doi:10.3115/1622110.1622120

Huenerfauth, M. (2005). American Sign Language natural language generation and machine translation. *ACM SIGACCESS Accessibility and Computing*, (81), 12-15.

Jaro, M. A. (1989). Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. *Journal of the American Statistical Association*, *84*(406), 414–420. doi:10.1080/01621459.1989.10478785

Jemni, M., & Elghoul, O. (2008). A system to make signs using collaborative approach. In *International Conference on Computers for Handicapped Persons* (pp. 670-677). doi:10.1007/978-3-540-70540-6_96

Joshi, A. K., & Schabes, Y. (1991). Tree-adjoining grammars and lexicalized grammars. *Technical Reports, 445*.

Jouison, P. (1990). Analysis and linear transcription of sign language discourse. In *Proceedings of the International Congress on Sign Language Research and Application*, Hamburg, Germany (Vol. 90, pp. 337-354).

Klima, E. S. (1979). *The signs of language*. Harvard University.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., & Dyer, C. (2007). Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions* (pp. 177-180). doi:10.3115/1557769.1557821

Krings, H. P. (2001). *Repairing texts: empirical investigations of machine translation post-editing processes* (Vol. 5). Kent State University Press.

Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press. doi:10.1017/CBO9780511615054

Liddell, S. K., & Johnson, R. E. (1989). American Sign Language: The phonological base. *Sign Language Studies*, *64*(1), 195–277. doi:10.1353/sls.1989.0027

Morrissey, S., & Way, A. (2013). Manual labour: Tackling machine translation for sign languages. *Machine Translation*, *27*(1), 25–64. doi:10.1007/s10590-012-9133-1

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticae Investigationes*, *30*(1), 3–26. doi:10.1075/li.30.1.03nad

Neidle, C., MacLaughlin, D., Bahan, B., & G LR, K. (1997). The SignStream project. American Sign Language Linguistic Research Project.

Newkirk, D. (1989). *SignFont handbook*. Edmark Corporation.

Och, F. J., & Ney, H. (2000). Improved statistical alignment models. In *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics* (pp. 440-447).

Och, F. J., & Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, *29*(1), 19–51. doi:10.1162/089120103321337421

Olivrin, G. J., & Van Zijl, L. (2008). South African sign language assistive translation.

Othman, A., El Ghoul, O., & Jemni, M. (2010). SportSign: a service to make sports news accessible to deaf persons in sign languages. In *International Conference on Computers for Handicapped Persons* (pp. 169-176). Springer.

Othman, A., & Hamdoun, R. (2013). Toward a new transcription model in XML for Sign Language Processing based on gloss annotation system. In *2013 Fourth International Conference on Information and Communication Technology and Accessibility (ICTA)* (pp. 1-5). doi:10.1109/ICTA.2013.6815317

Othman, A., & Jemni, M. (2011). Statistical sign language machine translation: from English written text to American Sign Language gloss. arXiv:1112.0168

Othman, A., & Jemni, M. (2012). English-asl gloss parallel corpus 2012: Aslg-pc12. In *5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon LREC*.

Othman, A., & Jemni, M. (2013). Othman, A., & Jemni, M. (2013). A Probabilistic Model for Sign Language Translation Memory. In Intelligent Informatics (pp. 317-324). Springer. doi:10.1007/978-3-642-32063-7_33

Othman, A., Tmar, Z., & Jemni, M. (2012). Toward developing a very big sign language parallel corpus. In *International Conference on Computers for Handicapped Persons* (pp. 192-199). doi:10.1007/978-3-642-31534-3_30

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics* (pp. 311-318).

Raghunathan, K., Lee, H., Rangarajan, S., Chambers, N., Surdeanu, M., Jurafsky, D., & Manning, C. (2010, October). A multi-pass sieve for coreference resolution. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 492-501).

San-Segundo, R., Montero, J. M., Macías-Guarasa, J., Córdoba, R., Ferreiros, J., & Pardo, J. M. (2008). Proposing a speech to gesture translation architecture for Spanish deaf people. *Journal of Visual Languages and Computing*, *19*(5), 523–538. doi:10.1016/j.jvlc.2007.06.002

San-Segundo, R., Pardo, J. M., Ferreiros, J., Sama, V., Barra-Chicote, R., Lucas, J. M., & García, A. (2009). Spoken Spanish generation from sign language. *Interacting with Computers*, *22*(2), 123–139. doi:10.1016/j.intcom.2009.11.011

Schembri, A. (2008). British sign language corpus project: Open access archives and the observer's paradox. In *Language Resources and Evaluation Conference*.

Speers, D. (2002). Representation of American Sign Language for machine translation.

Stokoe, W. C. (2005). Sign language structure: An outline of the visual communication systems of the American deaf. *Journal of Deaf Studies and Deaf Education*, *10*(1), 3–37. doi:10.1093/deafed/eni001 PMID:15585746

Stolcke, A. (2002, September). SRILM-an extensible language modeling toolkit. In Interspeech (Vol. 2002, p. 2002).

Stuart, M. T. (2011). A Grammar of SignWriting [Doctoral dissertation]. University of North Dakota.

Su, H. Y., & Wu, C. H. (2009). Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(7), 1305–1315. doi:10.1109/TASL.2009.2016234

Toutanova, K., & Manning, C. D. (2000). Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics* (Vol. 13, pp. 63-70). doi:10.3115/1117794.1117802

Valli, C., & Lucas, C. (2000). *Linguistics of American Sign Language: An introduction*. Gallaudet University Press.

Von Laban, R., & Lange, R. (1975). *Laban's principles of dance and movement notation*. Princeton Book Co Pub.

Winkler, W. E. (1999). *The state of record linkage and current research problems*. Statistical Research Division, US Census Bureau.

Zens, R., Och, F. J., Ney, H., & Vi, L. F. I. (2002). Phrase-based statistical machine translation. In Annual Conference on Artificial Intelligence (pp. 18-32). Springer.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., & Palmer, M. (2000). A machine translation system from English to American Sign Language. In Envisioning machine translation in the information future (pp. 191-193). doi:10.1007/3-540-39965-8_6

Achraf is a researcher and ICT Expert with a successful record of accomplishment and experiences in many universities and international organization by using information and communication technology (ICTs), with a great deal of dedication and vision. From 2017, he is a Senior Research Specialist on Assistive Technology and Accessibility at Mada Assistive Technology Center in Doha, Qatar. Previously, he was an ICT Expert Specialist at Arab League Educational, Cultural and Scientific Organisation ALECSO. According to his education background (Master and PhD), he was a member in many projects to make information more accessible to people with disabilities especially with Deaf communities. Also, he worked on big project called AlecsoApps to promote Arabic e-Content and Arabic Mobile Content. He has written many scientific articles and papers for many recognized and specialized conferences on ICT and accessibility.

Mohamed Jemni is a Professor of Computer Science and Educational Technologies at the University of Tunis, Tunisia. He is the Director of ICT at The Arab League Educational, Cultural and Scientific Organization ALECSO (www.alecso.org). He has been the General Director of the Computing Center El Khawarizmi, the Internet services provider for the sector of higher education and scientific research in Tunisia, from 2008 to 2013. He is a Senior member IEEE, member of the Executive board of IEEE Technical Committee on Learning Technology (www.ieeetclt.org) and he is co-editor of Springer Lecture Notes in educational Technology. At ALECSO, he is currently leading several projects related to the promotion of effective use of ICT in education in the Arab world, namely, OER, MOOCs, cloud computing and the strategic project ALECSP-APPS with three main components: ALecsoApps Store, AlecsoApps Editor and AlecsoApps Award (www.alecsoapps.com), aiming to provide a technological environment for the promotion of an emerging digital creative Arab Mobile industry, related to the fields of education, culture and science. His Research Projects Involvements during the last 28 years are Educational Technologies, High Performance and Grid computing and Accessibility of ICT to People with Disabilities. He published more than 300 papers in international journals, conferences and books, and produced many studies for international organizations such as UNESCO, ITU and ALECSO. Prof. Jemni and his laboratory have received several awards, including the Silver medal of the International Fair of Inventions in Kuwait 2007, the UNESCO Prize 2008 for the e-learning curriculum they developed for visually impaired, President's Award for the integration of persons with disabilities 2009 and the "World Summit Award (WSA) - Mobile 2010" in the field of social inclusion. In April 2012, his laboratory LaTICE received the Google Student AWARD and he has received the Best Communication Paper of the Web For All 2012. Prof. Jemni received the International Telecommunication Union Award: "ICT Innovation application challenge" during the WSIS Forum in Geneva in May 2013. He is member of the steering committee of G3ICT – United Nations, Global initiative for Inclusive Information and Communication Technologies and he is also the president of a Tunisian NGO created in June 2011: the Tunisian Association of e-accessibility (www.e-access.tn). He has launched many initiatives to promote ICT accessibility in the Arab region including the project of WCAG2.0 translation to Arabic (www.alecso.org/wcag2.0/) to promote accessibility of Arabic Web Content and the 2009 initiative for using ICT to develop Arab Sign language. His full CV is available at his web page: www.latice.rnu.tn/english/memb_jemni.htm