Journal of Theoretical and Applied Information Technology
10th January 2016. Vol.83. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# PROPOSED IF_THEN GRAMMAR TO TRANSLATE ENGLISH LANGUAGE SENTENCE TO AMERICAN SIGN LANGUAGE SENTENCE

[1]ASMAA M. HAMANDI, [2]ALIAA K. ABDULHASSAN, [3]HALA BAHJAT
Department of Computer Science, University of Technology, Baghdad, Iraq.
E _Mail: [1]asmaa_hamandi@yahoo.com, [2]hassanalia2000@yahoo.com, [3]hala_bahjat@yahoo.com

**ABSTRACT**

Sign Language is a language that used by Deaf community. One of the most widely used SL is the American Sign Language. In order to investigate the communication between Deaf people and others among them in the community, a machine translation system should be available for them in free use. The development of machine translation system that translate an English text into Sign language needs a grammar that producing Sign language from an input English text.

In this paper a proposed ASL grammar rules that used to covert English sentence into ASL sentence will be produced. The grammar rules have been introduced as an If-Then- Grammar.

**Keywords:** *Sign Language, NLTK, American Sign Language, Grammar Rules, Machine Translation*

## 1. INTRODUCTION

**A Sign Language SL** is a visual language which uses manual communication and body language to convey meaning instead of acoustically conveyed sound patterns. This can involve simultaneously combining hand shapes, arms or body, orientation and movement of the hands, and facial expressions to fluidly express a speaker's thoughts.

**Deaf Community** a group of persons who share a common means of communication (signs) that compose the SL [Cok, 94]. Wherever communities of deaf people exist, sign languages develop. SL is also used by persons who can hear, but cannot physically speak. SL has the same linguistic properties and uses the same language faculty as do spoken languages. Hundreds of sign languages are in use around the world according to different deaf communities. Some sign languages have obtained some form of legal recognition, while others have no status at all. The most widely used SL is the American Sign Language ASL. Therefore, SL is a language that has developed naturally over time among a community of users, Deaf people.

ASL has its own grammar, sentence construction, idiomatic usage, slang, style, and regional variations and the characteristics that define any language [Ric, 99].

Fingerspelling is the process of spelling out words by using signs that correspond to the letters of the word. In many ways finger spelling serves as a bridge between the sign language and the oral language that surrounds it.

There are lots of items when fingerspelling is used to translate it. The typical "these things are spelled" list includes such items as : people's names, places , titles, and brands [Wil, 11].

Producing machine translation systems to convert English text into Sign Language is very necessary on order to help Deaf people and their family and community.

In this work, a proposed method to convert English sentence to ASL based on building an If _ then _ Rule Grammar. The proposed method assumed to be easy and efficient in accurate translation from English written text to ASL written form. The ASL written form can used in future work to be displayed in one of the available ways of displaying the Sign Languages, that ways may be images, videos or animation.

Journal of Theoretical and Applied Information Technology
10th January 2016. Vol.83. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

## 2. ASL PHONOLOGY:

The term phonology is used by SL linguists to refer to the study of how signs are structured and organized. ASL signs have five basic parts **are:**

**Handshape** is the configuration the hand assumes when beginning to make a sign. The most frequently used handshapes are the letters of the American Manual Alphabet and the manual numbers, these handshapes are known as fingerspell.

**Palm Orientation** is the direction in which the hand is turned. The direction that the palm of the hand faces (up, down, left, or right) is a useful way of describing the orientation.

**Location** of a sign frequently contributes to its meaning. For example, many signs that denote feelings are formed near the heart, whereas signs related to cognitive concepts are formed near the head.

**Movement** is the direction in which a sign moves that may indicate the doer or recipient of the action. For example, if the sign HELP' moves from the signer toward another person, help is offered to that person. If the sign moves in toward the signer, another person might be helping the signer. The repetition of the movement may indicate several meanings such as the frequency of an action, the plurality of a noun, or the distinction between a noun and a verb.

**Nonmanual Signals, t**his nonmanual parameter occurs at the same time that the sign is being executed to contribute to its meaning. These nonmanual paramerters like: facial expression, head shaking, eye movement, eyebrow that give the emotions that associated with the sign, such that a complete thought about the sentence has been gives.

All the preceeded parameters can be illustrated in the example in figure(1). Figure(1) describes all the ASL Phonology through the example, the figure represents a word from ASL dictionary illustrated within it the five items of SL phonology. The figure represents the word "sorry" and all its synonyms (apologize, regret, remorse, repent), also illustrates the handshape, orientation, location, movement and nonmanual signals.

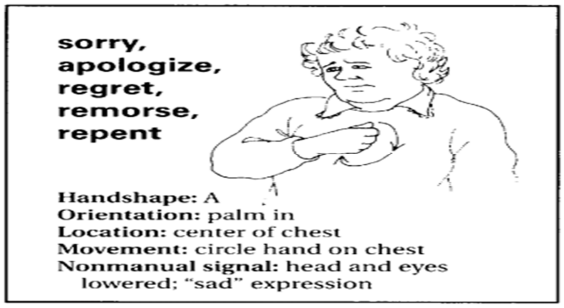


*Figure 1: a word from ASL dictionary [Ric,99]*

The sign is restricted with the five items of ASL phonology. Changing one of the five items will produce the representation of new different word with its synonyms.

## 3. ASL MACHINE TRANSLATOR (ASLMT)

ASLMT System: this expression can be used to describe the system that constructed to translate an English Text into ASL or another system that constructed to analyze the ASL signs into another language such as English language.

### 3.1. ASLMT Importance

Multiple things have been considered as problems that need the development of ASLMT in order to solve them.

**a. Enhanced Communication**

It helps to investigate an enhanced communication in a particular community, the Deaf community. Also for investigate the communication between the Deaf persons or hearing impaired inside their communities. [Eli,03], [Mar,09].

b. **Low Levels of reading capabilities for Deaf persons**

Many deaf people in the world have difficulties with reading and writing; in fact some of them do not read. A text-based system would make it impossible for the people to follow and understand a conversation in real-time.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

**c. Lacking for more Illustrations**

If spoken language is rendered as text, all the information related to the semantic-cognitive meaning, time, space, emotion, intensity of the signs and facial expression is lost .

**d. SL has its own unique grammar**

As the primary communication means used by members of deaf community, sign language is not a derivative language. It is a complete language with its own unique grammar . But the majority of automatic interpretation system ignores this important aspect therefore the interpretation lost the truth information meaning. [Hue, 05] [Moh, 08] [Meh, 08]

**e. No Standardization among Human Translators**

Further more if the dependant of the human translation for sign language is resulting in missing standardization in the translation, that is the differences among different persons as translators (such as: arm length, eyes shape, special movement) will affect the translation.

**3.2. Difficulties of ASLMT:**

According to the system that translates an English text into ASL there are many difficulties:

**a. Sign language does not represent spoken language**

It is a visual language that has its own grammar and structure that distinct from the other languages like English language.

**b.Different Levels of Deaf learning and communication**

Many Deaf people who do not acquire sign language from their parents, develop signing skills later in life. Most deaf people are bilingual, using a signed language and a written form of some spoken language.

**c. Different Signers Using SL**

Among signers, there is great variety in terms of hearing perception, language background, socialization to Deaf culture, native vs. late learning of sign language, style, and so on. What is also unique in Deaf communities is the daily reliance on interpreters, who can be informal makers of language policy.

**d. Lacking for Standard Written Grammar for ASL**

There is no existence for free standard written grammar for ASL, the only thing that exists is the description of the grammar from which the system has been derived the rules for this grammar.

**e. Lacking for Standardization for the output**

There is no standard output for the system. Some of ASLMT systems convert the written text as pictures, some of them work like a dictionary for word by word, while another form of output is as animation.

## 4. GLOSSING SIGNS

Glossing is like translating. A gloss of a sentence is a series of English words, written in small capital letters, that correspond to the signs in the ASL form. Parts of English, such as plural markers, past-tense markers, and prepositions, do not appear in glossing according to the grammar rules. The nonmanual features if any are indicated on a line above the sign glosses. Some basic conventions used for glossing are as follows[Cla,00]:

1. Signs are represented with small capital letters in English.
2. Fingerspelled words are written in small capital letters and preceded by the # symbol.
3. Nonmanual signals and eye-gaze are represented on a line above the sign glosses.

## 5. ASL GRAMMAR CATEGORIES

For constructing ASL sentence, some rules must be applied according to the ASL grammar categories. These categories are:

**5.1. Sentence Structure** and **Topicalization**

**Sentence Structure.** The basic word order for a sentence with an object is Subject-Verb-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Object, SVO, like the English sentence. **Topicalization** is the placement of the object at the beginning of the sentence, and is very common in ASL.

### 5.2. Time Words

Certain parts of language structure allow to show a difference among the present, the past, and the future. Sometimes it be used independent lexical items. For example, in English using the words tomorrow, yesterday, or soon to indicate the time of an event. If any of time words exists in the English sentence, it must be placed at the beginning of the sentence.

### 5.3. Negation in ASL

In ASL, negative sentences have specific nonmanual signals that include shaking the head from side to side. The negation of the sentence may be expressed by one of the two signs: NOT, NONE. According to ASL syntax, the negation words always come at the end of the sentence.

### 5.4. Conjunctions, Interjections, and Prepositions

**Conjunctions**, combining two ASL sentences is different based on the conjunction needed. For example the concept of the word AND does not exist in ASL. Sign a sentence, take short pause, and then sign the next sentence. While conjunctions such as OR and BUT have signs. **Interjections** like WOW! or OH! can be fingerspelled or they may have a sign like WOW. They are usually made into a separate statement. While **Prepositions** are not used in ASL, because they are shown in context.

### 5.5. Adjectives and Adverbs

**Adjectives.** ASL puts the adjective after the noun it modifies. This is differs from English, which requires that an adjective precede the noun it modifies. **Adverbs, i**n English are placed after the verb, whereas in ASL they are placed before the verb.

## 6. THE PROPOSED SYSTEM OVERVIEW

The proposed system takes the English sentence as an input, translates it into ASL sentence and gives the ASL sentence as an output using a proposed ASL grammar rules according to the ASL grammar categories.

As a preprocessing step of the system, is to apply the Natural Language Processing (NLP) to the Input English sentence (during NLP: process the English text for producing the part-of-speech (POS) tagging). For the purpose of the NLP, using NLTK toolkit in order to obtain trusted POS tagging. The Proposed ASL grammar rules will then be applied to the English sentence, according to the POS tagging, to produce the ASL sentence. Figure (2) shows the block diagram for the proposed system.

In figure (2) the sub steps of the system has been illustrated. Each of these sub steps has its important role. The importance of NLP is finding for each word its corresponding POS tag. The importance of the capitalization is to provide normalization that is very useful in searching and matching. In order to obtain the ASL sentence, a group of grammar rules has been applied, that represents the most important stage.

### 6.1. Proposed System Algorithm

Algorithm (1) describes the main steps of the system according to its general description and block diagram in figure (2). The main steps contains the most important step is applying the proposed ASL grammar rules.

Algorithm (1) (Translation stage: translating POS tagged English sentence into ASL sentence)
input: English sentence
output: ASL sentence
Step 1: Get the input sentence
str = readfile(sentence);
Step 2: NLP : finding the POS tag for each word in the input sentence
words= words-split(str);
tags= POS tags-split (str);
Step 3: Normalization for the words list, saving it into norm-words list
norm-words = capitalization(words);
Step 4: Applying the ASL Grammar Algorithm in order to produce the ASL sentence (algorithm 2)
ASL sentence= Grammar (norm-words, ASL grammar rules);
Step 5: Return ASL sentence

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The proposed ASL grammar rules have been described in algorithm(2). According to these rules some items must be deleted and other items must be reordered in the sentence, another items must be stayed at the end of the sentence and others must be brought to the beginning of the sentence. Algorithm (2) describes applying the ASL grammar rules. The proposed system has used the ASL grammar specifications for building (IF-Then-Grammar) Rules that converting the English sentence into ASL sentence. It is worth mentioning that, there no exists for such rules to be available for free use.

Thus, the proposed system allowed this proposed rule grammar for free use in scientific researches and available within reach to whom wants to use it, such as Deaf people and their family or friends or teachers or others in their community.

Algorithm (2) (Grammar sub-stage: applying ASL grammar to obtain the ASL sentence)

input: Norm-Words , POS tags

output: ASL sentence

Step 1: Searching the normalized words list and the POS tags list to find the items must be deleted in order to delete them

words= search-process(delete-items);

POS tags= search-process(delete-items);

Step 2: Searching the normalized words list to find the time items that must be brought to the beginning of the list

words= search-process(time-items);

Step 3: Searching the normalized words list and the POS tags list to find the items that refer to past tense in order to apply a past time flag at the beginning of the list

words= search-process(past-tense-items);

POS tags= search-process(past-tense-items);

Step 4: Searching the normalized words list and the POS tags list to find the items that refer to continuous tense in order to duplicate the verb in the list

Continue with Algorithm (2)

Step 5: Searching the normalized words list to find the items that refer to negation in order to apply a negation flag at the end of the list

words= search-process(negation-items);

Step 6: Searching the POS tags list to find the items that refer to adjectives and nouns in order to replace them in both lists

words= search-process(adjectives-and-nouns);

POS tags= search-process(adjectives-and-nouns);

Step 7: Searching the POS tags list to find the items that refer to adverbs and verbs in order to replace them in both lists

words= search-process(adverbs-and-verbs);

POS tags= search-process(adverbs-and-verbs);

Step 8: Searching the normalized words list to find dots in order to be deleted

words= search-process(dots-items);

Step 9: Now the produced list is representing the ASL sentence

ASL sentence= words ;

Step 10: Return ASL sentence

### 6.2. Proposed ASL Grammar Rules

One of the most important system contribution is the proposed grammar rules that have been applied to the English sentence with its corresponding POS tags in order to produce the ASL sentence.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

**1. Searching for Delete Items**

Searching through English sentence or its corresponding POS tags to find the items to be deleted according to the grammar rules. When the system found one or more of the these items, the word and its corresponding tag must be deleted from the POS tagged as well as the word lists in order to build the new ASL sentence.

**2. Searching for Time Items**

Searching through English sentence to find the items that have been referred to the time, these items can be summarized as: "MORNING", "AFTERNOON", "EVENING", "NIGHT" , "YESTERDAY".

when the system found on or more of these items, they must be brought to the beginning of the new ASL sentence according to the ASL grammar rules.

**3. Searching for Past-Tense Items**

Searching through English sentence or its corresponding POS tags to find the items that have been referred to Past-Tense. When the system found one or more of the these items, the word "YESTERDAY" must be added in the beginning of the new ASL sentence if it wasn't already exist, while the each one of the above items has its own processing. For example the verb in the past tense will be returned to its base tense in the third stage of the system, while the auxiliary verbs will be deleted in other stage of the system. Each item of the sentence must be processed with its corresponding POS tag in order to build the new ASL sentence.

**4. Searching for Continuous Items**

Searching through English sentence or its corresponding POS tags to find the items that have been referred to Continuous, these items can be summarized as ("IS", "ARE", "AM", "WAS","WERE") as words in the English sentence and must be followed by a word that its corresponding POS tag is ("VBG": it refers to verb, present participle or gerund, for example: encrypting interrupting erasing).

When the system found one or more of the above items with the condition of must be followed by present participle verb, the following subtasks must be applied: a. Deleting the auxiliary verb ("IS", "ARE", "AM", "WAS","WERE") and its corresponding POS tag. b. Duplicating the verb (present participle or gerund). These two subtasks have been applied according to ASL grammar in order to construct the new ASL sentence.

**5. Searching for Negation Items**

Searching through English sentence to find the items that have been referred to Negation, these items can be summarized as ("NOT" , "NONE" , "NEVER", "NOBODY" , "NOTHING" , "NOONE", "N'T").

When the system found one or more of these items, the following subtasks must be applied: a. Deleting these items for the sentence and its corresponding POS tag. b. Appending these items at the end of the sentence as well as its corresponding POS tag.

These two subtasks have been applied according to ASL grammar in order to construct the new ASL sentence.

**6. Searching for Adjectives and Nouns Items**

Searching through POS tags to find the items that have been referred to Adjectives. When the system found one or more of these items with the condition of must be followed by Noun, the system must exchange the position of the adjective and its noun, such that the noun must precedes the adjective in ASL sentence according to ASL grammar in order to construct the new ASL sentence. Switching the words in the word list must accompanied with switching to their corresponding POS tags in the tag list.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

**7. Searching for Adverbs and Verbs Items**

Searching through POS tags to find the items that have been referred to Adverbs. When the system found one or more of these items with the condition of must be followed by Verb, the system must exchange the position of the adverb and its verb, such that the verb must precedes the adverb in ASL sentence according to ASL grammar in order to construct the new ASL sentence. Switching the words in the word list must accompanied with switching to their corresponding POS tags in the tag list.

**8. Searching for DOTs**

Searching through English sentence to find the Dots, since the dot is not considered within the punctuations that are removed from the sentence in the first stage of the system. Removing Dots from the word list must be accompanied with removing its corresponding POS tags in the tag list.

**9. Return the ASL Sentence**

After applying all the ASL Grammar Rules to the English sentence, the new constructed sentence is an ASL sentence that is ready to be displayed in any way that be suitable to the Deaf people and their community. Thus the system provides them with the ASL sentence and display it as movies, such that one movie for each ASL word after searching for the movie in a suitable dictionary, that will be described in details in the next sections, as the third stage of the system "Matching Stage".

**6.3. Experimental Example**

In this section, an example of an English sentence as an input sentence, the steps of processing by applying the grammar rules in details, finally having the output ASL sentence.

**1. Input Sentence:** *" Huda, Muna, and Suha were good players yesterday."*

**2. Preprocessing Steps:**

a. Passing through preprocessing for getting POS tagging, the punctuations will be omitted, the sentence will be :

*"Huda Muna and Suha were good players yesterday."*

With its corresponding POS tags.

b. Splitting the sentence into words, construct list of words:

*Words = [[Huda],[ Muna], [Suha], [were], [good], [players], [yesterday.]]*

c. Capitalization, convert each letter into its capital form.

*Words = [[HUDA],[ MUNA], [SUHA], [WRER], [GOOD], [PLAYERS], [YERSTERDAY.]]*

**3. Applying ASL grammar rules.**

a. Searching for Delete Items.

The input sentence does not contain any item to be deleted, so the words list will not convert.

*Words = [[HUDA],[ MUNA], [SUHA], [WRER], [GOOD], [PLAYERS], [YERSTERDAY], [.]]*

b. Searching for Time Items.

The word [Yesterday] is one of the time items, so it must be brought to be at the beginning of the words list. The words list will be:

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [WRER], [GOOD], [PLAYERS] , [.]]*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

c. Searching for Past-Tense Items.

By searching for such items, the system will found the word [Were] that refer to past tense, but the words list will not convert because the word [Yesterday] is already exist at the beginning.

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [WERE], [GOOD], [PLAYERS] , [.]]*

d. Searching for Continuous Items.

Although the sentence contains the word [Were] that is referred to continuous only if it followed by verb in present participle tense(VBG), thus the word [Were] does not refer to continuous. So the words list will not convert, except of deleting the word [Were].

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [GOOD], [PLAYERS] , [.]]*

e. Searching for Negation.

The sentence does not contain any item that is referred to negation, so the words list will not convert.

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [GOOD], [PLAYERS] , [.]]*

f. Searching for Adjectives and Nouns Items.

By searching for such items, the system will found the two words [Good], [Players]. The adjective must follow the noun according to ASL grammar rules. So the words list will be changed.

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [PLAYERS], [GOOD] , [.]]*

g. Searching for Adverbs and Verbs Items.

The sentence does not contain any item of this type, so the words list will not convert.

*Words = [[YERSTERDAY], [HUDA],[ MUNA], [SUHA], [PLAYERS], [GOOD] , [.]]*

h. Searching for Dots.

Searching through words list to find the Dots. Removing Dots from the words list, will change the words list.

*Words = [[YERSTERDAY.], [HUDA],[ MUNA], [SUHA], [PLAYERS], [GOOD]]*

i. Return the ASL sentence is the sentence of last form of words list combining the words into one sentence.

*Words = [[YERSTERDAY.], [HUDA],[ MUNA], [SUHA], [PLAYERS], [GOOD]]*

**4. Producing ASL sentence** :

*"YERSTERDAY HUDA MUNA SUHA PLAYERS GOOD"*

### 6.4. System Evaluation and Discussion

The proposed system has been evaluated as any translation system according to two evaluation factors these are accuracy and time. The proposed grammar is the main part of the system to be evaluated.

#### 6.4.1. Building English- ASL corpus

The automated evaluation any translation system is based on building a corpus. Proposed system evaluation should be performed by building large English-ASL corpus that contains English sentence with its corresponding ASL sentence that is produced according to ASL grammar rule. The corpus also contains information about the POS tags for the sentences. The corpus consists of more than one hundred pair of sentences with its POS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

tag information. The corpus has been built for evaluation purpose only.

Since there is no existence of standard ASL grammar for free evaluation and comparing with the proposed system, the method for evaluating the proposed grammar will be available for future comparisons and evaluation metrics. To enlarge the constructed corpus, it would be time consuming, thus the corpus has been built for evaluation purposes only, and in future, it may be enlarged gradually over time to be a proposed standard English - ASL corpus available for free use. The English - ASL corpus that has been built for evaluation purpose is used to find the output ASL sentence corresponding to the English input sentence that has been produced by applying the proposed ASL grammar rules. Two evaluation factors would be used are accuracy and time.

### 6.4.2. Evaluation Factors of the Proposed System

The proposed system has been evaluated as any translation system according to two evaluation factors these are accuracy and time. The proposed grammar is the main part of the system to be evaluated.

#### a. Accuracy Evaluation Factor

For comparing the output ASL sentence for the proposed system with the ASL sentence of the corpus a similarity measure has been needed. One of the most common set similarity measures is the Jaccard Similarity Index, which is based on a simple set operations union and intersection. The standard Jaccard Similarity equation is (Equ. 1), using this measure with additional processing steps to find the evaluation accuracy measure [Wil, 03].

*Jaccard Similarity Ratio = Intersection(set A, set B) / Union (set A, set B) Equ(1)*

#### a.1. Evaluation Processing steps

Several processing steps have been applied to get the evaluation results. Algorithm (3) describes these steps.

Algorithm (3): Evaluation process

Processing the input sentence to find the index of the most similar sentence in the corpus.

Input: Input Sentence .

Output: The most similar sentence to the input one in the corpus.

step (1): Get the input paragraph or sentence

str = readfile(sentence);

Step (2): Remove the punctuations from the input text

str1= remove-punctuations(str);

Step (3): Remove English Stop Words from the sentence

str2= remove-stop-words(str1);

Step (4): Convert the input sentence into tokens

words= tokenization(str2);

Step (5): Normalization process by converting each token into Uppercase

words-norm= Uppercase(words);

Step(6): Searching the corpus to compute the Jaccard similarity ratio and finding the most similar one to the input sentence.

Semilar-sent=get-sent(Jaccardmax(words-norm,corpus-sentences));

Step(6): Return Similar-sent;

If the input text was a paragraph or a text file that contains several sentences, the algorithm has segmented it into individual sentences dealing with them one after another. The algorithm steps then would be applied to an individual sentence.

After reading the input sentence, a process of removing punctuations from the input sentence.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

After that, removing English stop words (words that are high frequency and low content like: but, is , or and if). Then, the normalization step by changing each letter in the sentence to its uppercase. Step (6) in the algorithm is to open the corpus to find out the most similar sentence in the corpus to the input sentence, the most similar sentence is identified by computing the Jaccard similarity.

For each input sentence, Algorithm (3) have been applied two times. First, applying the algorithm to find the most similar English sentence in the corpus to the input English sentence. Second, applying the algorithm to find the most similar ASL sentence in the corpus to the output ASL sentence that has been translated with the proposed ASL grammar. The success case is to return the same index in the two times. That means the translation process has been performed correctly. It is worth mentioning to know that most of the failure cases belongs to that the test sentence does neither exists in the corpus nor similar to any of the existence sentence in the corpus. This problem could be challenged by enlarge the corpus over time.

**a.2. Experimental Results**

Two types of experimental examples would be considered. First, in order to explain the algorithm(3) steps, an example of an English sentence would take place to find practically the effect of each step. Second, test text (group of test sentence) would be taken as input and results would be discussed.

**1. Input English Sentence**

Giving English sentence as an input to the algorithm and show the effect of each step. Input English Sentence: "In the eighteenth century, it was often convenient to regard man as a clockwork automaton."

**Step (1):** Get the input paragraph or sentence

str = "In the eighteenth century, it was often convenient to regard man as a clockwork automaton.";

**Step (2):** Remove the punctuations from the input text

str1= "In the eighteenth century it was often convenient to regard man as a clockwork automaton";

**Step (3):** Remove English Stop Words from the sentence

str2= "eighteenth century often convenient regard man clockwork automaton";

**Step (4):** Convert the input sentence into tokens

words= ['eighteenth', 'century', 'often', 'convenient', 'regard', 'man', 'clockwork', 'automaton'];

**Step (5):** Normalization process by converting each token into Uppercase

words-norm= ['EIGHTEENTH', 'CENTURY', 'OFTEN', 'CONVINIENT', 'REGARD', 'MAN', 'CLOCKWORK', 'AUTOMATON'];

**Step (6):** Multiple sub steps have been applied. These sub steps may be summarized as simple as possible to show only stemming and lemmatizing processes.

words-norm= ['EIGHTEEN', 'CENTURY', 'OFTEN', 'CONVINIENT', 'REGARD', 'MAN', 'CLOCKWORK', 'AUTOMATA'];

The list words-norm then would be matched with the all English sentences in the corpus to find the most similar one, and be returned as output.

**2. Input Test Text**

The results of the proposed ASL grammar are differ according to number of sentences, words, and the existence of the sentence itself or similar to it in the corpus or not. Selecting sentences for

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

testing operation depends on whether the test sentences were the same sentences of English-ASL corpus that has been build for evaluation purpose or nearing them, or the test sentences may be different of them. Different samples of test sentences has been selected giving different results. Table (1) and table (2) shows these different results.

Table (1) contains different cases ( English paragraphs) that are selected carefully, such that, most of the sentences in the paragraphs are similar to (not exactly the same) the sentences that have been saved in the corpus. Different sentences have been used, long sentences and short ones with different number of words.

*Table 1: Different Results for Different Selected Test Cases*

| No. of Sentences | No. of words | Max Jaccard for English Sentence Ratio | Min Jaccard for English Sentence Ratio | Max Jaccard for ASL Sentence Ratio | Min Jaccard for ASL Sentence Ratio | Total Success Ratio |
|---|---|---|---|---|---|---|
| 10 | 100 | 1.0 | 0.285 | 1.0 | 1.0 | 1.0 |
| 15 | 152 | 1.0 | 0.1428 | 1.0 | 0.25 | 0.9333 |
| 24 | 229 | 1.0 | 0.3333 | 1.0 | 1.0 | 1.0 |
| 25 | 230 | 0.8333 | 0 | 1.0 | 0 | 0.92 |

Maximum Jaccard Similarity for English sentences sometimes couldn't reach to (1.0), while for the ASL sentences could always reach to the maximum value(1.0), that belongs to the properties of ASL sentences (chapter two and three) that made the ASL sentences shorter than its equivalent English ones. The Minimum Jaccard Similarity may be the minimum value (0) for both English and ASL sentences, when there exists failure sentences in some cases. The Total Success Ratio, is the ratio that computed by dividing the number of success sentences by the number of all sentences for each case. The total success ratio may reach to (1.0) if all the sentences success and have a Jaccard similarity by finding its similar sentences (English and ASL) in the corpus.

*Table 2: Different Results for Different Random Test Cases*

| No. of Sentences | No. of words | Max Jaccard for English Sentence Ratio | Min Jaccard for English Sentence Ratio | Max Jaccard for ASL Sentence Ratio | Min Jaccard for ASL Sentence Ratio | Total Success Ratio |
|---|---|---|---|---|---|---|
| 9 | 93 | 0.1818 | 0 | 0.1818 | 0 | 0.2222 |
| 10 | 100 | 0.125 | 0 | 0.1666 | 0 | 0.3 |
| 13 | 170 | 0.3333 | 0 | 0.2857 | 0 | 0.2307 |
| 21 | 408 | 0.25 | 0 | 0.2222 | 0 | 0.3809 |

Table (2) shows the results for four cases in which produces random samples of paragraph from the web. Number of sentences is different from each case to another as well as number of words. It is obvious that the results in table (2) is less efficient than results in table (1). That belongs to the randomly selected English paragraph. Results in table (2) are still to be considered somewhat good because they referred to varietal the sentences that saved in the corpus. Such that there exists only more than one hundred sentences (because of the restricted time) and however that, the success results (even if it was few) is considered to be good mark that be given to both the grammar and the corpus.

According to results in both tables (1) and (2), each sentence is either computed as success of failure. The sentence computed as failure in one of two possibilities:

1. Couldn't find the similar sentence in the corpus.
2. Found an English sentence that is similar to the failure sentence, but couldn't find the ASL translation for it in parallel with the English one, or vice versa.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

According to the first possibility, the problem could be solved by enlarging the English-ASL corpus to contain more different sentences in order to find similar sentences for all sentence as more as possible. The second possibility is referring to an error in the grammar rules, or in the parallel English-ASL corpus. It is worth mentioning, according to all cases in the two tables (1 and 2) all failure sentences are belong to the first possibility, that all failure sentences couldn't be found in the corpus. That means, the grammar and the structure of the corpus have not recorded any error cases.

**b. Time Evaluation Factor**

As any translation system, the proposed system would be evaluated according to time factor. Since the system contains two sequential stages, take with consideration the time of each stage and the length of the sentence (number of words and total number of characters in the sentence).

*Table 3: System Stage Time in seconds*

| No. of words | No. of characters | 1st-stage Time in Seconds | 2nd-stage Time in Seconds |
|---|---|---|---|
| 12 | 63 | 2.36529 | 0.00044 |
| 11 | 57 | 2.23514 | 0.00054 |
| 10 | 54 | 2.26948 | 0.00048 |
| 11 | 67 | 2.38614 | 0.00054 |
| 7 | 43 | 2.33522 | 0.00142 |
| 7 | 42 | 2.30316 | 0.00044 |
| 4 | 17 | 2.26144 | 0.00048 |
| 13 | 66 | 2.23671 | 0.00056 |
| 99 | 513 | 2.38541 | 2.24371 |
| 161 | 958 | 2.43708 | 2.35889 |

Table (3) shows the execution time of different samples selected randomly. There is no large effecting to the number of words or characters, that because each stage consists of different sub stages that are required for all sentences whatever its number of words or characters, these sub stages such as stemming, tokenization, lemmatization. Using NLTK provides time saving for all of these operations. In addition to that, the number of words in the sentence refers to English stop words also have been computed (such as : a, the, to, in), while in most cases these words may be deleted in the first few sub stages. Number of characters also encounters the spaces and the punctuations.

## 7. CONCLUSION AND FUTURE WORK

### 7.1. Conclusion

1. Using the NLTK for NLP in order to obtain the POS tags of the sentence resulting into trusted system in shorter time.
2. If an error has been found in the translated system, it may caused by the small error ratio of the NLTK toolkit during finding the POS tags.
3. Using this system as an important tool for the deaf and hearing impaired persons in their community in order to understand each other.
4. According to ASL grammar, many rules have flexibility during the translation. The system try to be stable according to the most famous grammar categories for the ASL language.
5. Using Python programming language has the main contribution in producing these results (Tables 1, 2, and 3). It provides the using of NLTK in addition to provide reducing time consuming.

### 7.2. Future Work

1. As an important future work for this system is to find good method in order to display the ASL sentence. Such methods as pictures, videos, or animation. Displaying the ASL sentence required using big dictionary and big database and good method for matching.
2. It is very good if the system could be available on the web in order to be for free use by deaf communities.
3. The system may be improved to used for translation in specific subject, such as: weather forecasting, football matches. Such specification will make the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

vocabularies more limited, thus the error ratio will be small.

4. Developing the English _ ASL corpus by enlarging it with adding different sentences over time.

**REFERENCES:**

[1]. [Cla, 00] Clayton Valli , Ceil Lucas. (2000). LINGUISTICS OF maerican sign language. washington: clerc books Gallaudet University Press.

[2]. [Cok, 94] Cokely Dennis , Charlotte Baker-Shenk . (1994). American Sign Language: A Teacher's Resource . Washington: Clerc Books Gallaudet University Press .

[3]. [Eli, 03] Elizabeth Keating , Gene Mirus . (2003). American Sign Language in virtual Space: Interactions between deaf users of computer-mediated video communication and the impact of technology on language practices . In Language in Society. United States of America: Cambridge University Press, USA.

[4]. [Hue,05] M. Huenerfauth. (2005). "American Sign Language Generation: Multimodal NLG with Multiple Linguistic Channels", Proceedings of the ACL Student Research Workshop, pp. 37–42.USA.

[5]. [Mar, 09] Marek Hruz , and others. (2009). Input and Output Modalities used in a Sign-Language-Enabled Information Kiosk. Czech: University of West Bohemia, Pilsen, Czech Republic.

*[6]*. [Meh, 08] Mehrez Boulares, Mohamed Jemni. (2008). *Toward an example-based machine translation from written text to ASL using virtual agent animation", Research Lab. UTIC, University of Tunis, Avenue Taha Hussein, B. P. : 56, Bab Menara, Tunis.*

*[7]*. [Moh,08] Mohamed Jemni, Oussama Elghoul. (*2008). A System to Make Signs Using collaborative Approach. 11th International Conference, ICCHP 2008, Linz, Austria, July 9-11, 2008. Proceedings. Lecture Notes in Computer Science 5105 .pp. 670-677. Springer.*

[8]. [Ric, 99] Richard A.Tennant, Marianne Gluszak Brown. (1999). The american sign language handshape dictionary. washington: clerkBooks Gallaudet University Press.

[9]. [Wil, 03] William W. Cohen, Pradeep R., Stephen E. Fienberg. (2003). A Comparison of String Distance Metrics for Name-Matching Tasks. American Association for Artificial Intelligence. USA.

[10]. [Wil, 11] William Vicars, E. (2011). American Sign Language Fingerspelling & Numbers: Introduction. (Bryan Eldredge). USA.

Journal of Theoretical and Applied Information Technology
10th January 2016. Vol.83. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

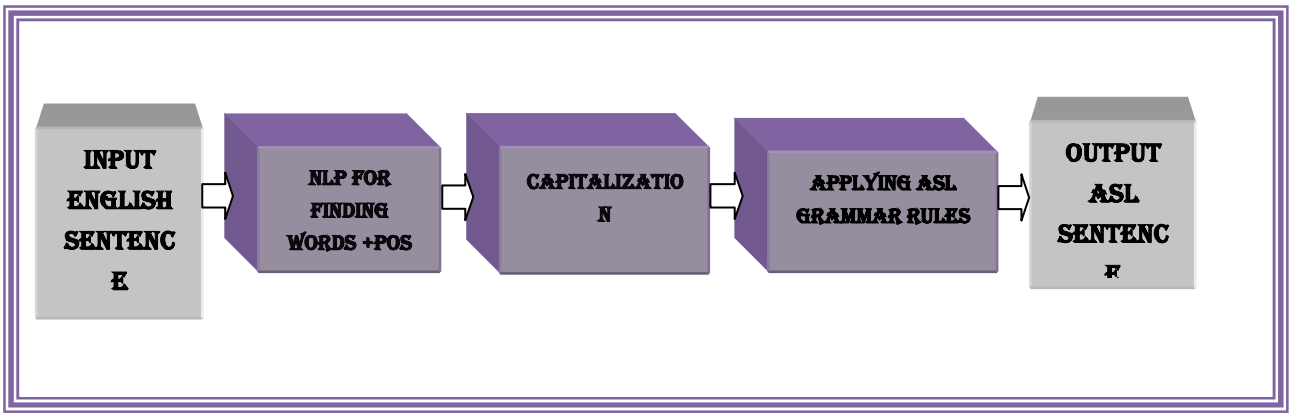


*Figure 2: Block Diagram For The Proposed System*