Support of Arabic Sign Language Machine Translation based on Morphological processing

Sawsan Asjea
Postgraduate Student (PhD.)
Computer Engineering Dept
Faculty of Electrical & Electronic
Engineering
University of Aleppo, Syria

O. Ismail
Professor
Computer Engineering Dept
Faculty of Electrical & Electronic
Engineering
University of Aleppo, Syria

Souheil Khawatmi
Associate Professor
Computer Systems and Networks
Dept.
Faculty of Informatics
Engineering, University of Aleppo,
Syria

ABSTRACT
This paper presents a morphological processing system as a part of arabic text to arabic sign language machine translation system. This morphological processing depends on Farasa analyzer tool, Stanford model and Arramooz lexicon. The characteristics of sign language are achieved to get intermediate arabic sign language sentences. Then these sentences are searched in a sign language dictionary word by word to display the related signs images if available, or to display letters of word using finger spelling alphabet images. The proposed system is tested on many non-vowelized arabic sentences, and good results and high accuracy are obtained.

Keywords
Machine Translation (MT), Morphological Analysis, Arabic sign language.

1. INTRODUCTION
Sign language is the native language of the deaf. The presence of translators from and into sign language is important to ensure communication between deaf and normal people. Machine translation between natural languages has been and remains a major challenge for researchers in the field of natural language processing and machine translation. This challenge increases when the source language and the target language are not identical at the syntactic and morphology levels [1].

Arabic language is very rich in morphological and syntactic structures. Each word in arabic has a certain pattern and morpheme parts which distinct words from each other. Arabic derivational and declension processes allow a wide range of verbs and nouns variation [2], Arabic linguistic resources are computerized as Arramooz lexicon [3], Almaany dictionary [4], and Azhary lexical ontology [5].

Arabic sign language is basically manual language, it has special structure and semantic features. Arabic sign language is made by combination of manual features as hands shape, position of hands in relation to body, directions of hands within space, and non-manual features, like face expression, mouth, and tongue movement[6].The vocabulary in arabic sign language is few and limited comparing with standard arabic. Unified Arabic Sign Dictionary contains 1500 sign in 2001, then 1,600 sign is added in 2007, so arabic sign language is still in developmental stages and many efforts are done to standardize it in arabic world [7].

A prototype of arabic sign language translation system is designed in [8] using example based EBMT approach, morphological processing and root extractor are added in this system for getting syntactic and semantic information in order to increase the accuracy of the system and overcome the limitation of small corpus.

A first version of machine translation system ATLASLang MTS 1 is presented in [9] based on rule-based Interlingua and example-based approaches, the system uses SAFAR Platform and ALKHALIL morph to extract the morphological properties of each word of the input sentence, then it generates a video sequence representing the sentence in arabic sign language, the input of the system must be completely vowelized.

An intelligent system to translate arabic natural language to arabic sign language is designed in [10] based on knowledge base and image processing, the knowledge base solves the problem of few number of words in the arabic sign dictionary compared with words in arabic language dictionaries.

Moreover morphological analysis is used in several applications as machine translation [11], Search Engine [12].

Researchers in the field of natural language processing has interested on designing morphological and syntactical analyzers to use them in improving the performance of natural language processing tasks [13][14]. Farasa is a fast and accurate arabic morphological analyzer tool depends on SVM-rank using linear kernels [15], Farasa tool has the ability of segmentation, part speech tagging, lemmatization, named entity recognition, and parser [16], the Stanford CoreNLP toolkit, an extensible pipeline that provides core natural language analysis is described in [17]. Stanford CoreNLP toolkit has analysis components supported for arabic like Tokenize, Sentence Splitter, POS tagger, Regex NER, Parser [18].

This paper is organized in the following sections, section 2 shows the characteristic of arabic sign language, and section 3 presents importance of the morphological search, whereas section 4 displays the architecture of proposed arabic sign language machine translation system and explains the design of morphological processing, and section 5 illustrates the results and comparison.

2. CHARACTERISTIC OF ARABIC SIGN LANGUAGE
Arabic sign language has a specific syntactic and morphological structures comparing with other languages. For example, the subject precedes the verb, and the most important word comes in the beginning of the sentence like adverbial of time and question keywords. Also, the preposition is not considered to be a syntactic component of

the sign sentence, because the prepositions are implicitly understood by the movement of the sign. Table1 shows several syntactic structures of the sentence in arabic sign language.

On the other hand, many morphological forms in arabic sign language are indicated by special signs, for example plural form is expressed by the sign of the word " a lot " "كثير", a human feminine is expressed by the sign of the word "girl" "بنت", and a dual form is indicated with the sign of the word "two" "اثنان", also the tenses of the verb are presented with special signs as the sign of the word "finish" "انتهى" for the past verb, the sign of the word "now" "الآن" for the present verb and the sign of the word "after" "بعد" for the future verb. Table 2 shows examples of these morphological forms in arabic sign language.

Table 1. Syntactic structures of the sentence in arabic sign language

No.	Arabic Language	Arabic Sign Language
1	Verb+ Subject+ Object	Subject+ Verb+ Object
2	Verb+ Subject+ Adjective	Subject+ Adjective+ Verb
3	Verb+ Subject+ Object+ Adjective	Subject+ Verb+ Object+ Adjective
4	Verb+ Subject+ adverbial of time	adverbial of time + Subject+ Verb
5	Why+ Verb+ Subject+ Object+?	? + Subject+ Verb+ Object+ "reason"
6	Where + Verb+ Subject+ Object+?	? + "place" + Subject+ Verb+ Object
7	Verb+ subject+ preposition+ noun after preposition	Subject+ verb +noun after preposition

Table 2. Examples of morphological forms in Arabic sign language

No.	Arabic word	Arabic sign language word	Morphological form
1	manager "مديرة"	manager + girl "مدير"+"بنت"	human feminine
2	workers "عاملات"	worker+ girl+ a lot "عامل" +"بنت"+ "كثير"	regular plural human feminine
3	library "مكتبة"	library "مكتبة"	inhuman feminine
4	books "كتب"	book+ a lot "كتاب"+ "كثير"	irregular plural masculine
5	farmers	farmer + a lot	regular plural
	"فلاحون"	"فلاح" + "كثير"	masculine
6	tow worker "عاملان"	worker+ tow "عامل"+ "اثنان"	dual
7	"يعمل" works	work+ now "عمل" +"الآن"	present verb
8	attended "حضر"	attend + finish "حضر"+ انتهى"	past verb

3. IMPORTANCE OF THE MORPHOLOGICAL SEARCH

Algorithms of the Morphological search depend on the morphological analysis of the words [11] [19].

Sign language dictionary contains the signs of the words as images, the arabic words which have the same lemma is indicated by the same image, so the arabic sentences need arabic morphological processing for getting the lemma of the arabic words. For instance, the sign of arabic word "worker" "عامل" is the same sign of the words "workers" "عاملون", "two workers" "عاملان", "the worker" "العامل", "the workers" "العمال" or "العاملات", "العاملون", so only one word have to be found "worker" "عامل" from morphological processing, then this word will be searched from arabic sign dictionary.

4. ARCHITECTURE OF THE PROPOSED ARABIC SIGN LANGUAGE MACHINE TRANSLATION SYSTEM

Arabic sign language machine translation system is used for converting text from arabic language (source language) into images of arabic sign language (target language) automatically. The implementing of this proposed system needs two tasks of processing, these tasks are:

1- **Syntactic processing:** which is used in creating the structure of the intermediate arabic sign language sentences based on the syntactic structure of arabic sentences and syntactic structure of arabic sign language sentences.

2- **Morphological processing:** which is used for achieving morphological forms in arabic sign language, and doing morphological search in arabic sign language dictionary.

So, when the arabic sentence is entered to system, morphological processing using arabic lexicon is done, and syntactical processing using syntactical templates will take place, then intermediate arabic sign language sentence is generated based on the combination results of morphological processing and syntactical processing, finally, every word in intermediate arabic sign language sentence is searched for in arabic sign language dictionary to display related sign images. Figure 1 shows the architecture of proposed arabic sign language machine translation system.

This paper basically focuses on morphological processing task by designing and implementing morphological processing system which is described in the next paragraph.
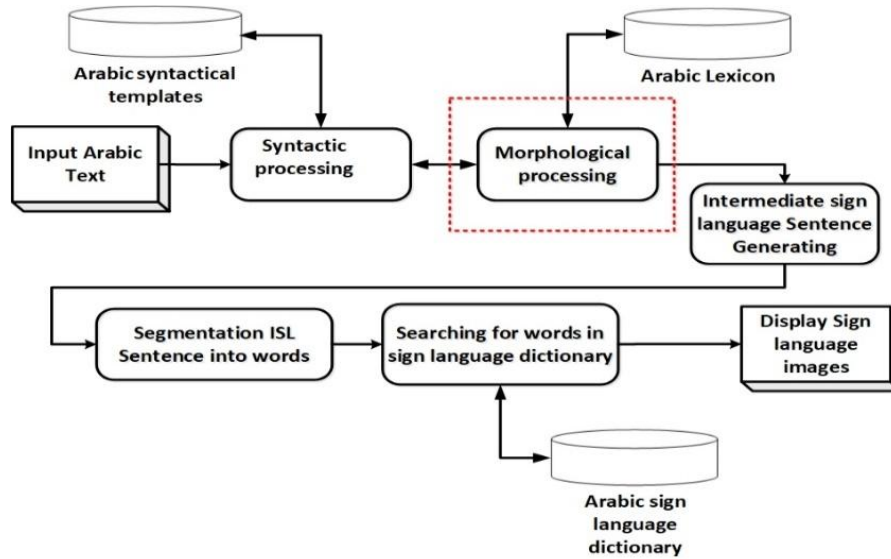
Fig 1: Architecture of the Arabic Sign Language Machine Translation System

4.1 Design and Implementation of proposed Morphological Processing System

Morphological processing system is essential part in arabic sign language machine translation system, and it is word based analysis. So, at first, the arabic sentence is split into words, and every word is passed to all phases of the proposed morphological system as shown in figure 2 as follows:
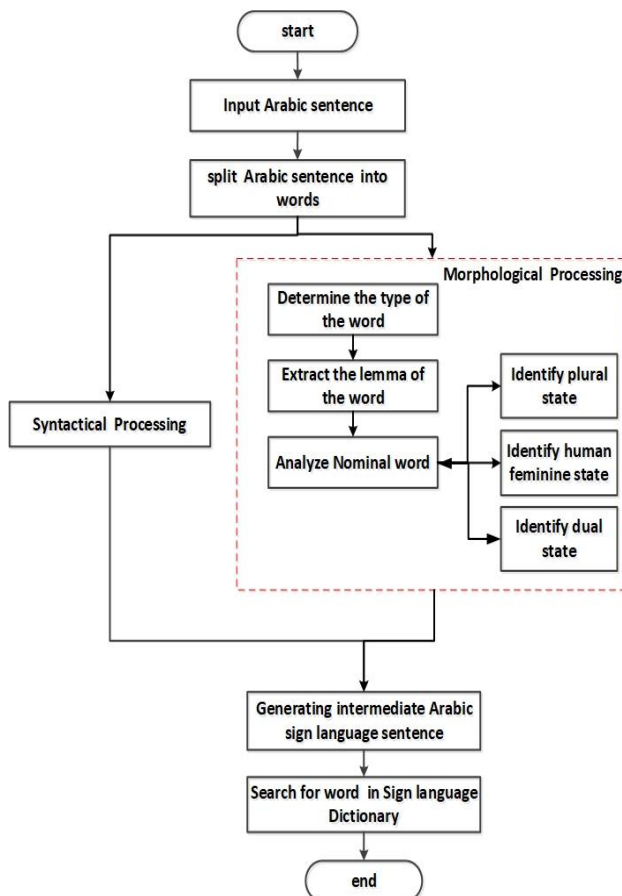

Fig 2: The phases of the design proposed morphological system

4.1.1 Determining the type of the word:

The type of the word in arabic is a name, a verb or a character. The names are distinguished from the verbs in arabic by different patterns, prefixes and suffixes.

The distinguishing process between names and verbs is a basic task to achieve the characteristics of words in the sign language, such as plural, dualization, feminization, and the tenses of the verb. Stanford POS tagger, the one of the most popular and accurate Arabic Part Of Speech (POS), is used in this stage as follows:

$$S = w0 …. wn$$
$$T = t0 …… tm$$

where, S is arabic sentence of n words, T is intermediate Arabic sign language sentence of m words.

R1#: If posstanford(wi)={VBD} then type(wi) = past verb.

R2#: If posstanford(wi)={VBP} then type(wi) = present verb.

R3#: If posstanford(wi)={VBN}and prefix(wi)={"ya" 'ي', "ta" 'ت', "noon" 'ن', "hamza" 'أ'} then type(wi) = present verb.

R4#: If posstanford(wi)={DTNN| NN| DTJJ| NNP| DTNNS,JJ,NNS} then type(wi) = noun.

Table 3 shows examples for determining the type of the arabic words using Stanford POS tagger.

Table 3. Examples for determining the type of the word using Stanford Part Of Speech model

Arabic Input Sentence	Stanford Part of Speech
The manager attended the party "حضرت المديرة الحفلة"	VBD DTNN DTNN الحفلة المديرة حضرت
The workers help the human "يساعد العاملون الإنسان"	VBP DTNNS DTNN الانسان العاملون يساعد
Pressure gauge device consists of pipes "يتألف جهاز قياس الضغط من أنابيب"	VBP NN NN DTNN IN NN انابيب من الضغط قياس جهاز يتألف

The skull protects the brain "تحمي الجمجمة الدماغ"	
The two doctors treated the children "عالج الطبيبان الأولاد"	

4.1.2 Extracting the lemma of the word:

Some verbal shifts occur in arabic when the verb is connected with pronouns, for example "he came" "جاء", "they came" "جاؤوا","I came" "جئت" and "he came back" "عاد", "he comes back" "يعود","we came back" "عدنا". Also, the forms of the nouns are changed when the noun is singular, dual or plural or when nouns are connected to pronouns. So the lemma of the noun or the verb is extracted to get the basic form of the word without any changes. In addition, the searching process in arabic sign dictionary will be done depending on lemma of the word.

Farasa lemmatization tool is used for extracting lemma of the arabic word because it is accurate, whereas Stanford can not. Table 4 shows some examples for extracting lemma of the words.

Table 4. Examples for extracting the lemma of the words using Farasa lemmatization tool

Lemma of Words in Farasa	Arabic Input Sentence
attend, manager, party "حضر" "مدير" "حفل"	the manager attended the party "حضرت المديرة الحفل"
worker, help, human "ساعد" "عامل" "إنسان"	the workers help the human "يساعد العاملون الإنسان"
pressure, gauge, consist, pipe "تألف" "جهاز" "قياس" "ضغط" "أنبوب"	pressure gauge consists of pipes "يتألف جهاز قياس الضغط من أنابيب"
skull, protect, brain "حمى" "جمجمة" "دماغ"	the skull protects the brain "تحمي الجمجمة الدماغ"
doctor, treat, child "عالج" "طبيب" "ولد"	the two doctors treated the children "عالج الطبيبان الأولاد"

4.1.3 Analysis of the nominal word:

The nominal word in this phase is analyzed to match one or two forms of plural, dualization, and feminization.

1- identify of the plural form:

The plural form is obtained based on the Farasa tool and the Stanford model for morphological analysis. Farasa is accurate in determining irregular plural form (P) of the word. Stanford is accurate in determining regular plural form (NNS, DTNNS) of word, so the result of Stanford and Farasa analysis is combined to determine the plural form of word that is analyzed. Two rules are proposed to overcome the mistakes that may occur when determining the morphological form of word as a plural instead of a singular.

The first rule R5 adds a condition to the plural word that the suffix of word is ("ya" 'ي', "waw" 'و', "alef ta" "ات" "ya noon"

"ين" "waw noon" "ون"), and the word doesn't equal to the lemma of the word.

R5#: if typestanford (wi)=noun and (posstanford ={DTNNS|NNS} || posfarasa=P) and suffix(wi) = { ون | ين | ات | و| ي } and lemma(wi)≠ wi then number(wi)=plural "جمع"

The second rule R6 adds a condition to the plural word that the word without "al" "ال" is must be a broken plural in Arramooz.

R6#: if typestanford (wi) =noun and (posstanford = {DTNNS|NNS} || posfarasa=P) and broken_pluralarramooz(wi without "ال") = true then number(wi)= plural "جمع".

Table 5 shows examples for determining the plural form in proposed system, Where the symbol (✓) represents a correct case and the symbol (✗) is a false case.

Table 5. Determining the plural form in the proposed system

word	plural form in arabic language	plural form in Stanford	plural form in Farasa	pro_posed rule	plural form in proposed system
farmers "فلاحون"	regular plural masculine	plural (✓)	singular (✗)	R5	plural (✓)
workers "عاملون"	regular plural masculine	plural (✓)	plural (✓)	R5	plural (✓)
writers "كاتبات"	regular plural human feminine	singular (✗)	plural (✓)	R5	plural (✓)
lands "أراضي"	irregular plural	singular (✗)	plural (✓)	R6	plural (✓)
children "أولاد"	irregular plural	singular (✗)	plural (✓)	R6	plural (✓)
manager "مديرة"	singular	singular (✓)	plural (✗)	R5,R6	singular (✓)
mad "مجنون"	singular	singular (✓)	singular (✓)	none	singular (✓)

For example, in table 5 the word worker "عاملون" is plural in Stanford and singular in Farasa so the rule 4 determines it as plural in the proposed system, because suffix is "waw noon" "ون" and the word does not equal the lemma of it, workers "عاملون" is not equal to worker "عامل". Also the word children "أولاد" is singular in Stanford but plural in Farasa so the rule 5 determines it as plural in proposed system because the word is in Arramooz lexicon as irregular plural (broken plural). The proposed rules are needed to limit the mistakes in

a decision making of many cases, as the word manager "مديرة", in spite of Farasa determines it as plural but it is not. If the word is not determined as plural in Farasa or Stanford then the rules are not applied, as the word mad "مجنون".

2- identify of the dual form:

Dual form is not supported in Stanford model, so the word is determined as dual using Farasa tool, or by ensuring that the suffix of the word is "alef noon" "ان", and the word is not equal to the lemma of it.

R7#: If typestanford (wi)=noun and (posfarasa=D || (suffix(wi)= "ان" and lemma(wi)≠ wi)) then number(wi)=dual.

Table 6 shows examples for determining the dual case using R7, where the symbol (✓) represents a correct case and the symbol (✗) is a false case.

Table 6. Determining the dual form in the proposed system

word	dual form in arabic language	dual form in Farasa	proposed rule	dual form in proposed system
two workers "عاملان"	dual	dual (✓)	R7	dual (✓)
two Painters "رسّامان"	dual	singular (✗)	R7	dual (✓)
Thirsty "عطشان"	singular	singular (✓)	R7	singular (✓)
brightness "لمعان"	singular	singular (✓)	R7	singular (✓)

For example, in table 6 the first word two workers "عاملان" is determined by Farasa tool as dual, but the word two painters "رسامان" is determined as dual because the suffix is an "ان" and the word "رسامان" is not equal to the lemma of word painter "رسام", but in the rest examples, the result of the rule is false because the words and the lemma of them are the same.

Notice that the dual in the case of accusative and genitive is ignored because the case markers are the same in the plural and the dual and it cannot distinguish from each other without vowels.

3- identify of the human feminine form:

The following rules are proposed to find the form of femininity and to distinguish the human feminine form.

R8#: If typestanford (wi) =noun and suffix (wi) ='ة' then gender1 (wi) = feminine.

R9#: If gender1 (wi)= feminine and feminable_valuearramooz (wi without 'ة') =Ta then gender(wi)= is human feminine

Every word in Arramooz lexicon has many properties, one of

them is "feminable", the value of the "feminable" property is Ta when the feminine is human and the value is "Ta" with double quotation for feminine which is not human like animal, or the value is blank for words which has the letter 'ة' as a basic letter. Some words has to be vowelized to distinguish it as human feminine or not like Syrian women "سوريَّة" and the country Syria "سورية". Table 7 shows examples for determining human feminine form of the word and comparing results of the proposed rules with Farasa tool results.

Table 7. Determining the human feminine form in proposed system

word	human feminine form in arabic language	human feminine form in Farasa	Pro_posed rule	human feminine form in proposed system
schoolgirl "طالبة"	human feminine	feminine	R8, R9	human feminine
manager "مديرة"	human feminine	Masculine	R8, R9	human feminine
skull "جمجمة"	non-human feminine	feminine	R8, R9	non-human feminine
bird "عصفورة"	non-human feminine	feminine	R8, R9	non-human feminine

4.2 Syntactic Processing:

Syntactical processing is done in previous works at the level of sentence structure. The syntactic processing studies the structure of arabic sentences, and the structure of arabic sign language. many researches utilized the syntactic processing and needed morphological analysis for translation as in [20], which used colored petri net (CPN) in the transformation stage from the grammatical components of the arabic sentence (the source), to the intermediate structure of the grammatical components of the arabic sign language sentence (target language).

4.3 Generating Intermediate Arabic Sign Language:

The generating of the intermediate arabic sign language sentences need to combine the results of the morphological and syntactic processing systems. The arabic input sentence is split into words, each word is processed and the output of morphological processing is arranged according to output of syntactic processing which is the grammatical structure of the sentence in arabic intermediate sign language. Table 8 shows examples of combination the results of morphological and syntactical processing to generate intermediate arabic sign language sentence.

R10#: If type(wi) = present verb then tj=lemma(wi)+ "now" "الآن"

R11#: If type(wi) = past verb then tj=lemma(wi)+ "finish" "انتهى"

R12#: If gender2(wi)= is human feminine then ti=lemma(wi)+"girl" "بنت"

R13#: If number(wi)= plural then tj=lemma(wi)+ "a lot" "كثير"

R14#: If number(wi)=dual then tj=lemma(wi)+"two" "اثنان"

where $wi \in W$ and $ti \in T$.

Table 8. Generating intermediate arabic sign language

Input Arabic sentence	Result of syntactical processing	Result of morphological processing	Output Arabic Intermediate sign sentence
the manager attends the party "حضرت المديرة الحفلة"	verb + subject +object→ subject + verb+ object "فعل+ فاعل+ مفعول به" ← "فاعل+ فعل+ مفعول به"	"attends" "حضرت" ← "attend" "finish" "حضر" "انتهى" "the manager" "المديرة" ← "manager""girl" "مدير" "بنت" "the party" "party"←"الحفلة" "حفلة"	manager girl attend finish party "مدير بنت حضر انتهى حفلة"

4.4 Searching in Arabic Sign Language Dictionary:

There are no digital dictionaries available for arabic sign language. Building digital arabic sign dictionary needs long time, effort and experience. A digital sign language dictionary is constructed by mapping the lemma of arabic word with the related image of the sign, jpg picture is used and they are obtained from Unified Arabic Sign Dictionary hardcopy. There is a concept of phrase in arabic sign language dictionary, the phrase consists of two or three words and it is indicated with one sign image.

The searching process of each word of intermediate arabic sign language sentence in the digital arabic sign language dictionary will be done as follows:

1) The word is not available in the dictionary: the word is divided into its constituent letters and the corresponding sign image of letters are displayed.

2) The word is available in the dictionary: check the matching's number and the search for phrase is tested.

4.4.1 Search Algorithm in Arabic Sign Language Dictionary:

When the searching process for the arabic word in the sign dictionary is done, there are two cases:

1) The first case:

- The word is found in one match, then the count of words in this match is checked.

- The count of words is one then the sign image of the word is displayed, else, the second and third word are searched for.

2) The second case:

- The word is found in more than one match, then the count of words for every match is checked.

- The count of words is checked in all matches, then the word is single/not phrase, or a part of phrase of two or three words.

The algorithm steps is shown in figure 3.
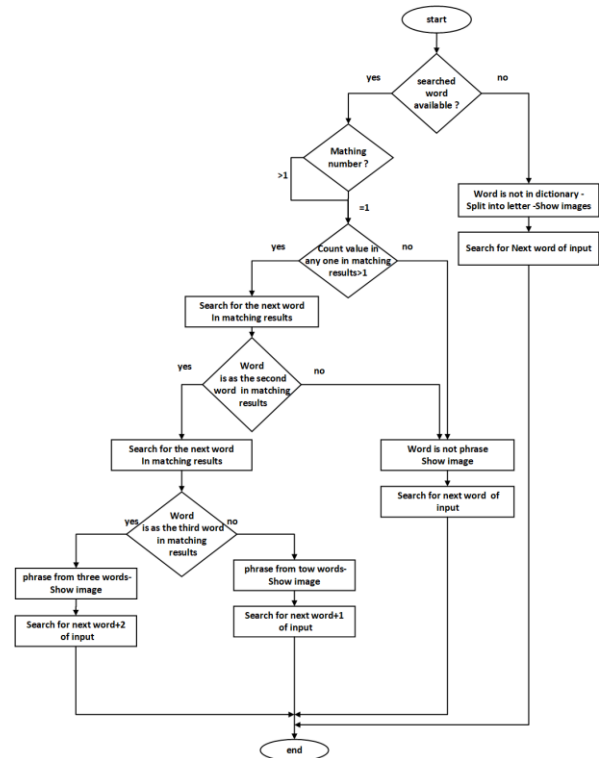
Fig 3: Search Algorithm in Arabic Sign Language Dictionary

For example, when the input text starts with the sentence "pressure gauge device" "جهاز قياس ضغط" and the word device "جهاز" is searched in dictionary, three matches are found.

1) "Device" "جهاز", the count of the words =1. In this match the count is one so it is not a phrase.

2) "Digestive system" "جهاز الهضم", the count of the words=2 , so phrase of two words is tested, the first word in match 2 is the same word in input sentence, but the second word in the match 2 is not the same word in input sentence, so it is not phrase of two words.

3) Temperature measuring device "جهاز قياس الحرارة", the count of words=3, so phrase of three words is checked, the second word in match 3 is the same of the input sentence, but the third word of the match 3 is not the same third word in the input sentence, so there is no phrase of three word.

4) "Pressure gauge device" "جهاز قياس الضغط", the count of the words=3, so phrase of three words is checked, the second word in match 4 is the same word of the input sentence, and the third word in match 4 is the same word in the input sentence, so it is phrase of three word.

After passing all matches and displaying the sign image of the phrase, moving to the next word in the input sentence is done. The next word is the word+1 if searched word is not a phrase, the word+2 if the searched word is identified as phrase of two

words, and word+3 if the word is determined as phrase of three words.

5. RESULTS AND COMPARISON

The proposed morphological system is applied to many non-vowelized arabic sentences. The results are combined with the results of the syntactic processing system in order to obtain the intermediate arabic sign language sentences. Each word of the intermediate arabic sign language sentence is searched in the arabic sign language dictionary to display the related sign images. The proposed system is implemented on a laptop with a Core i5 2.50 GHZ processor and Windows10 operating system, using java programming language. Stanford and Farasa software are invoked as free jar files in the NetBeans 8.0.2 environment.

Example1:

1) Input arabic sentence: The farmers implants the lands "يزرع الفلاحون الأراضي"

2) The sentence is split into words The farmers "الفلاحون"- implants "يزرع" – the lands "الأراضي".

3) Morphological processing:

   - **The type of words:** Implants "يزرع"is determined by R2 as present verb because it is marked VBP in Stanford. The farmers "الفلاحون" and the lands"الأراضي" are determined by R4 as noun because they are marked DTNN and DTNNS in Stanford.

   - **The lemma of words:** The lemma of words are "implant" "زرع", "farmer" "فلاح" and "land" "أرض" using Farasa.

   - **The nominal Analysis**: the word the farmers "الفلاحون" is determined as plural by R5, and the word the lands"الأراضي" is determined as plural by R6.

4) **Syntactic processing:** the arabic sentence has a structure of verb+ subject+ object, so it is transferred by previous work to the structure of arabic sign language which is subject+ verb + object.

5) **Generating Integrated Arabic sign language:** the subject is generated by R12 as "farmer" "فلاح" "a lot" "كثير", verb is generated by R10 as "implant" "زرع" "now" "الآن" and object is generated by R13 as "land" "أرض""كثير" "a lot".So the output is generated "farmer a lot implant now land a lot" "فلاح كثير زرع الآن أرض كثير".

6) **Search in Arabic sign language dictionary:** the words "the farmer" "فلاح", "implant" "زرع", "now" "الآن" are found as one match and the count of word is one so each word is not a phrase then the image is displayed, the words "a lot" "كثير", "land" "أرض" are not found in dictionary so the images of the letters ('ك', 'ث' , 'ي','ر') and ('أ', 'ر' , 'ض') are displayed.

So the output arabic sign language is displayed as follows.

"farmer" "فلاح"   "a lot" "كثير"   "implant" "زرع"

"now" "الآن"   "land" "أرض"   "a lot" "كثير"
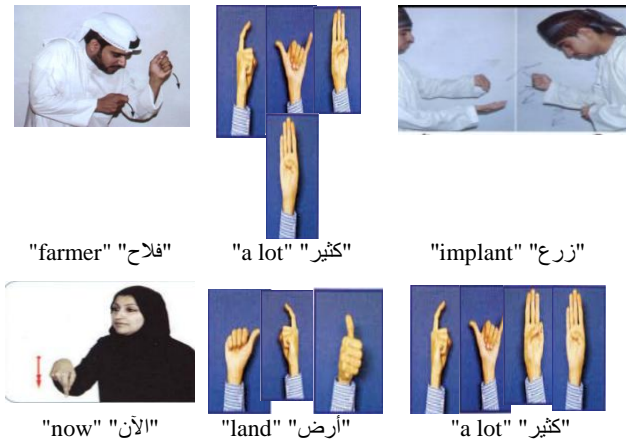
Example2:

1) Input arabic sentence The doctor used a pressure gauge in the clinic

"استخدمت الطبيبة جهاز قياس الضغط في العيادة"

2) Output intermediate arabic sign language sentence: Doctor girl use finish pressure gauge device clinic

"طبيب بنت استخدم انتهى جهاز قياس ضغط عيادة"

"doctor" "طبيب"   "girl" "بنت"   "use" "استخدم"

"finish" "انتهى"   "pressure gauge device "جهاز قياس الضغط"   "clinic" "عيادة"

In comparing the results in this research with others [9][10], it turns out that:

The topic	The current research	research [9]	research[10]
The input	non-vowelized sentences	vowelized sentences	some vowelized words
used utilities NLP tools	Farasa tool, Stanford model Arramooz lexicon	SAFAR Platform and ALKHALIL morphological system	none
methodology of morphological processing	morphological processing depends on Farasa and Stanford and proposed rules are designed to improve results	stem - based stemming, comparing resulted stem with pattern, and root- based stemming. Using ALKHALIL	-building knowledge base to solve a number of Arabic language

		morph for pre-processing text Using SAFAR to convert the output of analyzer from html to xml format	problems -extracting stem by key word processing
human feminine Processing	determine human feminine	determine feminine	determine feminine
word connected with pronoun Processing	transform the word to the lemma	transform connected pronouns to corresponding separated ones	transform the word to the lemma
search in dictionary	lemma based search	word based search or root based search	word based search or stem based search
morpho-logical cases	lemma- type of word-number-gender-human feminine- the tense of verb	type of word-number- gender-the tense of verb	lemma- stem gender-number- the proper noun.
output arabic sign language sentence	static images	2D animation images	static images

In according of table 9, the current system is easily suitable for arabic users where the input of the system is non-vowelized sentence, and more morphological features are determined like human feminine. The current system depends on Farasa and Stanford which give one result compared with ALKHALIL which gives many results, or maybe doesn't give any result. Lemma based search overcomes the root based search because more than one word has the same root as "martyr" "شهيد", "witness" "شاهد" and each of them has a different sign image in the sign dictionary. Usually, most dictionaries don't contain all forms of the words, so word based search is often considered not useful.

6. CONCLUSION

In this paper, a morphological processing system is proposed to achieve the characteristic of arabic sign language at the level of word, this system is a part of the proposed arabic sign language Machine translation system. The morphological system extracts lemma of the arabic word, identifies the type of word and determines the plural, dual and human feminine morphological forms. Many rules are proposed to use with Stanford model, Farasa tool and Arramooz lexicon, so the system overcomes the limitations and the weakness of morphological analyzers of the arabic language, i.e.,(inability to contain all the advantages of the arabic language).

A digital sign language dictionary is constructed by mapping the lemma of the arabic word with the related picture of the sign because the lack of digital arabic sign language dictionary.

The proposed search algorithm takes into account phrases in sign language. Many non-vowelized arabic sentences have been translated into arabic sign language with high accuracy

and good performance.

This work provides a contribution to morphological analysis development field of arabic language to suit the arabic sign language.

Arabic sign language dictionary will be developed in future work to display the signs by animation 3D avatar.

7. ACKNOWLEDGMENTS

8. REFERENCES
[1] Okpor, M. D. 2014. Machine Translation Approaches: Issues and Challenges. IJCSI International Journal of Computer Science Issues. 11(5). 159-165.

[2] Ryding, K. C. 2005. A reference grammar of modern standard Arabic (chapter 3.Arabic Word Structure: An Overview). Cambridge University press, 44-56.

[3] Arramooz Alwaseet: Arabic Dictionary for Morphological analysis, http://arramooz.sourceforge.net/ [Date Accessed 16-3-2019].

[4] The Meanings Dictionary, http://www.almaany.com/ [Date Accessed 23- 9 -2019].

[5] Ishkewy, H., Harb H., Farahat, H. 2014. Azhary: An arabic lexical ontology. International Journal of Web & Semantic Technology (IJWesT). 5(4). 71-82.

[6] Abdel-Fattah, M. A. 2005. Arabic sign language: a perspective. Journal of deaf studies and deaf education. 10(2). 212-221.

[7] Samrine, S., El Benali M., Standard Arabic Sign Language Rules. Qatar National Library ISBN: 99921-53- 41- 0.

[8] Almohimeed, A., Wald M., Damper, I., 2011. Arabic text to Arabic sign language translation system for the deaf and hearing-impaired community. The Second Workshop on Speech and Language Processing for Assistive Technologies. Association for Computational Linguistics, 101-109.

[9] Brour, M., Benabbou, A., 2019. ATLASLang MTS 1: Arabic Text Language into Arabic Sign Language Machine Translation System. Procedia computer science. 148. 236-245.

[10] El- Alfi, A., El- Basuony, M., El- Atawy, S. 2014. Intelligent Arabic text to Arabic Sign Language Translation for Easy Deaf Communication. International Journal of Computer Applications. 92(8). 22-29.

[11] Abu Shquier, M. M., Atoum, M. S., Abu Shqeer, O. M. 2017. Arabic to English Machine Translation. Proceedings of the New Trends in Information Technology (NTIT). 118-124.

[12] Hattab, M., haddad, B., yaseen, M., duraidi, A. 2009. Addaall arabic search engine: Improving search based on combination of morphological analysis and generation considering semantic patterns. In The second International Conference on Arabic Language Resources and Tools, Cairo, Egypt, 159-162.

[13] Lakhfif, A., Laskri, M.T. 2016. A frame-based approach for capturing semantics from Arabic text for text-to-sign

language MT. International Journal of Speech Technology, 19(2). 203-228.

[14] Boudchiche, M., MazrouI, A., Bebah, M., Lakhouaja, A., Boudlal, A. 2017. AlKhalil Morpho Sys 2: A robust Arabic morpho-syntactic analyzer. Journal of King Saud University-Computer and Information Sciences, 29(2), 141-146.

[15] Abdelali, A., Darwish, K., Durrani, N., Mubarak, H. 2016. Farasa: A fast and furious segmenter for arabic. Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, 11-16.

[16] FARASA : QCRI Arabic Language Technologies Tools & Demos, Qatar Computing Research Institute, http://qatsdemo.cloudapp.net/farasa/demo.html. [Date Accessed 16- 3 -2019].

[17] Manning, C., Surdeanu, M., Bauer, J. 2014. The Stanford CoreNLP natural language processing toolkit. Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations, 55-60.

[18] Stanford CoreNLP: http://corenlp.run/. [Date Accessed 16 -3 -2019].

[19] Khadim, S.M., Abdulmunim, M.E. 2013. Design a Computerized Lexicon for Machine Translation from Arabic to English. University of Technology. Baghdad, 191-202.

[20] Sawsan A., Ismail O., Khawatmi S., 2018. Prototype System Construction for Transfer Based Machine Translation Using Colored Petri Net Case Study (Translate Arabic language into Arabic Sign Language). Res. J. of Aleppo Univ., Engineering Science Series (2). 139.1-22.