# Automated Arabic-Arabic sign language translation system based on 3D avatar technology

**Ahmed A. Alethary[1], Ahmed Hussein Aliwy[1], Nabeel Salih Ali[2]**
[1]Faculty of Computer Science & Mathematics, University of Kufa; Najaf, Iraq
[2]Information Technology Research and Development Centre, University of Kufa, Najaf, Iraq

**Article Info**

***Article history:***

Received Aug 16, 2022
Revised Nov 3, 2022
Accepted Nov 21, 2022

***Keywords:***

3D avatar
Arabic language
Arabic sign language
ArSL
Deaf people
ESign
Machine translation

**ABSTRACT**

Arabic sign language (ArSL) is the natural language of the deaf community in Arabic countries. Deaf people have a set of difficulties due to poor services available. They have problems accessing essential information or receiving an education, communicating with other communities, and engaging in activities. Thus, a machine translation system of Arabic to ArSL has been developed using avatar technologies. Firstly, a dictionary of ArSL was constructed using eSign editor Software. The constructed dictionary has three thousand signs. It can be adopted for the translation system in which written text can be transformed into sign language. The dictionary will be available as a free resource for researchers. It is complex and time-consuming, but it is an essential step in the machine translation of whole Arabic text to ArSL with 3D animations. Secondly, the translator has been developed. It performs syntactic and morphological analysis and then applies a set of rules to translate an Arabic text into ArSL text based on the structure and grammar of ArSL. The system is evaluated according to the parallel corpus that consists of 180 sentences using the metric for evaluation of translation with explicit ordering metric for evaluation of translation with explicit ordering (METEOR) our system achieves a relative score of (86%).



***Corresponding Author:***

Nabeel Salih Ali
Information Technology Research and Development Center (ITRDC), University of Kufa
Kufa Street, Main Campus, University of Kufa, Najaf Governorate, Iraq
Email: Nabeel@uokufa.edu.iq

## 1. INTRODUCTION

Machine translation (MT) is a subfield of natural language processing (NLP). It is the translation process from one natural language to another using machines or computers. MT is one of the most challenging applications in natural language processing, resulting from many difficulties, such as differences in grammar, morphology, and semantic analysis [1]. Arabic language and ArSL are natural languages for normal and deaf Arabic persons, respectively.

The Arabic language is one of the most common languages in the world. The Arabic text is read and written from right to left, and there are three categories of Arabic language classical Arabic (CA) language, modern standard Arabic (MSA), and dialect. It consists of twenty-eight letters for all its varieties, where MSA is the most popular. CA is the language of Islam's holy book, "the Qur'an", but Arab people use MSA for reading, writing, and formal communication, such as magazines, books, television, and movies [2]. The Arabic language is a very ambiguous language with numerous challenges and complexities during the implementation of NLP due to the nature of the language. For example, the word "سنشاهدكم" (we will watch you) matches a sentence in English. The Arabic word may be analyzed as proclitic, prefix, stem, root, lemma, suffix, and enclitic [3]. Diacritic marks must be added in some circumstances to differentiate similar words

("دَرسَ" means studied, while "درسْ" means lesson). Arabic word input, especially for MT, should not be written without diacritics since it will be interpreted differently, making it very difficult to define the correct meaning.

Deaf or hard-of-hearing people have hearing loss, which prevents them from acquiring linguistic information through hearing and communication difficulties in society [4]. Moreover, Deaf persons have a problem in reading and dealing with written or spoken language. Deaf communities have numerous distinctive cultural aspects. They obtain language in various ways, based on their house circumstances where language construction plays an essential role in reading and writing learning. Therefore deaf people use sign language for essential occasions where communication must be evident [5]. Sign language (SL), it consists of manual features (MF) and non-manual features (NMFs). MF is performed by both or one hand in different shapes, locations, orientations, and movements. Facial expressions and body movements perform NMFs to embody the meaning of words and terms of a particular significance. In SL, the representatives are specific hand movements and shapes associated with facial expressions to express the meaning of these symbols, called signs.

Arabic sign language (ArSL) has grammar, sentence structure, dialect, characteristics, style, and regional variations like any other language [6]. ArSL is not universal in the Arab world alternatively it varies from one country to another and maybe differs in the same region. ArSL varieties include Saudi, Iraqi, Moroccan, Egyptian, Levantine, Yemeni, Libyan, Kuwaiti, Qatari, Emirati, and Omani sign language. In 1999, a dictionary consisting of approximately 3200 words was published in two parts produced by the League of Arab States (LAS) and the Arab league educational, cultural and scientific organization (ALECSO) to standardize the ArSL [7].

In this paper, i) We study ArSL and explain its orthography; phonology, structure, vocabulary, and grammar; ii) Construct a dictionary for ArSL using a 3D human avatar; and iii) Develop a translation system of Arabic text to Arabic sign language using modern technologies.

In recent years, researchers increased attention to sign language because different systems have been developed for various languages; these systems are based on grammatical rules, statistics, or examples. Many machine translators have been developed for the majority of languages in the world, but fewer of them are focused on translation systems to sign language. In this section, the related works will be introduced a MT system for non-Arabic language and Arabic languages.

For non-Arabic sign language, the work in Zhao *et al*. [8], developed the project of English language to American Sign Language (EL-ASL) at the University of Pennsylvania. It uses syntactic structure with the parsing of an English input text. The output of the linguistic portion was a written ASL gloss notation with embedded parameters such as Facial expressions and sentence mood as features associated with the individual word. The output of the synthesis module is animated using an avatar (animated human model). Safar and Marshall [9], the authors constructed an architecture of a translation system from English text into British sign language (EL-BSL) at East Anglia University. Their method adopted the CMU link grammar parser and WordNet to facilitate the design and implementation of the translation system.

Furthermore, they used prolog declarative clause rules to transform this linkage output into a discourse representation formation. A symbolic sign language (SL) representation script is generated using head-driven phrase structure rules. This script is the signing gesture markup language (SiGML), a symbolic coding scheme for the movements required to perform sign language. Moreover, the article by Segundo *et al*. [10] developed a translation system from Spanish speech into Spanish sign language (SS-SSL). The proposed system consists of four stages: i) speech recognizer, ii) semantic analysis, iii) gesture sequence generation, and iv) gesture playing. They used voice software developed by IBM for Spanish for speech recognition. They used the Phoenix v3.0 parser produced at the University of Colorado for semantic analysis and applied the generated rules for the sign sequence generation module. For output, an animated character is developed to represent sign language.

Further, the study Wang *et al.* [11] proposed a translation system on mobile devices for deaf people from the Chinese language to Chinese sign language (CL-CSL) with an animation system. The system includes three stages the first stage is to apply a semantic analysis algorithm that segments Chinese text into a set of sign word units. The second stage is searching signs based on text and storing signs in random access memory (RAM) to reduce seeking time. The third stage is to display and render a sign on the screen using a 3D avatar.

For the Arabic language, the Mohandes [12] developed a translation system from Arabic text into Arabic SL (ArL-ArSL). The system possesses a database of video files. The experts in sign language recorded each sign as a video sequence based on the unified Arabic sign dictionary. Firstly words are entered, then translated individually into their corresponding signs, and the following pre-recorded video clips are presented on a web browser. The system looks for each word in a database, and the corresponding video is immediately displayed if it is found. If the word does not have a sign in the database, the system spells its letter by letter and then shows the sequence of the obtained signs. Besides, Almohimeed *et al.* [13], used a chunk-based system to translate Arabic text to ArSL.

The system used google tashkeel to vocalize the text and Buckwalter's morphological analyzer to analyze the text. The latest stage is recombining ArSL sentences in a windows media video (WMV) format. The researchers constructed annotated ArSL corpus using the ELAN annotation tool with the guidance of three ArSL signers. The corpus includes 203 sentences with 710 separated signs. In addition, the research in Alfi and Atawy [14] proposed a translation system from Arabic text into ArSL. They tried to solve Arabic language problems such as synonyms, derivational, inflectional, diacritical, and plural based on a knowledge base. A rule-based technique is applied to convert an Arabic word into its equal sign. They solved the variations between Arabic and ArSL at the word level and ignored the structure and grammar differences at the sentence level. The proposed system displays the sign sequence using a set of images. Also, the authors Luqman and Mahmoud [15] proposed a rule-based translation system to convert Arabic text into ArSL.

Furthermore, the developed corpus consists of 600 sentences in the health domain. Moreover, suggest a gloss system to represent ArSL. The system performs a morphological analysis using The Madamira analyzer [16] and a syntactic analysis using the camel parser [17], then apply the rules to translate the sentence into a structure of ArSL. The system used GIF images to display signs.

We can summarize the following Based on our ArSL machine translation survey: i) The lack of linguistic studies on ArSL, mainly on grammar and structure, leads to a misunderstanding of language and makes the researchers construct unreliable ArSL systems; ii) Lack of available dictionaries for displaying ArSL. Even if some images and video data are accessible online, these data are not annotated, plus the signs are not segmented, which makes this data insufficient for use in sign language systems; and iii) Also, there is a problem with evaluating SL output which is a problem for the MT system in general.

## 2. BACKGROUND

### 2.1. Electromyography sensing

ArSL is similar to other sign languages in that they are mainly spatial gestural languages. This makes it a challenge to compare sign languages with their spoken. Several concepts were utilized to represent spoken languages, but they are inadequate for explaining sign languages. ArSL does not depend on the Arabic language structure and has a private lexicon and grammar rules [6]. Many rules must be utilized according to the ArSL grammar for generating ArSL sentences. This grammar will be explained in the following sections.

#### 2.1.1. Alphabet and numbers

ArSL alphabets are similar to the forms of Arabic letters, and this similarity can be found in the American and British sign language. One hand is used to represent the letters of the alphabet. ArSL alphabet introduces new terms and concepts and spells proper nouns, book names, street names, and other names. The ArSL has 35 letters, and each has five subtypes of symbol, which is used to represent movement. Figure 1 shows the ArSL alphabet representation.

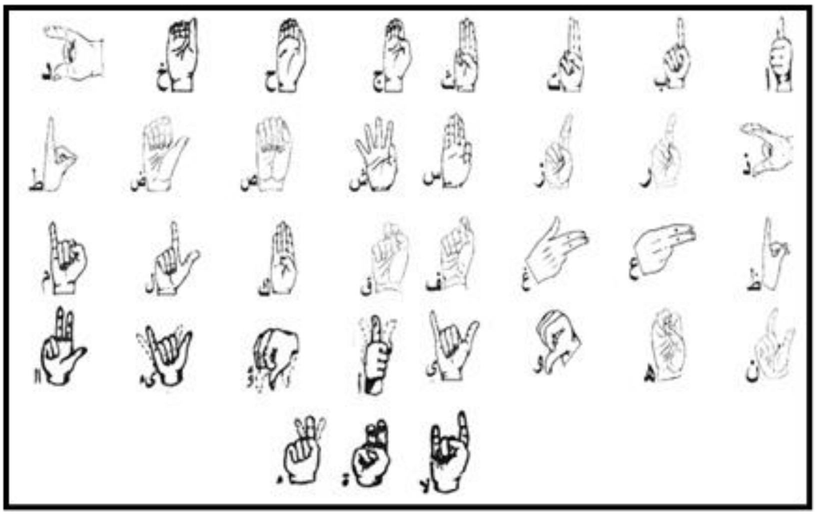


Figure 1. ArSL alphabet

Where numbers are handshapes symbols developed to help the deaf to recognize mathematical numbers depending on the handshape, orientation, movements, and mouth movement. In ArSL, numbers are not handled similarly to average numbers in the Arabic Language. Figure 2 shows examples of number signs in ArSL.

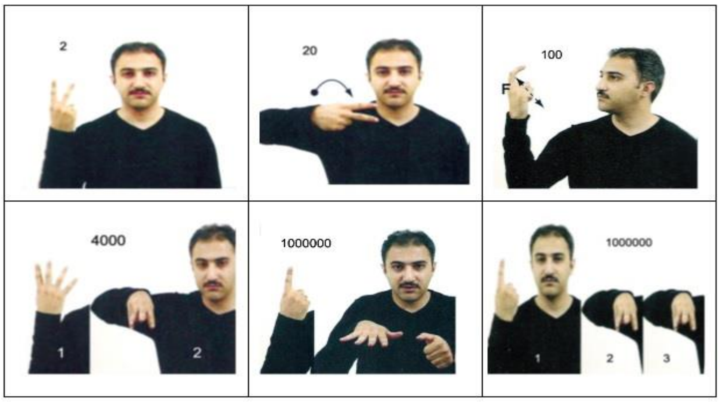



Figure 1. Examples of numbers in ArSL

#### 2.1.2. Nouns and pronouns

Deaf people in Arab countries are used fingerspelling for the proper nouns such as أحمد (Ahmed), علي (Ali). At the same time, the names of fruits, animals, and vegetables usually have a sign, and the sign of countries and regions are taken from deaf people in that regions. For pronouns, deaf people generally use the direction of their hand to obtain personal pronoun signs like أنا (me), أنت (you), هو and (he). If a pronoun is feminine such as هي (she), then the sign بنت (girl) will be added. Relative and conditional pronouns such as الذي (which) for male التي (which) for female have no sign, and they can be shown from the expression signs. The attached pronouns to verbs do not appear in the Arabic sign language and replace them with an addition sign, for example, the word درستُ (I studied) is translated into (أنا+درس) (Study+I). Deaf persons use possessives during their sign conversation to show possession by indicating whom they are speaking about.

#### 2.1.3. Plural and gender in ArSL

ArSL, like American sign language, does not change the structure of nouns to express plurality. There are different methods to display the plurality of people or objects in ArSL. To express dual nouns in ArSL, the sign of اثنان (two) is used after the noun. For example, the word رجلان (two men) is translated into (رجل+اثنان) (man+two). To express plurality in ArSL, the sign is repeated, or the sign of كثير (much) is used after the noun. For example, the word كتب (books) is translated into (كتاب+كثير) (book+much). Unlike the Arabic language, for indicating gender, ArSL used a separate sign (girl) after the noun to represent gender. For example, the noun طالبة (student) is translated into (طالب+ بنت) (student+girl).

#### 2.1.4. Time context

ArSL structure allows showing a different time context (present, past, and future). Sometimes autonomous lexical items can be used. For example, غداً (tomorrow) and أمس (yesterday) appear at the beginning of the ArSL sentence. ArSL does not use verb tense to indicate the sentence time but uses time signs at the beginning to indicate the tense. A time sign, such as قبل (before), is utilized for indicating past verbs, whereas signs such as الان (now) are used for indicating present verbs. For indicating future verbs, the sign قريبا (soon) is used.

#### 2.1.5. Questions

The question signs like ( ماذا ), what ( اين ), where ( متى ), when ( كيف ), how ( سبب ) or always come at the end and a question mark sign comes at the beginning of the sentence in ArSL. In addition to that, some question words in the Arabic language are substituted by others in ArSL for example, the word أي is substituted by هذه ام هذه this or this.

**2.1.6. Conjunction and prepositions**

Conjunctions such as و (and) and أو (or) combine two sentences or words in the Arabic language, while conjunction in ArSL does not exist, sign a sentence, take a slight stop, and then express the following sentence. Prepositions like الى (to), من (from), في (in), and others are not used in ArSL. An excellent opinion is to avoid prepositions during signing in ArSL since they are shown in the context.

**2.1.7. Sentence structure**

Arabic sentences can be classified as nominal and verbal sentences. The central word order for a verbal sentence in the Arabic language can be verb subject object (VSO), verb object (VO), verb subject (VS), verb object subject (VOS). In contrast, the word order for a nominal can be subject verb object (SVO) or subject object verb (SOV), While the structure sentence in ArSL is different. Most of the opinions addressed the sentence structure of the ArSL SVO or sometimes SOV.

**2.2. Hamburg notation system (HamNoSys)**

HamNoSys is a textual system for sign language representation developed in 1985 by the University of Hamburg, Germany. It is created depending on the Stokoe notation system [18]. The system can be used for any sign language and is developed for a linguistic description of signs, such as machine translation, sign generation, and construction of SL dictionaries. HamNoSys is also a basis for several avatar controls. It involves more than 210 symbols and is a language-independent system (any sign language can be described) for that, it can be used universally. HamNoSys notation system describes signs utilizing sequences of symbols. These symbols are represented in order (symmetry parameters, non-manuals expression, handshapes, palm orientation, location, and movement) [19]. Figure 3 shows the structure of HamNoSys for sign generation.

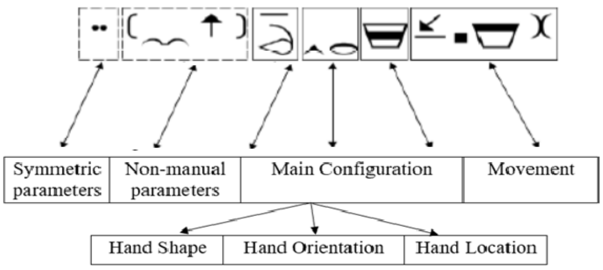



Figure 2. Structure of HamNoSys

**2.3. SiGML**

SiGML was constructed into a ViSiCAST project [20] as an essential component of a translation model for a sign animation system. The primary purpose of SiGML is to assist the interpretation of signing gestures in a method allowing them to be animated in live time utilizing an avatar. It is an extensible markup language (XML) scripting style constructed based on the structure of HamNoSys. It allows appending some features with a higher degree of exactness than HamNoSys, but building signs utilizing SiGML is time-consuming and challenging. However, the grammatical relationship between HamNoSys and SiGML is lossless SiGML can be transformed into HamNoSys, and HamNoSys can be converted back into SiGML. No information is lost in this process [21].

**2.4. ESign software**

ESign software has two applications: eSign Editor and JASigning Player. The eSign Editor application was produced in the eSign project to allow us to create signs in structures of HamNoSys and then convert them into SiGML files [22]. It has various advantages, such as several hand gestures, facial expressions, and lengthy sentences that can be contracted, such as several handshapes, hand movements, body movements, mouth gestures, and facial expressions; however, it needs experience in HamNoSys. Figure 4 shows the eSign Editor Interface while the JASigning player application [23] was created to test the signs produced by the eSign editor application.


Figure 3. Shows the eSign editor interface

## 3. THE PROPOSED SYSTEM ARCHITECTURE

The proposed system consists of two main components: constructing a bilingual dictionary using synthetic animation and the proposed translator of Arabic text to equivalent ArSL text, which is animated using a 3D avatar the system's architecture is illustrated in Figure 5.

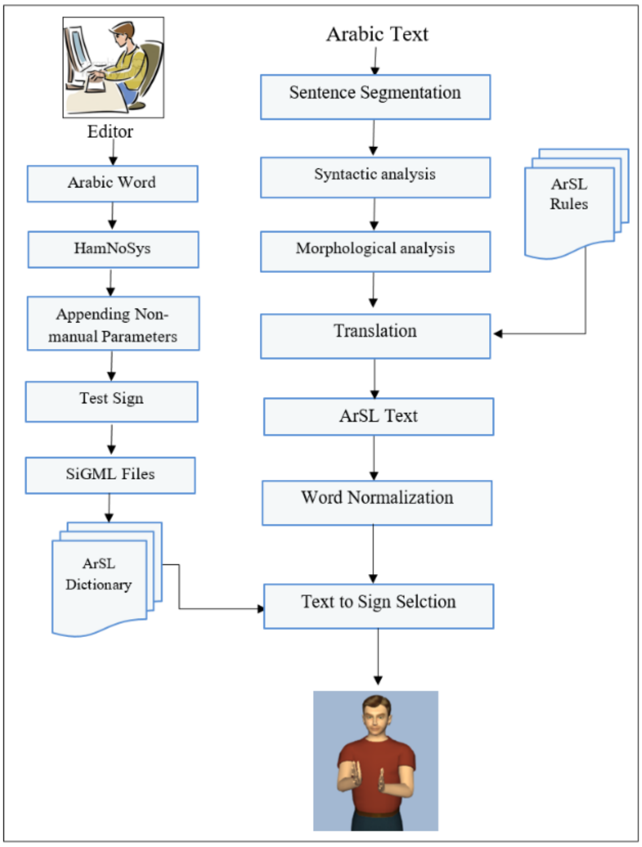



Figure 4. System architecture

### 3.1. Dictionary construction

ArSL suffers from a lack of resources; therefore, a dictionary of the Arabic language for ArSL has been constructed as a first step. Arabic words are transformed into HamNoSys using the eSign editor application, which uses HamNoSys as a notation system. HamNoSys was utilized to generate MFs, while NMFs have been appended utilizing face, limbs, and mouth in the eSign editor application.

After the description of MFs and NMFs, the signs will be experimented with to check if they fit utilizing the JASigning player application that was created to test the sign produced by the eSign editor. Once the sign is built and experimented with correctly, it will be saved as a SiGML file. Each generated SiGML file specifies an index number representing the name of this file. This index will be added to the dictionary beside the Arabic word lemma because well-known dictionaries such as the collins and oxford English dictionaries have generally used lemmas for listing words. For example: walk, walks, and walking are listed in the oxford dictionary under lemmas walk because it is a canonical form, while the other forms can be derived. For example, مدارس (schools), مدرستكم (your school), and المدرسة (the school) are all these forms considered to be variant forms of the lemma مدرسة (school). However, the sign language dictionaries need not be arranged any differently. The two lemmas will be taken in the case of composite words that combine two or more words and have to perform by one sign. For example, كانون الثاني (January) is made up of one sign so arranged as كانون+ثاني (January). Later, 3D avatar interprets the SiGML file scripts to the animated sign. The signs were designed based on the unified Arab dictionary produced by LAS and ALECSO. Table 1 shows the steps for generating HamNoSys for the word اثنان (Two), where Figure 6 shows the test of sign used JASigning player, and Figure 7 shows a SiGML file script for one word [24].

Table 1. Generating HamNoSys for word two إثنان

| Step | Hand Movement | Discerption |
|---|---|---|
| 1 | | Handshape selected |
| 2 | | Thumb finger added to the hand |
| 3 | ^ | Hand direction up |
| 4 | | Palm orientation to the back |
| 5 | | Hand location at the Shoulder line |
| 6 | )( | Hand contact with the body |
| 7 | ↑ | Straight hand movement to up |
| 8 | | Final HamNoSys |

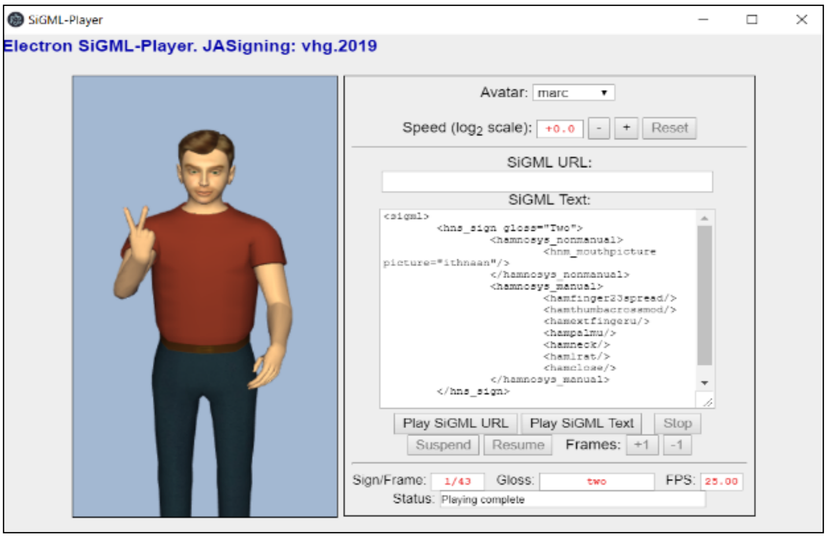



Figure 5. Test the sign two

### 3.2. The proposed translator

A rule-based translator has been developed. The system takes an Arabic text as input and returns its equivalent ArSL text, which displays as a 3D human avatar, as shown in Figure 5. The proposed translator structure consists of several stages: sentence segmentation, syntactic analysis, morphological analysis, translation, word normalization, text-to-sign selection, and sign animation. The rules used to generate the ArSL sentence are made on the variations between Arabic and ArSL, as illustrated in the previous section.

**3.2.1 Sentence segmentation**

The first stage of the translation starts by segmenting the Arabic text into sentences. The segmentation is mainly adopted on the punctuation mark for the Arabic language relatively straightforward task as in English [25]. Used punctuation marks for sentence segmentation are shown in Figure 8.

```
</sigml>

        <hns_sign gloss="اثنان">
                <hamnosys_nonmanual>
                        <hnm_mouthpicture picture="ithnaan"/>
                </hamnosys_nonmanual>
                <hamnosys_manual>
                        <hamfinger23spread/>
                        <hamthumbacrossmod/>
                        <hamextfingeru/>
                        <hampalmu/>
                        <hamneck/>
                        <hamlrat/>
                        <hamclose/>
                </hamnosys_manual>
        </hns_sign>

  </sigml>
```

Figure 6. The SiGML file script

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ; | ‘ | : | . | ، | ... | , | ? | “” | - | [] | = |

Figure 7. The used punctuation marks for text segmentation

**3.2.2. Syntactic analysis**

The Arabic sentences produced in the previous step are analyzed by Stanford parser [26] to obtain the sentence's grammatical structure. The parser has been developed to operate with many languages, such as English, Chinese, German parser, and Arabic. Arabic parser developed based on the Penn Arabic treebank. The structure of the sentence determines phrases (collections of words that belong together). The structure of the sentence is described as a parse tree. Figure 9 shows the parse tree for the sentence "ذهب الطالب إلى المدرسة" (the student went to the school).

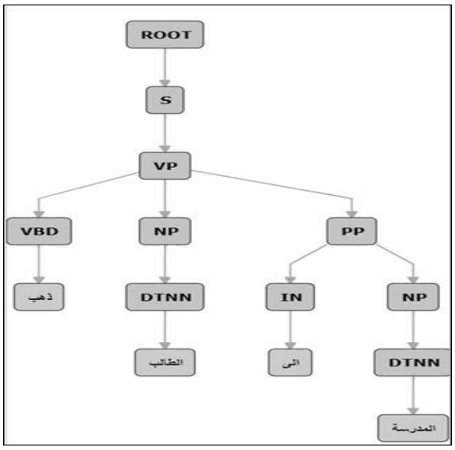



Figure 8. Parse tree generated by Stanford parser

### 3.2.3. Morphological analysis

Morphological analysis is significant for complex languages such as Arabic because it reduces different structures of a word to a single structure and gives various analyses of the word. Alkhalil2 Morpho system [3] was adopted to analyze the Arabic sentence and extract all features for each word. Alkhalil2 accepts the vowelized sentence as input, and the system's output is presented in the XML arrangement. Figure 10 displays an instance of analyzing the sentence "يَلْعَبُ الأولادُ بالكُرَةِ."(Children play with ball). Most Arab people are written a text without the usage of diacritics.

```
<?xml version="1.0" encoding="UTF-8"?>
- <morphology_analysis total_words="3">
  - <word total_analysis="2" value="يَلْعَبُ" w_id="1">
      <analysis additional_info="لَعِبَ" suffix="#" prefix="[ي: حرف المضارعة]" type="فعل مضارع مبني للمعلوم" impartial="مجرد"
        transitive="لازم" pos="ثلاثي مجرد مرفوع مسند إلى الغائب_هو، لازم" root="لعب" pattern="يَفْعَلُ" lemma="لَعِبَ" stem="يَلْعَب"
        vowled="يَلْعَبُ" a_id="1"/>
    </word>
  - <word total_analysis="1" value="الأَوْلادُ" w_id="2">
      <analysis additional_info="وَلَد" suffix="#" prefix="[ال: التعريف]" type="اسم جامد" pos="جمع مذكر مرفوع في حالة التعريف"
        root="ولد" pattern="أَفْعَال" lemma="وَلَد" stem="أَوْلاد" vowled="الأَوْلادُ" a_id="1" caze="في حالة التعريف "
        mood="مرفوع" gender="مذكر" number="جمع"/>
    </word>
  - <word total_analysis="1" value="بِالكُرَةِ" w_id="3">
      <analysis additional_info="كُرَة" suffix="#" prefix="[بِال: حرف الجر][ال:التعريف]" type="اسم جامد" pos="مفرد مؤنث مجرور في
        حالة التعريف" root="كرو" pattern="فُعَة" lemma="كُرَة" stem="كُرَة" vowled="بِالكُرَةِ" a_id="1" caze="في حالة التعريف "
        mood="مجرور" gender="مؤنث" number="مفرد"/>
    </word>
</morphology_analysis>
```


Figure 9. Example of an output of the Alkhalil2 Morpho system

Since they can understand diacritics from context, the Alkhalil2 Morpho system can take input without diacritics. However, it produces several analyzed outputs of each word by creating various suggestions about the missing diacritics. The system requires choosing one of these analyzed outputs. For example, the word "كتب" means books (noun) or write (verb), therefor Stanford POS software [25] produced by the Stanford NLP group at Stanford University was used. It supports several languages, one of them the Arabic language. It assigns parts of speech to each word in a sentence, such as a noun, verb, and adjective. Tagged the previous sentence by stanford POS the result: (يَلْعَبُ) VBP, (الأولادُ) DTNN), (ب) IN, (الكُرَةِ) DTNN. Based on the Stanford POS tagged, one of the analyses produced by the Alkhalil2 system is selected. At the end of this step, each word of the given sentence consists of a set of related analyses, such as (suffixes, prefixes, type, gender, number, and lemma).

### 3.2.4. Translation

This Module receives the output of the previous stage, a parse tree, and an Arabic word with its features (lemma, suffix, prefix, gender, number, type) as input and returns the ArSL sentence as an output for this module. The proposed system has adopted a set of (IF-Then) transformational rules for converting the Arabic text into the equivalent ArSL. Those rules can be explained as follows:

i) Syntactic reordering: the sentence structure of the Arabic language should be rearranged to agree with the structure of the ArSL sentence. This can be achieved by using the parsing tree produced by the Stanford parser. If the Arabic sentence were VSO, SOV, and VOS, the structure would be reordered into SVO.
ii) Special characters removal: these characters can be summarized as ( &, *, ،, -, ؛,\,:, ", #, **(, )**). Deaf people do not use these characters; therefore, they should be eliminated. This can be achieved using a type list produced by the morphological analyzer. If the system finds one of these characters, the character with its feature will be eliminated from the list.
iii) Pronouns transformation: according to the ArSL grammar, these attached pronouns should be transformed into their original form before verb-type words and after noun-type words. This can be achieved by using suffix and prefix lists produced by the morphological analyzer. The original form of an attached pronoun will be inserted before a verb and after a noun in the lemma list. For example, كتبتُ (I wrote) will be transformed into أنا كتب (I wrote), where كتابي (my book) will be transformed into كتاب أنا (the book I).
iv) Word gender transformation: the word gender (masculine and feminine) can be found using the gender list produced by the morphological analyzer. If the gender of a word is feminine, the word بنت (girl) will

be inserted after the word in the lemma list. For example, مديرة (manageress) will transform into مدير بنت (manager girl).

v) Word number transformation: word numbers (single, dual, and plural) can be found using a number list produced by the morphological analyzer. The word اثنان (two) and كثير (much) will be inserted after the word in the lemma list according to word number (dual, plural), respectively. For example, رجلان (two men) will be transformed into رجل اثنان (two men), where رجال (men) will be transformed into رجل كثير (a man much).

vi) Sentence time detection: two ways can detect the time of the sentence:
- From the words that refer to time, such as أمس (yesterday), غدا (yesterday)
- From the verb tense (past, present, future, command).

The system tries to find the word that refers to time first. These words can be found using a word list produced by the morphological analyzer. If the system finds it, the word will be shifted to the beginning of the lemma list. If the sentence does not have a word that refers to time, then the system will search the verb tense to detect the sentence time.

The verb tense can be found using the pos list produced by the morphological analyzer. These pos can be one of (past, present, or command). The morphological analyzer cannot detect the future tense; therefore, the affix list can be used for detection of the item that refers to the future tense as "س" for example, سأكتب (I will write). The words (قبل, الان, لازم, قريبا) (before, now, must, soon) will be inserted at the beginning of the lemma list according to the tense of the verb (past, present, command, future) respectively.

vii) Negation words transformation: These words refer to negation in an Arabic sentence. These words can be found using a word list produced by the morphological analyzer. If found, they will be replaced by the word لا (no).

viii) Questions words detection: question sentence can be detected from the question mark "؟". If the system finds the question mark, it will search for the question word using a word list produced by the morphological analyzer. If found, they will be shifted to the end of the lemma list.

ix) Words removal: deaf people do not use some words, such as prepositions and conjunction; therefore, they should be eliminated. This can be achieved using a type list produced by the morphological analyzer. If the system finds one of these words, the word with its feature will be eliminated from the lists.

As shown in the previous section, the translation module translates the Arabic sentence to agree with the arrangement of the ArSL sentence. This is done by removing, adding, shifting, and transforming the words. These rules have been applied according to ArSL grammar. The final output from this step is the ArSL sentence, where each word in the sentence is a lemma. The system presents the ArSL sentence and animates it using an avatar after searching for each ArSL word in a suitable dictionary that will be explained in the following sections.

#### 3.2.5. Normalization

Normalization is the method of removing some components from the text. It will be transformed into a single standard form [27]. In the proposed system, the words in the ArSL sentence will normalize. The normalization process will increase the performance of the searching and matching process. The normalizing process can be:

- Dialectal symbols are removed. For example (كَتَبَ, صُنْدُوقُ, بُسْتَانٌ) are normalized into (كتب, صندوق, بستان)
- Letters normalization (variant forms to one form conversion). For example, (أسد, إسلام, آمال) are normalized into (اسد, اسلام, امال).

#### 3.2.6. Text-to-sign selection

This stage of the proposed system accepts the output of the previous stages, the ArSL text as input, and gives the list of indexes of SiGML files. The search mechanism is based on the lemma in the ArSL dictionary, where each lemma is associated with its corresponding index (SiGML file name) in the dictionary. Since the dictionary used is small and does not contain all the words in the Arabic language, many will be recorded as unknown words. These unknown words are processed by the fingerspelling method; also, the numbers are handled in a particular way, as explained in the following sections.

#### 3.2.7. Fingerspelling translator

Fingerspelling means spelling words by using gestures that indicate the word's letters. A letter sign dictionary was constructed for that purpose. It is a small dictionary that includes 35 letters of the Arabic alphabet with their corresponding sign. This procedure can be achieved by searching the letter sign dictionary to determine the corresponding SiGML file for every letter of a given word and then combining these files in one file, which describes the translation of one word from the input sentence. Many items require a fingerspelling technique, such as proper nouns, places, and book names.

#### 3.2.8. Number to ArSL Translator

The numbers can be classified into two groups fundamental and derived. Raw numbers such as the numbers between 0 and 9 and the numbers 10, 20, 100, and 1000. These numbers have single sign representation, in contrast, the derived numbers are obtained from raw numbers described by joining the basic numbers' sign. For example, 165 is derived from 100, 60, and 5.

The Translator takes a number and then searches for the number - a sign dictionary was constructed for that purpose. This dictionary includes the primary number with its corresponding sign. If the number is found in the dictionary, then it gets the index of the SiGML file, which delivers to 3D avatar animation. Otherwise, the number is a derived number that extends to basic numbers, then searching a dictionary to get the indexes of SiGML files for the derived number and combines these files in one file, which delivers to 3D avatar animation.

## 4. RESULTS AND DISCUSSION

The system is developed using java programming language using a NetBeans development environment on the Windows 10 operating system. A coherent and simple system graphical user interface (GUI) has been designed to be a practical tool for people in the deaf community and their families who have close communication with deaf persons. The translator offers various 3D avatars provided by the JASigning system. The users can select one of these avatars for animating ArSL text. Plus, the JASigning system has many controlling options, such as controlling the avatar's speed, cycling (checking if the person needs to see the animation several times), and play stop options (beginning or stopping playing the avatar). These options were added to the system GUI to provide efficient system use. Figure 11 shows the interface of the system.

The constructed dictionary contains about 3000 signs, which means it is the largest known ArSL dictionary. Its construction has consumed time and effort. The individual sign is coded during a broad range of SL phonological parameters, including manual and non-manual parameters. Our dictionary was tested with multiple sentences of ArSL. These sentences also experiment with a dictionary of videos. The result shows that animation processing is faster than video processing results from the loading of large files of videos for each sentence to the cache memory. Two experts in sign language have been asked to evaluate the quality of generated signs in the range from1 to 5 for 100 signs selected randomly. The average rate was 4.3. The results are encouraging and can be utilized for communication with deaf people.

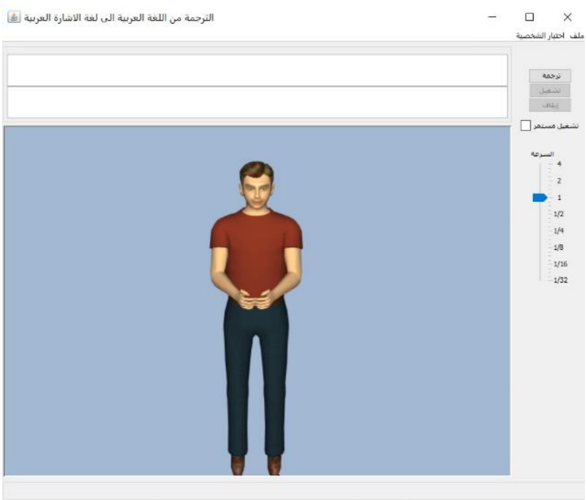
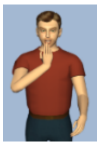
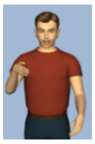
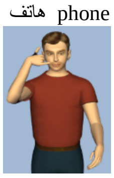



Figure 10. The system GUI

A parallel corpus containing about 600 sentences in the health domain [15] was used for system development and evaluation. The sign language part of the corpus was translated by two expert translators and two deaf people. The sentences in the corpus comprise approximately all types of Arabic sentences (nominal, verbal, and questions). The average sentence length is 5.5 words, where 70% of the corpus was used in developing the ArSL translator, and 30% was used to evaluate the system. Some of the words in the sign language part of the corpus are not in the lemma form, whereas the words in the generating ArSL sentences are in their lemma form because the constructed dictionary uses lemma. So these words transformed into lemma form. Table 2 shows the experiment results of translating four different sentences with their animation using a 3D avatar.

Table 2. Examples of translated sentences

| Arabic Sentence | ArSL Sentence with their animation | | | | |
|---|---|---|---|---|---|
| | مدرسة school | اتى came | انا I | brotherأخ | قريبا soon |
| سيأتي أخي من المدرسة<br>My brother will come from school |  |  |  |  |  |
| | ما what | انت you | | هاتف phone | رقم number |
| ما رقم هاتفك؟<br>what is your phone number? |  |  | |  |  |
| | ذكي smart | كثير much | بنت girl | طالب | student |
| الطالبات ذكيات<br>The students are smart |  |  |  |  |  |
| | 5 | 70 | 100 | انا I | طول height |
| طولي 175<br>My height is 175 |  |  |  |  |  |

Several metrics were invented to evaluate the translation systems in the latest decade, such as the metric for evaluation of translation with explicit ordering (METEOR) [28], bilingual evaluation understudy (BLEU) [29], and translation error rate (TER) [29]. The one that produces a good correlation with human judgment is the METEOR score.

The METEOR metric is dependent on the harmonic mean. It possesses many characteristics not located in other metrics, such as synonymy, stemming, and exact word matching. To estimate a METEOR score, firstly, we need to calculate unigram precision (P) (the number of unigrams in the candidate translation is also found in the reference translation to the whole number of unigrams in the candidate translation). Likewise, unigram recall (R) is calculated as (the number of unigrams in the candidate translation. They are also found to the whole number of unigrams in the reference translation, then Fmean is calculated as (1).

$$\text{Fmean} = \frac{10PR}{R+9P} \quad (1)$$

Fmean estimates the congruity to single words but not larger fragments appearing in the candidate and the reference sentence. For this, METEOR estimates a penalty (p) as (2).

$$\text{Penalty} = 0.5 * \left(\frac{\#\text{ chunks}}{\#\text{ unigrams\_matched}}\right)^3 \quad (2)$$

A chunk is a collection of adjacent unigrams in the candidate and the reference. The longer the adjacent, the fewer chunks there are. Lastly, the METEOR Score is calculated as (3).

$$\text{Score} = \text{Fmean} * (1 - \text{Penalty}) \quad (3)$$

Our Translator is evaluated using METEOR metrics using a parallel corpus comprising 180 sentences. The evaluation results are displayed in Table 3. The results are very encouraging and validate the rule-based systems better than the data-based approaches because of the inadequacy of data.

Table 3. METEOR metric scores for the proposed system

| Fold | Testing Sentences | Average Sentence Length | METEOR |
|---|---|---|---|
| 1 | 18 | 5.5 | 0.851 |
| 2 | 18 | 5.7 | 0.876 |
| 3 | 18 | 5.6 | 0.867 |
| 4 | 18 | 5.5 | 0.86 |
| 5 | 18 | 5.8 | 0.881 |
| 6 | 18 | 6.1 | 0.883 |
| 7 | 18 | 6 | 0.873 |
| 8 | 18 | 5.8 | 0.885 |
| 9 | 18 | 6.2 | 0.865 |
| 10 | 18 | 5.7 | 0.841 |
| Average | 18 | 5.8 | 0.868 |

## 5. CONCLUSION

In this paper, a translation system for Arabic text to ArSL has been developed. Based on our survey, there are very few resources that can present knowledge about sign language for helping the deaf communities in Arab countries. most of the proposed systems have displayed the ArSL sentence as an image which is now out of date in the era of computerization, plus most of the system performs a direct translation among Arabic and ArSL, which is lead to misunderstanding because of the structure languages are different. Therefore this work has concentrated on i) developing a new dictionary for ArSL using avatar technologies where signs are generated based on the unified Arab dictionary. Using 3D avatar technologies is much better than using human video because of the minor memory consumption, less conversion time, and supported translation system, an avatar can be varied according to choice. The dictionary can be adopted for the translation system in which written text can be transformed into sign language and can be utilized for the education of deaf people. ii) developed a machine translation system of the Arabic language to ArSL using modern technologies to help deaf and hard of hearing people communicate more efficiently with their families and hearing people. Moreover, it plays a role in assisting teachers in the sign language domain and people who do not know sign language to communicate with deaf persons.

## REFERENCES

[1] B. Kituku, L. Muchemi, and W. Nganga, "A Review on Machine Translation Approaches," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 1, no. 1, pp. 182–190, Jan. 2016, doi: 10.11591/ijeecs.v1.i1.pp182-190.

[2] N. Y. Habash, "Introduction to Arabic Natural Language Processing," *Introd. to Arab. Nat. Lang. Process.*, vol. 3, no. 1, pp. 1–185, 2010, doi: 10.1007/978-3-031-02139-8.

[3] M. Boudchiche, A. Mazroui, M. Ould Abdallahi Ould Bebah, A. Lakhouaja, and A. Boudlal, "AlKhalil Morpho Sys 2: A robust Arabic morpho-syntactic analyzer," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 2, pp. 141–146, 2017, doi: 10.1016/j.jksuci.2016.05.002.

[4] T. Dasgupta and A. Basu, "Prototype machine translation system from Text-To-Indian Sign Language," *Int. Conf. Intell. User Interfaces, Proc. IUI*, pp. 313–316, 2008, doi: 10.1145/1378773.1378818.

[5] L. Watson, S. Powers, and S. Gregory, *Deaf and Hearing Impaired Pupils in Mainstream Schools*. David Fulton Publishers, 2014.

[6] S. Semreen and M. Albinali, "The rules of Arab Qatari sign standardized language," *Supreme Counc. Fam. Aff.*, 2010.

[7] K. Al-Fityani, *Deaf people, modernity, and a contentious effort to unify Arab sign languages*, PhD diss. UC San Diego, 2010.

[8] L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, and M. Palmer, "A machine translation system from English to American sign language," *Lect. Notes Artif. Intell. (Subseries Lect. Notes Comput. Sci.*, vol. 1934, pp. 54–67, 2000, doi: 10.1007/3-540-39965-8_6.

[9] E. Safar and I. Marshall, "The Architecture of an English-Text-to-Sign-Languages Translation System," *Recent Adv. Nat. Lang. Process.*, pp. 223–228, 2001.

[10] R. San-Segundo, J. M. Montero, J. Macías-Guarasa, R. Córdoba, J. Ferreiros, and J. M. Pardo, "Generating gestures from speech," *8th Int. Conf. Spok. Lang. Process. ICSLP 2004*, pp. 1817–1820, 2004, doi: 10.21437/interspeech.2004-593.

[11] J. Wang, Y. Sun, and L. Wang, "Chinese sign language animation system on mobile devices," *Proc. - 2nd Int. Conf. Inf. Technol. Comput. Sci. ITCS 2010*, pp. 52–55, 2010, doi: 10.1109/ITCS.2010.19.

[12] M. A. Mohandes, "Automatic Translation of Arabic Text to Arabic Sign Language Mohamed," *AIML*, vol. 6, no. 4, pp. 15–19, 2006.

[13] A. Almohimeed, M. Wald, and R. I. Damper, "Arabic Text to Arabic Sign Language Translation System for the Deaf and Hearing-Impaired Community," in *Proceedings of the 2nd Workshop on Speech and Language Processing for Assistive Technologies*, 2011, pp. 101–109, [Online]. Available: https://aclanthology.org/W11-2311.pdf.

[14] A. E. E. El Alfi and S. M. EL Atawy, "Intelligent Arabic Sign Language to Arabic text Translation for Easy Deaf Communication," *Int. J. Comput. Appl.*, vol. 180, no. 41, pp. 19–26, May 2018, doi: 10.5120/ijca2018917081.

[15] H. Luqman and S. A. Mahmoud, "Automatic translation of Arabic text-to-Arabic sign language," *Univers. Access Inf. Soc.*, vol. 18, no. 4, pp. 939–951, 2019, doi: 10.1007/s10209-018-0622-8.

[16] A. Pasha *et al.*, "MADAMIRA: A fast, comprehensive tool for morphological analysis and disambiguation of Arabic," *Proc. 9th*

*Int. Conf. Lang. Resour. Eval. Lr. 2014*, pp. 1094–1101, 2014.
[17] A. Shahrour, S. Khalifa, D. Taji, and N. Habash, "CamelParser: A system for Arabic syntactic analysis and morphological disambiguation," *COLING 2016 - 26th Int. Conf. Comput. Linguist. Proc. COLING 2016 Syst. Demonstr.*, pp. 228–232, 2016.
[18] K. Kaur and P. Kumar, "HamNoSys to SiGML Conversion System for Sign Language Automation," *Procedia Comput. Sci.*, vol. 89, pp. 794–803, 2016, doi: 10.1016/j.procs.2016.06.063.
[19] L. Goyal and V. Goyal, "Automatic Translation of English Text to Indian Sign Language Synthetic Animations," *NLP Assoc. India*, no. December, pp. 144–153, 2016.
[20] J. A. Bangham *et al.*, "Virtual signing: Capture, animation, storage and transmission - An overview of the ViSiCAST project," *IEE Colloq.*, no. 25, pp. 23–29, 2000, doi: 10.1049/ic:20000136.
[21] Z. Kang, "Spoken language to sign language translation system based on HamNoSys," *ACM Int. Conf. Proceeding Ser.*, pp. 159–164, 2019, doi: 10.1145/3364908.3365300.
[22] I. Zwitserlood, M. Verlinden, J. Ros, S. Van Der Schoot, and T. Netherlands, "SYNTHETIC SIGNING FOR THE DEAF: eSIGN," in *Proceedings of the conference and workshop on assistive technologies for vision and hearing impairment, CVHI*, 2004, vol. 2004.
[23] S. Ebling and J. Glauert, "Exploiting the full potential of JASigning to build an avatar signing train announcements," 2013, doi: 10.5167/uzh-85716.
[24] A. H. Aliwy and A. A. Alethary, "Development of arabic sign language dictionary using 3D avatar technologies," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 1, pp. 609–616, 2021, doi: 10.11591/ijeecs.v21.i1.pp609-616.
[25] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," *Proc. 52nd Annu. Meet. Assoc. Comput. Linguist. Syst. Demonstr.*, pp. 55–60, 2014, doi: 10.3115/v1/p14-5010.
[26] S. Green and C. D. Manning, "Better arabic parsing: Baselines, evaluations, and analysis," *Coling 2010 - 23rd Int. Conf. Comput. Linguist. Proc. Conf.*, vol. 2, pp. 394–402, 2010.
[27] R. M. Sallam, H. M. Mousa, and M. Hussein, "Improving Arabic Text Categorization using Normalization and Stemming Techniques," *Int. J. Comput. Appl.*, vol. 135, no. 2, pp. 38–43, 2016, doi: 10.5120/ijca2016908328.
[28] A. Lavie and A. Agarwal, "METEOR: An automatic metric for MT evaluation with high levels of correlation with human judgments," *Proc. Second Work. Stat. Mach. Transl.*, vol. 0, no. June, pp. 228–231, 2007.
[29] K. Papineni, S. Roukos, T. Ward, and W. Zhu, "BLEU : a Method for Automatic Evaluation of Machine Translation," *Comput. Linguist.*, vol. 371, no. 23, pp. 311–318, 2002, doi: 10.3115/1073083.1073135.

## BIOGRAPHIES OF AUTHORS

**Ahmed A. Alethary** earned his BSc in the field of computer science from university of qadisiyah, diwaniya, Iraq, 2010; He completed his M.sc in university of kufa, najaf, Iraq, 2020. His research interests include Arabic NLP and what is located within this field of applications. He can be contacted at email: ahmeda.alethari@uokufa.edu.iq.

**Ahmed Hussein Aliwy** earned his PhD in the field of computer science from university of warsaw, Poland, 2013, He completed his M.sc in university of technology, Baghdad, Iraq, 2004 and BSc in Electrical Engineering, university of technology, Baghdad, Iraq, 1997. His research interests include NLP, text mining and what is located within this field of applications. He can be contacted at email: ahmedh.almajidy@uokufa.edu.iq.

**Nabeel Salih Ali** received his the BSc. in Computer Science from the University of Technology, Baghdad, Iraq, in September 2003, and the MSc. in Computer Science (Internetworking Technology) from the University Technical Malaysia Melaka (UTeM), Malaysia, in July 2015. Currently, He is a PhD student in Faculty of Electric and Computer Networks at University of Tabriz, Tabriz, Iran. He is working a manager of the scientific research division at ITRDC centre and Lecturer in ECE Dept, Faculty of Engineering, at the University of Kufa, An Najaf, Iraq. His research works in Computer Networks and Security. Besides, Healthcare Monitoring Systems, the Internet of Things (IoT), Machine Learning (ML), and Brain-Computer Interface (BCI). He can be contacted at email: Nabeel@uokufa.edu.iq.