Arabic to English Machine Translation of
Verb Phrases Using Rule-Based Approach

[1]Zainab Abd Algani and [2]Nazlia Omar
[1,2]School of Computer Science, Faculty of Information Technology,
University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[1]Department of Computer Science, Faculty of Science,
Omer Al Mukhtar University, Al Bayda, Libya

Abstract: Problem statement: Scientific translation represents an important stream in the current century due to explosion of the information revolution. The translation of scientific text is still limited in accuracy due to the fact that the scientific terms cannot be translated appropriately. Word order rules are very important for the generation of sentences in the target language whereas the word order in Arabic language is different from the order in English. Any Arabic Machine Translation (MT) system to English should be able to deal with word order. **Approach:** The aim of this study is to introduce-MT (Verbal Sentence rule based Machine Translation), an automatic system for Arabic verbal sentence of scientific text to English translation using transfer based approach. Verbal sentences constitute the majority of Arabic scientific documents. The system involves three phases: analysis, transfer and a generation phase. The transfer method is one of the rule based approach category and the most common technique used in machine translation system. **Results:** The system was trained on 45 verbal sentences from different Arabic scientific text and tested on 30 new verbal sentences from different domains. An experiment performed involves comparison with two other machine translation systems namely Syzran and Google. The accuracy of the result of the designed system is 93%. **Conclusion:** VS-MT has been successfully implemented and tested on many verbal sentences from different field of Arabic thesis. An experiment was performed which involves comparison with two other machine translation systems namely Syzran and Google. Our approach is efficient enough to translate Arabic verbal sentences of scientific text to English.

Key words: Machine translation, scientific text, Natural Language Processing (NLP), Verb-Subject-Object (VSO), Modern Standard Arabic (MSA), human translation, syntactic characteristics

INTRODUCTION

Machine Translation (MT) is an application of Natural Language Processing (NLP) and the area of information technology. It deals with the translation of human languages such as Arabic and English. The translation can be in one direction (uni-directional) as in translation from Arabic to English, in two directions (bi-directional) (Ghurab *et al*., 2010) as in the translation from Arabic into English and translation from English to Arabic, or in more than two directions (multi-directional) back and forth. The translation of natural languages by using machine has become a reality in the late of twentieth century (Hutchins and Somers, 1992). The aim of MT systems is to produce the best translation without human assistance. The

machine translation is intended due to the lack of accuracy in available MT systems such as Google, expensive human translation, the growing number of users on the Internet and for quick online communication. Arabic is one of the main languages considered in very early days of MT. The initial studies focused mostly on dictionaries and morphology. Few systems and research have dealt with Arabic language due to its syntactic characteristics which are different from Latin characteristics (Mokhtar *et al*., 2000). Scientific translation represents an important flow in the current century due to explosion of the information revolution that started from last decade. The amount of scientific information in foreign languages as well as Arabic language used all over the world as reported in FCCSET (1993) is that 50% of science and technology

Corresponding Author: Nazlia Omar, School of Computer Science, Faculty of Information Science and Technology, University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

literature. The Arabic sentence is generally classified as either nominal sentence or verbal sentence (Ryding, 2005). Verbal sentence is a sentence that starts with verb and the Verb-Subject-Object (VSO) is the default order in Arabic verbal sentences or more commonly used unlike English which allows SVO only. Word order rules are very important for the generation of sentences in the target language where the word order in Arabic language is different from the order in English. Any Arabic Machine Translation (MT) system to English should be able to deal with word order. The aim of this study is to describe VS-MT an automated system that translates Arabic verbal sentences of scientific text into English by applying transfer approach. Word order problem is also tackled in this system by rely on the grammar of both languages Arabic and English.

Related work: Nowadays the translation from and to Arabic has gained much interest from many machine translation researchers. Since the growing number of users on the Internet and the propagation of communication, this leads the researchers to focus more in Arabic works and try different approaches in order to improve the MT quality. Shaalan (2000) applied transfer based approach to develop MT system to translate Arabic interrogative sentence to English in agriculture domain. In the integrative sentences they use the imperative form of the verbal sentence. Salem *et al.* (2008) developed rule based approach using role and reference grammar based on Interlingua approach to translate from Arabic to English. They used the representation of the logical structure of an Arabic sentence in the proposed system. Their aim is to show how characteristics of Arabic language will affect the progress of MT tool. Recent study in Arabic by Shirko *et al.* (2010) translates Arabic noun phrases to English using transfer based approach. In addition, Shaalan *et al.* (2004) used transfer approach to translate English Noun Phrase (NP) into Arabic. As they mentioned, the NP translation is significant because NPs form the majority textual content of the scientific and technical documents. Lonsdale *et al.* (1994) applied Interlingua approach to build MT model to translate technical text from English to French. They mentioned that machine translation for scientific text is possible if we can come close to the problem in the correct way. Stalls and Knight (1998) translated named and technical terms from Arabic text to English based on statistical approach. They mentioned that the translation of named and technical terms is a problem when two languages that carry different alphabets are involved, such as Japanese/English and Arabic/English. Mokhtar *et al.* (2000) presented MT module based on transfer

approach using unification based grammar to translate English scientific text to Arabic. Up to our knowledge, the translation of Arabic verbal sentence of scientific text to English is exceptional as there is no other similar study so far.

Arabic Language characteristics: Arabic is a Semitic language, as most of researchers of MT found out about its derivational and inflectional rich morphology. The Arabic language written is in a horizontal way from right to left. It has 28 in which 25 are consonants and three are vowels. It has free word order. It is well–thought-out by much using the grammar. In addition, its characters are different from Latin characters which make it a difficult language to study. There are two forms of Arabic language i.e., Classical Arabic that used in Quran and Modern Standard Arabic (MSA) which is the general language of all spoken Arabic. It is also the form of Arabic used in TV, radio, newspaper. In this study we refer to the Modern Standard Arabic (MSA) type of Arabic.

Arabic part of speech: The sentence in Arabic language can be categorized in many parts. This category is called part of speech. A word can take different part of speech in different context (Albared *et al.*, 2010). As Attia (2008) pointed out, the conventional classification of Arabic parts of speech into nouns, verbs and particles is not enough for a full computational grammar. According to Salem *et al.* (2008) they classified the parts of speech into nouns, adjective, adverbs, verbs, demonstrative and others.

Arabic verbal sentence: Arabic verbal sentence is another type of Arabic sentence besides nominal sentence. It holds a verb and one or more participants where as the default word order in Arabic verbal sentence is: Verb (V), Subject(S) and Object (O), such as, " قرا علي الكتاب " " Ali read the book". Additional occurrence order is: Subject (S), Verb (V) and Object (O) such as," علي قرا الكتاب" "Ali read the book", but not a common order. The simple verbal sentences components are: verb, subject (الفاعل), direct object (مفعول به) and complement (مكمل الجمله) such as adjective and adverb. All these components, except the verb, can be missing. However, the verbal sentence can be constructed in a different way. It can have only a verb with subject; verb and subject and complement; or verb and complement and subject.

MATERIALS AND METHODS

System architecture: The system is based on transfer approach that translates Arabic verbal sentence of

scientific text to English. The steps below show the summary of the system:

- Input the Arabic verbal sentences
- 2. The preprocessing task starts and involves handling each sentence individually by performing loop. Tokenization process tokenizes each sentence into many tokens (words)
- The analysis phase (Morphological and Syntactic analysis) is where Arabic morphological analyzer devises information about inflected Arabic word as well as identifying some of its features and eliminate affixes (prefix and suffix) from it. The parser then starts to determine the structure of the sentence (relationships among parts of verbal sentence)
- After the analysis phase is completed, the transfer phase starts and involves lexical transfer where it searches for an equivalent English meaning of each word node in Arabic parse tree, by referring to the bilingual dictionary. Syntactic transfer involves transferring an Arabic parse tree to an English parse tree. It relies on the grammar of the two languages
- The generation phase is where the morphological generation constructs the inflected English word based on

English grammar rules: The syntactic generation involves English parse word tree to traverse and generate the final structure of the English sentence that is equivalent to the translated verbal sentence.

The overall process of the transfer approach is illustrated in Fig. 1.

The system needs to go through the preprocessing task before going to the analysis phase which involves the tokenization process. The purpose of the tokenization is to split the running sentence into tokens. The token is the smallest syntactic unit. It can be a word or a part of a word (Attia, 2008). The sentence " يؤثر نقص الاكسجين علي القلب "will be tokenized as: <يؤثر>, <نقص>, <الاكسجين>, <علي>, <القلب>. Once tokenized the sentence is ready for analysis stage. In analysis stage, the information about source language only is needed. Therefore, monolingual dictionary for Arabic has been used in this phase. The monolingual dictionary has a huge number of stems with its features which describe the stems and their Part Of Speech (POS). Morphological analysis extracts the stem of an inflected Arabic word and it could as well find the syntactic category of the stem (Mohammad, 2000). An example for morphological analysis is as follows:

ندرس س + ن + درس

الاعشاب ال + اعشاب

Syntactic analysis or parsing: the sentence is divided or analyzed into basic parts to find out its grammatical structure. The morphological analyzer returns to the parser the words in it singular form with a number of features such as the verb tense and number of a noun (Shaalan, 2000). The result of the parser could be represented in a tree of phrases calls parse tree where each phrase stands for the verb, subject and object. An example of Arabic verbal sentence parse tree is shown in Fig. 2. Considering the verbal sentence "يوجد عديد من المشاكل في سلوك الاطفال"the subject in this sentence is Adjective Phrase (AP) 'عديد من المشاكل'.

Essentially there are two transfers: Lexical transfer and Syntactic transfer, where lexical transfer exchange Arabic words to English words and syntactic transfer exchange parse tree of Arabic verbal sentence to equivalent English In addition, the bilingual dictionary is essentially in transfer method The proposed bilingual dictionary has a huge number of stems words for Arabic and English languages with all their features and part of speech. Figure 3 shows an example of lexical transfer.

Next are list of the transfer rules of Arabic verbal sentence into English structure representation, that the VS-MT system handled with some examples to be more obvious.

Transfer rule 1: This rule states that for any verbal sentence, it has the pattern: [V N1 N2 ADJ] and its verb carries future meaning. It should be transferred to the pattern: [PRON AUX V ADJ N2 N1] when translated to English.

For example, the verbal sentence "سنصمم شبكه ري جديده" which should be translated to "we will design a new irritation system". The transfer process of the parse tree of this sentence to the target language sentence is illustrated in Fig. 4.

Transfer rule 2: This rule states that for each verbal sentence, it has the pattern:[V N1 N2 N3] and its verb carries past perfect meaning. It should be transferred into the pattern: [N1 PREP N2 N3 AUX V]. This procedure is illustrated in the verbal sentence " افلح برناج انخفاض الديون " as an example and it should be translated to "program of decrease debt has succeeded". The transferring process to this sentence is shown in Fig. 5.

Transfer rule 3: This rule states that for each verbal sentence, it has the pattern: [V Broken plural N1 Broken plural ADJ] and its verb carry imperfect category and start with 'ن'.
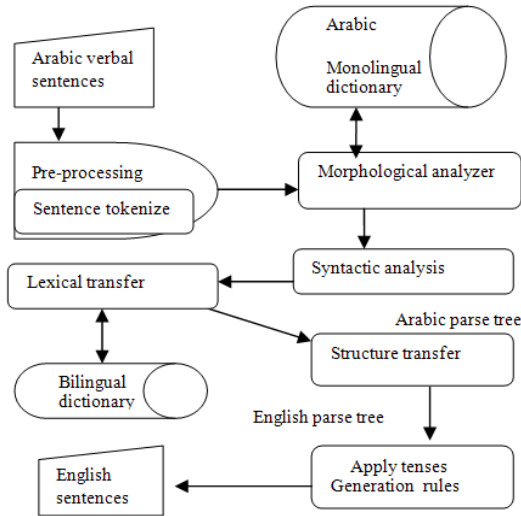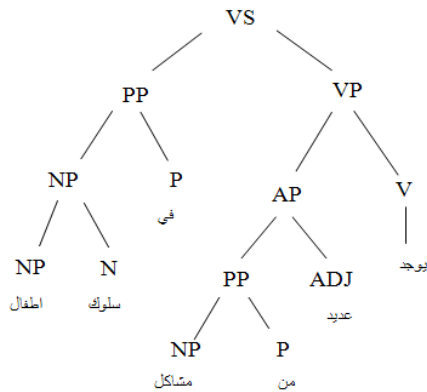
Fig. 1: The process of the system

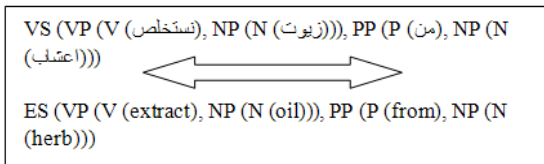Fig. 2: An example of (SL) parse tree

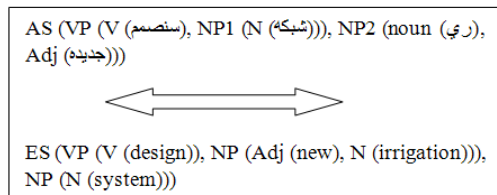Fig. 3: Lixical transfer

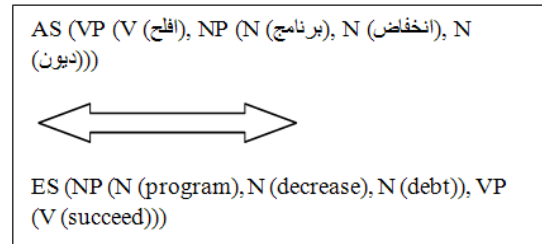Fig. 4: An example of transfer rule 1
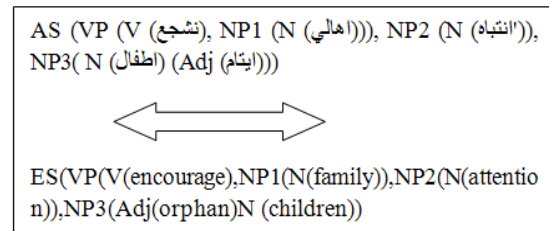

Fig. 5: An example of Transfer rule 2

Fig. 6: An example of Transfer rule 3

It should be transferred into the pattern: [PRON V Broken plural N1 ADJ Broken plural] as shown in Fig. 6 considering the sentence "نشجع الاهالي للانتباه للاطفال الايتام". Broken plural refers to plural noun.

Transfer rule 4: This rule states that for each verbal sentence, it has the pattern: [V N1 N2 ADJ] and its verb carry passive meaning. It should be transferred into the pattern: [N1 AUX V ADJ N2].

Transfer rule 5: This rule states that for each verbal sentence, it has the pattern: [V N1 N2 N3 ADJ] and its verb start with future character like 'س'. It should be transferred into the pattern: [AUX V N1 N2 ADJ N3].

Transfer rule 6: This rule states that this rule states that for any verbal sentence, it has the pattern: [V N1 N2 N3 ADJ]. It should be transferred into the pattern: [N1 V N2 ADJ N3].

Transfer rule 7: This rule states that for any verbal sentence, it has the pattern: [AUX V N1 PREP Broken plural N2] and its verb ends with 'نا'. It should be transferred into the pattern:[V N1 PREP N2 Broken plural].

Transfer rule 8: This rule states that for any verbal sentence, it has the pattern:[V N1 N2 Broken plural N3] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V N1 N2 N3 Broken plural].

Transfer rule 9: This rule states that for any verbal sentence, it has the pattern: [V N1 N2 PREP N3]. It should be transferred into the pattern: [N1 N2 V PREP N3 Broken plural].

Transfer rule 10: This rule states that for any verbal sentence, it has the pattern: [V N1 N2 PREP Broken plural] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V N1 N2 PREP Broken plural].

Transfer rule 11: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 N Broken plural2] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V Broken plural1 Broken plural2 N].

Transfer rule 12: This rule states that for any verbal sentence, it has the pattern: [V Broken plural N] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V Broken plural N].

Transfer rule 13: This rule states that for any verbal sentence, it has the pattern: [V N1 N2 N3] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V N1 N3 N2]. Otherwise it should transferred to [N1 V N3 N2]

Transfer rule 14: This rule states that for any verbal sentence, it has the pattern:[V N ADJ Broken plural] . It should be transferred into the pattern:[N V ADJ Broken plural]

Transfer rule 15: This rule states that for any verbal sentence, it has the pattern: [V Broken plural N] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V Broken plural N].

Transfer rule 16: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 N Broken plural2 ADV] and its verb start with 'س'. It should be transferred into the pattern: [PRON AUX V Broken plural1 N Broken plural2 ADV].

Transfer rule 17: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 ADJ PREP Broken plural2 ADV]. It should be transferred into the pattern: [PRON AUX V Broken plural1 N Broken plural2 ADV].

Transfer rule 18: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 ADJ PREP Broken plural2]. It should be transferred into the pattern: [Broken plural1 V ADJ Broken plural 2].

Transfer rule 19: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 N1 N2 Broken plural 2] and its verb start with 'ن'. It should be transferred into the pattern: [V Broken plural1 N1 Broken plural 2 N2].

Transfer rule 20: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 PREP N Broken plural2]. It should be transferred into the pattern: [Broken plural V PREP Broken plural N].

Transfer rule 21: This rule states that for any verbal sentence, it has the pattern: [V N ADJ ADV] and its verb start with 'س'. It should be transferred into the pattern: [PRON AUX V ADJ N ADV].
Transfer Rule 22: This rule states that for any verbal sentence, it has the pattern: [V N1 N2 ADJ]. It should be transferred into the pattern: [N1 V ADJ N2].

Transfer rule 23: This rule states that for any verbal sentence, it has the pattern: [V N1 PREP N2]. It should be transferred into the pattern: [N1 V PERP N2].

Transfer rule 24: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 N1 N2 N3 ADJ]. It should be transferred into the pattern: [Broken plural N1 V N2 ADJ N3].

Transfer rule 25: This rule states that for any verbal sentence, it has the pattern: [V Broken plural1 N1 PREP Broken plural2 N2]. It should be transferred into the pattern: [Broken plural1 N1 V N2 Broken plural2].

Transfer rule 26: This rule states that for any verbal sentence, it has the pattern: [V NPREP Broken plural] and its verb end with 'نا'. It should be transferred into the pattern: [PRON AUX V N Broken plural].

Transfer rule 27: This rule states that for any verbal sentence, it has the pattern: [V ADJPREP Broken plural] and its verb end with 'نا'. It should be transferred into the pattern: [PRON AUX V ADJ Broken plural].
Transfer Rule 28: This rule states that for any verbal sentence, it has the pattern: [V ADJPREP1 Broken plural PREP2 N1 N2] and its verb start with 'ن'. It should be transferred into the pattern: [PRON V ADJ Broken plural PREP2 N2 N1].

After the transfer phase finish the system will be go to the generation stage: Generation stage is commonly separated into two parts, syntactic generation and morphological generation. Morphological generation that generates inflected English word in its correct form

based on a set of grammar rules for English language where the English word passed from the transfer stage in its singular from some features. However, in the analysis phase, the Arabic words with their features are stored in a separate array. As happened earlier in the transfer phase, the main features for every Arabic word from the bilingual dictionary as well as its equivalent English words are stored in another array. These steps are carried out to prepare and pass them (the words) to the morphological generation. Furthermore, the tense generation rules that have been applied and the other generation rules that have been added in the process are explained in the following.

Tense generation rules: Each verb in Arabic has certain feature that represents certain tense when translated to English. From the prefix and the suffix that attaches to the verb, additional information can be known about the verb. The following points describe the tenses rules that have been generated in VS-MT system.

If the category of the verb is an imperfect verb, this means it refers to the present tense when translated to English. The following rules have been added:
[V (imperfect category and singular subject) V +'s' where V refers to verb. Such as: "" يؤثر نقص الاكسجين علي الماء " "lack of the oxygen affects on water".

V [(imperfect category and plural subject)] no need to add 's'.
Such as: "تنتمي الاعشاب للعديد من الفصائل" herbs belong to many groups".

If the category of the verb is perfect and the verb ends with 'نا', it becomes past perfect verb when translated to English with third plural subject pronoun. the following rule will be added:

[V [perfect category and ends with' نا']
'we' + 'have' + V]

Such as: لاحظنا→ noticed→ we have noticed.

If the category of the verb is perfect and does not end with 'نا', then when it is translated to English, it will carry past perfect meaning. The following rules have been be added:

[V [perfect category and singular subject]↔' has '+V].
Such as: "اظهرت الدراسه نتائج فريده" "the study has showed unique results"

[V [perfect category and plural subject] ↔ 'have' + V]
Such as: "اظهرت التجارب العديد من الفوائد" "the experiments have showed many benefits".

If the category of the verb is an imperfect verb and starts with future markers like 'سن', it refers to the future

meaning when translated to English, with third plural subject pronoun. The following rule has been added:
[V [imperfect categoryand starts with future markers]
↔'we' + 'will' + V]

Such as: سنستعرض → review → we will review

If the verb starts with 'ن', it refers to present tense with third plural subject pronoun, the following rule has been added:

[V [starts with 'ن'] 'we' + V]
Such as: نناقش → discuss→ we discuss

If the category of the verb is imperfect verb and starts with present continuous markers like the letters "ت" then the following rules will be added:

[V [present continues markers and singular subject] ↔ 'is' + V +'ing']

Such as: "تتحدث الباحثه عن التراث العلمي" "the researcher is talking about the scientific heritage".

[V [present continues markers and plural subject] 'are' +V +'ing']

Such as: "تتركز بحوث العلماء علي الاحتباس الحراري" "researches of the scientists are concentrating on the global warming".

If the category of the verb is imperfect verb and carries passive meaning when translated to English then the following rule will be added:
[V [imperfect verb and passive feature] ↔ 'is' or 'are' + V]
Such as: "تتميز خطوط الري عن القديمه" "lines of the irrigation are distinguished about the old ".

If the category of the verb is perfect verb and carries passive meaning when translated to English then the following rule will be added:

[V [perfect verb and passive feature]
↔ 'was' or 'were' + V]

Such as: "أعد النظام بطريقه جيده" " the system was prepared by good manner".

Noun and preposition generation rules:

The postfixes noun ("المضاف اليه"): If there is a compound noun where the first without the article "نكره" is undefined and the second "معرف" is defined, this is known as the postfixes noun.
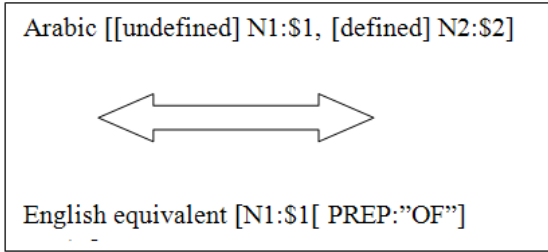
Arabic [[undefined] N1:$1, [defined] N2:$2]

⟷

English equivalent [N1:$1[PREP:"OF"]

Fig. 7: Generation rules for the postfixes noun

Arabic [(defined) N: $1, (defined) ADJ: $2]

⟷

English equivalent [(art:" the") ADJ: $1,N: $2]
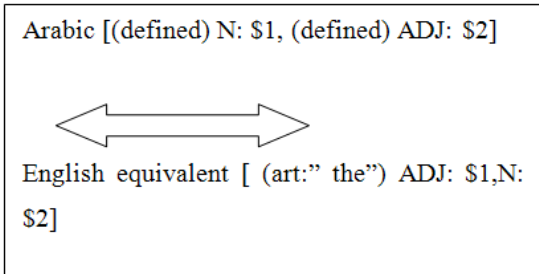
Fig. 8: Generation rules of defined adjective

Arabic [(undefined) N: $1, (undefined) N2: $2, (defined) N3:$3]

⟷

English equivalent [N1:$1, PREP:"of"N2:$2:N3:$3]
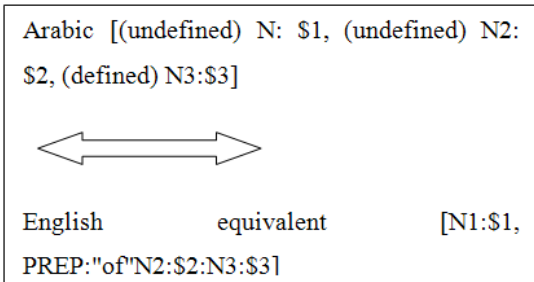
Fig. 9: Generation rules for undefined adjective

Arabic [(undefined) N: $1, (undefined) N2: $2, (defined) N3:$3]

⟷

English equivalent [N1:$1, PREP:"of"N2:$2:N3:$3]

Fig. 10: Generation rules of three consecutive

Arabic [NP1:$1 [start with PRER], NP2: $2 [start with the same PREP]]

⟷

English equivalent [NP1:$1,PREP, NP2:$2]

Fig. 11: Generation Rules Of Preposition
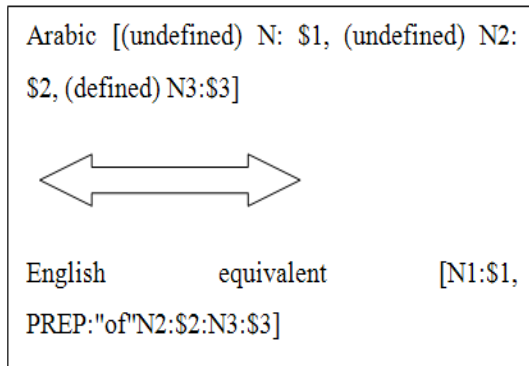
For example, the phrase "نقص الاكسجين "when translated, the proposition "of" will be added as shown in Fig. 7.

The Definition generation rules: If there are defined noun with defined adjective in the noun phrase such as" التعليم التقليدي" the article 'the' will be added before the adjective in the translation as shown in Fig. 8.

If there are undefined noun with undefined adjective in the noun phrase such as: مذيب كيميائي, the article "a" or "an "will be added before the adjective based on the beginning character of the adjective as shown in Fig. 9.

Generation rules of three consecutive nouns: If there are three consecutive nouns following each other, in which the first and the second are undefined while the third is defined such as "برنامج انخفاض الديون ", then the proposition 'of' will be added after the first noun in the translation as shown in Fig. 10.

Generation rules of preposition: If there are two following noun phrase that start with the same preposition the first one will be removed and the translation to it will be added before the second translated noun as shown in Fig. 11. Consider the sentence "نشجع الاهالي للانتباه للاطفال الايتام" "we encourage families the attention to the orphan children".

Generation rules of plural: The appropriate letters will be added based on the end of the English word as some letters have been removed from the end of English word before adding one of the plural letters to make the English word plural. This excludes irregular cases such as "children" will be stored in the database. Some example of plural English word:

Research   es   researches
Study   ies   studies
System   s   systems
Factory   ies   factories
Chapter   s   chapters

Syntactic generation the English parse tree traverses to produce the final structure which can then the English sentence be generated.

Input: نشجع الاهالي للانتباه للاطفال الايتام

Transfer: Encourage family attention orphan children

Output: we encourage families the attention to orphan children.

RESULTS AND DISCUSSION

Taking into account of the process of producing grammatically and syntactically acceptable words in the target language can lead to high quality MT system output (Zuntout and Guessoum, 2000). The idea of this experiment is to test the system on new verbal sentences to be convinced that the system output is acceptable. The evaluation methodology is applied on 30 verbal sentences from different Arabic scientific text and different domain. The methodology is based on making comparison between the outputs of the designed system for the test examples and the human translation for the input sentences. It also involves comparison with two other machine translation systems namely Systran and Google.

There are some mismatches test examples which arise some problems in the target sentence. The following specifies the problems that appeared in the target sentence.

Synonyms problem: This problem occurs because different synonyms of verbs and nouns are involved. For example, the noun "تأكد" could be affirmation, assertion or assurance and the verb "يحتوي" could be involve, include or contain.For example the word "fragmentation" in the translation of the sentence "نوضح كيفيه انقسام المجتمعات" "we explain how the society fragmentation" is not a common word.

Syntax problem: This problem occurs because some verb phrases in the same pattern come in different position when translated to English. For example, the verbs that carry future meaning may come before the subject. The sentence "يشكل قطاع التعليم اهميه كبيره " is translated to "A great importance of the education sector".

Tokenization ambiguity: This problem is due to the verbs that start with a certain subtokens such as "ن" are translated the same. The sentence " نجح برنامج انخفاض

"الديون is translated to "program of decreasing debt we has succeed".

Ambiguity of the preposition: This problem is due to some preposition can have two different meaning or can be omitted when translated to English such as the preposition"من"can be "of" or "from". The sentence "زرنا الكثير من المصانع" is translated to "We have visited a lot of factories"

The imperfect verb ambiguity: This problem is due to the translation of imperfect verbs in case of singular subject are treated the same. The sentence " يعتبر الماء مذيب كيميائي" is translated to "the water is considered a chemical solvent"

Ambiguity of the meaning: This problem is due to there are some verbs when translated to English become demonstrative pronoun with auxiliary verbs. For example, verb "يوجد" has two English equivalent meaning "there is", "there are". Such as: " يوجد علاقه ارتباط ايجابيه" translated to "there are a positive co relation".

The experiment performed involved 30 test examples and the score between 0- 0 was given based on the problems that appeared in the target sentence of each system. The score is given by human specialist in translation and it tests the differences among the human translation and the machine translation systems. An example of the experiment result shows in Table 1.

The entire problems that appeared in the target sentences of each system (Systran, Google and VS-MT) with their frequency occurrences are shown in Table 2. The table shows that the total of each problem for VS-MT system is less than the other two systems which means that VS-MT system gives better translation.

The percentage of the total score for each system is calculated by dividing the total score by 30 which is the number of the test examples. The system is evaluated out of 10. Table 3 illustrates the overall percentage of each system.

It seems that Systran performs worst in translation of Arabic verbal sentences. It has 57% accuracy only while, Google gave better translation when compared to Systran with 77% accuracy. On the other hand, VS-MT has higher accuracy of 93%. It is clear that, VS-MT system gain higher percentage than the other two systems, even with the six problems that occur in the target sentence of the test examples. Based on the linguistic rules that used in the designed system, the VS-MT system obtains good quality translation in comparison with the translation done by Systran and Google.

Table 1: An example of an experiment test

Verbal sentence (SL)	Translation	Arabic (MT)	Human translation problem	(No) Problem	Score
سنصمم نظام اتصال جديد	New regime of communication will design	Systran			
	We will design a new communication system	Google			
	We will design a new communication system	VS-MT	We will design a new communication system	2, 3, 4	7
نوضح كيفيه انقسام المجتمعا	Clarifies despotic his division gathered	Systran			
	Clarifis despotic his division gathered	Google			10
	We explain how the society fragmentation	VS-MT	We explain how the society split	1, 2, 3	10
				1,2	8
				1	9

Table 2: The Test Set Results with Total frequencies in each system

No.The problem	Total	Systran	Google	VS-MT
Synonyms problem	44	26	13	5
Syntax problem	39	27	10	2
Tokenization ambiguity	43	20	12	1
Preposition ambiguity	22	17	4	1
Imperfect verb ambiguity	4	0	0	3
Meaning ambiguity	2	0	0	2

Table 3: Experimental results

Machine translation	Systran	Google	VS-MT
Full score	171	231	279
Percentage	57%	77%	93%

CONCLUSION

In this study, an Arabic scientific text to English Machine Translation (MT) system has been discussed that is based on the transfer based approach. The system receives Arabic verbal sentences as an input. The necessary rules have been applied to identify their structural representation which can be helpful to identify the English structure representations. Then the target English sentences can be generated from such representations after applying the required grammar rules by considering the relational grammar of both Arabic and English. The rules are applied in each stage of the process from the input sentence until the output sentence is generated. There are several reasons that make transfer-based approach desired by MT community (Trujillo, 1999). These reasons are segments of transfer modules which can be useful when two languages that strongly related to each other are included, simple analysis and earlier development of grammar. The designed system can be used as a stand-alone tool and can be very well incorporated with a general MT for Arabic scientific text.

REFERENCES

Albared, M., N. Omar, M.J.A. Aziz and M.Z.A. Nazri, 2010. Automatic part of speech tagging for Arabic: An experiment using bigram hidden Markov model. Rough Set Knowl. Technol., 6401: 361-370. DOI: 10.1007/978-3-642-16248-0_52

Attia, M.A., 2008. Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation. Ph.D. Thesis, University of Manchester.

FCCSET, 1993. 1993. Machine Translation Technology: A Potential Key to the Information Age. 1st Edn., The Committee, Washington, D.C., pp: 54.

Ghurab, M., Y. Zhuang, J. Wu and M.Y. Abdullah, 2010. Arabic-Chinese and Chinese-Arabic Phrase-based statistical machine translation systems. Inform. Technol. J., 9: 666-672.

Hutchins, W.J. and H.L. Somers, 1992. An Introduction to Machine Translation. 1st Edn., Academic Press, London, ISBN: 012362830x, pp: 362.

Lonsdale, D.W., A.M. Franz and J.R.R. Leavitt, 1994. Large-scale machine translation: An interlingua approach. Carnegie Mellon University.

Mohammad, A.A.E.M., 2000. Machine Translation of noun phrases: From English to Arabic. M.Sc. Thesis, Cairo University.

Mokhtar, H.M.O., N.M. Darwish and A.A. Rafea, 2000. An Automated system for English-Arabic Translation of Scientific text (SEATS). Cairo University.

Ryding, K.C., 2005. A Reference Grammar of Modern Standard Arabic. 1st Edn., Cambridge University Press, Cambridge, ISBN: 052177151X, pp: 708.

Salem, Y., A. Hensman and B. Nolan, 2008. Implementing arabic-to-englishmachine translation using the role and reference grammar linguistic model. School of Informatics and Engineering, Dublin, Ireland.

Shaalan, K., 2000. Machine translation of Arabic interrogative sentence into English. The Pennsylvania State University.

Shaalan, K., A. Rafea, A.A. Mmonem and H. Baraka, 2004. Machine translation of english noun phrases into arabic. Int. J. Comput. Process. Oriental Language 17: 1-14.

Shirko, O., N. Omar, H. Arshad and M. Albared, 2010. Machine translation of noun phrases from Arabic to English using transfer-based approach. J. Comput. Sci., 6: 350-356. DOI: 10.3844/jcssp.2010.350.356

Stalls, B.G. and K. Knight, 1998. Translating names and technical terms in Arabic text. Proceedings of the Workshop on Computational Approaches to Semitic Languages, (CASL' 98), Association for Computational Linguistics Stroudsburg, PA, USA., pp: 34-41.

Trujillo, A., 1999. Translation Engines: Techniques for Machine Translation. 1st Edn., Springer, New York, ISBN: 1852330570, pp: 303.

Zuntout, R. and A. Guessoum, 2000. Arabic machine translation: A strategic choice for the Arab world. J. King Saud Univ. Comput. Inform. Sci., 12: 117-144.