

Evaluating Arabic to English Machine Translation

Laith S. Hadla
Department of Translation
Faculty of Arts
Zarqa University
P. O. Box 132222
13132 Zarqa - Jordan

Taghreed M. Hailat
Faculty of IT and CS
Yarmouk University,
Irbid 211-63, Jordan

Mohammed N. Al-Kabi
Faculty of Sciences & IT
Zarqa University
P. O. Box 132222
13132 Zarqa - Jordan

Abstract—Online text machine translation systems are widely used throughout the world freely. Most of these systems use statistical machine translation (SMT) that is based on a corpus full with translation examples to learn from them how to translate correctly. Online text machine translation systems differ widely in their effectiveness, and therefore we have to fairly evaluate their effectiveness. Generally the manual (human) evaluation of machine translation (MT) systems is better than the automatic evaluation, but it is not feasible to be used. The distance or similarity of MT candidate output to a set of reference translations are used by many MT evaluation approaches. This study presents a comparison of effectiveness of two free online machine translation systems (Google Translate and Babylon machine translation system) to translate Arabic to English. There are many automatic methods used to evaluate different machine translators, one of these methods; Bilingual Evaluation Understudy (BLEU) method. BLEU is used to evaluate translation quality of two free online machine translation systems under consideration. A corpus consists of more than 1000 Arabic sentences with two reference English translations for each Arabic sentence is used in this study. This corpus of Arabic sentences and their English translations consists of 4169 Arabic words, where the number of unique Arabic words is 2539. This corpus is released online to be used by researchers. These Arabic sentences are distributed among four basic sentence functions (declarative, interrogative, exclamatory, and imperative). The experimental results show that Google machine translation system is better than Babylon machine translation system in terms of precision of translation from Arabic to English.

Keywords—component; Machine Translation; Arabic-English Corpus; Google Translator; Babylon Translator; BLEU

I. INTRODUCTION

Machine translation means the use of the computers to translate from one natural language into another. Machine translation dated back to the fifties. Although the translation accuracy of online machine translation (MT) systems is lower than translation accuracy of professional translators, these systems are widely used by different people around the world due to their speed and free cost. The translation process in its own is not a straight forward task, for the order of the target words, and the appropriate choice of target words, essentially affect the accuracy of the outputs of the machine translation systems.

Online machine translators rely on different approaches to translate from one natural language into another, these approaches are Rule-based, Direct, Interlingua, Transfer,

Statistical, Example-based, Knowledge-based, and Hybrid Machine Translation (MT).

Nowadays, automatic evaluation methods of Machine Translation (MT) systems are used in the development cycle of Machine Translation (MT) systems, system optimization, and system comparison. The automatic evaluation of machine translation systems is based on a comparison of MT outputs and the corresponding professional human translations (Reference translations). Automatic evaluation of machine translation systems offers fast, inexpensive, and objective numerical measurements of translation quality. The first methods to automatic Machine Translation evaluation are based on lexical similarity. These are known as Lexical measures (n -gram-based measures) and are based on lexical matching between MT systems outputs and corresponding reference translations [1].

Bilingual Evaluation Understudy (BLEU) is based on string matching, and it is the most widely-used evaluation method to automatically evaluate machine translation systems, and therefore it is used in this study. BLEU is claimed to be language independent and highly correlated with human evaluation, but a number of studies show several pitfalls [1] [2]. BLEU measures the closeness of the candidate output of the machine translation system to reference (professional human) translation of the same text to determine the quality of the machine translation system. The modified n -gram precision is the main metric adopted by BLEU to distinguish between good and bad candidate translations, where this metric is based on counting the number of common words in the candidate translation and the reference translation, and then divides the number of common words by the total number of words in the candidate translation. The modified n -gram precision penalizes candidate sentences found shorter than their reference counter parts; also it penalizes candidate sentences which have over generated correct word forms.

Arabic language is a native language of over 300 million people, and it is the most spoken Semitic language. It is the official language of twenty seven countries, and it is one of United Nations (UN) official languages. Moreover, Muslims around the world use it to practice their religion. Modern Standard Arabic (MSA) is used nowadays in Books, Media, Literature, Education, official correspondences, etc. MSA is derived from Classical Arabic (CA). Arabic language is different from English Language, starting with distinctive features of Arabic script: Arabic language alphabets are twenty-eight, Arabic is written from Right to Left as other Semitic languages, Arabic letters within words are connected

in cursive style, short vowels are normally invisible, and finally Arabic language has no uppercase and lowercase letters (no capitalization in Arabic script) [3].

Many studies present different methods to improve machine translation of Arabic into other languages like Carpuat, Marton, and Habash [4], Al Dam, and Guessoum [5], Riesa, Mohit, Knight, Marcu [6], Adly and Al-Ansary [7], and Salem, Hensman, and Nolan [8].

This paper aims to evaluate the effectiveness of two free online machine translation systems (Google Translate (<https://translate.google.com>) & Babylon (<http://translation.babylon.com/>)) to translate Arabic to English. The necessary resources to accomplish this study like a dataset of Arabic sentences with two English reference translations are not found. Therefore this study includes a creation of a dataset consisting of 1033 Arabic sentences distributed among four basic sentence functions (declarative, interrogative, exclamatory, and imperative).

This study is organized as follows: section 2 introduces the related work, section 3 presents framework and methodology of this study, section 4 presents the evaluation of two free online machine translation systems under consideration using a system designed and implemented by the second author, section 5 presents the conclusion from this research, and, last but not least, section 6 discusses extensions of the this study and the future plans to improve it.

II. RELATED WORK

Three main categories are used to evaluate machine translation (MT): human evaluation, automatic evaluation, and embedded application evaluation [9]. This section presents a number of related studies to this study that means presenting studies concerned with automatic evaluation of machine translation quality only. Studies related to Bilingual Evaluation Understudy (BLEU) as a method to automatically evaluate machine translation quality are presented first. This section also presents some of the studies related to the automatic evaluation of MT that includes Arabic.

It is usual to have more than one perfect translation of a given source sentence. According to this fact Papineni et al. [2] casted BLEU in 2002 as an automatic metric that uses one or more reference human translation beside a candidate translation of an MT system. The increase in the number of reference translations leads to increase the value of this metric. BLEU metric aims to measure the closeness of a machine-translated (candidate) text to a professional human (reference) translation. BLEU uses a modified precision for n -grams at a sentence level and then averages the score over the whole corpus by taking the geometric mean, with n from 1 to 4. The BLEU metric ranges from 0 to 1 (or between 1 and 100). BLEU is insensitive to the variations of the order of n -grams in reference translations.

There are several studies in the literature presenting enhanced BLEU methods, and in this section only three of these are presented due to the limitation of space.

The first study is conducted by Babych and Hartley [10] and aims to enhance BLEU with statistical weights for lexical

items ($tf-idf$ and S) scores. This enhanced model helps to measure translation adequacy, and uses only one human reference translation, and it is more practical than baseline BLEU metric and more effective. Their enhanced model proposed a linguistic interpretation that relates frequency weights and human intuition about translation Adequacy and Fluency. They used DARPA-94 MT French–English evaluation corpus that has 100 French news texts, where average number of words in each of those French news texts is 350 words. Each French news text is translated by five MT systems, and four of these MT translations are scored by human evaluators. The DARPA-94 MT French–English evaluation corpus has two professional human (reference) translations for each news text. They concluded that their model is consistent with baseline BLEU evaluation results for Fluency and outperform the BLEU scores for Adequacy. They also concluded that their model is reliable if there is only one human reference translation for an evaluated text.

The second study is conducted by Yang, Zhu, Li, Wang, Qi, Li and Daxin [11] and proposed adopting proper weights to different words and n -grams into classical BLEU framework. To preserve the language independence in the framework of BLEU, they introduced only the information of the part-of-speech (POS) and n -gram length via linear regression model into classical BLEU framework. Experimental results of their study showed that this enhancement yields better accuracy than the original BLEU.

Chen and Kuhn [12] presented in their study a new automatic MT evaluation method called AMBER. This new method AMBER is based on BLEU, but it has new capabilities like incorporating recall, extra penalties, and some text processing variants. The computation of AMBER is based on multiplying Score by Penalty. The modification includes sophisticated formulas to compute Score and Penalty proposed by these two authors. This modified version of BLEU helps to get more accurate results (evaluations) than the results yield by the original IBM BLEU and METEOR v1.0.

Guessoum and Zantout [13] study presented a methodology for evaluating Arabic machine translation. Those authors evaluated lexical coverage, grammatical coverage, semantic correctness and pronoun resolution correctness. Their approach was used to evaluate four English–Arabic commercial Machine Translation systems; namely ATA, Arabtrans, Ajeeb, and Al-Nakel.

The impact of Arabic morphological segmentation on the performance of a broad-coverage English-to-Arabic Statistical machine translation was discussed in the work of Al-Haj and Lavie [14]. In their work, a phrase based statistical machine translation was addressed. Their results showed a difference in BLEU scores between the best and worst morphological segmentation schemes where the proper choice of segmentation has a significant effect on the performance of the SMT.

Professional human translations (Reference translations) are essential to use BLEU method, but not all automatic evaluation MT metrics need reference translations. One of these methods is a user-centered method introduced by Palmer [15]. Palmer's method is based on comparing the outputs of

This research is funded by the Deanship of Research in Zarqa University / Jordan.

machine translation systems and then ranking them, according to their quality, by expert users who have the necessary needed scientific and linguistic backgrounds to accomplish the ranking process. Palmer's study covers four Arabic-to-English and three Mandarin (simplified Chinese)-to-English machine translation systems.

Most of the people with Arabic as their mother tongue use dialects in their communications at home, markets, etc. One of these dialects is the Iraqi Arabic used mainly in Iraq. To automatically evaluate MT of Iraqi Arabic-English speech translation dialogues, Condon and his colleagues [16] conducted a study and concluded that translation into Iraqi Arabic will correlate higher with human judgments when normalization (light stemming, lexical normalization, and orthographic normalization) is used.

An evaluation of Arabic machine translation based on the Universal Networking Language (UNL) and the Interlingua approach for translation is conducted by Adly and Al-Ansary [7]. The Interlingua approach relies on transforming text in the specified language into a representation form that is language independent that can, later on, be transferred into the target language. Three measures were used for the evaluation process; F_{mean} , F_1 , and BLEU. The evaluation was performed using the Encyclopedia of Life Support Systems (EOLSS). The effect of UNL onto translation from/into Arabic language was also studied by Alansary, Nagi, and Adly [17], and Al-Ansary [18].

Carpuat, Marton, and Habash's [4] study is like our study in that it is concerned with translation from Arabic to English. Those authors addressed the challenges raised by the Arabic verb and subject detection and reordering in Statistical Machine Translation. To minimize ambiguities, the authors proposed a reordering of Verb Subject (VS) construction into Subject Verb (SV) construction for alignment only which has led to an improvement in BLEU and TER scores.

A good survey study was conducted by Alqudsi, Omar, and Shaker [19]. In their study, the issue of machine translation of Arabic into other languages was discussed. They presented, through their survey, the challenges and features of Arabic for machine translation. Their study also presents different approaches to machine translation and their possible application for Arabic. The survey concluded by indicating the difficulty of finding a suitable machine translator that could meet human requirements.

Hailat, Al-Kabi, Alsmadi, and Shawakfa [20] conducted a preliminary study to compare the effectiveness of two online Machine Translation (MT) systems (Google Translate and Babylon machine translation systems) to translate English sentences to Arabic. BLUE metric is used in their study to automatically evaluate the MT quality. They conclude that Google Translate is more effective than Babylon machine translation.

The study of Al-Kabi, Hailat, Al-Shawakfa, and Alsmadi [21] is the closest related study to this one, and it is an improvement too [20]. In their study they also use two free online MT systems (Google Translate & Babylon

(<http://translation.babylon.com/>) to translate English to Arabic, and a corpus consisting of 100 English sentences, and 300 popular English sayings were used. Al-Kabi et al. [21] study concludes that Google Translate is generally more effective than Babylon. The main differences between this study and our study are the size of corpus used in our study is larger than the corpus they collect, and this study is concerned with translation from Arabic to English not translation from English to Arabic.

The study of Al-Deek, Al-Sukhni, Al-Kabi, and Haidar [22] uses ATEC metric to automatically evaluate the output quality of two Free Online Machine Translation (FOMT) systems (Google Translate and IMTranslator). They concluded in their study that Google Translate is more effective than IMTranslator.

III. THE METHODOLOGY

To evaluate the two online MT systems automatically, first we constructed a corpus consisting of exactly 1033 Arabic sentences with two reference (professional human) translations of each Arabic sentence. The two reference translations were conducted by the first author and Dr. Nibras A. M. Al-Omar from Zarqa University. The size of our corpus is 4169 Arabic words, and the number of Arabic unique words is 2539. Table 1 shows the distribution of the Arabic sentences of the constructed corpus among four basic sentence functions. This corpus is uploaded to Google drive server in order to make it accessible to everyone wish to use it. Those who are interested in this corpus can download it using the following URL:

<https://docs.google.com/spreadsheets/d/1bqknBcdQ7cXOKtYLhVP7YHbvrlyJlsQggL60pnLpZfA/edit?usp=sharing>

TABLE I. DISTRIBUTION OF SENTENCES AMONG FOUR BASIC SENTENCE FUNCTIONS

Basic Arabic Sentence Functions	No. of Sentences
Declarative	250
Interrogative	231
Exclamatory	252
Imperative	250
Total	1033

The construction of the above corpus is followed by accomplishing the following main steps shown in Figure 1. BLEU method is used in this study to automatically evaluate the effectiveness of translation from Arabic to English by the two online Machine Translation (MT) systems (Google Translate and Babylon machine translation systems).

The main steps followed to accomplish this study are presented in Figure 1. First the source Arabic sentence is translated using (Google Translate) and (Babylon machine translation systems) and two professional human translations (Reference translations). Then, these Arabic sentences are preprocessed by dividing the text into different n -gram sizes, as follows: unigrams, bigrams, trigrams, and tetra-grams.

After that, the precision for Babylon machine translation system and Google machine translation system were computed for each of the four gram sizes. In the final step, for each of the four n -gram sizes, we compute a unified precision score for that size. These values are then compared to decide which of them gets the best translation.

In order to compute the precision score for each of the four n -gram sizes, we have to count first the number of common words in every candidate and reference sentence, and then we have to divide this sum over the total number of n -grams in the candidate sentence.

To combine the previous precision values in a single overall score (called BLEU-score), we start by computing the Brevity Penalty (BP) by choosing the effective reference (i.e. the reference that has more common n -grams) length which is denoted by r . Then we compute the total length of the candidate translation denoted by c . Now we need to select Brevity Penalty to be a reduced exponential in (r / c) as shown in equation 1 [2]:

$$BP = \begin{cases} 1 & \text{if } c > r \\ e^{\left(1 - \frac{r}{c}\right)} & \text{if } c \leq r \end{cases} \quad (1)$$

The computation of the final BLEU score is shown in formula (2) and it is based on Brevity Penalty (BP) shown in formula (1).

$$BLEU = BP \times \exp\left(\sum_{n=1}^N w_n \log p_n\right) \quad (2)$$

Where $N = 4$ and uniform weights $w_n = (1/N)$ [2].

This indicates that higher BLEU score for any machine translator means that it's better than its counterparts with lower BLEU scores.

Papineni, Roukos, Ward, and Zhu [2] study noted that the BLEU metric values range from 0 to 1, where the translation that has a score of 1 is identical to a professional human translation (Reference translation) [2].

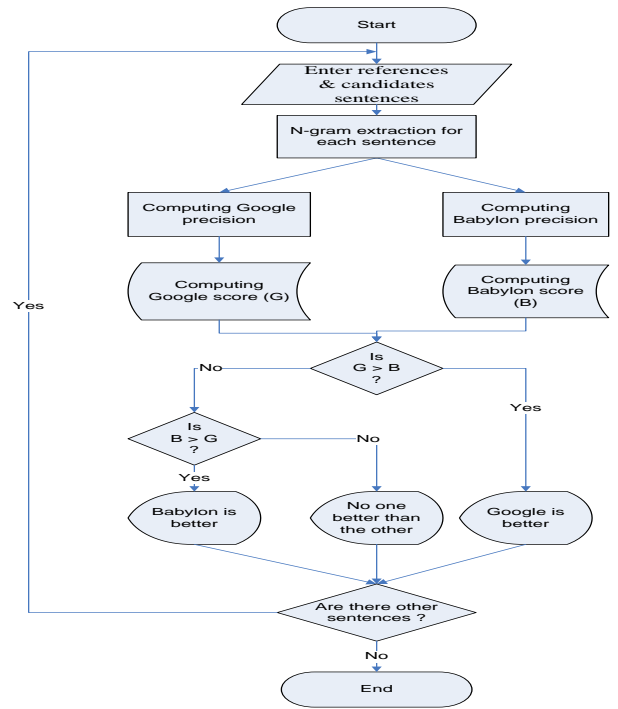


Fig. 1. Evaluation Methodology Flowchart

IV. THE EVALUATION

Many automatic evaluation methods of MT are proposed and used during the last few years, beside manual evaluation methods of MT. Bilingual Evaluation Understudy (BLEU) method is one of the well-known automatic evaluation methods of machine translation adopted in this study.

The following notes resulted from the conducted experiments on Google Translate and Babylon machine translation systems:

- 1) We noticed Babylon MT translates an Arabic word correctly to English, while Babylon MT ignores completely translating those Arabic words in other sentences.
- 2) Babylon MT could not translate the words that contain related pronouns "الضمائر المتصلة", for example the source sentence in Arabic: "سمعت كلمة اضحككتني", that was translated using Babylon as: "I heard the word اضحككتني"
- 3) Babylon machine translation system could not translate multiple Arabic sentences at one time, while Google Translate can translate a set of Arabic sentences at one time.

In our evaluation and testing of the two MT systems, we found that the translation precision is equal for both MT systems (Google and Babylon) for some sentences, but translation precision of Google Translate is generally better than translation precision of Babylon MT system (0.45 for Google and 0.40 for Babylon).

As a whole, the average precision values of Google and Babylon machine translation system for each type of sentences in the corpus are shown in Table 2. It is obvious that Google Translate system is generally better than Babylon machine translation system, but, as shown in table 2, Babylon MT system is more effective than Google Translate in translating Arabic exclamation sentences into English.

TABLE II. AVERAGE PRECISION FOR EACH TYPE OF SENTENCES

Type Translator	Declarative Sentence	Exclamation Sentence	Imperative Sentence	Interrogative Sentence	Average
Babylon MT System	0.3475	0.3686	0.5189	0.3588	0.39845
Google Translate System	0.4486	0.3378	0.5453	0.4668	0.449625

V. CONCLUSION

Arabic-to-English and English-to-Arabic MT have been a challenging research issue for many of the researchers in the field of Arabic Natural Language Processing (NLP).

In this study, we have evaluated the effectiveness of two automatic machine translators (Google Translate System and Babylon machine translation system) that could be used for Arabic-to-English translation and vice versa.

The accuracy of any MT system is usually evaluated by comparing its outputs to that of professional human translators, or professional human translators can manually evaluate the quality of translation. There is no standard Arabic-English corpus that can be used for such evaluations, therefore, we constructed a corpus and released it for free on the Internet to be used by the researchers in this field.

Although the collected data was relatively small in size, the well-known Arabic sayings usually presented a challenge for the machine translation system to translate them to English, and this problem faces us with these MT systems used to translate English sayings to Arabic.

Although the collected data was relatively small in size, the well-known Arabic sayings usually presented a challenge for the machine translation system to translate them to English, and this problem faces us with these MT systems used to translate English sayings to Arabic.

VI. FUTURE WORK

We plan in the future to study the effectiveness of other automatic evaluation MT methods like METEOR, ROUGE, NIST and RED.

We have tested our experiments on a relatively small corpus, and as part of the future work we are planning to build a larger corpus and release it to be used freely by different researchers in this field.

REFERENCES

[1] J. Giménez, and L. Márquez, "Linguistic measures for automatic machine translation evaluation," *Machine Translation* vol. 24, no. 3-4, pp. 209-240, 2010.
[2] K. Papineni, S. Roukos, T. Ward, and W.J. Zhu, "BLEU: a method for automatic evaluation of machine translation," in *Proceedings of the 40th*

Annual Meeting on Association for Computational Linguistics (ACL '02), Stroudsburg, PA, USA, pp. 311-318, 2002.
[3] K. C. Ryding, "A Reference Grammar of Modern Standard Arabic," Cambridge: Cambridge University Press, 2005.
[4] M. Carpuat, Y. Marton, and N. Habash, "Improving Arabic-to-English Statistical Machine Translation by Reordering Post-verbal Subjects for Alignment," in *Proceedings of the ACL 2010 Conference Short Papers*, pp. 178-183, Uppsala, Sweden, 2010.
[5] R. Al Dam, and A. Guessoum, "Building a neural network-based English-to-Arabic transfer module from an unrestricted domain," in *Proceedings of IEEE International Conference on Machine and Web Intelligence (ICMWD)*, pp.94-101, 2010.
[6] J. Riesa, B. Mohit, K. Knight, and D. Marcu, "Building an English-Iraqi Arabic Machine Translation System for Spoken Utterances with Limited Resources", In the *Proceedings of INTERSPEECH*, Pittsburgh, USA, 2006.
[7] Adly, N. and Alansary, S. 2009, "Evaluation of Arabic Machine Translation System based on the Universal Networking Language," in *Proceedings of the 14th International Conference on Applications of Natural Language to Information Systems "NLDB 2009"*, pp. 243-257, 2009.
[8] Y. Salem, A. Hensman, and B. Nolan, "Towards Arabic to English Machine Translation," *ITB Journal*, Issue 17, pp. 20-31, 2008.
[9] K. Kirchhoff, D. Capurro, and A. M. Turner, "A conjoint analysis framework for evaluating user preferences in machine translation," *Machine Translation*, vol. 28, no. 1, pp. 1-17, 2014.
[10] B. Babych, and A. Hartley, "Extending the BLEU MT evaluation method with frequency weightings," in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL '04)*. Association for Computational Linguistics, Stroudsburg, PA, USA, Article 621, 2004.
[11] M. Yang, J. Zhu, J. Li., L. Wang, H. Qi, S. Li., and L. Daxin, "Extending BLEU Evaluation Method with Linguistic Weight," 2008. *ICYCS 2008*. The 9th International Conference for Young Computer Scientists, pp. 1683-1688, 2008.
[12] B. Chen, and R. Kuhn, "AMBER : A modified BLEU, enhanced ranking metric," in *Proceedings of the 6th Workshop on Statistical Machine Translation*, Edinburgh, UK, pp. 71-77, 2011.
[13] A. Guessoum, and R. Zantout, "A Methodology for Evaluating Arabic Machine Translation Systems," *Machine Translation*, issue 18, pp. 299-335, 2005.
[14] H. Al-Haj, and A. Lavie A., "The Impact of Arabic Morphological Segmentation on Broad-coverage English-to-Arabic Statistical Machine Translation," vol. 26, no. 1-2, pp. 3-24, 2012.
[15] D. D. Palmer, "User-centered evaluation for machine translation of spoken language," in *Proceedings. (ICASSP '05)*. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp. v/1013- v/1016, 2005.

- [16] S. Condon, M. Arehart, D. Parvaz, G. Sanders, C. Doran, and J. Aberdeen, "Evaluation of 2-way Iraqi Arabic-English speech translation systems using automated metrics", *Machine Translation*, vol. 26, Nos. 1-2, pp. 159-176, 2012.
- [17] S. Alansary, M. Nagi, and N. Adly, "The Universal Networking Language in Action in English-Arabic Machine Translation," In Proceedings of 9th Egyptian Society of Language Engineering Conference on Language Engineering, (ESOLEC 2009), Cairo, Egypt, 2009.
- [18] S. Alansary, "Interlingua-based Machine Translation Systems: UNL versus Other Interlinguas", in Proceedings of the 11th International Conference on Language Engineering, Cairo, Egypt, 2011.
- [19] A. Alqudsi, N. Omar, and K. Shaker, "Arabic Machine Translation: a Survey", *Artificial Intelligence Review*, pp. 1-24, 2012.
- [20] T. Hailat, M. N. AL-Kabi, I. M. Alsmadi, E. Shawakfa, "Evaluating English To Arabic Machine Translators," 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT 2013) - IT Applications & Systems, Amman, Jordan, 2013.
- [21] M. N. Al-Kabi, T. M. Hailat, E. M. Al-Shawakfa, and I. M. Alsmadi, "Evaluating English to Arabic Machine Translation Using BLEU," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 4, no. 1, pp. 66-73, 2013.
- [22] H. Al-Deek, E. Al-Sukhni, M. Al-Kabi, M. Haidar, "Automatic Evaluation for Google Translate and IMTranslator Translators: An Empirical English-Arabic Translation," The 4th International Conference on Information and Communication Systems (ICICS 2013). ACM, Irbid, Jordan, 2013.

AUTHORS PROFILE

Laith Salman Hassan Hadla, born in Baghdad/Iraq in 1970. He obtained his PhD in Machine Translation from Al-Mustansiriya University in (2006),



his masters' degree was in stylistic translation from Al-Mustansiriya University in (1995), and his bachelor degree is in Translation from Al-Mustansiriya University in (1992). Laith S. Hadla is an assistant professor at the Faculty of Arts, at Zarqa University. Before joining Zarqa University, he worked since 1993 in many Iraqi and Arab universities. The main research areas of interest for Laith S. Hadla are machine translation, translation in general, and linguistics. His teaching interests fall into translation and linguistics



Taghreed M. Hailat, born in Irbid/Jordan in 1986. She obtained her MSc. degree in Computer Science from Yarmouk University (2012), and her bachelor degree in Computer Science from Yarmouk University (2008). Currently, she is working at Irbid chamber of Commerce as a Computer Administrator and previously as a trainer of many computer courses at Irbid Training Center



Mohammed Al-Kabi Mohammed Al-Kabi, born in Baghdad/Iraq in 1959. He obtained his Ph.D. degree in Mathematics from the University of Lodz/Poland (2001), his masters degree in Computer Science from the University of Baghdad/Iraq (1989), and his bachelor degree in statistics from the University of Baghdad/Iraq (1981). Mohammed Naji AL-Kabi is an assistant Professor in the Faculty of Sciences and IT, at Zarqa University. Prior to joining Zarqa University, he worked many years at Yarmouk University in Jordan, Nahrain University and Mustanserya University in Iraq. He also worked as a lecturer in PSUT and Jordan University of Science and Technology (JUST). AL-Kabi's research interests include Information Retrieval, Sentiment analysis and Opinion Mining, Web search engines, Machine Translation, Data Mining, & Natural Language Processing. He is the author of several publications on these topics. His teaching interests focus on information retrieval, Big Data, Web programming, data mining, DBMS (ORACLE & MS Access).