A Linguistic Investigation of Mobile Translation as Speech to Text From Arabic Into English

Assistant Lecturer: Muhammad Atteih Khalaf

Computer-assisted Language Learning

ABSTRACT:

The present study deals with linguistic investigation of mobile translation as speech to text from Arabic into English language.

This study consists of three sections; the first one consists of introduction , the statement of the study, the purpose of the study, the aims of the study and finally the hypothesis of the study. The second one consists of MT concepts, MT approaches, MT types, speech to speech, speech to text, text to speech, machine translation for Arabic texts,scientific, literary MTand evaluation of machine translation.The third one consists of conclusion and bibliography.

The major points for this study is that (MT) highlights the automatic output by electronic machineslike smart cell phone devices interpretations or other electronic devices.Nowadays all the machine programs must test the original text in the SL and to detect lexical and grammatical included in their lexicon storage.

The last section concludesthe area of MT has conventionalobjected its view at the technical translating or further informative versions, with the best focusing and other informative texts. Showing the general notion of making a globalact abletotransfer deep form as entertainment the orderly and expectedoriginal languages storagecut sentence.

Key words : (Machine Translation, Source Language, Target Language).

دراسة لغوية لترجمه الهاتف كمحاكاة للنص من اللغة العربية الى اللغة الانكليزية

م.م. محمد عطية خلف

تعليم اللغة الانكليزية بالحاسوب

وزارة التربية ــ مديرية تربية كركوك

Muhammadatteih@Gmail.com

الملخص:

تهدف هذه الدراسة إلى معرفة الدراسة اللغوية لترجمة الهاتف كمحاكاة للنص من اللغة العربية الى اللغة الانكليزية. تتألف هذه الدراسة من ثلاثة أجزاء، يتكون الجزء الأول من المقدمة، مشكلة الدراسة، الغرض منها وأهدافها. أما الجزء الثاني فيتألف من مفاهيم الترجمة الآلية، طرقها، أنواعها، محاكاتها للفقرة، محاكاتها للنص، الترجمة الآلية للنصوص العربيةوالأدبية، تقييم الترجمة الالية وأخيرا يتألف الجزء الثالث من الاستنتاج والمصادر.

إن الهدف الرئيسي من الدراسة هو إشارة الترجمة الآلية إلى المخرجات الآلية بواسطة الحاسوب أو أجهزة مماثلة كترجمة الهاتف أو الأجهزة الالكترونية الأخرى المتوفرة في الوقت الحاضر كبرامج الترجمة التي يجب أن تحتوي على قواعد لتحليل النص الأصلي وإيجاد تكافؤ قواعدي ومعجمي في ذاكرة قاموسها الالكتروني.

فان في الجزء الأخير يستنتج من الترجمة الآلية بان لها هدف في منظور ترجمة التقنيات أو بتعبير آخر النصوص المتكونة وبالتركيز على النصوص المتكونة الأخرى. وقد بدا جليا من تكوين الفكرة العامة والتمثيل العالمي بالقدرة على تغطية هذا العمق البنيوي كالتمتع بالتنظيم والتنبؤ باللغات الحية بشكل مستقل.

الكلمات المفتاحية : (الترجمة الآلية، اللغة المصدر ، اللغة الهدف).

Introduction:

The translation of natural languages by machine, firstdreamt in the 7th century has become a truth in the end ofthe twentieth. Computer programs have no ideal translations, for that there is anotion to human translator has not the capacityfor looking ahead; nor translations of moral texts, for the subhead and branches of verse are behind computational analysis; but translations of

technical manuals, medical reports, commercial prospectus, scientific documents and administrative notes.(Hutchins. 1995 : 1).

Machine translation history is effected from the borders and regulation of the 1950s and 1960s, the effect of the ALPAC documentin the middle of - 1960s, the restoration in the 1970s, the emission of trade and running systems in the 1980s, research during the 1980s, the modern advancements in research in the 1990s, and the increasing employ of regulations in the previous decade. This summarized history has the capacity of referencing only the main and most important systems and plans, and for more specific readers are indicated to the leaflets listed.(Ibid).

Popovic and Burchardt (2011) investigate the aims of MT and reported that thereare the two main targets of the proposed automatic approach for error analysis and ranking: To evaluate the ranking of mistakes over the mistake classes to detect that error kinds are essentially problematic for a given translation system. To evaluate the differences between the numbers of mistakes in each class for different translation products in order to compare translation system.

Sloculn,.(1985:3) points that the most instant division of translation objectives composed of information conquest versus spreading. The traditional example of the previous objective is reasonable-collecting: with masses of data to screen through, there is no money,time or stimulant to cautiously translate each document or paper by natural (i.e., human) means. Researchersstop with this problem: There area lot of things to read than can

be explained in the time obtainable, and having to work through sentences explained in other languages - when the abilities are reduction that any given sentence is real benefit - is not worth the work. In the previous decade, the lingua franca of science has been English language; this is becoming less and less true for many reasons, because of showing of nationalism and the expansion of technology all over the world. Therefore, scholars who depend on English are having best, hard retaining up with labor in their domain.

Vitek.(2000:1) claims thatthe employment of different idioms for the MT is makinglike "gisting" in the meaning of a "machine-translated sentence", in inversion with "gisting" to "custom translation" (i.e. human translation), etc. Machine-translated sentences have a denial on onespread in that no warranty is specified that the "translation" is good and there isn't sustenance supposed for the machine-translated making. Keepingversus a trail filed by an furiousclient who could think a good translation for his wealthy is particularlywanted when each one sentence hasmanytotalmistakes. Unless a translation is not done by a machine is good, it has the capacity of beinggood only coincidentally because the machine does not recognize the meaning of reliability.

Al-Mubark .(2015) says there are severalcases in translation from Arabic into English: Arabic sentences are notshort but long, Arabic sentences arenot simple but complex, the Arabic phrase is not clear but syntactically ambiguity and so complex for MTin addition to grammatical relationship, there are

maniple meaningsin Arabic words and terms and finally Arabic language has 28 letters and many of them have no equivalent in English.

1- Statement of the Problem :

According to the fast advancement in technology devices nowadays, English learners began torely on MT by mobile in their academic researches and also in their daily life because of the easiness of uses. It is noted that there is a mysterious and unacceptability in this type of translation.

2- The Purpose of the Study:

The main objective of this study is to explain the significance and necessity of mobile translation as speech to text from Arabic into English language and detecting the strength and weakness points in smart mobile phones translation as speech to text in a linguistic investigation in order to depend on them by all English learners and avoiding making the errors.

3- The Aims of the Study:

This study aims at identifying the significance and necessity of mobile translation as a speech to text and to explain Arabic language characteristics impact the advancement of a (MT) instrument from Arabic into English language. Growing the meaning of making a universal actare capable of transferring this deep structure.

2-1   Machine Translation Concepts:

MT points to the automatic output by electronic devices like cell phone translation nowadays, rules must be found in machine program to test the original sentence in the SL and to detect lexical and grammatical composed of its lexicon storage, and to form a new copyof the original sentence in the TL.(Abdul Hameed,1998:10, cited in Jabbar. 2013:348).

Hutchins (1986: 15,cited in Jabbar.2013:348) defines (MT) as "The implementation of devices in the translation of texts, from one natural language into another language" electronic devices and mobiles are going to load out any clearbasics of thumb operation that have done by a human user laboring in a regimental, but an sensible method. The electronic devices must get the results very quickly. (Turing,1951:1), cited in (Jabbar, 2013:348).

Al-Shehab, M. (2017) says that mobiles have many specifics that are being utilized for different pedagogicalexercises. Using of short messages services, (SMS), camera, GPS, WiFi, browsing,downloads, Bluetooth, voice calls and finally gaming makes it more effective. The following researchers Kizito, (2012)& Hoppe (2009) say that (SMS) systems can be used to help all foreign language learners and teachers and they have the ability to use SMS to communicate with group of learners or even groups from learners.

2-2   Machine Translation Methods

2-2-1Direct method:

The strategically employing in systems design the direct translationin the 1950s. The direct method is fundamentally a lexicon-based method in

thatelectronic devices plantranslates a word-for-word changing with some domestic harmony between language bracketsemploying a large bilingual lexicon. (Jabbar. 2013:348).

2-2-2 Interlingua method:

looking for more functional method than the direct translation referred to systems called interlingua in the 1960. Interlingua method relies on discussion which MT should be beyond refined linguistic information (semantics&syntax) include a knowing the meaning of the text. The reaction behind invention an interlingua was not short-lived belief that languages unlike extremely in their — surface forms, all of them share a popular—deep structure. From now on, grow the thought of originating a universal personification able to convey this deep structure as entertaining the harmony and expected orginal languages deficiency sentence independently. (Jabbar. 2013:348).

2-2-3 Transfer Method:

Transfer methodruns three phases: transmit (generation &analysis).

-First, the SL text is parsed into the source-language-specific intermediate syntactic structure.

-Second, linguistic principlesparticular to the language pair change this exemplification into an equivalent exemplification in the TL.

-The final TL text is produced. In another hand; The main notion of this method, there are three stages of translation: Testing of texts into SL, change

into TL, and composition of texts in the TL. The exemplifications produced by the testphase were goaled to solve grammar and lexical mysterious in question of the language, in opposite of the necessity for making other exemplification for idioms and similar forms. The test does not involve the parts that might have more than one comfortable in TL,(e.g. English know & Arabic اعلم اعرف. It is mission of change components to change clear SLexemplification into the convenientexemplification for a special TL. (Ibid).

2-3Machine Translation Types

2-3-1 Rule-Based Machine Translation (RBMT):

More complex than translating vocabulary to vocabulary, these regulations improve linguistic principleswhichpermit the vocabularies to set in several places, to getvarious meanings relying on context, etc. In 1954,the Georgetown-IBM experience was one of the first rule-based MTregulation and Systran and one of the first companies to improve RBMT regulations. RBMT methodology applies a set of linguistic principles in three various stages: test, change and generation. However, a rule-based system needs: grammartest, grammar generation, semantic testing and semantic decent. Mostly idioms, RBMT generates the target text given a source text following the steps.(Costa-Jussa. 2012:248-249).

2-3-2 Statistical Machine Translation (SMT):

CANDIDE regulation isthe most basic a more complicated form of vocabulary translation.New SMT regulations are phrase-based instead

ofvocabulary-based, and collect translations employing the overlap in stages. The basic aim of SMT is the translation of a text specified in some SL into TL.(Ibid).

2-4 Challenging for Arabic Machine Translation

Aransa .(2016:57) shows maintains that MT from any language into Arabic faces the same defiance as human translation between any other languages in addition to some particular issues belonging to Arabic language (such as missing of diacritics or short vowels). There are three basic problemsaccording to Arnold et.al (1993), there are: problems of mysteries, problems are appearing from lexical and structural different languages and finally multiword units like collocation and idioms. The researchers illustrate that the differences between human languages and the translation between similar languages causeddifficulty of translation andwouldn't be more difficult than the translation between non-similar. The researchers divide the difficulties of Arabic translation into several types; the degree of similarity of languages, the problem of ambiguity and human related difficulties.

2-5 Speech to speech

The first speech-only talking translation regulationis Verbmobil. Verbmobil's users have the capacity of picking up a standard smart mobile phone and speech commands and use voice dialing to start a dialog translation session .In contrary with the pastcopies of Verbmobil and other regulations in the C-STAR consortium, the process of the final Verbmobil regulation is perfectly hands-free without any push-to-speak button. Since the Verbmobil speech

translation server has capacity of being accessed by GSM mobile telephones, the system has the capacity of being used anytime and anywhere. No PDA, notebook, Personal Computer or other devices must be obtainable to use the Verbmobil translation service system, just a smart mobile phone for each talking partaker. As well as, there is no waiting time for mouse input and booting computers or keyboard to begin the Verbmobil system is needed— dialog translation has the capacity of starting immediately. In spite of the essential aiming of Verbmobil to upholding face-to-face talks, all partakers use their smart cell phones as translation instruments, Verbmobil has the capacity of being used for talks in that the partakers have not ability to see one another.(Wahlster.1993:1).

Tirronen. (2011:9) refers to the significance of speech as a communication method has created the need of the storage, capture and produce speech by electronic machines. Basics of automatic speech recognition and speech synthesis (or text-to-speech, TTS) are presented. TTS and ASR enable two-way vocal communication between machines and humans, that makes them mandatory process in all speech-to-speech MT systems.

Waibel, A. et.al (2003:370) refers to speech recognition as the most important part of the speech-to-speech interpretation operation. If decipher is not particularly made to continue on a PDA platform that has restrictedstore bandwidth and no fleetingspot, the confession will probably be very slowly for functional use. Multimodal Technologies have been operated on a small

footprint fast decipher HMM-based confession for many years and had important experiments in running on multiple languages and speech-to-speech interpretation regulation. The audio input system on PDAs is low fineness. Specified the hardware'ssize, it is popular that the audio channel has not an electrical noise from motherboard and the power supply, thus recordings on these electronic instruments are not clean. External digitizing of audio could be a choice in the long run, like off-system USB audio, or styling better shielding around the audio hardware, but the researcher's aim was to employcriterion PDAs such as substitutions are not obtainable nowadays.

2-6   Speech to Text

According to experiences as a teachers of English in secondary schools, the plurality of subheadsregulations explained in the literature depend on speech-to-text biasinstead of full automatic speech confession. The original speech must be simplified by re-speakers and speech confession machines are conformed to the captioned sound..(Meinedo. et.al. 2010:93).

2-7   Text to Speech

The fast advance in cell phone technology and rising number of cell phones users have attracted the awareness of cell phone usage advancers. There are many texts finding and translation implementations (Hsueh, 2011, Lue et al., 2010 as cited in Foong. 2013: 1794) from the market but they are in opposed of speech synthesizer. Those implementations have the ability to be used at anywhere and anytime because of the researchers' ability to be combined on the intelligent mobile phones and loadedaround. So, it is

compulsory to have internet connectivity for intelligent mobile phone users to work their intelligent mobile phones implementations in back-end server. Certainty, if free WIFI is not available, it will be costly to do roaming.(Foong. 2013: 1794).

Salami, (2014:3) tries to identify areas where MT has the ability to be useful by investigating what the kinds of texts such as fictional or non-fictional, the free online translation tool Google Translate (GT) has the ability to translate them most accurately. In so doing, GT's accuracy when translating these texts is evaluated. In order to be able to translate the results (e.g. constructions and/or words) MT systems generally have many difficulties.

Arthur, et.al (2014:190-191) define Google 'as an automatic MT service provided by Google'. It translates one written SL to another immediately or with English as a medium. Google translation exercised statistical MT by matching possibilities, using linguistic modeling, and statistical decision notion to detect the most often used translation. The most commonregulation of Google Translate system involved the phrase-based model with reordering and short text pieces, syntactic models and hierarchical, hierarchical phrase-based models. These models performed similar quality of translation. MT translation must be compared with human translation to evaluate the accuracy of translation and the SL at a deeper and more overall textual grade, involving the grades of the vocabularies, semantics, grammar, discourse and pragmatics.

The researchers may have an overall vision on the quality of MT with this multilevel comparison, as compared with human translation and the SL.

.

2-8 Machine Translation for Arabic Texts:

All the researches on accent Arabic MT such as working on normal language operation of accent Arabic texthave concentrated on running accent input words into MSA alternatives before interpreting from any language into English, and the researchers treat with inputs that containing on limitation section of accentvocabularies. The dialectal words normalization in a hybrid (statistical and rule based) MT regulation, by achieving a collection of character-and morpheme-grade mappings. The researchers translated the normalized source to English employing a hybrid MT or instead of that a Statistical MT regulation. The researchers examined their approach on proprietary test forms, controlling about BLEU point (Papineni et al., 2002 cited in Zbib. et.al 2012:50) raise on broadcast news / talking and about spots on web text. Salloum and Habash (2011 cited in Zbib. et.al 2012:50) decreased the ratio of accentout-of-vocabularies also by mapping their affixed morphemes to MSA equivalents (but didn't achieve lexical mapping on the vocabulary stems).(Zbib, et.al 2012:50).

2-9 Machine Translation for Arabic Scientific Texts:

Omar. et. al., (2012) developed an Arabic to English MT for both verb and noun phrases employingchange-based methods, for the noun phrases MT the researchers performed the syntactic reordering for this language couple,

the researchers performed suitable developments in translation fineness over belonged to methods, The researchers' methodsexaminedmany thesis headlines and newspapers from the computer science field. Average of the accuracy of their score was 94.6%. As in the verb phrase MT regulation, the researcher's survey was to show Verbal Sentence rule based MT, theregulationexamined on 45 verbal sentences from different scientific text in Arabic languageand trained on 30 modern verbal sentences from different fields. The researchers examined their regulationas opposed totwo different MTregulations, namely Google and Systran. Accuracy of the score was 93%. Salem et al. (2008:26) developed an Interlingua rule based method to interpret from Arabic into English named Arabic, that is concentrated on the function and Reference Grammar Linguistic Model (RRG), researchers employed the exemplification and the logical form of an Arabic sentence. Aiming was to explain how the characteristics of the Arabic language will impact the advancement of a (MT) instrument from Arabic to Englishlanguage.

2-10 Machine Translation for Scientific Text

Building perfect translation tools for scientific texts which are automatically adapted to different scientific areas is the challenge of the French National project. These systems will be carefullyinserted into the HAL open rcord, a multidisciplinary open-access recordthat was formed in 2006 to record publications from all the scientific society of French. As well as providing the automatic translations, the interface will also permit to post it the production, which itself will use to develop the machine translation. The improvement of

a phrase based statistical machine translation system (PBSMT) for the translation of scientific papers in many fields from French to English and vice versa. All these papers in the HAL archive may be scientific research papers(Lambert, et.al .2010: 3933).

2-11 Machine Translation for Literary Text

The meaning of literary MTcould seem at first to be a near-denial in expressions. The area of MT has classicallyobjective its vision at the translation of technical or informative versions, with the more powerful concentration on newswire and other informative versions attach to the aims the funders of the government. Although, the likelihood of literary MT is sympathetic. HT of literary versions is an frequently money - time - heavymission, but one that is a decisive part of the international regulation of translation literary interchange. From a technical point of view, since by explanation, literature is the art that employs language. Literary translation exemplifies may be more powerfulform of the MT problem.(Voigt & Jurafsky.2005:1).

2-12 Evaluation of Machine Translation

Translation quality evaluation ensures that post-edited translations are at least equal to familiar made translations. It refers to the significance of the information density, regarding the method, and structures of the linguistic source version when matching the performance of different translation junctures. Prose-like versions, like letters and company profiles, may be translated slower with MT translation, but the conclusion may be better

translation because of the translator's ability to concentrate on editing the version to be suitablewith its purpose contrary with concentrating on translating structures and words. Contrary that, it doesn't take much shorter to translate information-densityversions (like those depending primarily on bullet points) from scratch, that is why of good-quality MT translation has the ability to help so much. With these kinds of informative texts, editing for cohesion and linguistic style is much less significant.(Laubli.et.al .2013:89).

Conclusions:

Speech recognition as the most important part of the speech-to-speech interpretation operation. MT has the ability to be useful by investigating what the kinds of texts such as fictional or non-fictional, the free online translation tool Google Translate (GT) has the ability to translate them most accurately. In so doing, GT's accuracy when translating these texts is evaluated. Computers application and mobiles devices in the translation of sentences from Arabic into English language, these devices are intended to carry out any definite principles of special operations.Concept of making a universal exemplification able to convey this deep structure as entertaining harmony and foretell natural languages weakness of the sentences separately.In this research, Language explains how the characteristics of the Arabic language effect in the improvement of a (MT) device from Arabic into English language.The previous researches on accent in MT have concentrated on

composing accent input vocabularies and sentences, as well as, the dialect Arabic text including MT.The evaluation of MT accuracy must be compared with human translation in order to explain the average of the accuracy of translation. Most of modern smart mobiles phones are incorporated with browsing applications such as Internet explorer, Mozilla fire fox and finally Opera-mini.All learners and teachers have the ability to download translation video through mobile phone easily and present to students.

Bibliography

Al-Mubark, A.(2015). **Language and Culture.***The Journal of Literature, Language and Culture*. University of Jazan, Saudi Arabia.

Al-Shehab, M.(2017). **The Effect Of Using Mobile In Students' Ability In Translation From English Into Arabic At Jadara University.** *The International Journal of Applied Linguistics And Translation. Jadara University, Irbid, Jordan.*

Aransa, W.(2016). **Statistical Machine Translation of The Arabic Language.** *University of Maine, French.*

Arthur, H., Graesser, C & Cai, Z.(2014). **Comparison of Google Translation with Human Translation**. *Association for the Advancement of Artificial Intelligence.*

Costa-Jussa, M.(2012). **Study and Comparison Of Rule-Based And Statistical Catalan-Spanish Machine Translation Systems.***Computing and Informatics.*

Foong, M., Sulaiman,S., Kiu Ling, K. (2013)**. Text Signage Recognition In Android Mobile Devices.***Journal of Computer Science.*

Hutchins, J.(1995). **Machine Translation: A Brief History.** *Oxford: Pergamon Press.*

Jabbar, R. (2013). **A Linguistic Study of New Trends in Translation as a Speech to Act.***International Journal of Science and Research (IJSR)*

Lambert, P., Schwenk, Ch., Blain,F.(2010). **Automatic Translation of Scientific Documents in The HAL Archive.***University of Le Mans.*

Läubli1, S., Fishel, M., Massey, G., Ehrensberger-Dow, M., Volk, M.(2013). **Assessing Post-Editing Efficiency In A Realistic Translation Environment**. *Workshop on Post-editing Technology and Practice, Nice, September***.**

Meinedo, H.et.al.(2010). **The L2F Broadcast News Speech Recognition System.***VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop.*

Omar, N., Albared, M., Juzaiddin Ab. Aziz,M., Zakree, M. Nazri ,A.(2010)**. Automatic Part of Speech Tagging for Arabic: An Experiment Using Bigram Hidden Markov Model**. *In Rough Set and Knowledge Technology, pages 361–370. Springer,.*

Popovic, M & Burchardt.(2010). **From Human to Automatic Error Classification for Machine Translation Output**. *2010 European Association for Machine Translation.*

Salimi, J .(2014). **Machine Translation Fictional and Non-Fictional Texts**. *Stockholmes University, Swede.*

Sloculn, J.(1985).**A Survey of Machine Translation: Its History, Current Status, And Future Prospects.** *Computational Linguistics, Volume 11, Number 1, January-March.*

Tirronen, S.(2011).**Automated Testing Of Speech-To-Speech Machine Translation In Telecom Networks**. *Aalto University.*

Vitek, S.(2000). **Reflections of a Human Translator on Machine Translation.** *Translation Journal.*

Voigt, R., Jurafsky, D.(2005). **Towards a Literary Machine Translation: The Role of Referential Cohesion.***Center for East Asian Studies, Stanford University.*

Wahlster,W.(1993). **Mobile Speech-to-Speech Translation of Spontaneous Dialogs: An Overview of the Final Verbmobil System.***DFKI GmbH, Saarbrücken, Germany.*

Waibel, A.(2003). **Speechalator: Two-Way Speech-To-Speech Translation On A Consumer PDA.** *Eurospeech,2003.Geneva.*

Zbib, R., Malchiodi, E., Devlin, j., Stallard, d., Matsoukas,s., Schwartz, R., Makhoul, J., Zaidan,O., Burch, CH.(2012:50). **Machine Translation of Arabic Dialects.***Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.