A STUDY OF CONTEXT INFLUENCES IN ARABIC-ENGLISH LANGUAGE TRANSLATION TECHNOLOGIES

ZAILAN ARABEE ABDUL SALAM
RABIAH ABDUL KADIR

ABSTRACT

Social and cultural context is largely missing in current language translation systems. Dictionary based systems translate terms in a source language to an equivalent term in a target language, but often the translation could be inaccurate when context is not taken into consideration, or when an equivalent term in the target language does not exist. Domain knowledge and context can be made explicit by using ontologies, and ontology utilization would enable inclusion of semantic relations to other terms, leading to translation results which is more comprehensive than a single equivalent term. It is proposed that existing ontologies in the domain should be utilized and combined by ontology merging techniques, to leverage on existing resources to form a basis ontology with contextual representation, and this can be further enhanced by using machine translation techniques on existing corpora to improve the basic ontology to append further contextual information to the knowledge base. Statistical methods in machine translation could provide automated relevance determination of these existing resources which are machine readable, and aid the human translator in establishing a domain specific knowledge base for translation. Advancements in communication and technologies has made the world smaller where people of different regions and languages need to work together and interact.The accuracy of these translations are crucial as it could lead to misunderstandings and possible conflict. While single equivalent terms in a target language can provide a gist of the meaning of a source language term, a semantic conceptualisation provided by an ontology could enable the term to be understood in the specific context that it is being used.

Keywords: Machine Translation, Contextual Knowledge, Knowledge Representation

INTRODUCTION

Language forms the basis of human communication, and the modern world has made the world much smaller with the advancement of transportation and communication technologies. Cross cultural communication across different languages necessitates a rapid and accurate language translation facility to enable people to communicate with each other across the globe. The language of different communities is as vast and different as the individual customs and cultures practiced by those communities. As an example, the classical Arabic language is rich in meaning and provides specific terms to refer to specific context and meaning. A single term could have several meanings, some of which are useful and related to the context of the term being used, and some meanings are understood to be not relevant to the current utilization of the term. "The Arabic language is considered one of the most prolific languages in terms of language material" (Ishkewy et al., 2014). Arabic has been pointed out as "notorious for complex morphology", making the translation process difficult (Kadhim et al., 2013). "Language and culture are closely related and it is essential to consider both in the process of translation"(Braçaj, 2014). Any attempt to translate the Arabic Quran normally comes with a disclaimer by the translator to state "No translation of Quran can be one hundred percent accurate" (Tarazi n.d.; Zadeh n.d.), and their concerns can be understood. The Quran alludes to events without stating the historical background. Those who heard the quran at the time of revelation from the Prophet Muhammad were fully aware of the circumstances leading up to a particular verse (Haleem, 2005). Haleem (2005) further states that "Interpretation is further

complicated by the highly concise style of the Qur'an. A verse may contain several sentences in short, proverbial style, with pronominal references relating them to a wider context. Moreover, proverbial statements can be lifted from the text and used on their own, isolated from their context and unguided by other references in the Qur'an that might provide further explanation." This poses a problem when a Muslim who does not speak or understand Arabic tries to learn about Islam by reading translations of the Quran, as the translation might not provide the whole picture needed for the muslim to truly understand the word of Allah through the Arabic Quran. The obvious solution would be to state that a muslim should learn Arabic and be able to read and understand the holy scripture directly in the language that the Quran was revealed, but this solution has some practical issues, and while the learning of the language is taking place, a practicing muslim would need to understand and learn about Islam through translations and *tafseer*, or scholarly interpretations of the Quran and Hadith in their native language. A more comprehensive contextual representation of the concepts would assist the learner in grasping not just the gist of the concept, but a deeper and closer understanding of the concept to enhance semantic understanding.

Language translation poses many challenges, among them is the preservation of the meaning or context contained in the source language, and the attempt of finding an equivalent term in the target language, which may or may not exist. Kashgary (2010) contends "if equivalence is the essence of translation, non-equivalence constitutes an equally legitimate concept in the translation process." She asserts that different languages are strongly attached to different cultures, and the example of Arabic language is cited. The holy book of Islam, the Holy Quran, was conveyed to Prophet Muhammad (pbuh) in the Arabic language and many different translations of the holy book have been produced, but these translations only serve to convey the gist of the meanings of the holy Quran. Arabic, being rich in culture specific terms, often have no equivalence for its terms in the English language, and terms which are not lexicalized in the English language. Terms with no equivalence will need more than one term in another language to convey the contextual and cultural meaning. A simple example is the word Al-Lateef (اللطيف) in Arabic, which is one if the beautiful names of Allah. Google translates provides 38 equivalent terms for this word, which includes 'gentle', 'nice', 'kind', and 'soft', among others. Most Bahasa Melayu sources would translate it as 'yang Maha Halus' or 'yang lembut', to mean the subtle or soft. Native speakers of Arabic would have to explain the term in a sentence, and for the example of Al-Lateef it could be stated as "He has knowledge of all things, subtle and hidden, and He shows His immense mercy to His slaves in very subtle ways known to no one but Him." (Assayyid 2014). Another example is Niyyah (نيّة) which is commonly translated as 'intention' in dictionaries, and context translation engines such as Reverso provide examples of how this term is used in several Arabic sentences to show how it could be used in different contexts, but the meaning still circles around intention and purpose, which is deemed sufficient for general translation purposes. However, when scrutinized further for religious analysis and in depth understanding of a fundamental concept in Islam, scholars will highlight that this term also means 'seed'. Cultural and contextual information will elevate the concept of Niyyah or Intention to assert that any good intention is rewarded by God, but the intention is comparable to a seed, where a seed is planted in the heart, unseen and unheard by other people, emphasizing the sincerity and purity of the intention.   This shows contextual and background knowledge is often required to portray the true or closest meaning of a term, and that a single term in English is not sufficient to describe a single term in Arabic, as stated in the examples above.

"Non-equivalence is a fact among languages" (Kashgary, 2010). Different cultures possess different rules of appropriateness, and due to cultural differences native English speakers and Arabic speakers assign different degrees of severity to the same situation (Abdul Wahed Qasem Ghaleb Al-Zumor, 2011). "The study of language, culture and translation and the relationship

between them are crucial issues because of the importance of human communication in the world" (Braçaj, 2014). Braçaj (2014) further asserts that language is culturally embedded, expressing and shaping cultural reality, and the meanings of linguistic items, and it can only be understood when considered together with the cultural context in which these linguistic items are used. Translation is therefore always entangled in a set of relations that exist in both the source and target contexts (Braçaj, 2014).

With the advent of globalization, there is an increase in use of cross language research, but the focus on the influence of language differences and translation processes when working across languages has received little attention (Al-Amer et al., 2015). Twinn (as cited in Al-Amer et al., 2015) states that the reliability of the translation process is critical to the integrity of the results as this can be threatened if there are inaccuracies, such as when the data are analysed using thematic translation in the research process, where the social context is an integral part of the data analysis. This is supported by Kashgary (2010) and Al-Zumor (2011).

The common approach utilised today is dictionary based, where words that have more than one meaning will decrease the retrieval performance if the query translation returns an incorrect translation, or a translation equivalent which is out of context (Yahya et al., 2013). Dictionary based translations are limited where the search term becomes untranslatable, and this can be caused by several factors such as limitations of the dictionaries in terms of vocabulary, the processing required of inflected words, phrase identification and translation, and lexical ambiguity of the source language (SL) and target language (TL) (Pirkola et al., 2001). This paper is particularly concerned with the lexical ambiguity aspect due to the context of the word being used. Some efforts have been undertaken to solve ambiguity when using the dictionary based approach. Reddy and Hanumanthappa (2012) utilized the Greedy Algorithm to calculate a measure of cohesion between words in a phrase, where a low cohesion would indicate ambiguity for certain terms. The contextual information needs to be stored in a machine readable representation in order for machine translation results to produce more accurate and contextually relevant translations.

BACKGROUND OF THE STUDY

MACHINE TRANSLATION

Machine Translation (MT) has a long history which can be traced to the advent of utilization of computers in processing information in the 1950s, and a detailed historical account is given by Hutchins (1995). Most of these early efforts were based on the direct translation method using word- for-word translation with minimal or no analysis of the SL words (Bolia & Slyh, 2011). This long history involves many projects but are largely based on the three different levels of linguistic description; morphology, syntax and semantics. The late 1980s saw the emergence of the 'corpus based' and 'example based' approaches, which showed promising results, and statistical methods started gaining ground.

Machine Translation can be classified into two approaches: either by level of representation, or by sources of information. When classifying MT by level of representation, we can refer to three levels of direct, transfer, and interlingua approaches. Simple dictionary based translations fall under the direct approach where a single step of transformation between SL and TL produces an output. Rule Based MT (RBMT), Example Based MT(EBMT) and Statistical MT (SMT) fall under the transfer approach where three steps of analysis, transfer and generation is required to translate the SL to TL. The Interlingua approach requires only two steps, first is the analysis to transform the SL to interlingua representation, and second is generating the TL from this (Costa-Jussa & Fonollosa 2015). The challenge of the interlingua approach is that if more languages are to be supported by the interlingua and the types of text the system will have to

translate becomes broader, the more difficult it would be to design a sufficiently rich interlingual representation to capture all the word senses (Bolia & Slyh, 2011).

When classifying MT by sources of information, it can be classified by either rules (RBMT), which is linguistic based, or data, which is statistics based (EBMT and SMT). Figure 1 below outlines approaches of MT architectures, which involves rule based or corpus based MT. RBMT approaches require extensive knowledge of linguistics and Natural Language Processing elements. This approach works within limited example sentences, but faces difficulties; 1) when used on real world text, and 2) updating the rules to accommodate more real world cases and contexts often override the earlier rules; which worked well within their limited scope, but the updated rule might no longer accommodate the earlier scope and context (Bolia & Slyh, 2011).

Recent works shows that there is more interest in EBMT and SMT as there is more access to text corpora on the internet. EBMT utilizes pattern matching and performs translation by analogy, and SMT attempts to find the most probable translation given a SL term or sentence. SMT treats the translation of natural language as a machine learning problem (Lopez, 2012).
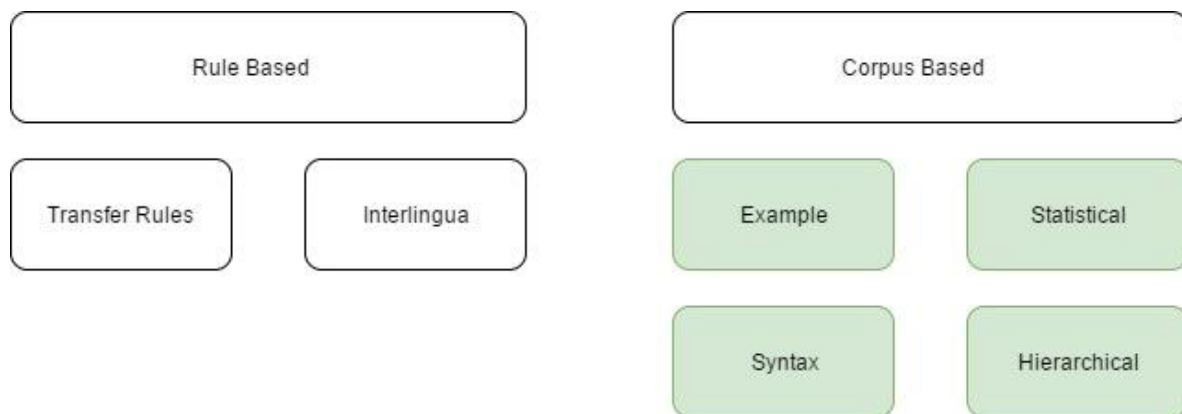
FIGURE 1. Machine Translation Architectures

Many recent efforts in MT employ a hybrid approach, combining rule based and corpus based techniques. Recent works in SMT show promising results, but are still not error-free (Peris et al., 2016). A contributing factor to failures of machine translation and large-scale intelligent decision support systems is the unavailability of "common sense" about the world and about language and its contextual use (Bolia & Slyh, 2011). It is proposed that MT techniques such as EBMT and SMT could be used on relevant existing Arabic and English corpora to determine relevance of terms and generate a knowledge base with contextual information, which could benefit language translation for a particular domain. MT approaches will need to be further researched to determine specific approaches which would most suit this purpose.

CONTEXTUAL INFORMATION

Context is defined as the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed. To fully understand the meaning of a word, we can look at the roots or origins to which the word was conceived. The word originates from combining the Latin terms *con* (together), with *texere* (to weave), which emphasizes the need to support a concept by weaving and binding it with other concepts or terms. Terminology identifies terms in a specific domain and denotes a concept in the domain. Specific terminology is used in a sentence and the term is attached to other ideas or concepts in the same language. This study discusses how terminology can be represented in knowledge bases to capture context, or in other words, semantic relations to other concepts or terms, and

advancements in machine translation efforts to improve on the accuracy of single term translation.

To know the translation and to analyze a word; knowing the root is essential. Arabic and semetic dictionaries requires knowledge of the root of the word in order to search for the word in the dictionary, and word origins are provided in most dictionaries to augment the reader's understanding of the term. The accuracy of a translation depends not only on the root of the word, but also the context which it is being used, particularly for single term translations. Recent works such as Gupta et al. (2016), Choi et al. (2017) and Brien et al. (2011) show that research on contextual information is still very much of interest and is an area with much work in progress. Several types of semantic relations can be defined for a term, such as synonym sets, taxonomic relations, lexical semantic relations, and transversal relations. Earlier work in the 90's looked at utilizing classical dictionaries for semantic representation, but conceptual relations were missing (Mallat et al., 2015). There was a need for other knowledge representation formalisms to account for relations other than synonymy. According to Genesereth and Nilsson (as cited in Gruber (1993)), in order for machines to 'read' or to distinguish semantic differences of terms in various languages, the terms need to be formally represented based on a conceptualization, which are concepts that are presumed to exist in a particular area and the relationships that hold them. An explicit specification of a conceptualization is known as an ontology (Gruber, 1993). One of the reasons to develop an ontology is to make domain assumptions explicit (Noy & McGuinness, 2001).

The problem of modeling context, defining terms and limiting or scoping domain knowledge to enhance semantics for specific purposes can be seen in wide domains of knowledge. Evermann and Wand (2005) describes their work to formalize the domain concepts for clarity in information systems design and development, in order to scope and limit the terms being used in the language to model their domain of interest using ontologies. Also within the scope of systems development, Storey (2005) utilized ontologies for classifying relationship verb phrases based upon the domain and context of the application for the purpose of comparing two conceptual database designs for integration and validation. In the field of image processing and recognition, Rasiwasia and Vasconcelos (2012) proposed a framework to context modeling based on the probability of co-occurrence of objects and scenes, and show that context based approach outperforms appearance and other existing representations in the tasks of scene classification and image retrieval. Where a wide variety of resources need to be utilized together, similar concepts need to be mapped. Ram and Park (2004) describe a Semantic Conflict Resolution Ontology (SCROL) which is a dynamic mechanism of comparing and manipulating contextual knowledge of each information source, and show that this is useful in achieving semantic interoperability among heterogeneous databases.

The need for contextual clarity is apparent in many fields of knowledge, and the most popular choice for contextual representation in most published works is ontologies. Coming back to the area of Machine Translation, terms which are deemed relevant or equivalent which cannot be confirmed by the machine learning system, can be searched manually by the translator, and this process should be captured by the system to encode the ontology with the relevance of the term in the TL to gain some contextual or social information of the SL terms. This will require that the information be stored in a machine readable form.
There have been attempts at utilising ontologies to automatically translate terms, such as the OntoLearn system for automated ontology learning which automatically translates multiword terms from English to Italian (Navigli & Velardi, 2003). There are other ontology based translation efforts (Helou et al., 2016; Yahya et al., 2013; Navigli & Velardi, 2003) which shows promising results, but also state the need for further work in this area.

An ontology would be able to explicitly represent the concepts to be understood by humans and also machine readable. According to Genesereth and Nilsson (as cited in Gruber

1993)), in order for machines to 'read' terms in various languages, the terms need to be formally represented based on a conceptualization, which are concepts that are presumed to exist in a particular area and the relationships that hold them. An explicit specification of a conceptualization is known as an ontology (Gruber, 1993). As ontologies are seen as a means of sharing and reusing knowledge (Hitzler et al., 2005), consolidating the knowledge from multiple ontologies could potentially assist in enhancing the definition of terms semantically. Specific ontologies are created for specific purposes, and would not be comprehensive in linking terms for general domains. By combining available ontologies, the pool of terms or resources would be widened, hence increasing the probability of the equivalent term being available and sufficiently described. One of the reasons to develop an ontology is to make domain assumptions explicit (Noy & McGuinness, 2001). The problem of modeling context, defining terms and limiting or scoping domain knowledge to enhance semantics for specific purposes can be seen in wide domains of knowledge. Evermann and Wand (2005) describes their work to formalize the domain concepts for clarity in information systems design and development, in order to scope and limit the terms being used in the language to model their domain of interest using ontologies. Also within the scope of systems development, Storey (2005) utilized ontologies for classifying relationship verb phrases based upon the domain and context of the application for the purpose of comparing two conceptual database designs for integration and validation. In the field of image processing and recognition, Rasiwasia and Vasconcelos (2012) proposed a framework to context modeling based on the probability of co-occurrence of objects and scenes, and show that context based approach outperforms appearance and other existing representations in the tasks of scene classification and image retrieval. Where a wide variety of resources need to be utilized together, similar concepts need to be mapped. Ram and Park (2004) describe a Semantic Conflict Resolution Ontology (SCROL) which is a dynamic mechanism of comparing and manipulating contextual knowledge of each information source, and show that this is useful in achieving semantic interoperability among heterogeneous databases.

The need for contextual clarity is apparent in many fields of knowledge, and a popular choice for contextual representation in most published works is ontologies. Coming back to the example of the term Niyyah or intention, a representation in an ontology could be envisaged in Figure 2 below. By combining data from root word dictionaries and other relevant Islamic sources of knowledge, the term can be related to other concepts and the context of the concept of an intention in Islam can be better understood.

Surveys on the literature of ontology matching and merging show that this area is very much relevant and various efforts published provide details on fully or semi-automatic tools to assist the human knowledge engineer in the matching process (Ramar & Gurunathan, 2016; Amrouch & Mostefai, 2012).
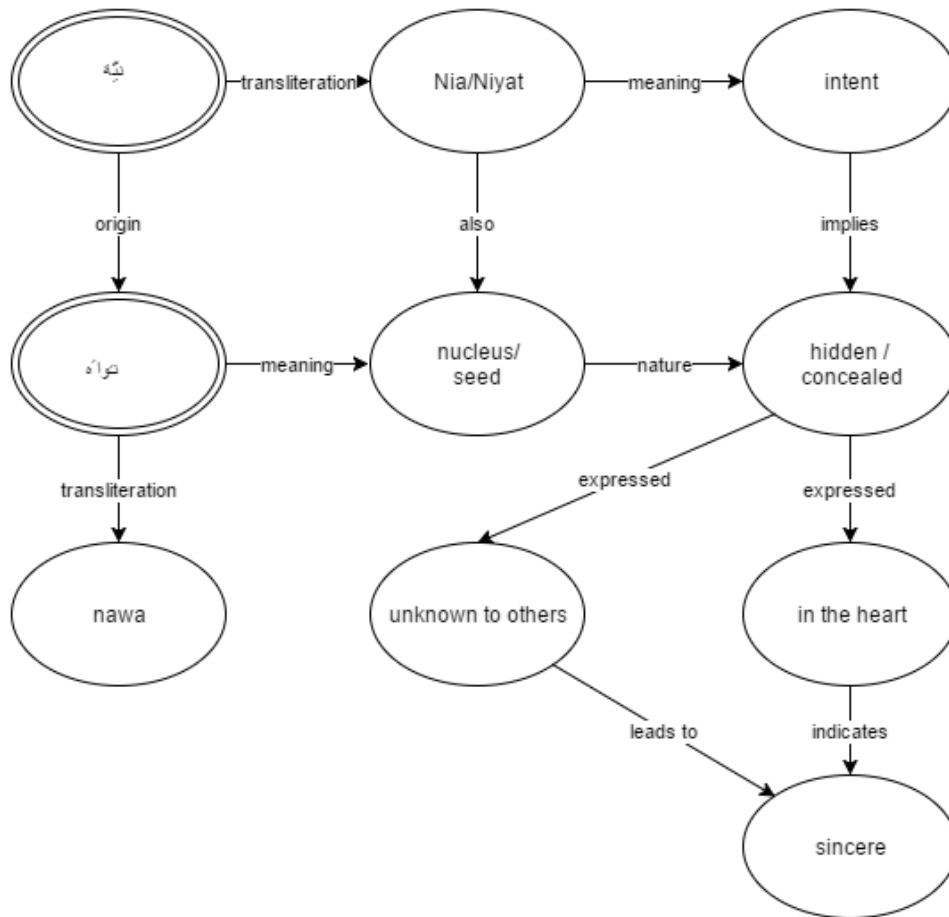
FIGURE 2. Sample Ontology Representation of An Arabic Term

DISCUSSION

Our ability to communicate with others across the globe has exponentially increased with the advent of the internet and communication technologies. The ability to learn new knowledge, explore new cultures, and interact with other people who speak a different language is hindered by the inability to understand the foreign language. Pew Research Center projected that by 2050 there will be near parity between Muslims (2.8 billion, or 30% of the population) and Christians (2.9 billion, or 31%), possibly for the first time in history. The uniting factor for this large population of Muslims would be the Quran and Hadith. Google Translate is a state-of-the-art machine translation tool which provides a billion translations a day for 200 million users (CNET 2013). Automated and Machine Translation accuracy or processes is a current problem and remains as an active research area, which can be seen in the number of recent publications related to this topic (Alshehab , 2015; Habash & Hassan, 2012; Helou et al., 2016; Kadhim et al., 2013; Maedche & Staab, 2005; Navigli & Velardi, 2003; Yahya et al., 2013)

The ability of computing systems to assist the human interpreter in language translation is apparent with the advent of various applications and technologies. Machines are able to assist if they are able to read the term just as humans do, albeit with significantly diminished level of understanding. Wang et al. (2011) describes "a framework for cross-language information retrieval that efficiently leverages statistical estimation of translation probabilities" (Wang et al., 2011). This could help by processing or preparing the document to establish relevancy of terms automatically, or semi-automatically with a human expert to fine tune the final translated document. These technologies may assist in the translation process, however, the role of the

human expert in ensuring the meaning is preserved is paramount. Translation is an interpretive process and not merely a direct message transfer from a source language to a target language (Al-Amer et al., 2015). Translations of metaphors or culture specific terms would require a human translator who is aware of the SL culture and sayings, to avoid a literal translation which does not make any sense to a reader of the TL . Terminology and concept mismatches between two ontologies require human knowledge and effort to resolve the ambiguities (Klein, 2001). At the present time, fully automated mapping between two or more knowledge bases is not feasible, due to the inadequacy of Natural Language Processing technologies, and manual mapping is an expensive and laborious task (Amrouch & Mostefai, 2012). Shvaiko and Euzenat (2013) assert that users should be involved in recommending relevant background knowledge in advance, be able to influence the selection and weighting of the various matching components, criticize aspects of intermediate results, and determine whether the final result is good enough to be put to use. They further state that little attention has been devoted so far to the realization of interfaces that actually allow users to become active in these ways. Systems should be developed on the basis of continual tests with final users, and the ultimate success criterion will be the extent to which the system has value for them (Shvaiko & Euzenat, 2013). Visual tools could assist the translator to automatically include relevant terms in the TL to widen the 'equivalence' of the term in the SL. This can be done by filtering and proposing terms which are considered to have an acceptable probability of relevance to the translator, in order to not overwhelm the translator with too much visual information. It is proposed that existing ontologies in the domain should be utilized by ontology merging techniques, to leverage on existing resources to form a basis ontology with contextual representation, and this can be further enhanced by using EBMT and SMT techniques on existing Quran translation and tafseer corpora to enrich cultural context that is captured in ontologies. Some background information could be appended to a verse or a single term in a verse in the ontology to enlighten the reader on the context of the revelation of the verse, which will clarify the verse and hence the term being translated or examined. Visual tools could then be utilized by human translators to verify the enriched ontology produced via MT approaches, to ensure that translation results produced by utilizing the ontology is accurate.

CONCLUSION

Social and cultural context is largely missing in current translation systems. As ontologies are able to store explicit descriptions of concepts, properties and restrictions in the domain, when combined with specific instances of terms or phrases in the domain, it is anticipated that ontology utilization will improve semantic relations to other terms and the translation results will be more comprehensive than a single equivalent term. Numerous practical implications can be envisioned from further study of this area. Language learners will not only get a clearer meaning of a term in all of its possible usages, but also a glimpse of the culture of the people of the source language being learned.

Different nations with different languages will continue to utilise and strengthen their native languages, and the need for language translation is seen as a perpetual need for mankind to continue to work together. Machine Translation provides an automatic approach to make translation facilities available to anyone, at any time. The accuracy of these translations are crucial as it could lead to misunderstandings and possible conflict. Domain knowledge can be made explicit by using ontologies. Statistical methods in MT could provide automated relevance determination of these existing resources which are machine readable. While single equivalent terms in English can provide a gist of the meaning of an Arabic term, a semantic conceptualisation provided by an ontology could enable the term to be understood in the specific context that it is being used.

REFERENCES

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M. & Salamonson, Y. 2015. Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies,* 54: 150–157.

Al-Zumor, A. W. Q. G. 2011. Apologies in Arabic and English: An inter-language and cross-cultural study. *Journal of King Saud University – Languages and Translation*, 19–28.

Alshehab, M. 2015. Two English Translations of Arabic Metaphors in the Holy Qura'n Mohammad Alshehab Jadara University-Jordan. *Arab World English Journal (AWEJ) Special Issue on Translation,* (4): 295–306.

Amrouch, S. & Mostefai, S. 2012. Survey on the literature of ontology mapping, alignment and merging. *2012 International Conference on Information Technology and e-Services, ICITeS 2012*.

Assayyid, M. M. B. 2014. *The Beautiful Names of Allah*. Riyadh.

Bolia, R. S. & Slyh, R. E. 2011. Representation and comprehension in machine translation and intelligent decision support. *IEEE Intelligent Systems,* 26(4): 40–47.

Braçaj, M. 2014. Reflection on Language, Culture and Translation and Culture as a Challenge for Translation Process. *Journal of Educational and Social Research*, 4(4).

Brien, S. O., Benjamins, J. & Brien, S. O. 2011. Integrating source-language context into phrase-based statistical machine translation. *Machine Translation,* 25(3): 239–285.

Choi, H., Cho, K. & Bengio, Y. 2017. Context-dependent word representation for neural machine translation. *Computer Speech & Language*, 45: 149-160.

CNET. 2013. Google Translate now serves 200 million people daily. http://www.cnet.com/news/google-translate-now-serves-200-million-people-daily/ [9 June 2016].

Costa-Jussa, M. R. & Fonollosa, J. A. R. 2015. Latest trends in hybrid machine translation and its applications. *Computer Speech and Language,* 32(1): 3–10.

Evermann, J. & Wand, Y. 2005. Toward formalizing domain modeling semantics in language syntax. *IEEE Transactions on Software Engineering*, 31(1): 21–37.

Gruber, T. R. 1993. Technical Report KSL 92-71 Revised April 1993 A Translation Approach to Portable Ontology Specifications by A Translation Approach to Portable Ontology Specifications. *Knowledge Creation Diffusion Utilization,* 5(April): 199–220.

Gupta, P., Costa-jussà, M. R., Rosso, P. & Banchs, R. E. 2016. A deep source-context feature for lexical selection in statistical machine translation. *Pattern Recognition Letters*, 75: 24–29.

Habash, N. & Hassan, H. 2012. Special issue on Machine Translation for Arabic: Preface. *Machine Translation,* 26(12): 1–2.

Helou, M. A., Palmonari, M. & Jarrar, M. 2016. Effectiveness of automatic translations for cross-lingual ontology mapping. *Journal of Artificial Intelligence Research* , 55: 165–208.

Hitzler, P., Krotzsch, M., Ehrig, M. & Sure, Y. 2005. What Is Ontology Merging?. *Web,* 1–4.

Hutchins, W. J. 1995. Machine Translation: A Brief History. *Concise history of the language sciences: From the Sumerians to the cognitivists*, 431–445.

Ishkewy, H., Harb, H. & Farahat, H. 2014. Azhary: An Arabic Lexical Ontology. *International journal of Web & Semantic Technology*, 5(4): 71–82.

Kadhim, K. A., Habeeb, L. S., Sapar, A. A., Hussin, Z. & Abdullah, M. M. R. T. L. 2013. An evaluation of online machine translation of arabic into english news headlines: Implications

on students' learning purposes. *Turkish Online Journal of Educational Technology*, 12(2): 39–50.

Kashgary, A. 2010. The paradox of translating the untranslatable: Equivalence vs. non-equivalence in translating from Arabic into English. *Journal of King Saud University – Languages and Translation*, 23 (1) : 47-57.

Klein, M. 2001. Combining and relating ontologies: an analysis of problems and solutions. *IJCAI'01 Workshop on Ontologies and Information Sharing,* 47(1): 53–62.

Lopez, A. 2012. Statistical Machine Translation. *ACM Computing Surveys,* 40(3): 1–447.

Maedche, A. & Staab, S. 2005. Ontology learning for the semantic web. *Intelligent Systems, IEEE*, 72–79.

Navigli, R. & Velardi, P. 2003. Ontology Learning and its Application to Automated Translation. *IEEE Intelligent Systems*, 22–31.

Noy, N. & McGuinness, D. 2001. Ontology development 101: A guide to creating your first ontology. *Development*, 32: 1–25.

Ram, S. & Park, J. 2004. Semantic Conflict Resolution Ontology (SCROL): An Ontology for Detecting and Resolving Data and Schema-Level Semantic Conflicts. *IEEE Transactions on Knowledge and Data Engineering*, 16(2): 189–202.

Ramar, K. & Gurunathan, G. 2016. Technical Review on Ontology Mapping Techniques. A*sian Journal of Information Technology*, 4(15): 676-688.

Rasiwasia, N. & Vasconcelos, N. 2012. Holistic context models for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 34(5): 902–917.

Reddy, M. V & Hanumanthappa, M. 2012. Dictionary Based Word Translation in CLIR Using Cohesion Method. *INDIACom,* 2–5.

Shvaiko, P. & Euzenat, J. 2013. Ontology Matching: State of the Art and Future Challenges. *IEEE Transactions on Knowledge and Data Engineering,* 25(1): 158–176.

Storey, V. C. 2005. Comparing relationships in conceptual modeling: Mapping to semantic classifications. *IEEE Transactions on Knowledge and Data Engineering*, 17(11): 1478–1489.

Tarazi, O. (n.d.). Allah's Words in Plain English. https://www.plainenglishquran.com/ [9 June 2016].

Wang, Jianqiang, Oard, D. W. 2011. Matching meaning for cross-language information retrieval. *Information Processing and Management*, 48(4): 631-653.

Yahya, Z., Abdullah, M. T., Azman, A. & Kadir, R. A. 2013. Query translation using concepts similarity based on Quran ontology for cross-language information retrieval. *Journal of Computer Science*, 9(7): 889–897.

Zadeh, H. Z. (n.d.). Tanzil.net. tanzil.net [9 June 2016].

Zailan Arabee Abdul Salam
Rabiah Abdul Kadir
Institute of Visual Informatics,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor
zailanz@gmail.com, rabiahivi@ukm.edu.my