Translation vs. Transliteration: Arabization in Scientific Texts

Grami Mohammad A Grami
King Abdulaziz University, Saudi Arabia
ggrami@kau.edu.sa

Abstract

This paper looks at the concepts of translation and transliteration in general and in scientific and academic texts in particular. In simple terms, the former refers to the process of finding equivalents in the target language (as opposed to the original language of the text), while the latter refers to writing the original word using the characters of the target language. The paper argues that translation works well in texts that explain, describe, detail, instruct and summarize while transliteration works better in concepts, processes, known procedures and proper nouns, to mention but a few. The paper suggests that the reliance on literal translation of terms and concepts can be counterproductive to the purpose of translation. Six computer science students were involved in a small-scale experiment. Tests were designed to determine which approach, Arabization or literal translation, is more efficient by measuring the time students took to complete certain tasks and whether students can trace the translated word back to its English origin. All participants were interviewed afterwards. Results showed that they preferred transliterated terms and that Arabic literal translation was not helpful. Results also showed that transliteration of scientific texts helped students understand faster and more accurately. The paper recommends a hybrid approach that employs both methods depending on what terms or processes are being translated.
Keywords*: Arabization, Literal Translation, Proper Nouns, Transliteration*

1. INTRODUCTION

A considerable amount of manpower, time and resources have been dedicated to translating texts from their original languages into other languages. In essence, this was attempted by translating existing books and articles with varying degrees of success. One persistent issue that crops up whenever attempting to translate an original text is what to do with terminologies and known phrases as well as processes and tools in more technical texts. Most translators opt for using a literal translation which does the job very well on the surface. However, this can be more often a well-meaning attempt to simplify something that does not require simplification in the first place.

One oft-cited project of translating terminologies into Arabic is teaching medicine in Arabic in Syria. The extra load of tens of thousands of new medical terms in Arabic along with their English/Latin counterparts can add to students' workloads for no apparent gain. The comprehensive translation is rarely used anywhere else, even in the Arab world, rendering it exclusive to the Syrian educational system (Al-Halawanii, Yani, & Kama, 2016; Sara, 2009)

The concept of translating knowledge in itself is very practical and useful in the sense that it helps students and learners save time and effort. This can be achieved as they are not constantly consulting dictionaries for new words. However, when the Arabic term used in place of its counterpart is so unfamiliar and opaque that students need explanation, then the entire reason for the translation in the first place is jeopardised.

Learners in the field in which the translations are attempted have to struggle with two sets of terminologies instead of one, namely the English terms in the field of study. Assuming that many students, especially in science, are very likely pursuing postgraduate degrees in English speaking countries or in universities where English is the lingua franca, being familiar with transliterated versions of the jargons and terms commonly used can be to their advantage.

The research question is: which approach produces better translations from English into Arabic that are more accurate, efficient and easier to understand when dealing with scientific terms in computer science, transliteration (Arabization) or literal translation?

To help answer the question two tests were designed which are a speed test and reverse translation. The former was developed to determine which approach generates texts in Arabic that are easier to understand by computer science students, and the latter helps determine the accuracy of the translation by matching the Arabic text to the original text in English. All participants were interviewed subsequently.

2. LITERATURE REVIEW

In this section, I attempt to identify the instances in which literal translation may not be as helpful as simply using the process of writing the sounds of the original word using the characters of the target language which is transliteration. The literature review also looks into the concepts of translation in general and the process of transliteration in more details. I will also attempt to identify the research gap and the rationale of the current study based on the findings of previous studies.

It must be noted that the focus here is on computer science texts because the participants of the study are all computer science students. This is not to assume that different majors should be treated differently but rather that certain scientific terms should be treated differently in translation regardless of the major. The concept here is that many scientific fields do contain certain terminologies, mostly from English origin, that are familiar with respective learners and experts in the field. Attempting to find equivalents in the students' first language (Arabic in this case) may not necessarily make texts easier for them.

2.1 Proper Nouns and Out of Vocabulary Words

Research shows that there are many instances in which literal translation may not be helpful when it comes to comprehension. In fact, the very reason underlying the concept of translation is to help readers understand the original text better and more efficiently. With this in mind, there are certain words that cannot or should not be translated, which are identified here.

Proper nouns are essentially references to unique entities and not a description of who/what they are or what they do. They are usually contrasted with common nouns which refer to a class of entities. Examples of proper nouns are Paris, Jupiter, John and April as compared to common nouns like city, planet, man and month. Some make a distinction between proper nouns and proper names where the former is limited to a singular form usually preceded by the article 'the' but this distinction is not always observed and nouns and names are used interchangeably. The term Out of Vocabulary - OOV - words can also be used to refer to these proper nouns, since it means words such as names of places and names of people, which are important content words in a sentence. The term OOV comes from speech recognition software that does not recognize proper nouns as part of the vocabulary. (AbdulJaleel & Larkey, 2003; Jespersen, 2013; Leech, 2006).

Proper nouns, therefore, should be transliterated not translated. Any attempt to find equivalents in the target language will simply yield awkward translations. My assumption here is that many scientific terms have moved from the common nouns category right into OOV and should be treated as such.

Since the focus here is on computer science, examples of these terms include names of products like Windows, Apple or processes like file dump and scanning.

I believe that the available literature does identify these terms but there is an apparent lack of understanding as to how to deal with them when it comes to translation. Very little is known of what readers expect from the translated text and what their preferences are. The study aims to fill this gap by directly studying the effects of transliteration and translation on students' comprehension and it also investigates which approach students prefer and why.

2.2 Transliteration and Arabization

Transliteration is the process of describing a word from a source language using the letters of the target language usually by converting the sounds of one script to its equivalent in the other (e.g. Cyrillic to Roman or Arabic to Thai). Transliteration is commonly used in the names of people and places. This is usually occurring when out of vocabulary (OOV) words are encountered and is very common with terms used in technology (Halpern, 2007).

According to Tsvetkov & Dyer (2015), transliterated words are among the four categories of vocabulary in a language and they call these words unassimilated. The remaining three being core words of the language, assimilated and semi-assimilated.

Arabization is the process of assigning words from other languages their equivalent phonemic Arabic characters and therefore it is essentially a form of transliteration. In this project, Arabization will exclusively refer to the process of transcribing English words using Arabic script. Words like Windows, modem, fax, tablet as well as names of people and places, are usually Arabized, not translated, and for a good reason as the paper attempts to argue.

Nonetheless, transliteration, especially Arabization, can be very challenging as Kharusi, Nafla, & Salman (2011) rightly observe. These difficulties stem from the significant phonological differences between Arabic and English including the absence of consonantal equivalents, the absence of letters representing short vowels in Arabic (diacritics are rarely used instead), different pronunciations of the same word due to different dialects, inconsistencies in orthographic representation of words, as well as a lack of a consistent universal rule governing the spelling of words, resulting in different variations.

Difficulties aside, the main reason behind transliteration is to give a fair representation of proper names in Arabic, one which allows for users to pronounce the word correctly without having to ponder its origin or meaning, which are irrelevant in this case. Transliteration makes words more recognisable regardless of language, which is helpful when users are not familiar with either the target or source language, as in the case of tourism.

Frequently used terms lose their descriptive qualities and become more abstract. They should be treated as proper nouns rather than lexical items. This calls for a different approach in translation since these terms have reached a level of abstraction beyond the realms of meanings (Longobardi, 1994; Manini, 1996; Pour, 2009; Zarei & Norouzi, 2014).

Translating English terms into other languages can be inconsistent and confusing. Examples, the operating system Windows © has been literally translated in Arabic textbooks to the words equivalent to its English counterpart. However, Macintosh Apple has not been literally translated at all although the equivalent of apple exists. This is inconsistent since both Windows and Apple are abstract brand names rather than lexical terms.

This transition towards more abstraction effectively turns these English words into proper nouns. If we continue with the Arabic examples, there are many biblical names which have equivalents in English (David, Joseph, Noah, Jesus, Mary ...). However, when referring to people whose names have equivalents in Arabic the standard practice is to use the original pronunciation of their names in their language of origin. A few exceptions may occur, mostly for humorous or comic purposes. By the same token, names of people that have lexical meaning (Shearer, Thatcher, Green ...) are never translated based on their meaning (Abdolmaleki, 2012; Homeidi, 2004; Kharusi et. al., 2011).

The literature is littered with examples of books translated into Arabic. Professional translators spent extensive time and effort creating new words in Arabic that a) do not sound similar to the word of origin as is the case in cognates, b) are completely unintelligible even

to experts in the field and c) will rarely, if ever, appear in other translations due to a lack of recognition among others with a similar linguistic and scientific background.

Finally, as observed in the previous section, there are many Out of Vocabulary OOV words in any language in the world and many researchers, including Tsvetkov & Dyer (2015), concede that transliteration should be the way to go when attempting to translate texts that contain these OOVs. The same can be applied to proper nouns according to Habash (2008) who also believes that transliteration is the proper course of action when it comes to translating texts containing such words.

I observed that the majority of recent research in the fields of OOV and transliteration (e.g; Chalabi, Morsy, Awadalla, El-Sharqwi, & Hassan, 2015; Habash, 2008; Tsvetkov & Dyer, 2015, to name a few) is more concerned with machine translation not the actual process of transliteration of OOV words. I struggled to find any recent study that investigates the impact of transliteration on the final product and readers' perceptions of the translated/transliterated text, which is important given the fact that the purpose of translating a text, automatically or otherwise, is to produce relevant texts in the potential readers' language that is easier to understand compared to the original text in the source language, usually English. I would therefore argue that the current study, albeit being limited and small-scale, is an attempt to fill a gap in the literature by involving the stakeholders in the processes of transliteration of OOVs.

3. RESEARCH METHODS

The small-scale study investigates whether transliteration of jargons helps students understand the translated text more than the other methods of translation.

3.1 Participants / Subject / Population and Sample

Six second- and third-year Computer Science students in two universities (both of which use translated textbooks in instructions) were included in the study. These students are referred to as A, B and C from King Abdulaziz University in Jeddah, Saudi Arabia and D, E and F from Umm Al Qura University in Makkah, Saudi Arabia.

No English proficiency tests were available but all students successfully passed the prerequisite General English courses before joining their respective departments. As teaching is conducted mainly in English by instructors in both universities who do not speak Arabic, it can be assumed that the participants' English is proficient at least in their computer science majors and the test that requires students to trace words back to their original English origins is viable.

3.2 Instruments

Two tests were designed to measure participants' attainments, preferences and expectations. Students were individually interviewed to reflect on their experiences and elaborate on their beliefs and preferences of translated texts.

The following tests were administered to all six students: Timed Speed Test: this test requires students to connect the definitions of certain processes and terms with both the translated words and the transliterated ones as quickly as possible. Reversed Translation Test: short sentences were given to see which method (translation vs. transliteration) retains

the original meaning of the translated text. Students were given one 50-word paragraph in Arabic and were asked to reverse translate it to see whether the literal translated texts or the transliterated ones better matched the original text.

Interviews: following the two previous tests, all participants were interviewed in their respective institutes. The interviews were semi-structured and mainly discussed the following questions: Which do you prefer and why? Are you willing to learn Arabic terminologies? How useful are they in the future?

3.3 Data Analysis

The quantitative data from the two tests were processed using SPSS. The descriptive data including means and standard deviation were generated to identify the points of difference between the two approaches in translation. The interviews were categorized and inferences were made when possible.

4. FINDINGS

The first test shows that all six students performed better in the speed test when connecting definitions to their transliterated terms rather than their literal translated counterparts. The following terms were used: file dump, scanning, the internet of things and format. It took all of the students less than one minute to connect to the transliterated texts while four students struggled with the terms 'file dump' and 'internet of things' citing the fact that the Arabic translations were misleading.

Table (1) Time in seconds to complete the test (Translation vs. Transliteration)

Student	Literal Translation	Transliteration
A	35	76
B	47	55
C	44	64
D	54	72
E	32	52
F	29	75
Avg	40	66

The averages show that students in the transliteration group took significantly less time to complete the task (40 seconds) compared to the other group which took significantly longer time (66 seconds on average) to recognize the terms literally translated into Arabic.

The second test does not show clear patterns since all students struggled with reverse translations. However, judging by the accuracy of the words used, it can be argued that transliterated texts showed better results since students were successful in writing the words correctly in English.

All six students were subsequently interviewed about their experience using a semi-structured approach and their responses were categorized as preferences, difficulties, expectations and recommendations.

5. DISCUSSION

It became evident that they overwhelmingly preferred the transliterated terms since they connected directly to the original text with which Computer Science students are already familiar. The majority of the interviewees did not approve of learning a parallel Arabic set of terms citing extra workload and effort that can be used in other areas instead. Finally, all students raised concerns about the usefulness of learning technical computing terms in Arabic.

Students A, B, and D stressed the fact that their studies in their respective universities are mainly in English and the majority of their instructors do not speak Arabic anyway. This favours transliteration since students are by default already familiar with the English words rather than their Arabic equivalent. Students A and E did point out that the Arabic translations are unheard of. They claim they cannot even find the meaning of some Arabic translations in dictionaries.

Attempting to find equivalents for each and every term in a given scientific text is neither realistic nor necessarily helpful to the target audience. One reason is the fact that specialists in a field build up a repertoire of terminologies over the years or over a course of study, regardless of their language of origin. Attempting to alter their "lingo" may result in unnecessary difficulties. Furthermore, creating unfamiliar translations can actually be counterproductive and serve the opposite purpose of the translation process in the first place.

The proposed alternative recognises the main purpose of translation, which is making the text more comprehensible to the potential reader. Yet it does not force awkward or unfamiliar terms in place of the original ones. A hybrid alternative seems to bring the best of both worlds, familiarity with common terms and ease of understanding the instructions.

Frequently used terms and even processes (scanning, presentation, memory dumping ...) in which the sound rather than lexical meaning is more familiar, should equally be transliterated not literally translated.

The reasonable compromise here is to stick to translations when there are no technical terms of processes and to use transliteration otherwise. The fact remains that the majority of the text is translatable and only a few words can be transliterated.

6. CONCLUSION

First of all, I would like to stress that this is a limited small-scale study. Therefore, any findings here should be treated as indicative not conclusive. The fact that the findings go in line with previous more established research is a plus point.

Although translation in itself should help potential readers access and navigate the text more easily, there are certain words that are better left alone or transliterated. These words are usually considered proper nouns in the sense that they have lost their literal meaning and rather became references to entities in their own right.

Despite the overwhelming evidence, even common sense, that proper nouns and names should not be translated but rather transliterated, there are still many examples to the contrary. It was found that students who naturally come across many proper nouns in their course of study do overwhelmingly prefer Arabization (transliteration) rather than translation. The fact that they are computer science students is significantly important

because of the wealth of new processes, procedures, terms and components that originated in English and have been translated in Arabic dictionaries.

This paper identified certain words in the field of computer sciences which demonstrate the concept of transliteration and why it works better than literal translation.

The two processes, transliteration and translation, work better together where the former is used in explaining and instructing students in their own language while the latter is used with common terms and processes. Transliteration also helps develop the repertoire of computer science students and allows them to become familiar with the terminologies of their field of study.

REFERENCES

Abdolmaleki, S. D. (2012). Proper Names in Translation: An Explanatory Attempt. *The Social Sciences*, *7*(6), 832–837.

AbdulJaleel, N., & Larkey, L. S. (2003). Statistical Transliteration for English-Arabic Cross Language Information Retrieval. *Proceedings of the Twelfth International Conference on Information and Knowledge Management ACM*, 139–146.

Al-Halawanii, A., Yani, A., & Kama, N. (2016). Problems Encountered in Translating Oxymora from English into Arabic 16, 2 – 15. *KONGSI*, *16*, 2–15.

Chalabi, A., Morsy, A. S., Awadalla, H., El-Sharqwi, M., & Hassan, S. (2015). *U.S. Patent No. 8,990,066*. Washington, DC: U.S. Patent and Trademark Office.

Habash, N. (2008). Four techniques for online handling of out-of-vocabulary words in Arabic-English statistical machine translation. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies*, 57–60.

Halpern, J. (2007). *The Challenges and Pitfalls of Arabic Romanization and Arabization*. Proc. Workshop on Comp. Approaches to Arabic Script-based Lang.

Homeidi, M. A. (2004). Arabic Translation Across Cultures. *Babel*, *50*(1), 13–27.

Jespersen, O. (2013). *The Philosophy of Grammar*. London: Routledge.

Kharusi, K., Nafla, N., & Salman, A. (2011). The English Transliteration of Place Names in Oman. *Journal of Academic and Applied Studies*, *1*(3), 1–27.

Leech, G. (2006). *A Glossary of English Grammar. Edinburgh University Press.* Edinburgh: Edinburgh University Press.

Longobardi, G. (1994). Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. *Linguistic Inquiry*, *25*(4), 609–665.

Manini, L. (1996). Meaningful literary names: Their forms and functions, and their translation. *The Translator*, *2*(2), 161–178.

Pour, B. S. (2009). How to translate personal names. *Translation Journal*, *13*(4), 1–13.

Sara, K. (2009). Teaching English at Damascus University Medical School. *Eastern Mediterranean Health Journal*, *16*, 653–664. Malaysia.

Tsvetkov, Y., & Dyer, C. (2015). Lexicon Stratification for Translating Out-of-Vocabulary Words. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, *2*, 125–331.

Zarei, R., & Norouzi, S. (2014). Proper nouns in translation: Should they be translated? *International Journal of Applied Linguistics and English Literature*, *3*(6), 152–161.