

# Reflection of Explicitation in Scientific Translation: Neural Machine Translation vs. Human Post-Editing

Ogareet Y. Khoury

Department of English and Translation, Al-Ahliyya Amman University, Amman, Jordan

**Abstract**—The present paper reports on the findings of an empirical comparative study on the extent to which explicitation is employed in the translation of a scientific text as conducted by Google Neural Machine Translation (GNMT) vs its post-edited (PE) version. A recent report released in English by the World Meteorological Organization in September 2023 was selected as the source text for the present study. The purpose of the study is to reveal how domain-specific acronyms and technical terms are lexically expanded (explicitated) in a GNMT output compared to its post-edited (PE) version as performed by a team of professional translators at a translation service provider in Amman-Jordan. Explicitation in translation can be obligatory or optional. The type of explicitation investigated in the present study is optional, pragmatic explicitation. The results show that GNMT has its limitations in dealing with scientific terms and acronyms in translating scientific texts from English into Arabic. In contrast, human post-editing explicitated domain-specific terms and acronyms producing a text with a higher level of readability and naturalness for domain expert readers and non-expert readers.

**Index Terms**—acronym, explicitation, GNMT, post-editing, technical translation

## I. INTRODUCTION

### A. *Explicitation in Translation Studies*

Explicitation has been defined as a translation strategy that makes the message in the target text clearer and more explicit than it is in the source text (Shuttleworth & Cowie, 1997; Hatim & Mason, 2004). The first theorists to introduce explicitation as a translation shift were Vinay and Darbelnet (1958) defining it as ‘a stylistic translation technique which makes what is implicit in the source language explicit in the target language’ (p. 342). It was argued that explicitating the source text (ST) contributes to the cohesion and coherence of the TT especially with regard to system-bound elements (e.g. Al-Dabbagh, 2008).

Features of explicitation include but are not limited to lexical expansion; i.e. inserting new meaningful elements in the TT that are absent in the ST or lexical specification in which a SL unit is replaced with a TT unit with a more specific meaning (Klaudy & Karoly, 2005, p. 15). Features of explicitation were also identified as a reflection of any textual segment that is encoded explicitly in the translation (Carston, 2002). The common ground among all definitions is that it tends to contribute to disambiguation, cohesion and coherence of the TT.

Explicitation is considered obligatory when it is driven by structural mismatches between the two languages and optional when it is employed as a strategy to enhance the readability and naturalness of the TT (Klaudy, 1998, p. 83). In other words, optional explicitation is associated with stylistic differences and preferences. A sub-category of optional explicitation is pragmatic explicitation which seeks to explicitate system-bound elements or culture-bound elements to overcome the differences in the world or cultural knowledge between recipients of source and target texts (Englund-Dimitrova, 2005). The present study investigates optional, pragmatic explicitation of system-bound elements in a scientific report.

### B. *Explicitation in Technical / Scientific Translation*

Translation of technical texts was defined as translation of texts written for special purposes including domains such as medicine, engineering, law, science among others (Wright & Wright, 1993). The fact that technical texts have their own jargons and specialized terminology makes their translation challenging (Muhiesen & Al-Ajrami, 2019). Thus, explicitation by lexical expansion becomes a necessity in specialized translation which seeks to expand domain-specific units with condensed information including highly technical terms, abbreviations and acronyms (Baker, 1992; Herget & Alegre, 2022). Acronyms were defined as abbreviations formed from the initial letter or group of letters of two or more words that are pronounced as words rather than separate initials (Pyles & Algeo, 1964). According to Al-Hamly and Farghal (2012), English opts for lexical reduction through abbreviations and acronyms for producing shorter discourse. In technical texts, acronyms are used to promote a sense of professionalism in domain-specific texts based on the assumption that it is written by domain experts to be read by experts (Barnett & Doubleday, 2020). However, when readership includes non-experts, translators may resort to lexical expansion of highly technical terms and acronyms to

simplify the text and enhance its naturalness. Some scholars even stressed the need for acronym expansion even for domain-expert readers (Hales et al., 2017; Thomas, 2021). Such a practice is more essential in case of translation from English into Arabic. While English tends to resort to acronyms especially in specialized domains (Al-Hamly & Farghal, 2012; Ulitkin et al., 2020), Arabic-coined acronyms are rarely found. Most acronyms used in Arabic discourse are either borrowed or transliterated from English (Kamil, 2020; Ameer & Altaie, 2009). Expanding a given acronym in the ST into its full form in the TT is a form of explicitation (Pakhomov et al., 2005; Anbaqi, 2009; Kasprowicz, 2010; Gumul, 2017; Abdelaal, 2019).

### C. Neural Machine Translation

Machine translation (MT) is the application of computer software or program to the task of translating a given text from one language into another. Examples of MT tools include *Google Translate (GT)*, *DeepL*, *Microsoft Translator*, *Bing Translator*, *Reverso* among others. Such software programs are built by encoding morphological, syntactic and lexical-semantic structures of SL and TL allowing a statistic-based mapping to match segments of the two languages (Hartley, 2009, p. 121). Based on the initial MT design, two approaches of machine translation emerged; a rule-based machine translation (RBMT) and statistical machine translation (SMT). While the former is based on linguistic knowledge retrieved from grammars and dictionaries, the latter which superseded the former used unit alignments from parallel corpora to overcome language irregularities (Koehn et al., 2017).

In 2015, Neural Machine Translation (NMT) was launched benefiting from artificial neural network for higher rate of accuracy (He et al., 2015). As a technological breakthrough in machine translation, it superseded the previous approaches of RBMT and SMT (Polakova & Klimova, 2023). In principle, NMT conducts end-to-end translation with SL encoder and TL decoder in which words are integrated or incorporated within their corresponding words. This means that languages are represented by an embedding vector and clustered in the embedding space (Tan et al., 2019). By way of exemplification, iron, copper and gold are related semantically and usually co-occur within certain contexts and therefore are grouped together in metal-related clusters. This means that whenever the word *iron* occurs within this cluster, it vectors NMT to translate it as a type of metal rather than a home appliance used for smoothing clothes. Word-embedding was originally drawn from semantic theories that can process the semantics of a given text in an effective way very close to human processors and creates a better text-context interaction (Koehn et al., 2017). Thus, NMT architecture was meant to enhance text-context interaction; an advancement which was reflected in GNMT. The present study uses (GNMT) output along its human PE version for investigating and analyzing reflections of pragmatic explicitation.

### D. Explicitation and Neural Machine Translation

Neural Machine Translation (NMT) is widely used in the translation industry to save effort and time (Koneru et al., 2023). In reflection to the market demands, machine translation and computer-assisted translation (CAT) as well as post-editing modules have been integrated into the study plans of translator training programs all over the world to prepare students for a competitive market (O'Brien et al., 2021). Subsequently, research on the efficiency of NMT vs human translation and the gaps that human translation can fill through post-editing have gained momentum.

There has been research on the development of NMT, its methods and efficiency (e.g. Tan et al., 2020). Several recent studies tended to compare between the quality of NMT vs human translation or post-editing (e.g. Khoury, 2024) in Arabic-English translation. To the best knowledge of the author, no studies have been conducted to compare explicitation in NMT vs. human translation between English and Arabic; a gap which the present study fills.

## II. LITERATURE REVIEW

Explicitation, as a translation shift received due attention in theoretical and empirical translation studies (Murtisari, 2016). Blum-Kulka (1986) studied the concept intensively through corpus-based comparative study of translations between French and English. Her study revealed that a remarkable addition of cohesive markers was noted as a recurrent translation shift in the TL from French into English and vice versa. As a result, Blum-Kulka (1986) hypothesized that explicitation is a translation universal.

Based on previous research, scholars conducted empirical corpus-based analysis to test the validity of the hypothesis (e.g. Shuttleworth & Cowie, 1997; Englund Dimitrova, 2005; Abdel Hafiz & Sokarno, 2004; Al-Khafaji, 2005). Most studies supported the notion of the universality of explicitation, especially in translation of professionals. Subsequent studies tended to compare reflection of explicitation in the translations of professionals vs. non-professionals revealing that non-professionals also resort to explicitation (e.g. Vesterager, 2017; Blum-Kulka, 2000; Al-Dabbagh, 2008).

Studies were conducted on different language pairs and text domains for reflection of explicitation. Studies included languages such as French, Danish, Hungarian, Norwegian, Portuguese, Arabic and English among others. Text domains selected for studies on explicitation included literary, legal, scientific and cultural discourse that are labelled as technical texts requiring specialized translation (Gotti & Sarsevic, 2006; Olohan, 2015). Translation of technical texts were found to reflect higher percentage of explicitation due to inclusion of system-bound terms that need to be explicitated for non-expert readers (De Beaugrande & Dressler, 1981; Dirchalidis & Patsala, 2018).

In the context of translation between English and Arabic, Baker (1992) investigated explicitation in translation of literary and culture-bound texts. Hjort-Pedersen and Faber (2009) investigated how explicitation is employed by trainee translators versus experienced translators in translation of technical/legal discourse between Danish and English. Vesterager (2017) conducted an empirical study on how explicitation is employed by professional vs. non-professional translators in legal translation and revealed that professional translators resort to explicitation more than trainees.

Most studies concluded that explicitation is a translation universal in translations of professionals and non-professionals with higher percentage in translations of professionals indicating that it is a practice that reflects expertise in translation. In his descriptive translation studies, Toury (1995) found a direct correlation between explicitation and readability of the ST and therefore producing a better quality TT. For Hatim and Munday (2004), explicitation clarifies the intended message of the source text which contributes to a more natural TT. Alzu'bi (2021) concluded that explicitation is a technique that adapts the TT and improves it. Most studies conducted on explicitation came to the conclusion that explicitation plays an essential role in inter-lingual and intercultural communication (Murtisari, 2016).

Explicitation techniques -whether the explicitation was obligatory or optional- were found to be reflected at four levels; syntactic, lexical, pragmatic and textual (Al-Anqabi, 2009). For instance, adding cohesive devices is deemed as a syntactic explicitation while adding footnotes and expanding condensed information (e.g. highly technical terms, acronyms and abbreviations) are considered lexical expansion (Baker, 1992; Shuttleworth & Cowie, 1997; Blum-Kulka, 2000).

Including neural machine translation in translation studies for quality assessment was reflected in several studies. The study carried out by Ahrenberg (2017) compared NMT with human translation in the use of connective devices revealing that human translation produces more cohesive texts. Some other studies compared performance of trainees vs. professional translators in post-editing GNMT outputs (e.g. Khoury et al., 2024) which revealed that experience in the job market can remarkably affect the quality of post-editing. Temizöz (2013) showed that experts of the subject matter post-edited technical terms in a better way than experienced translators. Béz (2018) revealed that while students of modern languages focus on syntactic post-editing, translation students post-edit all aspects of the source text.

The only study that compared features of explicitation in neural machine translation vs. human translation was conducted by Krüger (2020). It investigated explicitation in the translation of scientific text from English into German and concluded that human translation conducts more explicitation than NMT and thus produces a more natural text. Since explicitation indicates sensitivity to text-context interaction, the results of Krüger's study (ibid) served as a significant indicator of NMT underdeveloped performance in relation to explicitation as a translation universal. The present study contributes to the literature by providing the first study to investigate explicitation in NMT vs. human post-editing between English and Arabic.

### III. SIGNIFICANCE OF THE STUDY

The significance of the study lies in more than one aspect. Firstly, it contributes to comparative studies between raw NMT outputs and the PE version; a research scope that is still at its infancy. Secondly, it investigates explicitation which was not tackled in such comparative studies -to the best knowledge of the author- except for the study by Krüger (2020). Thirdly, such comparative studies have not been conducted in English-Arabic translations yet. The present study fills the gap by investigating reflection of explicitation in scientific translation as performed by GNMT vs human post-editing between English and Arabic.

### IV. STATEMENT OF THE PROBLEM AND RESEARCH QUESTIONS

As the industry today relies on NMT (O'Brien, 2021), machine translation needs to be part of translation empirical studies in terms of proficiency. As discussed in section 1, explicitation was found to be one indicator of professional translation. By comparing NMT performance with the post-editing of experienced translators, important implications can be driven at two levels: 1) the proficiency of raw NMT without explicitation and 2) what needs to be post-edited in the output to increase its proficiency. Therefore, the research questions for the present study are stated as follows:

- 1- What is the difference between GNMT and its PE version in expanding/ explicating condensed information units (domain-specific units and acronyms)?
- 2- What needs to be post-edited in a GNMT output in terms of expansion/ explicitation?

### V. METHODS AND MATERIALS

#### A. Source and Target Texts

A 48-page report titled '*United in Science 2023: Sustainable Development*' released by the World Meteorological Organization in September, 2023 was selected as the source text for the present study (found at <https://public.wmo.int/en/resources/united-in-science-2023>). The ST belongs to technical/ scientific domains or what is called specialized translation (e.g. Olohan, 2015).

The ST is an authentic text that was requested for translation into Arabic by a research development institute in Jordan. The translation was performed by a group of five professional translators in a translation service provider in

October, 2023. According to the translation service provider, the translation project was carried out by using raw GNMT output after which the output was post-edited by the five translators.

### B. Data Collection and Tools

For the purpose of the study, the only aspects that were investigated in the PE version were related to reflections of pragmatic explicitation in the form of lexical expansion of domain-specific expressions and acronyms before and after post-editing. Incidents of explicitation in the two outputs were analyzed qualitatively and quantitatively. The post-edited version was conducted by five experienced translators at a translation service provider as part of a translation project presented to a research development institute. Thus, the comparative study was built on the raw GNMT vs the PE version as provided by the translation service provider.

## VI. RESULTS AND ANALYSIS

In the following section, results and analysis are presented quantitatively and qualitatively. All domain-specific acronyms and terms were traced in the 48-page report. The report was found to use 55 domain-specific acronyms and 9 highly technical terms. Selection of domain-specific terms was based on their reflection at the UN digital library (<https://digitallibrary.un.org/record/452305?ln=en>). Sections 7.1 and 7.2 below present and discuss the translation of GNMT before and after post-editing in relation to reflections of lexical expansion as a form of pragmatic explicitation.

### A. Acronyms Processing in GNMT vs. PE Version

As mentioned previously, there are several techniques that lie within explicitation as a translation shift. One of which is lexical expansion (Klaudy, 2000). The source text used acronyms intensively which is typical of English scientific texts (Ammar et al., 2011; Al Hamly & Farghal, 2012).

The repetition rate of acronyms used in the report ranged between twice such as CORDEX (Coordinated Regional Climate Downscaling Experiment) and 185 times such as SDG (Sustainable Development Goals). In compliance with the norms of technical writing, acronyms are normally written in full form on first use (Mark, 2012). With reference to the ST, expansion of acronyms was provided at first mentioning; after which acronyms were used solely without their expansions which is typical of English technical writing as mentioned above.

#### (a). GNMT of Acronyms

The translation processing of acronyms in GNMT was found to lie within four main categories; 1) translating the full form of the acronym in Arabic; i.e. expansion in the target language, 2) borrowing the acronym as is in the SL alphabets without expansion, 3) transliterating the acronym into a TL-alphabet acronym without expansion. In addition, it was noticed that some acronyms that stand for more than one multi-word unit were inappropriately expanded; i.e. mis-explicated for the domain. Examples on each of the four categories are discussed below.

The analysis of GNMT reveals that 15 out of the 55 acronyms were properly expanded or explicitated into their TL full forms. This is equal to 27% of the total acronyms used in the ST. The rest were kept as acronyms either in the SL (borrowing) or TL alphabets (transliteration). Acronyms which were fully expanded into their Arabic full form in each mentioning throughout the report include IPCC, ILO, SDG which stand for *Intergovernmental Panel on Climate Change*, *International Labour Organization* and *Sustainable Development Goals* respectively.

On the other hand, acronyms such as EWSA (*Early warning for South Africa*), SE (*Safe Energy*), OSISAF (*Ocean and Sea Ice Satellite Application Facility*), SAA (*Vulnerability Adaptation Assessment*) were transferred into English acronyms in the Arabic TT without any lexical expansion. Apart from the fact that non-expert readers may not be familiar with a given acronym, such code-switching from Arabic into English in the TT is expected to break the flow of the Arabic discourse and thus block its comprehension.

The same resulted in transliterating the English acronym into Arabic alphabets which yielded a more puzzling TT since Arabic does not normally use acronyms in texts originally written in Arabic (Ameer & Altaie, 2009; Kamil, 2020). Transliterating acronyms into Arabic included units such *جي سي اف و جي جي دي دبليو وجو اون* as translation of GT (Gas Turbine), GCF (Green House Fund), GGGW (Global Green House Watch, GOAON (Global Ocean Acidification Observing Network) respectively.

The fourth category included acronyms that were mis-explicated since they stand for more than one multi-unit form. These were also found as disguised mistranslation since the explicitation was irrelevant to the domain. By way of exemplification, GNMT rendered the English acronyms IFRC (International Conference on Regional Climate), IOC (Intergovernmental Oceanographic Commission) AI (Artificial Intelligence), UC (Urban Climate) EEI (Earth Energy Imbalance) and FSS (Farmer Field Schools) into *الاتحاد الدولي، اللجنة الدولية الاولمبية، منظمة العفو الدولية، جامعة كاليفورنيا، مبادرة التعليم المصرية، خدمة الساتلايت الثابتة*. The back translations of the these full-form units are: International federal for IFRC, International Olympic Committee for IOC, Amnesty International for AI, University of California for UC, Egyptian Education Initiative for EEI and Fixed Satellite Service for FSS which were expanded out of their contextual meaning. Thus, the Arabic multi-word units given as explicitated segments to the English acronyms were semantically irrelevant to the subject matter of the scientific report. The fact that the ST provided the full form of acronyms on first use provides evidence of the intended meaning of the acronym. Such processing of acronyms by GNMT resulted into

further confusion to the Arabic reader since the acronym was translated properly on first mentioning and mis-explicitated in subsequent uses. In addition, some acronyms were completely mistranslated as they were confused in GNMT with English words. This includes REAP and WISER which stand for *Resource and Energy Analysis Programme* and *Weather Information Services* respectively were translated as words into *يجني و اكثر حكمة* in the meaning of *to reap a harvest* and *to have more wisdom*.

The analysis presented above on how GNMT architecture dealt with acronyms reveals more than one issue to be considered. Firstly, GNMT is trained on expanding acronyms from English into Arabic as it translated a percentage of 27% of the ST acronyms. However, it failed in expanding 73% of them with exception of the first mentioning. In addition, the analysis shows that GNMT is still under-developed to meet the contextual meaning in relation to acronyms. This is reflected in more than one example such as expanding IOC as International Olympic Committee instead of Intergovernmental Oceanographic Commission in a meteorology-specific domain.

Transferring acronyms from the ST into acronyms in the TT by borrowing or transliterating poses a cohesion-coherence problem for the Arabic TT. According to norms of translating acronyms from Arabic into English, only common English acronyms are borrowed into Arabic in their SL alphabets or transliterated into Arabic such as *اليونيسكو*, corresponding to UNESCO, FAO and NASA. Non-common acronyms are normally expanded in Arabic followed by the acronym in English between brackets (Al-Jarf, 2021). Evidence to this practice is supported by referring to parallel Arabic texts found at [www.un.org](http://www.un.org) and <https://arabic.euronews.com>.

#### (b). Post-Edited Version of Acronyms

The PE version of GNMT as conducted by the translators' team reflects different handling of expansion of acronyms. The 55 acronyms were expanded into their full forms in Arabic in every use of the acronym. The method of post-editing employed by the translators is expansion to the full form in Arabic followed by the English borrowed acronym between brackets without any transliteration of the acronym; i.e. English acronyms were not written in Arabic-alphabet acronyms, which was found to be consistent throughout the report. In addition, several acronyms were further explicitated within the expanded form above the lexical items for which they stand. For instance, acronyms such as GCF for Green Climate Fund was expanded and explicitated into *صندوق دعم التغيرات المناخية التابع للأمم المتحدة* (Back Translation: Green Climate Fund **of the United Nations**) in which UN reflects further lexical explicitation above the acronym initials.

With reference to Arabic parallel texts on climate found at [www.un.org](http://www.un.org), it was found that while all acronyms are expanded into their full form in Arabic, there is no clear unified methodology or norm in dealing with the expansion process. While some organization names are translated literally such as Green Climate Fund into *الصندوق الأخضر للمناخ*, other acronyms are expanded into a form that reflects the function such as Greenhouse Gases as *الغازات الدفيئة*. This implies that the PE version in the present study is reader-oriented that seeks naturalness of the translated text for expert and non-expert readers. Such feature is still far-fetched in GNMT and inconsistent in human translation available in parallel texts.

#### B. GNMT vs PE Version of Domain-Specific Terms

Domain-specific terms were defined as having significant meaning in a particular domain or belonging to a specific domain as part of its jargon (Kim & Cavedon, 2011) constituting a part of domain specific languages (DSL) in which the terminology is custom-tailored for a specialized domain (Pech, 2021). Selection and categorization of terms that are domain-specific was based on *the glossary of scientific and technical meteorological terminology, English-Arabic*, available at the UN digital library (<https://digitallibrary.un.org/record/452305?ln=en>). The scientific report reflected nine climate-specific terms which were cross-checked in the glossary. These included proper names of projects or software applications, e.g. Climpact and Femlink and other terms related to climate phenomena such as El Niño and eutrophication. Table 1 shows the translation provided by GNMT of these terms while Table 2 reflects the PE version of the same terms. Discussion follows the display of the Arabic translation of GNMT and human translation along with their back translations in Table 1 and 2.

TABLE 1  
GNMT TRANSLATION OF DOMAIN-SPECIFIC TERMS

	ST term	GNMT	Back Translation
1-	Anthropogenic Activities	نشاط بشري	Human Activities
2-	Climate Extremes	الظروف المناخية المتطرفة	Extreme climate conditions
3-	Climpact	Climpact	Climpact
4-	Cyclone Freddy	اعصار فريدي	Freddy Hurricane/ cyclone
5-	El Niño	النينيو	El Niño
6-	El Niña	النينيا	El Niña
7-	Eutrophication	التخثث	Eutrophication
8-	FemLINK	فيملينك	FemLINK
9-	Nowcasting	التنبؤ الانمي	Immediate forecasting

TABLE 2  
HUMAN TRANSLATION OF DOMAIN-SPECIFIC TERMS

	ST term	Human Translation	Back Translation
1-	Anthropogenic Activities	النشاط البشري المضر بالبيئة	Human Activities that harm the environment
2-	Climate Extremes	الظروف المناخية الحادة	Acute climate conditions
3-	Climpact	برنامج Climpact وهو برنامج لحساب المؤشرات المناخية ذات الصلة بقطاعات الصحة والزراعة والمياه	Climpact: a software which calculates climate indices related to sectors of health, agriculture and water
4-	Cyclone Freddy	اعصار فريدي وهو اعصار شديد ضرب جنوب المحيط الهندي في شباط عام 2023	Freddy cyclone: a severe cyclone whit hit southern Indian ocean during February 2023
5-	El Ni ño	النينيو: ظاهرة مناخية تحدث تحدث عندما يكون هناك انخفاض في كمية المياه الباردة التي ترتفع إلى سطح البحر بالقرب من أمريكا الجنوبية	El Ni ño: a climate phenomenon which occurs when there is a decrease in the amount of cold water that elevates to the surface in areas close to South America
6-	El Ni ña	النينيا: ظاهرة مناخية تحدث نتيجة برودة غير عادية لسطح المياه في المنطقة الاستوائية في المحيط الهادي	El Ni ña: a climate phenomenon characterized by unusual coldness of surface water of the Pacific Ocean
7-	Eutrophication	اثراء غذائي في المسطحات المائية	Richness of nutrients in water bodies
8-	FemLINK	فيملينك Femlink وهو مشروع يسعى لحماية المرأة (المهمشة) في الظروف المناخية الحادة	FemLINK: a project which aims at protecting overlooked women in extreme/ acute climate conditions
9-	Nowcasting	التنبؤ الاالي: تنبؤ للمستقبل القريب يتراوح بين بضع دقائق الى ستة ساعات	Nowcasting: near future forecasting covering a period of time between few minutes up to six hours.

Table 1 reflects the Arabic translations as provided by GNMT. It reflects two types of domain-specific terms; common climate-related terms and names of climatic phenomena, software applications and institutions. Common nouns were translated by GNMT without further explicitation. These include ‘anthropogenic activity’ and ‘nowcasting’ which were translated literally without any explicitation. Names of entities or phenomena such as El Ni ño, Cyclone Freddy, Climpact or Femlink were unrecognized by GNMT where some were transliterated and others were borrowed as acronyms in English alphabets without explicitation. This reveals that in addition to acronyms, NMT reflects drawbacks in recognizing some proper nouns (Hassan et al., 2018; L äubli et al., 2020). It may be argued that NMT are not trained to deal with proper nouns, however, some studies suggested injecting feature vectors into factored-transformer model to improve NMT of proper nouns (Takai et al., 2021).

With reference to human post-editing in Table 2, it can be noticed that proper common nouns were not only translated but further explicitated. For example, ‘anthropogenic activities’ was translated into النشاط البشري المضر بالبيئة which is back-translated into ‘human activities that harm the environment’ explicitating the negative impact of the activities. The same can be seen in the translation of ‘nowcasting’ that was explicitated in the PE version as future casting covering any period between few minutes up to six hours.

Interestingly, Arabic parallel texts found at <https://www.un.org/ar/chronicle/article/20126>, <https://www.un.org/ar/climatechange/reports>, <https://archive.unescwa.org/ar> and <https://news.un.org/ar/story/2023/03/1118817> revealed that domain-specific technical terms are translated into domain-specific technical terms without further explicitation (e.g. eutrophication: التخصث). Post-editing in the present study exerted further effort in explicitating domain-specific terms (Table 2) apparently to guarantee comprehension to domain-expert and non-expert readers.

Similar to the post-editing of acronyms, scientific terms were all further explicitated by the post-editors which does not comply with what is reflected in Arabic parallel texts. Therefore, data analysis provided in sections 7.1 and 7.2 reveal that human post-editing seemed to expand and explicitate both acronyms and highly technical terms to provide readability and naturalness to both expert and non-expert readers remarkably above the GNMT output and slightly above translations in parallel texts.

Based on the notion that explicitation contributes to a more readable and natural TT (e.g. Klaudy, 1998), it can be argued that the PE version in the present study rendered the text into a readable and more natural text for expert and non-expert readers. Therefore, the first research question stated in section 5 is answered as follows: PE version of GNMT output reflected explicitation of acronyms and domain-specific terms. Acronyms were fully expanded and further explicitated above the full form in some examples and domain-specific terms were explicitated by adding some explanatory lexical items to provide readability and naturalness to non-expert readers. Thus, the second research question is answered by building on the findings to suggest a thorough post-editing process for scientific texts from English into Arabic.

## VII. CONCLUSION AND RECOMMENDATION

Recent research emphasizes that encoder-decoder architecture in NMT has been designed to utilize recurrent neural networks that capture the meaning in relation to the context to perform better contextual meaning (Gao et al., 2022).

Despite the argument that NMT systems are booming in the translation industry, their development is still hampered by the lack of quality input, as well as the insufficient means of testing the quality of the translations (Shoshan, 2018). Based on this, the present study investigated the MT output vs its PE version to provide the feedback needed on the quality of the MT raw output.

The results of the present study show that vectors are not yet designed to deal with acronyms between English and Arabic in a systematic pattern and still need to be embedded in a similar pattern to recognize the semantic relationship between units. As for domain-specific terms, MT output was able to capture the meaning translating it into a domain-specific term, however, without any further explicitation which raises a question of whether there is a loss of meaning for non-expert readers.

The fact that GNMT expanded some acronyms and struggled with others reveals that the neural network through its encoder-decoder models are being fed with some Arabic-English acronyms but still lags behind when compared to human translation or post-editing. Apart from acronyms, GNMT is trained to deal with scientific terms from English into Arabic, however, the output remains confined to domain expert readers. In conclusion the Arabic-English encoder-decoder architecture in GNMT needs post-editing in relation to acronyms and domain-specific terms for scientific texts if the target readership includes expert and non-expert readers.

#### REFERENCES

- [1] Abdelaal, N. M. (2019). Subtitling of culture-bound terms: Strategies and quality assessment. *Heliyon*, 5(4), 1-27, e01411.
- [2] Abdel-Hafiz, Ahmad-Sokarno. (2004). *Lexical cohesion in the translated novels of Naguib Mahfouz: the evidence from 'The Thief and the Dogs'*. <http://www.arabicwata.org>, Retrieved on April 2, 2023.
- [3] Ahrenerg, L. (2017). Comparing machine translation and human translation: A case study. In *Proceedings of the Workshop Human-Informed Translation and Interpreting Technology*, pages 21–28, Varna, Bulgaria. Association for Computational Linguistics, Shoumen, Bulgaria.
- [4] Al-Anbaqi, N. (2009). The Role of Explicitation Strategies in Arabic. Doctoral dissertation, *Middle East University For Graduate Studies*, MEU Library Theses. [https://meu.edu.jo/libraryTheses/586cbc4960045\\_1.pdf](https://meu.edu.jo/libraryTheses/586cbc4960045_1.pdf). Retrieved on December 13, 2023.
- [5] Al-Dabbagh, U. (2008). The explicitation hypothesis and non-professional translators: A case study. *International Journal of Arabic-English Studies*, 9(1), 85-106.
- [6] Al-Hamly, M. and Farghal, M. (2012). English Reduced Forms in Arabic Scientific Translation: A Case Study. *Jordan Journal of Modern Languages and Literature*, 5(1), 1-18.
- [7] Al-Jarf, Reima. (2021). Derivation from foreign words and acronyms borrowed in Arabic. *Lingua. Language and Culture*, No 1/2021 Year Xix New Series, pp. 52-77.
- [8] Al-Khafaji, Rasoul. (2005). Translation-specific language universals: The case of Arabic translated texts. *Paper presented to the Atlas Fifth International Conference on Current Issues in Translation and Arabization*, 18-19, December, 2005.
- [9] Alzu'bi, K. (2021). Some Effects of Randomly Practiced Translation Strategies on Arabic. *Theory and Practice in Language Studies*, Vol. 13(7), pp. 1777-1784.
- [10] Ammar, W., Darwish, K., El Kahki, A., Hafez, K. (2011). ICE-TEA: In-context expansion and translation of English abbreviations. In Gelbukh, A. (Eds.) *Computational Linguistics and Intelligent Text Processing*. CICLing 2011. Lecture Notes in Computer Science, Vol 6609.
- [11] Ameer, A. and Altaie, A. (2009). *Errors made by the students at the departments of English and Arabic in acronyms: A contrastive study*. Babylon University Available at: <chrome-extension://efaidnbnmnbbpcjpcglclefindmkaj/https://cdnx.uobabylon.edu.iq/research/Acronyms%20in%20E%20&%20A.pdf>. Retrieved on 5 November, 2023.
- [12] B éz, M. (2018). Machine translation and post-editing: impact of training and directionality on quality and productivity. *Revista Tradum àtica*, 16(1), 24-34.
- [13] Baker, M. (1992). *In Other Words: A Coursebook on Translation*. London: Routledge.
- [14] Barnett, A. and Doubleday, Z. (2020). *The growth of acronyms in the scientific literature*. ELife. DOI: 10.7555/eLife.60080
- [15] Beaugrande, R. de and Dressler, W. (1981). *Introduction to Text Linguistics*. Longman.
- [16] Blum-Kulka, S. (2000). Shifts of cohesion and coherence in translation. In L. Venuti (Ed.), *The translation studies reader*, pp. 298-312. London & New York: Routledge.
- [17] Carston, Robyn. (2002). *Thoughts and utterances: the pragmatics of explicit communication*. 2nd ed. Oxford: Blackwell.
- [18] Dirchalidis, M. and Patsala, P. (2018). What makes a legal text 'legal? Implications arising from a corpus-based study. *Terminolox ú: A necessidade da colaboraci ón*, In Manuel Gonz áez Gonz áez, Mar ía-Dolores Sánchez-Palomino and In é Veiga Mateos, pp. 181-202. <https://doi.org/10.31819/9783964568090-009>
- [19] Englund Dimitrova, B. (2005). *Expertise and explicitation in the translation process*. Amsterdam & Philadelphia: John Benjamins Publishing.
- [20] Gao, R., Ponnuthurai, L., Suganthan, N., Zhou, Q., Yuen, K. (2022). Random vector functional link neural network based ensemble deep learning for short-term load forecasting. *Expert Systems with Applications*, 206(1), 1-12.
- [21] Gumul, E. (2017). Explicitation and directionality in simultaneous interpreting. *Linguistica Silesiana*, 38: 311–329.
- [22] Gotti, M. and Sarcevic, S. (2006). *Insights into Specialized Translation*. Peterlang.
- [23] Hales, A., Williams, K. and Rector, J. (2017). *Alienating the Audience: How Abbreviations Hamper Scientific Communication*. Available at: <https://www.psychologicalscience.org/observer/alienating-the-audience-how-abbreviations-hamper-scientific-communication>. Retrieved on November, 05, 2023.
- [24] Hartley, T. (2009). Technology and translation. In J. Munday (Ed.), *The Routledge Companion to Translation Studies*. Routledge: New York, pp. 106-140.

- [25] Hatim, B., & Mason, I. (2004). *Discourse and the Translator*. Pearson Education.
- [26] He, K., Xiangyu, Z., Shaoqing, R., and Sun., J. (2015). *Deep Residual Learning for Image*. Recognition. Available at: <http://arxiv.org/abs/1512.03385>. Retrieved on 29 October, 2023.
- [27] Herget, K. and Alegre, T. (2022). Analysis of explicitation evidence in technical translation. *Eighth International Conference on Higher Education Advances*, 29-42.
- [28] Hjort-Pedersen, Mette, and Faber, Dorrit. (2010). "Explicitation and Implication in Legal Translation. A Process Study of Trainee Translators". *Meta: Translators' Journal*, 55(2), 237-250.
- [29] Kamil, D. (2020). The problems of translating acronyms from English into Arabic. *Adab Al-Rafidayn*, 71(1), 182-208.
- [30] Kasprovicz, M. (2010). Handling abbreviations and acronyms in medical translation. *Translation Journal*, 14(2). <https://benjamins.com/online/etsb/publications/19949>. Retrieved on November 19, 2023.
- [31] Khoury, O., Alkhwaja, L., Ghnaim, F., Awwad, S. and Dudeen, H. (2024). Post-editing a google translated Output: Experienced translators vs. trainees. *Theory and Practice in Language Studies* (In Press).
- [32] Klaudy, K. (1998). Explicitation. In M. Baker & G. Saldhana (Eds.), *Encyclopaedia of Translation Studies* (pp. 80-85). London: Routledge.
- [33] Klaudy, K. (2000). Explicitation Strategies within Lexical and Grammatical Translational Operations. In: Lendvai, E. 2000. *Applied Russian Studies in Hungary*. Pécs: Krónika Kiadó 101-113.
- [34] Klaudy, K., & Károly, K. (2005). Implication in translation: Empirical evidence for operational asymmetry in translation. *Across Languages and Cultures*, 6(1), 13-28.
- [35] Krüger, R. (2020). Explicitation in neural machine translation. *Across Languages and Cultures*, 21(2), 195-216.
- [36] Koehn, P., Och, F., Marcu, D. (2003). Statistical phrase-based translation. In *proceedings of NAACL 2003*, pp. 48–55.
- [37] Koehn, P. (2017). *Neural Machine Translation*. Baltimore: Johns Hopkins University.
- [38] Koneru, S., Huck, M., Exel, M. and Niehues, J. (2023). Analyzing Challenges in Neural Machine Translation for Software Localization. Conference: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*. Kujamäki (Eds), Translation universals: Do they exist? (pp. 143-64). Amsterdam & Philadelphia: John Benjamins.
- [39] Muhiesen, E. and Al-Ajrani, M. (2019). Challenges in translating scientific texts from Arabic into English: a case study of Al-Muhiesen's the Acheulian culture in Jordan. *Dirasat: Human and Social Sciences*, 46(1), 322-328.
- [40] Murtisari, E. (2016). Explicitation in translation studies: The journey of an elusive concept. *The International Journal for Translation and Interpreting*, 8(2), 63-81.
- [41] Thomas, G. (2021). "Punctuation marks and abbreviations". In *Research Methodology and Scientific Writing*, edited by C. G. Thomas, pp. 455–73. Cham: Springer International Publishing.
- [42] O'Brien, Sharon. and Alessandra, Rossetti. (2021). Neural machine translation and evolution of the localisation sector, Implications for training. *Journal of Internationalization and Localization*, 7:1/2, 95-121.
- [43] Olohan, M. (2015). *Scientific and Technical Translation*. London and New York: Routledge.
- [44] Pakhomov, S., Pedersen T., Chute, CG. (2005). Abbreviation and acronym disambiguation in clinical discourse. *AMIA Annual Symposium Proc. 2005*; 2005: 589-93. PMID: 16779108; PMCID: PMC1560669.
- [45] Polakova, P., Klimova, B. (2023). Using DeepL translator in learning English as an applied foreign language – An empirical pilot study. *Heliyon*, 9(2023), e18595.
- [46] Pyles, T., & Algeo, J. (1964). *The Origins and Development of the English Language*. 1st Edition. New York: Harcourt, Brace & World, Inc.
- [47] Shoshan, O. (2018). *Make neural machine translation better, faster – a new way to measure NMT quality*. Available at <https://slator.com/make-neural-machine-translation-better-faster-a-new-way-to-measure-nmt-quality/>. Retrieved on 7 December, 2023.
- [48] Shuttleworth, M., & Cowie, M. (1997). *Dictionary of Translation Studies*. Oxford: Oxford University Press.
- [49] Takai, K., Hattori, G., Yoneyama, A., Yasuda, K., Sudoh, K. and Nakamura, S. (2021). Named entity-factored transformer for proper noun translation. *Proceedings of the 18<sup>th</sup> International Conference on Neural Language Processing*, pages 7-11. Silchar-India, December 16-19, 2021.
- [50] Tan, Z., Zhang, J., Huang, X., Chen, G., Wang, S., Sun, M., Luan, H. and Liu, Y. (2020). THUMT: An Open-source toolkit for neural machine translation. *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas Volume 1: Research Track*, 116-122.
- [51] Temizöz, Ö. (2013). *Postediting Machine Translation Output and Its Revision: Subject-Matter Expert Experts Versus Professional Translators*. MA Thesis. Universitat Rovira I Virgil. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-19437-5\\_4](https://doi.org/10.1007/978-3-642-19437-5_4)
- [52] Ulitkin, I., Filipova, I., Ivanova, N. and Babaev, Y. (2020). Use and translation of abbreviations and acronyms in scientific texts. *Innovative Technologies in Science and Education*, 210(1), 1-12.
- [53] Vesterager, A. (2017). Explicitation in legal translation – a study of Spanish –into-Danish translation of judgments. *The Journal of Specialized Translation*, 27(1), 104-123.
- [54] Vinay, J. P., & Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais*. [Title translation: Comparative Stylistics of French and English]. Paris: Didier.
- [55] Wright, W., & Wright, M. (1993). *Scientific and Technical Translation*. London: Routledge.

**Ogareet Y. Khoury** is an Associate Professor at the Department of Translation and the Vice-Dean of the Faculty of Graduate Studies at Al-Ahliyya Amman University (AAU). She holds a PhD Degree in Translation Studies from Aston University-UK with a 20+ years of teaching experience. Her research interests include, but are not limited to, *Translator Pedagogy and Training*, *Discourse Analysis*, *Literary Criticism* and *Philosophy* in which she published several research papers at top-ranked journals.