Ambiguity Resolution in Arabic Localization

Abdulfattah Omar, Iman El-Nabawi Abdel Wahed Shaalan, Wafya Ibrahim Hamouda

1 Department of English, College of Sciences & Humanities, Prince Sattam Bin Abdulaziz University.
2 Department of English, Faculty of Arts, Port Said University.
3 Department of English, College of Sciences and Humanities, Prince Sattam bin Abdulaziz University.
2 College of Humanities, Al-Azhar University, Egypt.
3 Department of Foreign Languages, Faculty of Education, Tanta University, Tanta, Egypt.

Corresponding Author: Abdulfattah Omar
e-mail: a.a.omar2010@gmail.com


ABSTRACT

Numerous studies have evaluated digital localization quality in different languages, including Chinese, German, Russian, and Spanish. Consensus has largely been reached that language clarity and accessibility, as well as linguistic consistency, are key factors for the success and reliability of the localization of software applications and websites. Although there is a sizable body of literature on the evaluation of localization quality, there is very little on Arabic localization and accessibility to Arab users. This study seeks to address this gap in the literature through an evaluation of the Arabic localization of learning management systems (LMSs). The development of digital transformation processes, and recently the outbreak of COVID-19, have resulted in LMSs being integrated in teaching and learning practices. It is thus important to evaluate their translation quality and linguistic consistency, these being essential for their success and reliability. The study conducted a survey to explore users’ perceptions concerning the translated content of LMSs. The results indicate that ambiguity and language inconsistency are the main problems with different LMSs, including Blackboard, Zoom, and Microsoft Teams. This is attributed to different reasons, including the peculiar linguistic features of Arabic, ineffective management of terminology, and a lack of consideration of dialectal differences and variations. Localization developers thus need to address the linguistic challenges and peculiarities of Arabic, including its morphology and writing system. Dialectal variations also need to be considered. The contents of an LMS should be adapted so that they are attractive to target users and easy to use. In this regard, localization developers should recruit the right people for language and translation tasks and jobs. Translators also need to realize the differences between localization and translation. The language should be accessible and easy to understand. Finally, the management of terminology and standardization should be considered.

Keywords: ambiguity, Arabic, COVID 19, digital transformation, linguistic consistency, localization, learning management systems (LMSs), software applications, translation

1. Introduction

In recent years, the development of software and web-based applications has progressed in an unprecedented manner in different fields, including communication, education, health, and marketing. The reasons for this are various, including globalization, digital transformation, and the emergence of COVID-19. The development of localization systems and services has occurred in parallel with the development of such applications. Although localization is essentially an additional investment that software developers and content owners make to increase the accessibility of their products and their commercial value, it is also considered one of the key factors for the success and reliability of applications and software products.
In education, various applications have been developed to address the increasing needs of users, educational institutions, and training centers around the world. These are usually released in different languages, employing localization systems so that speakers of different languages can use them easily and effectively. For the proper use of these applications, the quality of localization, and language clarity and consistency should be considered. Although there is a considerable body of literature on the assessment and evaluation of localization in different languages, including Spanish and Chinese, very little work has been done on the Arabic localization of technological applications in general and learning management systems (LMSs) in particular. In the face of this limitation, this study is concerned with evaluating the quality of localization of LMSs in Arabic.

The rest of this article is organized as follows. Part 2 defines the research problem. Part 3 provides a brief survey of localization evaluation and assessment systems. Part 4 describes the methods and procedures of the study. Part 5 reports the findings and part 6 concludes.

2. Statement of the Problem

Recent years have witnessed the development of several LMSs, including Blackboard Learn, Zoom, eFront, moodle, and Microsoft Teams, aimed at addressing the changing needs of learners, as well as those of learning and training institutions (Boldyk, 2020; Dias, Diniz, & Hadjileontiadis, 2013; Foreman, 2017; Keleş, 2016). These institutions have realized the importance of integrating LMSs into their teaching practice to improve learning and sustain a more knowledgeable and productive learning environment. In terms of a definition:

An LMS is a multiuser software application, usually accessed through a web browser. It helps organizations manage training events, self-paced courses, and blended learning programs. It provides automation that replaces rigorous and expensive manual work, saves time, and enables you to organize your data, and learners. It tracks and reports on training activity and results (Foreman, 2017, p. 9).

The global changes being witnessed today, and the importance of enforcing digital transformation processes in the education and training sectors pose many challenges for e-learning, the localization industry, and research in making online learning systems available to all users around the world (Auer & Tsiatsos, 2019; Omar & Almaghithwi, 2020). The recent shift in almost all educational and training institutions worldwide toward online learning due to COVID-19 has made it important for the manufacturers and designers of online learning systems to translate their global purpose into localized efforts. They need to adapt to the linguistic needs of users around the world. Localization developers thus need to provide consistent translations of the content of such systems into target languages so that users get the maximum benefits.

Although many LMSs are now being released in different languages, including Arabic, to ensure that they are delivered in the same way to all users globally, regardless of linguistic differences, many users still have difficulties dealing with the systems in various target languages. There is general dissatisfaction among users with many software applications and platforms, which contain large amounts of information and lengthy instructions that lack clarity and consistency. For the successful implementation of LMSs, language issues must be addressed and handled. In other words, localization quality is one of the main requirements for the localization process of an LMS. Reliable localization should be based on unambiguous and understandable language, the appropriate language level, standardization of terminology, provision of sufficient context to the translators, and validation of the target text (Börjel, 2007; Dunne, 2006; Esselink, 2000).

In light of this argument, this study seeks to evaluate the language quality of LMSs in Arabic. Although other issues, for example technical and commercial, are essential factors in localization, this study is limited to linguistic aspects. In other words, this study is only concerned with exploring how the Arabic translations of LMSs generated by localization systems are perceived by Arab users. The study carried out a survey to map users’ perceptions concerning the language quality of the Arabic versions of the LMSs used. For the sake of representativeness and reliability concerning the data collected and analyzed, users from different Arab countries, age groups, and social backgrounds were targeted.

3. Literature Review

In recent years, localization has been the focus of various disciplines, including natural language processing (NLP), language engineering, and computational linguistics (Kley, 2020; Mao & Fidan, 2009). This can be attributed to the increasing developments in the localization industry and research. Surveys indicate that users generally prefer to use software applications and websites in their native languages in spite of the dominance of English as a global language (Abbatbain & Khan, 2017; Casalegno, 2018; Reinhardt, 2019; Schrepp, Hinderkers, & Thomaschewski, 2017). In response to these requirements, localization has greatly evolved and has become a necessity rather than a luxury (de-la-Cova, 2016; Olvera-Lobo & Rodriguez, 2019; Zybato, 2009).

Generally speaking, localization is a process whereby a software application or website in one language is adapted to the needs and outlook of another. It can be defined as the translation and adaptation of software or a web product, which includes the software application itself and all related product documentation. It is derived from the
word “locale,” which in a technical context represents a specific combination of language, region, and character encoding (Esselink, 2000).

Localization is always associated with translation and the terms are even used interchangeably in some contexts. However, localization is not just a translation of a source version (document, web page, or multimedia) into a target language. In his analysis of the differences between localization and translation, Dunne argues that localization is not just about translation. It has a broader sense. Localization is more concerned with the adaptation of content to regional specificity, including consideration of cultural and functional requirements. In other words, localization is primarily marked by technical and commercial dynamics (Dunne, 2006).

Commenting on the technical dynamics of localization, Al-Mazrooa considers the scope of localization and translation to be completely different. In localization, for instance, the scope of a single product often encompasses a variety of text types: general texts, which present the product’s subject matter; legal texts, which present the company’s terms and conditions; and possibly, technical texts, which present messages from the operating system (Al-Mazrooa, 2018). As for the commercial dynamics, cultural and religious sensitivity should be considered in localization. In contrast, translators are normally asked to convey the meanings in the source language in translation, and may face restrictions in terms of the content of the material they are willing to translate (Cermák & Smutný, 2018; Keniston, 1997).

Failing to consider the sensitive nature of localization can have negative impacts on the accessibility, marketing, and value of software products. In this regard, software developers usually adapt the content of their products (including language, fashion, dress, and food) to enable better communication with other cultures. A recent example can be seen in the release of Nintendo’s game Paper Mario: The Origami King in July 2020. The localization team sparked a dispute about how it translated sensitive language from Japanese to Traditional Chinese (Pyner, 2020). Due to pressures from the Chinese government, the two Japanese words for “human rights,“ 人权, and “freedom,” 自由 were translated into Chinese as meaning “plain outlook” and a “peaceful life”.

In response to the increasing popularity of localization systems and services, a number of studies have been undertaken to evaluate the accuracy of localization systems in different languages, including Chinese, Hindi, Russian, and Spanish (Chao, Singh, & Chen, 2012; Chessa & Knauth, 2012; Esselink, 2000; Jimenez-Crespo, 2013; Mateo, 2014). These have mainly been concerned with measuring the quality and accuracy of the localization systems through comparing the source and target segments (Poirier, 2014). Despite the existence of prolific literature on localization evaluation and assessment, research on Arabic localization in general and the Arabic localization of LMSs in particular is very sparse. Indeed, to our knowledge, the only work evaluating Arabic localization of an LMS is Al-Mazrooa’s attempt to explore Arabic translation practices under the influence of the technical and commercial dynamics of localization (Al-Mazrooa, 2018). The study addressed the commercial and technical aspects of Arabic localization through three case studies exploring users’ perspectives on and attitudes toward the accessibility and use of the FIFA 15 video game, Blackboard Learn, and the Knorr website. Notwithstanding the contribution of the study to the Arabic localization literature, the results cannot be generalized to all LMSs as it was based only on Blackboard Learn and conducted solely in Saudi universities. In light of this limitation, this study seeks to bridge the gap in the literature through evaluating the localization quality of several LMSs in Arabic.

4. Methods and Procedures

For the purposes of the study, a questionnaire was administered. Participants were asked about their familiarity with LMSs and whether they used the Arabic versions of these systems. They were also asked about their experiences with the Arabic localization of these systems and whether they faced any problems in dealing with the Arabic version.

In total, 843 participants from 13 Arab countries, including Egypt, Libya, Saudi Arabia, Tunisia, and the United Arab Emirates, responded to the survey. The participants represented different age categories, social backgrounds, and nationalities. The survey targeted pre-university students, university students, teachers, and university faculty. The study included students, teachers, information technology (IT) employees, technicians, and faculty in both public and private institutions. As the study was essentially concerned with the Arabic localization of LMSs, it was appropriate to target only those involved in academic and learning contexts.

The responses indicate that more than 90% of the participants used LMSs. The majority had been LMS users for 1–2 years; only 15% had been using an LMS for more than two years. The most common LMSs used by the participants were Blackboard Learn, Zoom, eFront, moodle, and Microsoft Teams. The majority of the respondents indicated that they used just one LMS, normally the one imposed by their institutions. Only 32% indicated that they used two or more different LMSs.

Interestingly, 75% of the participants preferred to use the LMS in Arabic. However, around 27% indicated that they usually switched to the English version when having difficulties understanding the content in Arabic. Around 89% stressed that they faced problems with the use of the Arabic localization of the LMS. They attributed these problems mainly to ambiguity and language inconsistencies. They indicated that the terminology used in these
systems was in many ways ambiguous and not clear. Finally, the majority of respondents stressed that the dialect variations and peculiarities were not addressed in the localization systems.

5. Analysis and Discussion

The results indicate clearly that the participants were generally familiar with LMSs and the Arabic localization of these systems. This can be attributed to the adoption of different digital transformation processes over recent years by many educational institutions. The emergence of COVID-19, which has led to drastic changes in teaching paradigms, can also be suggested as one of the main reasons for the popularity of LMSs in Arab countries. In almost all Arab countries, LMSs have been integrated in the learning operations of schools and universities over the last two years.

The results agree with the bulk of the localization literature, finding that users prefer to use software applications and systems in their own languages. Although it was generally the case that the majority of the participants preferred using the LMSs in Arabic, the ratio differed significantly based on the country. Moreover, most reported that they faced linguistic challenges with these systems. Based on the participants’ perceptions, it is clear that ambiguity and language inconsistency are two main challenges in the Arabic versions of the LMSs. In many cases, the content is not clear linguistically, even though all the localizations are based on Modern Standard Arabic (MSA), which should be familiar in academic contexts.

According to Lobanov, ambiguity and inconsistencies have negative impacts on localization quality. He stresses that localizers have to put quality first. Lobanov defines localization quality as a state of the translated text, such that the language is correct, accurately reflects the idea of the original, takes into account cultural specifics, and is easy for the target audience to read and use (Lobanov, 2014). In other words, localization quality is achieved when the translated text reveals the idea of the original accurately in the target language (Lobanov & Hill, 2017).

In the case of the Arabic versions of Blackboard Learn, Zoom, eFront, moodle, and Microsoft Teams, however, there is a general problem with their understandability. For many users, the meanings and terminology are not clear. Indeed, most of the participants stressed that the localized version was not easy to use. This can be attributed to the peculiar linguistic system of Arabic. The distinctive morphological system, as well as the bi-directional, cursive, and context-dependent writing system, make it difficult for localizers to adapt their localizations to be suitable for the requirements of Arabic (Farghaly & Shaalan, 2009; Guellil, Saadâne, Azouaou, Gueni, & Nouvel, 2019; Omar & Aldawsari, 2020; Omar & Hamouda, 2020).

Another reason reported was the lack of consideration of dialectal variations. Although the text and content of the LMSs had been translated into MSA, which is the official language of the Arab world and used across all Arab countries (Abdelali, 2004; Holes & Allen, 2004), it was still difficult for many users to understand the meaning of the translated content. This can be attributed to the fact that technological terminology is usually translated differently in Arab countries. In spite of the efforts of the Arab linguists and translators to provide a standard Arabization of IT that can be used in all Arab countries for easy and effective communication between speakers from different Arab countries, there are different Arabic equivalents for the same English or French terms in the IT and telecommunications fields (Abdellatif, 2016; Hamed, 2016). Regional specificity is a critical factor that must be considered for successful, high-quality, and reliable localization.

A third reason is the lack of direct equivalents in translating IT terms into Arabic. In general, there are very few technological and scientific terms in Arabic compared to other languages, such as English and French (Argeg, 2015; Elsherif & Soomro, 2017; Soudi, Farghaly, Neumann, & Zbib, 2012). In this regard, it can be noted that in localization, many translators encounter the problem of multiple Arabic translations for the same term due to the multiplicity of translation schools and accordingly, differences in dictionaries. Such lack of agreement adds to the problems already evident in the linguistic environment, with differences in writing styles in terms of spelling and semantics, all the way down to the shape of letters and punctuation marks.

Arabic is rich in expressions that are similar, indeed seem synonymous. However, there are in fact subtle semantic differences between these terms (Benamoun & Bassiouney, 2017; Bloch, 1991; Glanville, 2018). Accordingly, placing more than one Arabic word against an English word does not mean that the use of any of them is equally valid, but rather that the meaning of the English word is used in more than one context, and each context has a different Arabic interpretation. Translators must note such differences and carefully take them into account when choosing one of the apparently synonymous terms. In localization, these differences should be considered.

Localization developers thus need to address the linguistic challenges and peculiarities of Arabic, including its morphology and writing system. Dialectal variations also need to be considered. The contents of the LMS should be adapted to be attractive and easy to use for target users. In this regard, localization developers should recruit the right people to language and translation tasks and jobs. Translators also need to realize the differences between localization and translation. The language should be accessible and easy to understand.

Finally, some sort of coordination is required between Arab language societies and departments of telecommunications and IT to address the problems related to the translation and Arabization of technological terms.
Efforts should be exerted to Arabize telecommunications and IT terms by developing an Arabic glossary in these fields. This can be considered an effective strategy for addressing the problem of contradictions and discrepancies in the Arabic translations of technological terminology. It is apparent that an English or French term can be translated differently into more than five Arabic equivalents. This problem of lack of consistency poses serious challenges for terminology management in the localization industry. The problem also has negative impacts on the spread of technological communication, information, and education in this field, especially as the issue of localization is taking on a greater significance than ever before: It has become linked to the economy and development at a time when the world is moving toward a knowledge economy, in the growth of which technology will play a major role (Hines, 2013; Singh, 2012).

6. Conclusion

This study has addressed the issue of localization quality in LMSs through investigating whether the language content and translated texts are clearly conveyed in Arabic. Localization quality is a key factor for the reliability, and successful and effective use of LMSs. This requires that language content, including words, terms and phrases, is consistently converted into Arabic. The findings of this study agree with the bulk of the localization literature, in that most users generally prefer to use software applications in their native languages. Indeed, in Arab countries, users still prefer to use software applications and LMSs in Arabic despite the dominance of English as a global language. The results clearly indicate that ambiguity and linguistic inconsistency are serious problems that have negative impacts on the localization quality of LMSs, and thus have adverse impacts on the effectiveness and reliability of such systems. It is suggested that localizers pay attention to the peculiar linguistic features of Arabic and develop effective strategies for terminology management. They should also recruit the right people for language and translation tasks and jobs. Finally, it is proposed that Arab language societies develop specialized glossaries that address the gap between Western and Arabic terminology in science and technology.

Acknowledgments

This publication was supported by the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University.

References

Argeg, G. (2015). The problems of translating medical terms from English into Arabic. (PhD PhD thesis), Durham University, Durham, UK.


