Internationalization Testing for a Multilingual Based Wireless Response System: A Case Study

Abduladim Ali
School of Computing and Engineering
University of Huddersfield
Huddersfield, United Kingdom
Email: Abuladim.ali@gmail.Com

Joan Lu
School of Computing and Engineering
University of Huddersfield
Huddersfield, United Kingdom
Email: j.lu@hud.ac.uk

***Abstract*— Recently, wireless technologies have had a tremendous influence on various advanced technology areas through e-learning environments such as the use of smart devices to engage learners and encourage them to interact effectively. Wireless communication technologies are being widely adopted within the education sector in order to deliver the best education support both virtually and globally. Today, introducing a multilingual Wireless Response System (WRS) application is not only an enormous challenge but also a complex one. The aim of this paper is to implement an internationalization-testing strategy using a WRS case study for multilingual speakers of Arabic, Chinese, French, German, Romanian, Spanish and Turkish. It also aims to evaluate the internationalization testing results as well as localization and cultural testing impacts. This paper also attempts to identify the various challenges that are associated with internationalization and localization testing based on smart devices, as well as introducing a globalization testing model for a multilingual system that includes each language's specific features, for instance: the direction of writing for some languages such as right language (Arabic and Urdu), and word spaces translating.**

Keywords*— *Mobile learning; WRS; Internationalization; Globalization Testing; Usability testing.

I. INTRODUCTION

In recent years, there has been continuous growth of mobile applications in various areas, from leisure to providing business needs. This is due to the tremendous growth in the diversity of smart devices. Therefore, advanced mobile application response systems have grown rapidly and this has led to intense attention on these applications by researchers and developers. In fact, various advanced technologies have been introduced for the users to facilitate and improve the effectiveness of teaching and learning [1]. The extensive everyday use of these applications has led to higher demands when it comes to quality and stability. Specifically, users of smart devices are more interested in utilizing easily accessible applications. These applications are required to be more scalable and to support various languages that is, they are supposed to be multilingual. According to John [2] fixing internationalization bugs costs 30 times more than other issues.

Mobile Learning (M-Learning) has had a significant impact on teaching and education sectors where M-learning is defined as learning across multiple contexts, through social and content interaction, using personal electronic devices [4]. One such benefit has been that students are more attracted to expanding the boundaries of higher education into an "anytime/anywhere" experience [3]. Therefore, a WRS in M-learning is necessary for students and teachers to interact with each other. The WRS as an M-learning application based on multilingual use has been developed by the XML Database and Information Retrieval (XDIR) research group at the University of Huddersfield in the UK [4]. WRS enables students and teachers to collaborate with each other by means of smart devices.

The WRS like many M-learning applications is based on advanced technology in order to operate on smart devices across a wide range of platforms [5] including: PCs, laptops, tablets, and smart phones.

This paper attempts to highlight several challenges , which are associated with the internationalization testing of mobile applications. This paper will also describe the effectiveness of the test strategy to evaluate the localization process testing with the help of a WRS mobile app case study. The internationalization testing was conducted for multilingual use in Arabic, Chinese, French, German, Romanian, Spanish and Turkish.

Though sample sizes larger than 30 and less than 500 are appropriate for most research studies [6], throughout this study a relatively small sample size has been considered as this parameter is not as significant compared with the respondents' interface and interaction with the investigated system. Thus, when a user with a particular language employs this system, he/she may find that customized text space is insufficient; this would of course be the result for all respondents in this language [7].

This paper is organized as follows: Section 2 presents the internationalization background of mobile applications. Section 3 presents works related to globalization testing. Section 4 presents the experimental methods of WRS. Section 5 presents the discussion and evaluation of the experiment result with statistical results for each language. Finally, conclusions and suggestions for future work are given.

 ISBN: 978-1-61208-479-4

II. BACKGROUND

This section highlights issues such as the state of worldwide mobile applications including the internationalization testing process, internationalization testing challenges, globalization testing, the definition of GUI testing, and cultural testing.

A. The state of worldwide mobile applications

According to Sandrini [8] and Buck et al. [9], in 2013 alone, twenty one billion mobile apps were downloaded globally. The profits through the sale of mobile applications are expected to be more than $330 billion by 2015 [10]. Therefore, every year the usage of mobile applications is rapidly increasing. Lu [11] stated that more than 15% of global internet traffic is expected to be from mobile users by 2015 as compared to desktop users. Hence, wireless smart devices play a key role in accessing a wide range of services and becoming an essential part of users' daily lives for banking, shopping, healthcare, education and games.

B. Globalization Testing

Internationalization and localization are subsets of globalization (abbreviated g11n). Globalization testing is the process of assessing the mobile application in order to identify the application's performance within specific culture and language input and output conditions [4]. As depicted in Figure 1, globalization has been categorized into three phases including: feature to be tested, localization testing, and local awareness testing [12].

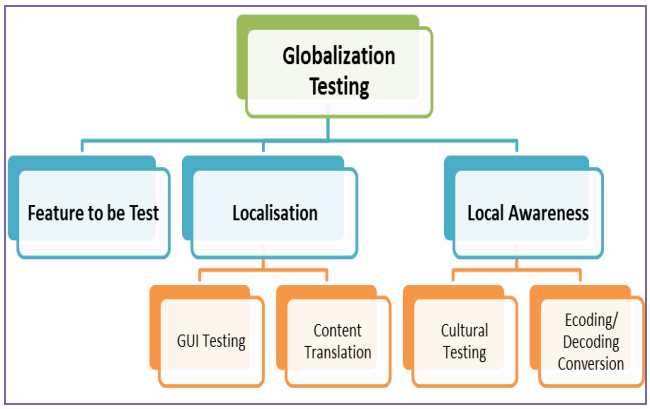

Figure 1. Globalization Testing Sub-tree

a) Internationalization Testing

Internationalization (abbreviated "i18n") organizes smart apps to make them easier for localization. Internationalization testing on the other hand is the process of evaluation for a single country, local, language, or culture [2].

i. Internationalization Testing Challenges

The internationalization testing challenges consist of issues and concerns related to the native users in different countries. Theses are in terms of language and culture specifications as depicted in table 1.

TABLE I. GLOBALIZATION TESTING ATTRIBUTES AND CHALLENGES.

Attributes	Challenges
Screen Size Limitation	Flexible layout to avoid expansion and contraction of text.
Screen Direction Layout	(RTL) and vertical text direction e.g. for Hebrew, Arabic, or Persian
Platform	Different OS IOS and Android
Context and Special Characters	The characters, symbols, labels, icons and buttons from language to language e.g., English to Chinese.
Data Handling	How the app handles data, warring messages and navigation
Regional Standards	Date, month and time e.g. Arabic or USA format (١٢ كانون الثاني ٢٠١٤).
Text Translation and Abbreviation	Space handling, e.g. content translated into German can take up to 30% more space than English
Culture Specification	Colour and image according to culture specifications.

ii. Internationalization Testing Process

The internationalization testing process consists of globalization testing with feature based testing, localization testing and local awareness testing as depicted in Figure 1.

iii. Feature Based Testing

Feature based testing is a process used to identify features of a product available in specific languages that have different content as compared to other languages and cultures. In fact, it is important to ensure that these features are turned on or off as required while switching locales [13].

b) Localization Testing

Localization testing is language verification testing, which is mainly focused on the appropriateness of the translation [14]. It is a process of asset verification within the application's user interface for the specific language contents and for the functionality state of the application, including: keyboard and mouse events, different GUI components e.g., menu bars, toolbars, dialogs, buttons, edit fields, and rich text [15]. In fact, 80% to 90% of localization issues are based on text translation quality issues, for example grammar mistakes, spellings, syntax, corrupted text, wrongly translated text, wrong language, and missing audio/text [16]. Moreover, linguistic accuracy is essential in globalization testing in order to identify those issues that have been introduced during the localization process.

c) Local Awareness

Local awareness testing is the process of detecting the translating formats of specific languages and cultures, for instance Arabic formats from right to left (e.g, ٩،٨،٧،٦،٥،٤،٣،٢،١،٠), as well as date, month, year, and time (e.g, ١٢ كانون الثاني ٢٠١٤). In fact, it is essential for each global mobile application to consider storing and retrieving documents containing locale-sensitive data such as changing the logic of all the formatting functions such as: date, time,

 ISBN: 978-1-61208-479-4

currency, numeric, etc. It must also follow the language specification formats in terms of text, number and image formats [17].

III. RELATED WORK

Many studies have examined several factors of internationalization testing such John [2], Meng and Lu [3], Lu [5], and Sekaran [6] though each of the studies has it is own contributions and limitations. Since 2007 when the first mobile application was released in app stores [6], countless technologies have been introduced to facilitate and improve the effectiveness of wireless response systems. Therefore, the internationalization of M-learning apps is essential for students and teachers to collaborate with each other, and to remotely enhance students' performance. Simkova et al. cited in Bastien [18] emphasized that M-Learning is certainly an interesting approach to learning, but is unfortunately still an unknown concept in many countries where it has not yet caught on.

Web course (WebCT) is an M-learning application which was criticized by Haller [19] due to the complexity of the application, which was contributed to by its complex architecture. Another example of an app that is challenging to use is the IELTS exam. This app has several individual applications as depicted in Figure 2, which consist of: the IELTS Reading app, IELTS Writing app, IELTS Speaking app and IELTS Grammar app. However, the individual apps are not useable by students. Because the apps are not all integrated into one application, students have to download each part of the English exam individually. In fact, these apps are suffering from a lack of a responsive approach.

Figure 2. IELTS Exam Mobile Application adopted from [7]

Moreover, Dhingra, cited in Haller [19] discussed various challenges that are associated with the localization and testing of mobile applications. He also described several elements and recommended that each element should be equipped with an effective test strategy for localization testing. He also stated that mobile localization should not just include translation, but should also include the user interface, and an adaptation of graphics based on locale or culture differences. [20].

IV. WRS METHOD OF EXPERIMENT

The proposed method for this research is based on the WRS case study and the eight identified attributes of internationalization testing requirements. It also takes into consideration several challenges of each attribute within each individual language that have been utilized in the WRS application. Translating the terms of the WRS application was based on an independent professional translating firm as professional translators.

The translated terms of every language were applied to localise the application by using Adobe Flex Builder 4.6 for the teacher phase, and Adobe Action Script® for the student phase. The WRS app was programmed in PHP and MySQL was used to retrieve the data.

For each language, there were 10 participants who carried out the globalization testing to evaluate the application based on their language attributes. The WRS testing was divided into two phases: student testing was the first phase, and the second phase assessed the teacher interface. Then the teacher and students interfaces were evaluated to understand how they experienced this application.

V. DISCUSSION AND EVALUATION

Internationalization testing for the WRS was carried out for seven languages: Arabic, Chinese, French, German, Romanian, Spanish, and Turkish. The calculation method for each language was: a number of questions were asked and the participants' answers were based on five evaluation matrices: excellent, poor, above average, below average, and average.

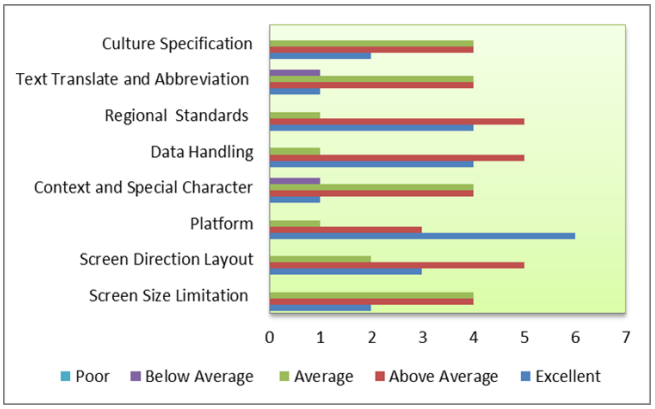

Figure 3. WRS Internationalization Test attribute for Arabic language

According to the test results, the internationalization testing for the Arabic language as depicted in Figure 3 highlights that 6 participants evaluated the WRS app as excellent for platform scalability, whereas 4 participants evaluated the app as above average in terms of text translation. 5 participants evaluated the WRS app as above average in terms of regional standards, data handling, and screen direction. 4 participants evaluated the app as average

 ISBN: 978-1-61208-479-4

as compared to only 2 participants who evaluated the app as below average.

The Chinese language was the second to undergo internationalization testing of the WRS as depicted in Figure 4. 5 participants evaluated the WRS app as excellent for platform scalability, whereas 4 participants evaluated the app as average in terms of culture specification and regional standard. In addition, 3 participants evaluated the app as above average compared to only 2 participants who evaluated the app as below average in terms of context and special characters. 1 participant evaluated the WRS app as poor in terms of text translation, context and special characters.

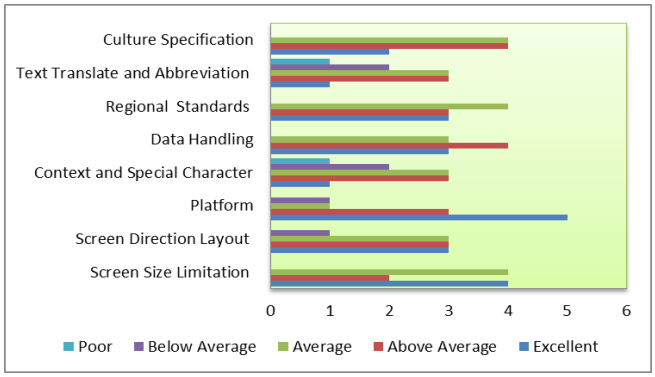

Figure 4. WRS internationalization Test attributes for Chinese language

The third language to undergo internationalization testing was the French language as depicted in Figure 5. Here, 7 participants evaluated the WRS app as above average for platform scalability, whereas 5 participants evaluated the app as excellent in terms of screen size limitation.

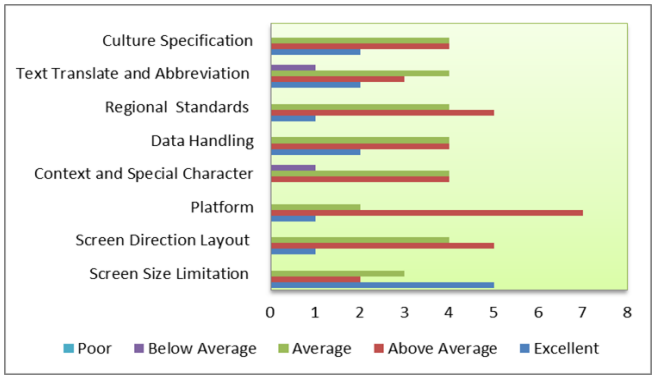

Figure 5. WRS Internationalization Test Attributes for French Language

4 participants evaluated the app as above average as compared to only 1 participant who evaluated the app as below average. Only 1 participant evaluated the WRS app as below average in terms of context and special characters.

Internationalization testing has also been carried out for the German language as depicted in Figure 6.

In this test, 6 participants evaluated the WRS app as above average for platform scalability, whereas 6 participants evaluated the app as above average in terms of culture specification.

4 participants evaluated the app as above average, as compared to only 2 participants who evaluated the app as below average in terms of context and special characters. Only 1 participant evaluated the WRS app as poor in terms of text translation, context and special characters.

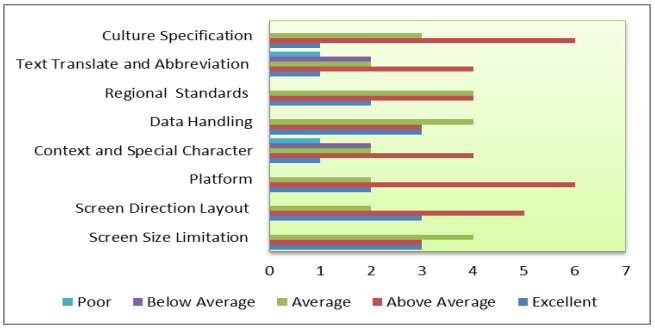

Figure 6. WRS Internationalization Test Attributes for German Language

The fifth internationalization testing was carried out for the Romanian language as depicted in Figure 7.

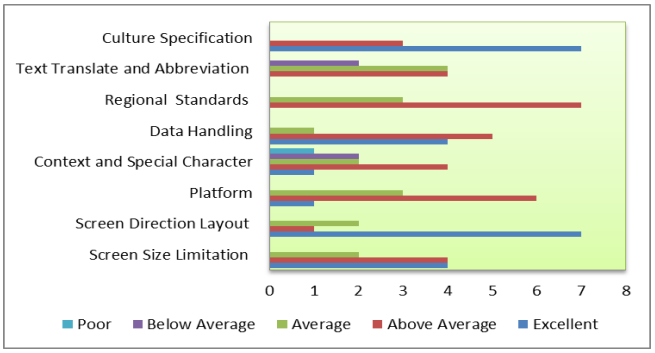

Figure 7. WRS Internationalization Test Attributes for Romanian Language

The majority of the participants evaluated the WRS app as excellent in terms of screen direction and culture specification; whereas 4 participants evaluated the app as above average in terms of text translation, context and special characters. 4 participants evaluated the app as excellent, 5 participants evaluated the app as above average, and only 1 participant evaluated the WRS app as average in terms of data handling.

The internationalization testing was carried out for the Spanish language as depicted in Figure 8.

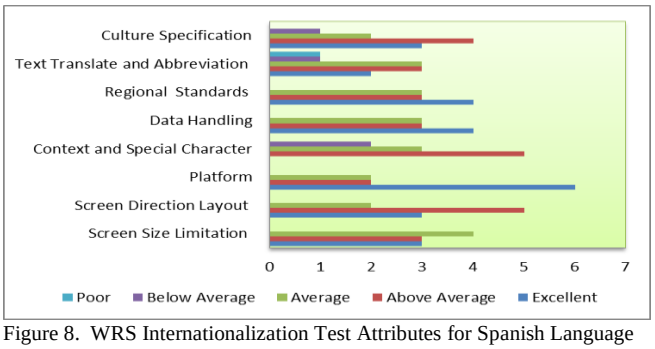

Figure 8. WRS Internationalization Test Attributes for Spanish Language

 ISBN: 978-1-61208-479-4

6 participants rated the WRS as excellent in terms of platform scalability. Whereas 5 participants evaluated the WRS app as above average in terms of context, special characters and screen direction layout. 4 participants evaluated the WRS app as excellent in terms of data handling, as compared to the other 4 participants who evaluated the WRS app as below average for context, special characters, text translation and culture specification.

The Turkish language has also been tested as depicted in Figure 9. 6 participants evaluated the WRS app as excellent in terms of platform scalability, and 5 participants evaluated the WRS as excellent for the screen limitation, screen direction and culture specification.

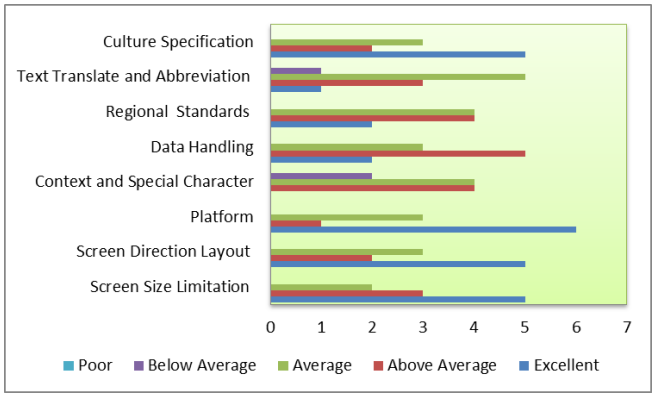

Figure 9. WRS Internationalization Test Attributes for Turkish Language

4 participants evaluated the WRS as average for regional standard, 3 participants evaluated the WRS as below average scale for context, special characters and text translation.

From the above analysis it is clear that the researcher dealt with 70 participants through seven different languages; the ratio of trust these participants gave to the application according to the offered attributes is depicted in the table below.

From the below table it is clear that the attributes that best won the confidence of the participants and were above average are Platform (70.8%), followed by Screen Direction Layout (70.3%) and Data Handling (70.3%), respectively.

TABLE II. THE RATIO OF TRUST GIVEN TO THE APPLICATION BY THE PARTICIPANTS, ACCORDING TO THE SPECIFIED ATTRIBUTES.

The Attributes	Average %	Total (N)
Screen Size Limitation	60.7	N = 70
Screen Direction Layout	70.3	
Platform	70.8	
Context and Special Characters	40.6	
Data Handling	70.3	
Regional Standards	60.7	
Text Translation and Abbreviation	40.5	
Culture Specification	70.0	

However, Context and Special Characters (40.6%) were below average. This is perhaps due to the difference in the size of the words and the number of characters per word for each language. Additionally, Text Translation and Abbreviation (40.5%) was lower than the average, this has been attributed to the translators failing to arrive at a precise and uniform translation for some of the terms, this is because each languages' words carry many meanings, and the translators can detect errors only when the application is used.

The ratio of trust that the participants gave to the application according to each language is in the following table:

TABLE III., THE RATIO OF TRUST GIVEN TO THE APPLICATION BY THE PARTICIPANTS ACCORDING TO EACH LANGUAGE.

The Languages	Average %	Total (N)
Arabic	70.1	N = 10 for each Language
Chinese	50.9	
French	60.0	
German	60.4	
Romanian	70.3	
Spanish	60.6	
Turkish	60.3	

From the above table it is clear that the best languages that won the confidence of the participants and performed above average were the Romanian language which achieved (70.3%), followed by Arabic (70.1%), Spanish (60.6%) and German (60.4%). The French Language (60.0 %) and Chinese language (50.9%) came in last, possibly due to the same reasons mentioned above.

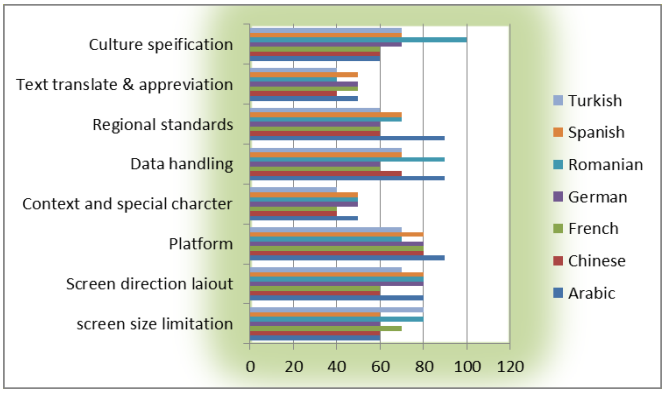

Figure 10. WRS Internationalization Test Attributes For All Languages

From figure 10 above, it is clear that there are some problems in the attributes (Context and Special Characters) as well as Text Translation and Abbreviation. The causes of these problems have been explained previously in the second and third tables.

VI. CONCLUSION AND FUTURE WORK

The aim of this research was to investigate two areas: firstly, it demonstrated the concept of the internationalization testing process, and secondly, it established a comprehensive method for internationalization testing techniques for a Wireless Response System (WRS) application in multilingual use. However, based on our

literature study we can highlight that international testing is a challenging task for mobile applications because it has a great impact on the quality of the mobile application in terms of globalization attributes. Our experiments were based on a holistic view of how to influence the WRS testing strategy for multilingual apps by identifying the challenges. In addition to this, we identified some of the key attributes and demonstrated the careful design of an internationalization testing strategy to overcome these challenges and develop solutions for performing localization testing of mobile applications.

In this paper we presented the globalization testing techniques including the internationalization and localization testing for the WRS case study evaluating globalization-testing issues, specifically within WRS apps. Our aim was to discover issues and errors in the selected languages. We also intended to focus on the key attributes for each language. Each language's attributes have been tested by native speakers of the selected languages in order to identify and highlight the drawbacks of WRS, as well as to help conduct different approaches for target audiences with varied cultures, regions, or screen direction.

For future work, we will attempt to use the internationalization testing approach in further validation experiments for more languages, followed by comparing the internationalization attributes in each of the languages in terms of complexity.

REFERENCES

[1] D. Banks, *Audience Response Systems in Higher Education: Applications and Cases: Applications and Cases*: IGI Global, 2006.

[2] S. K. T. John. *Localisation Testing, Build a Great Team.* Available: http://www.ministryoftesting.com/2014/08/localisation-testing-build-great-team/, 2014

[3] Z. Meng and J. Lu, "'Implementing the Emerging Mobile Technologies in Facilitating Mobile Exam System," in *2011 2nd International Conference on Networking and Information Technology, IPCSIT*, pp. 80-88 2011,

[4] J. Lu, R. P. Pein, G. Hansen, K. L. Nielsen, and J. B. Stav, "User Centred Mobile Aided Learning System: Student Response System (SRS)," in *Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on*, pp. 2970-2975. 2010,

[5] Lu, Zhongyu Joan, ed. "Learning with Mobile Technologies, Handheld Devices, and Smart Phones: Innovative Methods: Innovative Methods." *IGI Global*, 2012.

[6] U. Sekaran, "Research methods for business: a skill building approach," *Journal of Education for Business*, vol. 68, pp. 316-317, 2003.

[7] B. Esselink, *A practical guide to localization* vol. 4: John Benjamins Publishing, 2000.

[8] P. Sandrini, "Website localization and translation," in *EU-High-Level Scientific Conference Series MuTra*, 2005.

[9] C. Buck, C. Horbel, T. Kessler, and C. Christian, "Mobile Consumer Apps: Big Data Brother is Watching You," *Marketing Review St.Gallen*, vol. 31, pp. 26-35, 2014.

[10] J. Gao, X. Bai, W.-T. Tsai, and T. Uehara, "Mobile Application Testing: A Tutorial," *Computer*, vol. 47, pp. 46-55, 2014.

[11] J. Lu, "Student Response System (SRS)/Wireless Response System (WRS)–a Next-Generation Student Response System for Academia and Industry," 2011.

[12] B. Amen, M. Sardasht, and J. Lu. "Mobile Application Testing Matrix and Challenges", in *Fifth International conference on Computer Science and Information Technology* Sydney, Australia, 2015 pp. 27-40

[13] M. Kompf, "Usability and internationalization of information technology," *Education and Information Technologies*, vol. 11, pp. 187-189, 2006.

[14] M. Fetaji and Z. Dika, "Usability testing and evaluation of a mobile software solution: a case study," in *30th International Conference on Information Technology Interfaces*, 2008 ,pp. 501-506,

[15] V. Rajappa, A. Biradar, and S. Panda. "Efficient software test case generation using genetic algorithm based graph theory." 2008 First International Conference on Emerging Trends in Engineering and Technology, July 2008, pp. 298-303.

[16]Bernal-Merino, Miguel Á. "Globalization and Co-Creation: Trends in International Game Production" in A. Esser, M. Á. Bernal Merino and I.R. Smith *Media Across Borders: Localising TV, Film and Video Games* , 2016

[17] N. Albalooshi, N. Mohamed, and J. Al-Jaroodi, "The challenges of Arabic language use on the Internet" in *2011 International Conference for Internet Technology and Secured Transactions (ICITST), ,* 2011, pp. 378-382.

[18] J. C. Bastien, "Usability testing: a review of some methodological and technical aspects of the method," *International Journal of Medical Informatics*, vol. 79, pp. e18-e23, 2010.

[19] K. Haller, "Mobile Testing," *ACM SIGSOFT Software Engineering Notes*, vol. 38, pp. 1-8, 2013.

[20] J. Lei and J. Zhang, "Challenges and Future Directions in Sinic Education," in Y. Zhao (ed) *Handbook of Asian Education: A Cultural Perspective*, p. 134-156, 2010.

 ISBN: 978-1-61208-479-4