

languages under study.

ISSN: 2583-049X

Received: 22-01-2022 Accepted: 02-03-2022

The scientific contribution of Arabic to Information Technology research compared to other Languages

¹Arwa Aleryani

¹ Associate Prof., Computer and Information Technology, Sana'a, Yemen

contribution of the Arabic language to knowledge in

information technology by comparing to the contribution of

languages that have more speakers than Arabic and the

contribution resulting from languages that have fewer

speakers than Arabic. The results of this comparison showed

a sharp weakness in scientific research in the field of

information technology in Arabic, compared to all

Corresponding Author: Arwa Aleryani

Abstract

One of the pressing questions is whether one language is sufficient in the field of scientific research? A quick survey of research in scientific fields shows that we are facing the dominance of only one language, which is English, which is the majority of scholars deal with in scientific research. Does English take the language of science for granted, or is there someone who struggles to try to write scientific research in the local language?

The aim of the current research is to shed light on the

Keywords: Scientific research, Information Technology, Arab language

Introduction

When people engage in science, the language of communication they use tries to be more accurate and consistent. Science often gives technical words with specific meanings and also gives scientific meaning to words that may have different uses in everyday language (Hellinger & Pauwels, 2007)^[1].

One of the most important fields in applied linguistics is scientific communication, which deals with the link between language and science. In reality, inter-subject collaboration is what allows science to exist in the first place.

What is meant by "scientific communication" here is a set of discourses aimed at disseminating knowledge and research findings through communication procedures that vary depending on who acts as the transmitter and who is the transmitter to him. Changes in such procedures are mostly driven by societal factors such as economic, demographic, and political factors, which have ramifications for languages competing in scientific communication.

Although a common language is important in scientific communication, creating multilingual alternatives would enhance diversity while maintaining a communication channel. This effort should come from different actors and should not fall solely on the shoulders of English as a Foreign Language (EFL) researchers (Rami'rez-Castañeda, 2020)^[2]. Most of the reasons for selecting English mentioned by the Arabic researchers are the lack of references in Arabic in addition to the lack of highimpact journals that accept research papers written in Arabic (Aleryani, 2018)^[3]. In addition, many parts of the world, PhD students are required to publish a research essay in English in order to obtain the final degree (Nagano & Bukovszki, 2018)^[4].

Literature study

Scientific communication in the era of globalization is an excellent place to observe how most languages compare to English in terms of research and development. Current changes in scientific communication concern not only the language used, but also the sorts of texts, subject styles, and communication goals (Hellinger & Pauwels, 2007)^[1]. The author and the co-authors (Hellinger & Pauwels, 2007)^[1] discussed the researchers' apparent lack of interest in the issue although publication language is an inevitable feature of scholarly communication.

Tsunoda (Tsunoda, 1983)^[5] recorded the change in linguistic usage in the field of the natural sciences over an entire century (1880-1980) showing progress in the strong diffusion of English harmful to German and French which started slowly from 1930 and then it is exponential from the 1950s onwards.

Languages other than English are losing their scientific registers. The gradual spread of English as a major language of international scientific communication has been interpreted in many different ways by many scholars (Guardiano, 2007)^[6].



International Journal of Advanced Multidisciplinary Research and Studies

The paper presented a brief account of the scientific debate on such topics, focusing on the main stereotypes that have been created to provide explanations for the development of English as a language of science, and on the perception of non-English speaking scholars on the reasons for the dominance of English in the scientific literature and its disadvantages regarding speakers the originals.

Rami'rez-Castañeda (Rami'rez-Castañeda, 2020)^[2] studied the case of Colombian researchers in biological sciences. The aim was to overcome the negative consequences of English dominance that maintain the global science gap. The author reported that currently 98% of scientific publications are written in English, including researchers from EFL countries. The research aims to identify the disadvantages arising from the dominance of language in scholarly publishing by examining the additional costs that English communication creates in producing articles in the case of Colombia which is among the countries with the lowest level of English proficiency in the world. The author reported that it has been determined that more than 90% of scientific articles published by Colombian researchers are in English, despite publishing in a second language creates additional financial costs for Colombian researchers in addition to the problems with reading comprehension, ease of writing, as well as time and anxiety. Moreover, rejection or revision of these articles due to English grammar were reported by 43.5% of PhD students, and 33% chose not to attend international conferences and meetings due to the mandatory use of English in oral presentations.

Aleryani (Aleryani, 2018) ^[3] attempts to identify the challenges of using Arabic as a research language in the field of information technology. The results of the research found that most Arabic researchers agree on the existence of problems and obstacles to the use of Arabic in scientific research in the field of information technology. They admitted that English is the language of science, despite the difficulties some are faced in understanding which require translating because of the weakness in their English.

Hanauer and his co-authors (Hanauer, *et al.*, 2019) ^[10]. investigateed the additional burden that Mexican and Taiwanese English-speaking researchers observe when writing research articles in English as a second language (L2) compared to their L1 experience in scientific writing. The results revealed statistically significant differences between first and second languages in scientific writing with an increase in the burden on scientific writing in the second language consisting of an average increase of 24% in difficulty, 10% in dissatisfaction, and 22% in anxiety.

Establishing the Essential Data

The first step was to obtain the basic data for research purposes. The required data is the number of speakers of the selected languages, two languages of which are spoken by more than Arabic speakers and two with fewer. The next step is to determine five terms that are commonly used in scientific research in information technology and translate them into the chosen languages. Tables 1 and 2 show this data.

Table 1: Number of speakers for the selected languages

S. No	Language	Native Speakers	No. of Speakers as Second Language	Total
1	Spanish	471.4 million	71.5 million	543 million
2	English	369.9 million	978.2 million	1.348 billion
3	Standard Arabic		315 million	315 million
4	Portuguese	232.4 million	25.2 million	258 million
5	Russian	153.7 million	104.3 million	258 million

Source: Ethnologue (2021, 24th edition)- (Retrieved on 29 -01-2022)^[7]

Table 1 showed the number of native speakers of four languages which are Spanish, English, Standard Arabic, Portuguese and Russian. In addition to the speakers of these languages as a second language in the world. Our target language, which is standard Arabic, has 315 million speakers, we selected two languages that have above this number of speakers and two that have below it. Bearing in mind that the number of speakers in each language, included a percentage of children, illiterates (though this percentage varies between the selected countries), as well as a percentage of those who are not interested in the field of information technology, we assume the existence of these percentages in each language more or less, so the comparison remains valid.

Languaga	Term 1	Term 2	Term 3	Term 4	Term 5	
Language	IoT	AI	ML	NLP	CS	
Spanish	Internet de las	Inteligencia artificial	aprendizaje	procesamiento natural del	la seguridad cibernética	
Spanish	cosas	inteligencia artificiai	automático	lenguaje		
English	Internet of things	Artificial intelligence	Machine Learning	Natural Language	Cyber Security (CS)	
English	(IoT)	(AI)	(ML)	Processing (NLP)		
Standard Arabic	انترنت الأشياء	الذكاء الإصطناعي	تعلم الآلة	معالجة اللغة الطبيعية	الأمن السبيبراني	
	A Internet das		aprendizado de	processamento de		
Portuguese	coisas	Inteligencia artificial	máquina linguagem natura		Cibersegurança	
_			-			
Russian	Интернет	Искусственный	машинное	обработка естественного	информационная	
	вещей	интеллект	обучение	языка	безопасность	

Table 2: Information Technology/Terms by the selected languages

Table 2 showed translation of the selected information technology terms, which is the field under study, in each selected language, however, Arabic has different ways to

write these terms, so all the possibilities have been considered. The translation has been done by google translate.

International Journal of Advanced Multidisciplinary Research and Studies

Research Problem

Societies develop through the production of science in its original language, through original production or translation. Arabic-speaking researchers have reservations about using their language as a language in scientific writing, and one of their most important reasons is the lack of many readers of the Arabic language in the scientific field, and the other reason is the lack of Arabic sources that can be used. Therefore, the research relied on comparison with languages that have fewer speakers, so that we can identify the reasons behind the problem of the scarcity of scientific production in the Arabic language.

Research Aim

Nowadays, with increasing pressures on researchers from all over the world and from many language groups to publish their research, scientific publishing is dominated by the English language (Hanauer, *et al.*, 2019)^[10].

The current research aims to compare the scientific contribution in the field of information technology in Arabic with the size of the scientific contribution in languages whose number of speakers is higher than Arabic on the one hand, and on the other hand, with the scientific contribution from languages whose number of speakers is less than Arabic.

The goal is not to determine the volume of scientific production of these terms, but rather the goal is to compare the scientific contribution of knowledge published in the Arabic language compared to publication in the selected languages, so we will be satisfied with articles published since 2018 that will give indications for comparison.

Research Importance

As previously mentioned, societies develop with the development of the volume of their scientific productions in their original language, and thus the research aims to know the results of this comparison, which the researcher expects from the actual experience that it will not be good for the Arabic language. Thus, it is possible in the future to study the reasons and propose solutions that can contribute to increasing the Arab scientific production in the field of information technology.

Research Methodology

The methodology used in the current research is to find out the number of speakers of languages in the world. Then we choose two languages in which the number of their speakers is more than Arabic, and two languages in which the number of their speakers is less than Arabic. The languages which have higher speakers than Arabic are English and Spanish, on the other hand, Portuguese and Russian are the languages that have lower speakers than Arabic.

The next step is to identify the five terms of information technology that are being used more by the researchers, which are the Internet of things, artificial intelligence, machine learning, natural language processing, and cybersecurity. These terms have been translated by using Google Translate to the selected languages.

The next step was to search for each term in each language through Google Scholar, specifying that the term should be in the "title" to ensure that the article is affiliated with information technology, specifying the period from 2018, in order to know the scientific contribution in the past short period. In order to determine the volume of participation for these researches, the citation was determined for the first three papers that appeared in the research process.

Data collection and analysis

As a result of the presence of publications in the Arabic language where the term is written with slight differences, each term has been searched in all the forms in which it can be written. The possibility of writing the term incorrectly grammatically (with the possibility of that) has not been considered.

Term		No of Articles	Total	
	إنترنت الأشياء	32		
I-T	انترنت الأشياء	10	40	
101	إنترنت الاشياء	3	49	
	انترنت الاشياء	4		
	الذكاء الإصطناعي	57	68	
AI	الذكاء الصناعي	7		
	الذكاء الصنعى	4		
М	تعلم الألة	1	(
ML	تعلم الآلة	5	0	
NLP	معالجة اللغة الطبيعية	2	2	
	الأمن السبيبراني	32		
	الامن السبيبراني	5		
CS	الأمن الرقمي آ	7	55	
	الأمن الإلكتروني	9		
	الأمن الالكتروني	2		

Table 3: Number of articles for the selected terms in Arabic

Table 3 showed the number of publications in each possibility of terms writing.

Table 4: Number of Articles published since 2018

Language	Published articles in information technology since 2018 Retrieved on Tuesday 1-2-2022 at 8:00 am-10:00 am						
	IoT	AI	ML	NLP	CS	Total	
Spanish	732	3,330	713	85	23	4,883	
English	25,300	30,500	45,100	4,990	6,190	112,080	
Standard Arabic	49	68	6	2	55	180	
Portuguese	209	1,640	637	101	93	2,680	
Russian	706	2,550	371	21	1,710	5,358	

Table 4 showed the number of published articles about IoT, AI, ML, NLP, and CS in each language from 2018 to the present, in addition to the total publications for each language.



Fig1: Total of publications since 2018 in selected IT terms

Fig 1 gives a clear view of the severe weakness in the scientific contribution to information technology research written in Arabic. The Portuguese language has contributed

www.multiresearchjournal.com

nearly 2,680 publications despite it having a lower number of speakers than Arabic. The same for the Russian language

which has had 5,358 publications since 2018.

	The number of citations for the first three published articles, which indicates the					
	amount of benefit from the research					
	Retrieved on Tuesday 1-2- 2022 at 8:00 am- 10: am					
Language IoT AI		ML	NLP	CS	Total	
	A1:21	A1:105	A1:31	A1:	A1:0	312
Spanish	A2:16	A2:57	A2:7	A2:	A2:2	
Spanish	A3:19	A3:45	A3:9	A3:	A3:0	
	T:56	T:207	T:47	T:	T:2	
	A1: 128	A1: 431	A1: 813	A1: 673	A1:235	7,846
English	A2: 959	A2: 678	A2: 886	A2: 938	A2: 187	
English	A3: 452	A3: 584	A3: 438	A3: 297	A3: 157	
	T:1,539	T:1,693	T <i>:</i> 2,137	T <i>:</i> 1,908	T:569	
	A1:0	A1:0	A1:0	A1:0	A1:0	0
Anabia	A2:0	A2:0	A2:0	A2:0	A2:0	
Arabic	A3:0	A3:0	A3:0	A3:0	A3:0	0
	T:0	T:0	T:0	T:0	T:0	
	A1:124	A1:35	A1:7	A1:9	A1:4	288
Dorthonoco	A2:6	A2:31	A2:3	A2:2	A2:5	
Portuguese	A3:7	A3:40	A3:14	A3:1	A3:0	
	T:137	T:106	T:24	T:12	T:9	
	A1:165	A1:318	A1:149	A1:7	A1:351	1,649
Dussian	A2:134	A2:31	A2:22	A2:1	A2:85	
Kussian	A3:30	A3:57	A3:16	A3:0	A3:283	
	T:329	T:406	T:187	T:8	T:719	

Table 5: Citations

The number of citations can be considered a measure of a publication's usefulness, its influence. Hence, citations have been one of the current research interests. Table 5 shows the number of citations for each of the three articles that appear first in the search for each term as well as the total citations for these three articles.



Fig 2: Number of citations

Unfortunately, fig 2 shows the lack of citations from the first three papers that appeared in the search process. This result may indicate the weakness of the usefulness of these articles, even among the users of the Arabic language. While we find quotes in other languages are clearly present. If we focus on the languages below the Arabic language in terms of the number of speakers, we will find that Portuguese obtained 188 citations for the first three articles, while the Russian language obtained 1,649.

Discussion

Over the ages, nations have developed by translating science into their language so that everyone can read it and benefit from it, and most societies have been keen to publish and contribute scientifically to the local language, in order to ensure its presence in scientific contributions.

The English language has taken the lead in science, including information technology, and these contributions have been made by researchers from these countries who take English as their official language, moreover, there are distinguished contributions by researchers from other countries who use English as a scientific language. Furthermore, by the presence of indexes and influence factors, English-language journals are seen as the most important places to publish (Cargill & Burgess, 2017; Sheridan, 2017)^[9, 8]. As a result, the vast majority of available resources are written in English, and most high-impact journals are accepted only in English.

From the results of the current research, it is found that all countries (at least those under the current study) have publications of significant magnitude, and their research studies are cited in such a way as to motivate them to continue to use their languages in scientific research. But Arabic was the weakest for no apparent reason, especially when compared to the contribution from languages spoken by fewer than Arabic speakers.

Conclusion

The Arabic researchers should diversify their research production between English and Arabic to form a presence for the Arabic language in science. The problem of the scarcity of scientific resources for information technology in Arabic can be solved by doing research in Arabic with the use of available resources in Arabic or English. In addition to encouraging the scientific institutes to establish journals with high impact factors and highly rated for science research papers that accept the articles in Arabic. International Journal of Advanced Multidisciplinary Research and Studies

In general, engaging in L2 scientific production is a real burden on international researchers and should be understood regardless of anecdotal reports to the contrary. This fact is also supported by extensive research by other applied linguists illustrating the complexities and difficulties that scholars publish in English as a second language (López- Navarro, *et al.*, 2015; Hanauer, *et al.*, 2019)^[10]. The additional burden of scientific writing in a second language constitutes a linguistic injustice and a barrier to science that must be addressed, where the writing burden contributes to the feeling that English is a barrier to science writing (Hanauer, *et al.*, 2019)^[10].

From the previous research that was mentioned in the current research, it appears that there is a problem with the adoption of the English language as a single language for scientific research, but unfortunately this research showed that the problem for Arabic seems much deeper, especially with regard to information technology research. The Arabic language deserves to have a clear presence and an outstanding contribution in the field of science, especially information technology.

References

- 1. Hellinger M, Pauwels A. Handbook of Language and Communication: Diversity and Change. Handbooks of Applied Linguistics (HAL). Berlin, New York: Mouton de Gruyter. 2007; 9:525-554.
- 2. Rami'rez-Castañeda V. Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological sciences. PLoS One. 2020; 15(9):e0238372.

https://doi.org/ 10.1371/journal.pone.0238372

- 3. Aleryani A. Challenges and Opportunities of Using Arabic Language in Scientific Research in the Field of Information Technology from the Point of View of Arab Researchers, International Journal for Quality Assurance. 2018; 1(1).
- 4. Nagano RL, Bukovszki Spiczene E. PhD publication requirements and practices: A multidisciplinary case study of a Hungarian university. In M. J, 2018.
- 5. Tsunoda M. Les langues internationales dans les publications scientifiques et techniques. Sophia Linguistica, 1983, 144-155.
- 6. Guardiano C, Favilla M, Calaresu E. Stereotypes about English as the language of science, AILA Review. 2007; 20(1):28-52.
- Doi: https://doi.org/10.1075/aila.20.05gua.
 7. Ethnologue, 24th edition, 2021. https://www.ethnologue.com/guides/ethnologue200 (Retrieved on 29-01-2022)
- Sheridan CL. English medium "national" journals beyond the anglophone center: A qualitative study of multilingual scholars and their publishing decisions in Taiwan (Unpublished doctoral dissertation). Indiana University of Pennsylvania, 2017.
- 9. Cargill M, Burgess S. Publishing research in English as an additional language: Practices, pathways and potentials. Adelaide, Australia: University of Adelaide Press, 2017.
- 10. Hanauer D, Sheridan C, Karen Englander K. Linguistic Injustice in the Writing of Research Articles in English as a Second Language: Data from Taiwanese and

Mexican Researchers, Written Communication. 2019; 36(1):136-154.