

Language Selection and Query Formulation in Arabic Users' English-Language Searching

Ashwag Alasmari^{1 2}, Anita Komlodi¹

¹ University of Maryland, Baltimore County (UMBC)

² King Khalid University (KKU)

Abstract

Due to the overwhelming presence of English language sources on the Web, large numbers of non-native English speakers are forced to seek information in English to fulfill their information seeking needs. Arabic-speakers are among those who search in English as a non-native language. The objective of this work-in-progress paper is to explore the search behaviors and challenges of Arabs when searching in English. Data were collected in a diary study and focus groups. To mitigate the challenges of searching in English, our participants often turned to members of their own community to help navigate unfamiliar information spaces.

Keywords: Information-seeking behavior; online search; Arab web users; query formulation

doi: 10.9776/16586

Copyright: Copyright is held by the authors.

Contact: ashwag1@umbc.edu, komlodi@umbc.edu

1 Introduction

The Internet is a significant source of information, however, the majority of this information is in English while many searchers' native language is not English. In particular, Arabs represent a high proportion of overall web users, with approximately 135 million online ("Top Ten Internet Languages", 2013). Despite this large number, less than one percent of Web sources are available in Arabic (Chung, 2008). Consequently, Arab users often search in English. The purpose of this study is to explore of the search behaviors of Arab speakers when searching online in English.

2 Related Studies

A growing body of literature exists investigating the behavior of Arab language speakers as they search the Web and interact with search engines. This research has identified a number of interesting search patterns among Arab web users. One common search behavior is the use of short search expressions. Researchers have found that the majority of Arab searchers used keywords consisting of only one term across all languages used for searching (Al-Kabi, et.al, 2012; Tawileh, 2011). However, these researchers did not examine the differences when searching in Arabic vs. in English.

3 Data Collection and Analysis

In this study we used qualitative methods to explore search behavior in English. Participants were asked to record searches in diaries and recount them in interviews. A different group of participants was interviewed in two focus groups. Participants for both the diary study and the focus groups were students at the English Language Institute (ELI) of a US University and native Arabic speakers.

Participants in the diary study were asked to record their search behavior in a paper-based diary. The diary was designed to capture the intent of the web search, most important steps, and any challenges participants encountered. Participants were asked to capture screenshots, in addition to textual descriptions of the searches. Each participant was asked to record 10 searches in an eight-week period. At the end of the diary study, each participant was asked to participate in a one-hour semi-structured interview to discuss the searches they captured. We had 4 participants in this study.

Participants in the focus group were asked a set of questions about searching in English versus Arabic, the problems they encountered when searching in English, and any design recommendations that would improve searching in English as a non-native speaker. Table 1 lists the questions we asked in the focus groups. There were 11 participants in the focus groups.

Focus Group Questions
1. What topics do you search for in your native language and in English? Why?
2. What search engines or websites do you use in the case of searching and browsing in English?
3. How are your searching and browsing different in English and in your native language?
4. What problems do you find during searching and browsing in English?
5. How can search engines and websites help searching and browsing in English?
6. What advice can you give to other Arab students who have recently started searching for information in English (e.g. incoming first year university students)?

Table 1. Discussion Guide for Focus Group

Three forms of qualitative data were collected, diary entries, notes, and voice recordings of the interviews and the focus groups. Thus, the first step of the data analysis was to transcribe the voice recordings into Arabic. The transcriptions were in Arabic and the coding process was executed in Arabic. The transcripts were coded through an inductive and iterative process to identify the motive and general patterns of behavior (Merriam, 2009). Next, we present the resulting themes identified from the analysis.

4 Findings

4.1 Search Topic Drives Language Selection

Participants reported that their decision to use either English or Arabic is determined by the type of information they are seeking. For example, when searching for information related to Middle Eastern politics or religion, participants indicated that they searched in Arabic. One participant reported: "I will never look at religious matters in English" (Participant 1). Another participant commented: "When I would like to know about situation and news in Middle East, particularly Libya, Arabic is the more ideal language to access this kind of information" (Participant 2). The decision to search for these topics in Arabic was influenced by both the availability of content and cultural interpretation of these topics. Arab users in our study were more likely to search in their native language and seek out the shared point of view of Arabic sources. Participants also searched for information on Arabic sports, cooking, and music in their native language to ensure that the content they received matched their information need.

By contrast, the majority of search tasks in English were carried out for academic purposes to match their studies, which they were conducting in English. This is similar to findings of Chinese and Hungarian web users (Chu et al., 2015). Most of the diary entries were searches carried out for academic reasons. In addition, participants in the focus group declared that the main motive of searching in English is looking for academic materials relating to our study.

Participants also searched in English when they were looking for geographically situated information. All participants in the study have lived in the United States. Thus searching for information regarding their geographical location is one reason of searching in English. For instance, one participant used English to search for public safety information in Baltimore, stating, "since I am in Baltimore, I would like to familiarize myself with the city I have lived in" (Participant 3).

4.2 Searching as a Cultural Expression

Belonging to a particular cultural group plays an important role in searching for and finding information on the Web (Chu et al., 2015). The majority of participants reported that when searching for information they prefer to select sources recommended by their fellow countrymen. One focus group participant indicated, "When I would like to move to or visit a different state in United States, I would look for information that is posted by other Saudis. I know we understand each other and share the same culture. I will not look for what Americans say or think about a certain place" (Participant 5). Similarly, another focus group participant commented, "Not because we only struggle using English language, but we prefer knowing what my society would say about a certain topic." (Participant 6)

Social media is one of the ways participants reported searching for information as expressed by other members of their cultural group. In some cases, participants used social media as their primary search engine as it allowed them to see what other Arabs posted about a specific topic. One participant reported

that “using Google confuses me and I would rather use Twitter and Instagram to have more updated information from people I know” (Participant 8).

4.3 Challenges

The participants in our study struggled with forming queries and faced several other challenges while searching for information in English.

4.3.1 Query Formulations

Rozsa et al., (2015) have found that search expressions amongst non-native language searchers tend to be general, even in instances where they are looking for specific information. We found the same pattern among our participants. For example, when searching for the fat content of peanut butter, one Arabic diary-entry participant did not include the term “fat” in her/his query. In addition, we found in our participants’ searches, three-keyword queries are more common across all different searches, about 40%. Table 2 shows the frequencies of using different length in formulating search expressions. There are only three searches among 40 searches with one-keyword, which that contradicts with the previous finding that states that majority of Arab searchers’ keywords consisting of one term both in English and Arabic (Al-Kabi, et.al, 2012; Tawileh, 2011). However, contextual factors influence the length of search expressions, thus a direct comparison cannot be made.

Long of Query	Number of Occurrences
One- keyword	3
Two- keyword	7
Three- keyword	17
Four- keyword	7
Five- keyword	3
Six- keyword	2
Nine – keyword	1

Table 2. Type of Queries and Occurrences

4.3.2 Query Reformulations

The most common query reformulations were conducted to correct spelling mistakes. For instance, one participant indicated “even if I am not sure about vocabularies’ spelling, I would use it anyway and I trust Google to provide me with the correct spelling”.

Additionally, search expressions are sometimes formulated after browsing a number of search results. When a user fails to find what they are searching for, he/she will reformulate the search expression, as suggested in the following quote: “When I know the list of search results are not fulfilling my needs, I can have a quick overview of the list of retrieved search results. That helps me to formulate some keywords that I would use to change the current query”. (Participant 10)

4.3.3 Translating New Vocabulary to Arabic

In most cases, participants did not use new words without understanding their meaning. They preferred to translate any unfamiliar words encountered in the results to Arabic before using them in an English query. Additionally, when some participants opened a website, they would try to translate each new word in Google Translator to understand the context.

5 Conclusion

In this paper we presented an exploratory study of Arabic users’ search behavior. Foreign language searchers face challenges of creating queries where both vocabulary and grammar present challenges. These queries are often poor tools to navigate an unfamiliar information space. Users naturally turn to searching in their native language or finding their own cultural group to facilitate navigation in this information space. Future research is needed to more thoroughly explore behavior and the impact of language proficiency and familiarity with the culture of information producers in effectively searching for and using information.

6 References

- Al-Kabi, M., Wahbeh, A., Alsobeh, A., Ale'roud, A., & Alsmadi, I. (2012). Examining web search trends across Arab countries. *Arabian Journal for Science and Engineering*, 37(6), 1585-1593.
- Chu, P., Komlodi, A., & Rozsa, G. (2015) Online Search in English as a Non-native Language. Proceedings of the Annual Meeting of the American Society for Information Science and Technology. St. Louis, MO, November 6-10, 2015.
- Chung, W. (2008). Web searching in a multilingual world. *Communications of the ACM*, 51 (5), 32-40.
- Merriam, S. B. 2009. *Qualitative research: A guide to design and implementation*. John Wiley & Sons
- Tawileh, W. (2011). Exploring web search behavior of Arab internet users. In *Innovations in Information Technology (IIT), 2011 International Conference on* (pp. 313-318). IEEE.
- Top Ten Internet Languages - World Internet Statistics. (2013). Retrieved from <http://www.internetworldstats.com/stats7.htm>
- Rozsa, G., Komlodi, A., & Peng, C. (2015). Online Searching in English as a Foreign Language. WWW Companion: Proceedings of the 24th International Conference on World Wide Web. Pages: 875-880. doi>10.1145/2740908.2743007. Presented at the The 1st International Workshop on Multilingual Web Access (MWA 2015) - May 19, 2015.
- Wahbeh, A. H., Al-Kabi, M. N., Maabreh, M. A., Al-Smadi, I. M., & Al-Shawakfa, E. M. (2014). Studying and analysing students web search behaviours within three Jordanian universities. *International Journal of Continuing Engineering Education and Life Long Learning*, 24(3), 388-402.