

# Learning in Second-Language Searching

Anita Komlodi  
UMBC  
1000 Hilltop Circle  
Baltimore, MD 21250  
1-410-312-5665  
komlodi@umbc.edu

Nadia Caidi  
University of Toronto  
140 St. George Street  
Toronto, Ontario M5S 1C7, Canada  
+1-416-978-4664  
nadia.caidi@utoronto.ca

## ABSTRACT

In this paper we describe learning that happens when searchers look for information in a foreign language. A strong disparity exists between the language distribution of Web content and the representation of speakers of different languages among Web users. The proportion of Web content available in English is much larger than the proportion of native-English-speaking Web users. This disparity often forces non-native speakers of English to search for information in English. In this process they do not only learn the language, but also familiarize themselves with a new culture and information landscape. In this paper we describe when and how this learning happens and what tools and features can help the process.

## CCS Concepts

Information systems → Web searching and information discovery  
Information systems → Structure and multilingual text search

## Keywords

Search; Second-language search; Second-language acquisition; Information landscapes, literacies.

## 1. INTRODUCTION

Search systems have traditionally been developed to support retrieving information. Recently, the widespread availability and use of full text search systems have led to viewing these systems as learning environments [6]. Users interacting with information in search systems go through an iterative seeking and learning process. [11] They engage with rich full-text information, which informs their learning process and they modify their search behavior based on what has been learnt. [2] In this paper, we discuss the learning that occurs in searching when the searchers look for information in a foreign language. We use the example of searching in English as a foreign language throughout this paper.

A strong disparity exists between the language distribution of Web content and the representation of speakers of different languages among Web users. While English content makes up 55.5% of all content on the Web [14], native English speakers make up only 28.6% of all Internet users [7]. In contrast, Chinese

Search as Learning (SAL), July 21, 2016, Pisa, Italy

The copyright for this paper remains with its authors. Copying permitted for private and academic purposes.

first language speakers make up 23.2% of Web users [7] while Chinese language content accounts for only 2.8% of Web content [14]. This disparity forces many non-native English speakers (NNEs) to search in English to satisfy their information needs [8].

Non-native English speakers search in English for various reasons. Those living in their home country and immersed into their culture may search in English due to the lack of content in their native language. Those immigrating, temporarily or indefinitely, to an English-speaking country search in English to find information about their new environment [4]. While these two groups of searchers will have very different supports, challenges, and experiences, they will both learn about the language, culture, and new information environments through interacting with the search system.

While the same type of learning happens for native-language searchers to a lesser extent, the focus in this paper is on foreign-language searchers. The ideas presented here are meant to provoke discussion and spark ideas for future research. This paper is informed by several research projects [1, 3, 5, 9, 12, 15] led by the co-authors, however, this position paper is intended as a conceptual piece and does not include data.

## 2. LEARNING THROUGH FOREIGN-LANGUAGE SEARCH

### 2.1 Language Learning

The literature on NNEs searching behaviour is sparse. Yet, the existing literature points to several language learning processes, often observed and reported through user diaries and interviews [1, 3, 5, 9, 12, 15]. As online searchers seek information in a foreign language, they continuously acquire new language skills: they explicitly learn *new vocabulary*, often professional domain terminology. More implicitly, they also learn *grammar and sentence and text structure* through interacting with and interpreting foreign-language text. They develop both their *expressive language skills* through creating queries, and their *receptive skills* through reading and interpreting results and documents.

As searchers create queries, they learn vocabulary and grammar from query recommendations and the presentation of related queries. Identifying the appropriate terms to search on is the most difficult task for foreign-language searchers. Slight differences in meaning and usage can dramatically change the result set and impact the effectiveness of search. Finding the appropriate terms and learning about the slight variations in meaning and usage is a very important part of this process. Seeing the relationship between the query terms and the returned results sets helps searchers gain an understanding of the precise meaning of the words. If entering a query term returns a document with a certain

type of content, that content defines the term's meaning. The iterative nature of this process allows searchers to enter into a learning mode whereby more than the immediate retrieved items is being consciously or unconsciously registered.

Once results are returned, searchers read and interpret text in the foreign language. While searchers find this difficult and often rely on visual information to supplement their reading, this phase of the foreign-language search process presents fewer challenges than constructing queries. Reading and interpreting text with a very specific purpose can advance language learning greatly. Encountering new vocabulary in context helps learning its meaning. Interacting with grammatically correct and well-written text helps learning about the underlying structures of language.

The iterative cycles of reading and interpreting results and constructing queries can lead to enhanced learning through verifying meaning in the context of the results and interpreting content through reading [1, 3, 5, 9, 12, 15].

## 2.2 Learning About an Unfamiliar Information Landscape and Culture

Moreover, as these searchers interact with information in a foreign language, they also are exposed to sometimes unfamiliar information landscapes [10], which they have to learn to navigate. Different cultures have different values and notions associated with information and varying practices of writing and publishing information. Sources and genres of information can be unfamiliar as well, and thus searchers can have difficulty judging the credibility and authority of information. [15] By searching these information spaces, individuals get a sense of *what* the categories mean, in a particular context, and are able to understand the embedded values and relationships proper to a particular cultural context: for instance, how the healthcare or education system is conceptualized in one country vs. another, or what the foreign credentials recognition process says about labor market integration in a given country [3, 9]. Seeking information becomes a lot about learning about the new environment and how the system works [10].

## 3. FEATURES AND FUNCTIONS

*Language learning:* Several existing and proposed functions of search systems can enhance the searching and learning processes of foreign-language searchers. Query term recommendations and the presentation of related queries are both excellent tools for learning about new vocabulary. Providing easy and in-context translation and definition tools for query terms can support foreign-language learning as well. The user-controlled translation of results and resulting documents can also be very beneficial for enhancing reading comprehension skills. Translation tools are more widely available now both on search engine results pages (SERPs) and websites. While translating all text all the time will not enhance language learning, in-context and on-demand translation tools can be highly beneficial.

*Information landscape:* Representing the source and the reliability of the source on the SERP will improve the decision making process of these searchers [13]. Searchers can learn about reliable sources and by using these sources they can also find and learn markers of credibility on web information sources.

## 4. FUTURE RESEARCH DIRECTIONS

Searching in a foreign language is a common occurrence for Internet users, and provides a great opportunity to learn about the language and the culture of another country or cultural group. The language learning behavior of foreign-language searchers is an

area of research that requires further examination. Doing so will allow us to understand not only the practices involved with seeking, using and sharing information in a different language, but also identify learning opportunities and challenges for system design. Indeed, this understanding can lead to the development of more appropriate support tools to promote both finding information as well as the learning that often takes place through the process. This form of learning is increasingly being emphasized as a critical form of literacy, as evidenced by UNESCO's work in the area of global media and information literacy assessment [16].

## 5. ACKNOWLEDGMENTS

Our thanks to Google, Inc. for funding part of the research leading to this paper.

## 6. REFERENCES

- [1] Alasmari, Ashwag; Komlodi, Anita. (2016) Language Selection and Query Formulation in Arabic Users' English-Language Searching. iConference 2016. Philadelphia, PA. March 20-23, 2016. Poster Presentation.
- [2] Bates, M. J. (2002). Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research*, 3, 1-15.
- [3] Caidi, N., Komlodi, A., Lima Abrao, A., & Martin-Hammond, A. (2014). Collectively figuring it out: Foreign trained health professionals and labor market integration. *Libres*, 24(2), 118-131.
- [4] Caidi, N. Allard, D., & Quirke, L. (2010). Information practices of immigrants. *Annual Review of Information Science and Technology*, 494-531.
- [5] Chu, Peng; Komlodi, Anita; Rozsa, Gyongyi. (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. (long paper presentation)
- [6] Freund, L., He, J., Gwizdka, J., Kando, N., Hansen, P., & Rieh, S. Y. (2014, August). Searching as learning (SAL) workshop 2014. In *Proceedings of the 5th Information Interaction in Context Symposium* (pp. 7-7). ACM.
- [7] Internet World Stats. (2013). Top Ten Languages Used in the Web - December 31, 2013 (Number of Internet Users by Language). Retrieved April 1, 2015 from <http://www.internetworldstats.com/stats7.htm>.
- [8] Kralisch, A., & Mandl, T. (2006). Barriers to Information Access across Languages on the Internet: Network and Language Effects. In *Proceedings of the Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, 03. IEEE Computer Society. DOI: 10.1109/HICSS.2006.71.
- [9] Komlodi, A.; Caidi, N.; Martin-Hammond, A. Rayes, W.; Sundin, N. (2016). "Culturally-Situated Information Literacy: International Medical Graduates Navigating New Information Landscapes." In van der Velden, M., Strano, M., Hrachvec, H., Abdelnour Nocera, J., & Ess. C. (Eds.). *Culture, Technology, Communication: Common worlds, different futures? Proceedings of the Tenth International Conference on Culture, Technology, Communication (CATAC)*. London, UK, 15-17 June 2016, pp. 105-110.

- [10] Lloyd, A. (2010). *Information literacy landscapes: Information literacy in education, workplace and everyday contexts*. Cambridge, UK: Woodhead/Chandos.
- [11] Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4), 41-46.
- [12] Rayes, W.; Martin-Hammond, A.; Komlodi, A.; Caidi, N.; Sundin, N. (2016). An Informed Transition? International Medical Graduates Settling in the United States and Canada. In *Proceedings of the 79th Meeting of the Association for Information Science & Technology*, Copenhagen (Denmark), Oct 16-18, 2016.
- [13] Schwarz, J., & Morris, M. (2011). Augmenting Web Pages and Search Results to Support Credibility Assessment. In *Proceedings of SIGCHI* (pp. 1245–1254). New York, NY, USA: ACM Press.
- [14] World Wide Web Technology Surveys. (2015). Usage of content languages for websites. Retrieved April 1, 2015 from [http://w3techs.com/technologies/overview/content\\_language/all](http://w3techs.com/technologies/overview/content_language/all).
- [15] Young, A.; Komlodi, A.; Rozsa, G.; Chu, P. (2016) Evaluating the Credibility of English Web Sources as a Foreign-Language Searcher. In *Proceedings of the 79th Meeting of the Association for Information Science & Technology*. Copenhagen (Denmark), Oct 16-18, 2016.
- [16] UNESCO. (2013). *Global Media and Information Literacy Assessment Framework*. Available at: <http://www.uis.unesco.org/Communication/Documents/media-and-information-literacy-assessment-framework.pdf>