APPROVAL SHEET

Title of Dissertation: Information Seeking and Retrieval in English as a Non-native Language

Name of Candidate: Peng Chu
Doctor of Philosophy, 2017

Dissertation and Abstract Approved: _____

Dr. Anita Komlodi
Associate Professor
Department of Information Systems

Date Approved: _____

ABSTRACT

Title of Document: INFORMATION SEEKING AND RETRIEVAL IN ENGLISH AS A NON-NATIVE LANGUAGE

Peng Chu, Ph.D., 2017

Directed By: Associate Professor, Anita Komlodi, Department of Information Systems

A strong disparity exists between the language distribution of Web content and the representation of speakers of different languages among Web users. While more than one half of content on the Web is in English, native English speakers account for only about one-fourth of all Web users. This disparity forces many non-native English speakers (NNESs) to search in English to satisfy their information needs.

Information seeking and retrieval in a non-native language can present special challenges for users. There is insufficient research available on the information behavior of NNESs when they conduct online searching in English as a foreign language (EFL) or second-language (L2). Even less is known about their search strategies and interaction with systems and tools to accommodate their needs and assist their behaviors.

The purpose of this research is to further understand the NNESs information behavior and, subsequently, explore design solutions to support them. Due to the complexity of the binary nature of the research, the study process is carried out in two stages.

In the first stage, qualitative methods were utilized to extend the current understanding of NNESs search behavior in English. Findings in this stage shed light on behavioral patterns of language selection, search engine selection, query formation and reformulation, strategies, and browsing and filtering search engine result pages (SERPs). An iterative, or rather, spiral search process was observed and a user interaction model accommodating two query construction strategies was abstracted from these patterns.

Based on the previous study findings, the second research stage took place. First, multiple UI prototypes were designed by the researcher. A viable UI prototype, TranSearch 1.0, was reached, developed, and tested through a user testing study. Next, based on the user feedback and the researcher's follow-up usability inspection, the prototype was redesigned to reach a more functional version and include more features, TranSearch 2.0. Finally, a user study was conducted with two purposes, evaluating TranSearch 2.0 and providing further redesign ideas and suggestions from users' perspective. Results and findings not only provide information about the utility of the proposed solutions and design implications, but also further inform the model for NNESs information seeking behavior.

INFORMATION SEEKING AND RETRIEVAL IN ENGLISH

AS A NON-NATIVE LANGUAGE

By

Peng Chu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore County, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Acknowledgements

I would like to express my deep gratitude to Professor Anita Komlodi, my research mentor and advisor, for offering the opportunity, support, and assistantship for all the research projects I participated. Her guidance, encouragement, and critiques of the research and dissertation work are greatly appreciated. I would also like to express my great appreciation to Professors Ravi Kuber, Paul Leidig, Wayne Lutters, and Lina Zhou for serving on my committee, providing valuable advises and constructive suggestions, and accommodating me in their tight schedules. My thanks are also extended to my department, Information Systems, and Ms. Galina Madjaroff from the Erickson School for offering me the teaching assistantship. Finally, I would like to acknowledge the support and encouragement provided by my family throughout my study.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

More and more people search online in a language that is not their native language due to the limited availability or quality of content in their native languages. Most such users search in English as a second language. Information seeking in a non-native language can present special challenges for users. Language proficiency and skills, unfamiliarity of culture and society, and inconsistency in localized website icons, concepts, and models constitute an essential part in the context of searching in a non-native language, which in turn impact the search process and outcome.

Current research literature is insufficient to fully understand non-native English speakers (NNESs) online search challenge, behavior, and process. Even less is known about their search strategies and interaction with multilingual information retrieval (MLIR) systems and tools, as well as how these systems and tools can accommodate their needs and assist their behaviors. The dissertation research is carried out in two stages to 1) provide a better understanding on non-native English searchers' behavior; 2) explore and guide design practices of MLIR systems and UIs to assist their needs.

1.1 Motivation

Although English is known to be the most widely spoken language in the world, it is not with the most native speakers. English is merely the third largest language by number of native speakers, after Mandarin Chinese and Spanish. However, English

has the largest amount of non-native speakers (Exploredia, 2012). It is estimated that English has about 2 billion speakers, among which only 400 million are native (Crystal, 2003). In other words, more than one billion English speakers are non-native, which make a large community.

English is even more widely used on the internet. At an estimate, more than one half of all the content on the web is in English (W3Techs Web Technology Surveys, 2017). However, a strong disparity exists between the language distribution of web content and the representation of speakers of different languages among web users. While the majority of content on the web is in English, native English speakers account for only about one-fourth of all web users. NNES use internet search engines to carry out searches in a English for a variety of reasons: they may be looking for better content than what is available in their native language; they may be immigrants or international students in a new country, or travelers planning a trip to a country unknown.

Previous studies have identified NNES searchers to be a large user community (e.g., Kralisch & Mandl, 2006). When searching in a foreign language for content that was created in an unfamiliar culture, searchers face a myriad of problems: they may speak the language but not grasp the slight variations in meaning that will change their

search results drastically; they may not be aware of trustworthy sources; and they may face unfamiliar information architecture and design styles in websites.

1.2 Statement of Purpose

Though NNES online searchers account for a large community, current research provides insufficient understanding of their information seeking behavior. Even less is known about their interaction with online search engines, their preferred system functionalities and features, and the design guidelines to improve MILR systems and UIs to better accommodate their needs. The overall objective of this dissertation research is:

To extend current understanding of non-native English speakers' online search behavior by identifying and classifying their typical challenges and problems and addressing the causes and roots of these problems. In doing so, the goal is not to improve machine translation or retrieval technology, but to design for interaction and thus guide multilingual information retrieval systems and user interfaces to assist non-native English searchers' behavior and meet their needs.

1.3 Research Design and Study Procedure

The research design and study procedure are arranged based on the complexity of the binary nature of the research topic, as well as the research questions we aim to

address. The purpose of this research is to further understand the NNESs information behavior and, subsequently, explore design solutions to support them. Accordingly, the dissertation research is designed to be two stages with mixed research methods, an exploration and understanding stage and a design and evaluation stage. Both quantitative and qualitative data are collected and analyzed during the two stages. Our study bridges the gap between user studies of search behaviors and system design and development practices by closely coupling and linking the two.

1.3.1 Exploration and Understanding Stage

In the first stage, both quantitative and qualitative methods were utilized to extend the current understanding of NNESs search behavior in English. Two studies are carried out to explore and understand NNES search challenges, interactive behaviors, strategies, and entire search processes.

Findings in this stage shed light on behavioral patterns of language selection, search engine selection, query formation and reformulation, strategies, and browsing and filtering search engine result pages (SERPs). An iterative, or rather, spiral search process was observed and a user interaction model accommodating two query construction strategies was abstracted from these patterns.

1.3.2 System Design and Evaluation Stage

Based on findings of stage one, the second stage contains two studies on system design and evaluation. First, multiple UI prototypes were designed by the researcher. A viable UI prototype, TranSearch 1.0, was reached, developed, and tested through a user testing study. Next, based on the user feedback and the researcher's follow-up usability inspection, the prototype was redesigned to reach a more functional version and include more features, TranSearch 2.0. Finally, a user study was conducted with two purposes, evaluating TranSearch 2.0 and providing further redesign ideas and suggestions from users' perspective.

Both quantitative and qualitative data are collected and analyzed in the two studies of this stage. Results and findings not only provide information about the utility of the proposed solutions and design implications for MLIR systems and UIs to assist NNES users, but also further inform the model for NNESs information seeking behavior.

1.4 Dissertation Structure

This dissertation is organized as follows:

Chapter 2 examines three research areas related to our topic, Linguistics, Human Information Interaction (HII), and system design and development. The status of

these areas is summarized and my dissertation research is introduced by identifying our target audience and describing current research gap.

Based on the purpose of our research, Chapter 3 specifies the dissertation framework and methodology. Four sets of research questions are generated and introduced. Due to the complexity of the research questions, the entire study process is divided into two stages, consisting of four individual studies addressing one of each set of research questions. The mixed research methods and data collection techniques utilized are then explained. Quantitative and qualitative data collected are specified and data analysis methods are described.

The four studies are organized in the four following chapters, through Chapter 4 to 7. Though they have different focuses and address the four sets of research questions one at a time, the four studies are interrelated. Chap 4 and 5 address the goal of research stage one, that is to explore and extend the understanding of NNES searchers' behavior. The findings of stage one guide the studies in stage two, the goal of which is to explore and guide design practices of MLIR systems and UIs to assist NNES searchers' behavior and meet their needs. Finds of stage two verify and further inform those of stage one.

Specifically, Chapter 4 focuses on the behavioral differences when online users search in a non-native language. Particular attention is paid to query reformulation strategy, task performance, and user experience. Chapter 5 takes a step further and address what causes these behavioral differences, by studying NNES searchers' challenges and their root causes. The study takes a holistic view on the entire search process and abstracted a user interaction model to describe it. Two typical query formulation and reformulation strategies are identified and described.

Based on the findings of stage one, two studies are carried out during stage two and discussed in Chapter 6 and 7. The former, in Chapter 6, surveys and summarizes common design approaches for MLIR systems and UIs and introduces and describes our design process and prototype, TranSearch 1.0. A follow up user testing gathers user feedback and provides possible improvements to the prototype. In Chapter 7, a study aiming at extending the previous is conducted in two steps. First, based on the findings from the previous study, the prototype is redesigned and rebuilt by the researcher to reach a more functional version with more features, TranSearch 2.0, which will provide more insights than a prototype during evaluation. Subsequently, a user study is carried out 1) to evaluate the researcher redesign; 2) to engage users in further redesign. Results are discussed and the user interaction model abstracted in previous studies is reexamined. Final design implications are addressed.

Research methods, data collection and analysis, and results are discussed in each corresponding study chapter. Finally, Chapter 8 concludes the dissertation by summarizing the key findings and contributions of all the studies, discussing research limitations, and describing future work.

Chapter 2: Related Research

The aim of the dissertation research is to provide a better understanding of non-native English speakers (NNESs) search behavior and, subsequently, explore system solutions to assist them. This is a typical Human Centered Computing (HCC) and Human Computer Interaction (HCI) research topic and deals with human behaviors and computational systems. This chapter organizes our review of current research into two categories, human information behavior and system solution.

The research, inevitably, tackles human factors, search behavior, and online search systems. In this chapter, we examine the research literature from three disciplinary areas related to our research topic: Linguistics, Human Information Interaction (HII), and system design and development. Subsequently, we summarize the current status of the research related to our topic, identify the research gap, and introduce our research and study.

2.1 Target Audience

In this section, we identify and address our target audience, NNES online searchers. This specific group of Web users have two characteristics,

- They speak English but they are not native English speakers
- They conduct online searching in EFL or L2

To describe our target audience, we examine literatures from two areas accordingly, Linguistics and Web statistics and research. We select English as the foreign language that Web users search in because of the size of its speakers, the coverage, and the distribution.

2.1.1 Non-native English speakers

English as a global language (Crystal, 2003) is widely spoken and has often been referred to as a "world language" (Graddol, 1997), the lingua franca of the modern era. However, English is not the first language for the large proportion of its speakers. Linguistics researchers calculate that non-native speakers outnumber native speakers by a ratio of 3 to 1 (Crystal, 2003). It is estimated that English has 2 billion speakers, among which only 400 million are native speakers. In other words, over 1 billion English speakers are non-native (Exploredia, 2012).

While English is not an official language in most countries, it is currently the language most often taught as a foreign language. Correspondingly, Graddol (1997) differentiated three types of English speakers by relationship with the language. Figure 2-1 shows the three types of English speakers and the possible language shift among them.

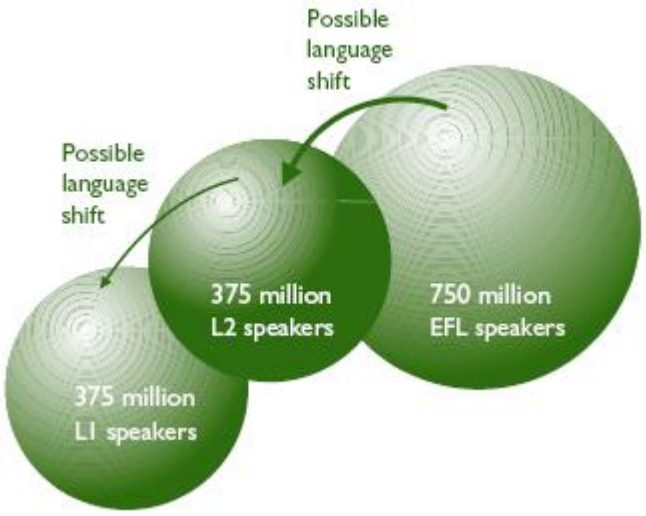


Figure 0-1. Three types of English speakers, adapted from Graddol (1997)

The three types of English speakers are first-language (L1) speakers, second-language (L2) speakers, and those who speak English as a foreign language (EFL). L1 speakers in his classification are "those for whom English is a first and often only Language". These native speakers live, for the most part, in countries in which the dominant culture is based around English. To distinguish between L2 and EFL, Graddol (1997) introduced geographic language usage areas. In EFL areas, such as China, English is used primarily for external (international) communication with speakers from other countries. While in an L2 area, such as South-east Asia, Africa, and the Caribbean, English is used for internal (intranational) communication.

The target audience of our research covers, regardless the geographic areas, all those English speakers whose native language is not English (NNESs). In other words, our

target audience speaks multiple languages and English is their non-native language. In a broader sense, our target audience is overlapping Graddol's (1997) L2 and EFL. The competence in English among our target audience varies from native-like fluency to extremely poor.

2.1.2 Web Content Language Inequalities

Inequalities in the availability of content in users' first languages exist on the internet (see Figure 2-2). Native English speakers make up 25.5% of all internet users (Internet World Statistics, 2017), while English content accounts for 51.8% of all content on the Web (W3Techs Web Technology Surveys, 2017).

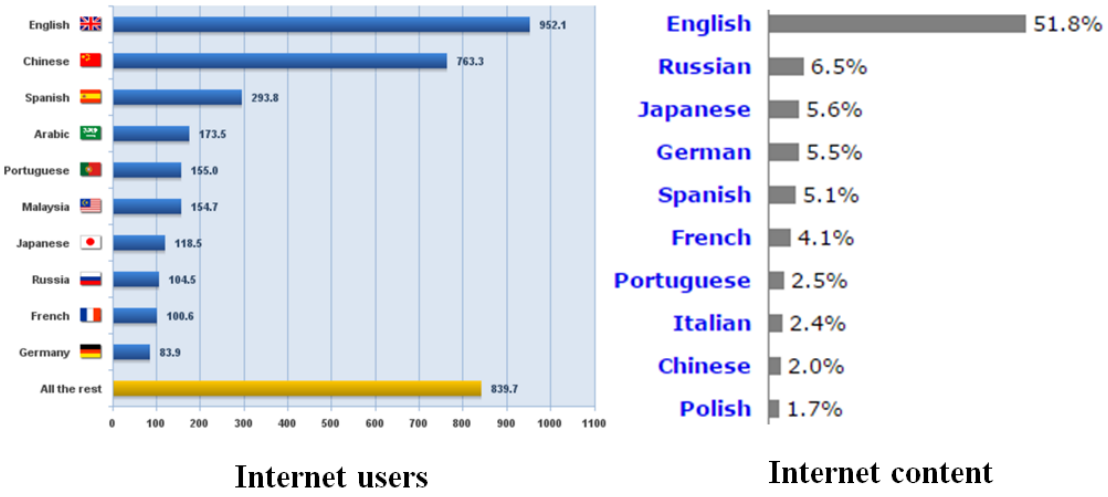
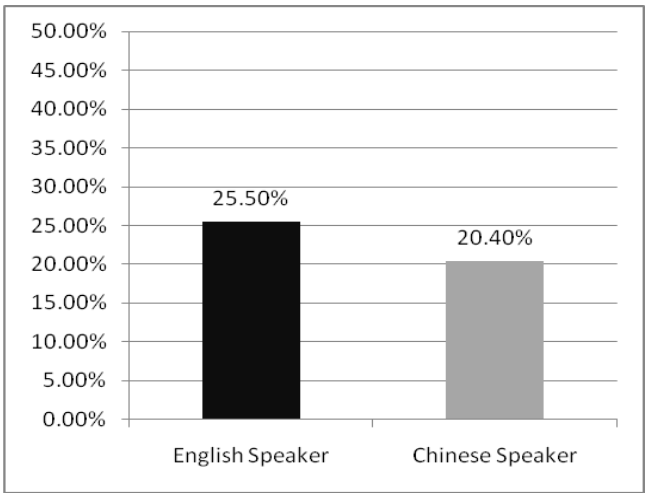


Figure 0-2. Internet users and content by languages, adapted from Internet World Statistics (2017) and W3Techs Web Technology Surveys (2017)

For an extreme example, Chinese first language speakers make up 20.4% of internet users (Internet World Statistics, 2017), while Chinese content accounts for only 2.0%

of Web content (W3Techs Web Technology Surveys, 2017). The disparity between the language distribution of Web content and the representation of speakers of different languages among Internet users is shown in Figure 2-2 and 2-3.

Panel A. **Percentage of Internet users**



Panel B. **Percentage of Web content**

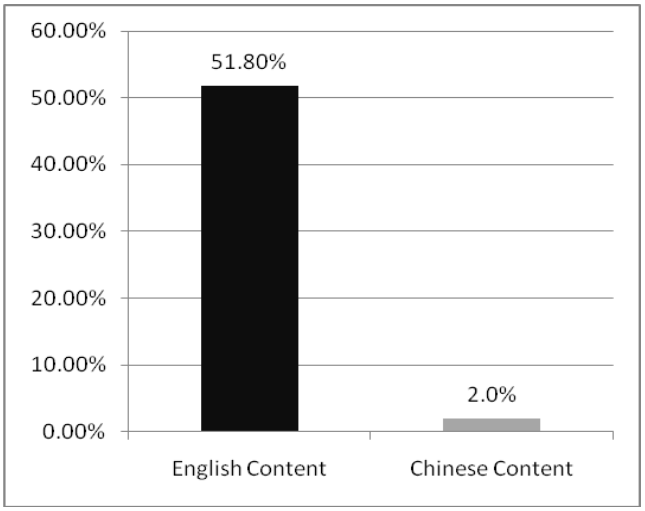


Figure 0-3. Percentage of Internet users and Web content: English vs. Chinese

The limited availability or low quality of content in other languages than English requires non-native English users to search in English to find content that does not exist in their native languages or that is of better quality in English than in their native languages.

2.1.3 NNES Online Searcher Community

The above Web content language inequality problem can impact a large user group, but estimating the exact size of the group is challenging. The estimate in this proposal was calculated by applying Internet penetration rates to the estimated number of English language speakers and learners around the world.

As discussed in the previous section, in 1997, it was estimated that English had over one billion non-native speakers (Crystal, 2003). In a study for the British Council, Graddol (1997) estimated the number of English language learners would grow to two billion in the next two decades. With the significant growth of English learners, there will be over two billion non-native English language speakers in about two decades. The largest portion of learners will be in Asia. According to the recent statistics in 2012, the Internet penetration rate in Asia is 27.2% (International Telecommunications Union/United Nations, 2012). Asia's Internet penetration rate is the second lowest in the world, only exceeding that of Africa. With Asia's current and relatively low penetration rate, 27.2% of two billion non-native English speakers will

result in 544 million potential non-native English language Internet users in the next two decades. This is a conservative estimate based on current Internet penetration rate for Asia which is sure to grow. This calculation also does not take into account the fact that Internet penetration rate is most likely much higher among English language speakers and learners, as they probably are more educated and have better access to internet and technology.

There are increasingly more people search online in a language, English in most cases, which is not their native languages due to the limited availability of content in underrepresented languages (Kralisch & Mandl, 2006). Berendt and Kralisch (2009) also found that underrepresentation of non-English languages is further strengthened by users' tendency to accept information in English. All in all, NNES online searchers accounted for and are forming a larger community.

2.2 Background

This sub section provides an introduction to the background of the related research. Previous research and development activities in the area of multilingual search usually focused on two areas: information seeking behavior (human-oriented) and system solution (system-oriented). Specifically, human-oriented research focused on gaining a better understanding of multilingual user search behavior and interaction with search tools and information on the web (e.g. Kralisch & Mandl, 2006; Berendt

& Kralisch, 2009; Reinecke & Gajos, 2014; Steichen & Freund, 2015). On the other hand, system-oriented research dealt with Information Retrieval (IR) tools and techniques, such as search query translation and disambiguation enhancement (e.g. Cao, Gao, Nie, & Bai, 2007; Gao et al., 2007; Magdy & Jones, 2011), search result content type and organization (e.g. Steichen, Ashman, & Wade, 2012), and user interface display personalization (e.g. Ghorab, Zhou, O'Connor, & Wade, 2013; Steichen, Ghorab, O'Connor, Lawless, & Wade, 2014).

Information Seeking, in general, is the entire process during which human seeks information. Case (2007) defined it to be "a conscious effort to acquire information in response to a need or gap in your knowledge". Human information need, information seeking, search process, and behavior, including the strategies employed when users are engaged in a search, are a complex cognitive process (Hearst, 2009). Numerous contributions have been made exploring this process (e.g., Belkin, Oddy, & Brooks, 1982; Sutcliffe & Ennis, 1998; Wilson, 1999) and there exists a large body of theories and models.

Similarly, the term of IR may have a very broad meaning. As exemplified by Manning, Raghavan, and Schütze (2008), simply getting a credit card out of your wallet so that you can type in the card number is a form of information retrieval. A rather general term of IR was defined as "finding material (usually documents) of an

unstructured nature (usually text) that satisfies an information need from within large collections" (Manning et al., 2008).

To narrow the topic down, in HCC discipline the information needs are usually supported by computational systems. Shneiderman, Byrd, and Croft (1997) defined information need in the context of HCC as "the perceived need for information that leads to someone using an information retrieval system in the first place". This definition addresses information need in terms of the search system and is constructed with IR systems in mind. In other words, HCI research studies the convergence of information seeking and IR, which is, to supports human-oriented information seeking process with technology-based IR systems.

There exist numerous theories and models. Wilson (1999) defines models of information behavior to be "statements, often in the form of diagrams, which attempt to describe an information seeking activity, the causes and consequences of that activity, or the relationships among stages in information seeking behavior". Those models within HCI are of more interest to us. Hearst (2009) compares and discusses the theoretical models of the search process that are most commonly referred to, the standard model, the cognitive model, the dynamic model, search as a sequence of stages, search as a strategic process, and sense making. All these theories and models

shed light on our research, however, none can be a perfect match and applies to our situation.

2.2.1 Information Seeking Behavior (Human-oriented)

This area of research focused on the study of multilingual Web users (e.g. Kralisch & Mandl, 2006; Berendt & Kralisch, 2009; Steichen, et al., 2014). Researchers were interested in understanding the reasons for utilizing multiple languages to search and the patterns of such search behaviors. The following three themes were found to be central to this avenue of research:

- Content Availability

Content available in many languages other than English was limited and of low perceived quality (Kralisch & Mandl, 2006; Aula & Kellar, 2009). Many searchers with sufficient English language skills searched for information in English to supplement their findings in other languages. Berendt and Kralisch (2009) even argued that underrepresentation of non-English languages was further strengthened by users' tendency to accept information in English.

- Language Proficiency

While polyglots would use any language to search in as long as they possessed sufficient proficiency in it (Steichen et al., 2014), searchers' varying levels of

language proficiency were found to impact their language selection and usage, search tools usage, search performance, and preference (Berendt & Kralisch, 2009; Marlow, Clough, Recuero, & Artiles, 2008).

- Other Contextual Factors

A series of contextual factors were also identified, analyzed, and proven to make a difference in multilingual searchers' language selection, usage, and preference. To name a few, domain of knowledge (Berendt & Kralisch, 2009; Clough & Eleta, 2010), domain of search topics (Steichen et al., 2014), and usage purpose (Steichen et al., 2014) individually or collectively had impacts on non-native speakers' search behavior.

These studies shed light on searching in non-native languages. However, they usually focused on language selection and factors impacting it, or the impact of a narrow set of factors on behavior. For example, Kralisch and Mandl (2006) used a user-centered approach to examine the first language's content availability and accessibility. They found that factors such as content-creation, link-setting, and link-following behavior will contribute to the under-representation of non-English languages on the web. Berendt and Kralisch (2009) in further studies found that under-representation of non-English languages is further strengthened by users' tendency to accept information in

English. They also identified a strong impact of both language proficiency and domain knowledge on the perceptions of utility and success of second language searching. First language and language proficiency had an impact on how users utilized language and query support tools (Marlow et al., 2008) and user interface preference (Reinecke & Gajos, 2014). Our study instead takes a more holistic and qualitative approach to better understand user behavior, search processes, and outcomes.

2.2.2 System Solution (System-oriented)

Bush (1945) popularized the idea of utilizing computers to store, search for, and retrieve relevant pieces of information. Automated IR systems have been introduced, developed, and improved ever since. It is commonly acknowledged that current IR systems and tools are usually technology-oriented and deal with algorisms. Thus, from its origin, IR systems face inevitable issues when people interact with them because human information seeking process is such a complex cognitive process (Hearst, 2009).

The search and interaction process between humans and computers was widely discussed by HII researchers (e.g., Salton, 1989; Shneiderman, Byrd, & Croft, 1998). Online search engines as IR applications are built on IR principles and rely heavily on

well-formed queries when searching for and retrieving information. Thus, they face challenges during the human computer interaction process.

One model drew our attention because it describes people's information need and interactive behavior with search engines. Broder (2002) pointed out in his "classic model for IR, augmented for the web" that people's web search process is a process consisting of a number of stages (see Figure 2-4). Firstly, a task or problem gives birth to an information need. Then people articulate the need or form a verbal description of that need. The subsequent process is an interaction cycle, which starts from formulating a query, followed by interacting with the search engine, examining retrieval results, and if necessary, reformulating the query and repeating the interaction cycle. Finally the process ends with a satisfactory result or user abandonment.

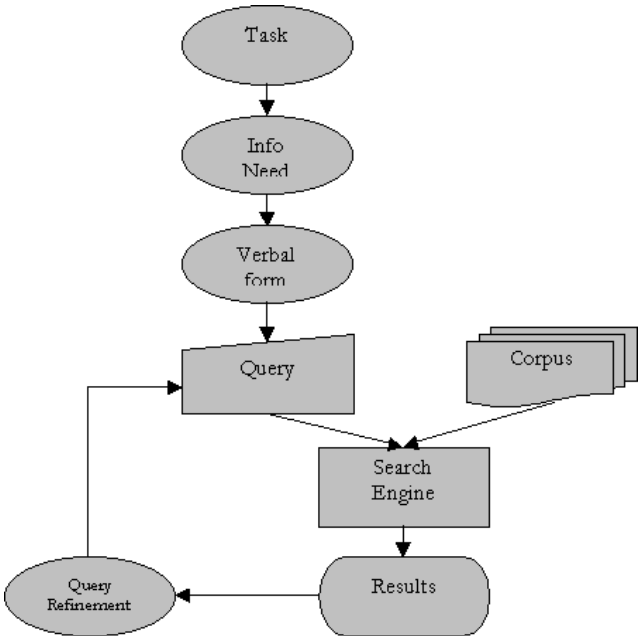


Figure 0-4. Broder's classic model for IR, adapted from Broder (2002)

As a standard model, it is rather universal and provides an understanding of people's searching behavior in general, searching in their native language. However, insufficient insights are given on search behavior when people search in multiple languages. When searching in multiple languages, contextual factors such as language proficiency, search purpose, and domain knowledge may individually or in interaction with each other change a searcher's behavior, strategy, and process. None of these is described in this model and this model may not work well in the context of searching in EFL or L2.

Context has been a popular topic in recent HCI and HII research. It is commonly acknowledged that understanding context is important for correctly interpreting user input and designing interaction (Berendt, 2007; Dourish, 2004). Language related

issues as well as other essential contextual factors should be incorporated into the models while designing IR systems.

IR research involving searching for and retrieving information in multiple languages was often referred to as “multilingual information access and retrieval” (e.g. Peters et al., 2005). Three overlapping terms were commonly used in this research area: Cross-language Information Retrieval (CLIR), Multilingual Information Retrieval (MLIR), and Multilingual Information Access (MLIA). During the interaction with search engines, CLIR consisted of entering queries in one language and retrieving relevant information in another (e.g. Grefenstette, 1998; Peters, 2001; Nie, 2010). MLIR was a broader term and embraced the concept of CLIR, because it dealt with "managing information access and discovery in multiple languages both monolingually and across languages"(Peters, Braschler, & Clough, 2012). MLIA was usually used in its broadest sense and addressed the problem of "accessing, querying, and retrieving information from collections in any language at any level of specificity" (Peters et al., 2012). We present the relationship and scope of these terms in Figure 2-5.

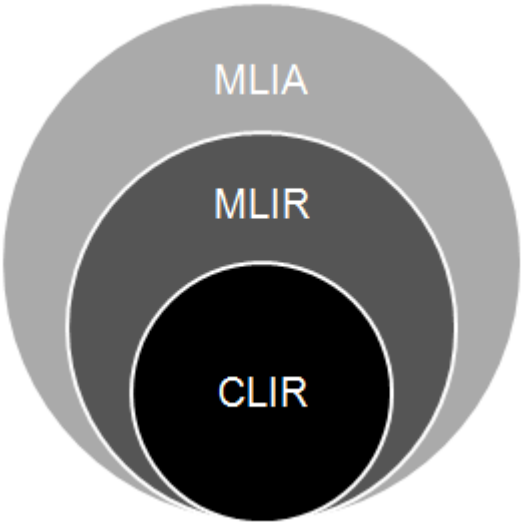


Figure 0-5. Relationship and scope of CLIR, MLIR, and MLIA

2.3 Current Focus and State

This sub section provides an outline of the current research state, contributions, and major findings of the two areas of interest to our research. Traditional laboratory IR theories and models are based on a test collection and test topics with binary relevance assessments. Performance of a system or a query is usually measured by recall and precision. Test queries are usually formulated automatically from test topics. Thus, traditional test environments and models are usually controlled to enable easy comparisons across various situations. However, these retrieval systems are designated for human users and such performance measures are insufficient to fully address the research question, especially in a cross-language context.

CLIR is, in the first place, built on user interaction to tackle the complexity of the issue, searching for texts written in foreign languages based on native language queries. Thus, it is generally accepted that CLIR tasks are split into two phases: query translation and the actually search. Researchers are increasingly focused on addressing HCI issues in these two phases. User studies with novel research design, instruments, methodologies, and performance measures started to emerge. In this subsection, we firstly review some current CLIR systems, subsequently look at performance measures, in terms of query translation and search user interface.

2.3.1 CLIR Online Search Systems

There has been research on developing functional cross-language online search systems, such as MULINEX (Capstick et al., 2000), MTIR (Bian & Chen, 2000), and Clarity (Petrelli et al., 2004; Petrelli, Levin, Beaulieu, & Sanderson, 2006). These systems represent current research and development practices and contributions in CLIR, but they differ in various focuses. MULINEX is a cross-language web search engine; MTIR is more of a query and document translation system; while as Clarity is a system focused more on user-centered design approaches.

With MULINEX, users type in a query and select the source language and the target language from its menu. The query is translated into the target language, then the translation with "back translation" are shown to the users. The back translations assist

the users in eliminating translations which are irrelevant to the intended meaning. The users select desired translation alternatives, and press the search button. The result list contains documents in selected languagges sorted by the estimated relevance. For each document, the language, title, URL, size, category and summary are displayed. The summary is presented in the document language, but it can be translated (Capstick et al., 2000).

MTIR is a Chinese-English IR system dealing with query translation and document translation. Bilingual dictionary and monolingual corpus-based approaches are adopted to select suitable translated query terms. A machine transliteration algorithm is introduced to resolve proper name searching (Bian & Chen, 2000). MTIR and MULINEX differ in many aspects: MTIR performs Chinese-English translation, while MULINEX has multiple language alternatives; MTIR is for both query and document translation, while MULINEX primarily translates queries and summaries of the document.

Clarity, on the other hand, takes a user-centered design approach and an interactive track to lead the development of a CLIR system. It shows how user interaction design evolves depending on the results of usability tests (Petrelli et al., 2004; Petrelli et al., 2006). Thus, their research provides a novel perspective of looking at CLIR system design, development, and evaluation.

Either from a system or algorism point of view or a user-centered perspective, these CLIR systems are all focused on system design, development, and evaluation. Research on the search process as a whole is insufficient and the understanding of search behaviors, patterns, and strategies guiding design efforts is lacking in the first place.

2.3.2 CLIR Query Translation

CLIR systems usually start with its first phase, query translation. Airio (2008) conducted a study to test whether query translation is beneficial in web retrieval. A total of 12 to 18 participants were recruited for each of the three language pairs, Finnish-Swedish, English-German, and Finnish-French. Each participant performed four retrieval tasks. The relative performance of the users' direct querying in the target language and the automatically translated (by either a dictionary-based system or a machine translation) target language queries is compared. The results differed depending on the language pair, but on average, the results of query-translation outperform the users' direct querying in the target language. Language proficiency also made a difference and query translation in web is beneficial especially for users with moderate and non-active language skills. As well, the dictionary coverage and quality of translation had an effect on the results (Airio, 2008).

MULINEX (Capstick et al., 2000) and MTIR (Bian & Chen, 2000) were based on different query translation algorisms, dictionary-based and corpus-based approaches. Both systems paid particular attention to the quality of query translation. MULINEX (Capstick et al., 2000) provided a tool, "Query Assistant", to allow users to choose appropriate translation. The back translations assisted the users in eliminating translations which were irrelevant to the intended meaning (see Figure 2-6).

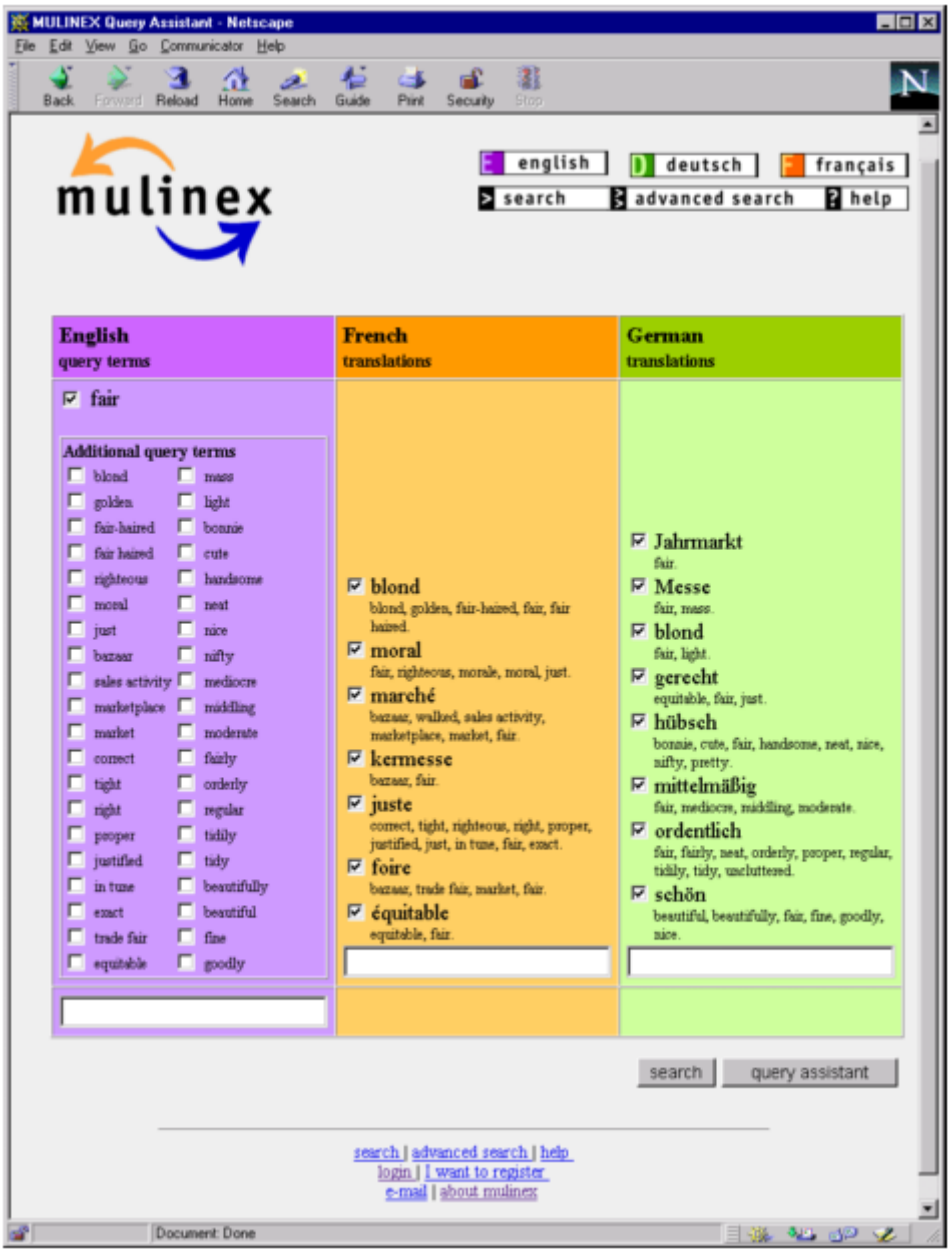


Figure 0-6. A query translation tool: Query Assistant of MULINEX, adapted from Capstick et al., 2000

The Query Assistant of MULINEX (Capstick et al., 2000) has its limitation, such as adding complexity and requiring more steps and user efforts. It may be useless, or even worse by confusing users, when users possess inadequate knowledge of the sensitive meaning of translated queries in the target language. Petrelli et al. (2006)

even argue in their research that users were not interested in, or were not able to, control the query translation steps. Though these results are arguable, they indicate that a simplified, lightweight, and straightforward solution is called on. Moreover, a systematic discussion of query formulation and reformulation strategy is lacking in query translation approaches.

2.3.3 CLIR Search User Interface

Although the underlying technology for achieving the two phases of CLIR, query translation and actual search process, is relatively well understood, the appropriate search result presentation is not. We argue there should be a third step in CLIR, result presentation, which we integrate in our search user interface discussion of this research.

Steichen and Freund (2015) presented a comparison of 5 different search interface designs for CLIR or multilingual search. The interfaces were analyzed and evaluated through a crowd-based experiment. Results indicated that "the common approach of interleaving multilingual results is in fact the least preferred, whereas single-page displays with clear language separation are most preferred". User proficiency and search content type also played a role in user preferences and different interfaces elicited different user behaviors. Figure 2-7 shows one of the CLIR or multilingual search user interfaces analyzed and evaluated.

Figure 0-7. A CLIR user interface, adapted from Steichen and Freund, 2015

Though their research took a first step by addressing user preferences in CLIR search result presentation, lots can be improved in designing a better search user interface. For example, their automatically translated query was fixed and could not be modified by users when it was not accurate or relevant. With the same query and translation, there was no flexibility to implement a query reformulation. Their study instrument and environment were very controlled. The queries utilized during their study were predetermined. Participants were not allowed to enter their own queries. There was only one active view and users did not have the flexibility to switch views or customize the interface. There was no flexibility of choose various search engines across languages. Microsoft Bing API was the sole search engine for all languages. Search engine performance may vary significantly when searches are conducted in different languages.

2.4 Summary and Research Gap

Our research is built on the current work and focused on the sub-area of HII, interactive information retrieval in English as a non-native language.

2.4.1 A Summary of Existing Research

Current research and development activities available in the area of non-native language searching focus on two sides. One is the human-oriented Information Seeking side. This group of research focuses on user study and understanding. But current research in this category usually focuses on why users pick a language to search in and browse information. The understanding on how they actually search and browse and the entire search process is in sufficient.

The other side is looking at system-oriented solutions. This area of research focused on enabling technologies such as tools and interfaces to support query formulation, translation, results merging, summarization, and presentation (e.g. Gao et al., 2007; Amato, Cigarrán, Gonzalo, Peters, & Savino, 2007). Though fruitful, this group of research was focused on backend translation and retrieval technologies and less attention was paid to user interaction issues of these systems and tools. More recently, researchers started to address these issues. For example, Steichen and Freund (2015) discussed user preferences in the interfaces presenting multilingual search results.

However, studies on NNESs search process, as a whole, and user experience (UX), user preference, and user satisfaction issues during it, are still less common.

2.4.2 Research Gap

A gap between the above-mentioned two groups of research exists. Our study bridges the gap between user studies of search behaviors and system development by closely coupling and linking the two. A better understanding of NNESs search behaviors, processes, and strategies provides guidelines for the design of systems and tools to assist these users.

Current research usually treats people who speak multiple languages as a whole group and focuses on their language choices and preferences (e.g. Kralisch & Mandl, 2006; Marlow et al., 2008; Aula & Kellar, 2009; Berendt & Kralisch, 2009; Clough & Eleta, 2010; Steichen et al., 2014). In other words, current research is, in terms of research scope, rather general and scarcely language specific. In terms of the nature of research question, current researchers usually focus on "why" users choose a language to search in. Our study, in this regard, varies in the following aspects:

We target NNES online searchers as a particular user community because of its size, coverage, and distribution. Our focus is not on the reasons searchers choose a language, but on "how" they actually search given a search language, English, is

selected. A research towards a better understanding of NNES searchers' behavior is lacking. Without this study stage, the interpretation of multilingual searchers' behavior is insufficient and it is too soon to jump to a conclusion on system suggestions.

Chapter 3: Framework and Methodology

This chapter provides an overview of the dissertation research framework and methodology. Firstly, towards bridging the current research gap, the dissertation research is introduced. Then, based on the purpose of the research, four sets of research questions are generated and described. Around these questions, research design and process are specified. Four studies are carried out in two stages due to the complexity of the research. These studies are described, and data collection techniques and analysis approaches are clarified.

3.1 Dissertation Research: An Introduction

Our study bridges the gap between user studies of search behaviors and system development by closely coupling and linking the two. Based on the nature of the research topic and questions, the dissertation study is carried out in two stages. The first stage consists of two exploratory qualitative studies and aims to provide a better understanding of NNESs search behaviors, processes, and strategies. Based on the findings, the second stage contains two studies on system design and evaluation and provides guidelines for MLIR systems and UIs to assist NNES users.

On multilingual searchers' behavior, our study extends previous research by providing further understanding and reasoning. Previous research usually focuses on language selection and impacting factors. Our study varies in research interests and moves a

step further. Though in our study we gather data on the causes participants choose English to search in and confirm the previous study findings, our research interests do not lie in why they choose English, but in how they actually search. To be specific, we focus on NNES searchers' challenges, behaviors, and the whole search process.

From system-oriented side, previous research was focused on backend translation and retrieval technologies of MLIR systems. More recently, researchers started to address user interaction and preference issues of these systems and tools. For example, a study by Steichen et al. (2014) presented findings from their survey and discussed the factors impacting multilingual web users' language choice and frequency. Based on these findings, they provided suggestions towards "Personalized Multilingual Information Access" systems. Similarly, Steichen and Freund (2015) compared various UI types and provided data on which one was mostly preferred by their participants. However, the studies provided insufficient knowledge on why users have such preferences, how exactly users interact with the systems, and what critical functionalities and features are from the user’s perspective. Our studies focus on user preference reasoning, NNESs system interaction process, as a whole, and users’ feedback on system usability.

The purpose of this research is to further understand the NNESs information behavior and, subsequently, explore design solutions to support them. Due to the complexity of

the binary nature of the research, the study process is carried out in two stages. In the first stage, both quantitative and qualitative methods were utilized to extend the current understanding of NNESs search behavior in English. Findings in this stage shed light on behavioral patterns of language selection, search engine selection, query formation and reformulation, strategies, and browsing and filtering search engine result pages (SERPs). An iterative, or rather, spiral search process was observed and a user interaction model accommodating two query construction strategies was abstracted from these patterns.

Based on the previous study findings, the second research stage took place. First, multiple UI prototypes were designed by the researcher. A viable UI prototype, TranSearch 1.0, was reached, developed, and tested through a user testing study. Next, based on the user feedback and the researcher's follow-up usability inspection, the prototype was redesigned to reach a more functional version and include more features, TranSearch 2.0. Finally, a user study was conducted with two purposes, evaluating TranSearch 2.0 and providing further redesign ideas and suggestions from users' perspective. Results and findings not only provide information about the utility of the proposed solutions and design implications, but also further inform the model for NNESs information seeking behavior.

3.2 Research Questions

The purpose of this dissertation research is to further understand the NNESs information behavior and, subsequently, explore design solutions to support them. Specifically, the study aims 1) to identify and classify foreign language searchers' typical challenges and problems, 2) to describe the causes and roots of these problems, 3) to abstract an interaction model and to prototype based on the model, 4) to provide design implications of MLIR systems through the findings.

The first stage of research aims at further understanding NNES search behavior by answering the following research questions:

- Research question set 1 – Addressed by study 1

  RQ1.1: What are the behavioral differences when online users search in a non-native language, specifically in terms of query reformulation strategy, task performance, and user experience?

- Research question set 2 – Addressed by study 2

  RQ2.1: How do NNESs search for information in English (search behavior and user experience)?

RQ2.2: What are the typical challenges of NNESs? What are the causes and roots of the identified challenges?

RQ2.3: What user interaction model can be abstracted to describe the typical search processes and characteristics of NNESs?

After addressing these research questions, the second stage of research focuses on system design and evaluation. The following research questions are discussed:

- Research question set 3 – Addressed by study 3

RQ3.1: What UI and tools can be created to support NNESs based on their challenges and search process (identified from our previous studies)?

RQ3.2: What are the strength and weakness of the UI and tools designed? Any improvement to the prototype?

- Research question set 4 – Addressed by study 4

RQ4.1: What do TranSearch 2.0 evaluation results indicate, in terms of system usability and user experience?

RQ4.2: Does TranSearch 2.0 accommodate the user interaction model and assist in user challenges and query formulation and reformulation strategies identified in previous studies?

RQ4.3: What are the user preferences and why? What are the crucial MLIR functionalities and features and why? What are the design implications can be provided to MLIR systems and UIs?

3.3 Research Design: Mixed Methods

Based on the research questions we are aiming to address, our research design involves mixed methods, both quantitative and qualitative. It is not uncommon to see mixed research methods in HCI research, especially when research questions involve both human and system. This section provides a general introduction to our research methods and guidelines. Individual study describes detailed methods and processes.

Quantitative research concerns with quantities or quantifiable data (sometimes referred to as objective properties). Researchers usually analyze the data with the help of statistics and are hoping the numbers will yield an unbiased result that can be generalized to larger population or establish the existence of associative or causal relationship between variables. On the other hand, qualitative research deals with qualities (sometimes referred to as subjective properties). Qualitative method asks broad questions and collects data in the form of words, images, videos, and etc. Researchers look for themes and describe in detail the themes and patterns exclusive to that set of participants. Qualitative research is often used as an approach to perform

exploratory research, further elucidate quantitative findings, investigate well-defined questions that involve human behavior, and etc.

In our study, qualitative and quantitative methods are used in combination, because our work in HII research deals with objective artifacts as well as people. Mixed methods are particularly useful to better understand human behavior, because "human activity is highly flexible, nuanced, and contextualized and computational entities such as information sharing, roles, and social norms need to be similarly flexible, nuanced, and contextualized" (Ackerman, 2000) and consequently, IS and HCC phenomena are too complex to capture with merely numbers.

Specifically, focus groups, interviews, search diaries, questionnaires, task observation, and activity logs are utilized to facilitate triangulation, gathering and analyzing data of multiple types, from multiple sources, and using multiple techniques. It is to get evidence from multiple sources in multiple ways. This is built on the fact that one data set trades off the limitations of another.

Single method, such as questionnaires, may suffer from poor question construction yielding inaccurate answers. More commonly researchers usually face no or unequal response. Some answers are inherently based on participants' understandings or out of their memories. Thus, errors or inaccurate answers inevitably occur. However, if we

make use of other techniques altogether, such as observing the participants while they are performing the tasks and making records, errors and inaccuracy will be decreased or explained by comparing different sets of data.

Both quantitative and qualitative data are collected in our studies. Quantitative data, such as search task time and likert scales, are collected and analyzed to measure user performance. Statistical analysis (ANOVA) is conducted on the data to identify statistically significant effects. To address our research questions, we collected and analyzed search behavior metrics, to name a few: searching time, number of new queries, number and types of query reformulations, number of search engine results pages (SERP) viewed, number of websites opened, number of links clicked within a website opened, likert scale evaluation.

Qualitative methods, on the other hand, are utilized in order to truly understand the roots of searching problems and not simply to observe their occurrence. Qualitative methods allow the researchers to immerse themselves in the experience of the participants and get a uniquely intimate understanding of the participants' perceptions and feelings. To observe natural behaviors, we allow participants to choose their own search tasks other than the prescribed ones. When searchers recreate their own searching in the demonstration exercise, more natural behaviors are observed.

3.4 Research Process and Data Collection: Four Studies in Two Stages

After research purposes, questions, and methods are finalized, four studies in two stages are carried out. In this section, the general research process is introduced (see Figure 3-1) and the individual studies are described. Detailed description is provided on participant recruiting, data types, and data collection techniques and procedures.

- Stage 1: To better understand NNES searchers' information needs and their interactive behaviors with online search engines

  Two studies are carried out with the focus on NNESs search challenges, behaviors, processes, query strategies, and user interaction model.

- Stage 2: To design, prototype, and evaluate a system solution

  Two studies focus on user preference reasoning, NNESs system interaction process, as a whole, and users' feedback on system usability, crucial functionalities, and features.

Figure 3-1 displays the introduction to the two stages and four studies, work breakdown structure (WBS), data collection and analysis timeline.

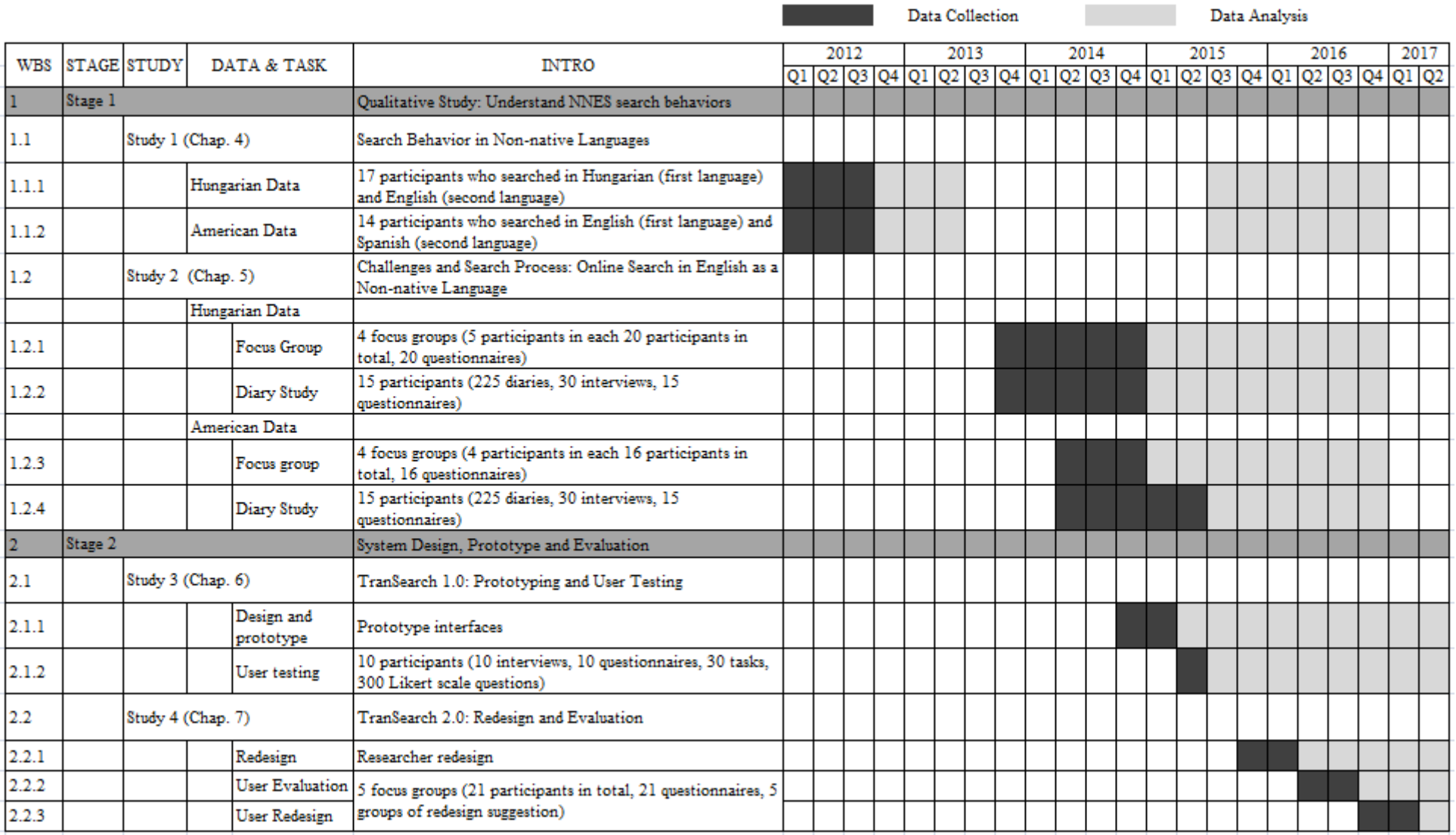


Figure 0-1. Research process, data collection and analysis timeline

3.4.1 Study 1 - Search Behavior in Non-native Languages

In this study, we focused on search behavioral differences in non-native languages. Attention was paid to how searchers picked keywords and conducted reformulation, as well as their searching strategies. Quantitative data on searcher performance such as searching time, results looked at on SERPs, websites click-through, and patterns were collected. Qualitative data on satisfaction, preference, and experience were gathered as well to analyze and uncover patterns. Two groups of participants were recruited:

- In Hungary: 17 participants who searched in Hungarian (native language) and English (second language)
- In the US: 14 participants who searched in English (native language) and Spanish (second language)

The participant selection criteria were that their foreign language proficiency was at a conversational level and they had conducted online search in that language. Native and foreign language search behavior data were collected in two countries, the US and Hungary, to facilitate comparison and diversity. The sessions in the two countries followed the same protocol. First, demographic information was collected on age,

gender, cultural background, and computer/Web experience. Next, the participants were asked to perform four information seeking tasks. They were asked to bookmark the result websites on which they found their answers.

Three of the tasks were prescribed and the forth was flexible to capture natural searching. The first task was a known-item information-seeking task. Participants had to find the CEO of IBM. This task served as a warm-up period for the participants and the answer was easy to find. The second and third tasks were subject driven medium-complexity tasks on travel planning for a family of four. Both participant groups received one familiar travel destination from their own geographical vicinity and the other from a foreign country. They were asked to search in their native language for the local destination and in the foreign language for the destination in the other country. The fourth one was defined in collaboration with the participants to reflect their own natural searching.

Both the native and foreign-language tasks followed the same template: "One of your friends spends his/her holiday in a city (destination varied based on the participant group as shown in Table 3-1) with his/her spouse and children ages six and eight…… He/she asks you to make some suggestions on what they should do on those two days……"

	Native Language	Foreign Language
US	Ocean City, MD, USA	Sucre, Bolivia
Hungary	Eger, Hungary	Ocean City, MD, USA

Table 3-1. Destinations for the native and foreign language tasks

While the participants were searching, their activities were logged. Video and audio data were recorded, along with eye movement data. At the end of the session, a semi-structured interview was conducted by the researchers with the participants about their general credibility evaluation process, as well as their searching experiences in different languages.

3.4.2 Study 2 - Challenges and Search Process: Online Search in English as a Non-native Language

This study focused on NNES searchers' challenges and the causes. Particular attention was paid to the entire search process, query formulation and reformulation strategies, and user interaction model. Two groups of NNES searchers were selected, native Hungarian participants (in Hungary) and native Chinese participants (in the US). Both groups comprised college and university students.

- New Chinese students at UMBC who recently arrived in the US (to minimize the impact of immersion)

- Hungarian students at the Budapest University of Technology and Economics (BUTE)

The participants were selected based on two criteria. First, they all spoke English at a conversational level. Second, they all regularly searched in English both to support their studies and their non-academic interests. These two criteria qualified them for the purposes of this study. They are all typical NNESs with a demand to search in English. However, they differ significantly in their cultural backgrounds and their exposure to and relationship with an English-speaking culture.

The nature of our research questions was seeking to explore phenomena. Thus, qualitative research methods were utilized and qualitative data were collected in order to truly understand the causes of challenges NNESs faced while searching and not simply to observe their occurrence. The research was designed as a focus group study and a diary study. The studies were conducted in two countries, Hungary and the US. The sessions in the two countries followed the same protocol.

For the focus group study, a total of thirty-six participants were recruited to complete eight focus group discussion sessions (20 native Hungarian participants in 4 groups in Hungary and 16 native Chinese participants in 4 groups in the US). For the diary study, a total of thirty participants were recruited (15 native Hungarian participants in Hungary and 15 native Chinese participants in the US). Each diary study participant completed fifteen diary entries of English searches and two separate interview sessions. Table 3-2 shows the arrangement of the two studies and data collected in the two countries.

Hungarian Data		
	Focus Group	4 focus groups (5 participants in each 20 participants in total, 20 questionnaires)
	Diary Study	15 participants (225 diaries, 30 interviews, 15 questionnaires)
American Data		
	Focus group	4 focus groups (4 participants in each 16 participants in total, 16 questionnaires)
	Diary Study	15 participants (225 diaries, 30 interviews, 15 questionnaires)

Table 3-2. Study arrangement and data collection

Specifically, data were collected through semi-structured focus group discussion, search diaries and follow-up interviews, and structured questionnaires with open-ended questions. The study processes are specified as follows:

- Focus group study

Four or five participants and one or two investigators engaged in each semi-structured focus group discussion. The native language (Hungarian or Chinese) of the researchers was the same as that of the participants and the focus groups were conducted in the native language of the participants. This ensured more openness and better rapport between the participants and the researchers, as well as provided more accurate cultural interpretation of the data during analysis. The study session, which took approximately one hour, was conducted in a user study laboratory and was video and audio recorded.

The investigators developed and used a discussion facilitating guide, which contained a list of open-ended questions and topics to be covered during the conversation, in a particular order (see Appendix 1: Focus group facilitating guide). The discussion followed the guide, but was able to follow topical trajectories in the conversation that might have strayed from the guide, if appropriate. New, usually follow-up and clarification, questions were added when necessary. Open-ended questions allowed participants to respond in their own words, rather than forcing them to pick from fixed responses as quantitative methods do (Mack, Woodsong, MacQueen, Guest, & Namey, 2005). The questions concerned search topics in all languages, strategies, query formulation, tools, and challenges (see Table 3-3 for sample questions).

Focus group, discussion guide
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 Chinese/Hungarian students who have recently started searching for information in English (e.g. incoming first year university students)?
Focus group, a follow-up question example
1. How do you come up with your search expressions while searching in English? Why?

Table 3-3. Sample questions in focus groups

At the end of the discussion, the participants were given questionnaires. After completing the questionnaires, they were given an opportunity to ask questions and make comments. They were then paid and dismissed.

- Diary study

To triangulate the focus groups, a separate diary study was designed to collect everyday life searching behavior data. In study 1, participants followed prescribed

searching tasks, which made it hard for the researchers to observe natural behaviors. In study 2, tasks were not prescribed and when searchers recreated their own searching in the demonstration exercise, more natural behaviors were observed. Thus, study 2 was not a lab study and no searching tasks or topics were assigned. Participants selected their own searching topics in English and could perform the tasks anywhere and anytime, as long as they could upload the diary entries to designated online folders (see Appendix 2: Diary study entry format). They were asked to create written notes and add screenshots. By doing so, we were trying to capture their most natural search behaviors in English.

The diary study participants were interviewed in two separate sessions (one in the middle and the other at the end of the diary study process) to reproduce the searching context, situation, behavior, and process. Similar to the focus group discussion facilitating guide, an interview guide was followed (see Appendix 3: Diary study interview guide). During the interviews, the investigator was also able to follow up topical trajectories in the conversation. Through all these instruments, we were trying to uncover search behavior differences, search challenges, and processes the participants had encountered and typical patterns of their behavior. The diary study participants were different than those who participated in the focus groups.

- Questionnaire

Structured questionnaires were used at the end of the focus group and diary studies to gather demographic information about participants, such as age, gender, and levels of education. However, open-ended questions were also included to gather qualitative data and triangulate the data collected through focus group and diary studies (see Table 3-4 for some sample questions).

Questionnaire, open-ended questions
1. Please list at least three topics for which you search in your native language.
2. Please list at least three topics for which you search in English.
3. What are the problems or difficulties arise from searching in English?

Table 3-4. Sample questions in questionnaires

3.4.3 Study 3 - TranSearch 1.0: Prototyping and User Study

This study focused on design and user study of a prototype, TranSearch, a MLIR system and UI that allowed the user to search in two languages at the same time, English and Chinese for the current prototype version. The study firstly reviewed and summarized previous work related to MLIR research and system development practices. Then, the prototype interface, TranSearch, was designed and developed

based on the findings of the studies of research stage one. The entire design process of TranSearch and how the UI prototype worked were explained.

After the prototype was designed and developed, a user study was carried out to evaluate it. Ten NNES participants were invited to conduct the user study sessions individually. The participants were selected based on two criteria. First, they were native Chinese speakers and spoke English at a conversational level, since the current version of TranSearch only supported Chinese and English language pair. Second, they regularly searched in English to support both their studies and non-academic interests. The two criteria guaranteed that TranSearch prototype was evaluated in a meaningful way.

Each user study session took approximately one hour. Participants were first asked to perform search tasks using TranSearch. Five tasks were initially devised in a pilot study (see Appendix 4 for the initial five search tasks), of which three (T1, T3, and T4 in the initial list) were selected to shorten the study session duration. The three selected tasks were typical search tasks that encouraged users to search in English, a location based restaurant recommendation task, a medical terminology lookup, and an English grammar and usage check (See Table 3-5). A word file was provided for

them to log their answers and the URL (URLs) of the website (websites) from which they found their answers.

Search Tasks
T1. One of your friends spends his/her holiday in New York City, USA with his/her spouse and children ages six and eight. He/she asks you to make suggestions on a Chinese restaurant and an American restaurant. Please list reasons you recommend.
T2. What is "Glucosamine" and what does it do?
T3. The usage of the most common prepositions of location: in, on, at.

Table 3-5. TranSearch user study search tasks

In a usability study lab, each study session engaged one participant who was equipped with a computer and was asked to perform the search tasks, using TranSearch 1.0 through a web browser (Chrome). At the beginning of the session, the investigator purposely did not explain how to use the TranSearch 1.0 interface to see how intuitive the interface was. When the participants were performing the tasks, the investigator observed participants' search behaviors and took detailed notes. If the participants had questions, the investigator would answer them. Otherwise, the participants worked on the tasks independently and without interruption.

Upon completion of search tasks and answer logs, a semi-structured interview was conducted. At the end of the session, questionnaires were used to gather demographic

information as well as complement the data collected from investigator observation, notes, and interviews. The following qualitative methods were utilized to gather data on prototype strengths, weaknesses, and improvement opportunities:

- Search task observation and investigator notes: The participants were asked to perform three searching tasks with TranSearch. Their behavior/screen were recorded and their eye movement data were collected

- Contextual semi-structured interviews: Each interview consisted of three parts. First, participants were asked to perform a demonstration for one of the search tasks they conducted (participants selected one out of the three tasks at their will). The purpose was to complement their search task logs and answers, as well as have them self-report how exactly they used TranSearch. Second, follow-up questions were asked about what they liked and disliked as the strengths and weaknesses of the prototype. The last part of the interviews was to ask broader questions on their suggestions to improve TranSearch

- Questionnaires: After the interviews were completed, the participants were asked to fill out a questionnaire containing demographic information (such as age and gender) and questions triangulating the interviews

3.4.4 Study 4 - TranSearch 2.0: Redesign and Evaluation

This study extended the previous and was carried out in two steps. Firstly, based on the previous study findings, TranSearch 1.0 was redesigned and rebuilt by the researcher, from a prototype to a more functional system and UI with new features, TranSearch 2.0. A query corpus, other than translation corpus, was introduced that enabled query recommendation and image query linking. The system enhanced user interaction during query construction phase and accommodated user preferences of search result displays.

Subsequently, a user study was carried out with two research purposes: 1) To evaluate the new functionalities and features of TranSearch 2.0; 2) To explore further redesign options and general user interaction needs of MLIR/CLIR systems. Thus, the user study was divided into two consecutive sessions. The first session was to evaluate the new features of the researcher redesign version of TranSearch with qualitative research methods. The second session was to engage participants to further redesign TranSearch and provide suggestions for the general design of user interaction of CLIR/MLIR systems.

Different than the previous, this study focused on the newly introduced query coups and its enabling features, user preference reasoning, and user redesign ideas

informing critical functionalities and features a general MLIR system and UI should possess. Due to the different research focus, instead of interviews utilized in study 3, focus groups were selected in this study because it was good for the redesign section in generating more ideas than individual interviews through the interaction of the participants. Moreover, this study focused more on exploration of general CLIR/MLIR interaction design needs, rather than solely system usability, and thus the focus group method was beneficial.

A total of twenty one participants took part in five focus groups, four of which were native Chinese speakers and one was other language speakers. Comparing to study 3, TranSearch 2.0 supported multiple languages and the participation of other native language speakers increased diversity of the participant sample. Two qualifying criteria were utilized for participant screening. The first criterion was that the participants' English (or another foreign language for native English participants) proficiency had to be at a conversational level. The second was that they had previous search experience in English or another foreign language so that they processed familiarity with English or another foreign language online search systems. Similar to the previous study, these criteria guaranteed the new features of TranSearch 2.0 were

evaluated in a meaningful way, as well as the quality of data of experienced users' redesign ideas and general system suggestions.

The two consecutive study sessions took approximately one hour and a half. Specifically, they were organized as follows:

- Evaluation session

Participants sat around a round table in a usability lab, each equipped with his/her own computer and asked to perform three search tasks individually and using TranSearch 2.0 as it was available to them over the web. The tasks were the same as those in the previous study and the same word file was provided for them to log their answers and the URL (URLs) of the website (websites) from which they found their answers. When they were performing the tasks, no instructions were given on the usage of TranSearch 2.0 interface to see how intuitive the interface was. Unless there were questions, the participants worked on the tasks individually without interruption. The investigator walked around the table, observe the participants' search behavior, and took notes. Collaboration was not encouraged among the participants so that each participant would try the UI independently and form their own opinion before the interactive group discussion.

Upon completion of search tasks and answer logs, a guided focus group discussion was conducted. The discussion started with how participants used TranSearch 2.0 by having each of them briefly describe a specific task they conducted (participants selected one out of the three tasks at their will). Follow-up questions were asked about their experiences and preferences of the new features and functionalities of TranSearch 2.0.

- User redesign and general suggestion session

Right after the evaluation session, the same group of participants worked on a redesign of TranSearch 2.0. The interactive setting helped brain-storming and generating more ideas through interaction among participants than individual interviews (see Figure 3-2).

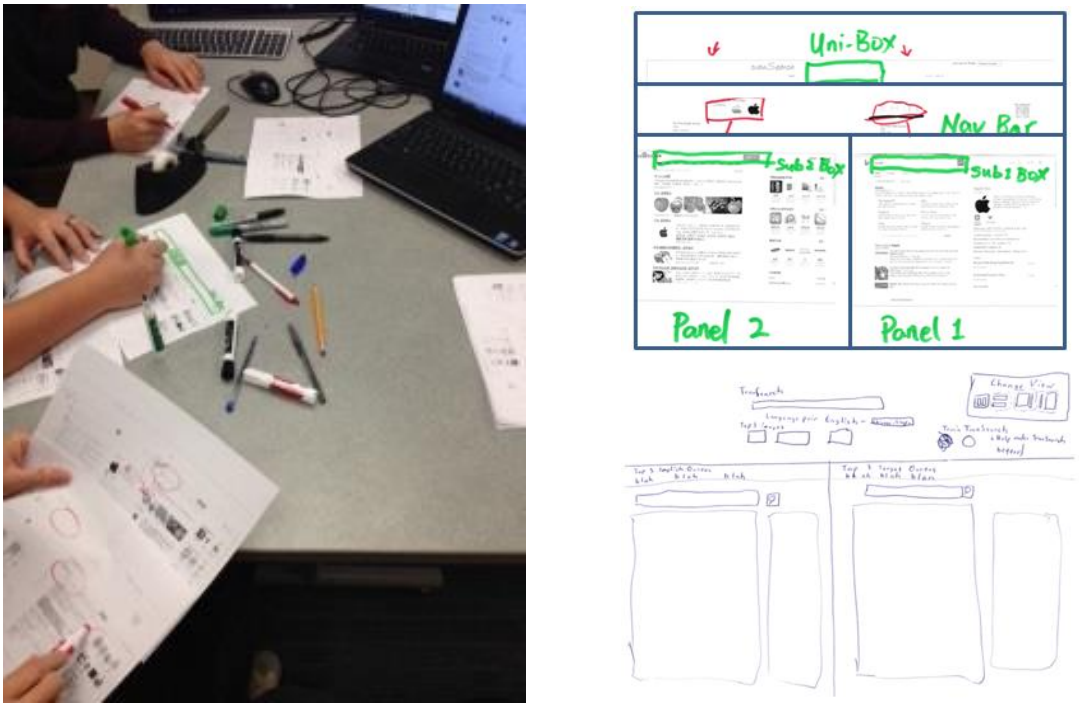


Figure 3-2. User redesign

Since each participant already interacted with TranSearch 2.0 without given instructions, it was hoped that this would create a realistic use scenario and each formed his/her own opinion. The redesign process started with a group discussion of their usage and experience of the UI. Next the participants verbally brain stormed their redesign ideas. Particular attention was paid to query translation, reformulation strategies, SERP comparison and contrast, and UI layouts. At last, the participants were provided with a set of printed screenshots of TranSearch 2.0 UI, together with markers, scissors, and glue sticks. Individually, the participants applied their ideas to those printed pages and created new designs by drawing, cutting, and pasting pieces

of the printed UIs. They shared their new designs with the investigator and other participants by a visual demonstration of their redesigns of the TranSearch UI (Figure 3-2). Participants commented on one another's design ideas.

The last part of the focus groups concerned the participants' general requirements and recommendations for CLIR/MLIR systems, as well as their deemed crucial functionalities and features. The data informed whether TranSearch 2.0 met their behavior and needs, as well as provided input on further design implications. At the end of the focus groups, questionnaires were used to gather demographic information, as well as complement the data collected from the focus groups.

3.5 Data Analysis

In the two research stages, a large amount of data were collected in both qualitative and quantitative format. Data analysis used to address the research questions inevitably involved both qualitative and quantitative methods. This section introduces the general analysis methods and guidelines. Detailed analysis of individual studies followed these methods and processes.

Specifically, in the first stage, quantitative analysis was conducted on searchers' behavioral differences, such as search task completion time, number of reformulation, and number of website viewed. The statistically significant factors were discussed.

On the other hand, qualitative analysis was also utilized, which allowed us to gain a deep understanding of the user behaviors as well as motivations behind the behaviors, especially particular user challenges and causes. The findings of this stage guided and informed the following design process. In the second stage, both qualitative and quantitative analysis methods were used in the user study and final evaluation in order to measure user performance and experience, reach design implications, and further inform the model of interaction behavior.

3.5.1 Quantitative Data Analysis

Quantitative demographics data were presented in each study. Other quantitative data, such as search task time and likert scales, were collected and analyzed to measure user performance as well as user experience. Statistical analysis, specifically a single factor ANOVA, was conducted on the data to identify statistically significant effects. To address our research questions, we collected and analyzed the following search behavior metrics:

- Searching time
- Number of new queries
- Number and types of query reformulations

- Number of SERPs viewed
- Number of websites opened
- Number of links clicked within a website opened
- Likert scale evaluation

Quantitative analysis methods were used in the second research stage for user study and final evaluation in order to measure user performance, reach design implications, and further inform the model of information behavior.

3.5.2 Qualitative Data Analysis

Focus groups and interviews were typical qualitative data collection techniques. They were voice recorded in the studies and transcribed subsequently. Depending on the individual study, the transcripts were in Chinese or Hungarian. The coding process were executed in Chinese or Hungarian, as well. Researchers who were native in Chinese or Hungarian carried out the analysis. Microsoft word and excel were used during the coding.

Investigators took notes during each focus group and interview. During the coding process, the detailed investigator notes were firstly coded to identify preliminary

concepts of interest. Next, the transcripts were coded through an inductive and iterative qualitative data analysis process. In the initial open coding phase, concepts that were related to the research questions of each study were identified. Following this step, axial coding took place to group the concepts of interest and create relationships between them. The method of constant comparison was used as new transcripts were added and the more recently emerging concepts and categories were compared to those previously identified (Corbin & Strauss, 2008; Merriam, 2009). The analysis process was carried out in iterations. The codes were finally classified into the identified themes, by which results and discussion were organized.

The resulting themes were then translated into English along with representative quotes for each concept and category. Depending on the individual studies, when there were two participant groups (e.g. Hungarian and Chinese), the themes identified from different groups were compared and the results were discussed as unified themes. Differences were presented and discussed if existed.

Chapter 4: Search Behavior in Non-native Languages

As the first step in stage one of the dissertation research, this exploratory study focuses on and examines behavioral differences in first and second language web searching. Query reformulations from 14 participants who searched in English (first language) and Spanish (second language) and 17 participants who searched in Hungarian (first language) and English (second language) are analyzed and compared. Results indicate that searching in a foreign language requires significantly longer time, more query reformulations, and more websites viewed. User feedback also indicates that people tend to utilize different search strategies when searching in a second language than those in native language search.

4.1 Introduction

Search result accuracy and user experience of foreign language searches suffer from the unfamiliarity with either the language itself or the culture of the country in which the language is spoken (Komlodi, Jozsa, Koles, & Hercegfi, 2011). The language proficiency of the searcher is an important contextual factor that can strongly impact the search experience and outcome. However, previous research (e.g. Kralisch & Mandl, 2006; Marlow et al., 2008; Aula & Kellar, 2009; Berendt & Kralisch, 2009; Clough & Eleta, 2010; Steichen et al., 2014) usually focuses on language selection

and its reasons and is insufficient to address the impact of searching in a foreign language on search process and outcomes. Thus, the goal of this study is to extend the previous work and examine search behaviors and identify potential differences in such context in college students' searching activities as they carried out searches in their native and a foreign language.

In this study, both quantitative and qualitative methods are used to examine 68 queries, 311 reformulations, 508 visited websites, as well as 31 interviews. Results are provided and discussed on both the task performance from the quantitative log data and the user experience from the qualitative interview data.

We focused specifically on the search tactics, efficiency, and user subjective satisfaction in second language searching. The study aims at answering the following research question:

> RQ1.1: What are the behavioral differences when online users search in a non-native language, specifically in terms of query reformulation strategy, task performance, and user experience?

The data collection and analysis followed the general methods introduced in Chapter 3 and details were described in section 3.4.1.

4.2 Results

In this study, male participants accounted for 68 percent, or 21 out of 31, of all the participants. The ages ranged from 19 to 62, with a mean of 23.7 and a standard deviation of 7.4 (see Table 4-1).

Demographics	
Ave. Age	23.7
SD Age	7.4
Male	68%
Female	32%

Table 4-1. Participant demographics

To address the research questions, we compared the following search behavior metrics between native and foreign language searching. We selected these metrics from the logs because we felt they would help us understand the users' search behavior:

- *Searching time (in seconds);*
- *Number of new queries;*
- *Number and types of query reformulations;*
- *Number of search engine results pages (SERP) viewed;*
- *Number of websites opened;*
- *Number of links clicked within a website opened*

Most measures are straightforward, such as search time, counting the number of SERPs and websites viewed. However, new query and Query Reformulation (QR) are essential concepts in this study, the definition and classification of which varied and were inconsistent in previous works. For instance, according to Rieh and Xie (2006), "QR is the product of the interaction between the user and the Information Retrieval system". Jansen, Booth, and Spink (2009) defined QR more specifically as "the process of altering a given query to improve search or retrieval performance".

To group the query reformulations identified in our own data, we combined and modified the classification systems used in the two studies (Rieh & Xie, 2006; Jansen et al., 2009). The definition of "New Query", "Generalization", "Specialization", "Reformulation", and "Content Change" were from Rieh and Xie (2006). "Synonym" and "Spelling Correction" were from Jansen et al. (2009). We then added "Regional English Variance". The resulting classification system is simple and the categories are mutually exclusive (Table 4-2). Queries and reformulations from both the US and the Hungarian data were classified into these categories.

	Specific Classification	Definition	Example	US		Hungary	
				Native L	Foreign L	Native L	Foreign L
New	New Query	A new query has no terms in common with the previous one.	IBM CEO ---> ocean city maryland	14	16	19	19
Reformulation	Generalization	Same query but at least one term less.	ocean city maryland activities for kids ---> ocean city maryland activities	4 (8.3%)	8 (6.1%)	5 (7.9%)	7 (10.1%)
	Specialization	Same query but at least one term more.	ocean city maryland activities ---> ocean city maryland activities for kids	9 (18.8%)	21 (16.0%)	14 (22.2%)	22 (31.9%)
	Reformulation	At least one term in common, at least one term changed. The changed terms are not synonyms.	ocean city activities ---> ocean city kids	33 (68.8%)	94 (71.8%)	35 (55.6%)	34 (49.3%)
	Synonym	At least one term in common, at least one term changed. The changed terms are synonyms.	ocean city activities ---> ocean city things to do	2 (4.1%)	6 (4.6%)	6 (9.5%)	1 (1.45%)
	Content Change	Same query but different content, "Web, Image, Video, Audio, News, and Maps".	Changing from “Web search” to “News”	0	0	2 (3.2%)	1 (1.45%)
	Spelling Correction	Correction of misspelling.	occean city activities ---> ocean city activities	0	2 (1.5%)	1 (1.6%)	2 (2.9%)
	Regional English Variance	Changing from British English to American English	ocean city maryland programmes, for kids ---> ocean city maryland programs for kids	0	0	0	2 (2.9%)

Table 0-1. Query and reformulation definition and data classification

A mixed method analysis was used examining both quantitative and qualitative data. In analyzing the quantitative task performance log data, a single factor ANOVA was calculated to identify statistically significant differences between native and second language searching. Specifically, significant effects were found in search time (in seconds), number of reformulation, and number of websites opened. For the qualitative data, grounded theory was used to explore user experience and search patterns while reviewing the interview transcripts.

4.2.1 Task Performance

We performed two rounds of statistical analysis on the quantitative data we gathered. The first round was based on the aggregated data (including both the US data set and the Hungarian data set) and focused on native vs. foreign language differences. The second round was to analyze the US data and the Hungarian data separately to compare and contrast the two countries.

For the aggregated data, we found significant language effects in the time spent and the number of reformulations (Table 4-3). This showed that searching in a second language required significantly longer time and more reformulations. The number of new queries was the same between the languages, however, for most users there was only one new query for each task and the rest of them were reformulations. We did

not find significant differences between the two languages in the number of SERPs viewed, the number of websites opened, or the number of links clicked within a website.

	Mean		***F***	***P-value***	***F crit***
	Native	**Second**			
Time spent	541.15	840.83	9.91	0.0026	4.00
Number of reformulations	3.58	6.45	6.94	0.0107	4.00

Table 0-2. ANOVA of the aggregated data

These results show that speakers of a second language struggle more with query formulation than with results processing. The large number of query reformulations shows that they had more trouble finding the right keywords and formulating the appropriate query, while the similarities in how they processed the results overall was very similar. However, when we analyzed the two data sets separately, variation in the results processing behavior was identified between the two countries.

For the separated data sets by country, in the US data we found the same significant differences as before and one additional significant factor. The number of websites opened (Table 4-4) was significantly higher in the foreign language. We did not find

significant differences between the two languages in the number of new queries, SERPs viewed, websites opened, or links clicked within a website.

	Mean		***F***	***P-value***	***F crit***
	English	**Spanish**			
Time spent	590.86	1048.43	9.15	0.0055*	4.23
Number of reformulations	3.43	9.36	12.10	0.0018*	4.23
Number of websites opened	7.14	12.64	9.51	0.0048*	4.23

Table 0-3. ANOVA of the US data

On the other hand, though the mean value of task time and number of reformulations in the Hungarian data followed the same trend, none of the variables were significantly different in the two languages. We saw three potential explanations for this which should be further studied.

- *Proficiency of the second language:* the Spanish proficiency of the US participants was slightly lower than the English proficiency of the Hungarian participants. We based our measurement of language proficiency on self report. In future research, we will measure this variable more precisely.

- *Previous experience:* Hungarian users reported having much previous experience in searching in English as a second language, while US users rarely searched in Spanish.

- *Quality, organization, design, and architecture of websites:* the quality, as well as the organization and architecture of Spanish websites were not as good as those of the English-language websites. There was more information in English websites and the websites were more usable. The Spanish-language sites occasionally introduced viruses or contained spam content.

Finally, the distribution of classifications of QRs was studied. Similar patterns were found in the US and Hungarian data. The subcategories of "reformulation" and "specialization" accounted for the majority of occurrence, approximately 80%, of the general QRs (Table 4-2). This shows that users are inclined to add or change terms in the process of altering a given query in both languages, but tend to do so more frequently in the foreign language.

4.2.2 User Experience

The semi-structured interviews about the searching experience in different languages were transcribed, translated (from Hungarian to English, for the Hungarian group), and coded. Findings from both the US group and the Hungarian group on searching

behaviors/strategies in a foreign language showed that while searching in a foreign language, users tended to report the following patterns:

- Language difficulties

Language barriers were seen in selecting search terms and understanding returned content in term of foreign language searching. This is the most obvious challenge, but there are several others.

- Searchers think more and initiate more specific queries in a foreign language

When searching in a native language, users usually selected general terms promptly and started to search and skim. For example, when searching in their native languages, they just initiated a query of "Ocean City Maryland activities" and started to skim. While during searching in a foreign language, they tended to think more and initiate more specific queries, like a museum or a cinema. They also refined their queries more frequently. A participant offered a typical situation: *"I had to think a little more about exactly what I was looking for. So instead of just going to a webpage and seeing what was there, I had to think of what kind of activities would work well for children on a rainy day. So I thought, ok maybe a museum and I typed in museums and so forth, where as for Ocean City (which was the search in my native language),*

I just typed in activities, which is a broader category... So I had to be more refined in the search for Bolivia (which was the search in a foreign language)" (Interview P7).

- Accept query recommendation of search engine more frequently

Based on researcher observation and participant self-reported interview data, participants used the recommendation features of the search engine more frequently when searching in a foreign language.

- Reformulate more often but click less

While searching in a foreign language, users were less confident in the terms to be used. Thus, they usually chose to reformulate the initial query and click less on the links of SERP. For instance, "*I just go back and try new keywords more often, instead of clicking on the results (search result links on SERPs)*" (Interview P4).

- Lack of familiarity with the domain knowledge of the search and what to expect as the search results

It would be beneficial to the search if users were familiar with the domain knowledge of the search, the expected categories of result sets which were related to the domain, and the localized information literacy. However, the participants indicated they usually did not possess such familiarity during the foreign language search. For

example, a Hungarian participant mentioned the unfamiliarity. *"mostly when I search in Hungarian, I know the sites of the topic and the part of the sites where I can find it. Also, if a site has a search field, then I use it. I don't do it on English sites, because I don't know what sets (categories of result sets) I can choose from (e.g. all events, arts and museums, concerts, live music, festivals and fairs), therefore the first set (all events) is the one I always choose. I don't have the knowledge of the element of the sets (the participant meant the sub-categories of event sets) that I can search within* (Interview P3).

- Lack of familiarity with trustworthy sources and the genres of websites available in a given culture

The difficulty of making credibility judgments was amplified while searching in a foreign language/culture. Additional efforts were required to judge the credibility of the opened websites.

- Lack of familiarity with design conventions other than some basic rules

Differences in the design conventions between the websites in the users' native language and those in a foreign language were also related to credibility judgments and an important difficulty (Ahmad, Komlodi, Wang, & Hercegfi, 2010). It made the

foreign language search even more taxing to estimate the quality of the sites in the foreign language.

- Look for content with less writing and resort to other sources and websites

This was partially caused by the language barriers. Users looked for websites where information was easier for them to understand, such as images, graphics, or videos. Occasionally, participants resorted to more than one source or website to look for information about the same thing in order to create a complete picture for it, since they were less confident while searching in a foreign language.

4.3 Discussion

Results of this study show that searching in a foreign language requires significantly longer time, more query reformulations, and more websites viewed. User experience also indicates that people tend to utilize different strategies when searching in a second language than those in native language search. In this section, we explore a few implications for future search engines and websites to better serve foreign language searchers.

Our results indicated that participants used the recommendation features of the search engine more frequently when searching in a foreign language. Thus, improving the

quality of the recommendation system of search engines would be useful. What is more, rather than showing recommendations only when initiating a query or entering a reformulation, utilizing a separate bar or tag to show relevant search queries and possible reformulations in a highly visible area of the search results page might be helpful. Additional information about the recommendations could be effective. For instance, other users' rating and ranking of the recommended queries, reformulations, or results could be added. The importance of customizable search user interfaces is also indicated. We believe customizability is essential to users' preference and experience.

In regards to websites, following guidelines for the international or cross-cultural design is crucial in making them easily accessible for foreign language users. Guidelines, such as using simple vocabulary and sentence structure at a lower reading level, applying intuitive designs and navigation structures, and avoiding culturally loaded imagery, can make important adjustments to websites. These adjustments make it easier to traverse and interact with foreign language searchers. One mostly mentioned feature is to encourage users to search within web pages and making it visible on web pages can be particularly helpful.

The findings from this study highlighted important differences between native and foreign language searching. The findings were derived from prescribed search tasks in a lab setting. In the next study we set out to gain a deeper understanding of the patterns and motivations of natural foreign-language searchers behavior. We plan to extend this study by examining the types of problems and challenges second language searchers face. Search tasks in the next study will be selected by participants and natural behavior data will be collected through real life search diaries.

Chapter 5: Challenges and Search Process: Online Search in English as a Non-native Language

The previous study focused on behavioral differences when online users search in a non-native language. This chapter takes a step further and addresses what causes the behavioral differences, by studying NNES searchers' challenges and their root causes. Particular attention is paid to search challenges, query formulation and reformulation strategies, as well as user interaction with online systems and tools. This study takes a holistic view on the entire search process. Search tasks in this study are selected by participants and natural behavior data are collected through real life search diaries.

Data are collected through a focus group study (8 groups of 36 participants) and a diary study (30 participants). Results indicate NNESs face a unique set of challenges that may not be present for native speakers when searching in English. Two typical query formulation and reformulation strategies are identified and described. A user interaction model is abstracted to address the iterative learning and reformulation search process of NNESs. Implications for design of systems and tools to assist this particular user group are discussed.

5.1 Introduction

The purpose of this research is to further understand the user behavior of NNESs when they conduct online searching in English and, subsequently, to provide implications for the design of systems and tools to support them. Specifically, we aim to answer the following research questions:

RQ2.1: How do NNESs search for information in English (search behavioral pattern and user experience)?

RQ2.2: What are the typical challenges of NNESs? What are the causes and roots of the identified challenges?

RQ2.3: What user interaction model can be abstracted to describe the typical search processes and characteristics of NNESs?

Based on the nature of the research questions we aim to address, a qualitative study was conducted. Specifically, data were collected through focus group discussion (8 groups of 36 participants) and search diaries and follow-up interviews (30 participants) in two countries, Hungary and the US. Two groups of participants were selected, native Hungarian participants (in Hungary) and native Chinese participants (in the US). They were all typical NNESs with a tendency to search in English. However,

they differed significantly in their motivation to search in English, cultural backgrounds, and exposure to and relationship with an English-speaking culture.

The data collection and analysis followed the general methods introduced in Chapter 3 and details were described in section 3.4.2. This chapter presents the results of the study and provides a discussion to better understand NNESs search challenges, processes, and strategies. User interaction models are abstracted and insights for design systems and tools to support NNESs are discussed.

5.2 Results

In this section, I describe participant demographics, as well as the resulting themes identified from the analysis. Results from the two sets of data (Hungarian and Chinese) are compared and discussed.

The focus group data collection and analysis followed the general methods introduced in Chapter 3. In this study, the focus group data were firstly analyzed. Then, the diary study data were utilized to triangulate the results and findings. The analysis process was carried out in iterations with separate languages (either Hungarian or Chinese). The resulting themes were translated into English along with representative quotes for each concept and category. The themes were finally compared and triangulated with diary entries and interviews. The themes identified from both the Hungarian data set

and the Chinese data set were compared and discussed. The results were presented as unified themes, although differences were noted.

5.2.1 Demographics

	Focus group study		Diary study	
Demographics	**Chinese**	**Hungarian**	**Chinese**	**Hungarian**
Ave. Age	29.6	22.3	26.3	24.1
SD Age	3.6	1.3	4.0	1.5
Male	9	7	9	9
Female	7	13	6	6

Table 0-1. Participant demographics

- Chinese Data Set

Four focus groups were conducted with four participants in each in the US. Fifteen participants took part in the diary study. All participants in this pool were born in China and were native speakers of Chinese. In self-report questionnaires, they all indicated they spent most of their lives in China and came to the US to pursue higher education, mostly graduate studies. Most had not lived in countries other than China and the US. The average age of focus group participants was 29.6 years and 26.3 for diary study (see Table 5-1).

All participants reported that they spoke English well enough to have a reasonable conversation with a local person in the US (since this was one of the study

requirements). About one third of participants stated that they learnt English as a foreign language during childhood and spoke it almost at native level. Most indicated they often used English in information search on the Web. Other than English, they did not speak or use another non-native language.

- Hungarian Data Set

For this participant pool, four focus groups were conducted with five participants each. Fifteen participants took part in the diary study. All participants were born in Hungary and were native speakers of Hungarian. Similar to the Chinese participants, they spent most of their lives in their native country. However, they had a lower average age (22.3 years for focus group and 24.1 for diary study) and age variation (SD = 1.3 for focus group and 1.5 for diary study; see Table 5-1) as they were mostly undergraduate students or recent graduates. They reported lower average level of education completed (BA/BS and high schools).

Comparing to Chinese groups, the Hungarian participants reported more diversity in their languages and countries of stay. All participants reported that they spoke English at a conversational level. Approximately ten percent of them stated they spoke it almost at native level. Some indicated they spoke other languages at a conversational level (about one third participants indicated they spoke German at a conversational

level; others indicated French, Spanish, and Russian). Forty percent of the participants stayed in countries other than their native country (from one month to fourteen months), which included the US, Austria, Switzerland, Mexico, Spain, Canada, Germany, and Bulgaria.

5.2.2 Search Behavioral Pattern and UX (RQ2.1)

The first research question pays particular attention to how NNESs search. Thus, this subsection presents results about the search behavior and process.

- Language Selection: Patterns of Mixed Language Use

Though searching solely in one language was not rare (such as English for music or movies, international news, and information for travel in an English-speaking country), alternating languages in the same search was mentioned by most participants, in both the Hungarian and the Chinese data sets. Participants in our research combined searching in their native language and English as it suited their needs. Searches in the two languages were complementary and often supported the searchers' interpretation of the information found in the other language. The patterns of language use fell into three categories:

1. Switching from the native language to English

Participants often mentioned two reasons for such a switch: They did not plan to search in English, but failed to find enough content in their native language; or they intended to search in English from the beginning, but they started to search in their native language to get an overview of the topic. Similar quotes could be found in both data sets.

2. Switching from English to the native language

After searching in English, participants sometimes switched back to their native language to verify the English results or further understand them. *"If we've found something in English, maybe afterwards we search for it in Hungarian as well to double-check or get a better understanding (on the English results)"*.

3. Intertwined language usage

Recall that the results in the previous sub-section showed some factors together made a difference in language selection. For example, topic and type of information went hand in hand and impacted language selection together, especially in the case of academic searches (searching in native language to get an overview and general information and search for a deeper understanding in English). *"I found information in English on Wikipedia, but the content was too much ... I sometimes looked back*

and forth at Chinese articles in order to have an overview". Another example was when they searched for online shopping products, such as computer parts, they searched in their native language for product description, specification, and reviews, whereas in English for prices and buying options. As such, queries in the native language and English often followed and complemented each another.

- Language Selection: Cultural Interpretation of Topics

In previous research, a series of factors were identified impacting multilingual searchers' language selection and usage: content availability in a specific language (Kralisch & Mandl, 2006; Aula & Kellar, 2009), language proficiency (Berendt & Kralisch, 2009; Marlow et al., 2008), domain of knowledge (Berendt & Kralisch, 2009; Clough & Eleta, 2010), domain of search topics (Steichen et al., 2014), and usage purpose (Steichen et al., 2014).

In addition to the simple impact of topic on language selection, we also found a culturally-mediated effect of the search topic. Our participants would choose the language of the search not just depending on the topic and the availability of the information, but also influenced by how they expected the topic to be represented in different cultures.

Chinese participants indicated a preference for searching in Chinese when they sought emotional support from the groups in which they were members, whereas they chose English to search for scientifically-based facts. When participants were selecting a language to search in, cultural affinity, a natural tendency to trust people with the same cultural background for personal feelings, made a difference. For instance, when a participant was pregnant and looked for discussions on symptoms of pregnancy and feelings, baby care topics, and prenatal symptoms, she indicated a preference for using Chinese in these searches because she believed the feelings of expectant mothers with the same cultural background were more relevant and easier to accept. On the other hand, she searched in English for treatment options, medication and side effects, as she indicated *"these topics in English seem more scientific"*.

Another example where the searchers' cultural ties influenced language selection was selecting language for cultural point of view towards political events. Chinese participants usually searched in Chinese for news about local political events that occurred in China. However, they also searched in English for the news of the same event to gain access to western voices and observe a different point of view on the events.

- Language Selection: Same Topic – Different Approach

We found that multiple factors interacted to influence language selection in other scenarios as well. For example, two major factors, topic and type of information (general introduction vs. specific questions), usually worked in concert. English was mentioned by most participants as the language used to search for academic materials. However, both Chinese and Hungarian were also used by the participants to search for academic topics. Why did they use different languages to search for the same topic? Because they searched for different types of content in the two languages. When searching in their native language, participants noted that they wanted an overview and general understanding. *"I just wanted to take a look at the academic topic, general discussion, issues, or research questions to gain some general understanding, because I read Chinese a lot faster"*. On the other hand, they looked for more specific information in English or aimed to get a deeper understanding of the same topic. *"I was trying to gain a deeper understanding and looking for references and papers for citation ... these papers and materials were more directly related to my own research"*. Recall that, in the example in the previous sub-section about the cultural factor, the participant was also searching in both languages for the same topic but for different content (emotional feelings in the native language and scientific

description in English). Thus, usage purposes (overview or detail) and content types or aspects of the topic collectively influenced searchers' language selection.

- UX: Lack of Confidence, Stress, and Attention

The vast majority of both Chinese and Hungarian participants indicated that searching in English as a non-native language was not as convenient as it was in their native language. This was consistent with the findings of previous research that indicated challenges existed in non-native language searching (Chu, Jozsa, Komlodi, & Hercegfi, 2012; Jozsa, Koles, Komlodi, Hercegfi, & Chu, 2012). Participants often described strain on their attention and an increased cognitive load. They also associated feelings of uncertainty and stress with English language searching: *"I didn't like searching in English for a long time. I know that I understand the content in English, but it bothered me. Especially when I search for an unknown topic… it makes me unsure…"*.

5.2.3 Challenges (RQ2.2)

Findings related to challenges are described and organized around various steps of the search process. We present searchers' challenges and typical behavior intermixed as they emerge around various search process steps.

- Challenge 1: Query Formulation

Query formulation was one of the first steps in the search process. It was indicated by most participants to be the most challenging job they faced while searching in English. We believed this was a common challenge for all NNESs. They faced this challenge because finding the appropriate keywords to search on was very difficult in a non-native language.

Some Chinese participants stated *"sometimes I had trouble expressing myself accurately in English"*. Others confirmed this and indicated it was especially difficult when it came to *"certain specialized or academic terms"* and *"common established expressions"* in English. A strategy was described to deal with this issue: *"the result might not be accurate if I used sentences instead of keywords when searching for information (in English) and I usually started with simple keywords"*.

Similar concerns were found in the Hungarian data. They tended to use only a few words in English as query terms, whereas in Hungarian they were more likely to use longer phrases and even sentences. They reported using fewer words and simpler expressions for searching in English, as these were easier to create than longer and more complex queries. *"In English, we'd rather use keywords. If I enter a whole sentence I may make grammatical or other mistakes, so I prefer entering an*

important keyword and I hope I will find what I seek. In Hungarian, I have a greater chance to find specifically what I am thinking of".

- Challenge 2: Query Reformulation

Similar to query formulation, our participants struggled with reformulating queries. It was especially difficult for them to pick an appropriate synonym of the initial term when reformulation was needed. As one of our Hungarian participants mentioned, *"there are a wide variety of Hungarian expressions for different meanings which are supposed to exist in English as well but we don't know them".*

NNESs faced this challenge due to the lack of specific vocabulary and an unawareness or insecurity about slight variations in meaning. Under such circumstances, translation tools, such as dictionaries or online translation, did not help much. For example, a Hungarian participant stated that *"Google Translate shows 2-3 expressions but I can't decide which one is the best or most commonly used".* To address this problem, Hungarian participants reported testing synonyms, querying similar expressions one after another to better understand differences in their meanings.

- Challenge 3: Viewing and Skimming Search Engine Result Pages (SERPs) and Websites

Our study confirmed the results from previous studies (e.g. Chu et al., 2012) that it took longer time to find results in a non-native language. Query formulation and reformulation took longer and so did viewing results. When it came to viewing and skimming SERPs and websites, our participants indicated that they read content in their native language faster than they do in English. For example, a Chinese participant stated, *"I read English content word by word, but I read Chinese content one paragraph at a time"*. The same situation was described by the Hungarian participants, *"It takes so much time… I have to open and look at several pages (on a SERP) to find what I really need"*.

Skimming a resulting website was even harder since it usually contained more textual content than a SERP did. As discussed above, our participants often described facing overwhelmingly large result sets of varying levels of relevance. Not surprisingly, it usually took them longer and they found it more difficult to process text in English, a non-native language. Three typical types of behavior were identified to deal with this issue in both datasets:

- Using the vertical search tools of search engines (e.g. using "Images" to get visual results which indirectly led them to textual information)

- Skimming text rather than reading the content of websites line by line (using "Ctrl+F" keyboard shortcut to easily locate keywords so that it was not necessary to read the full text of a webpage)

- Choosing one specific website and continuing to search inside to avoid facing the information overload of the result sets

NNESs faced these three types of typical challenges when conducting online search in English. The challenges intertwined during each step of the search process and made the entire process further complicated.

5.3 Discussion

The study findings sheds light on user challenges, experiences, and behaviors. In this section we summarize our key findings, provide further insights, and discuss implications.

5.3.1 Dilemma: Short vs. Long Query (RQ2.3)

The analysis of user challenge and experience identified a dilemma that existed during non-native English search. The dilemma involved the following two query length strategies:

- Formulating a short and general query is simpler, but it generates a large amount of results with lower overall relevance, and hence, requires the user to process more information in a non-native language, thus creating a more difficult challenge in the second half of the search process, browsing SERPs and websites.

- Formulating a long and specific query generates a smaller, but more relevant result set, and requires searchers to process less information in a non-native language. This requires users to create more complex and appropriate queries.

This issue is faced by native English speakers as well. However, NNESs' English language proficiency adds another dimension of complexity to the problem. Our focus in this paper is on the first scenario, formulating a short and general query, because this behavior was usually identified and described by our participants when determining the length of queries.

5.3.2 Query Formulation and Reformulation: Two Strategies (RQ2.3)

To tackle the challenges associated with query formulation and reformulation, participants used two different strategies: translation strategy and discovery strategy. The two strategies are presented and described in Figures 5-1 and 5-2, by adapting and modifying Broder's classic model for IR (Broder, 2002).

- Translation Strategy

Due to their relatively low English language proficiency level, English learners usually have difficulties finding keywords and forming proper and idiomatic English queries or expressions. The translation strategy is thus especially helpful to NNESs whose English language proficiency is less fluent or who are new to English searching.

Digital dictionaries and online translation tools or websites are usually resorted to when users are facing this issue. While most CLIR systems provide automatic translation, participants in our study usually translated queries by themselves. Some participants stated they were not aware of such CLIR systems or did not trust the automatic translation. Figure 5-1 presents and describes this strategy (the modification to the original Broder's classic model for IR, 2002, is highlighted).

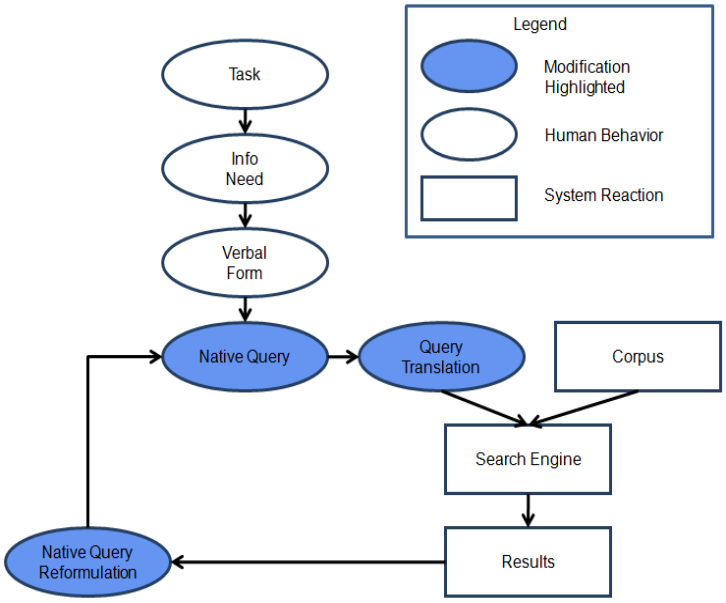


Figure 0-1. Query formulation scenario 1: Translation strategy

- Discovery Strategy

The second strategy is carried out solely in English. When searchers are not familiar with proper or idiomatic English expressions, they usually start with simple keywords and search. Subsequently, while browsing the SERP, they identify better English expressions for the query. Thus, they reformulate the original query by using the search expression identified through browsing the results and searching iteratively. In other words, a gradual, progressive, and spiral query formulation strategy is utilized (see Figure 5-2), during which searchers learn about the English vocabulary used in the search domain. Participants described and suggested the discovery strategy as, *"start your search with fewer keywords, then during the search, you can view and*

browse the results to modify and refine your original term. This will get more accurate results".

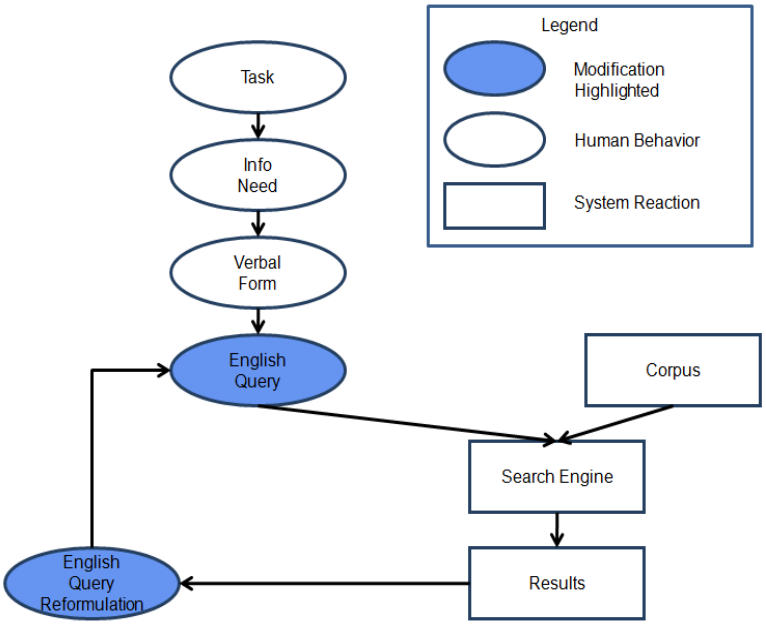


Figure 0-2. Query formulation scenario 2: Discovery strategy

5.3.3 User Interaction Model: Iterative Learning and Reformulation (RQ2.3)

The behaviors described in Figures 5-1 and 5-2 do not appear in isolation and are usually used in combination and iteratively within one search session. Figure 5-3 provides a user interaction model for the NNESs search process, which is an iterative and spiral process. In each iteration, NNESs utilize either the translation strategy (see left panel in Figure 5-3) or the discovery strategy (see right panel in Figure 5-3) to

formulate and reformulate their queries. By reading and learning from the results encountered, they gain a better understanding of the information they are searching for. Each iteration builds on the previous results and provides further understanding. While this is typical in information seeking in general, learning the vocabulary of the domain in a non-native language is emphasized and unique to NNESs in the process. Even if the focus of the search task remains the same, NNES searchers have to use the process to learn the language of the domain, as well as enhance their queries through the new knowledge. The difference between native and non-native search process highlights the fact that NNESs not just learn about the topic of their information need, but also the way English is used in that topic domain. At the end of the search process, NNESs either satisfy their information need or quit the search task.

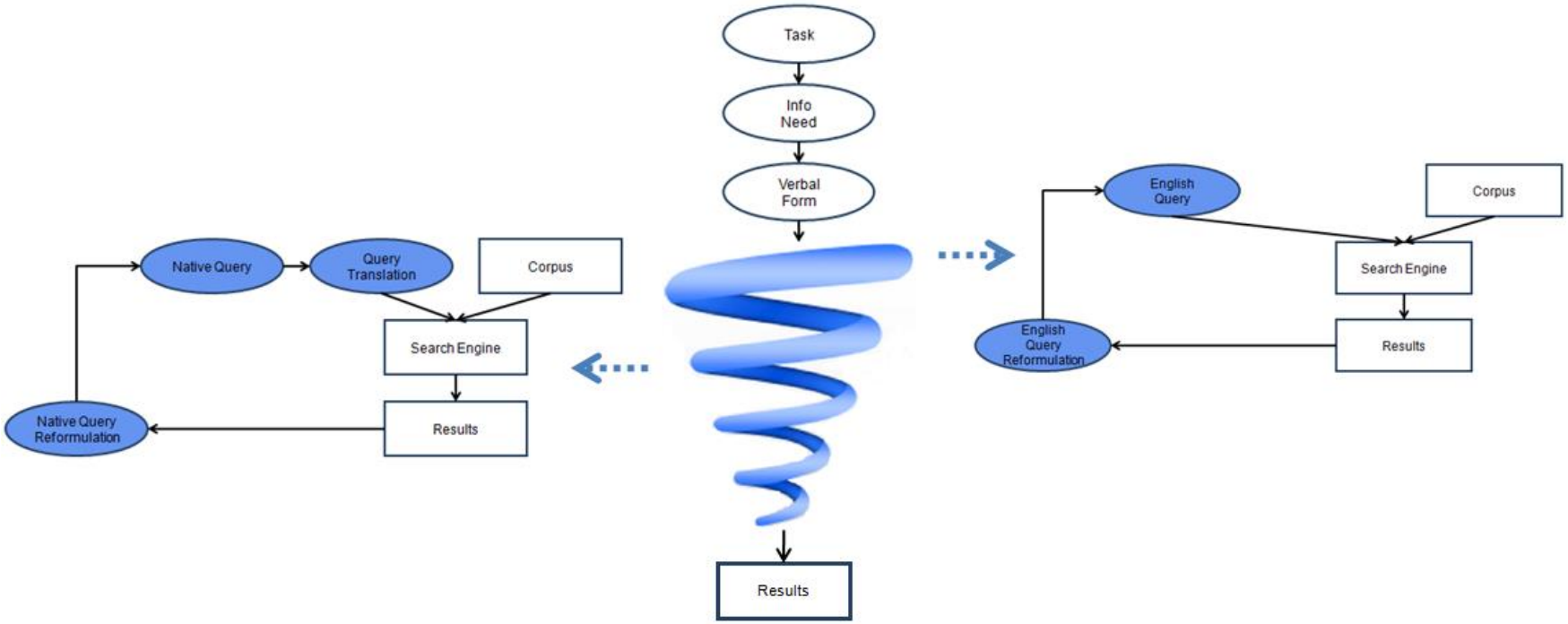


Figure 0-3. User interaction model: NNESs search process - iterative learning and reformulation

5.3.4 Implications

- Query Formulation and Reformulation

The challenges identified in query formulation and reformulation called for solutions to assist NNESs in shaping reliable and effective queries. Automatic query translation might fail when multiple synonyms with slight variations of meaning are applied. New tools for helping NNESs formulate queries are necessary.

- Result Presentation

Our participants described using alternate languages when viewing results. The complementary interaction of the two languages showed that, though not playing the same role, each language contributed something unique to the users' understanding of an information problem. Search engines allow for this integrated interaction to a certain extent, but better integration of results in multiple languages can be helpful.

- Interface Personalization

The results of this study indicated various user preferences of NNESs. The user preferences, in turn, urge variant personalization and customization. They may be provided on query formulation and reformulation (e.g. query translation and recommendation system), result presentation across languages (e.g. a personalized interface showing and linking results in different languages), and tools and features on SERPs (e.g. a customizable vertical search tool bar).

In summary, NNES searchers face language-based challenges both in the query formulation and reformulation phase and result viewing phase. They combine their native language and English while searching and use both translation and discovery/learning strategies to construct and reformulate queries. They also combine information sources in both languages when reading results and learning about the topic of the search. User experience, in general, indicates lack of confidence, stress, and extra attention during their searches. New system and interface solutions are needed to support these behaviors.

Chapter 6: TranSearch 1.0: Prototyping and User Study

Previous studies extended the understanding of NNES' information needs and search behavior, including challenges. Based on these findings, this chapter presents a prototype system and a user study to explore and extend current approaches of supporting non-native speakers' searching.

In this chapter, we review and summarize previous work related to CLIR and MLIR research and system development practices. Subsequently, we extend the current approaches by designing and evaluating a prototype, TranSearch, which is developed on the basis of our previous behavioral findings and the previous research on CLIR and MLIR systems.

TranSearch is a MLIR system and UI that allows the user to search in two languages at the same time, English and Chinese for the current prototype version. The system is built on previous CLIR and MLIR system research and integrates intuitive user interaction during the query construction phase and accommodates user preferences of search result displays. The design process of TranSearch and how the UI prototype accommodates the identified search challenges, patterns, strategies and processes are explained. Subsequently, the result of a follow-up user study are presented. Finally, the testing results are presented on prototype strengths and weaknesses and design recommendations are proposed and discussed.

6.1 Introduction

IR research responding to the need to search for and retrieve information in multiple

languages is often referred to as “multilingual information access and retrieval” (e.g. Peters et al., 2005). In chapter 2 section 2.2.2, we introduced and described three overlapping terms commonly used in this research area: CLIR, MLIR, and MLIA. During the interaction with search engines, CLIR consists of entering queries in one language and retrieving relevant information in another (e.g. Grefenstette, 1998; Peters, 2001; Nie, 2010). MLIR is a broader term and includes the concept of CLIR, because it deals with "managing information access and discovery in multiple languages both monolingually and across languages"(Peters et al., 2012). MLIA is usually used in the broadest sense and addresses the problem of "accessing, querying, and retrieving information from collections in any language at any level of specificity" (Peters, Braschler, & Clough, 2012).

These systems and interfaces tackle the challenges multilingual searchers face when searching for information in multiple languages. They usually address these problems by providing machine translation in the query construction phase and providing multilingual search result displays. Due to the complexity of the two phases, both depend heavily on user interaction and input. Current system and UI design approaches usually separate the two phases and focus on either query translation (e.g. Petrelli et al., 2006; Airio, 2008) or user preferences in result layout (e.g. Steichen & Freund, 2015).

In this study, a prototype multilingual search user interface, TranSearch, is designed and developed. Subsequently, a user study was carried out to test TranSearch prototype and provide insights of its strengths and weaknesses and possible

improvements, in hope of a future redesign. In addition, our study extends current research by explaining some of the reasons behind user preferences left unanswered by Steichen and Freund (2015). Specifically, our study aims at answering the following research questions:

RQ3.1: What UI and tools can be created to support NNESs based on their challenges and search process (identified from our previous studies)?

RQ3.2: What are the strength and weakness of the UI and tools designed? Any improvement to the prototype?

The data collection and analysis followed the general methods introduced in Chapter 3 and details were described in section 3.4.3.

6.2 Background Work

Traditional CLIR and MLIR research studies divide the entire search process into two phases, query translation and result presentation, and usually focuses on either improving query machine translation accuracy and search result quality (e.g. Gao et al., 2007; Amato et al., 2007), or the display of results (e.g. Steichen & Freund, 2015). Thus, we will review previous research on these two phases separately.

6.2.1 Query Translation

In CLIR and MLIR, query translation is conducted by either 1) user translation and thus "direct querying" by the user, or 2) fully machine automated query translation without any user intervention, "a delegate mode" (Petrelli et al., 2006; Airio, 2008).

Due to word-sense ambiguity, machine translation suffers from accuracy issues. The results of automated query translation often do not outperform the users' direct querying through human translation in the target language (Airio, 2008; Oard, He, & Wang, 2008). When constructing queries, users make substantial use of automated translations for unfamiliar languages, whereas they tend to write their own translations for familiar languages (Marlow et al., 2008). Petrelli et al. (2006) described and proposed a third approach, “a supervised mode”, machine translation with user interaction ability. In our opinion, offering user interaction during query construction is an improvement, so we select this approach, but the TranSearch user interaction type is different.

In terms of query construction methods, all CLIR and MLIR systems utilized one of the above three approaches. For example, MTIR (Bian & Chen, 2000), Steichen and Freund (2015), and 2lingual.com (2008) used the delegate mode , where as MULINEX (Capstick et al., 2000) and Clarity (Petrelli et al., 2004) fell into the category of the supervised mode. Our prototype selected the supervised mode because user interaction is important and can supplement machine translation. However, the supervised mode has a big challenge. What type of and to what extent should the user interaction be provided? User interaction adds system complexity and requires more steps and user effort. Current systems (e.g. MULINEX and Clarity) usually provide a query assistance tool having users select from translation synonyms. When users possess inadequate knowledge of the nuance in meaning of translated queries in the target language, adding improper user interaction, such as having users

choose from translation synonyms, may be useless, or even worse by confusing users, Petrelli et al. (2006) argued in their research that users might not be interested in or were not able to control the query translation steps. Our prototype, TranSearch, took a simplified approach by offering visibility of machine translated queries and user-editing option.

6.2.2 Search Result Presentation

Search result presentation can also be challenging for MLIR systems, since UIs display results in two or more languages. I surveyed the literature and created the following figure to show the classification. Figure 6-1 illustrates various options of UI types when differentiating search result presentation methods by languages:

- By pagination: Single page view (e.g. Braschler, 2004) vs. multi-page view, such as a tabbed view where each tab shows results from a different language (e.g. Clough, Al-Maskari, & Darwish, 2007)

- By collection merger: “Blended” vs. “non-blended” where the results are presented in separate lists by language or mixed together (e.g. Bron, van Gorp, Nack, Baltussen, & de Rijke, 2013)

- By layout: Panels vs. side-bar, where each panel or side-bar area shows a list from a different language (e.g. 2lingual.com, 2008)

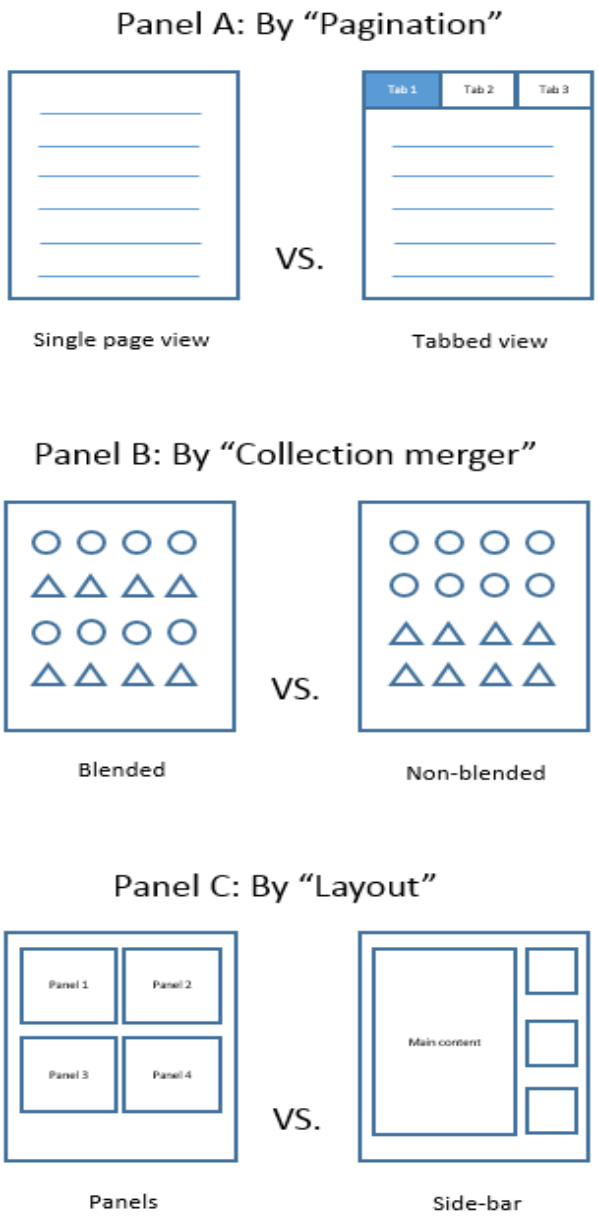


Figure 6-1. Search result user interface criteria and types

These UI types can be used in combination. Steichen and Freund (2015) presented a comparison of five different search result interface designs for CLIR or MLIR. Results indicated that single-page displays with clear language separation were most

preferred (similar to Panel C of Figure 6-1, a single page with panels or side-bars to display results by language). Thus, in our study we selected the single-page design with panels clearly separating the two languages. The panels of our prototype can be adjusted to accommodate user preferences. In the follow-up user study, we also examined the user preferences for panel placement in the single-page design and provided explanation for such preferences.

6.3 Design and Implementation of TranSearch 1.0 (RQ3.1)

TranSearch 1.0 is a MLIR interface prototype that allows the user to search in two languages at the same time, in this case English and Chinese (Chu & Komlodi, 2017). The prototype, built on results of the phase one of the research, highlights the importance of user interaction in the first phase of search process, query translation, and user-adjustable results displays in the second phase. Specifically, during query translation, TranSearch takes Petrelli et al. (2006) approach, "a supervised mode", machine translation with user-editing ability to facilitate user interaction. Its result display, the single-page design with non-blended panels clearly separating the two languages, accommodates user preferences of panel adjustment.

6.3.1 System Requirement Analysis

Our previous studies identified the following three most import challenges for NNES

searching in English (Chu et al., 2012). These challenges are taken into consideration for the TranSearch design:

Challenge 1: Query formulation. NNES struggled with constructing initial queries. Due to the language barrier, sometimes they had to think of a native language query and translate it into English with no confidence of the accuracy of the translation. In other times, they had an English query in their mind, but they were not certain that the query was an established English usage. Thus, the TranSearch design offers an initial text box, uni-box, that accepts queries in either language and the query is instantly translated to the other language by a widely used MT application (the MicroSoft Bing translator API). On the one hand, when a native language query is entered, MT assists and provides relatively high quality translation. On the other hand, when an English query is initiated, MT translates it back to the native language query, which can be used by the users to judge the accuracy of the English query they entered.

Challenge 2: Query reformulation. NNES had to create many reformulations before arriving at acceptable results. Thus, the TranSearch design includes flexible support for query reformulation. Users may reformulate their initial query in the uni-box, or in the text boxes below (sub-boxes) that hold both the initial and the translated queries. Reformulating the initial query will refresh those in the sub-boxes, whereas

modifying that in one of the sub-boxes does not change the other or the initial query. Searchers may find it easier to manage one query in one language and a different query in the other language. This also suits different topics. For example, a research topic related to professional or academic activities may be more precise in English, but a topic related to a personal interest can be more easily phrased in the searcher's native language. Automatic translation is provided but under user control (the supervised mode, Petrelli et al., 2006). Both queries can be edited by the user which supports the frequent reformulations observed when NNES search in English.

Challenge 3: Viewing and skimming SERPs: NNES also had difficulties viewing and skimming the result pages. The major reason was language barriers and they were not able to read as fast as they did in native languages. The TranSearch prototype provides an additional panel displaying search results in the native language. So users can quickly browse and skim SERPs in their native language and get an overview about the search task. They can also compare results in both languages and gain a better understanding. This functionality is especially beneficial for search topics that users are not familiar with, such as academic topics or medical terms. It further strengthens SERP viewing by incorporating a user-adjustable results display.

Our previous studies identified two query construction strategies, the translation

strategy (Figure 5-1 or left part of Figure 6-2) and the discovery strategy (Figure 5-2 or right part of Figure 6-2) and abstracted a user interaction model that incorporated both: iterative learning and reformulation (Figure 6-2) depicting the entire search process. Our prototype integrates these two strategies and support the iterative search process.

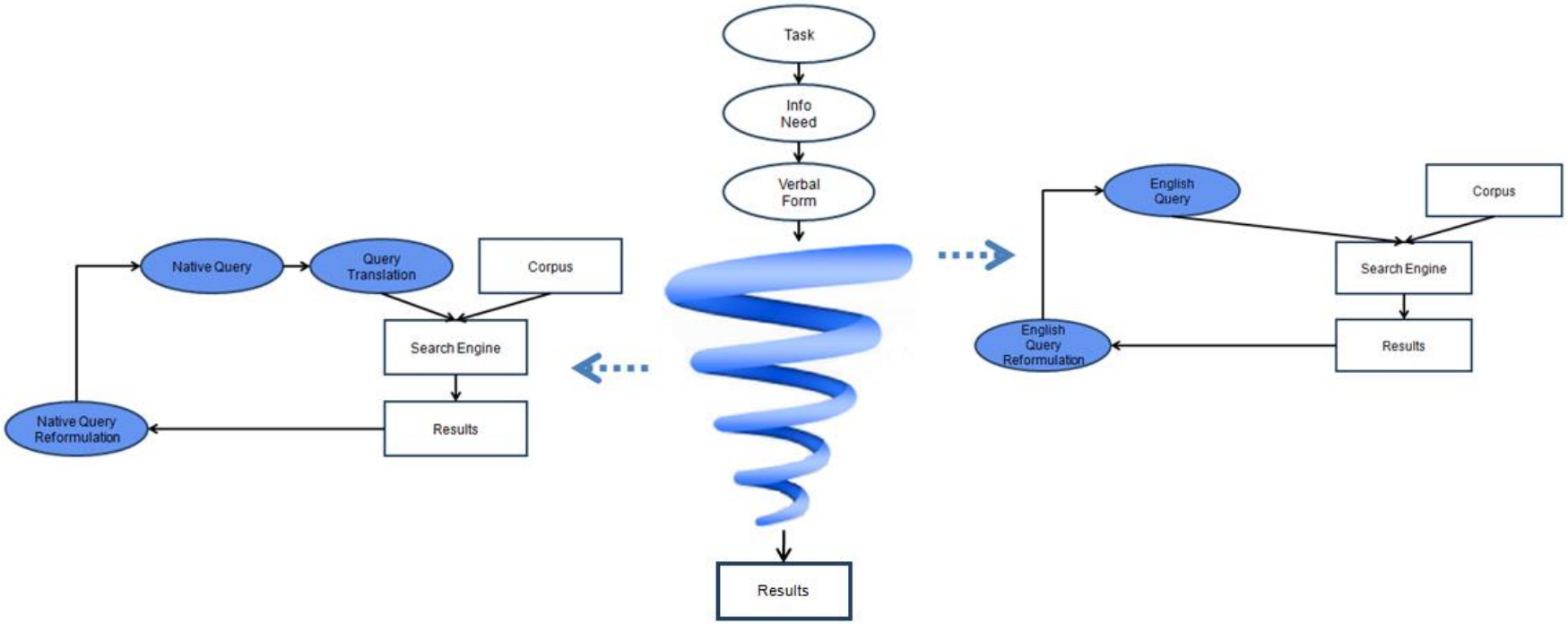


Figure 6-2. User interaction model: NNESs search process - iterative learning and reformulation (repeated Figure)

During query construction phase, TranSearch utilizes machine translation to accommodate users' translation strategy, which is enhanced by allowing user-editing to facilitate user interaction if machine translated queries are not accurate or do not suite the search context. For query discovery strategy, TranSearch's dual result display panels help users discover new query expression in results of both languages. Users may alternate modifying queries in the two panels. Separate but related iterations of search are carried out in the two panels and results of the reformulated queries are returned. By doing so, iterative search is well supported.

For result display, TranSearch applies single-page design with two panels clearly separating the two languages to simplify organizing result groups and facilitate result comparison and browsing. The two panels are independent. So when users change queries and get returned results in one panel, the other remains. This offers users the ability to control and keep track of the iterative progress of either panel. Depending on search task topics or contexts, users may focus on only one language and thus have a preference of panel size. Users can drag the border of either panel to expand it when they focus on results in that language.

The design of TranSearch takes all the requirements into consideration. The features and functionalities as a whole support the three user challenges, accommodate the two search strategies used in query construction, and address the iterative nature of NNES search process.

6.3.2 UI Design

After the system requirements are analyzed based on the previous study findings, the design of TranSearch 1.0 UI is finalized (see Figure 6-3) and implemented as a Web application written in Ruby language (with Microsoft Bing Translator API, Bing and Baidu search API).

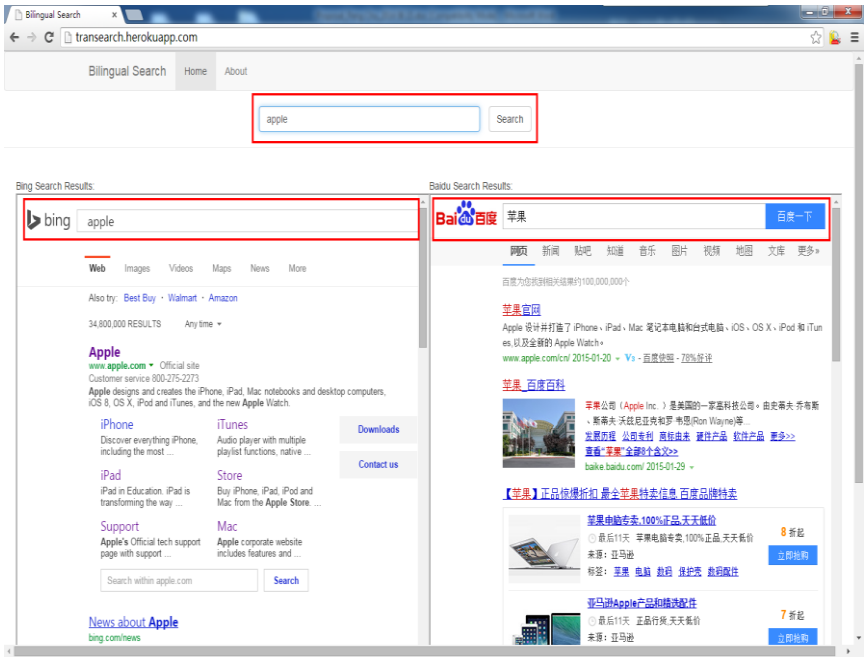


Figure 6-3. TranSearch 1.0 UI

TranSearch 1.0 accepts queries and returns results in English and Chinese. It has a top "uni-box" text box that takes in the initial query and recognizes its language. Below the uni-box, there are two "sub-boxes", one of which shows the original query and the other displays the machine-translated query in the other language (by using Microsoft Bing Translator API). Changing the text in the uni-box will refresh the two sub-boxes, whereas modifying one sub-box will not change the other sub-box or uni-box. This

kind of user interaction is utilized by online translation tools and dictionaries, such as Google Translate and YouDao digital dictionary, and can be intuitively perceived. The queries in the two sub-boxes are the input to two search engines (by calling Bing search API for English and Baidu search API for Chinese). Results of the two engines are returned and displayed in two parallel language panels under the corresponding sub-boxes. The width of the panels can be adjusted to accommodate user preferences.

Our prototype varies from other system UIs by integrating intuitive user interaction during the query construction phase and accommodating user preferences in search result displays. Specifically, it entailed the following functionalities and features:

1. Query construction

Specifically, TranSearch 1.0 provides the following functions during query construction phase (see Figure 6-3):

- Uni-box and sub-box allow more flexible query input

  The prototype allows users to enter queries and view the resulting documents in two languages: One universal search input box (uni-box) on top of the UI with language recognition (either Chinese or English query will be recognized); Two sub-boxes in parallel panels (displaying results in each language separately) within a single-page layout.

- Easy query formulation and reformulation in sub-boxes

  Our prototype allows users to modify queries either in the original language or

the translated format: Two sub search boxes below the universal search box automatically translate the query into both English and Chinese and start the search; The queries in these sub search boxes can be modified to assist query reformulation and do not override the query in the universal search box.

- Uni-box and sub-box help keep track of query translation/reformulation

Changing the uni-box will refresh the two sub-boxes, whereas modifying one sub-box will not change the other sub-box or uni-box. This kind of user interaction is utilized by online translation tools and dictionaries and can be intuitively perceived.

2. Result presentation

To give users control and accommodate user preferences in search result display, the size of the two panels can be adjusted. Specifically, the following functionalities are provided:

- Provide a default view that is preferred by most users in Steichen and Freund's study (2015): single page, non-blended, and double-panels
- Offer the flexibility to adjust the size of panels by dragging and dropping the border line of the panels
- Two panels work independently and utilize two search engines, Bing (for English) and Baidu (for Chinese)

Using Baidu for Chinese search was preferred by some users (Chu, Komlodi, & Rozsa, 2015; Rozsa, Komlodi, & Chu, 2015). But a search engine focusing on a specific language may not exist for all languages, TranSearch in future should offer users the ability to select preferred search engines. Currently, if the language is not Chinese, default search engine would be Bing.

6.4 User Study: Feedback on TranSearch 1.0 (RQ3.2)

During the subsequent user study, data were collected on how users used the TranSearch prototype, user feedback on strengths and weaknesses, and possible improvements.

6.4.1 User Study

Qualitative research methods were utilized during the user study, since they suited the nature of the research, exploring user interaction with the prototype and identifying and understanding system strengths and weaknesses. Specifically, investigator observation and notes on search task behavior and processes, semi-structured interviews, and structured questionnaires with open-ended questions were utilized. Study sessions were conducted with individual participants, leaving the researcher plenty of time to observe their usage of the prototype. Interviews and questionnaires gathered data answering the research question about system strengths and weaknesses, as well as providing reasoning for them and potential improvement opportunities.

Ten individuals were recruited to participate in the user study, of which six were female. The average age was 27.5 (see Table 6-1). Due to the fact that our prototype

was built on the English-Chinese language pair, all participants were chosen from native Chinese speakers. Two other qualifying criteria were utilized for participant screening. Participants needed to be able to speak English at a conversational level. The second criterion was previous English search experience and familiarity with English online search systems. To satisfy this criterion, participants needed to perform online search regularly in English. Those who were included in the study reported searching in English three to ten times per day (6.9 on average). The criteria guaranteed our prototype was tested in a meaningful way.

Demographics	
Ave. Age	27.5
SD Age	2.3
Male	4
Female	6

Table 6-1. Participant demographics

During the user study sessions, qualitative data were collected through search task observation, investigator notes, interviews, and questionnaires. The detailed search tasks, study procedures, and data analysis methods were described in Chapter3.

6.4.2 Results and Discussion

When participants were performing the search tasks with TranSearch 1.0, no introduction was given to the TranSearch interface, with the purpose to see how intuitive the interface was. All participants completed the tasks with only questions

on task instructions. The average amount of time spent to complete the tasks was 35 minutes and 47.5 seconds (or 2147.5 seconds, with a SD of 357.6).

- TranSearch Strengths (RQ3.2)

When talking about the strengths of TranSearch, most participants clearly stated that TranSearch was helpful when they were looking for information in English. Some indicated that TranSearch usage could be situational. For example, it was more useful for tasks that need to compare search results in two languages (e.g. Task 2, searching for an unfamiliar medical term) and less useful for tasks that do not need results in another language.

1. Query construction phase

    1.1. Automatic query recognition and translation

    Using MT to automatically translate queries is implemented by many systems. TranSearch also utilizes MT, the MicroSoft translator API specifically, to recognize and translate the query from one language into the other. While MT can have problems, participants in the user study found it useful. It helps the procedure of human translation and dictionary is no longer needed and, more importantly, the translation quality is acceptable. *"TranSearch saves my time translating keywords and translation is similar to a digital dictionary and acceptable"* (Interview P665; similar comments in interview P698, P695, P602, P648, P693).

TranSearch automatic translation was especially helpful when participants were not familiar with some terms, e.g. medical terms. *"I sometimes use digital translation dictionary software/APP (YouDao digital dictionary) to translate my keywords from Chinese to English when I do not know it in English"* (Interview P602).

1.2. Query modification history and reformulation assistance

The visibility of the machine translated queries and the flexibility to modify them are not readily available for many MLIR systems (e.g. 2lingual.com; Steichen and Freund, 2015). TranSearch, on the other hand, introduced the separation of the initial search input box (Figure 6-3. uni-box) and two additional modifiable sub search input boxes (Figure 6-3. sub-boxes). This feature was intuitively understood and used. The text boxes offered visibility of initial queries, translated queries, and reformulated queries, which was helpful.

"I can see what I entered and what were translated into...I can also modify the translation and do not have to change my initial terms every time...this is very useful" (Interview P693). *"It shows all the keywords (in both the uni-box and the sub-boxes) …very clear... helpful when I want to change keywords or go back"* (interview P602).*"Does it translate the search term? Oh, it puts it (the translation) in the box below it...that makes sense and I can see what search terms (in different languages) give me what results"* (Interview P639).

However, simply displaying the current translation and modification was not sufficient. Participants indicated that it would be more useful to provide a query modification history to track their queries entered, translated, and reformulated. This was not built in the current version of TranSearch, but would be a beneficial functionality of MLIR and could facilitate query reformulation.

2. Search phase

Most contemporary systems utilized only one search engine, while TranSearch was devised to make use of two engines, Bing and Baidu. It was designed this way since people tended to use different search engines when searching in different languages. For instance, results from a study on Chinese students searching in English (Chu et al., 2015) indicated that search engine selection depended on languages, specifically, searching in English with Google, while in Chinese with Baidu. Participants in this study had a preference in search engine selection.

"It has dual search engines, Bing and Baidu...This is how I search usually...I search for Chinese in Baidu and English with Google. I haven't used Bing very much though, but it is ok, but it would be even better if I could pick what (search engine) I use" (Interview P698).

3. Result display phase

As introduced in section 6.3, TranSearch is designed as a MLIR system with a

UI of single page view, non-blended, and double-panel (Figure 6-3). It was not surprising to see comments were around two aspects:

- Informativeness in two languages and two engines

"It is like hitting two birds with one stone... in two languages and with two search engines, both Baidu and Google (the participant meant Bing)" (interview P602). *"It gives me more information. Sometimes I do not really need the results in Chinese, but it is good to have them there if I needed"* (Interview P642).

- Assistance in comparison

"It is paneled. I primarily search for English, but TranSearch is useful to provide Chinese results as an auxiliary...sometimes comparison is so helpful that I did not even need to open any result websites (for search task T3, lookup Glucosamine)" (Interview P698).

- TranSearch Weakness (RQ3.2)

1. Situational usefulness

It was indicated by some participants that TranSearch was less useful, or not useful at all, when they only searched or looked in one language (for tasks that do not need to compare search results in two languages) or they went directly to a website (not searching at all).

2. Lack of familiarity

Some participants were used to traditional search UI and said TranSearch slowed them down due to unfamiliarity.

3. Limited panel visibility and resize adjustment

The panels separated the screen into two and sometimes the panel did not display the complete result title and short description. TranSearch offered the drag and drop resize adjustability to make viewing results in a panel easier, but this was not easily perceived.

Some participants preferred left-right panels on a desktop, whereas top-bottom panels on devices with smaller screens. However, TranSearch does not provide layout adjustment. An easy way of adjusting both panel size and layout was lacking.

4. Insufficient query formulation and reformulation assistance

Besides query translation, participants were expecting more assistance in query formulation and reformulation and disappointed at lack of such functionalities: *"TranSearch (in Task 3) is just like a digital dictionary, so it is only sort of helpful...with limited usefulness"* (Interview P600). Assistance could come in many forms, such as a recommendation system suggesting queues in other languages, or an image search functionality.

5. Lack of result set matching

TranSearch results between the two panels were not matched. It was expected to have some sort of result matching between the two panels.

6. Lack of engine selection

The fact that TranSearch did not offer the ability to select search engines had a negative impact on user experience. *"I will never use it (TranSearch) if it does not allow me to select my favorite search engine (Google)"* (Interview P665).

- Improvement (RQ3.2)

Three types of user preferences were identified for result display: panel size, panel layout, and an option to hide a panel. Preference of panel size referred to the ability to adjust the size of a panel to expand it for better readability. Panel layout demanded the system to offer the ability to change the UI layout, such as a left-right layout on bigger screen and a top-bottom layout on smaller screen. Hiding a panel was also needed by users when they completely focused on results in only one language.

Participants provided improvement ideas on TranSearch 1.0. Figure 6-4 shows a wireframe drew by a participant demonstrating improvement ideas. Based on the weaknesses, user study results suggest the most important improvement can be made in the following areas:

- Better query formulation and reformulation assistance, other than MT

- Offering users the ability to hide a language panel for certain tasks

- Offering better result panel size and layout adjustment

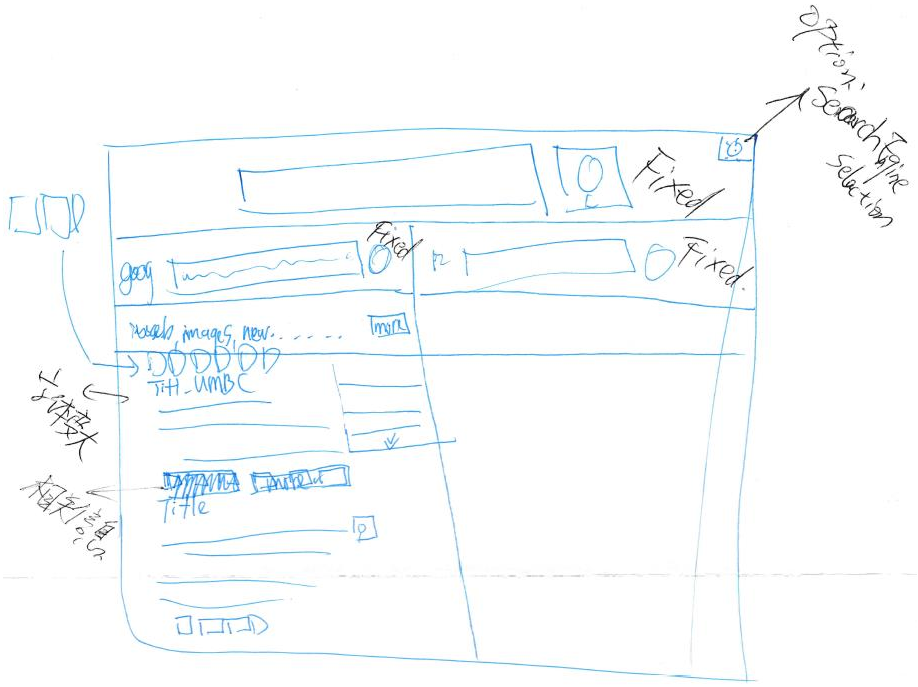


Figure 6-4. Wireframe drew by a participant during user study

Results from the interviews showed that about one half of the participants had never heard of a multilingual IR system or UI and it was the first time they tried out such a system and UI. Those who used such systems previously indicated they would use the systems again and those that were not aware of the existence of such systems stated they would like to try if there were established systems. Lacking usable and prevailing multilingual search systems calls for raising the awareness of such systems and research that provides insights on solution with appropriate affordances and features. *"I use traditional search engine, because I never heard of the search engine*

UI that you could search in two languages together" (Interview P665). *"I will use it, but I care about its popularity"* (Interview P642).

Chapter 7: TranSearch 2.0: Redesign and Evaluation

The previous study focused on prototyping, implementing, and evaluating TranSearch 1.0. A follow-up user study provided user feedback, which shed light on the prototype's strength and weakness. Based on the user feedback and study results, possible improvements to the prototype were discussed.

In this chapter, a study aiming at extending the previous is carried out in two steps. First, based on the findings from the previous study, TranSearch 1.0 is redesigned and rebuilt, from a prototype to a more functional system and UI with more features, TranSearch 2.0. Subsequently, a user study is carried out with two research purposes:

1. To evaluate the new functionalities and features of TranSearch 2.0 and explore further design options;

2. To explore general user interaction needs of MLIR/CLIR systems.

Different than the previous study, this study focuses more on the general CLIR/MLIR needs, and the critical functionalities and features of a successful CLIR/MLIR system and UI should possess. Evaluation results are discussed and the user interaction model abstracted in previous studies is reexamined. Based on the user redesign input, final design implications are addressed. This study aims at answering the following research questions:

RQ4.1: What do TranSearch 2.0 evaluation results indicate, in terms of system usability and user experience?

RQ4.2: Does TranSearch 2.0 accommodate the user interaction model and assist in user challenges and query formulation and reformulation strategies identified in previous studies?

RQ4.3: What are the user preferences and why? What are the crucial MLIR functionalities and features and why? What are the design implications can be provided to MLIR systems and UIs?

The data collection and analysis followed the general methods introduced in Chapter 3 and details were described in section 3.4.4.

7.1 Introduction

Results from the user study of TranSearch 1.0 informed the design of TranSearch 2.0 (see Figure 7-1 for the final wireframe, 7-2 for the default UI, and 7-3 for UI transformation of the new version, TranSearch 2.0).

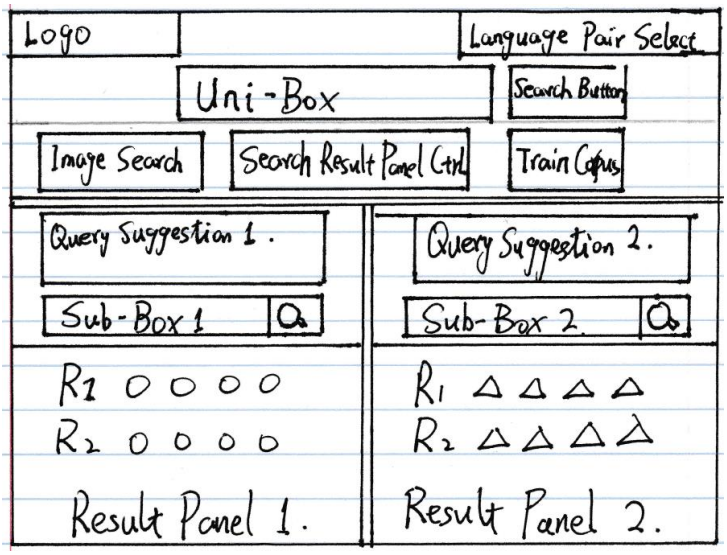


Figure 0-1. TranSearch 2.0 wireframe

The components and UI controls of TranSearch 2.0 (see Figure 7-2, numbered from top-left to bottom-right) are described below. Features that are new compared to TranSearch 1.0 are marked as “new”.

1. TranSearch logo: back to homepage when clicked (new in TranSearch 2.0)

2. Drop-down list: change language pair: English-XXX language (new)

3. Uni-box: enter and hold initial query

4. Search button: recognize initial query's language, translate it to the counterpart in language pair, perform search in two search engines, display result in two panels

5. Top three images associated with initial query (new)

6. Change view control: change the layout of result panels (four options: left-right view, top-bottom view, hide English panel, hide the other language panel) (new)

7. Train TranSearch: part of the functionality to build query corpus (new)

8. Top three English queries associated with the query in the other language (new)

9. Top three target language queries associated with the English query (new)

10. Sub-boxes: hold automatic resulting queries from initial query, can be modified by users without changing initial query or the other sub-box query

Figure 0-2. TranSearch 2.0 default UI

Panel A: Left-right

Panel B: Top-bottom

Panel C: Single

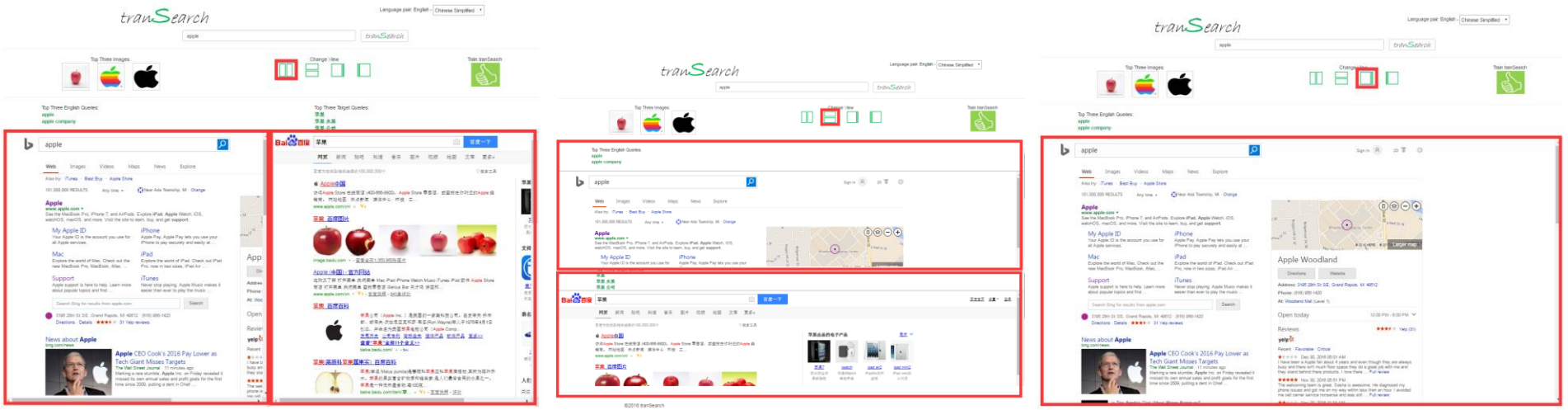


Figure 0-3. TranSearch 2.0 UI transformation

7.2 New for TranSearch 2.0

After the redesign, TranSearch 2.0 is implemented as a separate Web application from the previous version (rewritten in VB and ASP.NET language and with a backend database using Microsoft SQL Server) and entails the following new functionalities and features:

1. A query corpus as an auxiliary to MT

A straightforward approach, MT, has been utilized by most contemporary CLIR and MLIR systems and UIs (Grefenstette, 1998). However, MT quality faces a series of challenges, a major one of which is ambiguity resolution. Additionally, MT disambiguation can be even more challenging for query translation, since queries tend to be short, sometimes without internal structure and providing little or even contradictory syntactic analysis MT depends on (e.g. Grefenstette, 1998). Moreover, the application of MT in query construction of MLIR systems is based on the assumption that high quality of query translation will result in better search results.

Our previous studies and many other studies indicated this premise might not hold and relying merely on MT might not work for all MLIR query construction strategies and scenarios (e.g. Peters et al., 2012). In previous studies, we described two search strategies, translation strategy and discovery strategy, and abstracted a user interaction model, iterative learning and reformulation. Without user-editing interaction ability, MT only partially supports translation strategy. Moreover, if MT offers one definitive translation without variation, it does not support discovery

strategy.

Participants in the user study of TranSearch 1.0 also stated that more support for query construction and reformulation was needed. Thus, TranSearch 2.0 includes a query corpus, a backend database stores the query pairs and image links, as an auxiliary to MT. The data in the query corpus are gathered by "Train TranSearch", which is an optional functionality that allow users to manually adjust or input query pair, as well as an option to link an image, if they think the MT query translation is not accurate or translation itself is accurate but does not account for a good query. Figure 7-4 is the front end of "Train TranSearch" functionality.

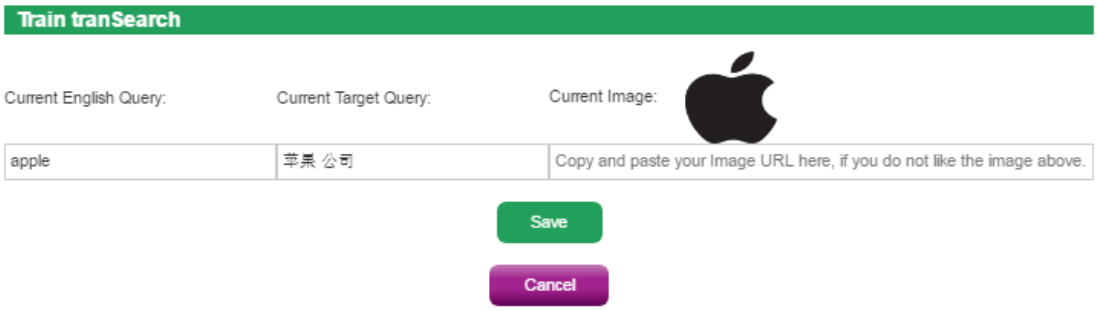


Figure 0-4. Train TranSearch functionality front end

Although corpus-based approaches are not rare in Machine Translation area (e.g. Jean, 2006; Okpor, 2014), a query corpus is different than MT corpus. In terms of content, a query corpus consists of queries in different languages, whereas MT corpus contains words or terms in different languages. Regarding purpose, a query corpus is to find queries in another language leading towards relevant search results, whereas MT corpus is used to find accurate translation. Zhang, Sun, and Min (2005) utilized

the Web corpus to translate the queries in cross-lingual information retrieval. They used the search engine to find out the corpus data in the target language on the Web by submitting the query in source language. In our opinion, they were still trying to find an accurate query translation out of the corpus and their corpus data were limited to search engine logs. Our query corpus does not require the query pair to be an accurate translation and allows image linking. Figure 7-5 shows the interdependency of the query pair and the linked image.

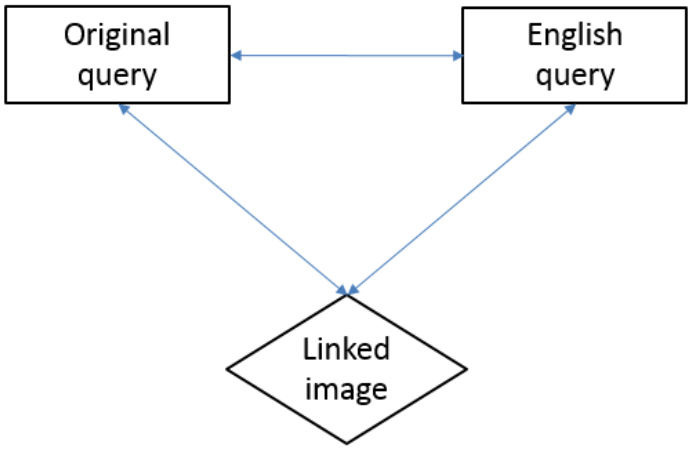


Figure 0-5. Interdependency of original query, English query, and linked image

Another difference is data collection for our query corpus. The data are gathered from direct user input with a purpose of finding the most relevant search results in another language when MT is not accurate of does not suite the search context. Figure 7-6 shows our query corpus database (developed using Microsoft SQL Server).

QueryID	en	zhCHS	es	hu	imageURL	humanCorpus
372	apple	苹果	NULL	NULL		0
373	gloucosamine	gloucosamine	NULL	NULL		0
374	apple	苹果	NULL	NULL		0
375	reatrant	reatrant	NULL	NULL		0
376	reatrant New york	reatrant 纽约	NULL	NULL		0
377	Nike	耐克公司	NULL	NULL		0
378	qpple	qpple	NULL	NULL		0
379	NYC top resturant	纽约顶尖餐厅	NULL	NULL		0
380	reataurant New york	纽约的餐厅	NULL	NULL		0
381	chinese reataurant New york	中国餐馆纽约	NULL	NULL		0
382	New York Chinese restaurants	纽约 中餐馆	NULL	NULL		0
383	Chinese restaurant in NewYork	chinese restaurant in NewYork	NULL	NULL		0
384	New York Chinatown Chinese restaurant	纽约 唐人街 中餐馆	NULL	NULL		0
385	reataurant New york	纽约的餐厅	NULL	NULL		0
386	reataurant New york	纽约的餐厅	NULL	NULL		1
387	Glucosamine	氨基葡萄糖	NULL	NULL		0
388	American restaurant in NewYork	在纽约的美国餐厅	NULL	NULL		0
389	how to use in on at	如何对在使用中	NULL	NULL		0

Figure 0-6. Query corpus database

1.1. Query recommendation (Figure 7-2, UI element 8 and 9)

The query corpus extending MT enables query recommendation based on the frequency of query pairs (English query and its counterpart language query) in the query corpus. When users input an initial query, query pairs in the corpus are filtered by this query. Then the top three most frequent counterpart of the query pairs associated with the initial query are returned as its query recommendation. The query recommendation serves as an auxiliary to MT and accommodates the query translation and discovery strategies and support the entire iterative search process.

1.2. Image querying (Figure 7-2, UI element 5)

Another feature that supports query formulation is displaying images that link to query pairs. The image querying functionality depends on data gathered from "Train

TranSearch" functionality. When users train TranSearch, they enter the query pair, with an option to link an image (by providing a URL of an online image). So, the query corpus database stores query pairs and the linked images. Thus, when users use TranSearch and enter an initial query, the linked images are presented on UI if they exist in the corpus. Clicking the image will load the linked most frequent queries in both languages and perform the search with them. Figure 7-5 illustrates that original query, English query, and linked image are interdependent in the corpus. For any given one variable, the system is able to return the most frequent other two variables.

2. Multilingual support (Figure 7-2, UI element 2)

TranSearch 2.0 supports multiple languages, each of which forms a language pair with English. There is a drop-down list that offers the ability to change language pair (English and XXX language) on top-right of the UI (Figure 7-4, component no. 2). The system utilizes the Microsoft Bing Translator API to recognize and translate queries from the other language to English, or the other way around.

3. Result display layout control (Figure 7-2, UI element 6)

The results of the evaluation of TranSearch 1.0 suggested that users needed more flexibility in controlling the layout of the panels. TranSearch 2.0 offers the control buttons to change the layout of result panels (four options: left-right view, top-bottom view, hide English panel, hide the other language panel). This is an easy and supplemental feature to the drag and drop resize function. It offers two options to modify: the number of panels shown (one or two) and the layout of the panels

(horizontal vs. vertical split).

7.3 User Study: Evaluation and General CLIR/MLIR Design

After the researcher redesign, a user study was carried out with two research purposes: 1) to evaluate the new functionalities and features of TranSearch 2.0; 2) to explore further user redesign options and general user interaction needs of MLIR/CLIR systems. Accordingly, the user study was divided into two consecutive sessions:

- The first session, evaluation:

Qualitative evaluation data were collected through investigator observation and notes of the search task processes and behavior, guided focus group discussion on the new features, and structured questionnaires with open-ended questions.

- The second session, user redesign:

An interactive group setting helped brain-storming and generating more ideas. Focus group discussion was also beneficial to suggest potential solutions and design ideas to problems identified. Data were first collected via TranSearch 2.0 UI reviews. Participants reviewed the interface and provided verbal feedback. Next, they brain-stormed and sketched improved designs for the interface. General suggestions on CLIR/MLIR system and UI were discussed.

A total of twenty one participants (see Table 7-1 for demographics) took part in five focus groups, four of which consisted of native Chinese speakers (Group 1, 2, 3, and 4) and one contained other languages speakers (Group 5: English, Arabic, and

Persian). Comparing to study 3, TranSearch 2.0 supported multiple languages and the participation of other native language speakers increased diversity of the participant sample.

Group	Participants	Language Pair (native - non-native)
Group 1	2	Chinese-English
Group 2	5	Chinese-English
Group 3	5	Chinese-English
Group 4	4	Chinese-English
Group 5	5	Arabic-English, Persian-English (2 participants), English-Japanese, English-Spanish

Ave. Age	30.0
SD Age	4.44
Male	13
Female	8

Table 0-1. Demographics

Similar to the previous study, two qualifying criteria were utilized for participant screening. The first criterion was that the participants' English (or another foreign language for native English participants) proficiency had to be at a conversational level. The second was that they had previous search experience in English or another foreign language so that they processed familiarity with English or another foreign language online search systems. These criteria guaranteed the new features of

TranSearch 2.0 were evaluated in a meaningful way, as well as the quality of data of experienced users' redesign ideas and general system suggestions. Non-native English participants reported searching in English for at least three times per day. Almost none of our participants had experience using CLIR/MLIR systems.

From the user study, qualitative data were collected through search task observation, investigator notes, focus group discussion, and questionnaires. The detailed search tasks, study procedures, and data analysis methods were described in Chapter3.

7.4 Results and Discussion

In general, the participants had a positive experience with the TranSearch 2.0 system and UI. All participants completed the tasks with no difficulties. They also confirmed the usefulness of the features built in the system and offered explanations as to why these were useful.

7.4.1 Search Topics (RQ4.1)

Participants indicated that search topics had an impact on system usage and TranSearch was beneficial especially to topics such as academics, news, sports, and shopping (see Table 7-2 for the top five topics mentioned). We counted the mentioned topics and analyzed them. These topics benefited from TranSearch because participants were not familiar with these topics or they needed to check and compare results in both languages. *"It would be helpful especially in areas and topics that I am not familiar with... for example, Chemistry"* (G1, 2, and 3).

Topics	Mentions
Academics	24
News/sports	21
Shopping	16
Health/Medicine	13
Culture	11

Table 0-2. Top five topics TranSearch is useful

Some of the participants mentioned that for some topics, Transearch was quite handy by offering different aspects of the same topic. For instance, when they were searching for a disease, they looked at English results for scientific explanation, whereas Chinese results for cultural aspects and patients' personal feelings. This confirmed the findings in Chapter 5 that for certain search tasks, users focused on different aspects in different language. TranSearch especially suited such user need by providing results in both languages.

7.4.2 Result Display Layout

As stated above, our results confirmed previous study findings that user preferences depend on tasks and context (Chu et al., 2015). While for some tasks searching in both languages is important, for others bilingual users would just use one language. Systems and UIs are required to provide flexibility to accommodate the various preferences. For certain tasks that users search and focus on only one language, displaying an additional panel may not be necessary and can actually be "noisy". *"Sometime I need only one panel (two panel are not desired)...It needs to be able to*

show whole screen, I mean to maximize and switch between language panels" (G2, 3, 4, and 5). TranSearch 2.0's new feature offering the ability to hide a result panel was very beneficial in this situation.

In a recent study, Steichen and Freund (2015) found that single-page displays with clear language separation were mostly preferred by their participants. Our study followed up by explaining why these were preferred. A single-page view with non-blended language panels facilitated result comparison, as well as kept results organized. *"The two panels are clear and so simple...I can compare results all together...on the same screen"* (G1, 3, 4). Participants indicated that they preferred such an interface because of its simplicity and visibility.

This study provided new insights that users' preferences were not fixed and might vary depending on devices and contexts. Some participants preferred left-right panels on a desktop, where as preferred top-bottom panels on devices with smaller screens. *"TranSearch paneled screen is small (when panels were horizontal), sometimes it is not able to display the whole results, I need to drag the scroll bar (vertically) to see all (because text was not wrapped)...It would be easier to see results in a top-bottom layout"* (G1, 4, and 5). In panels of top-bottom layout, the results would be fewer, but each result was entirely displayed.

7.4.3 Re-examine the user interaction model, challenges, and query strategies (RQ4.2)

Our previous studies highlighted the top three challenges for NNES searching in

English:

- Challenge 1: Query Formulation
- Challenge 2: Query Reformulation
- Challenge 3: Viewing and Skimming Search Engine Result Pages (SERPs) and Websites

Our previous studies identified two query construction strategies, the translation strategy (Figure 5-1) and the discovery strategy (Figure 5-2) and abstracted a user interaction model that incorporated both: iterative learning and reformulation (Figure 5-3) depicting the entire search process. In Chapter 6, we described how TranSearch 1.0 was designed to address these challenges, strategies, and the user interaction model. TranSearch 2.0 provided new features and functionalities to strengthen the support.

During query construction phase, TranSearch 2.0 introduced the query corpus, which enabled query recommendation and image querying. Their usefulness was justified in the user evaluation on the new features. These query assistance tools served as an auxiliary of MT and strengthened the translation strategy. Also, by offering additional recommended queries and linked images, they accommodated the discovery strategy. When users clicked on the recommended queries or linked images, another round of iteration of search was triggered. The iterative nature of search was supported.

For result display, TranSearch 2.0 offered layout control (Figure 7-2, UI element 6).

Four control buttons in the navigation bar provided users the ability to change the layout of result panels (left-right vs. top-bottom) or hide one of the two panels. This new feature further supported users' challenge of browsing SERPs.

7.4.4 Design Implications for CLIR/MLIS Systems (RQ4.3)

Our study identified a series of characteristics and functionalities that are crucial to the success of CLIR and MLIR systems and interfaces:

- Following established IR conceptual model and offering intuitiveness and simplicity

Online searchers are so used to market-leading search engines, such as Google, Bing, and Baidu, that following their established conceptual model is important for user acceptance. In addition, offering intuitiveness and simplicity is a key to the success of such system design. *"I like that it (TranSearch) is so simple that I know what I need to do without anyone showing me how"* (G1, 2, 3, and 4). Supporting search engine selection across language panels is also a necessity (user redesign G1, 2, 3, 4, and 5).

Using Baidu for Chinese search was preferred by some users in Chu et al. study (2015). But a search engine focusing on a specific language may not exist for all languages, TranSearch in future should offer users the ability to select their preferred search engines. Currently, if the language is not Chinese, default search engine would be Bing.

- Query translation, modification, and recommendation

Query translation needs to be clearly displayed and provide the ability to keep track of query translation and reformulation history. The searchers should be able to change the translated query terms (user redesign G1, 3, 4 and 5).

Machine translation is insufficient and query construction assistance, such as query recommendation and image link, is beneficial (user redesign G3, 4 and 5). Though building the query pair corpus is necessary for query recommendation and image link, having users “train TranSearch” does not seem feasible and other approaches to building the query corpus are called for (user redesign G1, 2, 3, 4 and 5).

One query recommendation feature that was not implemented in TranSearch 2.0 but the need for it was clear from the data is a history of the queries entered and modified in both languages. The searches are highly iterative and providing easy access to the list of queries in the session will help searchers keep track of progress and create new queries.

- Result display

For CLIR/MLIR systems and UIs, it is essential to provide users the ability to compare results in multiple languages. Offering additional results in another language works by providing auxiliary information. However, the extra cognitive load introduced by providing results in other languages also demands customizability, such as rearranging and resizing language panels. For example, if users do not want results in a language, they may close that panel. Flexibility and customization could accommodate various user preferences as well, such as allowing for changing the

layout across devices (user redesign G2, 3, 4, and 5). It is also important that user customization and preference (e.g. language selection, search engine selection, and panel arrangement) could be memorized by the UI by providing a login functionality (user redesign G3, 4, and 5).

The text boxes offering visibility of initial queries, translated queries, and reformulated queries are helpful. However, it would be more useful to provide a query modification history to track queries entered, translated, and reformulated. This is a crucial functionality to the usefulness of MLIR and facilitate query reformulation.

Some of the heuristic guidelines are particularly important to the success of CLIR/MLIR systems and UIs, such as visibility of system status, user control and freedom (user redesign G2, 3, 4 and 5). The fact that there are usually more UI elements of CLIR/MLIR systems, comparing to those of single language search engine UI, adds up system complexity. For example, each result display panel should have its own result website URL box and "going back/forward" buttons. They need to be easily perceived. Otherwise, navigation and orientation in multiple result panels can be difficult and confusing.

Though UI features and controls (e.g. query translation recommendation, image links, and result panel adjustment) are beneficial, users should be offered the ability to manage them. For instance, in certain context and for some tasks they are not necessary and users should have the ability to hide them or move them to the top or bottom of the UI to make room for the query boxes and search result panels, the

necessary parts of the UI (user redesign G1, 2, 3, 4, and 5).

- General implications for CLIR/MLIR systems

The usefulness of CLIR/MLIR systems is situational and depends on two factors, task topics and contexts:

- Task topics

  Some tasks do not need to search in another language or users determine that searching for tasks in a certain language would not provide useful results (e.g. searching Chinese restaurants in Persian did not make sense, G5).

- Task contexts

  MLIR UI needs to work across devices, especially on smaller screens. MLIR is more helpful for search tasks that are more complex, need additional result sets, or require result comparison (e.g. search for medical terms, G2, G3, and G5).

In some cases, users tend to focus on different aspects of a topic when searching in different languages. For instance, when searching for the restaurant recommendation task, participants focused on restaurant introduction in English, whereas personal feelings of other users in their native languages, which they indicated they tended to trust more (G2, G5).

Users need more control over the navigation and display control area. It was

suggested during user redesign session that users should have the ability to hide or move the additional UI control buttons to make the entire UI organized and save more room for result display (G3, G4, G5).

Being able to select the search engine users preferred is a required functionality (G1, G2, G3, G4, and G5). Also, the design of MLIR systems should follow accessibility standards. For example, the system needs to work with screen reader and alternative description text needs to be added to the web site (G5).

A query corpus may serve as an auxiliary to MT and enables query recommendation and image querying. However, collecting human input data for the query corpus is challenging. During the evaluation, though most of the participants affirmed the corpus was beneficial, very few indicated they would be willing to manually train TranSearch in real life usage (G3, 4, and 5). The biggest challenge to building the query corpus lied at finding a more automatic way or crowd sourcing of collecting query pairs and other corpus data, e.g. image link. Another problem with "Train Transearch" was that it needed to be more visible on the UI and provide quick and clear instructions to the users.

7.5 Summary

The current system solution, automatic machine translation of queries, is insufficient since it focuses only on query construction phase, let alone it sometimes may fail due to disambiguation failure and lack of search context. New system and interface solutions would need to assist in both query construction and result presentation. In

this Chapter, we redesigned TranSearch 1.0 prototype to provide intuitive user interaction in query translation and reformulation, as well as accommodate user preference in search result presentation in bilingual searching. A new approach, building a query corpus to extend MT, was introduced and tested. Its purpose was to support our previous findings about the entire iterative learning and reformulation search process and accommodate query construction strategies, translation strategy and discovery/learning strategy.

Various features of TranSearch 2.0 were evaluated in a qualitative user study in the two traditional CLIR phases, query translation phase and search phase. During the second half of the user study, participants were engaged to provide further redesign ideas and suggestions. Results highlight that 1) during query translation phase, intuitive user interaction, specifically, automatic machine translation with perceived user ability to edit and keep track of queries is especially beneficial; 2) in search phase, result displays facilitating comparison across languages, customizable result panels accommodating user preferences, and flexibility to switch result languages play important roles. It was indicated that user preferences might vary with different search tasks and contexts. UIs need to be flexible and accommodate these preferences. Simplicity, visibility, and customizability were the most valued system characteristics. Identified crucial functionalizes and features include search engine integration, query modification and recommendation, and result comparison. UIs need to accommodate user control and freedom, user interaction, and preference.

Chapter 8: Conclusion

In this chapter, I summarize the studies, provide design implications, and discuss limitation of my work and future research opportunities. The dissertation identifies three gaps in the previous research on NNES search behavior and system support: 1) NNES (particularly targeted) behavior has not been closely studied, findings and models are insufficient; 2) MLIR system support is not widely available to assist them, query assistance depends only on MT, interaction type and level are arguable; and result display UI preference reasoning is lacking; 3) Human behavior research is not linked to system design and development.

In order to close the gaps, I target and study NNES behavior using both quantitative and qualitative methods to gain a deep understanding that is insufficient previously. Subsequently, I design and evaluate query assistance tools and result display UIs in two iterations to arrive at more generalizable design guidelines. The system design is tied to design practices in previous literature and my behavior findings, though not all are captured in the prototype.

Accordingly, the dissertation research is designed to be two stages with mixed research methods, an exploration and understanding stage to study search behaviors and a system design and evaluation stage to assist them. The two stages have different research focuses: Stage 1 focuses on understanding NNES behavior, whereas Stage 2 aims at providing system and UI support. Stage 2 is built on the findings from Stage 1 and further informs the understanding and models generated from it. The dissertation study bridges the gap between user studies of search behaviors and system design

practices by closely coupling and linking the two.

8.1 Major Findings

The dissertation targets a specific user group, the large and growing community of non-native English speakers who conduct online searching in English. Existing research is insufficient to fully understand the search behaviors, habits, and processes of the target audience. Even less is known about their search strategies and their use of systems and tools to accommodate their needs and assist their behaviors. The major contribution of this study resides in two parts:

- Stage 1 (Studies 1 and 2) - Exploration and understanding stage

The dissertation extended the current research on information seeking behavior and provided a better understanding of NNES searching behavior in English. Results indicated that searching in a foreign language requires significantly longer time, more query reformulations, and more websites viewed. User feedback also indicated that people tended to utilize different search strategies when searching in a second language than those in native language search.

Findings in this stage also shed light on NNES search challenges and behavioral patterns of language selection, search engine selection, query formation and reformulation, strategies, and browsing and filtering SERPs. NNESs faced a unique set of challenges that may not be present for native speakers when searching in English. Three most important challenges were identified: query formulation, reformulation, and result viewing. Two search strategies, translation strategy and

discovery strategy, were described and a user interaction model, iterative learning and reformulation, was abstracted to depict the entire search process.

- Stage 2 (Studies 3 and 4) - Design and evaluation stage

I reviewed the literature and created the figure to show the UI classification. Based on findings of Stage 1, I designed and prototyped TranSearch 1.0, a MLIR system and UI that allowed users to search in two languages at the same time. The system was built on previous CLIR and MLIR system research and integrated intuitive user interaction during the query construction phase and accommodated user preferences of search result displays. The design process of TranSearch and how the UI prototype accommodated the identified search challenges, the two search strategies, and the user interaction model were explained. Through a user study, the prototype strengths and weaknesses and redesign recommendations were proposed and discussed.

Subsequently, based on the findings from the previous study, TranSearch 1.0 was redesigned and rebuilt, from a prototype to a more functional system and UI with more features, TranSearch 2.0. A user study was carried out 1) to evaluate the new functionalities and features of TranSearch 2.0 and explore further design options; 2) to explore general user interaction needs of MLIR/CLIR systems. This study focused more on the general CLIR/MLIR needs, and the critical functionalities and features of a successful CLIR/MLIR system and UI should possess. Results and findings not only provided information about the utility of the proposed solutions and design implications for MLIR systems and UIs to assist NNES users, but also further

informed the user interaction model for NNESs information seeking behavior.

8.2 Major Contributions

The dissertation contributions lie at three aspects:

- Search barriers

My dissertation extends previous research on NNES search barriers by closely investigating and separating language barriers and information literacy barriers. For language barriers, three most important language related challenges are identified for NNES when searching in English as a non-native language: query formulation, reformulation, viewing and skimming SERPs. Thus, the interaction and system design needs to provide new language support tools. For information literacy barriers, NNES searchers also face a series of difficulties in viewing SERPs and evaluating results. The system solution needs to indicate trustworthy sources and guide NNES searchers.

- Information seeking behaviors

The dissertation research provides a better understanding of NNES search behavior and describes patterns identified. When searching in English as a non-native language, NNESs spend significantly longer time, issue significantly more query reformulation, and viewed more result websites. They also mix language use when constructing search queries, viewing and skimming SERPs, and combining result sources.

Existing research literature on information seeking behavior covers the topic of query

construction and assistance. However, it usually focuses on tools and techniques, such as MT, assisting query translation. Research on query construction strategies that do not rely on translation, such as user query discovery strategy, is insufficient. When utilizing discovery strategy, NNESs face various challenges that native English searchers do not suffer, which needs further investigation. My dissertation extends previous research by closely comparing and contrasting query translation strategy and discovery strategy, generating behavior models to better describe NNES discovery strategy, and utilizing a user interaction model (integrative learning and reformulation model) to integrate both models and describe the entire search process. Discovery strategy cannot be fully supported by MT and requires new approaches other than translation.

- Interaction design

A systematic classification of system approaches and UI types is provided in my dissertation. In addition to system and UI design, I also focus on exploring the design for interaction. During query construction phase, TranSearch utilizes simplified interaction and requires less steps. A query corpus enables query recommendation and image linking and thus, serves as a different source than MT to assist query discovery strategy. For result display, TranSearch offers an easy switch to turn off multilingual mode and provides easy user control of panel layout and size, since these functionalities are indicated to be critical based on my findings on behavior.

8.3 Limitations and Future Work

One concern for this dissertation study is generalizability and transferability. In other words, research limitation lies in sample size and diversity of native language speaker groups. At the current study level, it is not feasible to verify the study findings are transferable to other native language groups than Chinese and Hungarian. In the future, I intend to expand the study sample population. The participant pool will be expanded both vertically and horizontally, by inviting more participants to the current participant pool and incorporating participants who speak other languages. The diversified data will not only increase triangulation, but also help us identify new themes.

English proficiency was one of the criteria of participant recruiting. It was recorded and identified from merely self-reported data. It can be argued that English proficiency varies in different aspects (e.g. reading and writing) and settings (e.g. conversational settings and academic settings) and potentially has an impact on searching behaviors and interaction with systems and tools. In future research, I will address these issues by utilizing a controlled English proficiency measure (e.g. a language test score).

Another limitation of this research is the lack of data to identify and separate the extent to which NNESs face challenges due to language difficulties or due to unfamiliar domain knowledge and cognitive issues (the latter may also be faced by native English searchers). Thus, attention in future work will also be paid to the differentiation of language difficulties from others, such as knowledge domain and

cognitive issues.

Another opportunity resides in analyzing cultural impacts. Cultural impacts start to draw attention of MLIR researchers (e.g. Chizari, 2016). However, there are no conclusive findings on the role that culture plays in information searching behavior of users. This dissertation touched the discussion on cultural aspects of search topics when searchers were selecting a language to search with and included a focus group of native speakers of languages other than Chinese and Hungarian to gather new insights. However, cultural impact was not the focus of this dissertation research and a future project could be dedicated to studying differences and patterns in the search techniques employed by various cultural groups and their interaction with online searching systems and tools.

A final opportunity is on the system solution side. Though the goal of this dissertation is not to build a fully functional MLIR system and UI, the solution can be improved in a number of ways. In future research, one direction is to improve user interaction to provide better query management, such as formulation, translation, and reformulation. From result presentation side, work could focus on result relevance, display, and correlation in multiple languages. During the query construction phase, though the proposed query corpus offers query recommendation and management, a viable way of accumulating query corpus records, not solely depending on MT or human efforts, is lacking. On the other hand, during result display phase, improvements can be made on the connection between results from two search panels or the integration of multi-

media search result types.

8.4 Summary

The disparity between the language distribution of Web content and the representation of speakers of different languages among Web users exists and many NNESs search in English to satisfy their information needs due to the lack or low quality of content in their native languages. The prevalent use of the Internet for information access means that knowing how to search in English has become an important element of information literacy globally. Current research is insufficient to fully understand the information behaviors and habits of NNESs when they conduct online searching in English. Even less is known about their search strategies, processes, and interaction with systems and tools to accommodate their needs and assist their behaviors.

Through four individual studies in two stages involving a total of 128 participants in two countries, the US and Hungary, this dissertation extends the current research by describing and interpreting the complex behaviors of NNES searchers, as well as testing solutions to assist them. The challenges, along with the root causes, impacting user performance and experience are addressed. The development and evaluation of an interaction model and a prototype UI enable us to provide insights for designing and building NNES search support systems.

NNES searchers face language-based challenges both in the query formulation and reformulation phase and the result viewing phase. They combine their native

language and English while searching and use both translation and discovery/learning strategies to construct and reformulate queries. They also combine information sources in both languages when reading results and learning about the topic of the search. User experience, in general, indicates lack of confidence, stress, and extra attention during their searches.

The current system solution, automatic machine translation of queries, is insufficient since it focuses only on query construction phase, let alone it sometimes may fail due to disambiguation failure and lack of search context. New system and interface solutions would need to assist in both query construction and result presentation. Simplicity, visibility, and customizability are characteristics of successful novel solutions. Identified critical functionalities and features include search engine integration, query modification and recommendation assistance, and result comparison aid. UIs need to provide user control, streamline user interaction, and accommodate user preference.

Appendices

Appendix 1: Focus Group Facilitating Guide

Focus group guide 分组访谈稿脚本

Introduction:介绍 ：

Dr. Anita Komlodi, UMBC

PhD student, Peng Chu, UMBC

Objective of the research: Use of English as foreign language in information search on the Web. 科研目的：网络信息搜索中英语作为外语的使用

Short introduction, first names. 简要介绍，研究员的名字。

First questionnaire, then conversation. 首先完成调查问卷，然后开始访谈。

On Paper: 纸面问题：

What kinds of topics did you search for in Chinese and in English? Why? 你曾经用英语和中文搜索什么主题？为什么？

What kind of search engines or websites do you use in the case of searching and browsing in English? 在用英语搜索和浏览时，你使用什么种类的浏览器或者网

站？

Orally: 口头问题：

How is it different your search/browsing in English and in Chinese? 你在用英语和用中文搜索或浏览中有什么不同？

What problems did you find during searching/browsing in English? 在用英语搜索或浏览过程中，你曾发现过什么问题或困难么？

How can websites and search engines help searching/browsing in English? 网站和搜索引擎如何帮助英语搜索或浏览？

What would you advise Chinese students who speak English but they have just started searching for information in English? (e.g. incoming first year university students) 你对刚开始搜索英语信息的讲英文的中国学生有什么建议吗？

Appendix 2: Diary Study Entry Format

Diary task template 日记任务模板

Instructions:指示：

1. This is your diary task template. You are encouraged to record your diary task description following this template form.这个模板是你的日记任务模板。你可以使用这个模板的格式来记录你的日记任务描述。

2. You do not have to use this exact format, but please make your diary-entries in a similar and consistent way.你不是必须使用与这个模板完全一致的模式，但是请使你的日记任务记录保持类似和一致。

3. Please provide as much detail as you could in your description to each circumstance. (We will provide a diary task example, your own diary task description should be similar in depth to the example, or at least as detailed as the example).在你的记录描述中，请为每种情况提供尽可能多的细节。（我们会提供一个日记任务的例子，你自己的日记任务记录描述应该与我们提供的例子在深度方面，至少在细节描述程度方面，保持类似。）

Diary task number (ordinal):日记任务编号（序数）：

Diary task date: 日记任务日期：

What task were you working on/what were you doing when the question came up?

(e.g. I was searching for literature for an university essay, Ergonomics lecture or I was searching for a song based on heard lyrics)你当时是在进行什么任务/当问题出现时，你当时在做什么？（比如，我在为大学课程的一篇文章，人体工程学课堂讲义，搜索参考文献，或者，我在搜索一首歌，基于我听到的歌词）

Where did you start the search? (e.g. Google, CNN, IMDB etc.) 你从哪里开始的搜索？（比如，Google, CNN, IMDB 等。）

What do you want to know? What kind of information do you want to find? 你想知道什么？你想找到何种信息？

What search expression had you planned to use? 你本来计划使用什么搜索表达？

Did you accept any recommendation that the search engine offered during the search? 在搜索过程中，你接受了任何搜索引擎提供的推荐吗？

What expression did you search with in the end? 最终，你使用的是什么搜索表达进行搜索的？

Please paste a screenshot of the first result page and copy its URL here.

请将第一页结果页面截屏并粘贴至此，并且请复制并在此粘贴它的 URL。

What result did you choose? Why? 你选择了什么结果？为什么？

Copy the URL. 复制并粘贴结果的 URL。

Paste the screenshot. 截屏并粘贴截屏。

Did you find what you were looking for? Was the opened website useful? 你找到你想要找的了吗？你打开的网页有帮助吗？

Did you see other result(s) in order to check the information? 为了检查信息，你看了其他结果吗？

If yes, which result did you look? Why? 如果是，你看了哪个结果？为什么？

Copy the URL. 复制并粘贴结果的 URL。

Paste the screenshot. 截屏并粘贴截屏。

Compare it with the previous result(s). 把这个结果与之前的结果进行比较。

Write your every previous experience regarding the used websites, if there is any.

如果有，请写下你所有之前的关于使用过的网页的经历。

Did you find what you were looking for in the end? 你最终找到你要找的了吗？

Were there any difficulties during the search? 在搜索过程中，有没有任何困难？

What kind of feelings did you have during the search, at all? 在搜索过程中，你有任何的感觉、感情或想法吗？

Please upload the completed diary task entry to your GoogleDrive folder (the one you were assigned to), with the name of file: diary task number (e.g. 1, 2, 3, ...), and format: .doc or docx (1.doc, 2.doc, 3.doc, ...).请将完整的日记任务记录上传至你的 GoogleDrive 目录（分配给你的目录），使用如下文件名命名：日记任务编号（比如 1，2，3，……）和文件格式：.do 或者 docx (比如 1.doc, 2.doc, 3.doc, ...)。

Appendix 3: Diary Study Interview Guide

Interview Guide 采访稿脚本

(0. Switch on the audio recorder) 打开录音设备

Analyzing searches 分析搜索

Please talk about this search. 请谈一谈这次搜索。(Opening the first documented search) （打开第一次的搜索记录文件）

What happened? 当时发生了什么？

What did you search for? 你当时在搜索什么？

Were there any difficulties during the search? 在搜索过程中，你遇到了任何困难么？

Did you find what you were looking for? 你找到你想要找的了吗？

Foreign language search habits 外语搜索习惯

What kinds of topics do you search for in English? 你用英语搜索什么主题？

Where do you search for it? 你从什么地方搜索这个主题？

Do you accept Google recommendations? 你接受 Google 的推荐、建议么？

What are your experiences regarding Google recommendations? 你对 Google 的推荐、建议有什么经历？(Quantity and quality of results, differences in search process compared to the native searches 结果的数量和质量，相对于母语搜索在搜索过程中有何不同)

Do you have any problems/difficulties during foreign languages searches? 在外语搜索过程中，你有任何困难或者问题么？

Do you have any experiences, observations regarding foreign language searches? 对外语搜索，您有任何经历，观察，体验么？

Do you use the Google translate? 您使用 Google Translate 么？ If yes, How? 如果是，如何使用？

Appendix 4: TranSearch Evaluation Tasks

Please finish these searching tasks. Give your answers and copy/paste the URL (URLs) of the website (websites) from which you find your answers:

Use the search engine you usually use for the first two tasks:

1. One of your friends spends his/her holiday in New York City, USA with his/her spouse and children ages six and eight. He/she asks you to make suggestions on a Chinese restaurant and an American restaurant. Please also list some reasons you recommend them.

你的一个朋友和他/她的爱人以及一个六岁和一个八岁的孩子在美国纽约度假。他/她请求你提供建议关于一个中餐馆和一个美国餐馆。 请你提出建议并且列出你推荐这些餐馆的原因。

2. A friend of yours is a fan of an American movie star, Tom Cruise. He/she wants some personal details about Cruise, such as his full name, birthday, height, 3 films he has starred in.

你的一个朋友是一个美国电影明星汤姆克鲁斯的爱好者。他/她想请你帮忙找一些关于汤姆克鲁斯的个人资料、简介，如全名，生日，身高，以及 3 部他演过的电影。

Use TranSearch for the following tasks:

3. What is "Glucosamine" and what does it do?

什么是"Glucosamine"以及它的用途？

4. The usage of the most common prepositions of location: in, on, at.

英语中表示地点的介词（in, on, at）的使用方法。

5. Search for an academic topic in your major. Please use a one-paragraph abstract to describe your search results/findings.

请搜索关于你的学术专业方面的一个主题，并使用一段摘要来描述你的搜索结果/发现。

Appendix 5: TranSearch Evaluation Interview Guide

Interview guide 采访稿

(0. Switch on the audio recorder) 打开录音设备

Analyzing searches 分析搜索

Please talk about this search. 请谈一谈这次搜索。(Opening the first documented search) （打开第一次的搜索记录文件）

What happened? 当时发生了什么？What did you search for? 你当时在搜索什么？

How did you select your search terms? 你如何形成搜索关键字的？ (If they do not describe the reasons for using those terms, also ask) Why did you use those words? 你为什么使用这些搜索词汇？

Were there any difficulties during the search? 在搜索过程中，你遇到了任何困难么？ Did you find what you were looking for? 你找到你想要找的了吗？

Did you change your search term and search again? 你是否更改搜索关键字重新搜索? If yes, how did you change your search terms and why? 如果有，你是如何更改的，为什么？

TranSearch Evaluation 评价 TranSearch

What do you think about the user interface, TranSearch? 你觉得 TranSearch 这个用户界面怎么样？(If they need prompts: What did you like about it? What did you dislike about it? Why? 如果需要提示：你喜欢什么？不喜欢什么？为什么？)

What is the difference do you think between the traditional search user interface and TranSearch? 你觉得 TranSearch 和传统搜索引擎用户界面有什么不同？

Does TranSearch help? Why and why not? TranSearch 有帮助吗？为什么？

Would you use TranSearch again? Why or why not? 你以后还会使用 TranSearch 吗？为什么？

(If they do not describe when/or for what types of searches they would use TranSearch for, also ask 如果没有明确表示何时或何种类型的搜索会使用 TranSearch，进一步提问) When would you use TranSearch and when would you use a traditional search interface? 何时你会使用 TranSearch，何时你会使用传统搜索界面？

What improvements would you make to the TranSearch tool? 你会对 TranSearch 工具做何种改进？

Bibliography

Ackerman, M.S. (2000). The intellectual challenge of CSCW: The gap between social requirements and technical feasibility. *Human-Computer Interaction*, 15(2-3), 179-204.

Ahmad, R., Komlodi, A., Wang, J. and Hercegfi, K. (2010). The impact of user experience levels on web credibility judgments. In *Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47* (Pittsburgh, Pennsylvania, 2010). American Society for Information Science.

Airio, E. (2008). Who benefits from CLIR in web retrieval? *Journal of Documentation, 64(5),* 760–778.

Amato, G., Cigarrán, J., Gonzalo, J., Peters, C., and Savino, P. (2007). MultiMatch – Multilingual/Multimedia Access to Cultural Heritage. *Research and Advanced Technology for Digital Libraries.* pp. 505–508.

Aula, A., and Kellar, M. (2009). Multilingual search strategies. *Proc. of the 27th international conference extended abstracts on Human factors in computing systems, CHI 2009,* pp. 3854-3870.

Belkin, N.J., Oddy, R.N., and Brooks, H.M. (1982). Ask for information retrieval:

Part I. Background and theory. *Journal of Documentation*, 38(2), 61-71.

Berendt, B. (2007). Context, (e)Learning, and Knowledge Discovery for Web User Modelling: Common Research Themes and Challenges. *Proc. from the Data Mining for User Modelling Workshop (DM.UM'07,)* UM, Corfu, June 2007.

Berendt, B., and Kralisch, A. (2009). A user-centric approach to identifying best deployment strategies for language tools: The impact of content and access language on Web user behaviour and attitudes. *Information Retrieval,* 12, 380-399.

Bian, G.W., and Chen, H.H. (2000). Cross-language information access to multilingual collections on the Internet. *Journal of the American Society for Information Science and Technology, 51(3),* 281–296.

Braschler, M. (2004). Combination Approaches for Multilingual Text Retrieval. *Information Retrieval*, 7, 1-2, 183–204.

Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, 36(2),3-10.

Bron, M., van Gorp, J., Nack, F., Baltussen, L.B., and de Rijke, M. (2013). Aggregated Search Interface Preferences in Multi-session Search Tasks. *Proc. of 36th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, 123–132.

Bush,V. (1945). As We May Think. *The Atlantic Monthly*, 176(1):101-108, July

1945.

Cao, G., Gao, J., Nie, J.-Y., Bai, J.(2007). Extending query translation to cross-language query expansion with markov chain models. *Proceedings of the sixteenth ACM conf. on information and knowledge management.* pp. 351–360.

Capstick, J., Diagne, A.K., Erbach, G., Uszkoreit, H., Leisenberg, A., and Leisenberg, M. (2000). System for supporting cross-lingual information retrieval. *Information Processing and Management, 36(2),* 275–289.

Case, D.O. (2007). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. Boston, MA: Elsevier/Academic Press.

Chizari, S. (2016). Exploring the Role of Culture in Online Searching Behavior from Cultural Cognition Perspective. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval (CHIIR '16)*. ACM, New York, NY, USA, 349-351. DOI: https://doi.org/10.1145/2854946.2854954

Chu, P., Jozsa, E., Komlodi, A., and Hercegfi, K. (2012). An Exploratory Study on Search Behavior in Different Languages. *Proc. 4th Information Interaction in Context Symposium (IIiX 2012),* Nijmegen, the Netherlands, August 21-24, 2012. pp. 318-321.

Chu, P., Komlodi, A., and Rozsa, G. (2015). Online Search in English as a Non-native Language. *Proc. Association for Information Science and Technology (ASIST 2015),* Missouri, USA, November 2015.

Chu, P., and Komlodi, A. (2017). TranSearch: A Multilingual Search User Interface Accommodating User Interaction and Preference. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI'2017)*, Denver, CO, USA, May 2017.

Clough, P., Al-Maskari, A., and Darwish, K. (2007). Providing Multilingual Access to FLICKR for Arabic Users. *Evaluation of Multilingual and Multi-modal Information Retrieval*, 205–216.

Clough, P., and Eleta, I. (2010). Investigating Language Skills and Field of Knowledge on Multilingual Information Access in Digital Libraries. *International Journal of Digital Library Systems*, 1(1), 89-103.

Corbin, J., and Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications.

Crystal, D. (2003). *English as a Global Language (2nd ed. First ed., 1997)*. Cambridge University Press. Cambridge, UK.

Dourish, P. (2004). What we talk about when we talk about context. *Personal Ubiquitous Computing*, 8(1), 19-30.

Exploredia. (2012). How Many People in the World Speak English 2013? (Retrieved December 31, 2014, from: http://exploredia.com/how-many-people-in-the-world-speak-english-2013/).

Gao, W., Niu, C., Nie, J.Y., Zhou, M., Hu, J., Wong, K.F., and Hon, H.W. (2007)

Cross-lingual query suggestion using query logs of different languages. *Proc. of 30th international conference on Research and development in information retrieval (SIGIR),* 463–470.

Ghorab, M.R., Zhou, D., O'Connor, A., Wade, V. (2013). Personalised Information Retrieval: survey and classification. *User Model. User-Adapt. Interact.* 23, 381–443.

Graddol, D. (1997). *The Future of English.* The British Council.

Grefenstette, G. (1998). *Cross-Language Information Retrieval.* Springer.

Hearst, M. (2009). *Search User Interfaces.* Cambridge University Press.

International Telecommunications Union/United Nations. (2012). Information and Communication Technology Data and Statistics. (Retrieved April 2, 2012, from: http://www.itu.int/ITU‐D/ict/statistics/).

Internet World Statistics. (2017). Top Ten Languages Used in the Web - March 31, 2017 (Number of Internet Users by Language). (Retrieved May 1, 2017, from: http://www.internetworldstats.com/stats7.htm).

Jansen, B., Booth, D., and Spink, A. (2009). Patterns of query reformulation during Web searching. *J. Am. Soc. Inf. Sci. Technol.* 60, 7 2009, 1358-1371.

Jean, S. (2006). *Boosting linguistic rule-based MT system with corpus-based approaches.* In Presentation. GALE PI Meeting, Boston, MA.

Jozsa, E., Koles, M., Komlodi, A., Hercegfi, K., and Chu, P. (2012). Evaluation of search quality differences and the impact of personality styles in native and foreign language searching tasks. *Proc. 4th Information Interaction in Context Symposium (IIiX 2012),* Nijmegen, the Netherlands, August 21-24, 2012. pp. 310-313.

Komlodi, A., Jozsa, E., Koles, M. and Hercegfi, K. (2011). Search Quality Differences in Native and Foreign Language Searching. In *Proceedings of the Human-computer Information Retrieval* (Mountain View, CA, 2011).

Kralisch, A., and Mandl, T. (2006). Barriers to information access across languages on the internet: Network and language effects. *Proceedings from the 39th Annual Hawaii International Conference on System Sciences,* IEEE Computer.

Mack, N., Woodsong, C., MacQueen, K.M., Guest, G., & Namey, E. (2005). *Qualitative Research Methods: A Data Collector's Field Guide.* FLI.

Magdy, W., Jones, G.J.F. (2011). An efficient method for using machine translation technologies in cross-language patent search. *Proceedings of the 20th ACM int. conf. on Information and knowledge management.* pp. 1925–1928.

Manning, C.D., Raghavan, P., and Schütze, H. (2008) *Introduction to Information Retrieval*. Cambridge University Press.

Marlow, J., Clough, P., Recuero, J., and Artiles, J. (2008). Exploring the Effects of Language Skills on Multilingual Web Search. *Proc. from the IR research, 30th*

European conference on Advances in information retrieval (ECIR'08), 126-137.

Merriam, S. (2009). *Qualitative research: A guide to design and implementation.* John Wiley & Sons.

Nie, J.Y. (2010). *Cross-Language Information Retrieval: Synthesis Lectures on Human Language Technologies.* Morgan and Claypool.

Oard, D.W., He, D., and Wang, J. (2008). User-assisted query translation for interactive cross-language information retrieval. *Inf. Proc. & Manage. 44, 1, 181–211.*

Okpor, M. D. (2014). Machine translation approaches: Issues and challenges. *IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 5, No 2, September, 2014. ISSN: 1694-0814.*

Peters, C. (2001). *Cross-Language Information Retrieval and Evaluation: Workshop of Cross-Language Evaluation Forum, CLEF 2000, Lisbon, Portugal, September 21-22, 2000, Revised Papers.* Springer.

Peters, C., Clough, P., Gonzalo, J., Jones, G. J. F., Kluck, M., and Magnini, B. (Eds.). (2005). *Multilingual Information Access for Text, Speech and Images*. Springer.

Peters, C., Braschler, M., and Clough, P. (2012). *Multilingual Information Retrieval: From Research To Practice.* Springer Science & Business Media.

Petrelli, D., Beaulieu, M., Sanderson, M., Demetriou, G., Herring, P., and Hansen, P.

(2004). Observing users, designing clarity: A case study on the user-centered design of a cross-language information retrieval system. *Journal of the American Society for Information Science and Technology, 55(10), 923–934.*

Petrelli, D., Levin, S., Beaulieu, M., and Sanderson, M. (2006). Which user interaction for cross-language information retrieval? *Design issues and reflections. J. Assoc. Inf. Sci. Technol., 57, 5, 709–722.*

Reinecke, K. and Gajos, K.Z. (2014). Quantifying Visual Preferences Around the World. *Proc. of SIGCHI Conf. on Human Factors in Computing Syst.*, 11–20.

Rieh, S.Y., and Xie, H. (2006). Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Inf. Process. Manage.* 42, 3 2006, 751-768.

Rozsa, G., Komlodi, A., and Chu, P. (2015). Online Searching in English as a Foreign Language. *Proc. 1st International Workshop on Multilingual Web Access (MWA 2015);* Florence, Italy, May 19, 2015.

Salton, G. (1989). *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Reading, MA: Addison-Wesley Pub.

Shneiderman, B., Byrd, D., and Croft, W.B. (1997). *Clarifying search: A user-interface framework for text searches.* D-Lib Magazine.

Shneiderman, B., Byrd, D., and Croft, W.B. (1998). Sorting out searching: A user-interface framework for text searches. *Communications of the ACM*, 41(4), 95-

98.

Steichen, B., Ashman, H., Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Inf. Process. Manag.* 48, 698–724.

Steichen, B., Ghorab, M.R., O'Connor, A., Lawless, S., and Wade, V. (2014). Towards Personalized Multilingual Information Access - Exploring the Browsing and Search Behavior of Multilingual Users. *Proc. from Conference on User Modeling, Adaptation and Personalization (UMAP 2014)*, Aalborg, Denmark, July 7-11, 435-446. DOI: 10.1007/978-3-319-08786-3_39.

Steichen, B., and Freund, L. (2015). Supporting the Modern Polyglot - A Comparison of Multilingual Search Interfaces. *Proc. of the ACM CHI Conference on Human Factors in Computing Systems (CHI 2015)*, pp. 3483-3492, Seoul, Korea, April 2015.

Sutcliffe, A. G., and Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10, 321-351.

Two lingual Google Search. (2008). Retrieved January 08, 2017, from http://www.2lingual.com/

Wilson, T.D. (1999). Models in information behaviour research, *Journal of Documentation*, Vol. 55, No. 3 (June), pp. 249–270.

World Wide Web Technology Surveys. (2017). Usage of content languages for

websites. (Retrieved May 1, 2017 from: http://w3techs.com/technologies/overview/content_language/all).

Zhang, J., Sun, L., and Min, J. (2005). Using the Web corpus to translate the queries in cross-lingual information retrieval. *Proceedings of 2005 IEEE International Conference on Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-KE '05.*