Theses and Dissertations

December 2016

The Information-seeking Strategies of Humanities Scholars Using Resources in Languages Other Than English

Carol Sabbar
University of Wisconsin-Milwaukee

THE INFORMATION-SEEKING STRATEGIES OF HUMANITIES SCHOLARS

USING RESOURCES IN LANGUAGES OTHER THAN ENGLISH

by

Carol Sabbar

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Information Studies

at

The University of Wisconsin-Milwaukee

December 2016

ABSTRACT

THE INFORMATION-SEEKING STRATEGIES OF HUMANITIES SCHOLARS USING RESOURCES IN LANGUAGES OTHER THAN ENGLISH

by

Carol Sabbar

The University of Wisconsin-Milwaukee, 2016
Under the Supervision of Dr. Iris Xie

This dissertation explores the information-seeking strategies used by scholars in the humanities who rely on resources in languages other than English. It investigates not only the strategies they choose but also the shifts that they make among strategies and the role that language, culture, and geography play in the information-seeking context. The study used purposive sampling to engage 40 human subjects, all of whom are post-doctoral humanities scholars based in the United States who conduct research in a variety of languages. Data were collected through semi-structured interviews and research diaries in order to answer three research questions: *What information-seeking strategies are used by scholars conducting research in languages other than English? What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?* And *In what ways do language, culture, and geography play a role in the information-seeking context, especially in the problematic situations?* The data were then analyzed using grounded theory and the constant comparative method. A new conceptual model – the information triangle – was used and is presented in this dissertation to categorize and visually map the strategies and shifts. Based on data collected, thirty distinct strategies were identified and divided into four categories: formal system, informal resource, interactive human, and hybrid strategies. Three types of shifts were considered:

planned, opportunistic, and alternative. Finally, factors related to language, culture, and geography were identified and analyzed according to their roles in the information-seeking context. This study is the first of its kind to combine the study of information-seeking behaviors with the factors of language, culture, and geography, and as such, it presents numerous methodological and practical implications along with many opportunities for future research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My sincere thanks to my advisor, Dr. Iris Xie, whose excellent advice and encouragement from my very first days at SOIS made this much easier than it otherwise might have been. My gratitude to my dissertation committee: Dr. Lynn Loewen, Dr. Ann Greer, Dr. Maria Haigh, and Dr. Johannes Britz, for their support and feedback through the dissertation process. To my classmates throughout my coursework, who were usually much younger than me but still accepted me and inspired me to keep up with them. A special note of gratitude to Dr. Melodie Fox whose advice helped me survive and succeed the prelim exam process.

I express my appreciation to the 40 participants of this study; talking to each of you was the most rewarding part of this project, and it would take dozens of dissertations to express everything I learned from all of you. I thank my transcriptionists – especially Megan, Shay, and Sam – for pounding out pages of typed interviews; I could not have done this without you. I also owe my success to my colleagues at Carthage College, especially the people in Library and Information Services who supported and encouraged me even when I had to miss work to attend class. Carthage is an amazing community, and I'm so very privileged to be part of it.

Finally and foremost, I thank my family. My father, Harold Hintzsche, helped me see that everything in the world is interesting if you take time to look at it. My mother, Alice Hintzsche, taught me what can be accomplished when a person is always "doing something constructive." To my sons, Ken and Jeff, because they are truly good young men, and I want them to learn the things I learned from my parents. And to the love of my life – my husband Mustapha – thank you for always believing in me. I did this because you believed I could and because you wouldn't let me give up. I want you to be proud; this is our accomplishment together.

x

Chapter 1 – Introduction

1.1. Background

How scholars conduct research is a complex topic, and many models have been built to represent, de-construct and understand their information-seeking strategies. But what if the resources they require to accomplish their research are not in English? An Asian historian may rely upon primary sources that exist only in Chinese, Japanese, or Punjabi to provide the information they need for their topic. Language scholars studying 12th century German texts, religious heretics from the south of France, nurses in South Africa, or Iranian dance require rare and specialized documents from the target countries and in the target languages as sources.

Past research recognizes that information seekers acquire knowledge from sources such as information systems, informal channels, and human colleagues (Ingwerson and Jarvelin, 2005). But what research strategies do scholars who rely on non-English sources specifically use, and which do they find most successful? For example, what strategies are most effective in researching Norwegian television, Latin American suffrage, or post-colonial Vietnam? What barriers and challenges do non-English researchers encounter in identifying and obtaining their sources? What if the needed documents exist only in another language and reside in archives on a distant continent (Madkour, 1975)?

The work of scholars, especially in the humanities, often requires exhaustive and deep investigation of many types of sources. Historians, for example, scan the environment for stimuli (books, periodicals, letters, diaries, newspapers, etc.) that are relevant to their research, then absorb and interpret those to create new knowledge. To do so, they must be able to identify, obtain, analyze, and interpret a vast body of literature related to their topic of study (Case, 1991). The prolific research output of such scholars is well known, leading us to conclude that any

potential barriers that result from language, culture or geography in the information seeking context must not be insurmountable. This dissertation explores the information-seeking strategies used by scholars who rely on non-English sources, the shifts made between strategies, and the roles of language, culture, and geography in the information seeking context.

1.2. Research problem and research questions

Existing research on information-seeking strategies among users in general or scholars specifically has not explored the roles that language, culture, or geography play in in the information seeking context or the shifts that scholars make between strategies as a result of these potential barriers. Most information retrieval studies have focused either on the characteristics of the user or the characteristics of the task. This study looks primarily at the characteristics of the sources and the topics. The scholars being studied have largely the same characteristics as other humanities scholars, having advanced degrees and similar research experience and ambitions. Their tasks are also much the same as other scholars – to find documents related to their topic of studies. What differentiates this group of scholars, and hence this study, from others is that their sources are in languages other than English. Studies related to language and information retrieval have not explored this research problem and have been limited to system-oriented studies, usually focusing on problems related to stemming or term matching in formal search tools. User studies related to language in the information-seeking context are virtually non-existent, representing a gap in the literature. This study specifically seeks to explore that gap, studying how scholars who rely on resources in other languages and originating in other countries choose the information-seeking strategies they use and when and why the make shifts from one strategy to another. The specific research questions are:

Q1: What information-seeking strategies are used by scholars conducting research in languages other than English?

Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?

Q3: In what ways do language, culture, and geography play a role in the information-seeking context, in particular, how do they create barriers that cause problematic situations?

1.3. Definitions

Because this dissertation combines the fields of information retrieval and language, there are two distinct vocabularies that come into play. Rarely do scholars from the field of IR have knowledge of linguistic terms, and language scholars are generally unaware of terms related to IR. In this dissertation, I try to be sensitive to both groups since they both comprise the audience of this study, and definitions of many terms and concepts are included throughout the dissertation. The most basic and significant terms used in this dissertation must be defined and understood:

- *Information seeking* is defined by Ingwerson and Jarvelin (2005, p. 259) as "the acquisition of information from knowledge sources, for instance, from colleagues, through (in)formal channels, and from an information system." This definition best fits this study in that it includes not only formal systems but informal channels and human entities as well. Further discussion of information-seeking, especially in contrast to information retrieval and information behavior, is provided in section 2.2.

- *Information-seeking strategies* are defined herein as activities undertaken by a researcher to advance an information-seeking task or meet an information-seeking need. Further explanation of information-seeking strategies is given in section 3.2.

- *Shifts* are defined as changes made by users from one information-seeking strategy to another in order to advance the information-seeking process. Further discussion of the concept of shifts can be found in section 3.4.

- *Barriers* are as any factor or element of a context that impedes research and therefore causes a problematic situation. Barriers are explained in section 3.5.

- *Language* as used in this dissertation includes the verbal and written representations of human communication. Because most of the research conducted by scholars utilizes written sources, much of the discussion here involves the more mechanical aspects of language such as alphabets and character sets, morphology, and language reforms. These concepts and terms are discussed in sections 2.3.1 and 6.2.

- *Culture* is defined by McGarry (1993, p. 59) as "the way of life of particular human groups as all the activities that this way of life entails." Cultural factors discussed by participants in the study include authoring and publishing, library and archival culture, political systems and climate, and societal norms such as bureaucracy and politeness. These are further discussed in sections 2.3.2 and 6.3.

- *Geography*, in the context of this dissertation, is concerned with the location of the needed materials, especially in relationship to the location of the scholar. Factors related to geography and its role in the information-seeking context are discussed in sections 2.3.2 and 6.4.

1.4. Research design and methodology

This dissertation explores the research questions by using qualitative methodology to explore and describe the information-seeking behaviors of scholars, which I define as academics having terminal, doctoral degrees in their discipline. The subjects are college and university

4

professors based within the United States. The great majority speak English as their native language; for others, their native language is the language in which their research is performed. Because of the nature of their disciplines, scholars were sought in the areas of history, languages, and other humanities. Humanities scholars are more likely than scholars in hard sciences, social sciences, or education to focus on a specific language, culture, or location, and then focus on sources in the related language. A total of 40 subjects participated in the study, which exceeds the number commonly thought to achieve saturation (Glaser, 1967; Seale, 1999). The group include scholars that represent 29 different languages and a wide variety of research interests. A table of the subjects is given in Table 4.1 in Section 4.3.

The subjects were chosen using purposeful sampling based on their active pursuit of research using non-English resources during the period of the study. Moreover, subjects were purposefully sought to represent a variety of languages with varying features. The time of the study for each subject varied from two weeks to four months, beginning in July 2015 and concluding in January 2016. This allowed for the subjects to conduct their research at their point of need, and it allowed time for both gathering and analysis of the data. Qualitative methods used for data collection included semi-structured interviews, research diaries, and questionnaires; these methods allowed for gathering of in-depth insight into the information-seeking behaviors and the information-seeking context in which these scholars conduct their research.

The collected data were analyzed using grounded theory and the constant comparative method (Glaser, 1965; Glaser, 1967). Data were analyzed using the three stages of open, axial and selective coding. During the open coding stage, taxonomies of information-seeking strategies, shifts, and the roles of language, culture and geography were built inductively from the data gathered from the subjects. During axial coding, the categories were refined to identify

the central phenomenon and the causal factors that impact it. Finally, at the selective coding

stage, theory emerged and a model was built to represent language, culture, geography,

strategies, and shifts in the information-seeking context. The specific methodology used is

described in Chapter 4 and the discussion of methodological implications, including the model,

are discussed in Chapter 7.

Qualitative methodology is especially well suited to an exploratory study such as this,

where the primary objective is to understand and document the information-seeking process

undertaken by scholars who use non-English resources. Inherent in each narrative is crucial

information about the scholars' customary practices, the experiences that have helped them

develop effective strategies, and best practices for humanities research in a field and context that

can present barriers at any time related to either language, culture, or geography. Further

justification of qualitative methodology more broadly and specific methods is provided in

Chapter 4.

1.5. Significance of the study

Some past studies have been conducted on how scholars in the humanities, especially

historians, conduct research in their fields. As on-line resources became more prevalent

throughout the 1990s, libraries, information scientists and research methodologists took an

interest in whether scholars would embrace these new technologies and whether newer search-

based methods would gradually replace more traditional methods. Information scientists have

studied research strategies used in the context of specific work and search tasks (Bates, 1989;

Xie, 2002, 2008) and have delved into shifts in research tactics within a search task (Xie, 2000;

Yuan and Belkin, 2010; Pharo, 2004). Others have identified factors used in determining the

selection of search tactics at various stages throughout a search task or analyzed the factors that

influence information-seeking in relation to situations and transitions (Xie and Joo, 2011; Pharo, 2004).

In the area of non-English information-seeking, some attention has been paid to the problems presented to IR systems by languages with complex linguistic structures (Bar-Ilan & Gutman, 2005; Hammarstrom, 2007; Lazarinis, Vilares, Tait & Efthimiadis, 2009). However, only a few user studies, including that of Meho and Tibbo (2003), have looked at scholars whose research requires the use of materials in other languages. These language studies did not analyze strategies and shifts, and studies that do focus on strategies and shifts have not taken into account language as part of the information-seeking context.

This dissertation has theoretical significance in its presentation of a new conceptual model for the categorization and visualization of information-seeking strategies and shifts. The information triangle was developed specifically to study these research questions, but it is a model that is highly adaptable to other types of studies of users' information-seeking behaviors. It is ideally suited to complement qualitative methodologies but could also be used in quantitative studies. Unlike previous models, the information triangle is designed to categorize the information-seeking strategies used by scholars (or any group of subjects) and to visualize the shifts that they make in various situations – especially problematic situations caused by encountering barriers.

The results of the study also present practical significance. The intended audience of this dissertation includes two major groups: scholars and information professionals. By learning from the information behaviors of their peers, humanities scholars can gain insight into common research contexts, best practices, effective strategies, and shifts that advance the information process. Likewise, information professionals have much to gain from recognizing the needs and

habits of this group of scholars. With a clearer understanding of the strategies most often used

by non-English researchers, the barriers they most often encounter and the shifts they make

among strategies, system designers, publishers, providers of information repositories, librarians

and archivists can provide improved tools, resources, and assistance. The results of this study

endeavor to provide valuable insight into the information-seeking behavior of these users toward

the end of providing an environment that will better contribute to their success.

Finally, the study also has methodological significance in the creation of a new grounded

theory research model. A model for language, culture, geography, strategies, and shifts in the

information-seeking context was developed using grounded theory and is presented in Chapter 7.

The model portrays how the various elements of language, culture, and geography impact the

information-seeking context and what aspects of the information-seeking process they most

directly impact and their roles in the context, especially as they create barriers. The model

emphasizes the impact of these barriers on choices of information-seeking strategies, including

the shifts made between strategies. In addition to the model, this study serves as a confirmation

and example of the effective use of qualitative methods – especially semi-structured interviews

and research diaries – in the field of information studies.

1.6. Dissertation structure

This dissertation is presented in eight chapters. Chapter 2 provides a comprehensive and

in depth literature review of concepts and previous studies to provide a foundational review of

the features of language related to the research questions as well as definitions related to

information studies. Chapter 3 presents the new conceptual model that was developed

specifically to analyze the results of the study. This conceptual model – the information triangle

– is a simple yet effective way to classify, categorize and explore the information-seeking

strategies used by scholars as well as the shifts that scholars make from one strategy to another. Chapter 3 also includes an in-depth literature review of prior studies and theory related to information-seeking strategies and shifts. Chapter 4 explains the research methodology used to undertake the dissertation study and the reasons for the choice of method. Chapter 5 presents the results and findings related to research questions 1 and 2 – strategies used by scholars and the shifts they make between strategies. This includes figures portraying information-seeking threads cited by scholars and mapped to the information triangle. Chapter 6 presents results related to research question 3 – the roles that language, culture, and geography play in the information-seeking process; an emphasis is made on barriers caused by these factors and the problematic situations they create. Chapters 7 provides a discussion of the findings and the implications: theoretical, practical, and methodological, including the new grounded theory model. Also included in Chapter 7 are the limitations of the current study. Chapter 8 discusses the key findings of the research and opportunities for future research.

1.7. Introduction summary

This introductory chapter of my dissertation shows that language, culture, and geography are factors that have been largely neglected in the study of scholarly information-seeking. It explains how qualitative methods were used to study the research behaviors of humanities scholars and how my new conceptual model was used to visualize the shifts that these scholars make between information-seeking strategies. Specifically, this dissertation explores the information-seeking strategies used by scholars who rely on sources in languages other than English, the shifts that they make between strategies, and the role of language, culture, and geography in the information-seeking context, especially in problematic situations.

Chapter 2 – Literature Review

2.1. Overview

This chapter presents a literature review of previous scholarship related to my research questions. Because the population being studied consists of humanities scholars who use resources that are not in English, we must understand the various aspects of language, culture, and geography that impact their research and the roles that they play in the information-seeking context. Studies of the effect of language on information-seeking strategies have been so rarely undertaken, previous literature must be investigated in two separate categories: 1) scholarship related to information-seeking strategies, including when and why shifts occur, and 2) scholarship that concerns the aspects of language, culture, and geography that might play a role in information seeking by scholars. Where scholarship in these two areas has overlapped is in the investigation of how the characteristics of various languages present challenges to IR systems and search algorithms; these are discussed in the section on language.

Much of the previous research related to the fundamental concepts and definitions of strategies, shifts, and barriers is presented in Chapter 3 in support of the discussion of my proposed conceptual framework – the information triangle. A discussion of previous models of information-seeking are also included in Chapter 3.

2.2. Information-seeking definitions

Information-seeking is a term that includes how users find, access, obtain, evaluate or otherwise interact with sources that support their research. Information-seeking, described by Ingwerson and Jarvelin (2005, p. 259) as "the acquisition of information from knowledge sources, for instance, from colleagues, through (in)formal channels, and from an information system," steps beyond the boundaries of formal information systems. Information-seeking is

purposive and can involve user interactions with both manual and computer-based information in order to achieve one's information goals (Xie, 2008; Xie, 2010). The study of information-seeking strategies belongs to a branch of information studies that focuses on user-oriented information retrieval. This more recent area of IR research considers the user as an integral component in the information-seeking process and system. Ruthven states that strong research in interactive information retrieval comes "not only from a technical knowledge of interactive systems development but also from a knowledge of people's search behavior and search context, including the environmental factors that influence behavior (Ruthven, 2008, p. 24; Fidel & Pejtersen, 2004)." Ruthven's statement also reveals how much of system-oriented research narrowly views information retrieval as all about searching: formulating a query, feeding the query into an IR system, and evaluating how well the system matches documents to the query. Not only does the user-oriented approach seek to include an understanding of the users' cognitive constructs during the information-seeking process, but it also allows us to break free from the focus on searching and include other types of information-seeking strategies (e.g. browsing) and how they contribute to the successful completion of information tasks. Along with this shift in emphasis from studying systems to studying users to studying the interactions between systems, users, and the information they seek has come a shift from quantitative research methods to qualitative and mixed methods (Xie, 2008).

Information-seeking can often be conflated with two related terms: Information retrieval and information behavior. ***Information retrieval*** (IR) – the narrowest of the three terms – is defined as finding desired information in an information store or database (Meadow, Boyce & Kraft, 1999; Large, Tedd & Hartley, 1999; Chowdhury, 2004; Wilson, 2000). IR is studied either from a system perspective or a user perspective and most often focuses on either the

formulation of queries used by users when interacting with a search tool or on the algorithms

used by the search tool to match documents to the user's query. ***Information behavior*** is the

broadest of these terms and includes all types of user-information interaction, including

information generation, sharing, management, as well as unintentional interactions with

information. Wilson (1999, 2000) is best known for his discussion of human information

behavior and defines it as "the totality of human behavior in relation to sources and channels of

information (Chowdhury & Chowdhury, p. 13)," or "those activities a person may engage in

when identifying his or her own needs for information, searching for such information in any

way, and using or transferring that information (Wilson, 1999, p. 249)," and he explains that it

can include interaction and communication with others and passive information reception from

sources such as television and radio, which the user may have no intention to act upon.

Marchionini (1995) draws the distinction between information retrieval and information-

seeking this way:

> "*Information-seeking* is a process in which humans purposefully engage to change
> their state of knowledge. The term *search* will be used to mean the behavioral
> manifestation of humans engaged in information-seeking and will also be used to
> describe the actions taken by computers to match and display information objects.
> The term 'information-seeking' is preferred to 'information retrieval' because it is
> more human-oriented and open-ended. Retrieval implies that the object must have
> been 'known' at some point; most often, whoever 'knew' it organized it for later
> 'knowing' by themselves or someone else. Seeking connotes the process of
> acquiring knowledge; it is more problem-oriented; the solution may or may not be
> found. For example, seeking spiritual enlightenment makes sense, but retrieving
> enlightenment does not. Retrieval is applicable to database management and most
> applied problems, but seeking is closer to question-answering or learning
> (Marchionini, 1995, p. 5)."

My dissertation focuses on information-seeking since the intent is to understand and document

the intentional behaviors of scholars looking for resources in languages other than English and

does not limit those behaviors to formal IR systems. Indeed, my research investigates the

relative efficacy of informal, manual search strategies as compared with formal systems and interaction with other humans.

The concept of information-seeking, therefore, falls between information retrieval and information behavior in its specificity, and includes the study of ***information-seeking strategies (ISSs)***. Because the concept of strategies is so key to the conceptual model proposed for this dissertation, they are defined and discussed in more detail in Chapter 3 – Conceptual Model. Likewise, literature that discusses the areas of shifts in information-seeking strategies and barriers to successful information-seeking is discussed in Chapter 3.

2.3. Language – Foundational aspects

While not widely researched by information science scholars from a user perspective, language, culture, and geography are all fairly widely recognized by scholarly researchers and library professionals in the field as "a barrier between the library and potential users (McGlamery & Ulrich, 2007, p. 1)." From a systems perspective, much has been written about the problems that various languages present to formal IR systems, and some modest exploration has been made into the area of cultural and geographical barriers to research.

A review of the literature supports the notion that scholars who rely on sources in languages other than English face three primary types of barriers in the information-seeking process: 1) linguistic (language) barriers related to identifying sources in other languages, including lack of knowledge of the language, the inabilities of formal search tools to deal with the linguistic aspects of some languages, and the lack of indexing of literature in the target language, 2) cultural barriers that may prevent scholars from gaining entrée into libraries, archives, and similar institutions whose collections hold the needed resources, and 3)

geographical barriers due to the needed resources being located far from the location where the scholar is conducting their research.

While language-related barriers are not the focus of any seminal or major work, a variety of individual studies and practical articles discuss the specific types of barriers from their own perspectives. As far back as 1975, Madkour recognized three major barriers that impede all documentation access worldwide: 1) the volume of documents to be stored, indexed, and disseminated, 2) geographical obstacles, and 3) language barriers. Some suggested solutions to language, cultural, and geographical barriers have been proposed and are now being promoted, including the concept of global English and cross-language information retrieval (CLIR) systems. While these concepts and movements may present some improvement for the global research community, they rarely, if ever, provide much assistance to scholars who conduct their research in languages other than English.

One of the areas in which previous research does exist in relative abundance is how language affects formal IR search tools. My research utilizes these studies to highlight the aspects of language that cause barriers and to create a context for the user perspective of language, culture, and geography barriers and the alternate strategies used by scholars to overcome or circumvent them.

2.3.1. Language barriers

To begin to understand language barriers, one must look at the aspects of languages that make them unique and therefore present challenges for both scholars and IR systems. Included among these aspects are:

- The graphical representation of language, including
    - The alphabets and/or character sets that they use

- o   The electronic representation of those characters within systems

- o   Orthographic variations within and between languages

- The morphological and syntactical properties of languages that affect how they can be parsed and stemmed by IR systems

- Discourse and lexical features related to the meaning of words and the general structures of documents of different types (Crystal, 2006)

- Dialect, nuance and lack of fluency by the researchers that impact relevance assessment and comprehension

Generally speaking, recent and current web search tools do not respond well to queries presented in languages other than English (Bar-Ilan & Gutman, 2005). Scholars who use languages other than English may often quickly shift from formal search-based strategies to strategies based on informal techniques such as browsing, citation tracing or to interacting with other humans. To understand these shifts, this dissertation looks at why language is such a significant barrier to formal IR systems.

IR systems match documents that they find in a database or on the web with the words in the user's query (Large, Tedd & Hartley, 1999). The documents must first be tokenized (broken into usable entities, usually words) and then indexed. Lazarinis, Vilares, Tait & Efthimiadis explain:

> "Search engines crawl the Web and fetch documents which are then indexed and included in their databases. Indexing of the fetched Web documents is a complex procedure which requires, among other specialized routings, identification of the language of the document, pre-processing of the texts, tokenization, stopword identification, stemming, and uniform handling of the morphological variances of the tokens (Lazarinis et al, 2009, p. 231)."

They continue to explain that well-established algorithms exist to automatically determine the language of a text. In web documents, the URL and the data in the HTML headers may indicate

15

the language or the country of origin. Recent tools such as Google Translate, that provide translation of web pages into a target language, have by necessity developed mechanisms to identify the original language of the page or document so that the translation can be performed. Some success has been found using N-gram frequencies to create specific fingerprints for individual languages (Hammarstrom, 2007). Despite these advances, several issues exist to the identification of document language. One of these is the distinction between dialects and closely related languages. While a URL domain name of .se can help distinguish Swedish from its .dk Danish counterpart, no such differentiation exists to aid in the identification of the several hundred regional languages and dialects of Indonesia (Lazarinis et al, 2009). The aspects of language discussed below impact not only the identification of the language but also how well it can be processed by formal IR tools such as web search engines, OPACs, and other library database tools.

2.3.1.1. Alphabets, character sets and orthography

When asking scholars about the barriers they experience in their research related to language, speakers of western languages – especially those that use a Latin alphabet – have to think harder to name very many. However, researchers who use languages less similar to English can quickly list several barriers, and they almost always begin with the written representation of the language. Since nearly all modern IR systems are based on keyword matching, barriers posed by the written forms of the language are formidable.

Several characteristics of written representation are significant. The most common writing systems for languages are: alphabetic, syllabic, and pictographic. The alphabetic forms are phonetically based where each grapheme represents a phoneme (Altmann & Fan, 2008). The most common of these are the Latin/Roman alphabet, Cyrillic, and Greek. Similar to alphabets

are abjads – consonant-only alphabets where vocalizations may be represented by diacriticals; examples of these are Arabic and Hebrew. Languages based on abjads are nearly always written from right to left, which is another factor that affects their ability to be interpreted properly by formal IR systems. A specifically problematic characteristic of Arabic – and other languages that are written using a cursive form – is that the alphabetic characters are connected together to form words and may have three different forms, depending on whether the character appears at the beginning, middle, or end of the word (Ahmed & Nürmberger, 2007).

Syllabic writing systems, including syllabic alphabets and syllabaries, use one character for each spoken syllable. In syllabic alphabets, the consonant is the main portion of each character, and the vowels can change using diacriticals. Syllables that consist entirely of a vowel sound have specific characters that represent them. Syllabic writing systems are common in central and eastern Asia and include such systems as Devanāgarī (the script used for Hindi), Thai, Hmong, Tibetan, and Balinese. By contrast, languages based on syllabaries have distinct symbols for each syllable where syllables with similar consonant sounds may have dissimilar graphemes.

Finally, pictographic or logographic writing systems utilize one or more graphemes for a single word. They may try to resemble the items that they represent, or they may be semanto-phonetic compounds, which means that part of the symbol gives a clue as to its meaning and part of the symbol provides a hint to its pronunciation (Ager, 2013). Modern examples of languages that use pictographic character sets include Chinese, Japanese, and Korean.

Since most IR systems were developed for alphabetic writing systems, they do not accommodate other systems well. Stated simply, "the problem with Chinese language retrieval is that Chinese characters – being pictographs rather than phonetic scripts – cannot be interpreted

17

as easily as English words by artificial intelligence (Chang, Morales-Arroyo & Spink, 2010, p. 309)." In the pre-processing of text documents, for example, the IR system must perform text segmentation, turning chunks of text into tokens (words or other entities) that will later be matched with search terms. Most systems rely on very rudimentary string processing logic that essentially says that any text between either a space or a punctuation mark is a word. However, various linguistic phenomena such as multiword expressions, contractions, enclitics attached to verbs, or variations in segmentation make this task difficult. For example, Germanic languages have a tendency to combine words together to create compound expressions that are not delineated by spaces but for which the individual pieces have significant meaning. In languages from Italian to Tagalog, clitics attach themselves to the main verb. Chinese uses no spaces between words at all. In nearly all of these cases, information systems developers have been able to improve the effectiveness of their systems by introducing new language-specific algorithms to perform segmentation (Lazarinis et al, 2009).

The difference in alphabets causes considerable difficulties for formal search tools, especially when the query entered and the desired web page(s) do not use the same script. Blanco and Lioma (2009) conducted tests of searches done to find the home pages of football teams from various countries and compared the success of those searches based on the language, the domain, and the script. When the queries were issued in the correct script for the language being searched, the desired pages were likely to come up in the first three entries. However, if a non-Latin language was transliterated and represented in Latin script, the desired pages were much less likely to be found. However, for especially difficult languages, such as languages that read right to left, transliteration into Latin script actually improved the chances of finding the desired pages, which may well indicate that search engines still need to improve their search and

matching capabilities in such languages. Queries performed in the incorrect domain for the search engine (such as Russian queries submitted to yahoo.es) were less likely to return the desired pages. However, in some Latin-script languages, the global domain of the search engine (e.g. google.com rather than google.fr) performed better, perhaps indicating that those search engine providers have spent more effort on their general algorithms than they have on domain-specific search algorithms (Blanco & Lioma, 2009, p. 344).

Since formal IR systems require the terms to be represented in their respective alphabets by machine code, a further complication is that of character sets. Even languages which use the standard Latin character set often utilize diacriticals which extend the character set beyond the original ASCII (American Standard Code for Information Interchange) representation which was based on a 7-bit system that could represent the 128 characters found in English. Approximately 180 8-bit standards have subsequently been developed to represent these extended character sets, making interoperability between standards and systems very difficult (Clavel, 2006). Even with the emergence of the Unicode 16-bit character set, which seeks to represent all of the characters of multiple alphabets (Greek, Latin, Cyrillic, Arabic, Hebrew, etc.) in one character set, and Unicode 32-bit which also includes character-based languages, there is often very little uniformity. Hence, while ASCII characters can be accommodated by a 1-byte (8-bit) character set, all other European languages and Middle Eastern languages require a 2-byte (16-bit) character set, and Asian languages require 3 or 4-byte (24- or 32-bit) character sets (Clavel, 2006).

Despite the advent and widespread adoption of Unicode, Clavel (2006) asserts that character sets are still an issue for the following three reasons:

- Special characters may not be indexed uniformly within scholarly and library databases

- Users may not have the necessary language support installed on their computers in order to enter the needed characters into a search engine or other search tool

- Records returned by the search process may not be represented in a way that is comprehensible to the user

Clavel (2006, p. 4) continues to explain that some special characters in languages such as Estonian, Hungarian, or Czech are "not reduced to their closest ordinary letter in the databases of the national libraries of the European New Member States." These variances in coding can and often do lead to lower recall rates when using these systems for information retrieval. In fact, the Second Workshop on Improving Non-English Web Searching (iNEWS) stated that "the main conclusion from the literature is that searching using non-English and non-Latin based queries results in lower retrieval success and requires additional user effort so as to achieve acceptable satisfaction levels. Furthermore, international search engines, like MSN Live, Google and Yahoo, are relatively weaker with monolingual non-English queries (Lazarinis et al, 2008, abstract)." As mentioned earlier, the identification of document language is one of the most fundamental barriers to non-English searching, and in the case of Asian languages, the existence of multiple encoding schemes can often be shown as the cause for low search effectiveness (Lazarinis et al, 2009).

As the IR systems available in English and some other western languages advance and allow for new features such as voice recognition in spoken queries, languages whose written representations are phonetically based will be at an advantage over character-based languages like Chinese since voice recognition is naturally done following phonetic patterns (Chang et al, 2010, p. 317). In the field of non-text document retrieval, encoding also causes havoc since file names of images, videos, etc., are encoded in Latin scripts even when the web site is in a non-

Latin language such as Russian or Greek, thus creating mismatched search result. "For example, the Polish query *pies* (dog) was falsely taken as the plural form of the English word *pie* and therefore no relevant canine images were retrieved. In addition, it was found that the absence of diacritics causes fewer relevant images to be retrieved (Lazarinis et al, 2009, p. 232)."

2.3.1.2. Morphology and syntax and their effects on stemming

Morphology can be explained as the grammar of words, and syntax is the grammar of expressions (Matthiessen, Teruya & Lam, 2010). Morphology is an important concept to IR systems which often utilize stemming to improve recall in information searching. Stemming is the process of reducing a word to its stem or root form; that is, we remove any affixes to derive the base morphological term (Fox & Fox, 2002). For example, the words *teacher, teachers, teaching, teachings,* and *teaches* all stem to the root *teach*. Before a stemming algorithm is applied, a user who types "foreign language teaching" into a search engine would not retrieve documents with the terms "teach foreign languages" or "foreign language teachers" in them. Obviously, the suffixes that are most often identified and removed in English (-s for plural, -ed for past tense, -er for comparative, -est for superlative, -ist and -er for professions, etc.) are not always those in other languages. Standard stemmers use pre-determined lists of word endings and other affixes, which means they are difficult to change, can run relatively slowly (Fox & Fox, 2002), and only work well for the language for which they were created. Therefore, different stemmers must be written for each language according to its own morphological patterns.

While English has a very simple morphology based on its lack of case or gender markers, some languages such as Czech have numerous endings for gender, number, and case for nouns and adjectives, resulting in 52 rules for removing these endings only in the "light" version of a

Czech stemmer (Dolamic & Savoy, 2009). Indonesian has approximately 35 standard affixes, which include not only suffixes but also prefixes, circumfixes, and some infixes, which can be iteratively combined; clearly, writing a stemming algorithm to meet these needs is a significant challenge (Lazarinis et al, 2009). Spanish has a highly complex inflectional morphology in which, for example, there are 3 types of regular verbs and over 30 types of irregular verbs with 118 inflected forms for each verb. Vilares, Alonso & Vilares (2008) recount how lemmatization – a process similar to stemming whereby words are reduced to their representative headwords (Lazarinis et al, 2008) – was used successfully to reduce conflation between Spanish verbs. A common issue with stemming is that a singular and plural may have completely different stems, as in Russian *chelovek* vs. *ludi* (man vs. men/people) (Bar-Ilan & Gutman, 2005) or in English *mouse* vs. *mice.* In languages with somewhat less predictable plural forms, such as *cheval* and *chevaux* (horse and horses) in French; removing the ending may result in a stem that is so short as to be ambiguous and create incorrect matches such as *cheveux* (hair.) Finally, many languages, including French, will prepend some definite articles to nouns; Hebrew can add not only articles but also most prepositions and some conjunctions to words as prefixes (Bar-Ilan & Gutman, 2005). Once these other particles are combined with the noun, they may be conflated with other words that have the same pattern.

In Arabic, the stemming and text-matching algorithms developed for European languages do not apply well because Arabic is morphologically more complex (Mustafa, 2004; Fattah, Ren, Kuroiwa, 2005). Some success has been achieved using hybrid N-grams rather than conventional stemming. N-grams are sequences of letters (characters) of N length derived from a word or phrase. For instance, the word *example* would divide into the following tri-grams: exa, xam, amp, mpl, ple or the following 4-grams: exam, xamp, ampl, mple. In many cases in

Arabic, word forms change not by adding suffixes or even prefixes but rather infixes (variations added in the middle of a word) and stem changes (Moukdad & Large, 2001; Mustafa, 2005). By breaking the words into N-grams of three or four characters in length (with the vowel diacriticals removed) and analyzing non-contiguous N-grams, variant forms of words with similar beginnings and endings can be matched. This technique was found to improve both recall and precision in searching (Mustafa, 2005).

When stemming and matching algorithms developed for English are used with other languages, the results are less than satisfactory. "If these tools are not suited to search in the specific language, the page might never be found. Thus pages in non-English languages have a much larger chance of 'being lost in Cyberspace…' If we want the web to remain a place for everybody, regardless of the languages they speak, an effort has to be made to provide these tools (Bar-Ilan & Gutman, 2005, p. 24)." Language- or nation-specific search engines may seem to have an advantage in dealing with their own language, but international tools such as Google or Yahoo have more manpower and a greater vested interest in order to handle additional languages, and their performance may actually surpass that of national search engines (Efthimiadis et al, 2007; Chau, Qin, Zhou, Tseng & Chen, 2008).

2.3.1.3. Lexical and discourse barriers

In addition to their graphical representation and their structure, words have semantic and lexical value that IR systems are not always able to distinguish. There are a variety of lexical issues that must be overcome by IR systems to allow for more effective searching. Large et al (1999) enumerated those that existed at the time:

- failure to recognize synonyms, more general, and more specific terms,

- failure to recognize and correct spelling errors,

- inability to navigate variants such as singulars and plurals (unless they stem easily) or variant spellings such as exist between British and North American English,

- the inability to disambiguate words with multiple meanings, including words that exist in multiple languages but have disparate meanings,

- problems that arise from sentence constructions that can have multiple meanings

For example, a person might search for information on "Italian sports cars." Unless they are programmed to do so (usually using authorities tables,) search engines will not match the search query with a page that contains the word *automobile* instead of the word *car*. Also, pages that use only the terms for specific types of cars (coupe, sedan, convertible) or specific brands (Ferrari, Alfa Romeo, Maserati) or a more general term such as *vehicle* will also not be listed in the search results. Second, if the words on the page exist in another language, numerous opportunities for mismatch occur. For example, a page in French would likely contain the term *voiture* or possibly *automobile*, whereas the word *car* in French is a false cognate and refers to a tour bus type of vehicle. Third, if the searcher mistyped *cart* or *var*, the system would be unable to identify those words as spelling errors. Fourth, variant spellings such as *color* and *colour* or *analyze* and *analyse* may not be considered matches. Fifth, words that have multiple meanings, such as *fixed* (repaired, stationary) or *fan* (enthusiast, ventilation device) require disambiguation in order to retrieve the correct results. Words may also have favored meanings between cultures even within the same language. For example, a person in the U.K. is referring to different phenomena when he/she speaks of *football* or a *boot*, than his/her counterpart in the U.S. False cognates in other languages further complicate this; for example, *pain* in English does not at all match in meaning with *pain* (bread) in French. Finally, as ambiguity exists at the word level, it

also exists at the phrasal level, like the old joke, "I see, said the blind man, as he picked up his hammer and saw."

All of the above issues have a common root: "considering only lexicographical aspects when trying to match Web pages and queries, ignoring semantic aspects. That is, search systems carry out a word-matching process instead of a concept-matching process (Garcés, Olivas & Romero, 2004, p. 565)." Because Hebrew is nearly always written without its vowels, homonyms (words with the same spelling but different meaning) are very common. In spoken Hebrew, the four letter combination of *mspr* has six different pronunciations and six different meanings, but all are written the same and thus are indistinguishable to formal search tools (Bar-Ilan & Gutman, 2005). Likewise, Arabic, being also based on abjads, has many opportunities for conflation in the written versions of words (Ahmed & Nürmberger, 2007). It is important here to remember that, while humans can distinguish between homonyms and other unmarked word pairs using context, search algorithms treat the words in user queries and the words that make a up a document as a "bag of words" and generally ignore the word order as well as the other aspects of the text that we would use to provide that context (Wang & Oard, 2012, p. 632)."

Since Large's research in 1999, a great many system enhancements and theories have been proposed to improve existing systems, overcome the aforementioned barriers, and thereby improve search results. Garcés et al (2004) list a number of different soft computing techniques being explored and implemented at that time:

- flexible adaptive sites that use Web patterns, user profiles, user behavior, or similar strategies

- the organization of retrieved documents into groups either using dynamic clustering or predefined thematic groups

25

- flexible query languages

- fuzzy association rules and fuzzy interrelations

- term interrelations stored in ontologies such as WordNet – a semantic net of word groups

- corpus-based search systems that rely on the probability of co-occurring concepts to disambiguate meaning

- "soft" disambiguation which creates a set of possible meanings and ranks them by relevance

- relative synonymy that uses concept-based vectors and a concept repository

- algorithms that consider the similarity between the query terms and the index term documents

- latent semantic indexing, which gives weight to terms that do not actually exist in a document because the term does exist in a similar document with co-occurring terms

While these are in development, they would need to be implemented separately for each individual language or group of languages in order to be effective in either monolingual or multilingual searching. It is not unlikely that system advances will overcome the barriers caused by language in the near future.

2.3.1.4. Dialect, nuance and lack of fluency

Comprehension of the sources being studied is important to any scholar; to a scholar using documents in another language, it can present a significant obstacle to their research. Madkour (1975) asserts that "only a perfect mastership of a language permits to bring out the inner meaning of a juridical text or to grasp the object and the real purpose of an invention hidden within the obscure literature of a patent." When the language and culture are not the scholar's own, it may add to the barriers they encounter. Specific issues include those of 1)

distinguishing and navigating dialects, 2) recognizing and understanding nuances of the language, and 3) assessing relevance of documents in light of less-than-optimal fluency.

Specific terms take on meanings as determined by the daily language of their users. Even similar terms between languages may have differing definitions between cultures. The language therefore expresses the thoughts and intent of the culture, and these subtle differences are hard to take into account when classifying resources or building tools to identify and access them (von Ungern-Sternberg, 2003).

However, many scholars master enough of a language to study its literature, its related cultures, and their histories. For many people worldwide, competence in a second language is not uncommon. In many cases, people have a "home" language, a "work" language, or a "school" language. They prefer to use a specific language to perform tasks in a particular context (Hansen & Karlgren, 2005). In Hansen and Karlgren's study, they investigated the speed and accuracy of relevance assessment based on language. The study observed native Swedish speakers with self-reported fluency in English who were presented with a list of documents from a search result. It was found that they were able to more quickly and more accurately assess relevance in their native language than they were in English. In their study of Dutch nurses searching in English on PubMed, Van Opstal, Vander Stichele, Laureys, and Buysschaert (2012) found that language factors presented barriers for effective on-line searching, including: lack of knowledge of domain terminology, lack of knowledge of the language of information and documentation, the interference of cognates, and poor spelling.

Language nuance can also be an issue for non-native speakers, mostly in assessing relevance and meaning when the desired documents are retrieved. In my pilot studies (Sabbar, 2013; Sabbar, 2016), two subjects recounted stories where either they or one of their colleagues

was unable to detect either irony, frustration, or sarcasm in the tone of a text, thus opening themselves to incorrect interpretations of the texts. To add to the complexity of this issue, historians and others doing research on past eras must be familiar with the nuances held by various words over time. For example, words used to represent various ethnic or social groups in the U.S. have changed dramatically throughout the past 100 or more years, and some words such as "gay" or "colored" may have not only had different meanings in differing historical contexts, but they would have elicited very different reactions and feelings as well.

As you can see from the above discussion, system-side studies are fairly common in the investigation of quantifiable language issues such as character sets or morphology, but they are less common in the discussion of more subtle issues such as nuance or context.

2.3.2. Cultural barriers

Culture can be described as "the way of life of particular human groups as all the activities that this way of life entails (McGarry, 1993, p. 59)." Language is one major component of culture, but there are many others. Among these are the authoring and publishing climate, indexing of literature within the language, library culture and conditions, political systems and climate, and societal norms including politeness.

2.3.2.1. Authoring and publishing

Not all languages and countries are equal when it comes to the creation and publishing of knowledge. Riley (1992) lists 19 languages that had 50 million or more speakers at that time and then contrasts that list with a list of only 13 languages that had produced over 10,000 books in that year. Only eight of the languages appeared on both lists: English, Spanish, Russian, Japanese, German, French, Italian, and Korean. The language with the most speakers worldwide – Chinese – and the language with the third most speakers – Hindi – did not make the second list

of languages with more books published. Of the widely spoken languages that did not make the publishing list, all except Portuguese are spoken in Asia or the Middle East. All five of the less-spoken languages that have higher publishing numbers are European: Dutch, Czech, Hungarian, Polish, and Swedish. Riley concludes that "the number of speakers is not a true reflection of the amount of printed information available in that language (Riley, 1992, p. 331)."

Web authorship requires less infrastructure than formal publishing, but the disproportions of authors vs. consumers between languages is still quite striking. Cairncross (1997) recounts the development of the internet and explains that in 1996, about 75% of all internet hosts were in the U.S., declining to 50% by 1997. Globalization has favored English, with four of twelve countries with the highest density of internet hosts being English-speaking and another five from northern Europe and Scandinavia where English is commonly spoken as a second language. Moreover, "the culture and not just the language of the Internet is also strikingly American. Its quirky blend of technocratic individualism, egalitarianism, and passionate resistance to government control all seem to many foreigners quintessentially American (Cairncross, 1997, p. 95)." While this history may have supported the concept of global English in the early stages of the internet, much has changed since the 90s.

Internet World Stats (2010) reported that 26.8% of Internet users are English speakers, followed narrowly by 24.2% Chinese speakers, 7.8% Spanish, 4.7% Japanese, 3.9% Portuguese, 3.6% German, 3.3% Arabic, 3% French, 3% Russian, and 2% Korean. All other languages together constitute a total of 17.8%. This represents a considerable redistribution from statistics cited by Bar-Ilan and Gutman (2005) from Global Reach (2004): 35.8% English, 14.1% Chinese, 9% Spanish, 9.6% Japanese, 3.5% Portuguese, 7.3% German, less than 2% Arabic (statistic not provided), 3.8% French, 2.5% Russian, and 4.1% Korean. I believe that what we

observe is an increase in internet use by languages often spoken in developing countries. As the proportion of non-English internet usage increases, search engines will increasingly need to accommodate queries in languages other than English (Efthimiadis, Malevris, Kousaridas, Lepeniotou & Loutas, 2007).

Best explained in the study by Meho and Tibbo (2003), a primary obstacle to conducting research that uses resources from developing countries is the lack of academic and research support to create and distribute knowledge on relevant topics. This is especially acute in stateless nations (e.g. Kurds, Tibetans, Chechens, and Basques) where potential literature by authors in their own languages may be repressed.

Finally, language can create its own cultural barrier. In a post-colonial era, people whose native languages are less spoken may need to master more common languages to advance in their own society. "To the extent that [language] can confer social, political, educational and cultural privileges, it becomes a barrier to those who do not or cannot enjoy such privileges. In this way language can become a basis for exclusion (Egbokhare, 2003, p. 8)." While children learn best in their own languages, they may need to master other languages to become scholars and authors. As users and beneficiaries of the world's knowledge, all efforts should be made to view knowledge as an expression of the culture in which it is created and to utilize it in the language that most naturally expresses that culture.

2.3.2.2. Indexing of literature in the target language

Even where literature exists in lesser spoken languages, it may not be findable using formal search tools. Even the best written and most valuable web sites are of no use to researchers unless they can be found. Numerous studies have shown that the main indexing services such as ISI (Institute of Science Index) or SSCI (Social Sciences Citation Index) have

strong Anglo-Saxon biases and often do not represent well non-English journals (Fernandez-Quijada, 2011; Johnson, 2006). This appears to be the same issue whether the journals are in the areas of humanities, social sciences, medicine, general science, or information science. "The ISI Citation Indexes have only included journals that are published in English. The dominance of the English language has been noted by a number of authors, highlighting the esteem attached to those LIS journals included in ISI's Social Science Citation Index, which are long established journals and mainly (but not exclusively) scholarly in their nature and American in origin (Johnson, 2006, p. 4)." In Nederhof's study of publications on modern language topics by Dutch scholars, he found that "the coverage of these source journals by ISI has an Anglo-Saxon bias (Nederhof, 2011, p. 119)" and also that less frequently spoken languages may be covered less well by citation indices. Specifically, less than half (20-40%) of the core journals in Dutch language and literature were covered, and that indexing of Scandinavian and Romanian language and literature sources by ISI were also problematic (Nederhof, 2011). Even Spanish, the second most common language in the communication category of the SSCI, had only 64 articles in 2009, compared to 1504 articles in English (Fernández-Quijada, 2011). An estimated 70% of scientific journals published in Latin America are not indexed by any service. This results in papers from the region in LIS being cited between 40% and 60% less than the world average (Ardila, 1999). Unfortunately, this imbalance seldom comes to the attention of or becomes the concern of the English speaking community (Johnson, 2006). However, when probed in studies, it can be documented that the foreign language barrier presents a considerable impediment to the dissemination and retrieval of information and revealed a need for library services and systems to address the problem (Riley, 1992).

 2.3.2.3. Library culture and conditions

In countries outside of the U.S., researchers and libraries alike are more likely to have to deal with multilingualism and multiculturalism and thus become more aware of how language and culture can present barriers between information providers and information seekers. Shoki and Oyelude (2006) looked at the language and cultural barriers experience by both information professionals and information seekers at public and university libraries in Nigeria. They state that "the language in which the information is presented and the cultural milieu from which the information emanates often interfere and become barriers to proper information flow and transfer. Information retrieval thus becomes difficult and information dissemination is made unsatisfactory (Shoki and Oyelude, 2006, p. 3)." Their study used grounded theory to derive a list of linguistic and other cultural barriers that presented users with difficulties in information retrieval. These included barriers such as:

- Ambiguous titles

- Inadequate subject descriptions in catalogs

- Lack of library materials of local color

- Organization schemes dominated by English and inadequately suitable for local resources

- Lack of available local resources

Meanwhile, their study also derived a list of linguistic and other cultural barriers that made information dissemination difficult for library staffs, including:

- Classification schemes that do not accommodate local publications or local languages

- Subject titles and classifications that are difficult to interpret

- Standard classification schemes (LCSH, AACR2) that are not broad enough

- Non-standard field modifications to classifications schemes by local librarians

Worldwide, libraries recognize that developing and maintaining collections that encompass multiple languages is both expensive and difficult (Karnes, Lajba & Shorb, 2007; Sexton, 2008). Therefore, most libraries develop collections only in the local language(s.) When titles in other languages are acquired, they require additional skill and effort to catalog. Whereas 75-80% of books in English can be copy catalogued from the records of the Library of Congress, only about 25% of books in major European languages have catalog records available (Kellsey, 2003). Titles in languages outside of Europe and North America are even rarer. As a result, scholars who rely on foreign languages sources must either: 1) access electronic resources remotely through databases, if they exist, 2) travel to the locations where the items exist (discussed later as a geographical barrier,) or 3) request that the resources be sent from the remote location by a library or other human intermediary. In few if any cases is international document delivery an easy solution. Considerable barriers, from the differing legalities of copyright between countries, to the lending rules of special collections, to the expense of the actual delivery often prevent scholars from acquiring sources through normal inter-library loan channels even in countries as otherwise accessible as France. While some commercial international document delivery services exist and some modest collaborations are beginning to take hold, the promise of international document delivery on a wide scale and with reliable results seems very far off (Gould, 2000).

2.3.2.4. Societal norms, bureaucracy, politeness and gaining entrée

Culture can be described as the context that gives meaning to language. Many everyday cultural factors can impact scholars' abilities to find or obtain sources in their target language. Callahan (2005) provides a dizzying number of factors that can impact a user's effectiveness when utilizing an interface from another culture. Intelligent and successful interface designers

take into account the understanding, perceptions, and sensitivities of the target culture. Something as simple as the choice of icons to use in the interface must be culturally appropriate. For example, American interfaces will use an image of a rural mailbox for "mail" and, by extension "e-mail." Trash cans and file folders are two additional icons that are culture specific. Mailboxes, trash receptacles, and instruments of document storage have very different shapes and forms in other culture. Callahan relates her own experience: "For example, as a cultural outsider, I was able to recognize an icon of a U.S. rural mailbox (because it had MAIL printed on its side) and understand its function, but failed to understand the significance of the red flag, which I interpreted as a decorative element (Callahan, 2005, p. 282)." The icons chosen, along with the colors and other aesthetic elements must be understood, and they must be considered appropriate. Layout of a web site or the nuances of the language may be entirely logical in one culture while at the same time being enigmatic or offensive to persons from other cultures. For example, a person used to a language the reads right to left will not view a web page in the same way or look for things in the same place as a person who reads left to right or vertically.

Different cultures have varying social and bureaucratic structures that one must navigate to achieve the desired outcome of his/her research. Crystal (2001) gives examples of the contexts in which language is used – from the courtroom to the street to the pub to the university. People growing up in one culture may occasionally violate the cultural norms for propriety and correctness in their own milieu, but people who did not grow up in the target culture may make many faux pas that will prevent them from gaining entrée to the locations, resources, and people that they should rely on. The vocabulary, tone, and even gestures associated with formal vs. informal speech can be a mystery for those unaccustomed to the cultural norms.

Many researchers rely on archives, including government documents as a basis for their study. Not all countries have open records policies, and those that do most often have a closure period during which access to records is restricted or not allowed. Scholars, including historians, must become familiar with these regulations and how they apply to specific types of records (Valge & Kibal, 2007). Even when records and other types of sources are open to research, additional barriers to access may exist related to gaining entrée. Many archives and libraries have policies about who can access and use the collection and/or the facilities, and there may be procedures that need to be followed to gain entrée to them. Archives of rare books and documents may have additional levels of security as well.

All of the scholars with whom I have spoken in my research could quite easily describe the requirements for gaining entrée to the libraries and archives they used, and many had stories of situations where access was either surprisingly easy, frustratingly difficult, or even impossible. However, there is virtually no literature or previous study concerning cultural differences – bureaucratic or social – that must be understood and navigated in order to access libraries worldwide. It is clear that credentials play a large role in many societies, but that others may rely more on nuances of human interaction and social status. A few of these phenomena, such as politeness, are discussed in the literature either in general or in the context of the workplace. Societal norms related to social status can be highly influential in determining the ease, difficulty or eventual failure of gaining access to libraries, archives, or other resources. In Japanese, for example, the concept of *keigo* – honorific language – is of great concern to both native speakers and foreign learners of Japanese. While a grammatical error by a foreigner may be viewed with amusement by a Japanese listener, failure to use the correct honorific language may result in an expressed or unexpressed instant emotional reaction. Often included in the

35

larger context of politeness and expression of respect, if the appropriate forms of honorific language are not used, foreigners approaching librarians or other professionals at archives or government institutions may be denied access for not having recognized the hierarchy or the relative status or power of the person they are addressing (Carroll, 2005). These phenomena are not unique to Japan or even to Asia. Scholars who need to approach individuals or institutions for information access would do well to learn how the concepts of directness, politeness, deferential language, hierarchy, social status and public self-image must be navigated in order to keep a keep interaction smooth and ultimately gain access to the resources they need (Yu, 2011; Carroll, 2005).

A Canadian study by Johnson and Duff (2004) found that history scholars often benefitted from and relied upon building relationships with archivists at the libraries and archives that they visited most often. It is not clear, however, that any similar user study has been done in relation to scholarly use of libraries and archives outside of North America or in the context of non-English research. It is recognized that foreign language expertise is often lacking among librarians in the United States, and that this creates barriers to the acquisition and cataloguing of resources in non-English languages and impacts the quality of assistance that can be provided by librarians to scholars who rely on those resources (Kellsey, 2003).

2.3.3. Geographical barriers

While Friedman (2005) argues that the world has become progressively flatter and tinier since 1492 and especially since a series of 10 events that began in 1989, his theories primarily revolve around the business world. He cites how accountants in India can be easily trained to do U.S. tax returns (Friedman, 2005, p. 14) and how large companies, small companies, and even individuals compete on a more level playing field in a global market. But what of scholarly

pursuits? His book outlines plans for companies and aspiring entrepreneurs and businessmen.

Yet, even in a flat world, he recounts airplane travel to distant locations to observe and interview

people in a very qualitative sort of way. Never does he indicate that the technologies of the time

have made the world so small that he can conduct his research from his desk or armchair at

home. The impact of geography, then, may be relative to one's means and flexibility. In his

article, Ghemawat (2007) presents the inverse view of the world as a very large place with

people in differing locations being less connected to each other or the "global community." He

states that the world's most globalized web site – Google – only reaches 28% of users in Russia

while its Russian competitors, Yandex (64%) and Rambler (53%) reach many more. Ghemawat

explains:

> "The biggest reason is the difficulty of designing a search engine to handle the
> linguistic complexities of the Russian language. In addition, these local
> competitors are more in tune with the Russian market, for example, developing
> payment methods through traditional banks to compensate for the dearth of credit
> cards (Ghemawat, 2007, p. 59)."

Ghemawat (2007) uses this example to explain that, while more and more people all over the

globe are more wired than ever, they are no more "global" and spend most of their time

connecting with their peers across town and their own local friends and family and only rarely

crossing national borders in their communication. The internet itself provides links between

cultures and countries, but even those do not necessarily encourage worldwide exploration and

information-seeking. Developing countries' web sites more often link to sites in richer, more

developed countries, and often to English-speaking countries. In contrast, seldom do developed

countries link to resources in locations in sub-Saharan Africa or locations in the Middle East or

South and Central America. Thus, the internet link structure encourages a more centralized and

directed pattern of information-seeking that creates a skewed global community that may not be

of help to many and certainly does not help to bring disparate cultures together (Brunn & Dodge, 2001).

Thus, despite the borderless nature of the internet that provides access to resources in many languages from the other side of the globe, most web users stay in local circles and stay in an on-line information space (Pharo, 2004) largely defined by a known language, familiar worldview, and societal norms (Huang & Chang, 2009). Huang and Chang studied the commonality of web site visiting between countries and hypothesized that the various factors outlined by Hofstede (2001) increase the degree of web site visiting commonality between two countries. Most of these are cultural, while some are demographic or economic. They are:

- Common language

- Common religion

- Similarity in the embracing of power disparity within culture ("power distance")

- Similarity in people's degree of anxiety experienced in unfamiliar situations ("uncertainty avoidance")

- Similarity in the culture's tendency toward individualism vs. collectivism

- Similarity in people's preference for assertive/competitive behavior vs. modesty/caring ("masculinity/femininity tendency")

- Greater numbers of internet users in both countries

- Shorter geographical distance between the two countries

- Similarity in the countries' levels of economic development

This dissertation specifically looks at scholars who need to utilize resources in other countries regardless of the commonality of web visiting between the U.S. and the other country. The point is that, if their need to access resources is uncommon (i.e. not many other people from

38

the U.S. use those same resources,) then the route to get to them is likely fraught with barriers. In the way that a frequently travelled road is often straighter and easier to navigate, a common usage pattern between scholars and resources will be more straightforward and effective, whereas unexplored territory and an unblazed trail take much more time and creativity to find and navigate.

Geographical distance can also have an effect on interactions between scholars and libraries. Brahme and Walters (2010) studied distance learning doctoral students compared to their on-campus counterparts and found that they value the physical library collection more, perhaps because they didn't have as ready access to it. They also found that distance students sought help from peers less often but sought help from instructors and library staff more often. Finally, while distance students demonstrated the same level of skill and knowledge of IR tools, they felt less confident and reported feeling "lost" more often. While their study presents a context different from that of scholars conducting their own research, the barriers, feelings, and strategies are very likely to be similar.

Sometimes, the resources needed by scholars simply aren't available electronically or through mechanisms like interlibrary loan, and scholars are required to travel to collections. For many scholars, travel to libraries and archives in the target country provides advantages over remote, electronic access, even if it were available. Perhaps one of the most appreciated aspects of digitized archival works is that scholars can preview **if** an item exists before travelling to the archive. The results of Anderson's longitudinal study confirm that researchers most often used e-mail and listservs to verify the existence and content of primary sources before travelling to visit the collection (Anderson, 1998).

The issue of availability of resources is especially prevalent in the area of government publications. Scholars of many eras in history benefit from government document repositories, but finding and identifying useful government documents is not an easy task. Pierce (2006) quotes a student who characterizes the problem this way: "The sheer number and variety of government publications is staggering, and it may be difficult to both find out that there's something interesting out there and sort it out of the haystack for all of the useless government documents (Pierce, 2006, p. 35)." Meho and Tibbo (2003) explain that, for scholars conducting research on stateless nations, such "government documents" may not actually exist. Similarly, countries that undergo frequent regime change may find that documents from specific eras have gone missing – either by accident, be neglect, or in an effort to erase or revise history in favor of the subsequent government.

In a study of librarians in Egypt and the U.S., Hover and Lu (2009) studied their inclination to access web sites in foreign languages or to click on the link to "translate this page" and use the machine-translated site. They included sites that were written in scripts unintelligible to the users such as Russian, Japanese, etc. They made comparisons not only between the Egyptian and American subjects but also between those who were monolingual vs. bilingual. What is significant for my research is that they recorded the reasons that subjects gave for clicking (or not clicking) on the links. While language was a factor, as anticipated, geographic and cultural considerations were a larger factor than they had thought. For example, when the topic to be searched for was about a Chinese film, subjects were more likely to click on a Chinese language link and have it machine translated than they were to click on a link in a language more intelligible to them (such as French or Spanish.) When searching on a science topic, the Egyptian librarians were more likely to click on "Science Afrique" in French because it

was related to their own continent of Africa. So, even when documents are available electronically, geography is a factor in information-seeking.

Even in cases where documents may be available on-line, many scholars – historians in particular – prefer to view the originals, explaining that context is often removed when documents are scanned (Duff, Craig & Cherry, 2004). Historians value more than just the words contained in the documents. Even carefully scanned copies do not reveal the type of paper, the condition of the volume or item, the ink used, and other factors that can have significance to historians. Graham's survey reports that historians often ignore online versions primary sources in favor of handling the originals even if travel is required to do so (Graham, 2000). Moreover, fifty percent of historians studied by Graham did not consider a digitized document to be equivalent to the original.

Of further import is the significance of browsing through a set or series of documents. While digitized documents tend to present themselves in isolation, archived documents are usually organized in boxes or files that represent collections. These collections represent a chain of events, an evolution of thought, a pattern over time, or any number of other connections that are significant to the historian but not evident from documents viewed in isolation. Viewing documents in a series can lead to the discovery of unexpected patterns that cannot be duplicated in online viewing (Pierce, 2006; Rimmer, Warwick, Blandford, Gow, Buchanan, 2007). In the research of Duff and Johnson (2002), the context of the documents is emphasized, including the reason for the collection, the provenance, the identity of the collector, and the relationship between the documents in the collection. "Historians must comprehend the records in their context rather than as separate disembodied items. Without this contextual information, the

historian could easily misinterpret the meaning or significance of the information in an individual record (Duff & Johnson, 2002, pp. 487-488)."

2.3.4. Global English and CLIR – Solutions that do not always fit

The concept of "global English" is sometimes touted as a solution to language barriers in domains from business to information retrieval. While it may be possible to legislate policy that pilots all over the world must speak English as the language of air traffic control, or while multinational corporations and international NGOs may be able to rely primarily on English, to expect scholarly research to be relegated to a single language – reflective of only a few specific cultures – is not only unrealistic, but it masks the beauty of the richly diverse world in which we live and the documents and resources that represent it. Kellsey and Knievel (2004) explain that, since the 1960s, less students in the U.S. are studying languages, and one might conclude therefore that today's scholars would be unable or unwilling to use non-English sources in their research. Upon conducting a longitudinal citation study, however, they found that the frequency of foreign citation has actually increased in nearly all areas of the humanities. The implications they suggest are not related to IR systems but rather to the collection development policies of libraries.

To think that all scholars can or would want to write in English is naïve. Of the 109 Nobel laureates in literature, 26 of them wrote in English (3 of whom were bilingual), while the others wrote in 24 other languages (nobelprize.org, 2013). In the sciences, even though it is relatively common for well-known and well-published scientists to write in English, many scientists from other countries are at a disadvantage to read and benefit from existing research reported in English, to then translate that knowledge into his/her local context, and/or to make the results of his/her work available both locally and on an international level (Meneghini &

Packer, 2007). These issues are exacerbated by current trends not only in publishing but also by database providers. In 2000, MEDLINE, for example included only 10% of articles in languages other than English, and they increasingly deselect journals from non-English speaking countries (Loria & Arroyo, 2005).

Janelle (1991) lists 12 languages that were spoken by 100 million or more people in 1987. He explains that "the 'babble of languages' complicates the vision of an interdependent world; however, communication among those of diverse cultural-linguistic backgrounds has been facilitated by the use of only a few languages in international dialogue. The United Nations recognizes only six official languages for its discussions and publications (Janelle, 1991, p. 56)." However, the idea that reducing the number of commonly spoken languages makes the world a smaller and flatter place is not a helpful concept for the scholar conducting research and relying on documents that may or may not be in one of these more mainstream languages. If the scholar is conducting research on the history of Japan, the literature of Colombia, religious culture in India, or the political parties in Germany, their sources will be in Japanese, Spanish, Hindi or Urdu, and German, respectively. No amount of promotion of the concept of global English will change the language of the documents and sources they rely upon. And even in cases where those languages are in the mainstream, the aforementioned language, cultural, and geographical barriers will still exist.

A second movement and area of study that is often seen as a solution to the language barrier is CLIR – cross-language information retrieval. Often called universal text retrieval, its goal is to return relevant text results regardless of the original language of the text and of the user query (Davis & Ogden, 2000). Many IR studies that look at language as a factor have focused on cross-language information retrieval (CLIR) systems. In fact, the factor of language is so

often associated with CLIR, that nearly every time I explain to someone that my dissertation focuses on language and its relation to information-seeking, a listener in the field of IR will say "Oh, cross-language information retrieval."

CLIR research often looks into the searcher's lack of knowledge of the language as a major language barrier in IR and how CLIR systems can help people find and utilize documents that exist in a language they do not speak or read. While we cannot dismiss that lack of knowledge of the language can be as significant a problem as lack of domain knowledge and/or lack of system knowledge, this is not the only problem that language poses to information seekers. Tools like Google Translate can translate entire web sites into a target language, possibly opening up bodies of knowledge to people in other languages. Admittedly, machine translation can result in "clunky" grammar and dubious word choices, and many people do not trust them (Hover & Lu, 2009; Notess, 2008). However, the translation algorithms are improving, and often, the results are readable enough to assess the relevance of the document. When documents containing multiple languages or a variety of languages are being searched, the tools needed to return relevant search results are still not able to cope well, and the performance of these tools does not always correspond to the information-seeking behaviors of scholars, so much work is left to pursue (Ha, 2014). Moreover, while document indexing was once done largely by humans who may or may not have knowledge of the languages in question, this indexing is now largely done by computer algorithms which, in turn, can use translation algorithms to make the indexed terms available to searchers in other languages.

It is my belief that sources must be created, exist, be indexed, and be findable in their native languages, and my focus is that understanding which information-seeking strategies are the most "language proof" based on the experiences of effective scholarly researchers is the first

step in building that understanding. Whether the solutions lie in improved systems; societal changes in domains like publishing, open access, knowledge dissemination; or improved services by information providers and repositories such as libraries, archives, and other similar institutions, there must first be a clear understanding of the aspects and properties of language, culture, and geography create these barriers.

2.4. Opportunity for new research

As mentioned earlier, the intersection of research between information-seeking research and language research is minimal. There is almost a total lack of existing research that focuses on how language, culture, and geography can impede the success of information-seeking researchers who do speak the language(s) in which they conduct their research. With the history of libraries and their development centering largely in the English-speaking world, the result has been a predominance of English-language resources and the development of systems that can catalog and retrieve these resources.

A fairly sizable corpus of system-oriented studies provide a solid explanation of how language issues impact the success of formal IR tools, but it does not begin to explain user research patterns or how language barriers affect scholars' choices of information-seeking strategies. This provides both a challenge and an opportunity for my research. Using the wealth of system-oriented studies will help me to understand and represent the language issues that affect scholars conducting their research. The lack of user-oriented studies means that there is ample opportunity to explore this uncharted area.

Specifically, very little prior research has been done regarding cultural and geographical barriers to the acquisition of sources needed to conduct research. Many of the previous models of information-seeking do not differentiate between the identification and acquisition of relevant

sources. For scholars in the U.S. using non-English and non-American sources, this distinction is significant. Even when a source can be identified and is known to exist, access to it is not at all a given. Whether the source is an Argentinian newspaper, a set of letters and family documents in Uttar Pradesh, science fiction stories in East Germany, or government documents in Mexico, knowing that they exist does not mean that they are accessible to scholars. In my pilot studies (Sabbar, 2013; Sabbar, 2016), nearly every scholar had stories to tell about the heroic lengths they had gone to in order to acquire the resources they had identified earlier. Some of these were related to navigating the cultural requirements of access to libraries and archives, and many were related to enlisting the assistance of human intermediaries to access the item in a remote location. Each of these experiences will help build a framework and description of the most effective strategies and how they help to overcome barriers of language, culture, and geography.

The study of information-seeking strategies is well suited to the analysis of language, culture, and geography as factors and barriers. Until now, no user-oriented studies have been undertaken to specifically study the information-seeking strategies used by scholars who look for and rely on resources in languages other than English. Many unexplored questions exist about how and why scholars select the strategies they use and how barriers created by language, culture, and geography create the problematic situations that cause alternate shifts of information strategies.

My research proposes a new conceptual model – the information triangle – to explore and document the strategies that scholars use in planned, disruptive, and problematic situations and the way that language and related factors influence those choices. Using qualitative methods to

collect and analyze data from scholars using a variety of languages in their research allows me to identify patterns and establish correlations between barriers, strategies, and shifts.

The practical implications of this research are a greater understanding of the special challenges and needs of non-English searchers working from the U.S. This increased understanding may be useful for developers of IR systems; it should be of significant interest to information professionals such as librarians and archivists; and it will certainly be of interest to other scholars who may indeed be able to improve their information-seeking success due to the increased insight into information-seeking in their unique situations.

3.1. Overview

This study makes use of a new conceptual model – the information triangle – developed as a result of my prior research, which allows for the description and categorization of the information-seeking strategies used by scholars throughout a specific information-seeking task as well as classification and visual representation of the shifts between strategies. (See Figure 3.1) While the three classifications that make up the sides of the information triangle correspond to some theories in the previous literature, they were derived primarily inductively and directly from the subjects in my previous pilot studies (Sabbar, 2013; Sabbar, 2016). In this dissertation, the list of strategies is inductively built based on participant input which supports the form of the triangle and the properties of the sides. The shifts described by the subjects are mapped from the data collected inductively using qualitative methods.

Figure 3.1 - The Information Triangle

The three categories that make up the information triangle are "formal system strategies," "informal resource strategies," and "human interaction strategies." This three-sided model most closely fits Ingwerson and Jarvelin's (2005, p. 259) description of information-seeking as "the acquisition of information from knowledge sources, for instance, from colleagues, through (in)formal channels, and from an information system." Choo, Detlor, and Turnbull (2000, p. 2) explain informal search as "involving a relatively limited and unstructured effort" and "formal

search as deliberate or planned effort to obtain specific information or types of information about a particular issue. Search is formal because it is structured according to some pre-established procedure or methodology."

3.2. Strategies

The first purposes of the information triangle are to categorize information seeking strategies used by the scholars in the study. It is therefore essential that we define the term "strategies" and understand how they are gathered and categorized.

3.2.1. Definitions

Beginning in the mid-1980s, various scholars have proposed definitions of information-seeking strategies and described their related aspects and components. Prior to this, some of the literature discussed only *search strategies*, which were largely viewed as specific to on-line searching, where the computer is the unifying factor rather than the human being who undertakes the search as part of an overall information-seeking process (Bates, 1979a, p. 306). However, beginning with researchers like Bates and Ellis, the focus began to shift to the needs, activities, and strategies of the users themselves.

In 1989, Ellis – one of the earliest researchers to focus on the information-seeking patterns of users as seen and described by the searchers themselves, refers to *information-seeking activities* but never actually uses the term *strategies*. Moreover, he does not define information-seeking activities but explains how these discrete activities are part of information-seeking patterns which he then breaks down into his six characteristics. An example given by Ellis of an information-seeking activity that begins part of an information-seeking pattern is to identify a key paper or seed document – a source that is relevant to the user's research or information need. Ellis doesn't consider **how** the scholar identifies the key paper, stating that it could be "either

49

one which the social scientist already knows or is told of, one which is on the boundaries of his existing interest, or one which is found or worth following up from a keyword search (Ellis, 1989, p. 176)."

Ellis' contemporary, Marcia Bates, utilizes the term "strategies," of which she lists six: Footnote chasing (backward chaining,) citation searching (forward chaining,) journal run, area scanning, subject searching in bibliographies and abstracting and indexing (A&I) services, and author searching. Many of these correspond to strategies identified in my pilot studies (Sabbar, 2013; Sabbar, 2016). For example, Bates' footnote chasing was generally referred to by my subjects as "citation tracing" – the act of using the footnotes and references from a seed article. Area scanning could be considered a superset of "visiting libraries" identified by my pilot studies. However, some of her strategies were more prevalent in the 1980s when scholars more often used print indexes and citation guides available in libraries. Bates does not actually define the information-seeking strategy as a term or concept, but she does imply that strategies change when the scholar changes either his/her technique or his/her source (Bates, 1989). She does not specifically point to techniques and sources as components of the information-seeking strategy, but we can draw some similarity between these and the methods and entities discussed by Xie (2008, pp. 233-236). That is, in order to classify information-seeking strategies, one may separate them based on the distinct techniques or the class of sources that they use.

Belkin (1993) discusses *information-seeking behaviors* as the behaviors we engage in to resolve a problem, respond to a gap in our life-path (Dervin, 1983) or resolve an anomalous state of knowledge (Belkin, 1993). He gives examples of these information-seeking behaviors: "asking a colleague for advice, browsing through journals to keep up-to-date, searching in a library for some specific information (Belkin, 1993, p. 56)," all of which are ultimately

50

interactions with texts. Belkin, Marchetti & Cool (1993) discuss information-seeking strategies

from a somewhat different perspective, giving examples of "searching for some known and

identifiable item(s); searching for items similar to some known item; searching for items on

some identified topic; looking around for something interesting among items; inspecting items

and their contents; identifying useful items by inspection; and browsing among item descriptors

and item organization schemes (Belkin et al, 1993, p. 325)."

Kuhlthau (1991) defines the related concept of the information search process (ISP) as

"the user's constructive activity of finding meaning from information in order to extend his or

her state of knowledge on a particular problem or topic (Kuhlthau, 1991, p. 361)."

Marchionini (1988) classified ISSs into two major groupings either as analytical

strategies, which are more intentional and systematic, or browsing strategies, which are less

formalized (Marchionini, 2006). This differentiation between searching and browsing is further

developed by Yuan and Belkin (2010a, p. 1988) who divide the strategies that they observed into

"those characterized by the method of searching, and those characterized by the method of

scanning, within the specific information behavior of access." This latter study contributes

significantly to my model and background in its treatment of multiple ISSs within the same

search task and its focus on the access stage of the information-seeking process. Yuan and

Belkin (2010a) specifically investigate the ability of IR systems to support different ISSs,

including browsing or specified searching.

To better understand the chosen concept of strategies, we should back up somewhat and

contrast strategies with the more basic concept of information *tactics*, defined by Bates (1979a),

as "a move made to further a search." *Moves*, are defined in turn by Fidel as "modifications of

query formulations… made by searchers to improve search results (Fidel, 1985, p. 61)." Shiri &

Revie also discuss moves in the context of searching and divide them into cognitive and physical moves (Shiri & Revie, 2003). For my research, however, this constraint of moves to the context of searching in formal IR systems is too limiting.

As for tactics, Bates (1979a, 1979b) identified 29 distinct information tactics that can occur during the information-seeking process as well as 17 idea tactics. At the time, the audience of these concepts was the information professional, perhaps because the tools available in 1979 were designed for expert use and not for the lay user. As such, the tactics proposed by Bates are intended for complex search situations rather than for known fact searches where ready reference tools may quickly provide the answer. She indicates that "search theory and practice may be advanced through a greater attention to the specifically human, psychological processes involved in searching, as distinct from the logical or formal properties of the process (Bates, 1979a, p. 205)." Bates continues on to define and contrast strategies and tactics, tracing the terms back to their usually military contexts. "Strategy deals with overall planning; tactics deals with short-term goals and maneuvers (Bates 1979a, p. 207)." Bates also expands upon the concept of tactics and moves, combining a complex sequence of moves and tactics in to "stratagem" which involve both information domains and modes of seeking (Xie, 2002). In a way, Bates foreshadows the concept of shifts, which will be discussed below and which constitutes a focus of my research.

Harter and Roger-Peters (1985) compiled a list of 101 tactics in six categories: overall philosophical attitude and approach, language of problem description, record and file structure, concept formulation and reformulation, recall and precision, and cost/efficiency. Shute and Smith (1993) researched and wrote at a time when effective research using search tools for bibliographic databases required a great amount of skill and often required an intermediary with

knowledge of the subject area being investigated. They promoted the theory that studying the content-related tactics of the intermediary is a useful endeavor for system developers. They point out, for example, that a person researching a chemical aspect of wastewater treatment would engage in tactics related to their knowledge of the types of terms that would improve the search results. The intermediary might know for example that "filtration" can also be expressed as "process for removing pollutant," and that they can then effectively use synonyms, Boolean operators, more general or more specific terms to achieve the desired results. Compared to tactics, strategies may include a series of actions and embody multiple dimensions.

In this study, I define information-seeking strategies as "activities undertaken by a researcher to advance an information-seeking task or meet an information-seeking need." My definition of strategies is a multi-dimensional one composed of two dimensions: methods and entities. Methods involve the types of actions and tactics undertaken by the researcher, and the entity involves the source of the information being sought. This definition will be explored in more depth in section 3.2.3 as it relates to the conceptual model of the information triangle.

3.2.2. Building the list of strategies inductively

The information triangle conceptual model is designed to provide a framework for present and future user studies, so it does not provide its own list of strategies. Rather, in each study, the researcher gathers a list of strategies that are enumerated by the participants. Two previous pilot studies – one related to scholars searching for sources in languages other than English (Sabbar, 2016), and one that studied scholars using primary sources (Sabbar, 2013) - resulted in very similar lists of strategies, but not identical. Most other models, many of which are explained below, have a pre-compiled list of moves, tactics, or strategies associated with

them. This inductive collecting of strategies from the subjects is well suited to the qualitative research methodologies used in this dissertation.

Since qualitative research collects data to describe and understand phenomena, theories and patterns arise inductively as data is collected. This is often referred to as "grounded theory" because the theory is built "from the ground up" based on the data that is collected (Gorman & Clayton, 2005). Grounded theory is based largely on the works of Glaser (1965) and focuses on the process of generating theory through the process of research (Patton, 2002; Glaser & Strauss, 1967). "Generating a theory from data means that most hypotheses and concepts not only come from the data but are systematically worked out in relation to the data during the course of the research (Glaser and Strauss, 1967, p. 5-6)." Grounded theory links theory to method, believing that how you study the world determines what you learn about the world. Because the researcher is closely entrenched in the real world, their observations and results are grounded in the empirical world – the world of on-going group life (Patton, 2002, p. 125).

Because the list of strategies is built inductively from the statements of the subjects, some of the strategies may be used frequently by numerous subjects, and some may be mentioned uniquely by one or two subjects only, creating a natural frequency distribution that will likely have a long tail. Also, because the list of strategies is collected from the data, the names of the strategies are in plain English, reflecting the actual terms used by researchers and not technical terms used by information scientists. For example, whereas some previous models have used scientific terms such as "chaining backwards," this dissertation uses lay terms such as "citation tracing" which are similar to those gathered from and understandable by the subjects themselves. Finally, while previous models have sought to enumerate a full complement of strategies, a list built inductively may not include all possibilities. In my pilot studies, for example (Sabbar,

54

2013; Sabbar, 2016), while many scholars talked about citation tracing, which is often called

chaining backward, no subjects mentioned its counterpart of chaining forward – either by that

name or by any other – so that strategy does not appear on the list and is not discussed or

categorized. A table of the strategies collected in this study is given in Chapter 5 in Table 5.2.

3.2.3. Three sides of the information triangle

Despite some common definitions of what a strategy is, there is no prevalent model of

how to classify or group strategies. Chen and Dhar (1991) classified strategies into five types

based on the cognitive processes of users: the known-item-instantiation strategy, the search-

option-heuristics strategy, the thesaurus-browsing strategy, the screen-browsing strategy, and the

trial-and-error strategy. Meanwhile, Pejtersen (1989) proposes three overarching ISSs: analytical

search, search by analogy, and browsing, while Belkin, Cool, Koennman, Ng, and Park (1996)

propose a taxonomy of search strategies: term strategies, database strategies, interaction

strategies, and search strategies. Xie (2008) discusses a multifaceted classification schema for

ISSs, which includes the following dimensions: 1) goal of the interaction (learn, select) 2)

method of interaction (scan, search) 3) mode of retrieval (recognize, specify,) and 4) types of

resources interacted with. Pharo (2004, p. 6) elaborates that "search task strategies may include

specification of which sources to use and the techniques that should be used in order to explore

them." Some of these classifications will be discussed in the context of dimensions, below.

This new conceptual model divides strategies into three categories. Each of the three

categories is described and explained below. Each takes into consideration the dimensions of

method and entity, although some ambiguity is to be expected.

Formal system strategies are those related to the "formal bibliographic apparatus" (Green,

2000, p. 202) in which the method of IR search tactics are used in conjunction with entities such

as periodical databases, OPACs, electronically indexed and searchable finding aids, digital libraries and repositories, and web-based resources. There are three characteristics that are generally used to categorize strategies as formal systems strategies: 1) search-based, 2) computer systems-based, and 3) having an interactive interface.

Numerous scholars have gone to some length to differentiate between formal and informal strategies. Most differentiate formal strategies as search-based and informal strategies as browse-based (Ruthven, 2008). Another common distinction is that formal strategies involve computer-based systems while informal/manual strategies do not (Wilson, 2000). It is overly simplistic, however, to say that all computer-based systems are uniquely search-based, since many system designers have seen the value of browsing techniques and have sought to incorporate browsing capabilities into their computer-based systems (Chang and Rice, 1993).

Xie (2008, p. viii) explains in more depth the components that make up an IR system, one of which is an "interface for user input and system output" that supports the presence of an interactive interface. Thus, the level and type of interaction help define formal systems strategies as compared to the other two categories.

Informal resource strategies include those that are traditionally considered in relationship to the entity of print sources, including: browsing; citation tracing; and using indexes, bibliographies, and static finding aids. Using informal resource strategies, scholars work directly with resources or their surrogates without the intermediary of an interactive IR system or interface. Strategies such as browsing web sites or bibliographies are considered informal even if they are performed on-line using a computer because the user does not formulate and present a query to a system which then returns a response.

Chang and Rice (1993, p. 235) focus on the method dimension and differentiate browsing from formal searching: "Browsing has been loosely described as a kind of searching in which the initial search criteria or goals are only partly defined or known in advance." They discuss three different types of browsing: search browsing, general purpose (semi-directed), and serendipity browsing. They further explain: "Browsing has been proposed as an alternative approach to IR that does not use Boolean operators or require specific search queries." Browsing is often used when the user "is unable to specify initial search requirements or is unfamiliar with the terminology of a domain of interest, or when he or she wishes to discover the general information content of the database" (Chang and Rice, 1993, p. 238).

The term "informal" is used not only by Ingwerson and Jarvelin but also by Green (2000) in her discussion of "informal techniques." Bates (1989), Wilson (2000), and Xie (2008) all refer to "manual" strategies, sources, or techniques, which are generally synonymous with the informal strategies use in my model.

Interactive human strategies include those that rely on consulting with human entities as either an intermediary or a direct resource, including archivists and librarians, colleagues, and other knowledgeable humans, whether directly or through electronic means. Attending conferences and interacting with other participants and obtaining sources from other humans are also interactive human strategies. While previous literature often differentiates search-based from browsing-based strategies, and a few prior models mention human sources in passing, few have classified information-seeking from human sources as strategies. When looking closely at the method dimension, I would propose that interactive human strategies are communication-based rather than browsing-based or search based.

A key differentiating characteristic of human strategies is the difference in the level of interactivity. In formal strategies, the interaction is predictable – beginning with a query from the user and a response from the system; iterative interactions with formal systems generally require re-initiation by the user. With informal resource strategies, the user is a consumer of the resources, and the interaction is generally one-sided with the user pursuing the strategy with no real participation from the resource being accessed. The resources – be it the library shelves being browsed, the citations being consulted, etc. – do not initiate interaction nor do they respond to queries.

In human interactions, there is no predictability similar to that of a formal system; a query to a human intermediary may result in subsequent questions to assist the user in formulating their needs or suggestions of alternate strategies rather than simply and directly providing a listing of relevant sources. If the human intermediary does not understand what the user wants, he/she will often pursue an exchange of question and answers until some satisfactory result is achieved. Interactive human strategies are also the only category where the intermediary – rather than the researcher – may initiate the information-seeking process or may provide added information at a later time without further prompting from the user. Human intermediaries also differ from systems (at least at the present time,) in that they can often understand the concepts and ideas that the researcher is looking for even in the absence of the correct key words and search terms.

3.2.4. Dimensions of information-seeking strategies

In my discussion of strategies, above, I have briefly mentioned **dimensions**. Belkin, Marchetti & Cool (1993) were among the first to consider dimensions that define ISSs. They explain that dimensions, also called facets, include: the *goal* of the interaction, the *method* of the interaction, the *mode* of retrieval, and the *type of resource* interacted with (Belkin, 1993, p. 62;

Belkin, Cool, Stein & Thiel, 1995). They were able to identify these by observing information-seeking behaviors in various contexts, and they posit that specific strategies can be categorized according to these aspects. They show two possible choices for each of the facets and then create a table of 16 possible prototypical ISSs to represent each of the possible combinations. For the facet of method, they contrast *scan* vs. *search.* The two specific goals are *learn* and *select*; the two modes are *recognize* and *specify*; and the two types of resources are *information* and *meta-information* (Belkin, Marchetti & Cool, 1993). Marchionini (2006) enumerates three kinds of search activities that correspond most closely to Belkin's goals: *lookup, learn* and *investigate*. The *lookup* search activity is the most basic, being used for fact retrieval, "known item" searching (e.g. finding the book "El Cid",) question answering (e.g. "What year was Emperor Hirohito born?,) and verification. These lookup activities are often performed using database management systems or library tools, which are best suited for this type of well-structured searching. Both *learn* and *investigate* fall into a realm that Marchionini considers exploratory searching, sometimes referred to as subject or topical searching (e.g. "What were the causes of the French Revolution?" or finding sources related to the rise to power of the Empress Dowager.) *Learn* encompasses knowledge acquisition, comprehension and interpretation, while *investigate* includes knowledge accretion, synthesis, evaluation, and discovery. While web search engines were initially best suited for lookup activities, users have come to increasingly rely on them for learning and investigation, to explore topics and broaden their knowledge on a given subject. One unique activity is that of social searching, which Marchionini places under the *learn* activity, stating that it results from people's aim "to find communities of interest or discover new friends in social network systems (Marchionini, 2006, p. 43)."

Xie (2008) expands the concept of the dimensions to include the dimensions of interactive intentions, dimensions of retrieval tactics, and dimensions of information-seeking strategies. Interactive intentions are the sub-goals that must be achieved by the user in order to complete his/her search goal or task (Xie, 2007; Xie, 2008; Xie, 2009). She names twelve distinct types of interactive intentions: Identify, learn, explore, create, modify, monitor, keep records, access, organize, evaluate, obtain, and disseminate. The contrast between the interactive intentions of identify and access is significant to my own research, since researchers looking for resources not in English may often identify an entity (i.e., they discover the existence of an information source, author, or reference to get started), but they may encounter insurmountable barriers when they try to access the item (i.e., they cannot gain access to an item because of its location.)

These interactive intentions are defined by entities and their attributes. Nine entities exist in the model: data/information, knowledge, concept/term, format, item/object/site, process/status, location, system, and human. These represent the entities with which the user intends to interact. Each entity of interactive intentions in turn has from one to five related attributes that specify the traits or elements of the entities. For example, an entity might be a concept/term, and its attributes could be broad, narrow, synonym, or parallel. Another possible entity is a human, of which the attributes could be professionals, experts, friends, or colleagues (Xie, 2008, pp. 231-234). The idea of the human as an entity is specifically developed and expanded further in this dissertation, looking at a researcher's interaction with a human as representing different strategies than when researchers interact with a system or an information source such as a journal or a book.

According to Xie (2000, 2002), retrieval tactics also have dimensions: methods, entities, and attributes, the model for which is constructed like that of interactive intentions. The methods that make up tactics are: scan, manipulate, specify, track, select, survey, extract, compare, acquire, consult, and trial-and-error. These are then accompanied by the same nine entities and accompanying attributes as found in the dimensions of interactive intentions. Thus, the difference between the interactive intentions and the tactics is that the methods of the tactics represent actual physical moves that happen rather than methods that represent the goal or desired outcome that the user is looking to accomplish. Some examples of methods related to tactics are: specify, which means that the user states their search statements such as identifying the author, title or subject to searched; select, which is the act of picking up an item from among a series or from a location, such as entering an item's number, choosing an item from a list of results, or pulling a book from amongst others on a shelf.

Limited prior research has been conducted on the dimensions of information-seeking strategies. For my conceptual framework, I have adopted two of the dimensions from Xie's (2000) model: methods and entities. The entities (resources) include information, information objects, and humans (Yuan and Belkin, 2010). The methods correspond to the tactics used by searchers in the research process, and generally include searching, scanning, or inquiring (Xie, 2008). Considering the dimensions of a strategy allows for shifts in strategies that may seem slight and even nested – like a strategy within a strategy – when the subject makes a change in only the tactic or the entity. For example a user may make a slight entity shift by using a Google search engine in a different language, by reformulating the query somewhat, or shifting their attention to another part of a formal system tool that provides chaining forward or chaining

backward capability based on the search results. The dimensions of entities and methods are included in the discussion of shifts, below.

3.3. Situations

In my discussion of shifts, below, I will correlate the types of shifts that users make with the types of situations they encounter. A *situation* is the user's perception of all the elements of the context at a specific moment based on his/her assessment of the interaction between his/her plan and the IR system or resource being used. Based on Hert's (1997) list of elements that make up situations, a situation can be defined as the product of interactions between users and IR systems (Xie, 2007). In this dissertation, we often refer to elements as factors. Whether called factors or elements, these items can be positive, negative, or neutral. The three types of elements proposed by Hert (1997) include: elements associated with the problem (subject area, specific requirements, point in the process, whether it is a group or individual project); elements associated with the system response (nature of retrieved sets, status messages, features of individual entries); and elements associated with users (knowledge, attitudes, expectations, emotions, etc.). Key elements of the situation include its social-organizational context, the user's plans, and his/her personal information infrastructure. The personal information infrastructure, in turn, includes factors such as his/her cognitive abilities, domain knowledge of the problem/task, system knowledge and skill level, and general knowledge of and expertise with information-seeking (Xie, 2007).

Two types of situations were originally discussed by Schutz and Luckmann (1973): 1) Routine situations are those in which things go as planned, everyday knowledge and strategies can be used, and results are fairly predictable. The user may employ a relatively consistent sequence of ISSs to fulfill his/her information need. 2) Problematic situations are the result of

unforeseen barriers, in which the user may need to adjust his/her intentions and/or tactics or

extend his/her knowledge of the domain or the system, shifting to alternate ISSs and adopting a

new plan to arrive at their information goal. Xie (2007) adds another type of situation: disruptive

situations. In a disruptive situation, the user may be sidetracked or moved along in a positive

way by an unexpected discovery. This may result in new or different ISSs being adopted, and

the user may eventually return to his/her plan, or they may continue on the path brought to light

by the disruption. Sometimes referred to as serendipity, these unplanned yet fortuitous

discoveries have been discussed by researchers such as Foster and Ford (2003).

3.4. Shifts among information-seeking strategies

The second purpose of the information triangle is to visually represent the shifts that

subjects make among strategies. It is then necessary to define the concept of shifts and how they

can be categorized.

3.4.1. Definitions

Since my study is largely concerned with why scholars change strategies when they

encounter barriers caused by language, we need to define, understand and incorporate the

concept of *shifts*. Previous studies, primarily those of Xie (2007, 2008), discuss shifts between

information-seeking strategies, which are key to my research since my primary focus is on why

and how scholars change strategies when they encounter barriers caused by language.

Kuhlthau's (1991) information-seeking process model is one of the earliest to bring up the

concept of shifts, discussing the change of stages in task performance. Belkin, Marchetti & Cool

(1993) state that "in systems not constrained to a single ISS, people often change from one

strategy to others during the course of a single information-seeking episode (Belkin et al, 1993,

p. 325)." Note that this particular discussion of shifts exists within the context of an IR system

rather than in a more holistic context of the array of tools and strategies available to users, including formal IR systems, informal resources, and human interaction. They do, however, provide examples of shifts within the information-seeking process that includes shifting from one type of IR system to another. In contrast, Belkin, Cool, Stein & Thiel (1995) do include activities and behaviors outside of the realm of IR systems such as scanning library shelves, then shifting to an activity such as consulting a catalogue to find the subject heading of a known item, and then searching an IR system using those terms (Belkin et al, 1995). In this dissertation, I define shifts as "changes made by users from one information-seeking strategy to another in order to advance the information-seeking process."

Xie (2008) has done the most extensive work in this area, breaking down shifts by their levels: shifts in interactive intentions and shifts in retrieval tactics. The four types of shifts in interactive intentions are:

- Type I - **Routine situations lead users to planned shifts**. Here, scholars shift between strategies (and sides of the information triangle) according to common research habits and planned information-seeking threads developed through their formal education or experience. For example, a scholar may often begin by searching a library on-line catalog (OPAC) to identify sources and then make a planned shift to visiting a library to acquire the print resources identified in the catalog.

- Type II - **Disruptive situations guide users to opportunistic shifts**. Here, a scholar using one strategy makes a positive, serendipitous discovery that leads them to another strategy. In this case, when the scholar visits the library to acquire the desired print resource, he/she may find alternate print resources that are also helpful, or a librarian may approach him/her to suggest additional resources.

- Type III - **Problematic situations turn users to assisted shifts**. Barriers and obstacles lead the user to a point where their chosen strategy is not yielding the needed information. The user must then "learn" by either developing system or domain knowledge in order to continue. In these cases, the IR system in use or an outside force may suggest or initiate the shift. If, when searching in an on-line catalog (OPAC,) the scholar finds that the resource is not available in their affiliated library, the OPAC system may present an option to request the item through inter-library loan.

- Type IV - **Problematic situations might also direct users to alternative shifts**. Barriers and lack of results again cause a shift, but this time it is the user themselves who comes up with the alternative strategy, tactic or resource to continue their research. If the item identified in the library's on-line catalog is only available in libraries that are overseas and too far away to get through inter-library loan, the scholar may ask the university bookstore or library if it can be ordered through an alternate book-buying service, or they may contact a colleague or surrogate in the target country to acquire it for them.

For the purpose of my study, I make no distinction between Type III and Type IV shifts because the party that initiates the shift is not a significant factor. Therefore, types III and IV are combined to include any type of shift that might be caused by a problematic situation. Any factor that causes a problematic situation is referred to as a barrier. We are most specifically interested in barriers related to language, culture or geography.

3.4.2. Mapping shifts to the information triangle

One of the major advantages of the *Figure 3.2 - Sample Information Triangle* information triangle model is its simplicity. The triangle is composed of three sides, one for each of the three categories of information-seeking strategies. As strategies are gathered inductively throughout my study (or any study, since the model is very flexible,) the strategies are placed along the appropriate sides of the triangle according to how they fit the definitions

Formal System Strategies

Informal Resource Strategies

Searching a scholarly database

Tracing citations

Interactive Human Strategies

Contacting an author

outlined in Section 3.2, above. A strategy such as "using a web search engine" would be placed on the left side with the formal system strategies. Informal resource strategies such as "citation tracing" or "browsing in bookstores" would go on the right. Strategies where the entity is human, such as "attending conferences" or "consulting an archivist or librarian" are placed along the bottom of the triangle with the interactive human strategies. This sample is mapped to the information triangle in Figure 3.2, above.

Each information-seeking thread is broken down into its component steps. For example, a scholar might start by searching in a scholarly database for a useful article on their topic. Once they find the article, they notice that the references include one or more sources that appear equally useful. This discovery is a disruptive situation that the causes an opportunistic shift to citation tracing. They may then make another planned shift to return to the formal systems strategy of searching a scholarly database to obtain one or more of those articles. If they are

unable to find a particular item using a formal system because

they encounter a barrier that the source is only available in

overseas libraries, they may make an alternative shift to the

interactive human strategy of contacting the author directly via

e-mail or other communication method. Arrows are inserted to

show the shifts and their directions. Arrows with solid lines are

used to show planned shifts. Arrows with dotted lines show

opportunistic shifts. Arrows with dashed lines show assisted or

Figure 3.3 - Sample Legend

alternative shifts. Each triangle includes a legend as shown in Figure 3.3 above. Strategies are

numbered within each triangle to make it easier to identify the order of the strategies in an

information-seeking thread. Any combination of strategies and shifts between strategies can be

mapped to the information triangle this way.
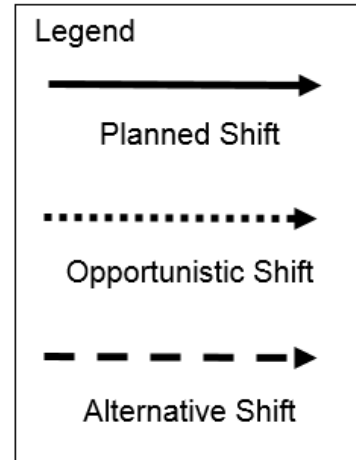
3.5. Barriers

The final concept to be discussed and defined is that of ***barriers***. A barrier can be simply

defined as any factor or element of a context that impedes research and therefore causes a

problematic situation. This concept is related to the third research question and how the role of

language, culture, and geography play, especially in problematic situations. In mainstream IR

and information-seeking literature, the term and the concept of barriers are very rare. In their

work on Cognitive Work Analysis, Pejtersen & Fidel (1998) use the term exactly once to talk

about student's inability to find information in a horticulture database because of a limitation of

the system that required search terms to be at least four characters long. Catalano (2013)

produced a list of six barriers encountered by students and the coping mechanisms that they

developed to overcome those. When encountering information overload, for example, students

might turn to strategies of citation chasing or focusing on subject specific databases or journals. When faced with time barriers, students show a preference for electronic materials, accepting convenience and ease of retrieval over the quality of the source, such as using Google Scholar for quick reference retrieval. The difficulty of constructing effective searches may be overcome by viewing tutorials, learning to use Boolean operators, or consulting with a librarian. When no time is available to learn new technologies, the coping mechanism is to avoid that particular technology. When students are unable to find relevant sources using the technologies at hand, they may turn to browsing in the physical library or consulting with another human such as an advisor or librarian. Finally, when keeping up to date presents a barrier, students may turn to e-mail alerts or browsing a table of contents (Catalano, 2013, p. 265).

A few very practical analyses have been made of barriers to information-seeking by diverse population, and barriers related to terminology are common. Lu, Lin, Chan, and Chen (2006) looked at how both language and terminology could prevent access to Chinese speaking lay users from retrieving the health information they needed. Between the language difference between Chinese and English and the lack of vocabulary knowledge of medical terminology, especially as related to MeSH terms, users are often unable to find information on such phenomena as Alzhiemer's disease since its popular name is Chinese is "dementia in aging people." Shoki and Oyelude (2006) looked at the language and cultural barriers experience by both information professionals and information seekers at public and university libraries in Nigeria. However, despite the term "language barrier" being a popular one in general, it is rarely discussed in relationship to information-seeking, information retrieval or information behavior.

The concept of barriers is inherent in Xie's (2000) discussion of shifts that users make in problematic situations. She draws the distinction between the shifts that users make when they

encounter a problem vs. the shifts they make when no problem exists. In most literature, however, shifts are seen as constructive moves that users naturally make to improve the effectiveness or efficiency of their information-seeking behaviors. Even though most researchers would readily admit that they exist, little is made of the barriers that cause the problematic situations that, in turn, lead to shifts.

In my own study of scholars, subjects often speak of the barriers caused by language, culture, and geography. When discussing language barriers, it is not the usual issue of comprehension, since my subjects already possess at least conversational knowledge and most often fluency or even native speaker knowledge of the language in question. Nonetheless, language creates problems for IR systems, especially when the language is not English. And even when IR systems are able to identify the existence of a source in another language, language, culture, and geography often throw up formidable barriers to the retrieval of the source, especially a primary source like a manuscript, newspaper, or government document from a foreign country. Further probing into the nature of language, cultural, and geographical barriers is warranted to provide insight in their causes, the shifts they cause, and the information-seeking strategies that prove most effective to overcome them.

3.6. Past theories and models

Numerous information scholars have built theories and conceptual models, contributing to the rich background of the field of information-seeking research, each of which has its specific strengths and limitations. In this section I will discuss models that have contributed to the basis on which my model is built or have properties that can be compared or contrasted to my own. Before delving into these models, however, we should understand what a model is and its purpose. Saracevic explains that

"The role of models is to depict the essential elements and relations of an object (system, process, entity, structure, idea…). A model is a choice of a representation, a given rendition, of a given object… There are many kinds of models; moreover, the same object can be depicted by several, often very different models. Thus, models themselves are a subject of examination and critique of how well they depict an object. A model represents an object in a certain way, but does not explain what goes on. A theory does that. In a sense, models are a 'weak' form of a theory, a possible prelude to a 'strong' theory. In these aspects lie both the strengths and weaknesses of models in general, and IR models in particular (Saracevic, 1997, p. 314)."

Wilson defines a model as "a framework for thinking about a problem [which] may evolve into a statement of the relationships among theoretical propositions (Wilson, 1999, p. 250)." The models and their approaches described here are selected based on several factors: 1) user focus, 2) focus on intentional research, 3) foundation for subsequent theories. Models explored herein in comparison to the information triangle include the following, as shown in Table 3.1.

Table 3.1 – Comparisons with prior theories

Theorist – Year – Model	Relationship to the Information Triangle
Taylor – 1968 – Levels of information need	Provides an early example of a user-oriented study. Defines four levels of information need, whereas the information triangle is only concerned with conscious and described information needs.
Belkin – 1977, 1978, 1980, 1993 – Anomalous states of knowledge, 1996 – Episode model of interaction with texts	Expands the concept of information retrieval to include information-seeking behavior. However, it considers humans as only generators and consumers of information and not as intermediaries.
Dervin – 1986 – Sense-making	Provides a methodology for framing research questions on human behavior including studies of diverse groups of users. The sense-making model is more suited for casual library users, and its focus on query formulation is narrower than the broader information-seeking focus of the information triangle.
Ellis – 1989 – Information-seeking characteristics	Ellis' six characteristics acknowledge the effectiveness of human sources as intermediaries. Also, many of the activities identified in his study have been identified as strategies in my pilot studies, perhaps because his study was of a similar population of mature researchers.
Bates – 1989 – Berry-picking	Focuses on strategies, many of which are similar to those discovered by my pilot studies, and her model takes into account manual sources in addition to formal systems.
Kuhlthau – 1991, 1999 – Information search process	Her focus on the information search process from a user perspective corresponds closely to my information-seeking

Theorist – Year – Model	Relationship to the Information Triangle
	threads. However, the information triangle is not concerned with the affective factors considered by Kuhlthau.
Marchionini – 1995 – Analytical vs. browsing strategies	Major influence in user-oriented information-seeking research. Makes the distinction between analytical vs. browsing strategies similar to the formal system strategies and informal resource strategies of the information triangle.
Wilson – 1999 – Information behavior, Information-seeking Context	Introduces the concept of intervening variables, similar to my concept of barriers, and recognizes the iterative nature of research.
Ingwerson and Jarvelin – 1992, 1996 – Cognitive model, Integrated IS&R (information-seeking and retrieval)	Recognizes the information acquisition is a central phenomenon, which is significant to my study because acquisition of sources is often the most problematic of scholars' information-seeking tasks.
Xie – 2008 – Planned situation IR model	Most significant basis for my study in that it presents the concepts of situations, shifts, and dimensions of information-seeking strategies.

Taylor (1968) derived his four **levels of information need** from working with reference librarians to understand how library users express their needs in the process of asking for help. The levels he proposes are: visceral (unconscious), conscious yet undefined and undescribed, formalized (rationally and formally stated), and compromised (expressed in a way that conforms to a system to which the request is presented.) Taylor's research also provided the basis and exemplar of user-oriented studies in which the user's thought processes are explored (Xie, 2008). My information triangle model is only concerned with conscious and described information needs since the users in my proposed study are all actively pursuing formal research.

Belkin (1977; 1982) expanded upon Taylor's (1968) theories as well as the work of Wersig (1979) and hypothesized that users develop "**anomalous states of knowledge**" when they encounter a situation in which they cannot resolve a question with the knowledge that they already have. In this approach, the cognitive role of the participants is primary, as is a "focus on the states of knowledge and problems, goals and intentions of the participants (Belkin, 1980, p. 133)." He studies how information is or can be transferred efficiently between the human generator and a human user through the intermediary of texts. He views the purpose of IR

systems to help people solve problems, supporting people's interactions with information, rather than to directly solve the problems posed to it (Belkin, 1996; Ellis, 1996). He describes the following IR situation, which is essentially the operation of a modern day search mechanism:

1. The user recognizes an information need and presents his/her request to the IR mechanism, which represents a collection of texts, and hopes that the IR mechanism will provide the texts needed to satisfy their information need.

2. The IR mechanism's task is to select and present texts that it judges best able to meet the user's information need based on the request as stated by the user.

3. The user then inspects the documents presented to him/her by the IR mechanism and uses his/her judgment to determine which documents are most relevant and whether they fulfill his/her information need completely, partially, or not at all.

At each end of this operation, both the generator's and the user's states of knowledge are impacted by their beliefs, intents, values, linguistic and pragmatic considerations, knowledge of the context of the information, and so on. On the user end, the user experiences an anomalous state of knowledge (ASK) and decides to use the IR system to resolve it. After going through the process outlined above, the user will decide whether the anomaly has been resolved, and if not, he/she may refine and re-present their request to the IR system, and the process becomes iterative. Thus, the user is part of the IR system itself (Belkin, 1980; Belkin, Oddy & Brooks, 1982; Belkin, 1993; Belkin, 1996).

Similar to Taylor, he likens the process of ascertaining the actual information need to a dialog like that used in a reference interview, calling a system that replaces the function of the human intermediary as "intelligent IR" (Belkin, 1996). He expands upon the concept of information retrieval to include information-seeking behavior as "an interactive process

characterized by people's interactions with texts" and that "the goal of IR systems is to support

the range of information-seeking behaviors (Belkin, 1993, p. 60)." While this expanded view is

a bit more appropriate for my study, it still views humans as only the generators or consumers of

information and not as a possible intermediary that helps people find information. The focus on

IR systems represents too limiting of a view.

Later, in 1996, Belkin proposed the **information-seeking episode** model of interaction

with texts, which could be described as an interaction with an information object (text) using

varying ISSs and hence requiring the support of different IR support techniques. This theory

supports the subsequent work of Yuan and Belkin (2010b) to assess the effectiveness of IR

support techniques for multiple ISSs, especially when used simultaneously.

Dervin's **sense-making approach** was first introduced in 1972, coming from the field of

communications and subsequently widely embraced by developers of IR systems (Dervin,

Foreman-Wernet & Lauterbach, 2003). Dervin's theory states that individuals will experience a

"cognitive gap" as they try to make sense of the data presented to them by their environment.

They engage in information-seeking and information-using when they are unable to otherwise

find their way through a particular situation until they have made sense of it. Each situation is,

therefore, unique and includes a plethora of details which create the context for the information

need. Dervin proposes a framework in which universals can be derived from the details because

they are ubiquitous to all sense-making situations. These universals are isolated by analyzing the

user's ability to "move through time" because of their unresolved information need (Dervin &

Dewdney, 1986, p. 507). The concepts of *situation, gap,* and *use* are key to the sense-making

model. The gaps can be expressed as the information needs or questions that the user has in the

time-space context (Xie, 2008). The use element is what the user does with the answer that they

find – "what the seeker/user hopes to do after crossing the bridge (Dervin & Dewdney, 1986, p. 507)." Therefore, information is not seen as a given commodity but rather depends on the user's context and use of it.

Similar to previous studies, Dervin proposes her sense-making approach in support of library professionals' support of users in their information-seeking. Dervin and Dewdney (1986) apply sense-making to the context of the reference interview to help librarians understand why the types of questions they ask help clarify user need. Since its inception as a communication and information-seeking model, sense-making has also been embraced as a methodology for framing research questions on human behavior (Dervin, 1992).

As in many of the early models, sense-making is highly involved in the stage of information-seeking where the subject is realizing their need, defining it, and expressing it in ways that let them act upon it. In my studies of scholars, these early stages of identifying an information need are of less interest. Scholars in my study – already possessing doctoral degrees and having written dissertations, articles, and possibly even books – have already developed research questions and have a firm grasp of the facts and knowledge that they are seeking. While Dervin's sense-making would apply well to a person casually entering a library looking for job-seeking materials or information for a school science project, my subjects have conducted prior research in their field, sometimes on the same topic, and are much more conscious of their precise needs and have extensive domain knowledge. Also, because scholars are often doing subject searches to explore and broaden their knowledge of a topic, the narrow focus of query formulation – query presentation – system response – result assessment presented by most early models is too limiting for my study.

Ellis' (1989) model of **information-seeking characteristics** provided the foundation upon which most subsequent studies of information-seeking strategies have been built and represents one of the earlier user studies. Ellis' study involved analysis of the information-seeking patterns of social scientists in order to derive a list of six characteristics. Ellis felt it was important to analyze the "perceptions of academic social scientists of their own information-seeking activities (Ellis, 1989, p. 172)." He interviewed social science scholars and used the information transcribed from the interviews to characterize their information-seeking "patterns." From those patterns, he inductively derived his six characteristics: starting, chaining, browsing, extracting, monitoring, and differentiating.

Since his work was done in the 1980s, it pre-dates web-based tools such as search engines, most OPACs, digital libraries, and on-line periodical databases. However, even at this early date, Ellis' work seeks to influence the design of information retrieval systems: "Six characteristics seemed sufficient to subsume the different generic features of the various patterns, and, at the same time provide the framework for a flexible model to underpin recommendations for information retrieval system design and evaluation (Ellis, 1989, p. 178)." This is helpful for the purpose of my dissertation because most other studies focus solely on IR in terms of users submitting queries to a computer system whereas Ellis acknowledges the effectiveness of human sources as intermediaries to help find information. Ellis' research brings out the use of informal contacts, the importance of which is frequently discussed by users as an information-seeking activity at the starting stage. Moreover, he states that "many of those interviewed said that their first step in finding information on a new or relatively unfamiliar topic would be to seek out people who knew something about the area and ask them for references to introductory works (Ellis, 1989, p. 180)." This more balanced and comprehensive view of information-seeking

provides a solid point of reference for my research. In fact, many of the "activities" identified by his study have been identified as strategies by my preliminary research of scholars' information-seeking strategies. Since Ellis' research also involves scholars as subjects, his characteristics are especially applicable to the more completely formulated information tasks of mature researchers.

Perhaps because Ellis and fellow researcher Marcia Bates operated in a time that was undergoing rapid change in the availability of electronic tools and resources for library research, their focus was on the strategies used involving manual sources used by scholars with an eye on how electronic system developers could make the electronic (sometimes CD-based but more often moving to an on-line internet environment) systems more closely approximate actual user behavior and habit. "We need to expand our understanding of these resources and the search techniques to use in them, and at the same time, expand our view of how the search interface should be designed to assist searchers with their new, complex tasks (Bates, 1989, p. 407)."

Bates proposed a model called **berry-picking** which, she argues, is closer to the actual behavior of information seekers than are the more traditional IR models of the time. The berry-picking model suggests that searchers reformulate their queries not only because of the need to find better sources and the lack of results they get from systems, but more often as a result of the new ideas and directions that result from each new piece of information they encounter in their research. This results in a "bit-at-a-time" retrieval pattern.

In her model, Bates proposes six information-seeking strategies: footnote chasing (backward chaining,) citation searching (forward chaining,) journal run, area scanning, subject searching in bibliographies and abstracting and indexing (A&I) service, and author searching. These strategies are admittedly very representative of the time in which the list was compiled; more recent lists would not likely make reference to subject and author searching separately, for

example. She explains that full-text was only in its infancy and most searching was being done with surrogates. She contrasts these with "real searches" in which the "search techniques change throughout, and the sources searched change in both form and content (Bates 1989, p 413)."

She posits that a researcher who has the widest variety of strategies at their disposal and can use them well will be the most effective researcher. It is clear that, in recent years, on-line database search tools have begun to incorporate many of the theories advanced by scholars such as Bates. For instance, at the time that she and Ellis published their studies, chaining forward was only possible by going to a separate publication, called a "citation index" that would list which more recent works had cited the title that the searcher was using. Recent systems have incorporated forward chaining in their interfaces as one of the options on the screen along with backward chaining, journal run, author search, and subject search. Some OPAC systems even include the "physical shelf view" that she promotes (Bates, 1989, p. 419).

Similar to Ellis, a strength of her research as it relates to my dissertation topic is that it takes into account manual sources in addition to formal systems whereas most recent studies have all but ignored the idea that scholars may still actually browse libraries, interact with colleagues, trace footnotes, etc. Note, for example, her resistance to the "tendency in information science to see browsing *in contrast* to directed searching, to see it as a casual, don't-know-what-I-want behavior that one engages in separately from 'regular' searching (Bates, 1989, p. 414)." Her belief that all information-seeking strategies are valid and intertwined provides a solid basis for my own research. However, her model is more rigid in that it provides a pre-set list of ISSs which applied more specifically to the time in which she worked. As new strategies come into being and others fall from favor with researchers, Bates' model would

require revision whereas the information triangle provides a flexible framework for inductively constructing a list of strategies as cited by subjects in a study.

Kuhlthau's (1991) research emphasizes the **information search process** from a user perspective rather than focusing on how queries best match the system's representation of text documents. Derived from an initial study in 1983, two subsequent longitudinal studies, and one wide-range study, Kuhlthau defines the ISP as "the user's constructive activity of finding meaning from information in order to extend his or her state of knowledge on a particular problem or topic (Kuhlthau, 1991, p. 361)." Similar to and building upon Dervin's sense-making approach, the user seeks to find meaning and form a personal point of view by combining information from new sources with what he/she already knows. The success of their journey to create meaning from information is evidenced by the creation of new knowledge that they share with others (Kuhlthau, 1991). This is particularly applicable to scholars, whose goal is most often is to create new knowledge in the form of a classroom presentation, conference paper, research article, or book.

Kuhlthau's model breaks the information search process into six stages, and at each of these she analyzes three different aspects: the affective (how the user feels), the cognitive, (what the user thinks) and the physical (the actions that happen). Also included in Kuhlthau's model are the tasks that would most likely move the user from his/her current stage to the subsequent stage. The stages are initiation, selection, exploration, formulation, collection, formulation, collection, and presentation. These are especially well-suited to undergraduate researchers, which were the population she studied.

Kuhlthau's framework of the information-seeking process is significant in that it brings together very specific stages and combines them with the three attributes of affective, cognitive,

and physical, creating a matrix of sorts upon which subsequent models, such as Saracevic's (1997) stratified model, could be built. Ellis (1992, 1996) also recognizes and contrasts the physical paradigm with that of the cognitive paradigm. He states that one can see these two paradigms as competing with each other but that the relationship between the two is actually much more complex. Ellis proposes that multi-paradigm research is necessary because all information-seeking has an inherent duality: it includes the people (authors, indexers, intermediaries and users) and also the things or artifacts (documents, document representations, abstracts, indexes, and databases.) One of the main tasks, therefore, of this multi-paradigm research is to simplify the cognitive aspects sufficiently as to make it work with the simpler characteristics of the IR system (Ellis, 1992, p. 60).

While not often cited as a model of major influence in user-oriented information-seeking, Marchionini's distinction between **analytical vs. browsing strategies** is significant to my research because it supports the distinction between formal systems strategies and informal browsing strategies. While his work is ostensibly in regard to electronic formats, he has much to say about the nature of strategies in general, including manual strategies.

> "We use a variety of gross strategies in information-seeking, including: consulting our own long-term memory; asking friends, colleagues or experts; consulting personal collections of books, periodicals and files; conducting empirical investigations; and applying formal systems. Formal systems include: libraries, research firms, government agencies, electronic networks, and the growing collection of information services that make up the information industry (Marchionini, 1995, p. 7)."

He further defines analytical strategies as well calculated and planned out in advance and dependent on specifying query terms, evaluating results and reformulating the query. Browsing strategies are, by contrast, opportunistic and require a more consistent cognitive alertness; they are more personal since the ability to identify resources that meet the individual's information need is crucial.

Numerous scholars have focused and expanded upon the analytical structures, including those that are not specifically search strategies: building blocks, (Harter, 1986), successive fractions (Meadow & Cochrane, 1981), pearl growing (Markey & Cochrane, 1981), and interactive scanning (Hawkins & Wagers, 1982). These strategies are concept-oriented, meaning that they seek to manipulate the concepts of the search topic. The ten strategies cited by Hawk and Wang (1999) are problem solving strategies related to web searching; they are: surveying, double-checking, exploring, link following, back and forward going, shortcut seeking, engine using, loyal engine using, engine seeking, and metasearching (Xie, 2008). Conversely, browsing strategies include scanning, observing, navigating, and monitoring (Marchionini, 1995).

To understand Marchionini's perspective, as well as that of his contemporaries, it is important to note that Liebscher and Marchionini (1988) were writing about electronic information-seeking during the very infancy of hypertext. At that time, hypertext was seen as a way to link "fragments of text" to facilitate browsing from one to another (Marchionini & Shneiderman, 1988), and this capability would not be available to users until 1992. This was in contrast with the command-line or "green screen" search tools of the day such as some OPACs and database tools like Gopher or Carl Uncover. It would not be until 1994 that web search engines would begin to appear to bring search capabilities to hypertext (Schwartz, 1998). From Marchionini's viewpoint as a computer scientist, it was easy to dichotomize these two concepts. Ellis (1996) also contributes to the discussion of hypertext, agreeing that it facilitates browsing but also that it can create disorientation and cognitive overload at least partially caused by multi-tasking and the user's need to keep track of a number of simultaneous tasks. This view would likely not be discussed among today's users who are so accustomed to both web surfing and multi-tasking that they are unlikely to complain of cognitive overload.

Wilson (1999) presents several models of **information behavior** over a range of years, primarily to show the relationship between information-seeking, information retrieval, and information behavior more generally. These models are connected together and become increasingly dependent on the preceding ones and progressively delve deeper into the intricacies of information-seeking and searching behaviors. Wilson's 1981 model diagrams 'information-seeking behavior' as a suggested alternative to the many 'information need' theories of the time. It begins when the user recognizes his/her own information need and then pursues the help of either formal or informal information sources or services in order to resolve that need. If the process is successful, the person will proceed to use the information; if it is unsuccessful, the user will pursue other resources, including other people, to again attempt to resolve his/her need. The resulting diagram is a variation on a programmer's Do-While loop, demonstrating the iterative nature of information-seeking (Wilson, 1999, p. 251). Wilson then proposed a second model of **information context** in 1981 which includes two important distinctions: 1) the information need that the user experiences is secondary to some other primary need, and 2) in their information-seeking process, they will encounter barriers to success. He characterizes the basic need as being either physiological, cognitive, or affective and states that the context of the need includes not only the person but also their social role and their environment which could be political, social, economic, technological, etc. The barriers, then, arise out of the same contextual elements: personal, role-related, or environmental (Wilson, 1999).

Wilson's 1996 model is based on the basic framework of his 1981 model. The concept of barriers is expanded to become *intervening variables*, which can include factors that either impede or facilitate the information-seeking process. The fact that Wilson uses the term and concept of barriers is significant to my research and relates to my own concept of barriers.

Ingwersen's **cognitive model** (1992, 1996) is one of the primary interactive IR models, which means that it goes beyond the traditional IR analysis of how well systems return results relevant to a user query. These interactive IR models became more common in the 1990s as information scientists began to seek to understand and explain why and how user information needs and behaviors change and how they affect the information-seeking process (Xie, 2008). Thus, these models are user-centric models and not IR algorithm-centric.

Central to Ingwersen's model is the *cognitive viewpoint*, which he most often refers to as "the view." Ingwersen states that the trends in research at the time were moving from one of access-orientation to one of accessibility and use of stored knowledge, embracing a more holistic human driven approach to knowledge transfer (Ingwersen, 1992, p. 13). The view, then, is more epistemological than practical, having its roots in Wersig's (1992) concept of *knowledge for action.* Nonetheless, the view and the model include practical implications for system design because the intermediary in the model can be either a human being or a machine, and therefore the system becomes one of the cognitive actors.

Given the dynamic, temporal nature of this model, it is understandable that any information-seeking task may have its own context based on the current needs, cognitive state, and environment of the various actors at that time. "The turn" toward a cognitive model, a concept popularized by Ingwersen and Jarvelin (2005) becomes possible when the existing body of research has built enough generalizable data – both empirical and analytic – to support more tailored investigations of cognitive and behavioral interactions between the actors as proposed in Ingwersen's cognitive model (Ingwersen, 1992). This foundation paved the way for **integrated IS&R (information-seeking and retrieval)** as proposed by Ingwersen and Jarvelin (2005).

Integrated IS&R recognizes that both information-seeking and information retrieval are subsets of information behavior as explained above (Wilson, 1999), and that information acquisition is a central phenomenon of information behavior and includes IS&R (Ingwersen & Jarvelin, 2005). One of the significant advances from Ingwersen's prior model is that, rather than the central actor always being an individual, that entity could also be a group (team) or a role that the user takes on, and as such, the other elements within the context also change (Xie, 2008, p. 185). Ingwersen and Jarvelin also recognize the influence of a great number of human actors within the model, including not only the information seekers who engage in various behaviors and bring various cognitive models to bear, but also the information authors, indexers, system designers, database designers, and publishers and aggregators who decide what information sources will be available to whom and at what cost. All of these other actors also bring to the situation their own cognitive world views, expertise, and expectations (Ingwersen & Jarvelin, 2005; Xie, 2008).

Xie's **planned situational IR model** combines many of the foundations of prior models, including 1) both macro- and micro-level interactive IR models, 2) user-oriented approaches, 3) a planned model from cognitive science and situated action, 4) the level of user goal/task and the relationships between levels of user goals and levels of tasks, 5) empirical interactive IR studies in varying digital environments, and 6) Xie's own research in interactive IR (Xie, 2008, p. 217). Xie's model goes into much more depth than previous models to characterize information-seeking strategies by associating them with the user's goals and tasks. The planned situational model takes into account a variety of factors:

- Levels of user goals and tasks and their relationships
- Dimension of work and search tasks

83

- Personal information infrastructure

- Social/organizational context

- Dimensions of information-seeking strategies

- Dimensions of plans and situations

- Shifts in search goals and tasks and the factors that cause them

The concepts of dimensions of information-seeking strategies, the dimensions of plans and situations, and the shifts in goals, tasks, and strategies are discussed in the definitions section, above. Goals and tasks are inseparable and make up significant components in interactive IR, especially in digital environments, which is Xie's primary context for her model. The planned situational model is suited to situations involving academic, personal, or work tasks. Xie breaks tasks into several dimensions: stages of the task, timeframe of the task, nature of the task, organization of the task, type of task, and domain of the task.

Xie derives much of her theory of the stages of tasks as they are explored and explained in part by Kuhlthau (1991) and Vakkari and colleagues (Vakkari, 2001; Pennanen & Vakkari, 2003; Vakkari, Pennanen, & Serola, 2003). The timeline of the task is explored in a workplace context by Xie (2006) as a significant factor when determining plans and tactics. Xie defines timelines as either extremely urgent (within a half hour,) urgent (within 24 hours,) or non-urgent (more than 24 hours,) showing that users must develop alternate plans and change ISSs in order to complete their information-seeking needs in extremely urgent and urgent situations. The definitions of timeline can easily vary in differing contexts. In my own research, where scholars conduct primarily long-term research projects, most tasks fall into the non-urgent timeline; however, scholars may also have more immediate (urgent) information needs when preparing for a class or meeting a publication deadline.

The nature of the task can be defined by any combination of three factors: its structuredness, its familiarity, and its situation. Byström and Järvelin (1995, p. 194) propose a classification of tasks based on its a priori determinability (structuredness) and its difficulty or complexity. Byström and Järvelin further concluded that task complexity had a direct correlation to the types of information, the information channels, and sources required to resolve the underlying information need (Xie, 2008). A second aspect of the nature of the task is the user's familiarity with the task, which could include routine tasks, typical tasks, and new tasks.

Finally, the situation of the tasks comprises the third aspect of the nature of the task. This can include the difference between a work task, a scholarly task, or a personal information need. The origination of the task indicates whether it is assigned by and external party such as an instructor or work supervisor, or whether it is self-generated. Numerous factors related to motivation, enjoyment of the task, etc. depend on the origin of the task. The type of task can include the classic types of known fact, known item, or subject searching. The domain of the task specifies the general area in which the task falls, which could include scholarly research, shopping/purchasing, medical research, or travel planning. Even the most casual user will quickly recognize that the domain context of the information need is highly influential in the choice of information-seeking strategies and tactics.

Another concept included in Xie's planned situational model is the personal information infrastructure, which includes the various types of knowledge and skills that the user possesses. As explained in the section on shifts, above, the components of the personal information infrastructure include the user's cognitive abilities, his/her domain knowledge, his/her system knowledge, his/her knowledge of information-seeking, and his/her overall general knowledge. A user who possesses superior knowledge in any of these areas is generally more successful,

especially in digital environments, than users with less skill and knowledge. For example, a user who understands the general workings of search engines is more likely to be successful because they will know that most on-line systems do not "understand concepts," meaning that, if they search for "butterfly," the search engine will not return texts that contain the words "cocoon," "moth," "insect," etc. unless they also contain the word "butterfly." Novice users might, for example, spend unnecessary time searching for words in alternate order; since most systems use a "bag of words" concept, the order of the words makes no difference in either the results returned or the order in which they are presented. Numerous studies cited by Xie have looked at searching knowledge, habits, and styles and their effect on the success of the searches (Xie, 2008).

Similar to Ingwersen's cognitive model as well as the field of Cognitive Work Analysis, Xie's planned-situational model includes consideration of the user's social-organizational context which defines the environment in which users interact with IR systems, primarily the work domain. This was explained above in the explanation of CWA. Interaction activities are the interactions between the actors and their social-organizational context. Since the context is not static, the dynamic interaction activities (including task activities, decision activities, and strategy activities) may vary. Culture is also one of the factors in the social-organizational context (Callahan, 2005) which is of major interest to my own research.

Of all of the models discussed above, Xie's model is the one that forms the broadest basis for my own research as it is the only one that includes the concepts of shifts and the dimensions of information-seeking strategies. However, Xie's research focuses almost entirely on digital environments, and the very premise of my own research is that users will often gravitate toward informal resource strategies or interactive human strategies because those strategies are less

susceptible to the barriers of language, culture, and geography. As such, Marchionini's model, which differentiates between analytical vs. browsing strategies, provides significant support for my own model of the information triangle. In addition, both Ellis' information-seeking characteristics and Kuhlthau's model of the information-seeking process help establish a series of steps providing a scaffolding framework upon which I will map shifts between information-seeking strategies from one side of the information triangle to another.

3.7. The need for a new model and a new study

As I undertake the study of scholars who rely on sources in languages other than English, none of the existing models is exactly appropriate. Most existing models were built to study the effectiveness of IR systems or how users utilize and react to IR systems. These do not represent a broad enough perspective for my study since many scholars utilize strategies and resources that do not at all rely on formal IR systems, but rather utilize informal strategies, manual resources, and interaction with humans. Similarly, most other models include a pre-set list of strategies, tactics, or moves that users engage in to accomplish their tasks. My model seeks to provide a more flexible framework that is unconstrained, allowing for inductive building of a list of information-seeking strategies unique to each study, each era, or each unique group of subjects.

Yet other models involve aspects of subjects' information-seeking processes that are not significant factors for my research. Since the focus of my study is how subjects identify and obtain sources for their research, it is mostly the physical aspects of the search – the actual actions that are undertaken – that are of interest. To a lesser degree, the cognitive aspects are of minor interests, seeking to understand the types of decisions that scholars make in selecting strategies. The affective aspects of a user's experience are not germane to my study.

For the type of study being undertaken in my dissertation, the information triangle has the following beneficial qualities:

- Correct focus: The model allows for the visual mapping of the strategies used in a single information thread and the shifts between them. It allows the researcher to build visual representations of what users actually do and then verbally describe the barriers that cause the shifts being mapped.

- Flexibility: the model allows for the inductive collection of strategies from the subjects of a study rather than relying on a pre-existing list of moves based on either prior research or information retrieval theory. Rather than presenting a static representation of what users usually do, it provides a blank slate on which to map the strategies that a specific user has actually pursued then abandoned in hopes of pursuing another more successful strategy, and so on. It can be used to map individual shifts on a more micro level or an entire set of shifts from the initial strategies through the last that leads to the fulfillment of the information need.

- Simplicity: Many prior models are complex, attempting to map numerous aspects of user behavior. This model focuses in on strategies and shifts to the omission of information about mental state, intentions, and affective factors. Those can be described in an accompanying verbal (written) narrative, if desired.

Likewise, there is also a strong need for the study itself. The pilot studies (Sabbar, 2013; Sabbar, 2016) have helped me to hone the appropriate methodologies, but those have been in anticipation of a full-scale, thorough, and rigorous study that will explore and draw a clear picture of the information-seeking practices of scholars who rely on non-English resources. The information gathered from this dissertation study will help validate the conceptual model and

provide understanding for system developers, content providers, information professionals, and other scholars in the field.

Chapter 4 – Methodology

4.1. Summary

The purpose of this dissertation study is to explore three research questions:

Q1: What information-seeking strategies are used by scholars conducting research in languages other than English?

Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?

Q3: In what ways do language, culture, and geography play a role in the information-seeking context, in particular, how do they create barriers that cause problematic situations? To do this, I conducted a qualitative study using human subjects, documented their information-seeking strategies, shifts and related factors. The gathered data was then analyzed to answer the research questions. This chapter describes the research methods used and explains why qualitative methods were best suited for this study.

4.2. Justification of qualitative methods

Before embarking upon a study in information science or any other field and selecting the most appropriate methodologies, a researcher must understand and evaluate the premises that, in all scientific inquiry, guide the choice of methodology. First, we must define *methodology* in contrast with *methods*. Bogdan and Taylor explain that:

> "The term *methodology* in a broad sense refers to the process, principles and procedures by which we approach problems and seek answers. In the social sciences the term applies to how one conducts research. As in everything we do, our assumptions, interests and goals greatly influence which methodological procedures we choose. When stripped to their essential, *most debates over methods are debates over assumptions and goals, over theory and perspective* (Bogdan & Taylor, 1975, p. 1)."

Over two decades later, Strauss and Corbin define methodology as "a way of thinking about and studying social reality (Strauss & Corbin, 1998, p. 3)." Bailey (2007, p. 63) defines methodology as a "term that refers to the larger research design that one follows when engaging in research, rather than just the specific methods used for collecting data." When gathering data for development of theory, any and all appropriate methodologies should be considered and chosen based on their suitability for the given situation. "The major criterion for selection of methodologies should be accuracy of description, and the data gathering methods chosen should present the most accurate description of phenomena to enable analysis and theory-building (Grover & Glazier, 1985, p. 250)." Moreover, methodologies should be chosen based on their appropriateness for a given stage of the research.

Qualitative *methods*, which can be defined as "a set of procedures and techniques for gathering and analyzing data (Strauss & Corbin, 1998, p. 3)," emanate from and are a result of the nature of qualitative methodology. The methods used in qualitative research collect descriptive data, using them to describe events that happen in the subjects' contexts. This is typically done using the participants' own words along with the observations of the researcher him/herself. Because we often equate verbal expression as a description of reality, verbatim records are very important to qualitative research. Participant observation is often used to place the researcher as close to the subject as possible and within the same context (Gorman & Clayton, 2005).

Denzin and Lincoln (1994) explain that, since qualitative methodology does not belong solely to one discipline, the specific methods used by qualitative researchers may well vary from one discipline to another. For example, researchers in literary studies may use textual analysis while researchers in anthropology utilize ethnographic methods such as participant observation,

and sociologists may rely on the analysis of personal documents or case studies. Some researchers come from disciplines (e.g., anthropology) or have philosophical orientations (e.g., phenomenology) that traditionally make use of qualitative methods. Yin (2009) explains that studies can be pursued for one or more purposes: exploratory, descriptive, or explanatory, and that any of the individual methods can be used for any of them. He explains that some methods correspond well to specific types of research questions such as "what," "why," or "how."

Several interrelated premises can help determine the choice of methodology and the subsequent methods. These can include their appropriateness for the type of study being conducted, but much has to do with the researchers themselves. The worldview of the researcher is key as well as his/her preferences, temperaments, and experience.

The first premise for choosing qualitative methods is their **suitability to the nature of the research problem and the research questions**. Qualitative methods can be used to explore substantive areas about which little is known or about which much is known to gain novel understandings (Stern, 1980). "Qualitative methods can be used to obtain the intricate details about phenomena such as feelings, thought processes, and emotions that are difficult to extract or learn about through more conventional research methods (Strauss & Corbin, p. 11)."

Generally speaking, qualitative methods allow researchers to document and understand behaviors that cannot be quantified about the people being studied (Berg, 2001, p. 7). More generally, Bailey describes field research as "well suited for… descriptive or exploratory research. Generating theory is another frequent motivation for engaging in field research. Evaluation researchers… are among those who often choose this method to generate data from their projects. Field research is sometimes used in the search for cause-and-effect relationships (Bailey, 2007, p. 4)."

In the area of information studies, where user experience is the yard stick by which nearly all successes and failures are measured, qualitative research methods are especially well suited. Sandler states that "qualitative research methods… are loosely defined as those techniques with contribute to the in-depth description and understanding of experiences and interactions occurring within libraries (Sandler, 1992, p. 174)." Qualitative methods can be especially effective for the study of information-seeking behavior by researchers since they focus on viewing the experiences of users and librarians from their own perspective and attempt to explain why participants in a library setting or engaging in an information-seeking task react the way they do (Mellon, 1986, p. 160). Similar to Kuhlthau's (1991) study of the information search process, my study explores and describes the processes used by scholars who rely on sources that are not in English, including primarily their actions and reactions throughout the process. The nature of my research questions is to understand the experiences of the researchers that are the subjects of my dissertation study. While this exploration of behavior is not as sensitive as many qualitative studies that examine the experiences of people in problematic situations such as illness, family distress, etc., the information-seeking behaviors in question lend themselves well to "getting out into the field and finding out what people are doing and thinking (Strauss & Corbin, 1998, p. 11)."

A second premise that guides the choice of methodology and subsequently the specific methods used is the **researcher's theoretical perspective**. Researchers who look at the population or phenomena being studied from a positivist perspective will seek to understand the causes of those phenomena from a more societal basis. Conversely, researchers who approach the field from a phenomenological or interpretivist perspective are concerned with understanding human behavior from the subject's own context and understanding (Bogdan & Taylor, 1975;

Gorman & Clayton, 2005; Miles & Huberman, 1994; Bailey, 2007). These phenomena are often complex, descriptive of society or human behavior, and not easily described in quantifiable terms (Liebscher, 1998; Bailey, 2007). While quantitative analysis views the context as a controllable entity and therefore rarely a variable to be studied, the qualitative researcher actively enters the context and seeks to collect and make sense of the data based on their own experience of the context (Gorman & Clayton, 2005). Qualitative researchers in the interpretivist mindset ask what is it that people do, how they do it, what purposes and goals are served by their behaviors, how those behaviors influence and are influenced by their social world, and what these actions and reactions mean to the subjects. "The goals of field research for scholars who use an interpretive paradigm involve empathetic understanding of participants' day-to-day experiences and an increased awareness of the multiple meanings given to the routine and problematic events by those in the setting (Bailey, 2007, p. 53)."

Because the context and the world in which the subject exists are so crucial to the understanding sought by qualitative researchers, the perspective is also a holistic one. Each subject is not merely an isolated case but rather viewed as a contributing part of a whole (Bogdan & Taylor, 1975.) Likewise, each aspect of the situation (the variables) are all deemed worthy of interpretation and analysis as part of the individual's world whereas quantitative methodologies would seek to control many of those variables to study only isolated pre-chosen factors. In the study of information-seeking and information retrieval, the social and organizational context of the user is seen as a major factor in determining information behavior (Xie, 2007; Xie, 2008), so the qualitative approach of seeing the context as a whole is very useful. Indeed, an individual's group membership influences that person's values, cognitive framework, and communication patterns, and qualitative methods provide a way to study the behaviors of groups by studying

those of selected, individual members of that group (Grover & Glazier, 1985, p. 249). "Society is to be understood in terms of the individuals making it up, and individuals are to be understood in terms of the societies of which they are members (Meltzer, Petras & Reynolds, 1975, p. 2)."

Meanwhile, qualitative researchers are also realists, believing that the social phenomena that they observe exist not only in the minds of their subjects but also in the objective world and that "lawful and reasonably stable relationships are to be found (Miles & Huberman, 1994, p. 4)." As such, they derive constructs that underlie individual and social life, and they embrace the epistemological position that what they can learn of the social world exists independently of the researcher him/herself (Bailey, 2007). And they are "naturalists," conducting intense and/or prolonged research in the field and observing normal situations that reflect the everyday lives of individuals, groups, cultures, or entire societies. Thus, the attempt is to gather data on the perceptions of their subjects through close attention and empathetic understanding. They believe that there is a real and natural world that exists and can be documented (Wolcott, 1982; Miles and Huberman, 1994, pp. 5-7; Esterberg, 2002, pp. 13-14; Donmoyer, 2001, pp. 176-177). This process is, by nature, more time-consuming than quantitative research; it requires greater clarity of goals during the research design process; and the data gathered cannot be analyzed automatically or calculated by a computer program (Berg, 2001).

The positivist paradigm begins with theory and uses deductive reasoning to develop a testable hypothesis from the theory. Data is then collected and quantified to see if they support the hypothesis empirically or not. Positivist researchers favor quantitative methods that allow them to collect data and measure the phenomena using questionnaires, inventories, and other demographic data (Bailey, 2007, p. 51). In contrast, phenomenological (interpretivist) researchers favor qualitative methods such as participant observation, interviewing, or diaries to

95

understand the phenomena in the subjects' own words and actions (Bogdan & Taylor, 1975).

Qualitative researchers believe that the meaning of events, interactions, and other phenomena can best be understood through the eyes of the actual participants. The delicate balancing act exists: being able to observe and document the phenomena without influencing it by one's own presence (Gorman & Clayton, 2005).

Another common division between quantitative and qualitative methodologies is that they commonly support deductive and inductive research, respectively. The quantitative researcher often collects data to deductively support or prove one or more hypotheses that he/she has already postulated. By contrast, the qualitative researcher collects data to describe and understand phenomena, letting theories arise inductively as data is collected. This is often referred to as "grounded theory" because the theory is built "from the ground up" based on the data that is collected (Gorman & Clayton, 2005). Grounded theory is based largely on the works of Glaser (1965) and focuses on the process of generating theory through the process of research (Patton, 2002; Glaser & Strauss, 1967). "Generating a theory from data means that most hypotheses and concepts not only come from the data but are systematically worked out in relation to the data during the course of the research (Glaser and Strauss, 1967, p. 5-6)." Grounded theory links theory to method, believing that how you study the world determines what you learn about the world. Because the researcher is closely entrenched in the real world, their observations and results are grounded in the empirical world – the world of on-going group life (Patton, 2002, p. 125). Grounded theory can often be conflated with inductive methodology in general, and it is nearly always tied to qualitative methods. However, Glaser (2000) asserts that grounded theory is a general method and can be use with any type of data. He also warns

that not all qualitative research actually utilizes grounded theory and should not be thought to do

so (Patton, 2002; Glaser, 2000, p. 7).

A third premise in the selection of methodology and specific methods is that of the

temperament, preferences and/or experience of the researchers. Some people are more

oriented and temperamentally suited to doing this type of work. Glesne & Peshkin (1992, pp. 2-

3) recount "Alan's Story" about choosing qualitative methods:

> "In terms of doing research, I knew that I was not at home in the world of
> numbers long before I realized that I was at home in the world of words. It was
> less of a case of rejecting numbers than feeling rejected by them. They made me
> feel like an outsider; they created walls that I could not penetrate. Thus,
> becoming a qualitative researcher did not begin with a principled, anticipated
> rejection of one paradigm in knowing acceptance of another. I began by
> understanding what I could not comfortably do and gradually learned where, in
> fact, I could be comfortable."

My own attraction toward qualitative methods stems largely from my natural tendency toward

inquiry as well as direct communication and interaction with people.

All interviewing requires a rapport between the interviewer and the interviewee;

however, the interviewer should maintain neutrality in the interview process (Bogdan & Taylor,

1975). That is, the words spoken by the interviewee should not change the interviewer's

perception or evaluation of the interviewee as a person. "Rapport is built on the ability to convey

empathy and understanding without judgment (Patton, 2002, p. 366)." A good interviewer is

alert to establish and promote rapport by demonstrating interest in the subject's responses,

listening carefully, and being genuinely curious and open. Other interviewer attributes that can

influence the success of the interview process are: 1) the ability to anticipate and prepare for the

interview process; 2) a certain naïveté that places the researcher in a true learner role; 3) an

analytic nature to listen for interesting comments and observations that could lead to follow-up

questions; 4) the ability to be both in control of the interview and yet simultaneously submissive;

97

5) a "nonreactive, nondirective, and therapeutic" stance similar to neutrality as explained above;

6) patience and the ability to probe gently to draw out more in depth information, turning pro forma answers into quality data (Glesne, 1992, pp. 78-87).

Much as Jane Goodall eventually interacted with and eventually influenced the apes at Gombe despite not having planned to, no researcher is truly isolated from their subjects. In fact, "participating in relationships with members of the setting provides the basis for the interpretive process considered so central to field research. These relationships supply the foundation for what field researchers come to know in the setting (Bailey, 2007, p. 73)." So the ability to build rapport effectively must be part of the personality and being of the researcher. "Honesty, openness, friendliness, and a willingness to get along are usually the best qualities to exhibit when you first undertake contact with participants in the field, and with time most people will respond positively to genuine concern and interest in them (Neuman, 1991, p. 349)." Some of the researcher's characteristics that seem the most superficial and unimportant can have unexpected or even undetected effects. These might include: physical appearance, punctuality, politeness, physical health, level of expertise, outgoing nature, level of calm, cheerful demeanor, professionalism, academic credentials, tone of voice, etc. These, along with instrument design, can work together to determine what level of participation, insight, honesty, and even candidness that the researcher can elicit from his/her subjects. For example, when gathering my data on how scholars using other languages go about identifying and accessing their sources, the actual topic of their research is not highly significant. Also, in the data I gather, my own anecdotes and observations about language are not worth transcribing in the end. However, both serve a purpose as illustrated by an interview conducted during one of my pilot studies where a participant was wary of my questions at first. This was a telephone interview, and the subject

98

and I did not know each other nor had we met. He asked up front what I was going to do with the data, and he seemed very hesitant and reserved, as if I were somehow judging his answers to be right or wrong. But as he spoke about his research, and I asked him a few specific questions about his topic in response to his comments, I could sense a change in his demeanor. He spoke with increasing enthusiasm and assertiveness and eventually felt very much more at ease. In listening to the recorded interview later during transcription, I could only conclude that the "useless" part of the conversation had permitted him to feel like an expert in his field and to have confidence that I was interested in his *perceptions and opinions*, and that this was not a test of his knowledge of research methods or library jargon. The ability to inspire confidence and openness, then, is a key trait that matches researchers with appropriate qualitative methods. In short, the methods not only fit the research; they fit the researcher as well, and the researcher fits the methodology.

Qualitative researchers need the attributes of appropriateness, authenticity, credibility, intuitiveness, receptivity, reciprocity, and sensitivity (Corbin & Strauss, 1998, p. 6). Of course, these researcher attributes are not automatic, nor are they completely innate. They develop with practice, and they will not develop if the researcher focuses solely on the procedures and applies them in a mechanistic manner. The methods should serve as a means to bring their vision into reality (Strauss & Corbin, 1998).

In information science, both quantitative and qualitative methods are used frequently. Selecting the appropriate methodology for an individual study can rely on some very basic factors and questions:

- Which method best corresponds to or supports the hypothesis or research question being tested or explored?

- In the existing literature, what methods have been applied to similar types of problems or questions?

- Which methods will result in the strongest statement of results at the conclusion of the research?

- Which is more important: controlling independent and extraneous variables to make highly replicable and generalizable results or explaining why a certain phenomenon or condition exists (Losee & Worley, 1993)?

More recently, the divide between components of quantitative and qualitative research is softening, and that the concept of mixed methods has come into much wider acceptance to the point of being trendy. Creswell explains that "the situation today is less quantitative *versus* qualitative and more how research practices lie somewhere on a continuum between the two. The best that can be said is that studies *tend* to be more quantitative or qualitative in nature (Creswell, 2007, p. 4)." Creswell also explains that the "three elements of inquiry (i.e., knowledge claims, strategies, and methods) combine to form different approaches to research. These approaches, in turn, are translated into processes in the design of research (Creswell, 2007, p. 5)." Experienced researchers often recognize that there can be a need for both types of data, and that a mix of methods and data types can complement each other and help overcome the weaknesses inherent in each (Gorman & Clayton, 2005).

Not all qualitative theorists agree, maintaining that it is preferable to preserve the purity of methods. Guba and Lincoln (1988) argue that each approach has its own internal consistency and logic which should discourage the mixing of differing inquiry modes or data collection strategies. Ultimately, it is the researcher's choice whether to preserve methodological purity or pursue the gathering of the most relevant possible information by combining methods creatively

(Patton, 2002, p. 252). In my study, only qualitative methods are used. This will be discussed in greater length in the results and discussion sections.

4.3. Sampling

The study was conducted using human subjects, all of whom are humanities scholars employed at colleges or universities in the U.S. with doctoral degrees actively pursuing research on a topic in the humanities and who rely on resources in languages other than English. Humanities scholars were chosen for this study because their subject areas (mostly history, language, and religion) more often rely on primary sources and sources in diverse languages than scholars in fields such as hard sciences, education, or social sciences. Scholars in the humanities – especially historians – rely on a wide variety of document types over diverse time periods, many of which are less likely to exist on-line than more current documents. A total of 40 subjects participated in the study to completion, identified and recruited using purposeful sampling. Table 4.1, below, lists the participants, and Tables 4.2 and 4.3 provide a summary of languages and disciplines represented in the study, respectively.

Purposeful sampling, defined as the selection of a sample on the basis of its contribution of information rich cases for in-depth study (Patton, 2002,) was used for this study, choosing "representatives from within the population being studied who have a range of characteristics relevant to the research project (Gorman & Clayton, 2005, p. 128)." Since qualitative research focuses on depth and not breadth, it relies on a smaller sample, where cases are chosen for the express reason that they are believed to represent the larger population being studied. Subjects, groups, etc. are chosen because they represent *information-rich cases* for in-depth study, from which researchers hope to learn a great deal about the issues under evaluation. "Studying information-rich cases yields insights and in-depth understanding rather than empirical

generalizations (Patton, 2002, p. 230)." Also called theoretical sampling by Glaser and Strauss (1970), purposeful sampling aids in the discovery of categories, defining their properties, and assessing their inter-relationships. It is important to remember that one of the goals of qualitative research – as well as one of the measures of its validity and success – is the ability to effect change. Therefore, if the researcher hopes to improve services to scholars conducting research using sources in languages other than English through his/her research, it is more important to understand the challenges faced by a modest number of diverse subjects and the ways that they have successfully overcome language barriers than it is to know what percentage of researchers experience such-and-such an issue or use specific strategies. To find appropriate subjects, "you decide the purpose you want informants (or communities) to serve, and you go out to find some (Bernard, 2000, p. 176)."

Purposeful sampling can be accomplished using any one or more specific strategies to identify and select information-rich cases (Patton, 2002, pp. 230-244; Bailey, 2007, p. 65). In my dissertation study, participants were selected based on their ability to provide in-depth insight into scholarly information-seeking strategies when using non-English sources. The qualifications for inclusion in the study were fairly significant and included:

- Competence in at least one language other than English and the ability to use that language for research purposes

- Status as a humanities scholar - faculty member at a college or university with a doctoral degree in their field

- Active pursuit of research during the study period using non-English resources

- Residence and home base in the United States

102

Table 4.1 – Participants

Subject	Discipline	Research Language(s)	Native Language	Gender
S2	History	Japanese	English	F
S3	Modern Languages	German	German	M
S4	History	German, French, Latin	English	M
S6	History	French, Latin, Italian, Catalan	English	F
S7	Modern Languages	Spanish	Spanish	F
S8	History	Chinese	English	M
S10	History	German	English	F
S11	Modern Languages	French	English	F
S13	Modern Languages	German	English	M
S14	Modern Languages	French	English	M
S15	History	Portuguese (Brazilian)	English	F
S16	Modern Languages	German	English	M
S17	Dance History	Persian, Croat, Russian	English	M
S18	Art History	Japanese	English	F
S19	Classics	Ancient Greek, Latin	English	F
S20	Classics	Latin, Ancient Greek	English	M
S23	Classics	Syriac, Coptic, Ethiopic, Greek, Arabic	English	M
S24	Modern Languages	Italian	English	F
S25	Modern Languages	Spanish	English	M
S29	History	Irish	English	M
S30	History	Xhosa, Swahili	English	F
S31	History	Czech, Slovak	English	F
S32	History	French	Hebrew	M
S33	History	Japanese	English	F
S34	Modern Languages	Spanish	English	F
S35	History	Ottoman Turkish	English	F
S36	History	Russian	English	F
S37	Religion	Persian, Arabic	Persian	M
S38	History	Arabic, French	English	M
S39	History	Punjabi, Urdu, Persian	English	M
S40	Modern Languages	Norwegian	English	F
S41	History	Chinese	English	M
S42	Modern Languages	French	English	F
S43	Information Studies	Afrikaans, German, French, Dutch	Afrikaans	M
S44	History	French, Vietnamese	English	M
S45	History	Russian	English	M
S46	History	Dutch	English	F

S47	History	Chinese	English	M
S48	History	Spanish	English	F
S49	Archaeology	Egyptian, Coptic, German, French, Arabic	English	F

The number and variety of the subjects chosen contributes to the richness of the data. In many studies, the researcher seeks to study a very homogenous group in which each subject is considered as anonymously and as uniformly as possible. In this study, as explained in the theories above, I specifically sought a variety of subjects from different disciplines within the humanities that represent as wide a variety of languages and cultures as possible. This is of significant importance since pilot studies have indicated that the further the language is from English - linguistically, culturally, and geographically – the more barriers will be encountered by the scholar. To conduct a study that values, embraces, and leverages diverse viewpoints and perspectives, it is important to note that it matters quite a bit who the participants are. Rather than attempting to preserve anonymity or pursue uniformity, this studies relies on a group of subject that represent a diversity of languages, cultures and fields. Lists of research languages, native languages, and disciplines are given below.

Table 4.2 – Languages represented with frequency

Language	No. of subjects using	No. of native speakers
Afrikaans	1	1
Ancient Greek	1	
Arabic	3	
Catalan	1	
Chinese	3	
Coptic	2	
Croat	1	
Czech	1	
Dutch	1	
Egyptian	1	
English		35
French	9	
German	6	1
Hebrew		1

Language	No. of subjects using	No. of native speakers
Irish	1	
Italian	2	
Japanese	3	
Latin	4	
Norwegian	1	
Persian	3	1
Portuguese (Brazilian)	1	
Punjabi	1	
Russian	3	
Slovak	1	
Spanish	4	1
Swahili	1	
Syriac	1	
Turkish (Ottoman)	1	
Urdu	1	
Vietnamese	1	
Xhosa	1	

Table 4.3 - Disciplines represented with frequency

Disciplines	No. of participants
Archaeology	1
Art history	1
Classics	3
Dance history	1
History	21
Information studies	1
Modern languages	11
Religion	1

Recruiting of subjects began in July 2015 and concluded in December 2015. The majority of subjects were recruited via email, a sample of which is shown in Appendix A. I had anticipated that would be able to re-connect with many of the humanities scholars from prior pilot studies as well as other scholars already known to me, but in fact, only four of the participants are from previous studies, and only those four participants were known to me previously. Moreover, while I had anticipated that snowball sampling would be useful – learning

about one participant from another – this did not happen, and all of the participants in the study were identified by me from their college or university web sites or through listservs. From my prior experience with my pilot projects, I learned that recruiting participants from major, well known public universities and Ivy League institutions is very, very difficult and not worth the effort. Since I work for a private liberal arts college, I began by contacting history, language, and religion faculty from similar liberal arts institutions by email, identifying and recruiting 14 participants in this manner. Contacting faculty at a few Wisconsin public universities resulted in one additional participant. A humanities listserv/bulletin board post recruited three subjects. I then endeavored to contact scholars from larger or higher profile liberal arts colleges, identifying four more participants. At this point, I had identified primarily participants in more common languages, especially German, Spanish, French, and classical languages. In order to represent a wider breadth of languages, I began to search for people who use specific languages such as Irish, Swahili, Arabic, Hindi, Russian, and other Germanic and Asian languages. In several cases, I placed phone calls to target faculty mostly from nearby and liberal arts institutions, leaving voicemail as needed. This proved to be surprisingly effective, giving a much higher rate of acceptance than email. The total number of subjects recruited specifically for the language they use is 13, of which about nine were recruited by phone.

It is relevant to note that one of the requirements was that the subjects be actively involved in research during the study period, and specifically in the information-seeking stage. Some prospective subjects were interested in the topic but were either not actively involved in research or were at the data analysis or writing stage of their project. Of those that participated, two (Subjects S10 and S15) actively engaged in archival research in their target country between the initial and final interview. The others were limited to research activities that they could

conduct from or near their home base in the U.S. This may account for a skew in the data recorded in the diaries that includes less actual on-site, overseas archival research and more of other strategies conducive to use in the U.S.

The 40 participants recruited exceeded my goal of 30 and exceeds the number of cases traditionally recognized as needed for saturation in qualitative research. In qualitative research, data saturation is said to be reached "when there is enough information to replicate the study, then the ability to obtain additional new information has been attained, and when further coding is no longer feasible (Fusch and Ness, 2015, p. 1408)." In determining the minimum number of cases for a valid study, Lincoln and Guba (1985) recommend a satisficing strategy, adding cases until no new observations are being made with the additions of new subjects and cases. Likewise, Bailey proposes a rule of thumb for how many people to interview: Start with 20 and add "until you have at least five new cases that fail to add anything new to the analysis (Bailey, 2007, p. 64)." What I found, however, was that the diversity of the participants that I had sought also made saturation virtually impossible to achieve. The more languages and cultures that were added to the study, the more unique phenomenon were brought up. For example, participant S6, who studies religious heretics in Southern France found that the calendar used in that region in the year 1315 began on or around March 25, so the documents she was searching for in the local archive we misclassified by one year due to the dating discrepancy. One would not expect to hear of this same language or cultural barrier from any other participant no matter how many were recruited unless perhaps the study were limited to scholars of medieval Western European history.

However, if the phenomena are grouped broadly into categories such as "cultural differences over time due to changes in government, colonization, invasion, etc.," the calendar

phenomenon could be grouped with barriers encountered by scholars studying Algeria, Punjab, Croatia and other regions that have experienced regime change over the centuries. This view is consistent with Patton (2002, p. 246), who states that "the solution is judgement and negotiation. I recommend that qualitative sampling designs specify minimum samples base on expected reasonable coverage of the phenomenon given the purpose of the study and stakeholder interests. One may add to the sample as fieldwork unfolds." To illustrate this challenge, it is worth relating that, after I had finished gathering data from the 39 subjects that I had initially identified and engaged, and I was struggling with the concept of saturation and how I would justify my sample size despite each case revealing new phenomena, I came in contact with a scholar of Egyptology who was willing to participate in my study. Her participation would not help achieve saturation; on the contrary, it would most likely contribute and describe several new phenomena related to the pictographic nature of the Egyptian language in use during the time period she studies. Suffice it to say that to add this new data to the sample as the fieldwork unfolded is consistent with the purpose of the study – to broadly describe the information-seeking strategies used and shifts made by these scholars and to explore the role played by language, culture, and geography in the information-seeking process. Especially in regard to research question 3, the more languages, countries, and cultures are added to the study, the richer the data will be, and the further the hope of saturation will slip from our grasp.

4.4. Data collection methods and procedures

Data were collected from the subjects using two primary qualitative methods: semi-structured interviews and diaries. Each subject was asked to participate in four activities in the following order:

1) An initial semi-structured interview about their research, the sources they use, and their usual information-seeking strategies,

2) A questionnaire that included some demographic information as well as specifics related to the sources, languages, and locations used in their studies,

3) A diary to record their research activities during the period of the study, and

4) A final semi-structured interview in which the subjects were asked questions about the research conducted during the research period, their insights into which strategies were most successful, and the role of language, culture, or geography in their research context.

The goal of the initial semi-structured interviews was to learn about the subject's field of study and their usual and planned research processes and strategies. At or around the time of the interview, each subject was given the structured diary form, and its use was explained to him/her. Each subject was asked to fill out the diary form for two or three discrete information-seeking tasks during their study period. Only a modest number of participants – primarily those interviewed in the summer and in December – were able to complete the task in the requested time frame; most took considerably longer to complete the process. Two of the scholars conducted research in their target country during the period of the study, which gave special insights into archives usage. Regardless of the variations in their timeline, the process allowed for the subjects to conduct their research at their point of need, and it allowed time for both gathering and analysis of the data.

Interviewing was the primary methodology used in my study, chosen for its effectiveness in gathering data on the subjects' perceptions and thoughts over and above their behavior. Interviews were conducted purposefully to obtain specific information about their research and to explore, describe, and explain their behaviors (Olszewski, Macey & Lindstrom,

2006). In recent years, with the proliferation of interviews in the media on topics from weather

to Hollywood scandals, interviewing as a way to gather reliable data is viewed with skepticism,

especially in comparison to direct observation.

> "The issue is not whether observational data are more desirable, valid, or
> meaningful than self-report data. The fact is that we cannot observe everything.
> We cannot observe feelings, thoughts, and intentions. We cannot observe
> behaviors that took place at some previous point in time. We cannot observe
> situations that preclude the presence of an observer. We cannot observe how
> people have organized the world and the meanings they attach to what goes on in
> the world. We have to ask people questions about those things (Patton, 2002, p.
> 341)."

Of the three main types of interviews (*unstructured, structured,* and *semi-structured,*) my

study utilized *semi-structured* interviews, which use an *interview guide* to lend some structure to

the interview, but preserve flexibility to add my own questions, make up my own probes, skip

questions, or change the order, based on the flow of the conversation or any of a number of other

factors. The semi-structured format allows for the discussion of topics that emerge that are of

interest or importance to the subject. Moreover, it allows me freedom to add my own thoughts

and comments (Bailey, 2007; Patton, 2002). Semi-structured interviews have many of the same

benefits as structured interviews, such as keeping the process focused so that efficiency is

maximized. The data is also more easily analyzed than in unstructured interviews since each

interview follows largely the same configuration.

My study explored the subjects' perspectives, experiences, and behaviors, so it was

crucial that the questions be open-ended, permitting answers to be expressed in the subjects' own

words and allowing them to tell me their stories. Multiple choice and Likert-scale questions do

not promote such responses, and questions that correspond to these formats - "How difficult is it

to get into libraries in Japan?" will lead the respondent to a narrow array of answers. True open

ended questions might include "Tell me about your experiences at libraries in Japan." Care must

also be taken with "why" questions. While understanding the "why" is one of the key goals of qualitative research, asking a direct "why" question is most often not the best way to discover the level of causality that the researcher is trying to ascertain. Instead, the researcher may want to ask a more specific question. For example, rather than asking "Why do you use primary sources in your research?" the researcher may want to ask "What aspect of primary sources make them preferable for your research (Patton, 2002, pp. 363-365)?"

My semi-structured interviews were organized in several loose sections. The first few questions engaged the subject in open-ended discussion of his/her topic. These allowed me to understand their topic, and more importantly, they put the subject at ease, helped create a rapport between myself and the subject, and encouraged him/her to speak openly, freely, and with confidence, which was important as I moved on to the next section. The next question asked what types of resources they utilize, followed by two or three questions about HOW they usually identify and obtain resources. These questions were intended to establish a baseline of planned strategies used in routine situations. The remaining five or six questions asked the subjects to relate examples of specific situations where they found sources, experienced barriers, overcame obstacles, and/or specifically used primary sources. These questions were intended to encourage the subjects to tell stories, from which I could gather examples of strategies used in disruptive or problematic situations. Appendix B shows the questions used for the initial interviews.

Over time as my research progressed, in order to explore emerging concepts discovered by grounded theory and the constant comparative method (Glaser, 1965; Glaser, 1967), some especially helpful and rewarding questions emerged that were used when applicable, including (in no particular order):

- How do you know what archives or libraries contain the sources that you need?

111

- In what situations did you pursue specific language skills in order to pursue your research?

- When and in what situations do you find it beneficial or necessary to travel to the target country to get the sources you need?

- What are some of the interesting stories that you have related to accessing and using archives and libraries in your target country, culture, or language?

- Do you use the Google search engine (or other search engine) of your target language or Google.com or both? Do they result in finding the same sources or different ones?

- What challenges do you encounter finding sources for your topic related to how old or how recent the sources are?

- In what ways do you search differently for sources for your own research as compared to sources for your students to use in their coursework?

Shortly following the initial interview, each subject was asked to fill out an on-line **questionnaire** that includes demographic and source usage questions. While questionnaires are not common among qualitative methods, this is a simple and effective way to gather very specific information about what languages they use; what libraries, archives, institutions, or other locations they visit; what specific sources they rely on; and so on. Since the subjects of my study speak a variety of languages, gleaning names of libraries, cities, authors, subjects and sources verbally from interviews and getting the words and spelling correct can be challenging. At the same time, the questionnaire, created and distributed as a Google Form, provided minimal demographic information, including their native language and past education. A sample is shown in Appendix C. Demographic data such as age and gender were also collected.

The secondary qualitative method that was used is a **research diary** in which the subjects recorded their process, the steps they took, and so on, for at least two information-seeking threads that they pursued. They were asked to record not only their actions but also their thoughts, the barriers they encountered, and their changes in direction and/or strategy. When filled out attentively and conscientiously, a diary can provide an accurate and useful log that documents the research process. The diary is used in place of direct observation in information science research because direct observation would be problematic on several levels. First, observing research would be highly time consuming for the researcher since many scholars may work for hours at a time while making only infrequent choices and shifts that are of interest to my study. Second, as explained earlier, much of the salient activity that happens during research occurs silently in the minds, feelings, attitudes, and intentions of the subjects. Such diaries have been used successfully in studies of information-seeking by Xie (2006) and Kapusniak (2013).

Utilizing a structured diary allows the subject to record the behaviors of interest at the time they occur and to classify and comment on them as they happen. Admittedly, there are some risks involved in relying on self-reporting, but these are deemed minimal in comparison with the overall effectiveness of the diary technique.

During two prior pilot studies, I utilized a diary format. The first diary was less structured and designed to allow for many different types of notes over a period of approximately two months. The primary issue encountered with that diary form was the high amount of variability between subjects in filling out the diary. Some wrote very general descriptions of a whole research session that lasted one to four hours, while other subjects recorded very granular moves as they occurred. Secondarily, a few of the subjects found it to be too much effort and did not complete or use the diary. The diary format used in the second pilot

study was more structured and designed to document more closely the individual strategies chosen, the barriers encountered, and the shifts made within a single information task. In the second study, I asked each subject to fill it out for only one information task, and those who completed the diary did so with better uniformity and adherence to the intent of the instrument. Moreover, the data from the structured diary was extremely easy to code and analyze since some portions of the diary use checkbox options that correspond to the types of shifts that are of interest to the study. When a more structured diary is used and closely followed by interviews of the subjects, it can provide a very helpful mix of uniform, quantifiable data and anecdotal, descriptive narrative that explains it.

The diary format used for my dissertation study is very similar to that used in the second pilot with only some terminology changes for clarification. The header portion asks the subjects to record the information task they undertook and the general context of that research. The primary portion of the form, in which the participant's steps are recorded, is in a table format. The first row of the form itself asks them to record the first strategy and first resource they used in pursuing their information task. The next row gathers information about the situation and the shift that led to the next strategy. The third row is for the second strategy, and so on. The subject was asked to fill out as many rows as necessary to record their various iterative steps until they either complete or abandon their information quest. Appendix D shows the diary pages proposed for this study, and Appendix E shows a sample of a completed diary page received from a participant.

Most of the participants did very well filling out the diary form, using it as it was intended with little or no instruction other than a sample. A few people not only filled it out well but recorded complete stories about what they encountered and why they made shifts from one

strategy to another. I believe that having provided the diary form electronically made it easier

for subjects to type in their comments and stories.

After the subject completed his/her diary, a follow-up interview was conducted to collect

the narrative to accompany the diary plus other insights, experiences, or stories that are relevant

to their research strategies and processes. The second interview was shorter and had two primary

sections. The first two or three questions focused on the subjects' research processes during the

study period as recorded in the diary. Also included was a crucial question about whether they

believed any of the barriers they encountered in their research were related to language, culture

or geography. The remaining two or three questions ask for the subjects' suggestions of how

systems, libraries, and other entities could serve them better and what advice they would give to

less experienced researchers in their field about how to conduct research more successfully.

These last two questions were universally asked of all participants since they illicit extremely

insightful information. The final interview questions are shown in Appendix F.

The interviews, questionnaire, and diary methods complement each other and help

provide a more complete and accurate representation of the research process.   The more

objective nature of the diary form and the survey complement the much more subjective and in-

depth nature of the semi-structured interviews.

4.5. Data analysis

Coding is one of the first steps in data analysis. "The analysis of qualitative data

involves creativity, intellectual discipline, analytical rigor, and a great deal of hard work (Patton,

2002, p. 442)." It is a misconception that coding and data analysis begin when data collection is

complete. Naturalistic inquiry encourages the evaluation of data as it is collected. As patterns

emerge, themes may begin to appear and shed light on the continuing direction of the research

itself. This can be called a dialectic process which involves the collection of descriptive data, analysis of the data, and then proposing a generalization based on the data (Grover & Glazier, 1985; Bailey, 2007). This is the basic tenet of Glaser's constant comparative method (1965), in which the researcher considers each incident as it occurs in comparison to categories established based on analysis of prior data points. The purpose of the constant comparative method is to generate grounded theory more systematically by using explicit coding and analytic procedures while allowing for the development of theory throughout the process. Moreover, it is designed to aid researchers to generate a theory that is integrated, consistent, plausible, close to the data, and in a clear form.

The steps of the constant comparative method are: 1) The analyst codes an incident using explicit coding and analytic procedures at compares it to previous incidents already coded in that category. As incidents are added, the analyst will likely develop and adjust their theories. 2) As cases and observations are added to the coding collection, the analyst begins to compare the various categories (criteria) with each other and to integrate the categories and their properties. They develop an overall hierarchy or theory of relationships between them. As this is done, the analyst will likely develop and adjust their theories. 3) Theory is delimited to keep its development from becoming too unwieldy, keeping it confined and parsimonious. It is also used to determine the scope of the theory and how it might be more broadly applied to a wider range of situations. Moreover, this stage helps the analyst to commit to a specific definition of the theory and a specific set of categories. 4) At the end of the process, the analyst has coded data and a written theory. (Glaser, 1965; Seale, 1999).

Often, while still in the field, the researcher will begin to develop ideas for how to make sense of the data. These ideas (insights) that occur during work in the field are both part of the

data collection stage and the beginning of the data analysis stage. However, these thoughts should be allowed to arise naturally, and while they may influence (and hopefully improve) both the theory and the methods being used, they should not interfere with the openness of naturalistic inquiry, and the researcher should not rush to make premature conclusions (Patton, 2002, p 437). Some theorists view coding as a relatively mechanistic process of turning words into numbers or categories. However, Miles and Huberman maintain that coding is much more of an artful, creative process.

> "Coding is analysis. To review a set of field notes, transcribed or synthesized, and to dissect them meaningfully, while keeping the relations between the parts intact, is the stuff of analysis. This part of analysis involves how you differentiate and combine the data you have retrieved and the reflections you make about this information (Miles & Huberman, 1994, p. 56)."

Olszewski et al. (2006) describe the work of coding of interviews in the steps of creating agreement, maintaining the integrity of the codes, and then completing the work. Creating agreement refers to the process in which the coders agree upon strategies and standards such as categories and methods. The negotiations involved in that process help preserve the integrity of the codes since an agreed-upon system is more likely to be followed than one that is devised unilaterally and/or quickly. Finally, completing the work means that the researchers analyze the entire transcript or corpus of data, regardless of how complex or controversial the process can be. These three processes can compete and conflict with each other, each having its own goal of either precision, agreement, or expedience (Olszewski et al, 2006).

Qualitative data are voluminous: possibly hundreds of pages of descriptions, transcriptions, etc. "At first, the data are just a 'bunch of words' that, like numbers, reveal meaning only after being compiled and organized (Bailey, 2007, p. 125)." These data must be made into manageable and meaningful observations. Often the term "data reduction" is used to

describe coding, indicating that its goal is to both reduce and simplify the data. Developing a

manageable classification or coding scheme – the first step in qualitative analysis – allows the

researcher(s) to turn the volume of raw data into a more organized, more digestible, smaller sets

that can be easily retrieved when they are needed (Bailey, 2007, p. 127; Chi, 1997). The process

involves identifying, coding, categorizing, classifying, and labeling the primary concepts in the

data, thereby determining what parts of the content are significant. Here, we should contrast

typologies, which are "classification systems made up of categories that divide some aspect of

the world into parts along a continuum" and taxonomies, "which completely classify a

phenomenon through mutually exclusive and exhaustive categories (Patton, 455)."

A first step in the process involves "tidying up" the data by making sure all of the

documents and sources are organized, missing data are filled in, labeled, and so on. A second

step involves finding items by looking at frequency (things that occur often), omission (things

that do not occur at all), or declaration (things specifically brought out by the participants

themselves) (LeCompte, 2000). Using recurring regularities in the data to see patterns allows for

the creation of categories that then should be evaluated based on both internal homogeneity and

external heterogeneity. This means that the data within the same category fit together, resemble

each other, and have common properties, and that the data in each category differs in a

noticeable and marked way from the data in other categories (Patton, 2002, p. 465; Chi, 1997).

The organization of the data within categories should not be fuzzy, unassignable or overlapping

(Guba, 1978, p. 53).

Qualitative research is often related to inductive creation of theory. Similarly, inductive

data analysis involves the discovery of the patterns, themes, and categories from the data. Once

patterns, themes, and categories are derived inductively from the data, the researcher may begin

to formulate theoretical propositions and hypotheses, which is actually a deductive process (Strauss & Corbin, 1998). The portion of coding that is done inductively is often called "open coding," which indicates that the coding scheme and the researcher are open to the discovery of new thoughts and patterns (Patton, 2002). Strauss and Corbin (1998, p. 223) liken open coding to working on a jigsaw puzzle. The person working the puzzle may begin by organizing the pieces by edges, colors, etc., then looking for pieces that are similar and fit together. The researcher does similar operations with the collected data.

Through coding and data analysis, the researcher interprets meaning from the data. The researcher looks at each interview, case study or observation and asks, "What does this mean? What does this tell me about the nature of the phenomenon of interest? The analyst works back and forth between the data… and his or her own perspective and understandings to make sense of the evidence (Patton, p. 477-8)." This may include cross-case descriptive comparisons or possibly causal explanations – the "why" – or it may portray a holistic picture of what the phenomenon or setting is like. The researcher moves back and forth between the actual phenomenon of interest and his/her abstractions about the phenomenon, between the descriptions of the events and his/her interpretations about it (Patton, 2002, p. 480).

The data collected during my study were analyzed using grounded theory and the constant comparative method (Glaser, 1965; Glaser, 1967) as described above. The data was analyzed using the three steps of open, axial, and selective coding, but it must be said that most of these steps overlapped in chronology during the course of the study. The data from interviews and diaries were taken from their initial recorded form, segmented into distinct units of information, and assigned labels to define and organize them. Table 4.4, below shows the scheme used for coding the data related to the three research questions.

Table 4.4 – Sample of coding scheme

ID	Types of Strategies	Types of Shifts	Types of Roles

　　　　Several tables within a Microsoft Access database were used to segment and categorize the data according to the above scheme. These categories of data were compared and contrasted in search of similarities and differences. As categories were identified, codes were refined. The following tables were included in the coding scheme:

- Participants – Information on the participants, including languages, disciplines, pseudonyms, contact information, and information from the consent form.

- Progress – Tracking information on the activities completed by the subjects during their participation in the study. Columns were also included for subsequent steps such as transcription, data coding, etc.

- Categories – Abstract representations of events, objects, actions or interactions, identified as significant and distinct units of information. Strategies, barriers, and related insights were gathered from interviews and diaries. An example is provided in Table 4.5, below.

Table 4.5 – Types of categories: Sample coding table

ID	Category Name	Description	RQ	Subjects mentioning
26	Strategy: Searching Scholarly databases	Using scholarly databases (library, electronic)	Q1	S3, S7, S8, S11, S16, S18, S19, S20, S23, S25, S30, S34, S36, S37, S38, S40, S41, S42, S49
27	Strategy: Private collections	Using books or materials that the scholar owns	Q1	S3, S6, S7, S8, S10, S13, S15, S17, S19, S23, S24, S29, S30, S31, S32, S33, S34, S35, S36, S38, S40, S44, S46, S48, S49
28	IR: Progress in digitization	More and more sources are being digitized		S3, S4, S8, S15, S18, S23, S25, S35, S36, S41, S47
29	Barrier: Accuracy of abstracts	Abstracts may not match content based on culture or language		S3, S13, S20

ID	Category Name	Description	RQ	Subjects mentioning
31	Strategy: Contacting authors & family members	Finding authors, families of authors and others	Q1	S4, S16, S18, S19, S25, S29, S32, S38
32	IR: Serendipity	Finding things by happy accident		S17, S18, S19, S29, S32, S36, S38, S45, S46
33	Lang: NEED to travel overseas	Must visit the target location to obtain sources		S7, S10, S13, S18, S23, S25, S36, S41, S44, S47

- Quotations – Direct quotations taken from interviews and/or diaries that illustrate specific themes. These are then related to the corresponding theme in the database.

- Strategies – Strategies were coded in two different ways. During the initial coding of the data, a table was created to record which subjects used what strategies as evidenced in the interviews or diary. The strategies are listed along the left as row titles, and the subject IDs are listed along the top. The strategies are grouped by category but otherwise shown in the order in which they were coded during data analysis. The cells are then filled in to indicate which subject used which strategy and in what interview or diary; this is shown in Table 4.6.

Table 4.6 – Types of strategies: Sample coding table

Participant ID	S2	S3	S4	S6	S7
Formal					
Library OPAC (incl. WorldCat)		Int2		diary1, diary3	diary1, diary2, Int1
Scholarly database		Int1, Int2			Int1
Google Scholar/Books	Int1		Int2	yes	diary1, diary2, Int1
Web search engine	Int1, Diary1	Int2			diary2, diary3, Int2
Online archival site/finding aid					
Discovery tool (e.g. Summon)					
Social Media or streaming	Int1, Diary1			Int2, Diary?	
Wikipedia		yes			
Search a publisher, bookseller site					
Informal					
Visiting libraries		Int1	Int1, Int2	Int1	Int1
Visiting archives - outside US		Int1	Int1, Int2	Int1, Int2	Int1

Participant ID	S2	S3	S4	S6	S7
Visiting archives - inside US			Int1, Int2	Int1, Int2	
Citation tracing				Int1, diary1	
Visiting bookstores					
Private collections		Int1		Int1, diary1	Int1
Known web site	Diary	Int1	Int2	diary1	
Archival finding aid (print, non-searchable)				Int1	
Newspaper (print or web)	Int1			diary3	yes
Journal run (browse specific titles)					
Using a print index or bibliography				Int1	
Human					
Asking a librarian / archivist		Int1	Int1	Int1, diary1	
Asking a colleague (expert)				Int1, Int2	
Contacting author or family member			Int1	Int1	
Use a research assistant					
Forums and listservs, blogs					
Learn from a senior scholar					
Attending conferences		Int1	Int1		diary1, Int1
Social media	Int1			Int1, diary1	
Interviewing human subjects					
Hybrid					
Inter-library loan		Int1, Int2		Int1, diary3, Int2	Int1
Purchase a copy online				diary3, Int2	

- Shifts - A chronological log of information-seeking threads, including strategies and shifts, was coded from the diaries. Included in this table is a classification of each strategy according to the information triangle and each shift according to the three types. A sample of an information-seeking thread with three strategies and two shifts is provided in Table 4.7, below.

Table 4.7 – Types of shifts: Sample coding table

ParticipantID	Circumstance	StratShiftType	StrategyName	StratShiftLocation
S15	Reviewed notes from my doctoral research "I kept my notes from PhD research, didn't use this particular source for my dissertation but always found it interesting. To work in AN,	Informal Strategy	Private collections	Diary2

ParticipantID	Circumstance	StratShiftType	StrategyName	StratShiftLocation
	need to know document folder #s to expedite search."			
S15		Planned Shift		Diary2
S15	Contacted AN (Archivo National) to setup an appointment for research early in my research trip	Human Strategy	Ask an archivist	Diary2
S15	Was not able to work on dates initially proposed. Switched days	Alternative Shift		Diary2
S15	Arrived at AN and was able to access folders requested	Informal Strategy	Visit archives overseas	Diary2

In order to use the data from the interviews, they were first transcribed into text files. Then the transcripts were read and themes were coded inductively. Specific attention was paid to information that answers the research questions, especially the information-seeking strategies mentioned and experiences related to the role of language, culture or geography. Since the strategy list was derived inductively, it does not necessarily match any pre-determined list from other scholars' studies, nor do they exactly match either of my pilot studies. In the initial coding phase, over 50 strategies were coded. As the data was analyzed, any new strategy or theme that was not previously mentioned was added to the list. This is an example of grounded theory and the constant comparative method (Glaser, 1967). Demographic data other than gender were not coded since they are not considered a factor in my study.

Axial coding to condense and refine the data reduced the number to 30 identifiable and distinct strategies, organized into four types according to the information triangle. Table 5.2 in Chapter 5 lists and defines the strategies used by participants in this study.

As the interviews and diaries were read during initial coding, nearly 600 quotations were selected. During preliminary coding, quotations were selected based on their suitability to

describe the phenomenon in question. During axial coding, quotations were re-analyzed and aligned with the category that they best describe. Some quotations originally considered were eliminated, and some that had been set aside were reconsidered and flagged for use.

Shifts and strategies were coded from the diaries, and well-defined shifts were selected to be illustrated on the information triangle and used as exemplars in the results section. 25 triangle diagrams were created from the diary entries and two from interviews with particularly clear narrative of an information-seeking thread. Those diaries in which the information-seeking thread contained only one or two strategies (and therefore zero or one shift) were not mapped to triangles. Also from the diary data, Table 5.3 was built showing the counts of which strategies were used first (initial strategies) and which strategies occurred last (final strategies.)

During the selective coding process, categorized data were further analyzed to identify the central phrenomenon, causal factors, context, strategies, and intervening conditions (Flick, 2009). In this process, a model emerged that explores how the three factors of language, culture, and geography impact the information-seeking process of scholars who use sources that are not in English, including the strategies used and the shifts made between them.   The model is presented in Chapter 7.

Inter-coder reliability was achieved by discussing coding schemes with a fellow researcher and having her code various samples and compare results. Options for coding – especially in the area of categories – were discussed until consensus was reached.

4.6. Data validity and reliability

In order for any research to be truly worthwhile, it must demonstrate **validity** or trustworthiness, as well as **reliability,** and it must **contribute to the body of knowledge** in which it is performed. The assessment of research validity is crucial both in evaluating previous

research and in planning for and conducting one's own research. "Validity refers to the extent to which something actually measures what it is intended to measure (Gorman & Clayton, 1993, p. 25)." It must be recognized, however, that many of the types of validity tests that apply to quantitative research, such as replicability, the testing of hypotheses, causal questions, randomized experimental design, and objective procedures, are not concepts that apply well to qualitative research.

Researchers must address questions about reliability and generalizability in addition to validity. Reliability "refers to the extent to which repeated employment of the same research instrument… produces the same result (Fidel, 1992, p. 39; Gorman & Clayton, 2005)" or the consistency of findings over time (Bailey, 2007, p. 180). Generalizability means that the results as drawn from a sample can be applied to the overall population of that group. Based on Guba's model (1981), Krefting (1991) indicates that the notion of reliability in quantitative research corresponds to that of dependability in qualitative research, both of which have the goal of establishing consistency. In qualitative research, and in this dissertation study specifically, reliability may not imply the generalizability or consistency over a larger or broader group of subjects. Thus, the discussion of reliability is inherently present in the discussion of validity.

Traditionally, the validity of a qualitative study has been demonstrated by showing the degree to which the researchers' claims about knowledge correspond to the reality being studied (Cho & Trent, 2006, p. 320), that is, the extent to which the study measures what it intends to measure (Bailey, 2007). However, there is not always agreement among qualitative scholars on what constitutes validity. Seale (2007) discusses the tension between hard and fast rules for validity and the desire to present research in ways that the intended audience will be attracted to. He proposes a model similar to Campbell and Stanley's (1966) conception of external and

125

internal validity, where an inner dialogue concerns the internal logic such as a strong relationship between claims and evidence, and the outer dialogue concerns the external relations of a research project – its relevance to practical and political projects, its consequences, uses and overall purpose as well as context for which a study is intended and to which it contributes knowledge (Seale, 2007, p. 380; Bailey, 2007).

Cho and Trent (2006) suggest two general approaches to validity: transformational and transactional. Transactional validity involves the researcher and the subjects revisiting the collected data, their feelings, experiences, values, and beliefs to improve the level of accuracy and consensus. One of the techniques is called "member checking" and is also discussed briefly on the list below; it allows for misunderstandings to be identified and revised. Transactional validity can also involve triangulation, described as the "use of multiple methods… [to] partially overcome the deficiencies that flow from one investigation or one method (Denzin, 1989, p. 236)." Triangulation is also explained further below.

Transformational validity is described as "a progressive, emancipator process leading toward social change that is to be achieved by the research endeavor itself (Cho & Trent, 2006, p. 321-322)." In this approach, the basic tenet is that qualitative research should result in the betterment of society, specifically social justice, so that the validity of the research is proven by the change it effects. One key concept in transformational validity is self-reflexivity. That is, the researcher identifies, describes and analyzes the meaning that they give the research results based on their social constructions and multiple meanings. Thus, this "catalytic validity" can be described as "the degree to which the research empowers and emancipates the research subjects (Scheurich, 1996, p. 4)."

The interpretivist researcher will rely on the validity of his/her methodologies as proof that the data collected and the conclusions drawn are worthwhile and correct. To do this, they may use the lens of the researcher – enforcing their own internal validity by revisiting and re-checking their data, their methodology and their conclusion – and the lens of the participants – verifying the accuracy of perceptions and conclusions drawn from the data with the subjects themselves (Creswell & Miller, 2000).

Since the analysis of qualitative data cannot disprove bias nor can it be easily prevented using sampling techniques, researchers must take special care to protect against bias. One such type of bias is the holistic fallacy in which the researcher interprets events as more patterned and congruent than they actually are, eliminating or ignoring the outlier data which may actually be of great value in understanding the entire phenomenon. Similarly, in elite bias, the researcher ascribes higher importance to data from those subjects that were more articulate or higher in status and underplays the data from less well-spoken and lower status subjects. Finally, when a researcher "goes native," they become so in tune with the informants that they lose their perspective and ability to properly categorize and analyze the data correctly (Miles & Huberman, 1994, p. 263).

In place of the concept of validity, some qualitative researchers from the interpretivist view prefer to assess *trustworthiness*, which involves "conducting and presenting the research in such a way that the reader can believe, or trust, the results and be convinced that the research is worthy of his or her attention (Bailey, 2007, pp. 180-181). Gorman and Clayton (1993) provide three rather straightforward and tangible criteria to be used to assess qualitative research: 1) well-designed and sound procedures, 2) reliability, and 3) data analysis (Gorman & Clayton, 1993).

In the area of procedures, it is crucial that the procedures and instruments used be described in detail and related to accepted qualitative practice (Miles & Huberman, 1994).

In data analysis, a description of analysis methods and coding schemes will help establish the validity of data. Gorman and Clayton also propose a list of specific steps to assure validity of qualitative research from that perspective:

- Triangulation: collecting data from different sources or using differing methods

- Full documentation of data, including a 'chain of evidence' that references pertinent data sources such as interview or observation transcripts, source documents, etc.

- Documentation of the connections between the collected data and the conclusions that can be drawn from them

- Deliberate inclusion of data even when it does not support the hypothesis or observed conclusions

- Willingness to consider alternate explanations for the observed phenomena

- Self-reflection on the part of the researchers

- Review of the documented results by participants or other independent observers to confirm that the report provides a fair and accurate portrayal of the event, situation, or phenomenon, often called "member checking."

- Awareness of limitations of both the gathered data and the generalizability of the study (Gorman & Clayton, 1993, p. 25-26).

Busha and Harter (1980) compiled a list of questions for the evaluation of research in librarianship, some of which are more targeted to quantitative research, although many are considered universal. Gorman and Clayton (2005) then organized those questions into categories that correspond somewhat well to the criteria that I discuss as relevant to the assessment of

128

research validity: research strategy, instrument design, sampling, coding analysis of data, and presentation of findings. A subset of those questions is given below, structured according to the usual format of a research paper.

Overview (abstract and introduction, research strategy)

- Is the purpose of the research explained and its importance noted?

- Is its theoretical basis presented?

- Has relevant previous work in the area been drawn upon, and related to the present study?

- Is the use of a qualitative approach discussed and justified?

- What is the research question or focus of the study?

Methodology (instrument design, sampling, coding, data analysis)

- Are the instruments used by the investigator adequate reflections of the conceptual variable of the study?

- Are the data collected from different perspectives or obtained from different techniques?

- What sampling methodology was used? Was it well chosen, explained, and justified?

- What techniques were used to analyze the data? Were they appropriate to the investigation?

- Are the processes used for identifying patterns and key ideas clearly described, and do these appear likely to have arisen from the data themselves?

- Is the methodology explained well enough as to be replicable by subsequent studies and researchers?

Presentation of the findings (findings, discussion, conclusions)

- Are examples of the data obtained included?

- Do the discussion and conclusions directly relate to the research questions or focus?

- Are the conclusions and recommendations made clearly based upon the findings, and do they seem reasonable?

- Did the investigator make recommendations for future study?

- Was the research report written in a factual, straightforward, honest, and lucid manner? (Gorman & Clayton, 2005, p. 21-23).

In most discussions of qualitative methods, there is a defense of the methodology itself as valid in the face of critics who would characterize it as non-scientific (Berg, 2001). It can be said that the results of a qualitative study do not demonstrate statistical significance, but this does not negate their value or validity. Instead, qualitative findings are evaluated by their *substantive significance*. The researcher presents the findings and his/her conclusions about them, and readers and users of the study in turn make their own judgments as to their significance. To do so, four types of questions can be asked:

1. How solid, coherent, and consistent is the evidence in support of the findings?

2. To what extent and in what ways do the findings increase and deepen understanding of the phenomenon being studied?

3. To what extent are the findings consistent with other knowledge? (Findings that confirm existing theory have confirmatory significance. Findings that generate new theory or knowledge have innovative significance.)

4. To what extent are the findings useful for the intended purpose?

This dissertation adheres to several of the standards of validity described by Cho and Trent (2006), including aspects of both transactional and transformational validity. When the coded data was analyzed and organized into the results chapters, I engaged in "member checking" by sharing and revisiting results with the subjects that were the most often, most

extensively and most significantly cited. They were asked to confirm that their experiences were accurately and fairly portrayed. They were also invited to suggest revisions, ask questions, and critique both content and theory. Some minor revisions in content were made as a result of participant feedback.

In addition, considerable thought was given during the entire course of data gathering, coding, and presentation to transformational validity, especially as it relates to the betterment of society. If a goal of qualitative research is to effect change and improvement in the researcher's field, then there is evidence that the diary form specifically helped the subjects to become more aware of their information-seeking behaviors which may in turn help them to become more effective. Subject S41 explains:

> "The whole concept of writing things down and thinking about where I'm going to find information and how I'm going to find information was interesting because it did make me think about how limited, normally, I look for information. You know, I just got the basics here, and these are the places I go to, and if it's not there, I think I'll go find something else. I thought, 'Well wait a minute, there's got to be a lot more places I could be looking,' and that forced me into… I found a really interesting book -- a Chinese history manual -- and its chock full of links and places – archives to look for more information. That will be a great resource and thanks to your study, I think it will help me immensely in the long run."

The above comment is not the only indication that this dissertation has already caused some self-reflection for many of the participants. Also of benefit to some were tips and stories that I was able to share both from my experience and from the experience of other subjects in the study. For example, it will be explained in Chapter 6 that many scholars who use non-English sources are not aware how the language of the browser can impact search results or how using "alternate Googles" from other countries can improve search results. Both of these topics came up in a handful of discussions between myself and participants, and it created several "aha moments" for them.

While this dissertation does not involve concepts of social justice, it has great potential to improve the understanding of the unique challenges and contexts faced by scholars who use non-English resources. Both information system designers, information professionals, and other researchers in this same group can benefit from understanding the information-seeking behaviors of people in this group. As this is a new topic in the field of information retrieval, the dissemination of this dissertation may begin conversations and promote awareness of the roles of language, culture, and geography in the information-seeking context. Those conversations will lead in turn to information sharing and ultimately more effective and successful research by scholars who use non-English sources.

4.7. Summary of the methodology

The qualitative methods outlined in this chapter were used to explore the answers to my research questions. 40 participants were recruited based on their status scholars in the humanities actively pursuing research using sources in languages other than English. I used semi-structured interviews and diaries to discover and record what information-seeking strategies were used by the scholars, including which are most often chosen first, and which most often occur last. Gathering data inductively and using grounded theory allowed me to identify themes that demonstrate when and how language, culture, and/or geography affect scholars' choices of strategies as well as the shifts they make between strategies in routine, disruptive, and problematic situations.

Table 4.8, below, shows a summary of the research questions and methods used for data collection and analysis in relation to each question.

Table 4.8 – Methodology summary by research question

Research questions	Data collection methods	Data analysis methods
Q1: What information-seeking strategies are used by scholars	Interviews, research diaries, questionnaire	Taxonomies of strategies

conducting research in languages other than English?		
Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?	Research diaries, interviews	Open coding; taxonomies of shifts; mapping to information triangle diagrams
Q3: In what ways do language, culture, and geography play a role in the information-seeking context, especially in the problematic situations?	Examples were gathered from the interviews.	Open coding; taxonomies of factors related to language, culture, and geography

In addition to providing insight into the strategies used by scholars, the shifts made, and the role of language, geography, and culture, the finding presented herein demonstrate the usefulness of the information triangle as a tool for visualization of the information-seeking process.

Chapter 5 – Results: Strategies and Shifts

5.1. Overview

This chapter presents the results of the study as they pertain to research questions 1 and 2, regarding the information-seeking strategies used by scholars and the shifts that they make among the various strategies. Strategies are defined in this study as "activities undertaken by a researcher to advance an information-seeking task or meet an information-seeking need," and shifts are defined as "changes made by users from one information-seeking strategy to another in order to advance the information-seeking process." Results pertaining to research question 3 will be presented in Chapter 6. The two research questions covered in this chapter are:

Q1: What information-seeking strategies are used by scholars conducting research in languages other than English?

Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?

Results related to information-seeking strategies used by scholars are presented in section 5.2 and are organized according to the four categories of strategies represented by the information triangle as discussed in Chapter 3. Results related to shifts between strategies are then presented in section 5.3. The gathered data resulted in the identification of three categories of shifts, which resemble those previously described by Xie (2008): planned shifts, opportunistic shifts, and alternative shifts. Throughout Chapters 5 and 6, quotations from subjects are included to illustrate the findings. The majority of the quotations are taken from the interviews with a few – especially in the discussion of shifts in section 5.3 – taken from the diaries. In cases where the quotation comes from a diary, its source will be noted when presenting the quotation. The quotations used herein are selected based on how well they illustrate and portray the strategy,

shift, or barrier being discussed. Some subjects are quoted quite often, while others are quoted

less; only minimal effort was made to include at least one quotation from each subject.

Rationale for inclusion of quotations include their ability to:

- Express a strategy, quote or barrier that is representative of a common phenomenon

  across the subjects

- Illustrate the phenomenon of interest in a way that facilitates the understanding of the

  reader

- Provide an example of a unique phenomenon in that subject's information-seeking

  context or across the pool of subjects

To assist in navigation, the structure of this chapter is provided in Table 5.1, below:

Table 5.1 – Chapter 5 structure

Section	Title	Description
5.1	Overview	General information about this chapter
5.2	Strategies	Overview of types of information-seeking strategies
5.2.1	Formal System Strategies	Examples of formal system strategies used by participants in the study
5.2.2	Informal Resource Strategies	Examples of informal resource strategies used by participants in the study
5.2.3	Interactive Human Strategies	Examples of interactive human strategies used by participants in the study
5.2.4	Hybrid Strategies	Examples of hybrid strategies used by participants
5.2.5	Initial strategies	Discussion of what strategies are most often used first
5.2.6	Final strategies	Discussion of what strategies are more often used last
5.2.7	Summary of strategies	Summary of strategies as discussed
5.3	Shifts	Overview of the concept of shifts
5.3.1	Mapping shifts to the information triangle	Explanation of how the information triangle is used to portray the shifts made
5.3.2	Planned shifts	Planned shifts made in routine situations
5.3.3	Opportunistic shifts	Opportunistic shifts make in disruptive situations
5.3.4	Alternative shifts	Alternative shifts made in problematic situations
5.3.5	Complex information-seeking threads	Discussion of information-seeking threads that are complex and involved many strategies and shifts
5.3.6	Summary of shifts	Summary of shifts as discussed
5.4	Chapter summary	Summary of chapter contents

5.2. Strategies

The first research question asks: What information-seeking strategies are used by scholars conducting research in languages other than English? As stated in Section 3.2.2, I define information-seeking strategies as "activities undertaken by a researcher to advance an information-seeking task or meet an information-seeking need."

In this study, strategies are categorized into four types according to the conceptual model of the information triangle. The four types of strategies will be explained in detail in the sections below. They are: formal system strategies, informal resource strategies, interactive human strategies, and hybrid strategies.

The list of strategies shown in Table 5.2, below, was built inductively based on the data gathered from participant interviews and diaries. Since the list is based on actual responses rather than on pre-existing lists from previous literature and studies, it may include strategies that do not exist elsewhere, and it may omit strategies that appear in previous studies. It is also important to note that the strategies are listed based on user definition rather than the specific types of moves that they may represent in the field of IR. For example, in most prior IR discussions, there is much talk of chaining backward (citation tracing) and chaining forward. In this study, only one respondent mentioned chaining forward, and that strategy had not been mentioned in either of my prior pilot studies. Similarly, prior studies most often considered browsing in general to be a strategy without any differentiation based on the entity being browsed. To the subjects of this study, the entity is crucial and hence browsing in libraries, bookstores, archives overseas and archives in the U.S. are all regarded as distinct strategies.

Since this study does not use quantitative methods or analysis, the frequencies of strategies was not analyzed according to any user characteristics; this could certainly be done in

136

future studies to uncover any patterns that might relate to discipline, age, years of study, language, geographical area of study, etc. In this dissertation, the subject's discipline is a part of the information-seeking context but not one of the factors actively being studied.

The definition of information-seeking strategies used in this dissertation is multi-dimensional, composed of the dimensions of method and entity. The method dimension is related to the type of action taken or tactic used. Common action/tactics include searching, browsing, consulting or asking, but there can be others. When using the information triangle, the method dimension is a major determiner of the strategy type. For example, strategies with a method of searching are classified as formal system strategies. The entity is related to the source of the information; it is also used to differentiate among strategies and can help determine the strategy type on the information triangle. For example, as explained above, the four strategies of browsing in libraries, browsing in archives in the U.S., and browsing in archives overseas, and browsing in bookstores all share the same method of browsing but have separate entities, so they are considered herein as distinct strategies. Any strategy that has an entity that is a human being is classified as an interactive human strategy.

The strategies are presented in this section first according their type and then according to the frequency with which they were mentioned in the study. Table 5.2 reflects the same order as the numbered sub-sections in which the results are presented below. In addition, Appendix G shows the strategies used in a tabular format by the discipline of the subjects.

Table 5.2 – Types of strategies

Section	Strategy	Definition	Researchers Mentioning	#
Formal System Strategies:				
5.2.1.1	Search library OPACs	Searching for books using a library online public access catalog, including WorldCat	S3, S6, S7, S8, S10, S11, S13, S16, S18, S19, S20, S29, S30, S31, S33, S34, S35, S36, S37, S38, S39, S40, S45, S46, S48, S49	26

Section	Strategy	Definition	Researchers Mentioning	#
5.2.1.2	Search the web using a search engine	Using a web search engine such as Google to search for relevant web documents	S2, S3, 27, S8, S11, S14, S16, S18, S20, S23, S25, S31, S32, S33, S34, S36, S37, S38, S40, S41, S42, S43, S45, S46, S48	25
5.2.1.3	Search scholarly databases	Searching for scholarly articles in periodical databases such as Ebsco, Jstor, Project Muse, ERIC	S3, S7, S8, S11, S16, S18, S19, S20, S23, S25, S30, S34, S36, S37, S38, S40, S41, S42, S49	19
5.2.1.4	Search Google Books or Google Scholar	Using Google Books and/or Google Scholar	S2, S4, S6, S7, S13, S19, S20, S25, S31, S32, S34, S41, S43	13
5.2.1.5	Search an online archival finding aid	Searching a finding aid or web site provided by or in an archive	S10, S14, S29, S31, S46	5
5.2.1.6	Search in a discovery tool (e.g. Summon)	Searching in my library's discovery tool	S11, S25	2
5.2.1.7	Searching social media or streaming sites	Searching social media postings (e.g. Tweets) or streaming sites (YouTube)	S2, S14	2
Informal Resource Strategies				
5.2.2.1	Visit and/or browse libraries	Consulting or browsing collections in a library	S3, S4, S6, S7, S8, S10, S11, S16, S17, S18, S19, S23, S29, S31, S32, S33, S34, S36, S37, S38, S39, S40, S41, S43, S44, S47, S49	27
5.2.2.2	Visit and/or browse archives outside the U.S.	Consulting or browsing collections in archives outside of the U.S.	S3, S4, S6, S7, S8, S10, S11, S15, S16, S17, S23, S24, S30, S32, S35, S36, S37, S38, S39, S40, S41, S44, S45, S46, S48, S49	26
5.2.2.3	Consult private collections	Using materials already in the person's own possession	S3, S6, S7, S8, S10, S13, S15, S17, S19, S23, S24, S29, S30, S31, S32, S33, S34, S35, S36, S38, S40, S44, S46, S48, S49	25
5.2.2.4	Citation tracing	Following the footnotes, citations, and bibliography of one source (often called a seed document) to identify other useful sources	S6, S8, S10, S11, S17, S18, S19, S20, S24, S29, S31, S32, S34, S35, S38, S39, S40, S41, S42, S44, S45, S46, S48, S49	24

Section	Strategy	Definition	Researchers Mentioning	#
5.2.2.5	Consult known web sites	Consulting web sites known to the researcher, including static library indexes, listings, bibliographies	S2, S3, S4, S6, S10, S13, S16, S19, S20, S23, S29, S30, S34, S35, S37, S38, S41, S42, S45, S46, S47, S48	22
5.2.2.6	Visit and/or browse archives in the U.S.	Consulting or browsing collections in archives in the U.S., including nearby major research universities	S4, S6, S8, S11, S13, S17, S18, S30, S31, S32, S33, S35, S36, S37, S38, S39, S42, S45, S46, S47, S49	21
5.2.2.7	Read newspapers (on-line or in print)	Reading and consulting on-line or print newspapers	S2, S6, S7, S14, S15, S18, S25, S29, S30, S31, S32, S33, S34, S36, S38, S45	16
5.2.2.8	Visit and/or browse bookstores	Consulting, browsing and/or purchasing physical resources in bookstores	S11, S13, S16, S17, S18, S25, S34, S38, S40, S44	10
5.2.2.9	Consult archival finding aids	Using archival finding aids – print or static online, non-searchable	S6, S10, S24, S29, S31, S35, S36, S44, S45, S46	10
5.2.2.10	Use a print index	Consulting a print index	S6, S13, S16, S20, S23, S35, S36, S37, S39, S49	10
5.2.2.11	Journal run	Browsing multiple issues of a specific journal previously identified as useful	S16, S18, S36, S40, S45	5
5.2.2.12	Chaining forward	Using a citation tool to see what newer sources have cited the current source	S2	1
Interactive Human Strategies				
5.2.3.1	Consult a colleague or other expert	Consulting with a person in the field by means such as e-mail, phone, in person, etc.	S6, S15, S17, S18, S23, S29, S30, S31, S33, S34, S35, S37, S38, S40, S41, S43, S44, S45, S46, S49	19
5.2.3.2	Attend a conference	Attending a conference and interacting with presenters and/or presentation	S3, S7, S8, S13, S15, S16, S17, S25, S30, S31, S32, S34, S38, S39, S41, S43, S48	17
5.2.3.3	Ask a librarian or archivist	Asking for help from a librarian or archivist	S3, S4, S6, S10, S11, S13, S20, S29, S32, S34, S35, S38, S45, S46	14
5.2.3.4	Use social media to contact other people	Sending a message or posting on social media such as Facebook, Twitter	S2, S6, S14, S19, S24, S36, S37, S43, S48	9

Section	Strategy	Definition	Researchers Mentioning	#
5.2.3.5	Contact an author or family member	Contacting the author of a source directly or his/her family member or friend	S4, S16, S18, S19, S25, S29, S32, S38	8
5.2.3.6	Learn from a mentor or senior scholar	Learning from a mentor such as an advisor or other senior scholar	S10, S16, S18, S20, S30, S31	6
5.2.3.7	Use forums, listservs, or blogs	Consult or post on a web-based forum, blog, or email listserv	S16, S19, S24, S35, S38, S43	6
5.2.3.8	Use a research assistant or send someone	Engaging the help of a research assistant or surrogate to access documents in your place	S29, S30, S35, S41, S44	5
5.2.3.9	Interview human subjects	Conducting interviews of humans as primary sources	S4, S11, S29, S30	4
Hybrid strategies				
5.2.4.1	Use interlibrary loan (ILL)	Requesting materials from another library through interlibrary loan	S3, S6, S7, S8, S10, S11, S13, S15, S16, S18, S19, S20, S25, S29, S30, S32, S33, S34, S37, S38, S40, S42, S43, S44, S45, S46, S47, S48	28
5.2.4.2	Purchase a copy online	Utilizing an on-line book vendor to purchase a source	S6, S11, S13, S16, S18, S25, S34, S36, S40	9

5.2.1. Formal System Strategies

As defined in section 3.2.3, **formal system strategies** are those in which the method of searching is used in conjunction with entities such as periodical databases, OPACs, electronically indexed and searchable finding aids, digital libraries and repositories, and web-based resources. The three characteristics most often used to categorize strategies as formal systems strategies:

1) Search-based,

2) Computer systems-based, and

3) Having an interactive interface

In other words, these strategies are typified by the activity of a user entering search terms into a search tool and the tool/system subsequently returning search results. The user then assesses the

relevance of the results and may choose to explore specific results or re-formulate his/her query and search again. All of the strategies classified as formal system strategies follow this paradigm. The tactic of searching differentiates these strategies from informal resource strategies, in which the tactic is browsing.

One might think that all scholars would prefer to use searchable systems rather than other types of strategies for the convenience factor, and while this is often that case, scholars may also "settle for" searching on-line because distance and the uniqueness of their field does not allow for the human contact that they would prefer. Subject S33 explains:

> "I tend to use, I would say, electronic resources most of all, precisely because I am so far from Japan. If I were working in a Japanese university, I could probably go down the hall and ask a colleague. It doesn't really work so well here when there aren't that many of us who, you know, speak Japanese well enough to do research in it."

Seven different formal system strategies were used and mentioned by the participants a total of 92 times. They are presented in the sections below starting with the most frequently mentioned.

5.2.1.1. Searching a library OPAC

Searching in a library on-line public access catalog (OPAC) – including WorldCat – was mentioned by 26 subjects – the greatest number of all formal system strategies. As shown in Table 5.3, below, it is one of the most commonly used initial strategies, providing scholars with a reliable way to identify secondary sources and learn where and how they might obtain them.

Most of the participants used WorldCat specifically and for good reason. S36 says "I do use WorldCat quite a bit because my books are so rarely in my home collection, and WorldCat makes it so easy to quickly see where the book is." S3 adds, "The advantage of WorldCat is that it gives me location." S8 provides this language-specific assessment, "Even fairly obscure

Chinese materials are pretty accessible through WorldCat." Subject S18 voices her high praise

for WorldCat and provides the following language-related rationale:

> "WorldCat has really been a great resource for some years now because it actually handles non-Western characters. It certainly doesn't include many libraries in Japan, but if I'm trying to figure out the title of a book or something, and I don't know the title, WorldCat can be very helpful and productive in helping me narrow what it is that I'm trying to find."

Another reason that WorldCat is especially helpful to non-English researchers is explained by

subject S33, citing a common barrier for Japanese scholars of knowing the correct representation

for proper names, as explained in section 6.2.4. She says:

> "I use (WorldCat) quite a bit because when you go online to search on WorldCat, you can get the English language version of the citation, but there's also a little button you can click on to get the original character version, so you can toggle between the Japanese and the English, and if you're looking up an author's name, if you've got the name of a work that they've written, you just enter in the title and then go between the Japanese and the English, and you can see very easily how that name is supposed to be written."

Of course, there are other OPACs beside WorldCat. Subject S48 uses her own college's

OPAC which is built on WorldCat and includes worldwide results along with local holdings.

Others, such as subject S8, rely on OPACs from their graduate school libraries, "I did my

graduate work at Harvard, so I had a pretty good idea of what the library has. But they have an

excellent online catalogue, so that's something that makes things quite a bit easier."

A few subjects mention using OPACs and union catalogs in their target country. Subject

S30 gives an example: "For the government documents, I went to the archives in Pretoria and

also the library in Cape Town after doing a search in South Africa's national archive catalog

which I'm pretty sure includes all university and government archives." Subject S3 add that

"because some of the articles and some of the reviews are not in common journals here, there's

also a kind of WorldCat for serials in Germany called DigiZeitschriften so I also double-check with that."

5.2.1.2. Searching the web using a search engine

Not surprisingly, searching the worldwide web with a search engine is among the most popular formal system strategies with 25 participants mentioning its use. Web searching can be used to identify both primary and secondary sources, and is also a commonly used initial strategy, as shown in Table 5.3. Web searching can be used more effectively than many other formal system strategies when searching in languages other than English, and it is effective from nearly any location. Finally, web searching is helpful whether the topic of research is ancient, recent, or in between. Subject S34 gives an example of researching very recent films from Spain:

> "Say I want to research a film that I haven't really investigated. I will take the film and Google information about the author, about the film, or about the reception of the film. Through Google, it will take me through articles, like newspaper articles from Spain like El Païs that have been written. I will look at the reviews first."

Despite the popularity and value of searching the web, many academics feel a bit self-conscious that perhaps "Googling" is a less reputable research strategy. The comment made by subject S11 both explains the virtue of web searching and belies the reputation that it has compared to scholarly databases:

> "You can't be a snob about your searches. You can't assume you have to use academic databases to get where you're going. I think it's really important not to be above Googling and just know that it's often a two-step process. You have to find what you need, and then you have to verify its validity."

Subject S15 provides another example of how searching the web complements the scholarly strategies that she tried first:

> "Author #3 I found only as a general Google search. I couldn't find them on any of the academic search engines, not even Google Scholar, but just in Google. I

searched for this author, and her work came up, and then I was able to find the Brazilian journal in which she had published the article, and they had archived all of their articles on-line, so I was just able to access the PDF immediately."

While nearly every subject that mentioned web searching spoke specifically of Google, often using it as a verb, subject S11 continues on to discuss the value of multiple search engines, "I use different search engines because they turn up different things. The ones I use most are Google, Yahoo, and Duck Duck Go." Moreover, some larger or more technologically advanced countries have their own search engines that provide mainstream alternatives for scholars researching in those cultures. Such is the case with Yandex in Russia, Baidu and Sogou in China, and Naver in Korea. Subject S36 gives an example:

> "I'll actually go to Yandex or something, like a Russian search engine rather than Google if the results that I'm getting aren't answering my question fast enough. I do just use the main Google page, but I change my font and type in Russian and I find that that usually works pretty well."

One aspect of web searching that is unique to scholars who use sources in other languages is the phenomena of "other Googles." Google has various versions of their search engine that use alternate languages and have search algorithms, stemming algorithms, and interfaces customized for those languages. These can be accessed by using the domain name (e.g. google.fr, google.de, google.mx, etc.) Subject S11 gives an example of researching the history of Le Marais – a historically Jewish neighborhood in Paris: "Because 'Marais' doesn't (mean anything in English,) when you use that in google.com, it doesn't turn up things in the same depth as if you go to google.fr." Subject S34 explains further:

> "Start with google.es in the actual country; there you're getting texts from Spain. I think it's very good to go into their newspapers. They write a lot more about culture in their local newspapers than Americans. There's even a Google for Catalonia, for the Barcelona area. I definitely use that."

In fact, scholars who use Spanish have a unique challenge and opportunity. While scholars of French all gravitate to google.fr, scholars of German to google.de, and scholars of Japanese to google.jp, Spanish is spoken in numerous countries, each of which have their own Google. Subject S25 studies Latin America, and he mentioned using various Googles from Spanish-speaking countries, including Mexico, Argentina, Venezuela, and so on.

In the course of my two previous pilot studies and this study, I found that not all scholars use the alternate Googles (Sabbar, 2013; Sabbar, 2016). For some, they feel they find enough results in google.com, so they do not look beyond. Shortly after interviewing subject S25 as explained above, I conducted a Skype interview with subject S7. She indicated that she mostly used google.com and sometimes google.es – logical in that most of her research is in Iberian topics. As I explained how subject S25 used the various Latin American Googles, I could see her typing on her computer, after which she said "I just found a blog on google.ar that I can use."

Even fewer scholars understand the role that the browser plays in alternate language searching. If the user sets either the language of their computer or the language of their browser to the target language, Google will interrogate the browser and computer and match the Google interface and algorithms with the user's preferences (Sabbar, 2010), but very few users understand this interaction. For scholars of Japanese, this can cause some mysterious but ultimately explainable search results. When asked if she used google.jp, subject S33 related a challenge that was also told to me by a subject in my first pilot study (Sabbar, 2016):

> "One of the biggest problems I have is sometimes when the characters in question that I've entered are actually common in Chinese too, because there are a number of websites in Chinese online, it seems to dwarf the number in Japanese, and a lot of times it's common in Chinese - I will get only Chinese websites up until the third or fourth page of hits, so that's kind of irritating."

This problem arises because one of the character sets of Japanese is based on Chinese, so when entering those characters into google.com, Google is unable to discern if the user is searching in Japanese or Chinese, so it will return results in both languages. Since China has a larger population and more web sites, Chinese results will logically outnumber the Japanese results. Based on what I had learned during my information studies coursework and a prior project, I was able to explain to her that, if she set her browser to Japanese, Google would know that she is searching in Japanese and would not return Chinese language results (Sabbar, 2011). This problem is similar to interference from cognates and homographs when searching in many European languages (Bar-Ilan & Gutman, 2005). Most scholars do not have an opportunity to learn about this complexity on their own, so it is not surprising that language can become a barrier when using web browsers and search engines.

5.2.1.3. Searching scholarly databases

19 of the subjects talked about searching in scholarly databases for the purpose of identifying secondary sources related to their research. Subject S42 tells about her use of two specific databases: "The first thing I would do is go to Ebsco or MLA to research what has been written and sort of take it from there." MLA is mentioned by name exclusively by language scholars (as compared to other humanities scholars such as historians) in this study, including subjects S3, S7, S25, S34, and S42. Subject S7 explains how MLA is her first preferred source for scholarly articles:

> "If I'm looking for articles, MLA. Many times I will say, 'I want to know everything that has been written about this.' So I go to the MLA bibliography and I search there, and after that, I go to Google Scholar and search there because I realize that MLA does not classify everything, so Google Scholar gives me a few added resources."

146

Subject S34 explains why MLA was once the only database that she used but that she has

recently expanded into additional databases based on training provided by her school's

librarians:

> "In grad school, we're just told to go to the MLA database. But then learning and
> going to the workshops our library has… it's very small, but they're very good
> with us. I've learned how to use other databases other than the MLA."

Jstor was specifically mentioned by subjects S3, S25, S40, S41, and S49. Subject S38

also uses LexisNexis, saying, "For the French Algerian book, I have sources off of LexisNexis,

and I'm getting French sources in particular there and a few Algerian francophone newspapers."

Scholars conducting research in Spanish may have some advantage, as many libraries

subscribe to Spanish language databases to support their Spanish language courses or even

Hispanic populations on their campuses. Subject S25 says "If it's Spanish, I will also go to

HAPI, which is the Hispanic American Periodicals Index." Subjects S7 and S34, whose Spanish

research topics focus on Iberian literature and media, both talked about Dialnet, which is a free

database from the Universidad de le Rioja in Spain.

Greek and Latin classics scholar S20 tells how the best database for his discipline

recently became more accessible:

> "The second within the category of digital search tools is classics databases, and
> the key one of these is L'Année Philologique – the Philological Annual. That
> one, up until 2010 or thereabouts, was only available in French and only for a
> very expensive subscription, but now it's available through Ebscohost, which now
> my library has access to, which is a lot cheaper than the proprietary interface."

Similarly, Egyptology scholar S49 tells that there now exists a searchable database for her

discipline:

> "There's something called the Online Egyptological Bibliography, and that will
> give you a complete bibliographical listing up to whatever's been published in the
> last year, so that's gotten to be fantastically useful. It's searchable, so if you look

for 'palace,' it will return 25 pages worth of anything that has palace in the title. Or you can search by author, so truly searchable."

A newer database resource not mentioned in either of my pilot studies but mentioned in this study by three scholars (S15, S20, S48) is academia.edu. Subject S15 explains how academia.edu not only provides secondary sources but also creates an assisted shift to the interactive human strategy of contacting other scholars:

> "Brazil has become a real hotbed for scholarship. If you want to talk about informal networks, I use academia.edu, and I've actually gotten materials from Spain, from Brazil, from Argentina, by just going and actually seeing if that scholar has an academia.edu and contacting them directly."

In addition to scholarly databases, which are usually defined as containing articles from peer-reviewed academic journals, there are other various databases that scholars can search. Subject S46 states simply, "There are a few Dutch databases for images if I am looking for images." Classicists and Biblical scholars such as subjects S19, S20 and S23 speak about databases that contain annotated texts that they frequently consult. Subject S19 gives an example: "There are also online databases of Greek texts which I can use. Then you expand the search from just that text to people in the same genre and in the same time period so that you can get a broader intellectual picture." In regard to databases of religious texts, subject S37 says, "The fact that we have a theocracy in Iran, they have invested a lot in creating databases that have made work easier."

Even though scholars know of useful databases, they sometimes do not have access to them, usually due to cost factors. A recurring theme among scholars using sources in Japanese – both in this study and in my pilot studies (Sabbar, 2013; Sabbar, 2016) – are the costs of accessing Japanese resources. Subject S18 gives an example:

> "There's a couple of different kinds of databases and sources that I would look at. So, first and foremost the National Diet Library in Tokyo, the equivalent of the

Library of Congress, has databases that you can search online. There's also an article database called Find Me for Japanese language materials, which is also pretty comprehensive, but not being a research institution and having access to paid databases, I always wonder how much I'm missing."

Subject S19, a scholar in classics, also finds expense to be a barrier:

"The biggest, most important thing is that for those databases that are not open-source or that aren't free, having subscriptions to them is key… A lot of the bigger ones are being done now by those major academic publishers that are very expensive, and the texts themselves are very expensive, so you can find yourself in a situation where you can't access them digitally or hard copy for certain periods of time."

Another commonly cited factor related to cost is the idea that the scholar may be the only one at his/her institution that would use a particular database, so the institution will not consider the subscription fees to be worthwhile. Subject S38 explains:

"A lot of good databases and libraries are just too expensive, and I don't want to invest in, let's say, a North African database that might be used by a few students. It depends what field you're in, but these fields are not only understudied but also under-researched in regards to databases; there isn't enough there. I'd like to get the Economist Intelligence Unit; that's very expensive. (A local large research university) used to have it, but it was too expensive for them... You have libraries but not the resources to invest in esoteric fields."

5.2.1.4. Using Google Scholar and/or Google Books

13 subjects indicated that they use Google Scholar and/or Google Books. Many use these as part of their routine information-seeking process for identifying published sources, often as a first step. Subject S41 says "One of the first things I do is Google or Google scholar or if I'm looking up just bits of information, then hopefully I will find some lead to the book, and from the book I will find footnotes." Subject S19 adds, "Sometimes I find new primary source material through other scholars, and a lot of times that happens through a Google Scholar search." Even more often, scholars use Google Scholar or Google Books as a second or third strategy to explore what other sources may be available in addition to what they found in scholarly

149

databases or a library OPAC. Subject S31 explains how Google Scholar figures in her frequent

thread of strategies:

> "If WorldCat doesn't give me enough, I will go usually to Google scholar and will try to do very specific Google searches, and then if that doesn't work, I email some of my former advisors who also work in the region and say 'Do you know anything about this source or have you ever come across anything like this?' And I'll try to tap into my professional network."

Subject S6 relates how she ended up in Google Books as a result of a Google search that

was a bit of a last resort:

> "I was looking for an inquisitorial sentence. Four people were being condemned, were being sent to the stake in 1318, and I knew that it existed because I mention it in my book, and I even give a reference, but somehow in all the moves that I've done since then, I can't find the original document – the Xerox of the document that I once upon a time certainly had. It's in this giant, enormous, folio-sized book from the 18th century, and I thought this is really going to be a drag because my library doesn't have a copy, and there's unlikely to be any libraries in driving distance that have a copy. So I really didn't quite know what I would do. On a whim, I Googled the name of the book, and it's on Google Books. I was so surprised. And that meant that, within 3 minutes, I was looking at the sentence in the 18th century edition."

For her specific information-seeking needs related to the history of Czechoslovakia, subject S31

finds Google Books and Google Scholar helpful. "A lot of the publications, the early

newspapers from say the 1910s and 20s that were produced by immigrant populations are

available through Google, so I can get on Google Books or Google Scholar and pull up those

publications."

5.2.1.5. Searching an on-line archival site or finding aid

Since a number of scholars in this study use archives and archival sources (see sections

5.2.2.2 and 5.2.2.6), many of them also use archival finding aids. In many cases, archival

finding aids cannot be considered a formal system strategy because they do not employ actual

search techniques but are rather a print index or document. Five of the subjects indicated that

they had access to a searchable archival finding aid tool. It is notable that, while many subjects

know and use the term "finding aid," others do not but rather describe them in functional terms.

Another interesting observation is that four out of five users of this strategy are historians, with

the fifth subject is in modern language; see Appendix G.

On-line searchable finding aids can work much like an OPAC, as subject S10 explains a

searchable resource in Germany that provides some limited information on finding aids and

related archival collections:

> "There is a central website you can go to that links to all the online German
> catalogs. You can go there and search online for the records, but that doesn't get
> the records, that just get you the location, but it also doesn't get you much about
> the records. Some have a little summary of what's in there, some just say there are
> collections of these, and its twenty feet long, and you have no idea what's in
> there. And some of them don't have any online catalogs at all, many archives in
> Germany don't have any catalogs online at all. In fact, Augsburg doesn't have
> any catalogs online."

5.2.1.6. Using discovery tools

A relative newcomer on the scene of academic libraries is the discovery tool – a

bibliographic search tool that searches across multiple academic databases and holdings of a

specific institution. Examples would include Summon, Ebsco Discovery Service (EDS), Primo,

WorldShare, AquaBrowser, and others (Chickering and Kang, 2014). Researchers use a

discovery tool in the same circumstances that they would use scholarly databases; they can save

time by using a discovery tool rather than searching individual scholarly databases. Two

subjects in the study spoke of their use of a discovery tool. Subject S25 describes this new

strategy in functional terms, "Our library has switched to this new thing where step one in the

search process is just entering your basic terms in a search bar, and then it does a search of all

kinds of different databases. That is where I found author #2." He went on to cite his library's

discovery tool by name and describe when he used it and what resulted:

"The first wave of results on Summon, which was the first step, came back with a lot of English language sources that were mainly superficial – you know, mainstream newspapers or journals with just really brief mentions of the film's existence, or in the best case scenario, some kind of a review, although not necessarily or usually an academic review."

5.2.1.7. Searching social media or streaming video sites

Another strategy mentioned for the first time in this study but not in my prior pilots is the searching of either social media or streaming video sources. It was mentioned by two subjects in this study. Here, we must explain that subjects can use social media as a formal strategy – actively searching in these tools for content that they provide – or they may use social media as a human strategy – to connect to other scholars to ask questions using a method of "ask" rather than "search." This section only discusses social media used as a formal strategy. Subject S2 states, "It's sort of a new compendium of sources – Tumblr, Twitter, YouTube – that I never would have thought to use four years ago, five years ago." Scholars actively teaching in modern languages often spoke of finding video sources on YouTube or other video repositories that they could use for their courses. Subject S14 explains his information-seeking task related to the French courses he teaches:

"So I look for shorter… I would say documentaries between a half an hour and an hour long that go along with the reading that repeat the stuff so that they can (learn it.) And I also want to provide them with images in their heads that will allow them to engage more vividly with the history."

5.2.2. Informal Resource Strategies

As defined in section 3.2.3, **informal resource strategies** include those that are traditionally considered in relationship to the entity of print sources and are characterized by methods that allow the user to interact with the resources directly, using the tactic of browsing or possibly consulting. Entities can include library holdings, archival holdings, scholarly

publications, personal collections, indexes, bibliographies, known web sites, and static finding aids. When using informal resource strategies, scholars work directly with resources or their surrogates without the intermediary of an interactive IR system or interface. Strategies such as browsing web sites or bibliographies are considered informal even if they are performed on-line using a computer because the user does not formulate and present a query to a system which then returns a response. Based largely on the work of Marchionini (1995), the distinction between search-based and browsing-based activities is key to differentiating between formal system strategies and informal resource strategies. Each of the 12 informal resource strategies involves the subject looking through an entity, be it an archival fonds, a library collection, his/her own private collection, or the citations in a seed document; no search interface is present, and the subject does not formulate or submit a query.

Browsing is particularly conducive to serendipitous discoveries, and section 5.3.3, which discusses opportunistic shifts in disruptive situations, provides examples of opportunistic shifts that result from the use of informal resource strategies.

Subject S44 explains why he favors informal resource strategies over both formal system strategies and interactive human strategies in his research:

> "For me, (informal strategies) tend to be the most important because almost nothing has been digitized that's Vietnamese, at least not for the period I'm interested in. While there are contemporary periodicals online, the ones that my author published in haven't been digitized at all, so I still do most of my research using published findings aids or language materials I have collected myself. In terms of speaking to people about things, one of the things I get particular about in Vietnamese history is that there are very few of us who do it, at least who use Vietnamese language materials for research, and because there are very few people, the scope of what can be done is very large, so, few people tend to have overlapping expertise... Sometimes you can ask people, but only for general things. Most of the time, if it's something very particular you have to find, you have to do most of the work yourself."

12 distinct informal resource strategies were mentioned 196 times throughout the study –

more than in any other category; these are presented in the sections below.

5.2.2.1. Browsing in libraries

In my prior pilot studies, "browsing in libraries or archives" was considered as one

strategy (Sabbar, 2013; Sabbar, 2016). However, the subjects in this study treat libraries as

distinct from archives, and they distinguish between archives overseas (in their target country or

culture) from those in the United States. The strategy of browsing in libraries is most often used

in the middle of an information-seeking thread, generally when a secondary source was

identified either by searching an OPAC or by citation tracing. Alternately, some subjects find it

useful to browse in a library periodically to scan for new or previously unidentified sources.

The specific strategy of browsing in libraries was mentioned by 27 subjects – the most of

any informal resource strategies and more than any of the formal system strategies. Subject S38,

who studies colonial and post-colonial North Africa, explains his approach to libraries:

> "It may not sound too adventurous going to the library, but once you get there,
> you have to use your imagination and imagine what would be your ideal source. I
> tell my students when I bring them to the library, 'This is a book, basically, and
> you look at these.' I think that's important too, the physicality of research and
> actually carrying five or six books, and actually having that weight."

Subject S7 explains why she specifically chooses to browse the collection of a nearby large

research university library:

> "I go to the library and see what else they have there that I wasn't looking at…
> I'm seeing what they have, like if I go to the U of M(innesota) library. I don't
> look in the catalog sometimes; I go to the section of what I like, and I browse –
> old fashioned browsing, you know. A Medievalist can do that today. Just do the
> old-fashioned way of browsing through the books. I kind of enjoy that, too."

In fact, subject S7 recounts one of the most inspirational anecdotes in the study,

demonstrating the power of accessibility in libraries to support and inspire research:

"When I was studying in Spain, I wasn't that interested in research. Why? Because our libraries are closed stacks. You will have to wait in line for 20 minutes and go to this person and say 'I would like these many books,' and she says 'No, you can only get three.' You couldn't bring them home… it wasn't conducive to research. That's what I found, so I wasn't interested in it. Then I came to the US to the University of Minnesota where you can touch the books – go through aisles and aisles of books. That was it. That's what made me go into my Ph.D. and keep going. Many hours in the library, just browsing."

A common recurring theme among scholars was visiting a nearby university library – often within a few hours of where the scholar now works, but always closer than travelling to libraries overseas. Twelve subjects – S6, S7, S11, S16, S18, S19, S30, S31, S36, S37, S42, S47 – indicated they use this specific library browsing strategy. In some cases, this was the same library where they did their graduate work. Subject S11 explains:

"I'm an hour, hour and a half from Penn State. That's where I got my PhD and grew up, so I know that library inside and out, so I go there and make periodic trips over to Pattee – the library of Penn State, to comb the stacks, and you can also get a 24 hour visiting research pass to use their online databases, which they have many more than we have (in our college library.)"

The above quotation points to another reason to visit nearby university libraries – to utilize subscription databases that can be accessed by IP while on that campus. Subject S36 elaborates:

"In the United States, there's the big problem of, now that I'm not at a big flagship campus, I do not have access to the databases that I would want. I could see why (my institution) would not have (access to) Russian journals all the way from the 20th century to the present in searchable text in Russian…I can't have those. I have to go to Madison to use those."

Specific destinations in the East include the Weidener Library at Harvard (S6, S8, S37), Yale (S31), Penn State (S11,) and University of Maryland (S33). Destinations in the Midwest include the University of Chicago (S39, S49), UW-Madison (S36), Ohio State (S19), University of Minnesota (S7, S8) and the Center for Research Libraries in Chicago (S30). Further west, destinations include the University of Washington (S40, S47). Subject S19, who studies ancient

Greek texts, is fortunate to live in close proximity to the Center for Epigraphical and

Paleographical Studies at Ohio State University, which is about 30 miles away. She says,

> "I sometimes make trips down to Cincinnati, which has a special endowed library.
> It has a lot of material on classical archaeology, most of which I can borrow
> through Ohiolink, which is a godsend. It's the thing that makes Ohio the envy of
> the library world. But, um, sometimes they don't lend certain things, so a trip must
> be made."

Similarly, subject S31 explains that she now lives closer to Yale University than before, so this

allows her to visit the archives even if she is uncertain that they will have sufficient materials to

make the trip worthwhile:

> "(My project is) about the World Student Christian Foundation or the general
> housing crisis in the post war period in Prague, but I couldn't find any secondary
> sources. The primary sources were like jackpot. Where I'm living in New Jersey
> now, I have so many archives around me, where I didn't before. I went through
> WorldCat and found that the Yale Divinity Library has this World Student
> Christian Federation collection that's only like an hour and a half away…. The
> archivist said they might have some material, but… they had like two boxes of
> materials that were all about this project. Half of them were in the Czech
> language; they had all these publications produced by the students living in the
> student housing project. They had photographs of it and the total walking tour of
> it. I couldn't have asked for anything more… about this housing project; they
> literally had everything about it. And if I was still living in Indiana, I don't know
> if it would have taken the risk of flying out here to see what they had, but the fact
> that I'm only an hour and a half away and could take a half-day to gamble on it."

Subject S49 indicates that there are a limited number of Egyptology libraries in the U.S. When I

asked how she knows what libraries will have sources she can use, she replied:

> "By word of mouth primarily. Egyptology is generally a small community.
> There are a limited number of Egyptology archives or resources or libraries. The
> University of Chicago has one; Yale has one… University of Pennsylvania, Johns
> Hopkins, Berkeley, and UCLA. Those are the major ones… the University of
> Toronto also has a program."

So effective is this strategy and powerful is the attraction to these major libraries that a

scholar may relocate to be nearer to one, as subject S37 did earlier in his studies:

"I figured out when I first started my dissertation was that this is a pain to order all these sources through inter-library loan. So my advisor and I talked, and I moved to Boston for two years just to do my dissertation research at the Weidener library, (where I had) all the sources at my fingertips. I would sit at a carrel at the 6th floor of the Weidener library. I didn't really need my personal library a lot because they basically have everything you need."

Since only a few of the subjects in this study are employed by larger research

universities, most of their local collections do not carry the specialized materials that they need

for their own research. However, many of them use their own library's collection and tools as

well. Subject S40 praises her own college library:

"It's such an incredible gift to be at (my college) because of the historical connection and to be focused on Norway - the historical connection here is so rich. And there are so many people that were so committed to making sure that Norwegian language resources and Norwegian-American resources were preserved here. And so I think there are tons of resources in terms of basic historical background that I can find here in our library."

Subject S38 also praises his university's library, "I've been very impressed, and what we have is

good. When it comes to number of books on the Middle East and northern Africa, I think we

now top Notre Dame." Moreover, scholars such as the subjects in this study can influence their

institution's library collections to make them more useful to themselves and their students, as

subject S46 explains: "I think there's more Dutch stuff and things on decolonization – more

things on my particular general topics – since I have been at (my university) because my

librarian has ordered them for me and for my students and so on."

Finally, some subjects mentioned visiting libraries overseas, such as subject S10

explains:

"While I'm (in Germany,) I can go to the local library for sources that are in German that are just harder to get in the States. When I'm in the States, I can borrow them for three weeks, and then I have to send them back. Some are hard to get and some you can't get at all. So there are some secondary sources easier to get (in Germany.)"

157

5.2.2.2. Visiting and browsing archives outside of the U.S.

The population of this study was specifically chosen because they often require resources that are found only outside of the U.S. Not surprisingly, visiting and browsing archives overseas is one of the most commonly used strategies, with 26 subjects saying they use this strategy. Of the various final strategies shown in Table 5.3, it is the second most common, pointing to the reality that there is very often no other way to obtain sources than to visit archives in the target country.

In addition to the examples given here, many other instances of overseas archival research are discussed in Chapter 6 as they relate to the factors of culture or geography.

Subject S10 says, "Everything I have done research on, I could not have done without going into a German archive. I have not written anything that doesn't have at least some basis in manuscript sources only available in archives that have not been digitized."

Subject S44 explains his strategy for conducting research in Vietnamese archives:

"In Vietnam, there are four national archives: two in Hanoi, one in a central hill town called Dalat, and one is in Ho Chi Minh City, which used to be Saigon, and for my research, because my first project was about the history of Saigon, almost all my sources were in the archive in Ho Chi Minh City. So I sort of knew from the beginning that that would be where my sources would be, and I spent about a year there. There are now some published guides in Vietnamese and French for three of those archives, so you can at least have some sense of where you want to go and how you want to orient yourself before you get there. In addition, there is the French colonial archive which is in Aix-en-Provence in the south of France, which is a very easy place to use that has an enormous amount of material as well. They have some online guides to those archives. There's nothing online for Vietnamese archives, but for French stuff there are some finding aids that are now accessible."

All of the historians and most of the other humanities scholars in this study utilize archival sources, including newspapers, government documents, organizational records and personal papers, and many of those are found primarily in overseas archives. Newspapers are

one of the most common sources sought in archives – both overseas and in the U.S. Because the use of newspapers by scholars is so common and so complex, and because its entity is distinct and unique, it is treated as its own distinct strategy in section 5.2.2.7.

Another commonly used source in overseas archives is government documents. Subject S39 accesses Indian parliamentary documents in the National Archives in New Delhi, India. Subject S48 accesses historical government documents on women's suffrage in Mexico City's Archivo General de la Nación. Subject S30 searches health records and related government documents in various archives in South Africa. Subject S7 uses centuries-old official travel records in Spain. As subject S30 also explains, not all government documents are in library archives, but they are found in the target country: "I also went to some government offices looking for maps and other reports. In my research on the nurses now, I contacted some people in the department of health asking where I might find some records." Subject S36 uses a variety of government sources in his research on post-revolutionary France:

> "I've used trial records. Some of my research includes records of trials of individuals who were key functionaries in the Reign of Terror, but after the Reign of Terror, they are put on trial. I work with their trial records, so that includes affidavits, testimony, their own writings, evidence written in support of the prosecution or of their defense, stuff like that. I've used municipal records because I have chapters that deal with transformation of mass graves of victims of the terror into expiatory monuments, commemorative chapels. So the records of that take me into the 1830s - from 1794 to the 1830s - and those are records in several sites in Paris, in the city of Lyon, and the city of Orange, which is in the department of the Vaucluse in south of France."

Organizational documents are also a target of research in foreign archives. Subject S46 provides several examples:

> "The archive where I've done a great deal of work in Holland is the International Institute for Social History. It's a conglomeration, or rather it collects a lot of leftist archives, not just from Holland, but from around the world. So I know if there's something involving socialist, communist, or something not affiliated but on the left, those materials for either individuals or organizations, they have been

left to that archive, not exclusively, however. There are some in private collections. Some organizations are on file with national archives. Another, for example, is on file with the Protestant Documentation Center at a Dutch university. One of the archives I will be going to is the Catholic University Archives where it's the archives of the Catholic political party. So it depends; sometimes it's very obvious where those materials are going to be held, and sometimes it's really not. It's up to the organization, at least I've found, in Holland, to decide where they want those materials to go."

Regardless of the document type, many of the participants discuss the importance of using a browsing tactic when working in overseas archives. Subject S29, who conducts work primarily on Irish immigrants, explains:

"I think (it) requires a lot more browsing and a lot less (searching.) You know, what we were talking about catalogs and searchability and making the perfect search engine – until that's possible, a lot of research still requires trying to find where you think something might be and then having the good luck of being able to browse through the entire collection."

Subjects S6, S10, and S24 all talk about the practice of going through entire fonds (sets of archival records) in a specific archives as standard practice. Subject S13, who conducts research about migrant culture in Germany, tells why it is so important to browse archives overseas, "It's also important to be in country. A lot of stuff is available online but not everything. You know the need to go to a library, the need to go to an archive... to be actually in those libraries where you might bumble into people and bumble into things you never (knew were there.)" Subject S30 elaborates on how browsing allows scholars to conduct research on both current and future projects simultaneously:

"I was doing my research on the first project which included one of the community programs, which was a health clinic in one of the villages, and as I was going through the research I said to myself, 'I want to know what these records are like, what this collection is like so I can keep an eye out for a future project.' As I was going through the collections in the local archives, I started noticing there was quite a bit on clinic and hospitals and nursing, and I thought I am going to start keeping track of this stuff; this might be something interesting to follow up on."

5.2.2.3. Using items in private collections

A strategy that has not been cited in previous scholarship is that of using private collections. In this study, this unique strategy was mentioned by 25 participants – as many as said they use web searching. Because non-English and foreign sources can be difficult to find and even more difficult to obtain, scholars may opt to make or own a copy to make it easier. These may be sources that they purchase (in overseas bookstores, for example,) copies they make when they visit foreign archives and libraries, or copies provided to them by colleagues. They may also include both primary or secondary sources, and documents or artifacts in any language. Subject S46 gives an example:

> "I was looking for records and lists and stuff and things that, and even though they were in Dutch, they were quite easy to find… not necessarily online, I already had them from trips to Holland and so on. So if it wasn't for that, I would not have been able to find these. I had gotten them on different trips."

She then continues to explain how important this strategy is in her research:

> "I was going over materials that I already had, and some of them are quite simple. Some of them are just newspapers, but they were Nazi newspapers, and I realized how physically dependent I am on those copies because those are not available online. If I didn't have something in my records and didn't have the original, I would have to go to Holland to see it or somehow get it through interlibrary loan."

It is somewhat more common for classicists and scholars studying much older topics to have copies of seminal works or common reference books in their own libraries, while scholars who study more current topics – especially language scholars – are those that frequently make purchases in bookstores, as explained below in section 5.2.2.8.

5.2.2.4. Citation tracing

Tracing footnotes, citations, and bibliographic notes has been an effective strategy for decades, being mentioned in Bates' study also under the name of "chaining backward." Citation tracing was cited by 24 participants; it is used occasionally as an initial strategy but more often in

161

the middle of an information-seeking thread. Citation tracing is one of the most "language-proof" of all strategies in that the seed document may cite sources in any language from any country, culture, or location. Citations can include primary or secondary sources, and the primary benefit of citation tracing is that it can lead to the identification of many sources in a short amount of time.

Subject S15 looked at citations in works by her colleagues in the field: "I found it was really helpful to connect with my colleagues that do Brazilian history and not only on a personal level but to read what their research was. So I really cannibalized a lot of footnotes and bibliographies." Subject S20 gives another example:

> "I come up with an idea reading from the Latin or Greek text, and I will think 'Oh, I think I have seen this before. I think I will try to find a recent article published in a good, peer-reviewed journal.' And I try to get a good, substantive article from that in the past couple of years on my poem or on my specific topic and then look at their bibliography to begin to gather my resources."

Subject S20 continues on to describe one of the most fascinating aspects of citation tracing mentioned in this study – the clustering of source languages depending on the language of the seed article being used:

> "I've found that, when I read an article to gather citations, an English article will give me lots of English, some Italian, sometimes some German. But even when I think I've got it comprehensive and I'm reading my last German article out of Gnomon, (like fifty articles for every language,) that that article will cite two or three or four things in German that I have missed, and same with Italian and French, that each of the languages, the articles tend to include same language citations that I have missed for all my other searches – digital or otherwise."

When I asked for clarification about the reason for this clustering, subject S20 explained:

> "But there's also a reason for the clustering, sort of the underlying reason. Methodologically, there are differences in what kinds of work different countries do on these same texts. There are lots of exceptions and fringe areas overlapping like Venn-diagram circles, but Italians, at least in my field, tend to be primarily interested in either anthropological readings or in catalogs, and Germans are primarily interested in sort of philological things, looking at word usage and same

with the Spaniards. Most of the articles in Spanish in my field are analyzing word usage within a text or cross-text. And the French tend to be sort of either doing kind of shallow close literary reading kind of stuff or a theoretically high flying sort of cerebral analysis of texts."

When asked how the subjects know what archives have what materials that they need for their study, many of them indicated that they identified sources using citation tracing. Subject S45 explains:

"In the projects on Soviet and Russian history, there are only going to be a handful of archives that have material. Depending on what particular century or centuries that you're working on, it's going to be in one or the other archives. So sometimes when you read an article, you look at the archives they use to get an idea of what's out there. Citations give you a good idea of the name of the archive and the specific types of materials they might have in there and then you can build from there. Then you can always email someone, but these days you can just go on the website and you can just find whatever you need."

Sometimes, there may be a particular seminal work that provides a wealth of citations specifically pointing to archival collections, as subject S8 explains:

"For the Paris archives, there's a history of the mission in Sichuan that was written by the archivist of the Paris archives about a hundred years ago. It's very old-fashioned in its approach, but it does provide me with a lot of references to particular documents I want to look up."

Subject S32 describes a specific challenge related to using citations from older sources:

"I write about the late 18th century (in France,) and there are references to sources, to primary sources, eye-witness accounts, in 19th century books. But 19th century books do not use the citations that we use now. If they use references at all, they are often virtually impossible to track down; I have no idea what that citation means. It is often frustrating."

5.2.2.5. Consulting known web sites

What differentiates formal system strategies from informal resource strategies is not whether they are conducted on-line, but rather whether they use a search tactic or a browsing tactic. This means that consulting known web sites is classified as an informal resource strategy since the scholar generally browses the site. Marchionini (1995)

163

considered the use of hypertext to click links and move from web page to web page was a browsing activity, and that same characterization is used in this study as well. 22 of the subjects discussed visiting specific web sites that were known to them or that they discovered in their information-seeking process. This strategy is often used in searches for known facts. Subject S47 says, "There are a number of web sites that talk about digital Chinese history because there's a lot of stuff now that is coming on-line. So every year or so, I will go in and look at these web sites on-line." Similarly, subject S42 studies a specific French author and reports, "There's an organization for the study of Yourcenar in France, so I would just go to their web site, and they have listings of publications and conference proceedings and things like that that are pretty readily available. She's a pretty big deal, so it's pretty organized." Finally, of her study in ancient Greek texts, subject S19 says, "I look at book review sites in France and in Germany just as part of my looking for sources because some of the stuff that I work on in particular is still very popular with the Germans and the French, and now, increasingly, I'm having to do more work in Spanish scholarship."

5.2.2.6. Browsing archives in the United States

Not all research using foreign sources needs to be conducted in overseas libraries and archives. Many universities, organizations, and other libraries and archives in the U.S. have built collections of foreign and non-English documents to serve scholars at their own institution and across the U.S. Many of the examples in section 5.2.2.1 regarding browsing in nearby research university libraries also apply to archival research in the U.S. Visiting archives in the U.S. was mentioned by 21 participants. Subject S36 gives an example, "I'm going to Washington to work in the national archives and also the NASA archives for a new project on Soviet communication

164

satellites." Even though using U.S. archives was discussed by 21 of our subjects, it can sometimes be overlooked as scholars focus on overseas research. Subject S11 admits, "What I didn't think about early on was the sources that I had in the United States," speaking of her discovery of relevant video interviews in French at the U.S. Holocaust Museum.

5.2.2.7. Consulting newspapers on-line or in print

16 subjects indicate that they use newspapers – print or on-line – for their research. In fact, over time, consulting newspapers has involved a variety of media, from the original paper copies to microforms, to electronic on-line surrogates. While this is a common strategy mentioned by both historians and modern language scholars, not surprisingly, it is not mentioned by any of the classicists or the religion scholar in this study, as shown in Appendix G. Subject S31 provides an example: "So newspapers, I usually start with, from major cities like Prague, Berno, Bratislava, I start with those. As I dig deeper, organizational records and personal papers." Subject S32 explains that newspapers used in his research on French history can be accessed in any of three different formats:

> "I use a wide variety of primary sources. They include newspapers from the 18th century, printed material. The period I'm talking about is between 1794 and 1804/1806. At a certain point in my research I looked at them on microfilm. When I was in France, I consulted them on hard copy, and now a lot of them are available online."

Some of the subjects indicate that the only way to access the newspapers they need is to travel to an overseas archive. Such is the case with subject S45, "Often I'm interested in newspapers, and I just know I will have to go to Russia to look at the newspapers. The ones I want usually aren't available in the United States."

Subject S30 explains that a colleague had informed her that many African newspapers were available in the U.S., specifically at Northwestern University and the Center for Research

Libraries in Chicago, and this allows her to access them in the U.S. and reserve her time in South

Africa for other types of research:

> "For most resources I need to be in South Africa, but there are some microfilmed documents of newspapers and some government documents, there are some resources I got in the United States. The Center for Research Library had a pretty good collection. I've been to Chicago just to go there and read microfilm newspapers. I've ordered a number of microfilms from CRL and spent about a week at Northwestern reading South African newspapers."

Similarly, Subject S15 found Brazilian newspapers have been put on-line. "The national library

in Brazil, in Rio, has also digitized all their newspapers, and I can access their newspapers. So

there are (other) things that I am going to reserve to do when I'm not in Brazil."

Scholars who research contemporary topics – including faculty in modern languages

looking for documents for their students to read – may rely more on on-line news services and

newspaper web sites for current articles than on using newspaper archives. French teacher and

scholar S14 says, "I run into stuff every day in the press that is relevant to the things that I

teach."

5.2.2.8. Visiting and browsing in bookstores

A strategy that is decidedly unique with scholars whose topics take them outside of the

United States is that of visiting and browsing in bookstores. From Spain to Japan and nearly

everywhere in between, ten participants report unique, vibrant, and helpful bookstore cultures

that lamentably do not exist domestically, and tell how they browse in bookstores there. Subject

S34 tells of her usual strategy in Spain:

> "I will go to the bookstores, and I will browse and see if there are any new things. They have more kinds of bookstores in Spain than we have here – a lot more specialized book stores. If you're looking for 19th century things, they might have a bookstore that has a lot of old texts or manuscripts or copies and replicas of manuscripts where you can purchase it -- not only the library archive that you can visit. For example, in Malaga, Spain, which is in the southern part, they have a bookstore that has academic journals you can buy. So you can go through, if you

166

already know what you need, you can go through the journals and pick up the academic journal you need, which is a lot different from here."

Subject S36 conducts her research in Russia, extolling the value of bookstores there:

"When I was there, I would always go to the bookstores because they have the certain big capital city type major bookstores – particularly BiblioGlobus in Moscow - but also some of the special (ones.) Universities have bookstores there that are not quite like our university bookstores; they're really about publications by the faculty at those universities. I try to hit all of those when I'm in Moscow and buy whatever is relevant to my work. I've found a ton of stuff that I wouldn't have known about through doing that – the really more recent publications. Again, it's very serendipitous. I'm sure there are things that I'm missing, but it gives me a decent sample."

Despite cost factors of purchasing books in this way, subject S44 relies on bookstores:

"What usually happens is I have to wait until I'm in Vietnam again, and in Vietnam there are some second-hand book vendors who will hunt for things for you, and they charge a great deal to do this, but that's a good way to get something you don't have access to in the United States."

In addition to bookstores, scholars who use film in their work rely on video stores in the same way. Subject S25 finds these sources particularly invaluable for films and explains how this strategy involves a time-consuming browsing tactic:

"When I went to that conference in Buenos Aires in November, I spent days walking around that city and visiting I don't know how many stores, probably 30 or 40. And… Latin American cinema as an industry is so dominated by the Hollywood machine that locally made films are not usually considered commercially viable or important, so they kind of get eclipsed even locally in these video stores. So I have to go to very specific video stores that only have or that specialize in national films, and there's no electronic database; there's no way of finding out ahead of time who has what film. You have to physically go into each of these little shops and look through stacks and stacks and stacks of DVDs, and it takes hours. But I have found many films from the 50s on that are considered important fundamental films for Argentine cinema, for Mexican cinema, that come up in academic literature about those topics all the time, but they're almost impossible to obtain commercially from the U.S."

5.2.2.9. Consulting archival finding aids

In section 5.2.1.5, we discussed searching finding aids as a formal system strategy. However, in many cases, the finding aids in archives are not searchable but rather a static list. These might be in a word processing document on a web site, or they may be a print list or bound volumes only available when visiting the archive itself. An archival finding aid is similar to a print index – a different informal resource strategy described in section 5.2.2.10. Ten of the participants – mostly historians – in this study talked about the use of static archival finding aids. Such is the experience for subject S44, who conducts research in Vietnam:

> "There's a finding aid for the archives that I use in Saigon, but it is something like eighty bound volumes that have been put into a word processor, but they have been printed out and they are in more or less random order. That is to say that they are not categorized by provenance or year or contents. So you just have to start with volume one and read all the way through. It would be good, if the Vietnamese were willing to share either the Word files that they're based on or provide web access to the computer catalog in the reading room."

Subject S36 consults a print compendium of finding aids for various archives in Russia:

> "The main resource, still now, that got started I think in the early 90s was to publish the catalogs of the Soviet archives in editions that you could access and that libraries could buy. There is actually a physical set of volumes of all finding aids, not specific, but of the collections, the subdivisions of the collection, the names and addresses of the actual physical archives, who wrote them, stuff like that. It is very comprehensive and excellent."

On-line finding aids are especially helpful in planning for archival visits and making sure that the trip will be worthwhile. In this case, subject S31 was looking for sources related to the Studentsky Domov in Czechoslovakia, and she explains how she was able to use WorldCat to locate an on-line finding aid, and then her use of a static on-line finding aid led her to shift to the interactive human strategy of contacting the archivist:

"Searching them through WorldCat let me see where their archival holdings are. I found them at (Yale) Divinity Library, and then if you go to Divinity Library, they've published all their finding aid stuff online now, which is great because I don't have to search through a bunch of random catalogs. I can just literally look at their finding aids. I found that they had a few bullet pointed listings in one of the boxes for Studentsky Domov. I emailed the archivist, 'Is this legit? Are those there? What is this? Because I want to come see it.' The archivist was really helpful at facilitating my visit. So all the web sources worked the way they were supposed to work to get me there."

The quality, completeness, and organization of finding aids can depend on many things, including funding, time, available technology, and the pride taken by that library or culture in its archives. Since some of the earliest archival standards were pioneered by the Dutch (Cook, 1997), the observation by subject S46 is perhaps not surprising:

"I tend to rely a great deal on finding aids. The Dutch create really detailed, very accessible finding aids in PDF form that I will often be working with simultaneously as I am working through the collections. I will download them and go back to them repeatedly. Some are more in detail and more helpful than others, but for the most part, I have found really cohesive and helpful finding aids, far more helpful than anything I have seen usually in American or German collections."

Despite the possible value of on-line finding aids for planning purposes, subject S35 cautions that they may not be as complete or helpful as the finding aids in the archive itself:

"In Turkey, a lot of these major repositories, their documents or manuscripts have internet sites, and you can go to them. What I've found is that they don't work very well and they don't have everything that they might have. For example, I went three years ago to Turkey and to the main repository of the documents, even though they had an online guide, it wasn't until I actually got to the archive itself, that I actually got the guide in the sense of that it was a more detailed guide. There's a certain amount, there is a lot you can do before you go, but there's some things you don't know until you're there."

5.2.2.10. Using a print index

In many cases, scholars fall back on the use of a print index to help them to identify and locate sources in libraries and archives. This is similar to the strategy of using an archival

finding aid, except that a print index may describe an entire library collection or even resources

not locally held. For scholars using non-English resources and/or collections in other countries,

print indexes are fairly prevalent, especially where digitization and automation are not yet

common. Print indexes were mentioned by ten participants, including three of the four scholars

of ancient texts in this study. Scholars who conduct research in Asia often report the use of print

indexes. Subject S39 explains his experience when using the National Archives in India: "There

are printed or handwritten indexes to the collections. Your search mechanism is leafing through

those; it's not at all mechanized." Subject S37 adds "For me to have some of these

compendiums – bibliographical compendiums – that I needed to actually identify books, works,

people. They were all at the Weidener – like 20 volumes, and I can't really order that through

inter-library loan."

5.2.2.11. Journal run

A journal run occurs when a user browses multiple issues and articles within a specific

journal title previously identified as useful. Five participants in this study talked about the

journal run and it is efficacy. This is also one of the strategies cited by Bates and shown in Table

7.1. Subject S40 explains a common phenomenon where there is a particular journal that is

recognized as the most prestigious in a particular field:

> "The *Journal for Scandinavian Studies* is the biggest journal here in the United
> States for the field. In Scandinavian the journal is *Edda*, and that is a fabulous
> journal as well as the *Journal of Ibsen Studies*. These are the main places that I
> make sure to keep my finger on when thinking about research."

Subject S16 studies German literature and explains how he browses through particular titles

rather than searching:

> "I went straight through our Subscription to Journals portal at (my institution.) So
> instead of going through WorldCat, I went through our journals list. And some
> are like Oxford journals or Taylor and Francis. We're a subscribing institution, so

instead of burrowing down through WorldCat, I just go straight to the journal. So I went to *Holocaust and Genocide Studies*, and *Central European History* and *German Studies Review* and a couple other places like that. I just look through the last three or four years of their tables of contents. So that's a sort of not-searching, really, but that would be like browsing, I guess. I'm browsing through the table of contents of the last several years of, you know, half a dozen journals."

While the journal run most often occurs using a journal title that is familiar to the scholar and commonly used in the field, there are times when serendipity plays a role, and the scholar finds a previously-unknown and useful journal. Subject S45 explains how this can happen in his studies of the Jewish experience in Russia:

> "Sometimes serendipity, you just come across material you didn't know existed because you're looking at one particular source… I did one project where I had a student who knew Russian who looked at a particular journal from the 1920s with a lot of illustrations, and it was very intriguing, so I decided that one day when I was in Saint Petersburg to look at the same journal to see what was there and I was stunned by how much material I could find, so I just read fifteen years of this one journal looking for a particular kind of illustration."

5.2.2.12. Chaining forward

Chaining forward is defined as "finding sources that cite the current source." It is exactly the opposite of citation tracing, which can often be called chaining backward – so named because the sources go back in time whereas chaining forward helps the user identify newer sources. In both of my prior studies, this strategy was never mentioned. It may have been more common in a previous time when print citation indexes were more prevalent, and many of today's on-line tools do not provide this capability, or scholars do not know how to use it within tools where it does exist. In this study, only subject S2 used chaining forward. She had been able to find a Washington Post article that talked about cultural appropriation as related to Katy Perry's geisha routine on the 2013 American Music Awards. She was then able to find posts on Tumblr and Twitter that referenced that Washington Post article.

5.2.3. Interactive Human Strategies

The third type of strategies on the information triangle is interactive human strategies. As defined in section 3.2.3, **interactive human strategies** include those that rely on a method of consulting or asking with the entities always being human. These may include archivists, librarians, colleagues, authors, or other knowledgeable humans, whether directly, passively through conference attendance or blogs, or through electronic means. Subject S45 explains:

> "You even just go in and talk to the archivist and say, you know, you want to check out this. Or I've had librarians say, 'You want to go to this archive and find this material.' It doesn't even have to be colleagues in the west; it could be professionals, even Russian colleagues as well who will tell you where you can get access because they have more experience working in the archives."

One of the unique properties of interactive human strategies is the interactivity itself. Unlike either formal system strategies or informal resource strategies, when another human is involved, that human may actually initiate contact with the researcher, make a suggestion of a helpful resource, or otherwise facilitate and advance in obtaining a source all without being asked to so so. The largest number of examples are given in regard to archivists, who are often credited with making unsolicited but helpful suggestions, as Subject S30 states:

> "I had a day or two, just two days at this archive and the archivist knew what I was looking for and said 'Did you know we had these records?' And I had somehow missed them or hadn't searched for them, and they were really valuable in getting me the information I needed."

An archivist may also be able to suggest other archives where the scholar can find helpful sources, as subject S8 demonstrates:

> "For example, when I first used the archives in the Missions Étrangères in Paris, the archivist mentioned to me that the Lazarist Archives nearby had things that I would be interested in, so it was just that kind of personal connection. And the Lazarist Archives are not as open as the Missions Étrangères. So having a kind of a connection, a reference, from the archivist of the Missions Étrangères gave me access to the Lazarist Archives in Paris."

Subject S15, in her pursuit of Brazilian history topics, expresses her concern about loss of human contact as sources are digitized:

> "I know that there's this push to digitize and to make stuff accessible but, wow, when you lose that human interaction… Those folks are experts; they know that stuff like the back of their hand, and so it's really helpful. 'Hey have you thought about looking here? Have you thought about looking here?'"

In addition to archivists, it may be colleagues who take initiative to suggest or obtain sources, as subject S35 explains:

> "I was tracing a footnote, and I (thought), 'This is about the individual I'm trying to trace and here's this really good source and its only located at this one archive in Turkey, and that archive happens to be closed for several years for whatever reason… I can't get to that.' I was talking to another friend about this, and she said that it had been published, so I didn't even need to go to the archive. The footnote I was going from was an old enough footnote… so I was able to look it up in a library catalog, and I checked it out. Some things you just really have no way of knowing unless someone tells you."

Nine different interactive human strategies were mentioned a total of 88 times; each of these is discussed in more detail below.

5.2.3.1. Consulting a colleague or other expert

The most common of all interactive human strategies is consulting with a colleague or other expert in the field, mentioned by 19 participants in the study. These contacts with others may be in person, by phone, by email, or by social media. (Note that this strategy involves a targeted connection with a known person, as compared to strategies of using blogs or social media, as discussed in sections 5.2.3.7 and 5.2.3.8, which involve an untargeted communication to a group of people in hopes that someone will be able to provide assistance.)

This strategy is most often used at or near the end of an information-seeking thread when other methods of obtaining a source have failed. Subject S44 explains the crucial benefits of

keeping in contact with other scholars and colleagues who conduct similar research, especially in

advance of overseas research trips:

> "I think the best thing is to identify someone who has done something similar and talk to them. The only reason I… was effective at all when I got to Vietnam or that it wasn't much harder than it was, was that I had been fully warned and prepared by people who had been there previous years, so they had recent experience and had a good knowledge of how things work. Every time I go back I try to keep in touch with graduate students who have been in Saigon, even saying that I am going to be there and they had been there earlier, just to see if procedures have changed, because they do change these things and don't post it online. So you can save yourself a lot of headache by talking to someone who will tell you the library is closed on Fridays so you don't have to go those days. So it is important for people to be in contact with people who have been doing research very recently."

Subject S42 describes the value of having a community of other scholars that can help to provide

leads and assistance:

> "When I have spoken with people who do similar things, not necessarily in this context, it's just been like a wealth of information, like I didn't think of that, or suggestions, and that's just extremely useful. I think also just having more of an intellectual community away from your place of employment is nice and invigorating and gives you more energy to maybe extend yourself more than you think you can do given your time and energetic strengths."

A third crucial role of colleagues is to assist scholars in obtaining documents that they have

already identified, as subject S8 explains:

> "And one of the challenges, of course, was that I was here, and the journal was in the archives in Paris…But the archives are not yet online, and so, going to Paris for a weekend just wasn't practical. It did work out; (a colleague) actually sent me copies of JPG files--photocopies of the parts of the journal that we worked with."

Subject S41 provides another example, based on his research on Chinese economic and political

history, including research on flour milling in China:

> "I was also lucky in that the original mill that imported this machinery from Milwaukee was still in existence so I went there to look at their archives which they wouldn't allow me to do, but the academy of sciences in Shanghai had somebody who was actually researching this flour milling, as well, and he provided some information to me."

One form of collegial human interaction is working on a research project with a colleague and helping each other find sources, as subject S36 demonstrates:

> "I'm working with another scholar. It's a joint project… and my role is very much as the Russian speaker because he is Swedish and works primarily on western European transnational broadcasting. He's written about Russian things but he doesn't have Russian, so my job is to be the one who digs through the Russian archives on behalf of this project, because I can. He, however, has been working on transnational broadcasting satellites for a while, so a lot of the secondary work he tends to do... So we share articles over email and suggest things, sometimes he has actually mailed me books from Sweden which is kind of silly, but on the other hand sort of easy."

Other forms of human interactions with colleagues and experts in the field are more serendipitous, happening through casual conversations or informal meetings, as Subject S32 relates:

> "I have many examples of the human element. A good friend of mine who is a historian of the revolution in France came to visit… and stayed with me. When we were talking about one of the guys who I was writing about who was put on trial, she called my attention to that Jean Paul Sartre, a French philosopher, had begun to write a play about this guy but never finished the play. I had no idea that existed. I took a note, ordered stuff through interlibrary loan and it turned out the notes of this unfinished play by Jean Paul Sartre are housed in the Harry Ramsey Center at the University of Boston."

5.2.3.2. Attending a conference

17 scholars in this study talked about conference attendance as a helpful interactive human strategy used in their research. While it is generally not the goal of attending a conference to either identify or obtain sources, that is often a benefit of interacting with colleagues and others in the setting of a conference, as subject S17, a dance historian, says, "I make it a point to go to conferences in my field and in that way, word gets around and someone comes up to you and says 'Do you know about this article? Because they know it's your area."

Subject S38 provides a specific example of how conference attendance led to an opportunistic

shift to consulting a known web site:

> "I have been working on the French Algerian book, and I realized how important
> going to conferences is. You meet so many different people… I was having
> dinner on the board of directors, and I was sitting next to this Algerian, and I
> started talking about Bennabi with him because I was wondering if there was any
> sort of memorium on the day of his birth, and he said that Bennabi is too
> controversial for people to pay a lot of attention to. But he talked about this
> publishing house in Switzerland who see Bennabi as an influence, which was
> something I was unaware of. So I went online and found this place – another
> place to investigate and research – and it was a pretty good site."

5.2.3.3. Asking a librarian or archivist

A common interactive human strategy is to ask or consult with a librarian or archivist; it

was mentioned by 14 of the participants. It is most often used in conjunction with the related

informal resource strategies of visiting a library or archives, but not always, as contact can also

be made via phone, email, or even letter to get assistance from a professional. Most of the

participants in this study are from smaller private colleges, and they indicate that their own

librarians are quite willing to provide assistance, as Subject S13 explains in his research in

German migrant culture: "I knew this book had to be available online. It was, but it took a

reference librarian looking with me for a while to actually find the thing just because...it was

hard to figure out how to search for it." Subject S11 explains that the information-seeking

expertise of her university librarian was very helpful in her research on various French topics

despite the librarian's lack of knowledge of the language and how important it is to collaborate

with others to find the most helpful sources:

> "It was really hard to find articles or publications talking about this, and it was
> hard to find and know what search words to use. Finally, we went over to the
> librarian who, like I said, doesn't speak French, but she's specialized in her field
> and knows about search words, and she was incredibly helpful. So, we all sit in
> front of our computers, myself often late at night, and I think we forget about, the
> computer just does not lead to collaboration, it leads people away from

176

collaboration too often, and I think that human collaboration is just huge, and we shouldn't forget about that part of it."

Sometimes, the librarian being consulted is not at all local to the researcher. Instead, it may be a librarian in a remote, overseas location who has knowledge of a source and its location. Subject S3, in his research on philology, gives an example:

"I did this years ago with a lost Latin poem that by sheer luck of fortune, I happened to track down. The reason I took that poem is because I happened to read in one of my resources that it would be nice if we had a decent edition of this thing. There is one manuscript that's been lost since the war. Its presumed gone or last place was… So I just wrote and said, 'You guys ever heard of this?' They said, 'No, but I've got a librarian friend at the Uniwersytet Jagielloński in Poland. She might be able to help you.' I wrote, and she wrote back and said, 'Oh yea we have it.' (So I asked,) 'Can I get a copy?'"

Despite the appreciation of the subjects in this study, King (1994) found that scholars, especially in history, can be skeptical of the knowledge and expertise of their librarians, but they more often have high praise for archivists. Subject S30 tells how instrumental archivists are in identifying helpful resources in her research on nursing in South Africa, "I would say that, in my case, in the beginning, it was more asking other people, and they would be pointing me to databases, or talking to archivists, and they would tell me what was in the archive and what I might be looking for. Archivists are extremely important in that sense."

Subject S29 further explains how an archivist's knowledge of their collections makes them a powerful resource:

"On arrival, the archivist turned out to be much more helpful than the online records. Though the archivist was not closely familiar with the individual periodical or its contents, she knew that it was part of a larger collection at Fordham of such periodicals. This was not indicated in WorldCat, of course, or even in the Fordham online record, but the archivist knew this and could therefore bring out boxes and boxes of similar periodicals from the 1920s in Irish."

Similarly, Subject S10 tells how proactively helpful an overseas archivist can be where she conducts research in Germany:

"For a researcher coming here all the way from America, they're often extremely helpful. I had a guy in one archive who was practically trying to do my research for me. I would show up, and he would have these documents pulled out and marked already of what I might be interested in."

5.2.3.4. Using social media to contact people

A new strategy that was not discussed in my prior two pilot studies but was mentioned by nine participants in this study is contacting people on social media, primarily Facebook. This is distinct from the previously mentioned formal system strategy of mining social media in that there is no searching involved but rather the goal is to establish human contact through social media venues. It is also different than contacting a colleague or expert in that a posting on social media is an untargeted contact not aimed at a specific person. This strategy seems to be applied most often for convenience and when the researcher does not know which colleague or expert will have the needed source or information. Subject S6 aptly explains:

"I find more and more of my friends doing this. They post on Facebook. We all have friends who are scholars in our own fields… Just the other day, a friend of mine who is in Vienna and doing research said 'Does anyone have a copy of this one thing? I need to know on what page something is.' He hadn't written that down when he'd see the original. Within 15 minutes, someone had answered and said 'I have it. I'll send it to you.'"

Subject S19 has a similar strategy, and subject S34 explains how she leverages conference attendance in conjunction with multiple social media outlets:

"I use Facebook to connect with other colleagues or peers of mine that I have met at conferences and stuff like that. 'Do you know the source for that?' Or LinkedIn, if I don't have them on Facebook, I will go on there and message them. Email of course, to find sources for ideas when I'm researching."

Subject S37 indicates that there is a Facebook group on-line for scholars in his area, and subject S36 also says, "There's a Facebook group of foreign scholars doing research in Moscow where you can ask questions."

5.2.3.5. Contacting an author or family members

Sometimes, the easiest way to obtain a known source is by contacting the author directly. This strategy is used in many of the same circumstances as contacting a colleague and is often a strategy of last resort when a source is impossible to acquire using any other strategy. Subject S19, a classicist, states simply, "Sometimes I know the person who wrote it, so I just email that person and say, 'Can you get me the article?'" This interactive human strategy, mentioned by eight participants, can lead to results that exceed expectations, although finding out how to contact the author may not be that simple, as the example from subject S32 illustrates:

> "I read a book by another woman who is not a scholar, in French. Her ancestors… one of them was guillotined during the revolution, and she wrote a story about her family and the burial place where her ancestor was buried during the revolution. I… didn't know how to contact her, so I contacted her publisher who put me in touch with her. She ended up inviting me to her house in Paris and she gave me permission to use her private archives. It includes letters, diaries of her ancestor… She was kind enough to do that."

Sometimes, this strategy involves contacting family members of an author, artist, collector, etc. Subject S18 describes a series of creative steps that she took to find the family members of a Japanese photographer and request access to their personal collections:

> "I found (sources) by very primitive searching techniques in the photography museum, in their database, literally by searching the period name and photography studio. I came across a catalogue of that studio in the faraway location, and then I actually went to the Yellow Pages of the phonebooks in the public library and sent a letter to the descendant of the studio founder. Just to give you the chronology, the studio was founded in 1870, but it's pretty interesting that the family members are still alive; one of them is still a photographer. I sent a snail-mail letter, and he invited me to come down there and gave me the limited publications that had been produced about the studio, all of which had been produced by the local prefectural government, so they included things that otherwise I would not have known about… That personal connection is actually what facilitated the project."

It is significant to note that, almost always when this strategy is mentioned, it is because it was the final strategy and resulted in the scholar obtaining sources they were looking for and sometimes additional sources of which they were previously unaware.

5.2.3.6. Learning from a mentor or senior scholar

A special case of contacting a colleague is the strategy of learning from a mentor or senior scholar. This strategy, mentioned by six participants, could possibly be combined with contacting a colleague, but the relationship of advisor or mentor seems to be special. In the case of a mentor, the scholar is very familiar with what the senior scholar knows, and he/she may come to rely on the mentor not just to recommend or provide sources but also for general direction and advice on how to develop research strategies. The relationship is more instructional and less likely to be sporadic. In fact, advice from a mentor is more likely to influence the direction of a scholar's research over the long term. Subject S10 speaks of consulting with her dissertation director as a dissertating grad student, and subject S3 talks of the influence of his "Doktorvater" (Doctor-Father or advisor) during his graduate studies. As already quoted in section 5.2.1.4, subject S31 relies on "some of my former advisors who also work in the region" for assistance obtaining sources. When asked what advice he would give to younger scholars, subject S16 replied, "Find a mentor. Find somebody who can show you some show and tell. Don't try to learn all that stuff by yourself; watch over somebody's shoulder while they show you how to download and store index webpages. That's so efficient later on."

Once the mentor relationship is established, it may be the mentor that contacts the scholar to provide leads on helpful resources, as subject S30 explains, "One of my former professors at Michigan State University, who is also one of the Afrikaans librarians there, he is always looking at what's coming through the library and letting me know if he finds something related to my interests, and that has been extremely helpful."

5.2.3.7. Using forums, listservs, or blogs

Very similar to the social media strategy, six subjects indicate that they use other more traditional on-line communication tools – email listservs, blogs, and on-line forums – to contact colleagues and other human sources in their information-seeking process. As with social media, human contact through a forum, listserv or blog is non-targeted, so the researcher does not need to know who has the information they seek or how to contact that person. Subject S24 discusses available options which admittedly blur the lines between available communication tools:

> "For example, the Venetians have two different listservs. One is a Facebook listserv, another one is a Yahoo group, and so if people have questions, and even people outside of the group. For example, somebody wrote from the Smithsonian or somewhere saying 'I'm checking the label for an exhibit, can you tell me in the big basilica, what are the figures on the mosaic?' and had a series of questions about what was in the mosaic, and within about six hours had expert scholarly answers to all of the questions. That's a huge, huge deal when you're wandering into a new area."

5.2.3.8. Using a research assistant or surrogate

Some scholars – especially from larger universities with graduate programs – have access to a research assistant who may be able to search for, identify, or obtain sources for him/her. Each one of the five participants who use this strategy have another person go to an archive or library in a foreign country on their behalf. Most often, the envoy then transmits the item back to the scholar electronically, opening possibilities for obtaining sources that were previously impossible. Subject S44 provides an example: "Sometimes I will use a research assistant who can locate things in Vietnam and either digitize them or email them to me." Similarly, subject S30 recounts, "There were some European organizations that funded some community programs in South Africa, and I thought that I could not myself travel to the archive. I knew they were there, there was this catalog online, but I ended up paying someone to get those for me and send them to me."

The vast majority of subjects in this study do not have access to a research assistant, as subject S6 explains, "I couldn't get a research assistant because they don't know the languages, and that is just so common in my field." However, three other scholars (S29, S35, and S41) talk about sending someone else to a specific location to retrieve a source for them. Subject S41 makes arrangements with colleagues on an exchange model:

> "The challenges are getting access to the material. Like you said earlier, I can identify them, but then I have to go to China and find them. I regularly go back and forth, but more often than not, it's on my own time. That in itself is one of those limiting factors and that's why I rely more and more on people who are there, and have them work for me, and I do the same for them."

Subject S35 relies on a friend and colleague on location in her field:

> "From reading somebody's footnotes, I knew that there were two versions of a particular famous Turkish Ottoman source and the more finished version had been published, and I had a copy of it, but the earlier version hasn't been published, but there are different manuscripts of it. I happened to have a friend in Turkey; she was working at the particular library where these manuscripts were located. I was able to identify what particular section I needed so she didn't have to look through the whole thing. She was able to locate what I wanted. They just made a digital copy and then she emailed it to me. It was that personal intervention, because I could never had done this from the United States with the people at that particular library in Turkey to get what I had wanted or have a person there to do it."

Subject S29 explains how technology and changing policies have made this strategy more valuable than ever:

> "The other thing that's been helpful in the last five or eight years, more and more repositories and libraries are willing to let people photograph sources. And so now you can ask a colleague to go into a library and photograph the 5 or 6 pages that you need, and libraries are willing to grant permission for that. And I think that's great because it both preserves the source - it doesn't have to leave the library – and if the photographing can be done in a non-intrusive way, it's great. It allows for consultation from afar."

5.2.3.9. Interviewing human subjects

The final interactive human strategy, mentioned by four participants in the study, is to use human subjects directly as sources for their study (as this dissertation does.) This strategy is not

particularly common within this study since many participants' topics are too old to have living witnesses, or their topics are literary. In prior pilot studies, scholars in political science or social sciences were more likely to use human subjects than were humanities scholars (Sabbar, 2013; Sabbar, 2016). Subject S38 provides one example of using human subjects:

> "I got in contact in Washington with one of Bennabi's former students. So I would go to Washington for meetings and then I would try to find and meet with these guys, and I did this multiple times, so oral interviews and research, verbal research became a big part of this research."

5.2.4. Hybrid Strategies

Hybrid strategies are those that do not categorize simply into one of the three sides of the information triangle. Their methods and/or entities are ambiguous and somewhat unpredictable. The two strategies gleaned from this study in this category are using inter-library loan and buying a copy on-line. For example, the inter-library loan strategy might consist of filling out an on-line form, clicking a button when using an OPAC, filling out a paper form in a library, or contacting their ILL librarian. The delivery of the item might be physical or electronic. When purchasing a copy on-line, they may use search techniques, but the transaction is not really search based, nor does it utilize browsing or human contact to any great extent. While there are not many strategies in this category, they are relatively widely used.

5.2.4.1. Using Inter-library loan

With 28 participants mentioning it, inter-library loan is the most frequently discussed strategy in all of the categories combined. It is also notable that ILL is the most often used final strategy shown in Table 5.3 since it almost always results in obtaining a resource. It is never an initial strategy since ILL can only be used once a source has been identified; it provides no assistance in the search and identify stage. Subject S25 provides this example of how effective it is in obtaining sources:

183

"So the Spanish language sources that I ended up requesting through inter-library loan – they all materialized within a week. So those were good, and I actually got yesterday an Italian language source as well. One of my… I think of the four inter-library loan requests I put in for this search, I think that two were in Spanish, one was in Italian, and one was in French. And so I'm just missing the French one at this point."

Many of the subjects had high praise for their librarian's abilities to obtain sources on their behalf via inter-library loan. Subject S16 says, "The most important person in my scholarly life is my interlibrary loan person. I mean, he's just like a magician. So, given ILL and in-house support and funding, you know, I don't really have to go anywhere. I can get everything I need."

As often as the subjects of this study use ILL, and as pleased as they are with its efficacy, they are well aware of its limitations, including the fact that it is not available in Europe (and most likely other regions,) as subject S46 explains:

"European archives are especially difficult. If they don't have a book there you are kind of out of luck unless you go there yourself and physically get it. Sometimes libraries aren't accessible without a membership or if you are not of that nationality, so it's never quite as simple as it should be to get materials in Europe as it is in the United States… I think in Holland at least, and this goes to Europe, too, as a whole, it's so impossible to do interlibrary loan. Interlibrary loan in the United States is so easy, and it is so difficult to do in Europe."

Another limitation of ILL is that not all libraries will loan the documents that they own, whether because they are rare, fragile, etc. Subjects S35 and S39 each identified useful documents from Turkey and India, respectively, in the archives at the University of Chicago, but neither was able to obtain it through inter-library loan. Subject S35 says,

"These documents have been published and printed, and libraries own them, but even then most of those are non-circulating. You can get them in the United States as opposed to going to Turkey. Like the University of Chicago has published documents, but they don't circulate. You could not get them through interlibrary loan."

And subject S39 relates a similar experience:

"I found through WorldCat a listing for one of these. I think this particular one was an early 19th century Persian version of this horse-buyer's guide which is held by the University of Chicago, and because of its age, they wouldn't lend it, so in order to actually use it, I had to physically be at the library to go in there and look at it."

Subject S6, in her research on southern France, had a comparable experience with books from Harvard's collection and more generally with foreign books:

"I actually did go to Harvard quite a number of times for my dissertation to try to get stuff. And I was generally able to get it. And sometimes it was stuff that they couldn't get for me on ILL. Harvard has it, but they won't loan. And that's typical. It's very rare that they'll lend out a book, so that's a real challenge. Of course, the books that are published in foreign countries are very often not available on ILL, so that makes it hard."

Subject S13 expresses issues both with loan limitations and the fragility of documents:

"Although I think even there the acquisition politics are probably different than they were 10 years ago. But you know, it is sometimes hard, and some of the stuff is obscure and rare, and they don't like to lend it out. That's different for contemporary materials versus 19th century materials. The other issue is that the 19th century was the golden age of print, because they discovered cheaper printing technologies, but they also discovered wood based paper which deteriorates. So that has, you know...that made some things difficult."

Another problem related to ILL borrowing is the time limitation as expressed by subject S6, "ILL is really good, but of course you can't keep things for very long, so especially if it's something in a foreign language, especially if it's in Latin or it's a very long volume, you really are in trouble."

5.2.4.2. Purchasing a copy on-line

The second hybrid strategy – purchasing a copy online – was mentioned by nine participants in this study. It is often used when the researcher would otherwise visit a bookstore but travel is not possible; or it may be used when other strategies such as visiting a library or requesting the item through inter-library loan fail to produce the source. Subject S18 explains how buying a copy figures in to her strategies:

"A second strategy is to actually purchase the thing - if it's a journal or a book, to try to purchase it myself. Japan is really very well-developed in terms of bookshops and secondhand bookshops, and there is a website that is kind of like Abe books in the U.S., I guess; there's a website in Japan that collectivizes… you know, many different secondhand bookshops will list their goods through this one particular site. They will not all ship overseas, but some of them will. So if I think that I really need something, and I can't find it, sometimes I do that if I can't wait until the next research trip to Japan."

If immigrant populations for the target culture exist in the U.S., this may facilitate purchase of materials in that language or from that culture, as subject S36, who does research in Russian, explains, "I bought DVD collections that you can get through Brooklyn-based book importers that serve immigrants."

The primary obstacle to the effectiveness of on-line purchase are cost factors; subject S25 compares costs for the sources he seeks in central and South America, "It's 30 bucks to buy the book there. Amazon will charge you $300 to order it from the US." Another barrier that can arise to disrupt on-line materials purchase is bureaucracy and currency, as subject S40 expresses,

"You can order them online, but the customs issue of, like, trying to get them here, and, you know, they don't take American credit cards, and so, like, who's going to pay for them? … Actually last year, I tried to order 'Frozen' in Norwegian. I teach a class on Nordic myths and fairytales and folktales, and so we did a myth analysis of Hans Christian Andersen stories and then watched 'Frozen,' and I ordered it, and I have a Norwegian bank account, so I paid that way, and it got caught in customs for six months, and then they ended up never actually sending it."

5.2.5. Initial strategies

Of special significance in this study are the strategies that participants choose to use first. Initial strategies are selected by the participants based on their research expertise because it is believed that these strategies will be most effective in quickly leading to success. When a scholar chooses an initial strategy for their research, they anticipate that they are and will be in a routine situation – following a series of one or more strategies to obtain a desired source. Table

5.3 in section 5.2.6, below, provides counts of which strategies were used as initial strategies in

the diaries. Of those information-seeking threads portrayed in the diaries, the strategies most

commonly chosen as initial strategies are the formal system strategy of searching an OPAC and

the informal resource strategy of using private collections. Nearly as popular is the formal

system strategy of searching the web using a web search engine. Overall, formal system

strategies are much more likely to be used first; in this study, formal system strategies were

chosen as initial strategies 24 times, compared to 15 times for informal resource strategies and

three times for interactive human strategies. As explained in the discussion of inter-library loan,

hybrid strategies are rarely if ever used first because they do not help in the identification phase

of the information-seeking process. Table 5.3 will be described further in section 5.2.6 which

discusses final strategies.

In addition to the data gleaned from the diaries, the participants were also asked during

initial interviews about their first steps in identifying and obtaining sources for their research.

The quotations below provide a sampling of answers from the interview data.

As stated above, formal system strategies are popular starting places for scholars in this

study. Subject S20 tells how he uses three different formal strategies for any given task:

> "I sort of have three different categories of sources which I start with, which will
> vary depending on the project, especially by how I come to the idea. In any
> project, I will always do all three. One is sort of formal and informal digital
> search tools: Google, Google Scholar, and nowadays Academia.edu can be
> helpful, or at least I can get some texts on just a Google search that can be useful,
> so that's the first one which is just the general access."

Subject S33 prefers to start by searching in the WorldCat OPAC, explaining that it will lead

conveniently to the next strategy: "The first place I would probably go would be to WorldCat to

see what's out there, like what exists, and WorldCat's nice too because it will tell you what exists

and what of those items exist at your own library, and then you can also, if it's not there, you can

187

do interlibrary loan." Subjects S41 and S42 search scholarly databases first. Subject S41 records the following in his research diary:

> "I found a reference on Jstor to an article on Qing whose bibliography listed a relevant dissertation and primary sources in Chinese of imperial correspondence. Jstor is a go-to first step for information on a specific topic. Articles either provide the information I seek or lead to other relevant/useful works."

> Three subjects – S24, S29, and S45 – explain how and why they begin with citation

tracing. Subject S24 says of her research on Renaissance Italy:

> "In the very beginning, what I usually do is go to the published work of the best scholars I know who use archival resources, because some people do and some people don't. In fact, I very often look first for historians rather than literary people because back in the early 20th century it was much more common for literary critics to use archival resources, particularly for these far away historical periods. But lately that hasn't been so common, so if I'm looking at more recent research I actually can more often tend to read works by historians that use archival sources just to see what they've come up with."

Subject S29 elaborates on how citation tracing from secondary sources will lead to relevant

primary sources on Irish immigrants in New York:

> "The first step again would be that you're kind of thinking of a topic and you think there might be some primary sources or you're not sure, you would start with the secondary literature. Maybe you are aware of a couple of studies that have been definitive on that subject. You would read those thoroughly and you would really pay attention to the bibliography and the footnotes, and you would make a complete catalog of what you had looked at and maybe you would get a sense of what primary sources they had been able to cover, and then you would go to their sources – their primary sources – and do that as a second step."

> Subject S29 and S46 both start directly with the informal resource strategy of browsing in

archives. Subject S29 explains how working with the primary sources first can eliminate any

bias that might exist in the secondary sources, but he goes on to explain that this requires that the

scholar know what archive(s) contains the sources that he/she needs. Subject S46 begins in

archives she is aware of in her target country:

"Most of my work is in Dutch and is held in Dutch archival collections… Mostly, because of the nature of my collections, I don't wind up researching them or finding out where they are via some central database. I have a number of Dutch collections where I know my materials will be found, and I usually start from there. I don't know of a general search engine that allows me to look at everything in Dutch collections."

5.2.6. Final strategies

Table 5.3, below, provides counts for the initial and final strategies as represented in the collected diaries. 43 different diaries were collated into information-seeking threads, and that data is provided below. Note that not all strategies are listed in the table if they were not represented in the diaries as an initial or final strategy. Another important note is that not all of the information-seeking tasks represented in the diaries involved obtaining a resource or document. In some cases, the task was a known fact search such as finding the name of a person. This would explain why strategies such as searching an OPAC, which don't normally lead to obtaining a source, can be a final strategy.

Formal system strategies appear as final strategies a total of 13 times; informal resource strategies 17 times; interactive human strategies four times; and hybrid strategies eight times. This is in understandable contrast with initial strategies since informal resource strategies (especially visiting a library or archive) are more often credited with helping scholars obtain a previously identified source.

Table 5.3 – Initial and Final Strategies

Strategy	Category	Frequency initial strategy	Frequency final strategy
Search library OPACs	Formal	8	3
Search the web using a search engine	Formal	7	4
Search Google Books or Google Scholar	Formal	3	3
Search scholarly databases	Formal	3	2
Search in a discovery tool (e.g. Summon)	Formal	2	
Search in an online archival site or finding aid	Formal	1	
Search social media or streaming sites	Formal		1

Strategy	Category	Frequency initial strategy	Frequency final strategy
Utilize items in private collections	Informal	8	1
Consult known web site (including Wikipedia)	Informal	3	6
Visit and/or browse archives outside of the U.S.	Informal		6
Citation tracing	Informal	3	
Visit and/or browse archives in the U.S.	Informal		3
Consult archival finding aids	Informal	1	
Use a print index	Informal		1
Ask a librarian or archivist	Human	1	2
Consult a colleague or other expert	Human	1	1
Attend a conference	Human	1	
Use social media to contact other people	Human		1
Request an item through inter-library loan (ILL)	Hybrid		8

5.2.7. Summary of strategies

The sections above discuss the strategies used by scholars who conduct research in languages other than English to answer research question 1. 30 distinct strategies were discussed as gathered from the user data: seven formal system strategies, twelve informal resources strategies, nine interactive human strategies, and two hybrid strategies. Examples are provided of each. These sections also provided a discussion of which strategies most often occur first and which commonly occur last.

5.3. Shifts

Research question 2 explores the shifts that scholars make among strategies in varying situation types. The model consists of three types of shifts that correspond to three types of situations: planned shifts made in routine situations, opportunistic shifts made in disruptive situations, and alternative shifts made in problematic situations. The sections below portray specific types of shifts and use the information triangle to map those shifts.

The participants in this study are well aware that, for the overwhelming majority of information-seeking tasks, success will not result from a single information-seeking strategy. Here, it is important to note the common IR concept of a known fact search, a known item

search, and a subject search (Marchionini, 1995). If the information-seeking task is to find a specific fact, (e.g. "Who were the Dutch delegates to the United Nations in 1946?" as sought by subject S46) then a single strategy is more likely to suffice. A known item search might include finding a specific source such as a book containing the English translation of a 13th century Latin text (Subject S6), and these tasks may also be possible to accomplish in a single strategy. However, a topic or subject search seeks to find information about a topic, and even if there is success in a single strategy, advanced scholars such as those who participated in this study most often engage in exhaustive research, looking for ALL that is written about a topic, not just the first thing they can find. In these cases, they will often look for one source to lead to another until they have gathered multiple sources related to their topic. Moreover, many of the scholars who participated in this study indicated that they have multiple concurrent research interests for which they may gather sources and leads as they arise. Subject S41 gives an example:

> "Every once in a while, something I am looking up will spark an interest in something else… for projects on the back burner. I'll be looking up something for this textbook, and I'll find something on China in the mid-1800s, or another paper I'm writing is on Chinese attitudes and relationships with the Hmong in the late teens, so the periods all overlap, but the topics don't. So it's kind of nice to be looking up one thing and say 'Oh look at this,' and pick that up and throw it in the file so I can look at it later."

5.3.1. Mapping shifts to the information triangle

One of the most significant contributions of this dissertation to the field of information studies is the conceptual model of the information triangle. It is designed to help visualize the shifts that users make between information-seeking strategies when pursuing information-seeking tasks as explained in Chapter 3. The basic triangle diagram places the formal system strategies at the upper left, the informal resource strategies at the upper right, the interactive human strategies below the triangle, and the hybrid strategies in the center of the triangle. The first strategy used

in an information-seeking thread is placed in the appropriate place on the triangle and numbered

with a 1. The first shift is then portrayed on the triangle as an arrow that starts at the first

strategy and ends at the location of the second strategy, which is noted with a 2, and so on. The

design of each arrow is significant. Planned shifts are represented by solid arrows; opportunistic

shifts are represented by dotted arrows; and alternative shifts are represented by dashed arrows.

A legend is included with each triangle diagram to make interpretation easier. A sample legend

is shown as Figure 3.3 in section 3.4.2.

Figure 5.1 - Blank Information Triangle

The diaries collected in the study, as well as a few of the quotations, were mapped to the information triangle, and select examples are given below, organized according to the types of shifts. Most information-seeking threads consist of multiple types of shifts, so examples may include multiple arrow designs. Within the situation types, the examples

are organized according to the category of initial strategy: Threads that begin with formal system

strategies will be presented first, then threads beginning with informal resource strategies, then

threads beginning with interactive human strategies. In all, 25 triangle diagrams were created

from diaries having three or more strategies each. While each of those 25 diagrams presented

unique patterns and had various meritorious aspects, only 14 were selected for inclusion in the

results, below, for reasons of space. The 14 examples that are presented below were selected

based on their ability present a clear and straightforward example(s) of the type of shift being

discussed. In some cases, triangles that were not selected had ambiguous strategies or shifts that were classified at one type by the participant but described as another type. Triangles with clear and straightforward data were chosen for the sections below.

5.3.2. *Planned shifts in routine situations*

Planned shifts are made in routine situations. These are shifts between strategies made based on usual research habits developed through their formal education or experience or on a logical progression identified by the user in advance. Planned shifts take place when no unexpected event – positive or negative – arises to change the intended progression.

Since scholars usually anticipate that it will require more than one strategy to accomplish their information-seeking task, they learn and develop progressions or threads of information-seeking strategies that they commonly use and rely upon. Subject S20, a scholar of ancient Greek and Latin texts, explains: "That is something that I have learned about myself that characterizes my work regardless of language or even topic or domain: that I tend to follow this same sort of patterns when it comes to doing work in anything in classics." These are considered planned shifts and are used in routine situations where research proceeds as anticipated.

5.3.2.1. *Planned shifts in threads beginning with formal system strategies*

As explained in section 5.2.5, above, the data in the study shows that formal system strategies are most often chosen as initial strategies with web searching and OPAC searching being the most frequently used. Subject S31 explains a progression that is commonly cited by the participants in this study. She begins with formal system strategies to identify one or more seed sources; then she mines the footnotes and references using citation tracing; then she contacts people related to the identified sources in order to obtain them.

> "That model is definitely my mode of research right now. That's what I do. I've got my formal resources, my search engines where I'm searching and mining

193

online. Then the pillaging of other people's footnotes and their sources and their bibliographies, trying to mine that way. And then people who have written books… There was a woman who wrote a book where she talked about… Czechoslovakian war brides in the interwar period who were working with the American volunteers. So I emailed her and said 'I am doing stuff on this, do you have any more information because this is a good story and I could use some help?'"

Mapped to the information triangle as shown in Figure 5.2, this information-seeking thread covers all three sides of the information triangle, and all of the arrows in the diagram are solid lines used to represent planned shifts.

In this example, subject S31 talks about her "search engines," but it is uncertain if she means web search engines or more scholarly searching systems such as OPACs or scholarly databases. However, it is clear that she is referring to formal system strategies that involve searching, so the direction of the arrows on the triangle is correct.

Figure 5.2 – Example of planned shifts (S31)
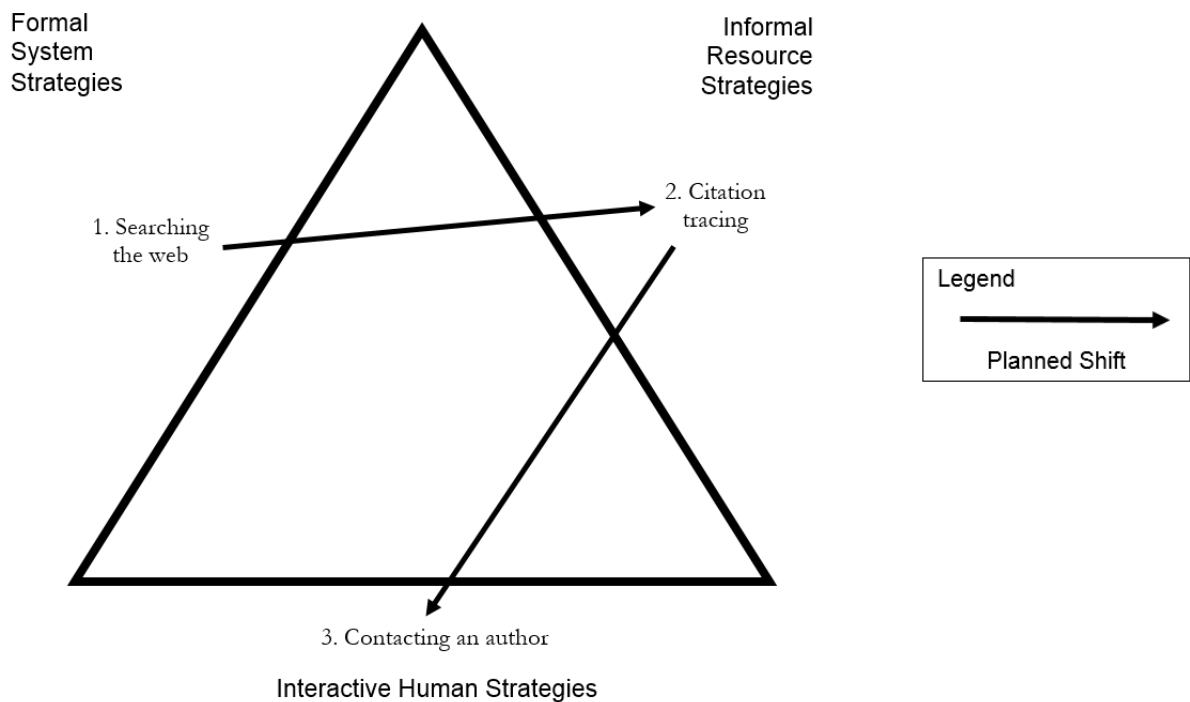
Subject S37 follows a progression of planned shifts that begin with formal system strategies, continue to hybrid strategies, and on to informal resource strategies, all illustrating exhaustive research. Quotes provided her are taken from his research diary. First, he conducted a web search to get "information on the life and works of Mulla Muhammad Tahir Qumi (d-1686) about whom I was writing an article. Searched the web – Google search engine in English and then in Persian." Saying that "I knew I can find only so much on the web, (I began) searching a database of Persian journals and then Jstor for English journals," using a formal system strategy of scholarly databases. Wanting "more info from books," he then "requested books and manuscripts from Majlis Archives in Iran via ILL." Finally, he "looked for a more comprehensive picture" and "checked the bibliographic sources on known bibliographies on-line

Figure 5.3 – Example of planned shifts (S37)
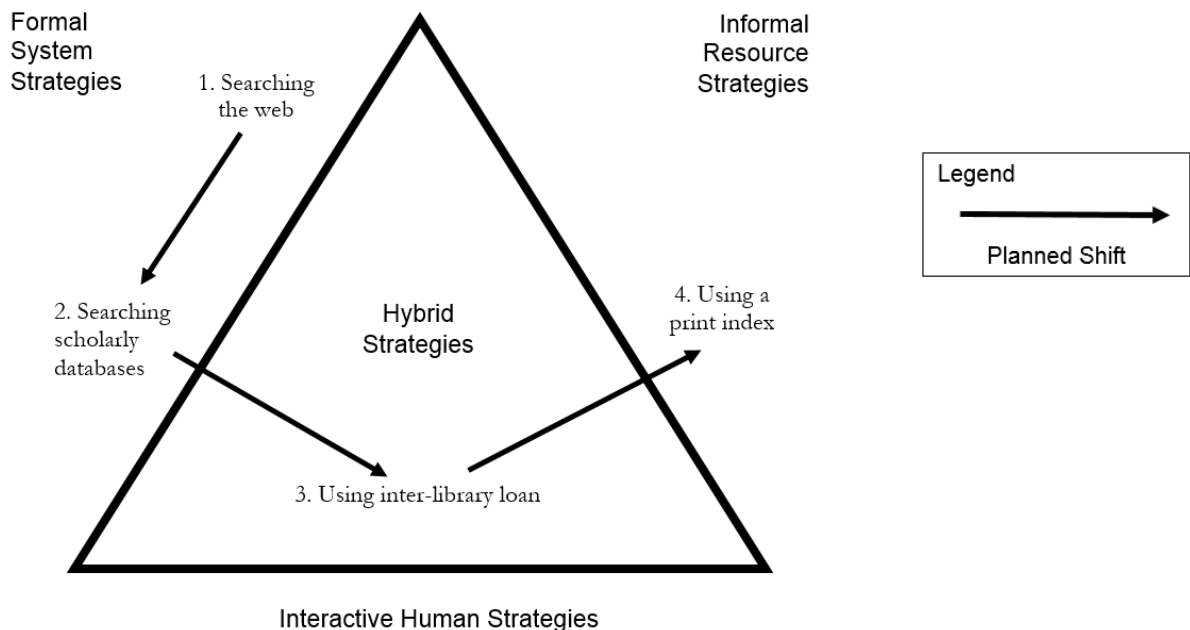
in Iran to see whether I had missed anything written by him." Since this last resource is described by subject S37 as a "comprehensive catalogue of Persian and Arabic manuscripts in

195

Iran," this strategy is grouped with using a print index or bibliography even though the resource is on-line. This thread is depicted in Figure 5.3, above.

5.3.2.2. Planned shifts in threads beginning with informal resource strategies

Planned shifts may also begin with informal resource strategies, especially with private collections. As shown in Table 5.3, above in section 5.2.6, informal resource strategies were used as initial strategies 15 times in the diaries (compared to 24 times for formal system strategies and 3 times for interactive human strategies.) Information-seeking threads documented by subjects S10, S15 and S46 all begin with using private collections and then making the planned shift to citation tracing.

Subject S46 sought to expand her list of sources for a project she had written about previously, and describes her steps in her research diary. Her project it to find the "Dutch Nazi Party's imperial plans during the period of 1931-42. I located physical and electronic versions of documents I had collected at the NIOD in Amsterdam a few years ago, as well as documents I collected more recently. Included paper photocopies." From her own private collection, she made a planned shift to another informal resource strategy of citation tracing. "Limited my focus to those I had already cited/referenced/used in this particular piece of writing and began gathering secondary sources…" She then describes her next planned shift and subsequent formal system strategy, "Because I wrote this piece a few years ago, I needed to incorporate more recent scholarship or at least make reference to such scholarship. Here I searched databases, such as WorldCat to find any recent Dutch language studies of the Dutch Nazi party, Dutch imperialism during WW2, etc." Figure 5.4, below, portrays this information thread.

Figure 5.4 – Example of planned shifts (S46)
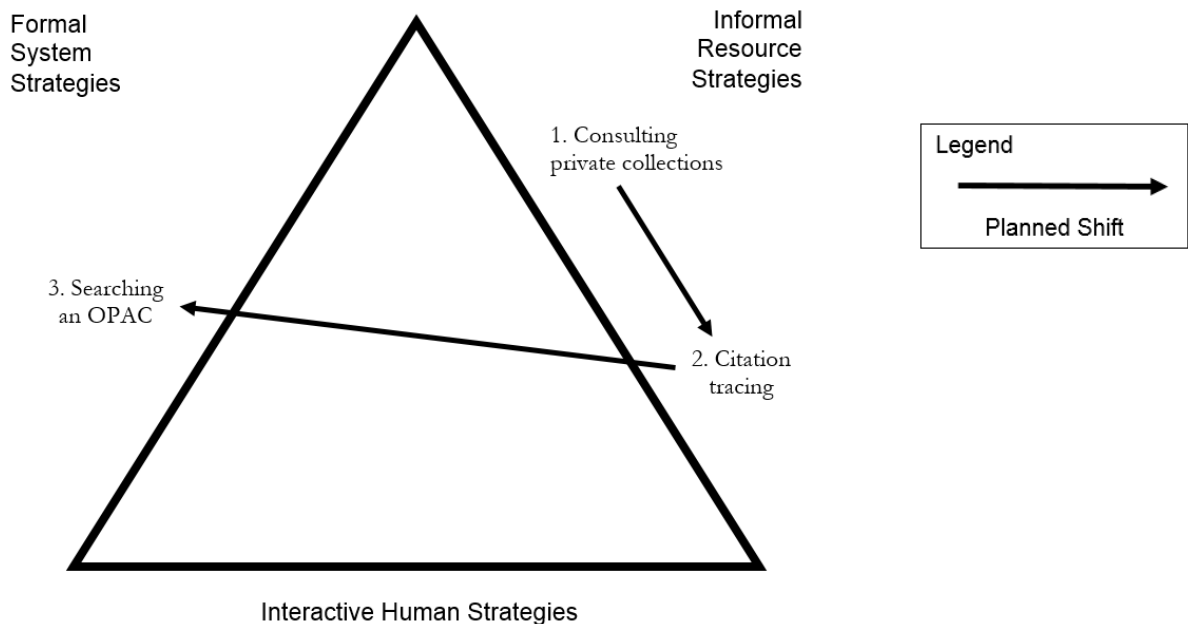
5.3.2.3. Planned shifts in threads beginning with interactive human strategies

Of all of the drawn information triangles, only two begin with interactive human strategies, and both of these only involve planned shifts. Subject S7 begins with the interactive human strategy of conference attendance where a conference panelist talked about the concept of homo ludens (people having fun,) a topic related to her research. This is a good example about how leads and sources on nebulous topics are harder to come by using formal topics because the keywords are often too general to result in a list of search results high in precision. However, armed with the new search terms from the conference presentation, she shifts to the formal system strategy of searching the web using Google. Regardless of the results found when searching the web, subject S7 makes a planned shift to the formal system strategy of using WorldCat and another planned shift to the formal system strategy of Google Books and Google Scholar in order to identify additional relevant resources. See Figure 5.5, below.

Figure 5.5 - Example of planned shifts (S7)

Formal System Strategies

4. Searching Google Scholar / Books

Informal Resource Strategies

Legend

———————→

Planned Shift

3. Searching an OPAC

2. Searching the web

1. Attending a conference

Interactive Human Strategies
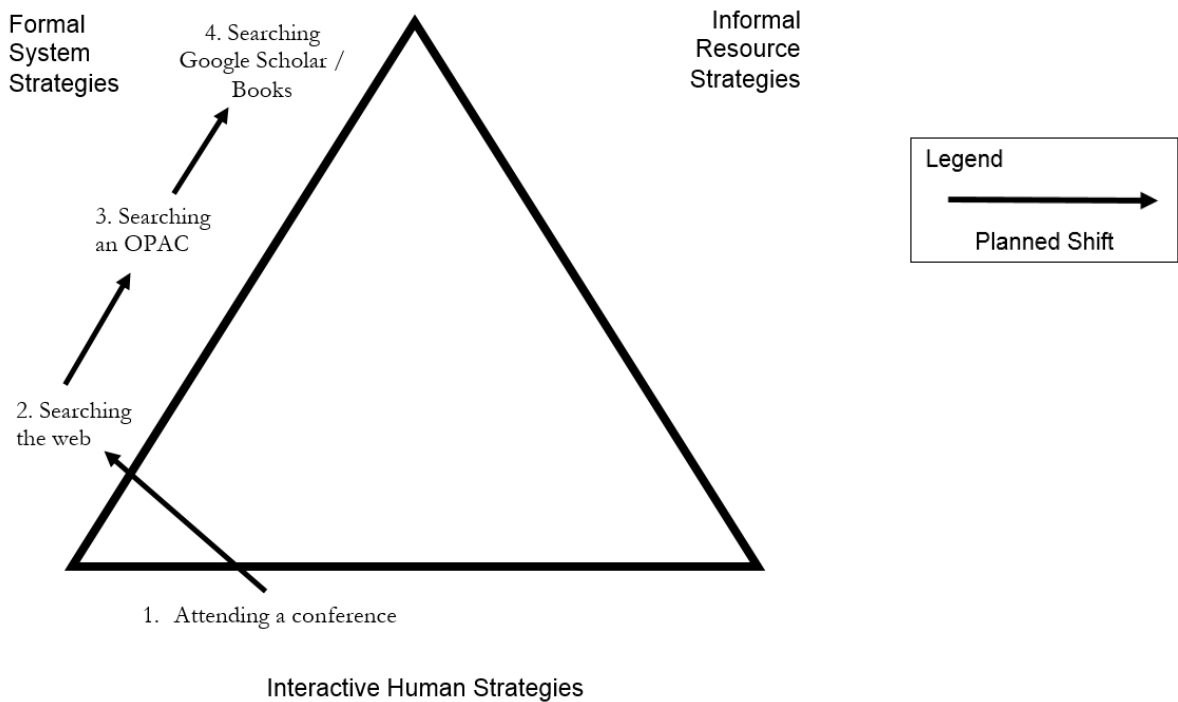
Subject S29 began his chain of planned shifts with the interactive human strategy of contacting a colleague when he heard of a relevant collection of sources by "word of mouth from a colleague." He made the planned shift to the informal resource strategy of using an on-line archival finding aid. He then made another planned shift, explaining in his research diary, "I was in a hurry and did not want to spend too much time browsing through the entire collection if it was too big, especially if a folder-level description was available and this could help identify more promising material." So he made his next shift to the interactive human strategy of asking an archivist, sending "email to the archivist to ensure the collection was available." Finally, he proceeds through his planned shift of visiting the archive in the US to obtain the sources. While this is the last strategy, it comes with a twist: "I could see from the finding aid that there was only one box of material. However, in the reading room, once the box was presented, it was

198

clear that it had only been minimally processed. It was foldered, but the folders didn't

correspond to the paper finding aid." This thread is shown in Figure 5.6, below.

Figure 5.6 – Example of planned shifts (S29)
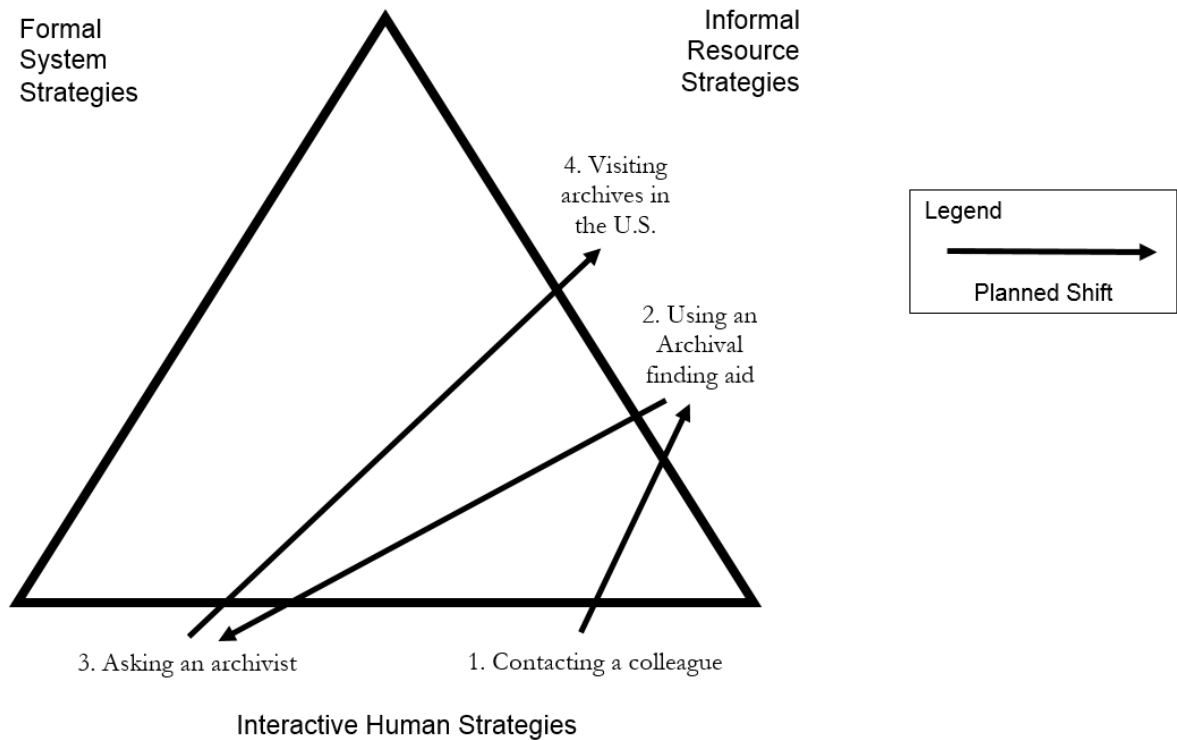
5.3.3. Opportunistic shifts in disruptive situations

Opportunistic shifts are made in disruptive situations. This occurs when a scholar

using one strategy makes a positive, serendipitous find that leads them to another strategy.

Many of the subjects in this study spoke about their serendipitous discoveries. This section

documents threads that contain opportunistic shifts. While it is relatively common, as shown

above, for an information-seeking thread to consist of only planned shifts, it never occurred in

this study that an entire information-seeking thread consisted only of opportunistic shifts. In

fact, of the four information-seeking threads portrayed in this section, two had only one

opportunistic shift, and two had two opportunistic shifts each. In three of the cases, the

opportunistic shift occurs first, followed by either alternative or planned shifts, and in the

remaining case, the opportunistic shift occurred last after two alternative shifts.

5.3.3.1. Opportunistic shifts in threads that begin with formal system strategies

Portrayed below are two very different information-seeking threads. Each of them

contains one or more opportunistic shifts, and each begins with a formal system strategy, but

they are otherwise quite distinct.

First, subject S34 provides the example shown in Figure 5.7, below, as described in her

research diary. She first pursued the formal system strategy of searching in Google Scholar Spain

to "find out if there is anything new on Pablo Aranda, author of Ucrania." She describes her

Figure 5.7 – Example of opportunistic shifts (S34)
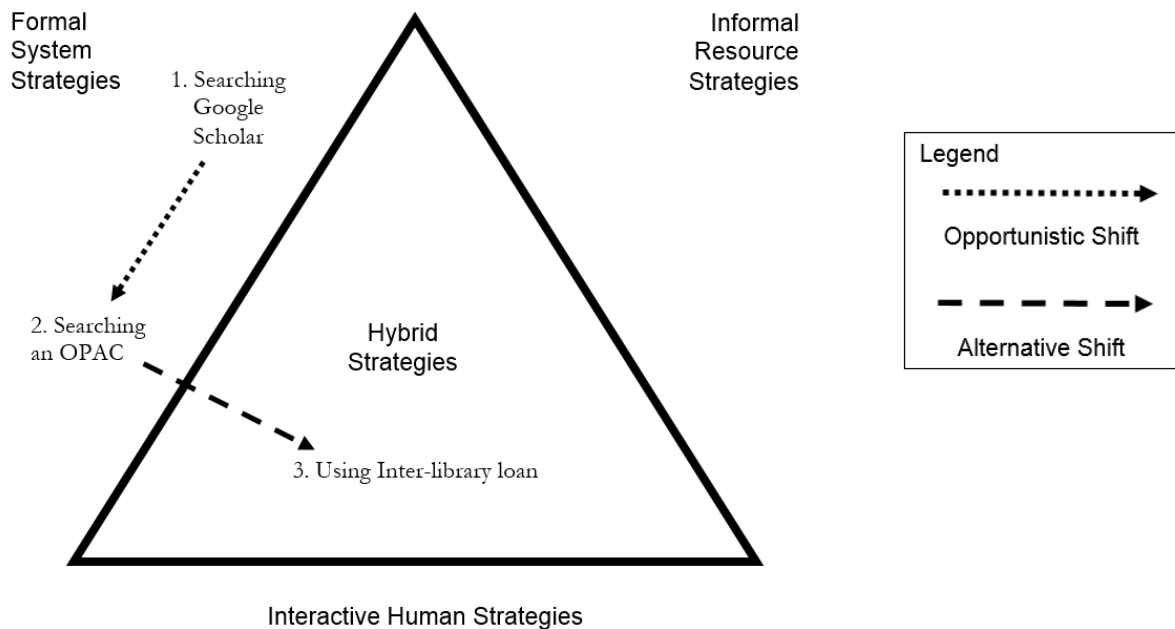
fortuitous find and opportunistic shift to the formal system strategies of using an OPAC, "I was

pleased to find that more people have written on this author since I first wrote about Ucrania in

2008. I wanted to find these articles. I went to (my university) catalog, and found them." Since

the catalog only informed her of the existence of the articles but did not provide the full text of the articles, she codes her subsequent shift to inter-library loan as an alternative shift.

Second, subject S29 benefitted from two opportunistic shifts in his search for materials on Irish language and New York. Although this information-seeking thread begins with a formal system strategy, that is not the starting place of the opportunistic shift. Rather, the first shift – the one that results from the formal strategy – is an alternative shift, and the opportunistic shifts come later. He describes his first strategy in his research diary: "I had come across the name of a periodical in the Irish language published in New York from the 1920s in WorldCat while browsing in the subject area of 'Irish Language and New York.' After search then BROWSED results by subject area." This narrative actually describes two strategies linked by an alternative shift. First, using WorldCat is classified here as the formal system strategy of using an OPAC based on the user's mention of the search first. Then, S29 explains the language barrier that caused the search tactic to fail and led to a browsing tactic within the same entity: "The title entry was not findable in WorldCat because the original library record did not render the Irish correctly (it used Gut instead of Guth, the missing h a result of script used in the periodical.)" When the search failed to immediately turn up the name of the periodical, subject S29 continued to use WorldCat but browsed the search results and used links to applicable subject headings. Because WorldCat provides the link to the subject headings, this could also be considered an assisted shift, but because S29 had to use his knowledge of the language and the tool, it is classified as an alternative shift here. When the subject begins to browse in WorldCat, this is a shift in the dimension of method – from searching to browsing, and the strategy is no longer a formal system strategy; it is classified as the informal resource strategy of using (browsing) a

known web site. This strategy allowed the subject to identify the title of the journal but not to

access the material so the thread continues as shown in Figure 5.8, below.

Figure 5.8 – Example of opportunistic shifts (S29)
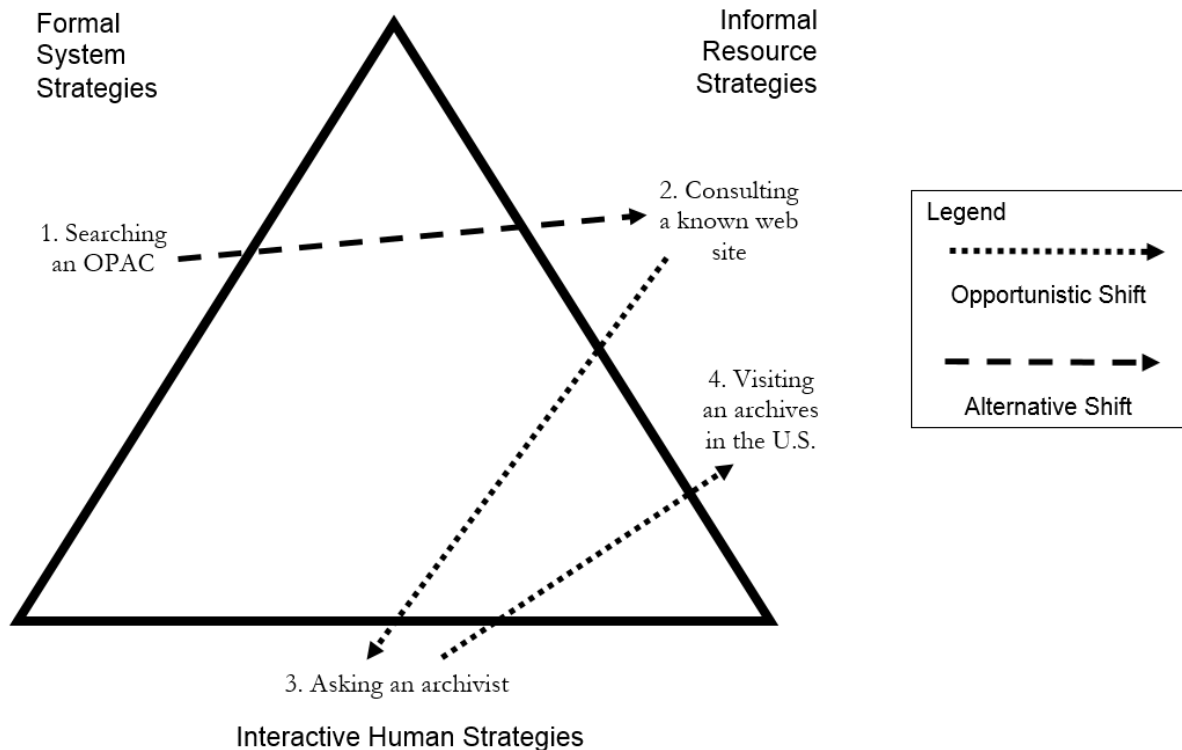
Scholars often express that their serendipitous discoveries come as a result of browsing

tactics and informal resource strategies. It is when subject S29 shifts from searching (formal) to

browsing (informal) that the first disruptive situation and opportunistic shift occurred:

> "This was simply a fortuitous discovery as I had not heard of this periodical title,
> 'Guth na Ghaedhilge' and just happened to find it in Worldcat. Luckily, it had
> been catalogued well enough to note its publication location, NY (a tip off to me
> that it might be valuable). The record pointed to Fordham University."

As a result of his fortuitous find, subject S29 contacted the archivist at Fordham – an interactive

human strategy – leading to the second disruptive situation and the final informal resource

strategy of visiting an archive in the U.S. In his research diary, subject S29 describes his contact

with the archivist:

> "On arrival, the archivist turned out to be much more helpful than the online
> records. Though the archivist was not closely familiar with the individual
> periodical or its contents, she knew that it was part of a larger collection at
> Fordham of such periodicals. This was not indicated in Worldcat, of course, or
> even in the Fordham online record, but the archivist knew this and could therefore
> bring out boxes and boxes of similar periodicals from the 1920s in Irish. The
> archivist was willing to bring me the entire collection, known as the McLeese
> collection, in three boxes. This was a huge help as it enabled me to browse all the
> periodicals and select the periodicals that fit my framework. There was no
> finding aid indicating place of publication, so this ability to browse was essential.
> Eventually, I discovered that the majority of holdings were published in Ireland,
> making them less useful, but for the two titles published in New York, I was able
> to find great new material for the project."

5.3.3.2. Opportunistic shifts in threads that begin with informal resource strategies

As stated above, browsing activity is often credited as conducive to serendipitous

discoveries, so it is not surprising that several information-seeking threads documented in the

diaries begin with informal resource strategies. Of the two threads represented here, one of them

begins with private collections, and one with use of a known web site. In both cases, the first

strategy results in a positive disruptive situation and causes an opportunistic shift. First, as

recorded in his research diary, subject S46 was conducting an "investigation of United Nation's

treatment of colonial conflict in Indonesia during the period of 1945-1949." The first step of her

information-seeking task was to "locate the names of two Indonesian members of the Dutch

delegation to the United Nations." She explains:

> "I reviewed Dutch-language primary source materials from a large edited volume
> of documents; I had made copies of relevant sections a few months ago, and now
> I returned to these. I remembered seeing mention of their names when I first
> located these sources, so I knew exactly where to begin. I knew I needed to
> compare these names against those appearing in a Dutch collection (one I had
> located in archives years ago). I located this list of Dutch delegates, compared
> against the two Indonesian names I had just located, and began to compile a
> complete list of Dutch/Indonesian members of the delegation."

Consulting the documents in her personal collection led to her fortuitous discovery, "I realized that I was, in fact, looking at two very different groups of delegates: one represented the Netherlands to the United Nations General Assembly, the other represented the United Nations Security Council." Her discovery led her to another informal resource strategy of checking a known web site, "I tried to confirm this information by checking it against a listing of delegates as made available by the UN. I searched the UN's materials as contained on their website; it's publicly available, but it contains decades' worth of relevant materials searchable by year, keyword, etc." This strategy is considered to be the use of a scholarly database, albeit not a periodical database, because it uses a method dimension of searching and provides government documents and artifacts. This resulted in a problematic situation that resulted in an alternative shift to the formal system strategy of searching the web: "I did not find this information on the UN-held materials, so I performed a broader internet search to see if some organization listed delegates to the UN meetings in 1946. I did not find this." In the process of looking for the list, subject S46 found an unexpected source: "I located a number of British Pathé newsreels/short films recently made available on YouTube (the company has been steadily digitizing and uploading its archived materials) portraying both Dutch and Indonesian members of the United Nations Security Council, during sessions." This disruptive situation led to an opportunistic shift to the formal system strategy of searching in a streaming media site – YouTube – which led to obtaining useful resources for her research. This information-seeking thread is illustrated on the information triangle in Figure 5.9, below.

Figure 5.9 – Example of opportunistic shifts (S46)
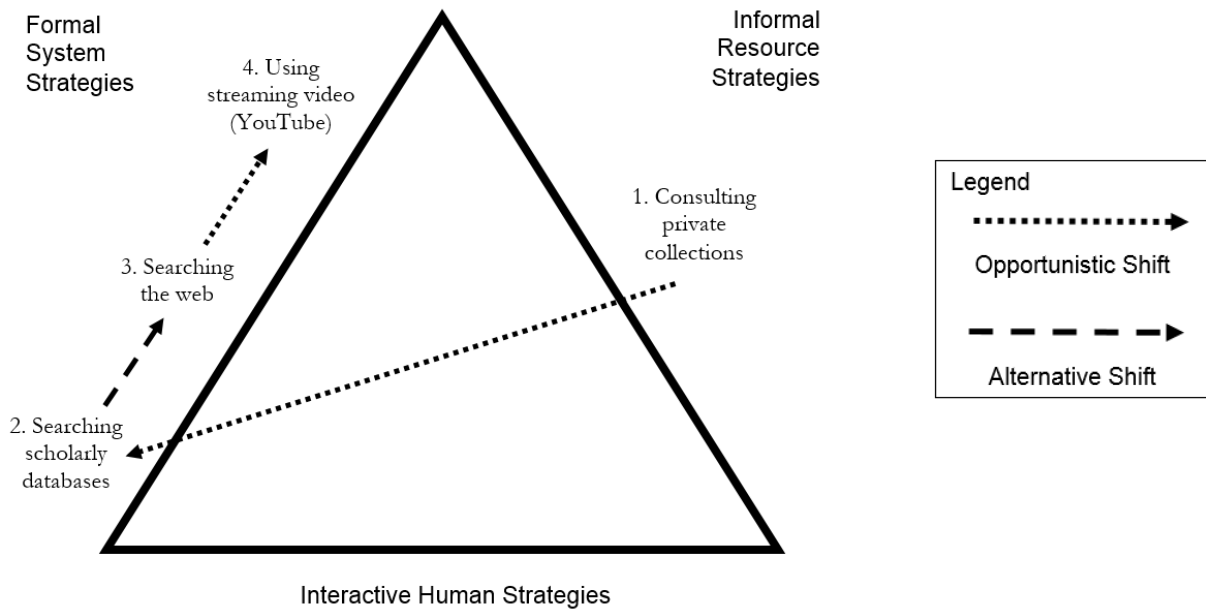
5.3.4. Alternative shifts in problematic situations

Alternative shifts are made in problematic situations. Situations caused by barriers and lack of results create problematic situations in which the researcher must shift from one strategy to another. If a system or resource provides or suggests an alternate strategy, then this is considered an assisted shift, but if it is the user themselves who comes up with the next strategy, then it is an alternative shift. In this study and in my pilot study, assisted shifts are so rare that when they do occur, they are grouped with alternative shifts.

Three information-seeking threads that contain alternative shifts are mapped to the information triangle in this section: two that begin with formal system strategies and one that begins with informal resource strategies. No threads that included alternative shifts and began with interactive human strategies were documented by the participants. Unlike the section on opportunistic shifts, there are ample examples of threads provided in the diaries that consist entirely of alternative shifts; all three portrayed here are a series of alternative shifts. Also

interesting is that two of the three threads have a final strategy of using inter-library loan. In the examples below, the factors that cause the problematic situations are highlighted.

5.3.4.1. Alternative shifts in threads that begin with formal system strategies

Participants provided five diaries and a few other simpler examples of alternative shifts that begin with formal system strategies. Two of the diaries have been mapped to the information triangle and are provided here.

It is a common pattern for information-seeking threads that begin with searching an OPAC to continue with an alternative shift since the sources being sought are most often not held by the scholar's local library, nor are they digitized and available online. In a somewhat different and unique example of a thread beginning with searching an OPAC, subject S11 describes in her diary how she undertook a known fact search to find the English title of a French book "Les Heures Longues" by the author Colette. Her information-seeking thread is illustrated in Figure 5.10, below. She first attempted to use WorldCat, explaining, "If you enter the original title and change language filter to English, it's pretty good at coming up with translation even with substantially different titles." In this case, however, the translated title did not appear. This problematic situation could, for example, be attributed to the language factor as it relates to library systems since it indicates that the resource in French is not associated in the WorldCat cataloguing record with its English counterpart. Because of her problematic situation, Subject S11 then made an alternative shift to another formal system strategy and "Googled Colette and the French title, trying to find professional organization web site (Amis de Colette, for example.)" This was not successful; "such a site did not come up quickly." Finally, she made another alternative shift to the known web site of Wikipedia "to see if the Colette entry included a full list of works, including works in translation."

Figure 5.10 – Example of alternative shifts (S11)
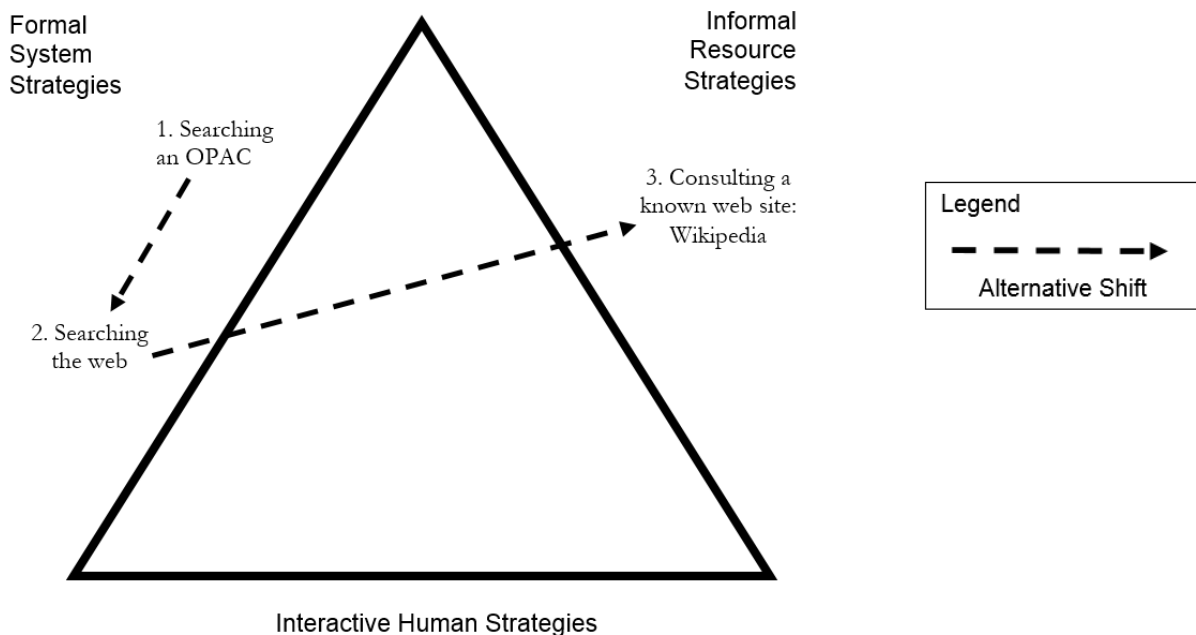
The second example from the diaries begins with searching a scholarly database. Subject S23 also pursued a known fact search, "trying to locate use of a word in the Book of Job in Syriac-Christian commentators." See Figure 5.11, below; the quotations are from his research diary. He began by searching a scholarly database – a comprehensive Aramaic lexicon. As in the prior example of S46 searching the UN web site, this is considered a database rather than a known web site because the method is searching rather than browsing. Subsequently, he "needed to browse translations of the text of commentaries on Job to locate ways that commentators use this word. These commentaries are not available in searchable form," so he made an alternative shift to the informal resource strategy of using his own private collections where he had PDFs of older translations from the 19th century. While in this process he states that he "recognized that certain words occur in Syriac version of Job that are completely different from the original text in Hebrew. Recalled article by advisor on these words." At this

point, his information-seeking task become a known item search, for which he makes another

alternative shift, this time to the formal system strategy of searching the web to "locate article by

(his) advisor. Used Google for relevant bibliographical information." When the full text of the

article was not available on-line, he made a final alternative shift to the hybrid strategy of

requesting the item through inter-library loan. The three problematic situations that caused the

alternative shifts are all approximately related to the inability to obtain the required fact or item

using that particular strategy; the information or source is not available on-line or in his

collection. The inability to obtain an identified resource is very common among the subjects of

this study.

Figure 5.11 – Example of alternative shifts (S23)
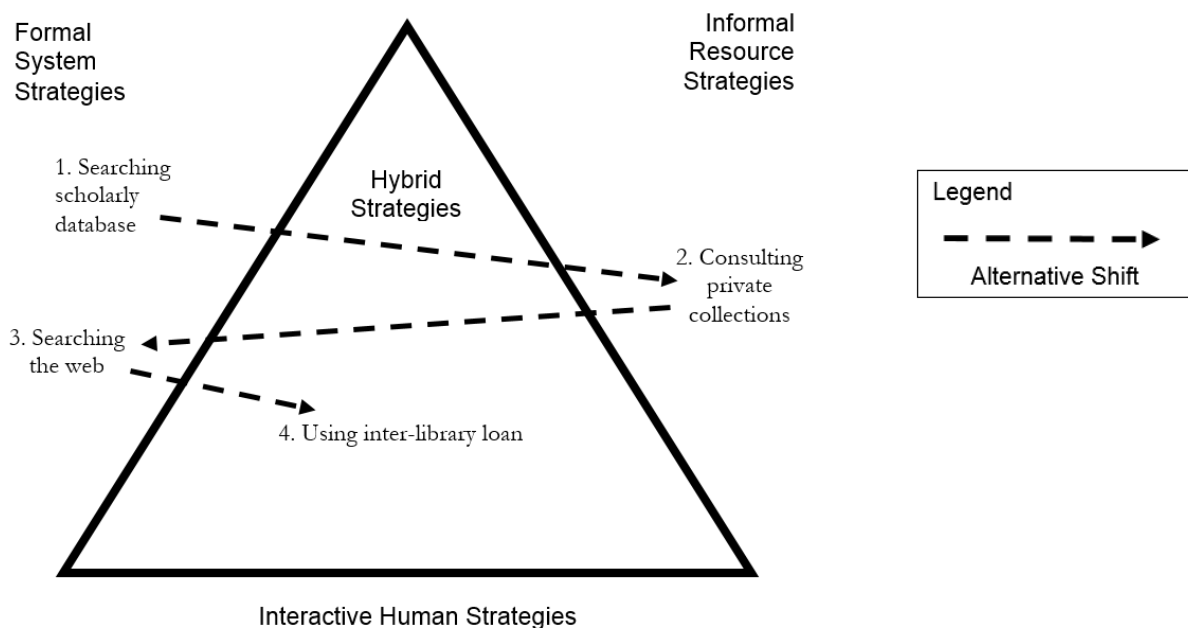
5.3.4.2. *Alternative shifts in threads that begin with informal resource strategies*

The diary thread presented here also portrays a common pattern: beginning with the

informal resource strategy of citation tracing and then shifting to the formal system strategy of

using an OPAC. Most often, the shift between those two strategies is a planned shift, but in this

case, subject S29 finds that he has insufficient citation information and so considers the situation as problematic and the shift as an alternative shift. Subject S29 recounts a thread to support his known item search for three books. He begins by following "previous footnotes and citations from Zotero, my personal research library, etc." He describes his intentions and the cause of his problematic situation and eventual alternative shift in detail:

> "In the case of one of the books - Douglas Hyde's *Turas go Meiriceá* (aka My Visit to America) - I knew in general of its existence but discovered I hadn't collected its bibliographic information in my Zotero library. I thus needed to head to WorldCat to see what editions existed, and also whether there might be translations. I often like to take stock of translations of Irish material because English versions often get more citations and if I come across a specific mention of a fact in Hyde's book in another scholar's writing, I sometimes need to compare the original Irish to ensure there isn't a mismatch in the Irish vs. English wording. Why a mismatch? Sometimes, especially with early Irish material published in 1920s and 1930s, the Irish state, in sanctioning translations, liked to clean up or soften material in the original Irish. Furthermore, sometimes a very recent English translation/edition has scholarly apparatus that can prove a speedy way to understand the background to a primary source in Irish."

However, his next strategy – searching in an OPAC – also did not yield the anticipated result but did result in identifying some helpful sources: "Using Worldcat, I discovered (to my surprise) that there was not an English version. Moreover, there were no reprints. Only the original 1937 publication in the sean-chló (old font) Irish." Despite his concern that libraries might "not lend older publications, I made the ILL request and crossed my fingers," making an alternative shift to the hybrid strategy of inter-library loan. This final strategy allowed him to obtain the resources he sought, "Happily, all three ILL books arrived. However, especially with rare material, and this happens a lot for Irish material, the lending time was very short." Subject S29s information-seeking thread is shown in Figure 5.12, below.

Figure 5.12 – Example of alternative shifts (S29)
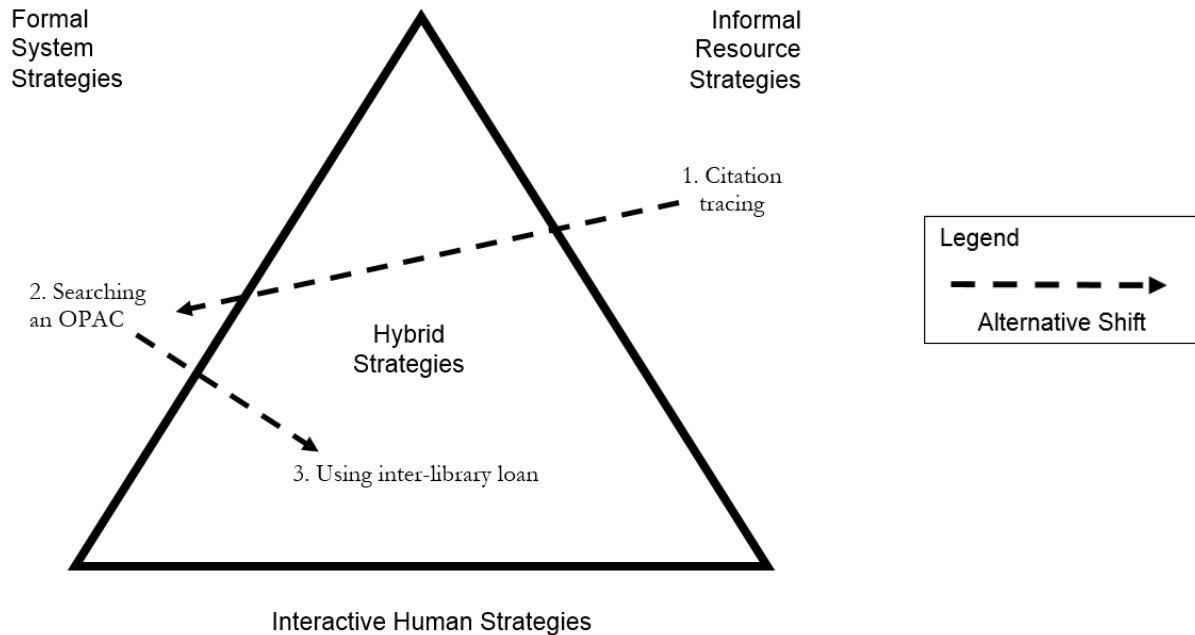
5.3.5. Complex information-seeking threads

Most information-seeking threads consist of between two and five strategies and their intervening shifts. However, in a few cases, the scholar needs to undertake a complex information-seeking thread that may include upwards of six or more strategies. Usually, this is symptomatic of an information need that is very difficult to fill and entails a great number of alternative shifts due to barriers. Three such complex threads are discussed below; most include multiple types of situations and shifts and multiple categories of strategies.

Subject S6 describes one of her information threads as an "epic saga," explained both verbally in her final interview and documented it in her research diary. Her verbal account of her known item search is as follows:

> "The first one was sort of epic because it went on and on and on, trying to find
> this thing, which was a book – a recently translated edition – a Latin edition of a
> 13th century text, and the press didn't have it on their web site, they didn't answer
> their phone… I checked Worldcat, and there was one copy available in the entire

United States, and then I posted on Facebook seeing if somebody had a copy and they could make copies of specific pages for me, and (my friend had it but) was unable to do that. I then found out that the sections of the book had been translated into Italian, and it turned out that I owned the book. I had it on my shelf, and I never really noticed exactly what was in it. One friend actually called the press and ordered the thing, so a couple weeks later, I got the book in the mail. So I was able to get the exact quote in Latin that I really needed. But then I found out that someone else had done a transcription of the Latin text and had put it up on the web, so I now have this embarrassment of riches of having one version on the web and this whopping tome that I now own. Then, I was going to have to translate those quotes into English, and it was really hard, complex theological Latin, and I had heard through the grapevine that the fellow who did the edition was also making a translation into English that had not yet appeared. So I tracked down his email address. First I tried an old email, and then I emailed another scholar and friend who had his (current email address,) and so I emailed the guy and said 'Hi, I haven't seen you since 1998, and I know it's not published, but any chance you'd just share with me Chapter 10?' And within an hour, I had a response from him with Chapter 10 of the English translation."

Figure 5.13 – Example of complex information-seeking threads (S6)
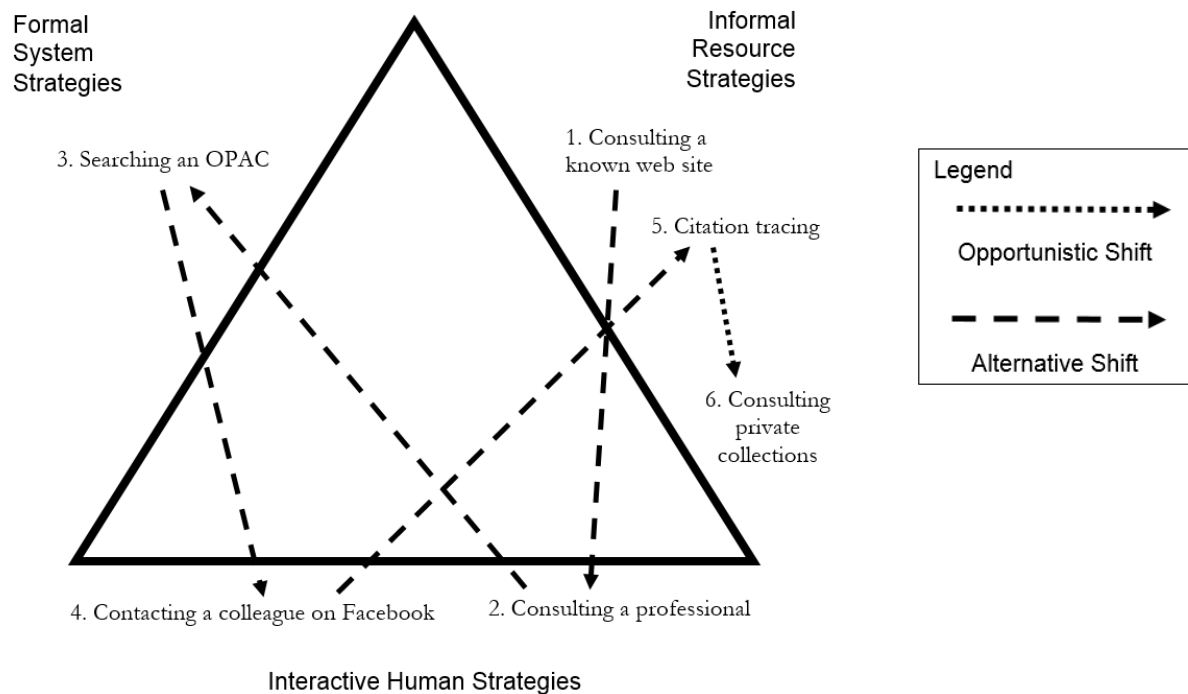
The information triangle shown in Figure 5.13, above, is coded from the diary form, so it is not

exactly the same as the interview account. Subject S6's epic saga includes six distinct strategies

on all three sides of the information triangle. Her thread crisscrosses the triangle multiple times based on problematic situations until the end, where her unexpected discovery causes a disruptive situation and an opportunistic shift leading back to using her own private collection. She attributes many of her problematic situations to the need to find her source in a specific language and the inability to obtain an identified source.

A second complex example is provided by subject S41 in Figure 5.14 below. The diagram is unique in that multiple arrows emerge from the first strategy and one of the subsequent strategies. Unlike most of the participants in the study – who at least for the sake of the study – followed a relatively linear path going from one strategy to another, subject S41 performed multiple shifts and pursued multiple paths based on the results of his first strategy. Note that there are multiple strategies labelled #2 for that reason.

Subject S41 begins by searching the scholarly database Jstor, making some fortunate discoveries as documented in his research diary. First, he says that he "found a reference on Jstor to an article on Qing whose bibliography listed a relevant dissertation and primary sources (in Chinese) of imperial correspondence. Jstor is a 'go to' first step for information on a specific topic. Articles either provide the information I seek or lead to other relevant/useful works." That first discovery led to the informal resource strategy of citation tracing, which began a multi-step information-seeking thread described below. However, during that same Jstor search process, his findings prompted multiple follow-up activities that each used different strategies, each from a different category. One led to the hybrid strategy of inter-library loan: "Also identified a dissertation from the article in the first step. Requested that through ILL." Yet other search results led to the informal resource strategy of visiting a library: "Also noted several books to pick up from (my own university) library." Each of the above shifts were considered

212

opportunistic shifts by subject S41. Other results that proved more elusive led to an alternative

shift to the interactive human strategy of contacting a colleague: "Also sent email to historians in

China for possible assistance."

Figure 5.14 – Example of complex information-seeking threads (S41)
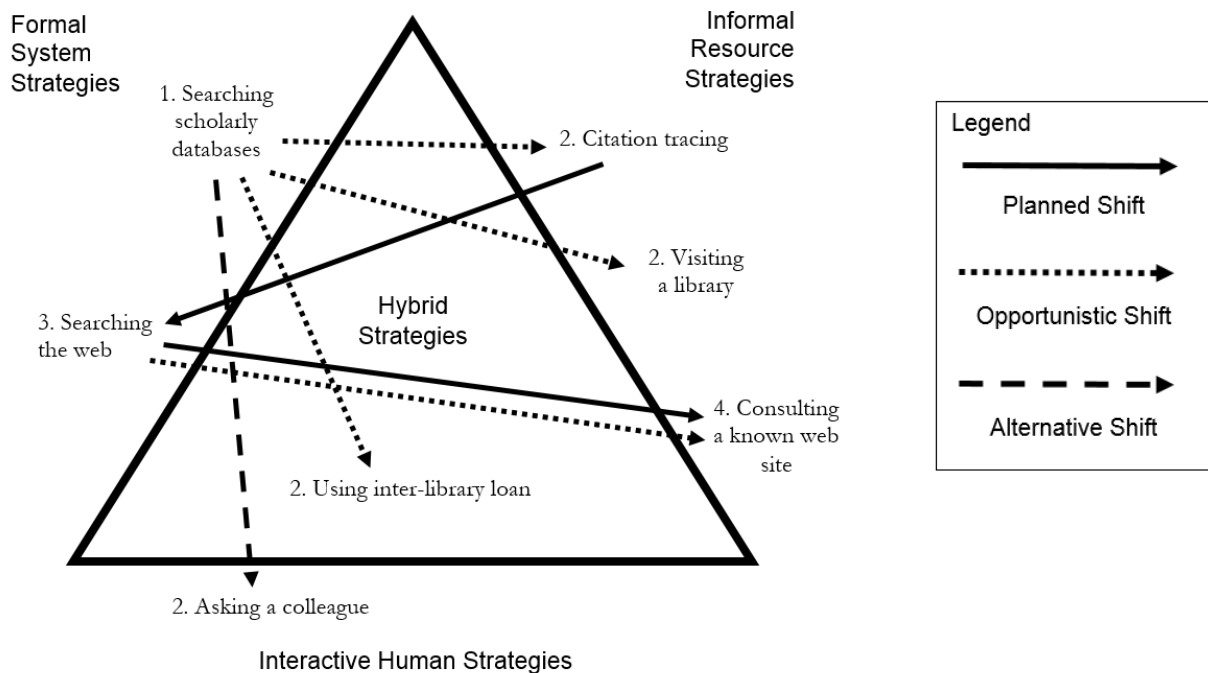
The thread that began with scholarly databases and shifted to citation tracing led to a

planned shift to the formal system strategy of searching the web. "The Jstor article's bibliography

listed a set of secret memorial at the National Palace Museum archives. So I Googled the title of

the materials I was hoping to read to see if it might possibly online." Based on his findings in

Google, he made another planned shift to the informal resource strategy of visiting a known web

site: "Results including the Historical Archives of China and the National Palace Museum web

site." He visited the first site and then made an alternative shift to visiting the second. "The

National Palace Museum web site was not helpful. The First Historical Archives held more

promise," which ended the thread. However, there is another branch in this thread as well; note

two different arrows – one solid for the aforementioned planned shift from web searching to visiting a known web site and one dotted which represents the opportunistic shift from web searching to a different known web site, as the subject explains: "Actually, while searching for information on the 'tangent' topic, I found a web site with useful material related to my original search."

In summation, subject S41's epic search began with a single initial strategy of searching scholarly databases, yet it ends with five final strategies: Three are informal resource strategies, two of which are different known web sites and the other is visiting libraries. One is the interactive human strategy of asking a colleague, and one is the hybrid strategy of inter-library loan. In all, this thread included seven different strategies in all four categories and eight shifts of all three types.

The third complex information thread is provided by subject S34 in her second diary and consists of eight steps, six distinct strategies, and one strategy that is a shift in entity only. It is shown in Figure 5.15, below. Of the seven strategies, all but one of them is a formal system strategy. In other ways, her thread is common in that it begins with the formal system strategy of searching an OPAC and ends with the hybrid strategy of inter-library loan. She states her information-seeking task as the subject search of "looking for articles about the 1999 Spanish film 'Flores de otro mundo.'" She begins by searching her "university library catalog which connects to WorldCat through their system. I used the catalog because our system also links to the online articles and books that aren't literally housed in our library. It is the easiest way to search." She indicates that she makes an alternative shift to searching the web when "only three of the entries actually had either 'Flores de otro mundo' or 'Iciar Bollain' in the title. I wanted to find more articles." She describes her web searching and ensuing planned shift to the formal

system strategy of searching a publisher's web site: "I decided to search on Google for the actual webpage of *Studies in Spanish & Latin American Cinemas* to see if there were some more on the film. I clicked on the publisher site of the journal. I went to the Advanced Search option: http://www.intellectbooks.co.uk/search/book-Form/."

Figure 5.15 – Example of complex information-seeking threads (S34)
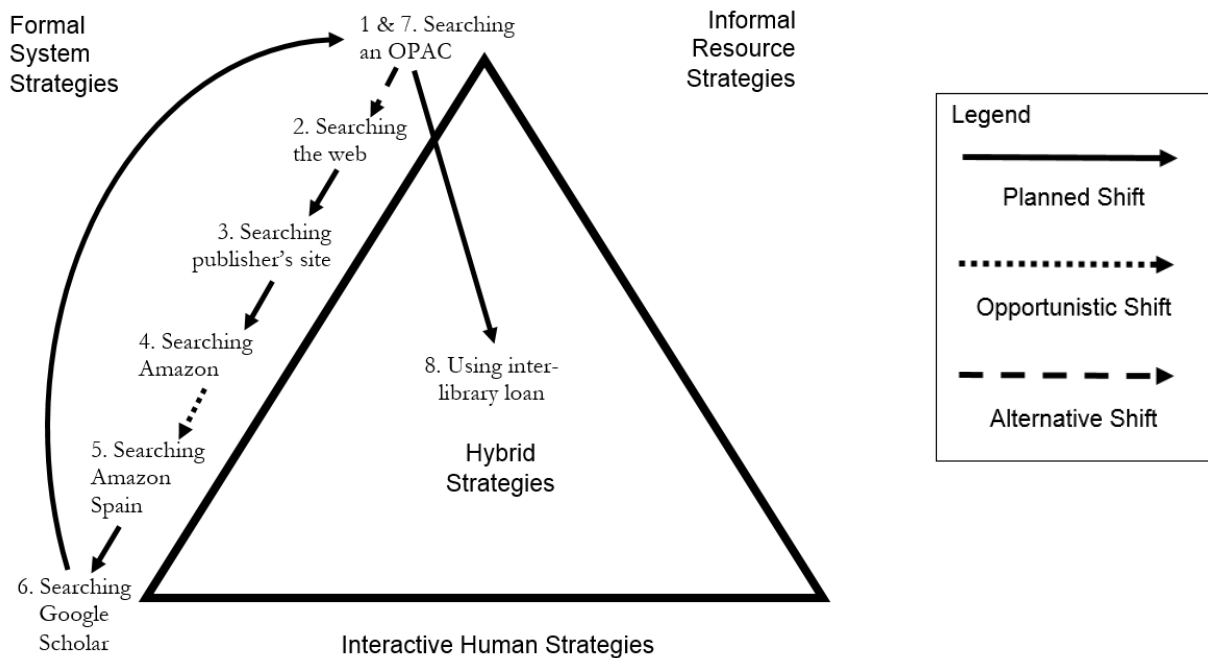
She then makes another planned shift to searching on Amazon. This is not considered a hybrid strategy because it is not clear that she intends to buy a copy on-line but rather to identify possible sources. This strategy leads to an opportunistic shift to explore additional research interests. She says, "I was curious to see if there was anything on Amazon, just because in the past I have found the oddest texts about something I am researching. I searched for 'Flores de otro mundo' on Amazon.com. I found a film for teachers in Spain about how to teach ethics using the film. It was $27 to $147. This has nothing to do with my research per se but I was curious to see how the film could be used didactically in Spanish for foreigners in Spain." Her

next shift entails only an entity change from Amazon.com to Amazon.es in Spain. "I searched the same 'Flores de otro mundo,' but in the Spanish version of Amazon—Amazon.es. There I found a text about the filmmaker and the screenwriter talking about the film. This is cheap and looks interesting, so I saved it on my list to buy when I go to Spain this December. And I will probably buy the lesson plan book. These aren't really related to my research but I have a future goal of creating my own film guides for a Spanish cinema class." She then makes a planned shift to Google Scholar Spain, the reason for which she describes as "I obviously got sidetracked and need to go back to looking for articles. I searched (https://scholar.google.es/) for 'Flores de otro mundo' to get citations. I recognized three new articles I had not found on the Library website. One of them was in German so I was curious to learn what it said, so I Google Translated it, and realized it was the translation of the book I found on Amazon with a different cover photo." Nearing success, she makes a planned shift and doubles back to searching her university library catalog to find the titles identified in Google Scholar and request them through inter-library loan, which is her final strategy.

5.3.6. Summary of shifts

The sections above summarize the shifts made by scholars to answer research question 2. Shifts were discussed by type: planned, opportunistic, and alternative. Examples are provided of each. For each type of shift, examples of information-seeking threads were taken from diaries and interviews and mapped to the information triangle. The commentary provided by the participants was then used to explain the shifts as illustrated.

5.4. Chapter summary

This chapter provides study results in support of the first two research questions. In response to research question 1, subject responses were analyzed to identify 30 distinct

216

strategies: seven formal system strategies, twelve informal resources strategies, nine interactive human strategies, and two hybrid strategies. Examples are provided of each strategy as illustrated by participant quotations and commentary. In response to research question 2, shifts were coded from both interview and diary data, and the information triangle was used to map and illustrate the strategies and shifts that occur in the various information-seeking threads. Per the format of Q2, the results were organized based on the shift types that correspond to routine, disruptive, and/or problematic situations. Considering the strategies and shifts together gives us a fairly accurate idea of the information-seeking strategies used by scholars who conduct their research in languages other than English and how and when they shift from one strategy to another.

Chapter 6 – Results: The role of language, culture, and geography in the information-seeking context

6.1. Overview

The third research question represents the point where the science of information-seeking meets the study of language, culture, and geography and makes this dissertation a new and singular study of a previously unexplored area. It explores the roles of language, culture, and geography in the information-seeking context and how they can create barriers. A barrier is defined in this study as any factor or element of a context that impedes research and therefore causes a problematic situation. Research question 3 reads as follows:

Q3: In what ways do language, culture, and geography play a role in the information-seeking context, in particular, how do they create barriers that cause problematic situations?

This chapter explores each of the three aspects and how they influence scholars' choices of information-seeking strategies and how they may contribute to problematic situations. Table 6.1, below, outlines the structure of this chapter. As in Chapter 5, quotations provided in this chapter are from subject interviews unless otherwise designated.

Table 6.1 – Chapter 6 structure

Section	Title	Description
6.1	Overview	General information about this chapter
6.2	Language	Aspects of language and their role in the information-seeking context
6.2.1	Graphical representation of the language	How languages are written, including alphabets, diacriticals, transliteration
6.2.2	Differences in linguistic properties	Properties of language such as morphology, syntax, and lexical features
6.2.3	Language reforms	Changes in language over time
6.2.4	Issues with proper names	Challenges related to proper names in various languages
6.2.5	Less common languages	Challenges related to less frequently spoken languages
6.2.6	Required language expertise	Scholars' language knowledge and skill
6.2.7	Summary of language	Summary of language and its role in the information-seeking context

Section	Title	Description
6.3	Culture	Aspects of culture and their role in the information-seeking context
6.3.1	Need to experience the culture	The role of culture in informed scholarship
6.3.2	Authorship and publishing culture	How cultures promote authorship and publication
6.3.3	Internet culture	How cultures foster a mature internet presence
6.3.4	Conditions in libraries and archives	Problematic conditions in libraries and archives including facilities, material availability, and organization
6.3.5	Political considerations	Political factors affecting access to materials, including government change, rules of access by foreigners, etc.
6.3.6	The role of culture in gaining entrée	Aspects of culture that impact access to materials
6.3.7	Summary of culture	Summary of culture and its role in the information-seeking context
6.4	Geography	Aspects of geography and location and their role in the information-seeking context
6.4.1	Need to travel to collections	The need for and advantage of accessing sources in remote locations
6.4.2	Location and distribution of archives	Where archives are located in the target location and the impact on information-seeking
6.4.4	Need to plan for overseas travel	Considerations made by scholars when planning for overseas travel
6.4.5	Summary of geography	Summary of geography and its role in the information-seeking context
6.5	Comparison with domestic research	Comparison by subject of non-English and domestic research pursuits
6.6	Chapter summary	Summary of chapter contents

6.2. Language

Of the three factors considered in the information-seeking context by this study, language is the most fundamental. The subject of this dissertation and the selection of the participants revolves around their use of non-English sources, and many of them were recruited specifically based on the language they use in their research. This section focuses on the aspects of language that influence scholars' choices of strategies and can cause problematic situations in the information-seeking context that necessitate shifts. Results herein are organized using roughly the same structure as the discussion in section 2.3 of the literature review. Since the data is

collected and coded inductively, only aspects of language that were discussed by the participants are included here, and this section also includes several unique aspects not discussed in the literature review.

6.2.1. Graphical representation of the language

As discussed in the literature review, each of the various linguistic aspects of language can complicate research in those languages. Perhaps the simplest and most rudimentary of these is the physical written representation of a language. Since scholarly research is conducted in the vast majority using written texts, the role of written language can pose a significant barrier to scholars using languages other than English, especially those that are written in non-Latin character sets.

The sections below present additional aspects of written language that may cause barriers as well as the strategies and behaviors that scholars have found to mitigate those barriers.

6.2.1.1. Alphabets and character sets

The question of what alphabet is used to represent a language can present a formidable barrier to scholars. This section describes general complexities of alphabets and character sets and provides a context for exploring how they interact with formal system strategies in subsequent sections. When scholars use multiple languages with different character sets, they are even more susceptible to the problematic situations caused by the handling of those alphabets by formal systems and search algorithms. Subject S39 gives an example that illustrates the complexities of various languages in India, especially as related to their written representations. As explained in the sections below, each of the languages then requires distinct keyboard layouts, stemming algorithms, etc., each of which can present a barrier to information-seeking. Moreover, formal systems may not collect, index, or support multiple character sets or

languages, so the scholar may need to know and use different tools for each language. When the script of a language changes over time, documents in the same language but from different time periods may not be findable even when they exist in a single database or tool.

> "This is complicated by the issue of script. Punjabi, as it is taught in Canada and the US and perhaps in the UK, it is taught through a script called Gurmukhī, which was developed in the 16th century by one of the Sikh gurus as a means for committing to writing the verses that were composed by earlier gurus of the Sikh community, particularly Guru Nanak, the founder. That script is derived from what I'll call a business class short-hand, much the same for Gujarati… Certainly, one could write down Punjabi in the Devanagari script used for Sanskrit and Hindi and, as it's turned out, it has been used, although not as widely. To use the modified Arabic that's used for Urdu can also be used for Punjabi as well. I'd say that's more common in late 20th century and 21st century Pakistan than it is for older periods."

Subject S49, who studies Egyptology, explains how script differences have occurred over time in the Egyptian language. The intricacies of these changes must be mastered by scholars in order to locate and interpret sources related to their research, and formal tools must take into consideration the evolution of the written language in order for one search string to return results in that language regardless of the time period:

> "Let's say there are five periods of Egyptian language study: Old Egyptian, Middle Egyptian, Late Egyptian – which covers the New Kingdom and the Late period, so the problem is that the language phases don't one-to-one correlate with the chronology phases. And then there's demotic, which is a different script as well, and then Coptic. So for old, middle, and late, there are two types of scripts: the hieroglyphs, which tends to be the fully formed pictures, and then the cursive form called hieratic. And then demotic is just a late development out of hieratic, so it's even more cursive and even more abbreviated. And then Coptic is the Egyptian language written in the Greek alphabet with a handful of additional signs to represent sounds that they didn't have symbols for."

In the above example, subject S49 deals with two types of writing systems that are most difficult for formal bibliographic systems and search engines to handle. Pictographic writing systems utilize one or more graphemes for a single word; examples include not only Egyptian hieroglyphics but also modern Chinese. Since most IR systems were developed for alphabetic

221

writing systems based on phonetics, they do not interact well with pictographic languages (Chang, Morales-Arroyo & Spink, 2010). Cursive alphabets are somewhat less problematic in that they are phonetically based, but they can still cause difficulties because the alphabetic characters take on different forms depending on where they appear in the word and what other characters are connected to them (Ahmed & Nürmberger, 2007).

Three different scholars work with languages that use an Arabic script but are not Arabic itself. Subject S35 says that her "current project is in Ottoman-Turkish. It differs from modern Turkish in that it's written in Arabic script rather than in the Latin alphabet, so that's one way it differs. Language changes over time." Subject S37 explains that "Persian uses the Arabic script with additional letters…but then there are some nuances in the way things are written in Arabic that are not affected in Persian... If you don't know those nuances you might not find your results." Subject S39 encounters a very specific issue:

> "I find the modified Arabic script that's used for Urdu and sometimes Punjabi is sometimes more difficult to work with because already the script is notorious for leaving out short vowels, and because there are more vowel sounds in north Indian languages than both Persian or Arabic, the script has to rely on combinations of letters in order to produce particular kinds of sounds."

Not all scholars experience the same amount of difficulty with the written representation of their target language, as subject S36 states, "Russian is really regular, it's an alphabet like English, there's a limited amount of letters, and it's not that different."

6.2.1.2. Electronic representation of characters and alternate keyboards

Any language that uses characters that do not exist in English will require a U.S.-based scholar to represent those characters when typing on their computer to enter search terms into formal systems. This is most often done by installing specialized keyboard layouts on the

scholars' computer(s). Subject S31 explains why she uses an alternate keyboard, and she

explains the ongoing challenge of doing so:

> "I will use a Czech or Slovak keyboard when I search for some of the key terms,
> and that helps a little bit… In Czech and Slovak there's so many diacritics, there
> are so many letters we don't have in our American English, so even adapting your
> keyboard and using the proper diacritic, it took a lot of practice in the beginning
> getting used to that. It took time."

Similarly, subject S16 needs an alternate keyboard to represent the characters he needs for his

research on a specific author, "Marinicz has a diacritical above the c which is not on the US

keyboard. I had to download the Croatian keyboard." Subject S45 also talks about the

advantage of using an alternate keyboard for her research in Russian, "Russian has become much

easier because you can just get the keyboard up, and even some of the finding materials in

Russian have made life so much easier because you can access stuff and find out where things

are very quickly."

Some subjects utilize multiple keyboards for the multiple languages they use. Subject

S23 indicates that, after many years of using multiple keyboards in Biblical languages, he code

switches automatically between keyboards as he begins to think in those languages. Subject S37

talks about why using multiple alternate keyboards is so important: "I have a Persian keyboard

and an Arabic keyboard, and I just switch between them. I find that something that is not helpful

at all is to search for Persian or Arabic sources with Latin characters. It's a completely different.

If you use the ordinative characters, it opens up a completely different part of the web."

6.2.1.3. Diacriticals

Even for languages written in Latin scripts, diacriticals can present challenges when

conducting research, especially using formal system strategies. Subject S25 conducts research in

Spanish and provides an example:

"I've been doing this for almost a decade now between being a grad student and a faculty member, and I still don't remember which search engines must be devoid of diacriticals. I think it's MLA International bibliography where you have to leave them out. And I think it's HAPI – the Hispanic American Periodicals Index – where you have to have them in. And I think Jstor is the only one that I use that is efficient enough to bring up search results regardless."

Subject S31 explains how part of her process to learn the language was learning when to use the

diacriticals in the Czech language:

"As I'm learning and becoming more familiar, I know when to expect the diacritical over the y, or that this doesn't look correct, it should have something. I know more now what to look for, but in the beginning, there are words that if you spell with the r versus the diacritical; it changes the entire meaning of the word. So if I would be putting in and forget the diacritical or use the wrong one, it would disrupt my search results unless I knew to look back and double check that. It would be very frustrating."

A specific kind of diacritical are the reading marks that originally appeared in the Syriac

Bible. Subject S23 explains their importance and how they were omitted in more recent editions

for reasons of technological incompatibility:

"Take, for example, the Syriac Bible. It's been published in multiple editions over the last 150-200 years. Everybody order on Amazon a copy of the printed Syriac Bible called the Bashita but, just to give you an example, many of the punctuation marks and the rhetorical remarks that were used to emphasize certain readings – whether something should be an exclamation mark or a question mark – there's a whole series of these marks used for reading that go back literally almost unchanged to the year 600 within the manuscript tradition. This makes it one of the longest-lasting traditions of reading marks in Christian history. But because those individuals who published the Syriac printed Bible lacked access to many of these marks or lacked fonts for many of these marks, (they) never made their way into the printed editions of the Syriac Bible."

6.2.1.4. Alternate alphabets in library systems

In addition to whatever other challenges alternate alphabets can present, they can cause

problems specific to library tools and systems either because the tools do not support the

language; because the documents have not been catalogued or represented in a way that makes

them findable; or because search algorithms cannot match search terms with the documents.

224

Subject S37 gives an example, "Working with non-Latin script and sources that are in Persian or Arabic has been really a major source of frustration - library catalogs not having the original Persian and Arabic searchable."

Subject S39 explains specifically how alternate alphabets and transliteration affect library systems in his research on the history of Punjab:

> "Indexing is kind of a big issue in the sense that, particularly with the Library of Congress, they have a set method for transliteration which in south Asian languages in generally based on Sanskrit. In Sanskrit, the way they transliterated it, there are words that would end with a final short vowel, an 'a,' which in many modern languages is simply dropped."

6.2.1.5. Transliteration

Users who are not familiar with alternate alphabets have a tendency to try to search using transliterated search terms. The success of transliteration in searching can be mixed. Subject S41 finds it useful to search in the transliterated form of Chinese – pin yin:

> "A lot of times I just use pin yin for search terms and that'll get you where you need to go. It's pretty easy to use on any computer, but you can also use keystrokes and compose your characters using keystrokes. There's a relatively limited number of those and you can write all the characters on those. It's interesting because I don't think of it that much, just the physical keying in."

But transliteration brings its own challenges. There may or may not be standards for transliterating non-Latin alphabets, or there may be multiple standards, as subject S25 explains:

> "It just got me thinking how, if I were to rely more on Russian language sources, how would I go about navigating, because there are different ways in which Cyrillic letters in Russian are transliterated into Russian? There's the French way, and the American way, and the British way, and they're all different."

Moreover, any given search tool, system or cataloguer may or may not have adhered to existing standards, which presents difficulties for searching algorithms used in formal resource strategies. Subject S39 provides an example:

225

"The real trick is you have to get the spelling right for your search term, particularly if you are searching by title because if one doesn't necessarily know either the official library of congress transliteration, or a catalog at an institution hasn't used the official transliteration, you may or may not be able to get something that is actually held by an institution that contributes to this database."

The result is that scholars may abandon the use of transliteration in searching, which may render some systems entirely unusable. Subject S37, who conducts research in Arabic and Persian – both of which read right to left – says, "I've tried a number of times to use English transliterated words in my search, and I have not come up with a good result. I've basically done away with that practice. I don't do it anymore."

6.2.2. Differences in linguistic properties

The morphological, syntactical, and lexical features of language can create barriers, and prior studies in information retrieval have delved into how difficult it may be, for example, to conduct stemming in languages that are morphologically complex (Lazarinis et al, 2008, Lazarinis et al, 2009; Mustafa, 2004; Mustafa, 2005; Fattah, Ren, Kuroiwa, 2005; Bar-Ilan & Gutman, 2005; Dolamic & Savoy, 2009; Moukdad & Large, 2001). In this study, the participants are by and large not linguists, nor are they aware of the IR principles related to stemming or search algorithms. While it is highly likely that many of the failures of formal systems strategies to return adequate search results could be attributed to the linguistic qualities of the target language, the subjects of this study rarely discuss those barriers in linguistic terms. Instead, they provide examples for some of the more basic cases. Subject S19 gives an example of how the rules and grammar of other languages can conflict with accepted citation practices: "German has their strict rules on capitalization, but if you follow APA or MLA, they expect you to capitalize the first letter of every major word, or you don't capitalize them at all except for the first word of a sentence, but then you've got the German wrong."

226

6.2.3. Language reforms

The issue of changes in language over time was so often mentioned by study participants that it merits a section of its own. This phenomenon is usually called language reform, especially when the changes are done explicitly and purposely by a government entity. Since the majority of the subjects in this study are historians, along with several classicists, they work with older versions of languages, but this does not obviate the need to understand more modern and recent versions as well. Subjects experience changes in the written representation of the language, changes in vocabulary and grammar, and so on. Subject S29 offers a particularly poignant explanation of how language reforms in Irish create multiple types of barriers, especially when using formal search strategies:

> "One of the common problems that's specific to Irish is that, for historical sources, there was a major reform in the spelling of Irish in the 20th century, and it was fairly significant as to the changes to the conventions of spelling such that terms from the 19th century or even early 20th century are very hard to match in a keyword search. It's a matter of missing letters compared to the current spelling, and probably the most notable one is that the Old Irish spelling uses a punctum over many consonants to indicate a phenomenon known as lenition, which is a softening of the consonant. In the 20th century, that punctum was replaced with an h, so if you do a keyword search with an h, which is really the only way you can do it because the punctum is not meaningful to search engines, and somebody else has somehow rendered their catalog term using the punctum, you'll miss that term. There's no doubt about it. And even in WorldCat, you see people have used the Unicode character for say a "d" with a punctum above it, or they'll do something that looks like it – they're trying to find the character that looks like it, but it's not really the phonetical meaning of that term… of that consonant, and so of course you're missing those, and that's always a concern, so you sort of look at strategies to get around that."

Subject S18 indicates that multiple aspects of Japanese have changed from earlier forms, "19th- and early 20th-century materials, the grammar is different than contemporary Japanese. The characters are also different than contemporary Japanese."

When the writing system of language is set more recently, it can have more logic imposed on it from the beginning, making sure, for example, that each written character represents a specific sound. Subject S44 explains how this happened in Vietnam, "The Portuguese began to urbanize it, but the French basically standardized the Latinization of Vietnamese in the late 19th century so by the 1920s and 30s, you get a lot of newspaper and periodicals written in Vietnamese using modified letters of the Latin alphabet." Nonetheless, language resists stagnation and permanency, and it evolves despite well-meaning canonization and a recently imposed system of writing. Such is the case of Xhosa – a click language of South Africa. Subject S30 explained to me that its written form was more recently set and therefore generally unproblematic, except in the case of evolving proper names. She explains:

> "For example, there's a place in the Eastern Cape, South Africa called Mthatha, and in the past it was spelled U-M-T-A-T-A and after, in the past 10 years they changed it to M-T-H-A-T-H-A, so instead of it being a U, the beginning letter is now M, but there are very few cases of that and if I'm searching I would just know to put in both ways, but most of the time it would be the old spelling."

Some of the phenomena that have caused language reforms and their effects on the information-seeing context are included in section 6.3.5.2 because they are related to government and regime changes.

6.2.4. Issues with proper names

Very closely related to both written representations of language, language reforms, and government/regime change are the challenges related to proper names. Throughout my pilot studies and this study, issues related to proper names have been cited on numerous occasions as causing problematic situations (Sabbar, 2013; Sabbar, 2016). This is most often the case when the research language has a non-Latin character set, but there can be many reasons that searches for proper names fail in search engines, databases, and OPACs. With so much research being

done by scholars in Spanish, one might think that such phenomena in Spanish would be rare, but

subject S7 provides an example of how languages that have been written for centuries may have

been inconsistent:

> "When I do colonial work, I do have to look because there are different spellings
> for the name of a person. It could appear in different spellings, and Old Spanish
> wasn't… well, you know, it wasn't really 'old Spanish,' but Spanish grammar
> wasn't fixed back then like it is today, so people would spell different names. So
> a scribe could have written the name with one spelling, another scribe with
> another spelling, and then when you search it, it only pops up with that spelling,
> not another. So I have to write the different variations of the name of a person
> when I am doing colonial work. At times I've done it with medieval too, because
> that happens with medieval Spanish as well."

Subject S6 provides a similar example of a person's name that could be represented in multiple

languages and therefore may not be findable in search tools:

> "For example, this book which someone might think would be labelled as having
> been written by Peter Olivi. If I were asking for it, that's what I would have
> looked for, if I didn't know where to find it. I would have said it was written by
> Peter Olivi, but it turned out that the guy wrote the name in Latin… The author of
> the book itself is Petros Johannes Olivi, which is the Latin form of the guy's
> name. Everyone has been arguing what this guy's name is. Is it Peter John Olivi?
> And the answer is 'no' because it's the genitive of John so it's Peter of John Olivi,
> but you put it in French because he's from the south of France. Or you put it in
> Occitan because he's from the south of France. It's really just a mine field."

Subject S20, who studies classics, finds similar issues with names in Latin. When these figures

are referred to by name in secondary literature, their names are rendered into the language of that

secondary source, and those vary from language to language. His go-to strategy facilitates

searching for Latin names across languages – something that more general scholarly databases or

web search engines would not do:

> "If I'm just looking for scholarship on an author, an ancient author, then
> (L'Année Philologique) does that for me, they basically use the Latin form of the
> author's name and tag every article on it. This author I was looking for Lucretius,
> which is in English, is spelled the same way as the Latin name Lucretius, but the
> French call him Lucrèce, and the Italians call him Lucrezio and the Germans call

him Lucrez. That's like four different spellings, but however it is spelled, it will be tagged with the Latin name of it, so L'Année will do that."

In addition to names of people or places, another variation of issues with proper names is that of acronyms. Subject S31 finds that searching for acronyms in the target language yields different results than searching for the English acronym when searching in formal tools:

"If I used the proper names in Czech, it will usually give me the most specific information… Sometimes I use the English translation of it or the English phrasing of it, but that's only if I'm trying to get an American perspective or British perspective on one of the agencies. So, for example, the YWCA abbreviated in the Czech, IFKA, whereas the Brits and Americans call it the YWCA. If I want to look up Czech attitude or publications, I look up IFKA instead of YWCA."

The above quotations provide examples of barriers encountered when searching for known names. Another, more culturally interesting barrier related to proper names is not being able to find out what the name is. In my first pilot study and in this study, scholars of Japanese nearly consistently express challenges in finding names and their canonical representation in Latin characters (Sabbar, 2016). Subject S33 explains the phenomenon well:

"One of the biggest (problems) for the project I'm working on now, because I'm using a lot of archival sources from the 1940s, the problem of locating the readings for Japanese place and people names. This is a particularly thorny problem in Japanese studies. The readings of characters used to write Japanese names for both places and people are highly irregular. For a famous person, it can be relatively easy to find the correct reading because these people have Wikipedia pages or other biographical dictionaries and things that one can rely on, but failing that, if it's just sort of a regular person whose name happens to appear in, say, a newspaper article, it can be difficult, if not impossible, to locate the perfect… pronunciation that they would have used to say their own name correctly."

She then provides a specific example of looking for a place name in Japan:

"So I put that in, and plus 'Hiroshima,' and I wound up finding a tourist website that had a map on it, and the map was conveniently labeled in both Japanese characters and Romanized characters because it was for tourists, right? I was able to match the characters that I had on hand with a Romanized reading of the characters, and that sort of clarified the issue for me."

230

Although not as problematic, subject S36 retells her challenge finding the full name of a

Ukrainian cinematographer:

> "There was someone who I'm still not completely sure I got right; there was a
> Ukrainian director… Her name was off by only one letter, and the time period
> was about right for her; her first film was made right around the time that this all
> took place in 1973, so that was my best guess and I put a footnote about it in the
> thing. But all the others, and there were 10 or 12 of them, I managed to be very
> confident that I had found the right first name for them. Russian names are very
> standardized; there aren't many of them."

6.2.5. Less common languages

In the recruiting stage of this study, I reached a point where I specifically sought scholars

who utilize a wider variety of languages, including less spoken languages. Not only do these add

to the diversity of linguistic qualities, but they present a unique set of challenges as well. A

common phenomenon related to minority languages is the hesitance of institutions to fund

resource acquisition or research in those languages. Scholars of Norwegian, Xhosa, Irish, Dutch,

Czech, or Hindi will likely not only be unique to their institutions, but they are also unique

regionally and even nationally. Subject S29, who conducts research in Irish, explains that

"money is scarce for minority language sources. So people say 'Well, how much usage are you

going to get out of it? We can't, as a grant body, in good conscience, fund something that is

going to have so few users around the world.'"

As with other language aspects discussed in this section, the phenomenon of less

common languages has impact specifically in the use of formal system strategies, because

minority languages are more likely to be miscoded and misrepresented in library systems.

Subject S29 gives an example:

> "That's the other problem, they may have been mis-catalogued because they're in
> a less common language. It's sort of a diversity of sources because you don't
> know what might have Irish language content in it. It might not even be
> catalogued as such. It could be catalogued as an all-English source."

Additional examples of the impact of less spoken languages are also included in the discussion of authorship and publishing culture in section 6.3.2, below.

6.2.6. Required language expertise

A defining characteristic of the participants in this study is that all of them speak and read the languages that they use to conduct their research. Of the factors related to language, culture, and geography, this aspect of language expertise is the only one that could be considered a characteristic of the subject and not of the more general context. Unlike most other aspects of language discussed herein, the issue of language expertise can and does affect the researchers' use of strategies in multiple categories. For example, a researcher's knowledge of the language directly affects their choice and use of keywords which in turn impacts the success of formal system strategies. When using informal resource strategies characterized by browsing, language expertise can affect relevance assessment of sources (Hansen & Karlgren, 2005). As discussed in section 6.3.4, gaining entrée and getting help from archivists can be made more daunting when the researcher's spoken language expertise is lacking.

In only a few cases, discussed below, is the research language also the native language of the researcher. In the remaining cases, the scholar learned that language as a part of his/her education. With the language scholars, the pursuit of language expertise more often preceded their research on specific topics. Conversely, for historians and classicists, it was the topic that first interested them, and they then undertook the study of the language in order to conduct research in that area. Such was the case for subject S29:

> "Initially I was very much interested in the everyday life and the culture of 18th and 19th century Ireland, and I felt that it was important to know something about the language and those communities in order to really understand it because of course it was a prevalent part of the culture. So that was my initial drive to learn the language and use it for research. It wasn't something that I started with or that

I had coming into graduate school and doing my training as an historian. It's something I picked up early on that I had to work very hard to gain in order to work with the sources."

For some scholars, a conversational level of knowledge or even a reading knowledge of a language goes a long way. Subject S17 explains, "For obtaining the information I needed, I had to learn Serbo-Croat. For purposes of being able to read the research that had been done as well as be able to speak with informants in the former Yugoslavia." Yet the scholars realize that they must master the primary language of their research extensively. Subject S19 says:

"If you're publishing or writing scholarship on a classical antiquity, if you don't know your source language well enough to detect things like irony, like sarcasm, like humor, which are the hardest things to detect in a language, then you shouldn't be doing the research. But in terms of the scholarship on other languages, for me it's functional. It's pure functionalism. Can I get the gist of what is being said?"

Scholars using languages that are no longer spoken need to develop reading knowledge of that language and a speaking knowledge of the language that is spoken in the primary research location. While these may be similar (e.g. ancient Greek and Modern Greek, Latin and Italian,) they are often different enough to present a challenge. Such is the case for subject S35 and her use of Ottoman and modern Turkish:

"I was actually studying Turkish in order to be able to read Ottoman-Turkish, so I don't think I could do that until my second year of graduate school, but I think by my second year of graduate school I was taking classes in Ottoman-Turkish. The modern Turkish, of course it's helpful to know when you go to Turkey and you're doing research, but it wasn't really my main focus. I actually wish I knew modern Turkish better, because when I do go to Turkey, it's difficult."

Knowledge of the language currently spoken in the target culture and location can indeed open doors and help the researchers to make connections, as subject S15 experienced. "At least within Brazil, Portuguese speakers in general are really, really appreciative when you speak the

language. And so even if you feel like you're stumbling, you know, flex your Portuguese muscles. It's really helpful. It also establishes that relationship."

For the subjects in this study and other scholars studying topics outside the U.S. and languages other than English, having language expertise not only allows them to read sources and talk with people about their research, but it also provides a window into the culture that they are researching. Subject S30, who studies nurses in South Africa using the Xhosa language, explains why she believes "knowing the Xhosa language is so important to understand the society these things are in. You really learn a lot about a people and a place by learning the language."

Finally, not only do these scholars have to develop language expertise for a culture that is not their own, they need to maintain it from afar – from their homes and offices in the U.S. Subject S31 expresses this challenge:

> "I would say that I have an intermediate fluency in both Czech and Slovak, the two primary languages I use, but I don't have access to keep up with that language study. I have to take initiative to do that and work with my former advisors to speak the language or to keep in contact with my fellow students from my classes. Otherwise, I find myself getting out of practice, and that doesn't help me continue to develop the skills I need to go through the sources."

6.2.6.1. Conducting research in multiple languages

Some areas of research require expertise in multiple languages since the primary texts are in one language and secondary sources are in various other languages. This is especially true of scholars in the classics: Greek, Latin, and Egyptian. Subjects S19, S20, and S49 all indicated that they were required to develop competency in multiple languages. Subject S20 explains:

> "Classics, as one of the earlier disciplines in the modern academy and also an interdisciplinary one, is transacted in a lot of different languages. So in order to do research on any of my topics, I have to work in scholarship in English and German and French and Italian and occasionally in some topics in Spanish. I haven't had to deal a whole lot with Polish, Russian, or Modern Greek but those

234

are three that, depending on the subject, some of my colleagues have to deal with and work in and research. The subject matter is Latin and Greek and the scholarship is in those modern European languages."

Subjects who study primarily religious texts (S23 and S37) often need to study multiple languages, including Syriac, Coptic, Aramaic, Greek, Arabic, and Persian.

Some scholars change or expand their research interests over time, so they undertake new languages to explore those multiple interests. Subject S43 says:

"I use Afrikaans quite a bit, and then I often read my news in French and German. Because I took the time to translate my research to Afrikaans, in doing so and engaging in Afrikaans, I was able to further my research because it exposed new parts to me which otherwise would not have been evident. I find that really valuable to use a lot of languages, and Dutch is really close to it, so sometimes I use Dutch and then sometimes I do vocabulary searches in other languages."

Another common reason to need to know multiple languages is to study regions that have been colonized or changed hands at different times in history. This is the case for subjects S6, S30, S38, S43, S44, and S46. Subject S44 states, "My first book focuses on the social history of poverty in Saigon during the colonial period, and the main languages I use are French and Vietnamese and increasingly also classical Chinese." He explains the cause of this need,

"Vietnamese was spoken by Vietnamese people, but because they were under French colonial rule from 1859 through to 1954, a lot of the documents from the colonial administration are written in French. French was the language of administration, and educated Vietnamese sometimes read and speak in French, but typically Vietnamese."

Subject S38 gives an example of the effects of colonization and his multiple interests:

"My research deals with North Africa; it deals with France; it deals with Islamism. Languages I use are French, Arabic, and occasionally Spanish when referring to ex-Spanish colonies in North Africa. I also do Byzantine studies so I use Greek, but my middle Greek isn't that good, my Modern Greek is better. But I essentially use French, Arabic, Greek, and Spanish."

And subject S39 says the following about a specific research project he conducted in and around

Pakistan:

"The bulk of the sources I used were in English, but there were some sources I consulted which were in Persian, although that particular Persian document was composed in the early 1860s and was actually commissioned by a district office, so the vocabulary in that is heavily Persianized although the beginning is heavily Arabicized. There are also one or two things that were actually printed in Urdu or written in Urdu, but again, the bulk was in English."

In many fields of Latin American studies, scholars find that using only Spanish can exclude some potentially useful sources. Expanding their research to include Brazilian Portuguese opens up additional sources, as described by subject S25:

"In comparison to academic sources in Mexico, for example, or Argentina, or Chile, it is generally a lot easier to find full PDFs of academic articles written in Portuguese at Brazilian universities within just one or two steps of doing a search on-line. Whereas Argentine academic literature on films tends to be less open, not as internet-accessible, and harder to access."

6.2.6.2. The factor of native language

Of the participants in this study, only four of them (S3, S7, S37 and S43) conduct research in their native language. While conducting research in one's native language gives a definite advantage in terms of expertise in the current form of the language, it may not help the researcher overcome barriers related to language characteristics that cause issues with search algorithms, nor does it help in obtaining far away or rare sources in those languages.

Subject S18 explains a language related barrier that she (as a non-native speaker) experiences that she believe applies to a lesser extent to native speakers. She discusses the difficulty of reading older versions of Japanese script: "A lot of Japanese people also can't really read the stuff, but yea, they have the native fluency, so they can kind of parse the sentence or whatever in a way that I cannot do." Native language skills can also help when communicating with librarians, archivists, colleagues, and authors in the target country or culture. Subject S10 relates a challenge that she encounters:

236

"There's not much of a cultural barrier for me in Germany. There's a little bit of one because sometimes they don't understand me. I speak with an accent, but most of the people are used to working with foreigners, and in the town of Augsburg, they know me."

Finally, the relationships developed by natives in the culture may provide some advantage when relying on colleagues and others in the field. Subject S37 gives his perspective on conducting research in him home country of Iran:

"Part of what happened to me, also, is that, because I'm from there, I have a wide network of friends, so I can ask them, if I have a source that's kind of in a library somewhere with no digital version, I can actually ask a favor and they would go and scan that for me."

He continues on, however, to state that he does believe that scholars from the U.S. can build a similar network of friends and colleagues on which they can rely. "When you are there, you develop these networks, so it doesn't really matter if we are Iranian or not."

6.2.6.3. Dialects and nuances in meaning

If developing adequate language expertise were not complicated enough, the issue of dialects, nuances, and other variations in the language can obfuscate meaning in ways that make research more difficult at the stage of assessing and utilizing the resource. Non-native speakers often find it challenging to identify nuances in language such as sarcasm, humor, irony, etc. They may also lack the cultural context to identify or feel the significance or visceral reaction that native speakers and people in-country would attribute to specific words, phrases or texts. Subject S46 provides this example:

"My Dutch is ok, but there are times that I just don't have the context. I just don't know why something is a big deal or what the conflict is that people have with each other as a non-Dutch person... Some of that essential background might be apparent to some people, and that's when I will Google things and go to the secondary sources and play around with them and see what I can find out, usually that's my problem that I can't figure out why something is so significant because I lack that context."

She also provides a specific example of specialized language required for her research and how it thwarts the use of formal system search strategies and creates a problematic situation that causes her to shift to the interactive human strategy of asking an archivist:

> "My first larger topic was looking at Dutch resistance organizations during WWII and their views of the Dutch colonies, and for that I did look at a lot of German materials, and unless you know the particular Nazi coding - Nazi language, versus the Dutch or the more colloquial Dutch understandings of certain terms, you can't really search for them. They're not going to show up the same way. At that point you would have to ask the archivist and say, 'OK, I'm looking for this. Where do you think I would find this?' Because searching for those terms would actually lead you nowhere."

A similar phenomenon to that of nuance and specialized vocabulary is colloquial language. Scholars who study language in order to pursue research most often learn the language with a very academic vocabulary. Learning colloquial language is not easy, even when spending some time in country. Subject S30 provides an example:

> "Some of the titles are in Xhosa and you have to know what you're looking at, you have to read the language to know exactly what it is, so that in the past when I have been looking for other things has been a little bit of a challenge and then also the title is known colloquially differently as it would be cataloged as."

Another type of variation in language is dialect. Most languages have multiple dialects within them that may have multiple variant linguistic properties: primarily lexical, orthographic and semantic. When dialects intervene in the information-seeking context, it can mean that there are multiple, albeit possibly similar, languages to learn. Subject S40 explains the situation of Norwegian:

> "Norwegians were really hesitant to commit to any written system of language, and so there are technically two official languages in Norway - one that's kind of the standard Danish-Swedish influence, and the other one is this guy that traveled around Norway in the 1840s recording dialects, and then he just wrote them all down into a system. And they call that Nynorsk or 'new Norwegian.' So when it comes to grammatical structures and written orthography, you never quite know what you're going to find. And about 15% of the populations speaks Nynorsk, and about 85 does standardized kind of Danish-Norwegian… There are about 40

238

dialects in Norwegian. So while you can write either the standard form or this new Norwegian, there are then dialects. So from whatever region you're from, it's pretty clear when you start speaking to people. And the written Norwegian will be different from their spoken Norwegian."

In some cases, the dialect may provide an opportunity rather than a barrier to research, as subject S24 recounts:

"I'm an expert in the history of the dialect, and so I started reading the playwright's text because they were written in dialect. So that's given me an edge that almost no other Americans have. Venetians have it, they can read their own dialect because dialect stays the same, but Americans have a lot of trouble with it because of the features, and it's hard. You have to learn just by immersing yourself in it. No one teaches it. Americans make a lot of mistakes when they characterize or summarize a document because, just to give one small example, the third person singular and the third person plural have the same verb ending, and so people don't realize, they don't look close enough at the subject. And Venetians will include a subject pronoun sometimes, but not always, and so you have to read very carefully."

6.2.6.4. Gaps in language skills

As stated earlier, each of the subjects of our study have expertise in the main language(s) of their research. However, research has a way of branching into additional languages or of requiring deep knowledge of a specific vocabulary that may extend beyond the researcher's training or expertise. In cases like this, scholars may rely on colleagues to assist with translation, adding an interactive human strategy. Subject S46 provides an example, "Much of this project is going back and forth between Dutch and English; there's some Indonesian in there that I can't read very well, and I work with one of my friend's graduate students if there's something I need translated." Subject S31 adds, "There will be certain Czech or Slovak that may not directly translate and… I sort of rely again on kinship networks because of my limited language fluency maybe compared to other people who are completely fluent in both languages or maybe native in more than one language. So that becomes a big issue."

239

In addition to relying on colleagues for translation help, a few subjects mentioned the use

of electronic and on-line translation tools, with languages scholars being the most skeptical of

their effectiveness. Subject S43 says:

> "I use Google Translator too. There are some other translation tools I use as well;
> an online dictionary that at the word level is very helpful because it gives you
> synonyms and antonyms and breaks down the roots of the word. When I'm
> working with philosophical stuff, that is really great. It expands my understanding
> of how the words are put together and how the thought is developed."

The topic of translation tools came up when asking subject S38 how systems could help

him better. He replied:

> "Something that I value quite a bit, are translators - online translators - and
> they're not very good. I'd like to see more improvements in translators. I might
> translate and it doesn't sound quite right and I say 'Well let me look up this
> word.' I have my dictionaries here, but that's something that could be helpful."

A concept that appears to be unique to scholars of Japanese is the pocket translator.

According to subject S33, they are much more sophisticated than what people tend to associate

with tools for tourists:

> "I have an electronic dictionary that looks like a little calculator, and I purchased
> it in Japan. They're very common in Japan; I never see them in this country, for
> whatever reason. It's got I think upwards of a hundred different types of
> encyclopedias and dictionaries, you know, on the little chip that runs it. It has
> encyclopedias. It has character dictionaries. It has English-Japanese and
> Japanese-English dictionaries. It has dictionaries of sayings, you know, like,
> aphorisms. It has all name dictionaries. It has various kinds of dictionaries
> housed within it, and that is often useful, too."

6.2.6.5. Librarians without language skills

A commonly cited theme from these subjects is that, although their local college and

university librarians are generally willing to help and have advanced information-seeking skills,

they are not knowledgeable of the language(s) used by the scholar. Subject S4 explains how

difficult it can be to obtain sources when the library staff do not speak the language and how the situation worsens as monolingualism in the U.S. increases:

> "In the case of my original work in Germany, I was dealing with highly developed professionals, and they always knew where things were or should be. When I was dealing with the German stuff in St. Louis and in New Orleans, and normally the person you're dealing with doesn't know that language. So you have to use strategies to get the information you want. And it's often very difficult because in some cases you're dealing with people who don't even think that these things should exist or do exist, and it becomes more difficult with time as the monolingualism of American academics increases."

Similarly, subject S11 states how lack of language expertise precludes any really substantive assistance by her institution's librarians, "As far as getting French sources, I think the big things are that I can't call on a librarian at my university, so I can't use their expertise except in the most general way." It is perhaps notable that most of the scholars who talked about monolingualism being an issue are those who use more commonly spoken languages such as French, German or Spanish. Other scholars who use less common languages (e.g. Irish, Persian, Vietnamese, Dutch) do not even bring it up because they have absolutely no expectation of support in their languages.

6.2.7. Summary of language

The above section shows that language plays a significant and complex role in the information-seeking context. Many aspects of language, especially its written form, can create problematic situations primarily when using formal system strategies. Language issues encountered while using formal system strategies most often create barriers for scholars at the identification stage. Yet other aspects of language, including language expertise, can pose challenges to understand or assess relevance once a source is identified and obtained. To navigate their way through these challenges, scholars who use non-English sources have developed their own language expertise and may rely on that of colleagues as well. Most often,

241

information tools and information professionals in the U.S. provide very little assistance in overcoming language-based barriers.

6.3. Culture

The factor of culture is the most complex of the three factors discussed in this chapter, and it affects many aspects of the information-seeking context. The problematic situations caused by cultural factors most often appear in relation to informal resource strategies and interactive human strategies. The sections below explore the various aspects of culture discussed by participants in the study.

6.3.1. Need to experience the culture

Before we begin to explore the role of culture in the information-seeking context and how and when it contributes to problematic situations, it is important to hear what scholars have to say about the importance of not only understanding but experiencing the target culture. Subject S30 speaks of the need to understand the local perspective – in this case by reading local newspapers in their native language:

> "I wanted to read the newspaper from a local perspective, looking for any news reports on health issues, on nurses, how nurses were viewed in the community, public opinions as well as news. The Xhosa newspaper would give me more insight into the African or Xhosa point of view rather than the English."

In order to delve into their research areas, many scholars have discovered and spoke about the importance of being "in country" and experiencing the culture directly. In addition to providing more convenient access to primary sources, being in the target culture provides a number of advantages, as subject S40 explains:

> "Apply for as much funding and grant money as you can find and get yourself to Scandinavia. When you are studying source material, there are linguistic codes, there are norms, cultural practices, all associated with academic work that you cannot experience when you're in the United States. If you're going to be somebody who wants to spend their life researching and thinking about another

part of the world, it is incredibly important to conduct a lot of that research in country."

Subject S31 further explains that experiencing the culture directly allows the scholar to write and portray their research with more authenticity:

"Going overseas lets me get integrated into the culture and the space. So for a project right now where I'm writing about a building being built in the city. If I've never been to the city, I feel like my description of what's happening isn't as authentic as someone who can go there, travel there, and experience what Prague really is – what it means to be Czech. I think there's something about the intangible culture, the physical space, just being there, gives you a different level of understanding of what you're writing about, even if it's not directly related to finding the sources."

Subject S32 describes the limitations and dangers of only relying on formal system strategies and the value of pursuing informal resource strategies in the target country to develop a connection with the location and the subject:

"I remember warning the graduate students about the overreliance on online research… There are parts of historical research … that's about developing what I would almost call an intuitive sensory connection to your subject… You develop an intuition when you're immersed for a while and that requires immersion in the place. What happens is that students have less and less funding available to them because more funding agencies say 'Hey, you can get everything on the internet. Why should you go there?' And I think many historians think that's unfortunate."

Two eloquent and poignant examples were provided by subjects S38 and S41 about the intangible and precious effects of being in country. Subject S38 illustrates the value of experiencing the culture first hand:

"You have to go to get a taste of the place at least. You have your documents, fine, but there are intangibles associated with the documents or with information, and it's good to experience them. I remember interviewing a pied noir woman, and she said the sun is different here compared to Algeria, and sure enough, when I went to Algeria after that, she was right. The sun was different. It was beautiful; it was nice and warm, and it was brighter. Maybe that's the romantic in me… you can do a lot of work online without having to go anywhere, you know. But occasionally, I think it's important to go someplace."

Subject S41 relates how important it is for historians to have experience with the context of what they study. His target culture is quite far from that of subject S38, but his observation illustrates the same point:

> "What it's like to walk through a bamboo forest, it's a different sensation than walking through a fir tree forest. Knowing the culture, knowing the geography, having been there, I think it fills in a lot of gaps when you're doing historical research or teaching about that. It can provide… you're getting in character but also changing it from a two-dimensional to a three-dimensional world. There's so much historical information you fill in if you have been in the country, particularly if you live there instead of traveling through there."

6.3.2. Authorship and publishing culture

An aspect of culture that can impact both the existence of and access to sources is the authorship and publishing culture of the target country or culture. As explained in the literature review, only a few countries and languages have active and vibrant authoring and publishing cultures, and most of those are European. When asked if some of the articles she accesses are written in Swedish or Danish in addition to her target language of Norwegian, subject S40 points out that the publishing culture in Scandinavia may be moving toward Global English:

> "I think there is a huge push for publication in English. The biggest Norwegian literary journal – *Edda*… they accept publications both in English and in Norwegian in order to hopefully open up the scholarly community a little bit, but also… it sends the signal that the primary publishing language is English. And when you go to, like, the Society for the Advancement of Scandinavian Studies, the big conference in the United States, every paper is given in English."

She continues on to explain how collaborative the Scandinavian countries are in sharing their research with each other, but that outside of Scandinavia, their research is hard to find:

> "Across the board between Norway, Sweden, and Denmark, and even Finland and Iceland, there's this great… the scholarly community in that part of the world is so fantastic and so collaborative, and the joint projects between universities are so, so, so impressive. But if you're going to try and extrapolate any of those resources to the larger world, it's really hard to find them."

For scholars studying non-Western cultures, they will need to rely on in-country primary sources, but the secondary literature may or may not emanate from that same country, culture, or locale. This is largely dependent on how vibrant or significant the research and publishing culture is for that topic within that country. Subject S36 explains:

> "There's a real imperial thing happening where Western scholars, the countries they study provide the primary sources and Western scholars provide the secondary sources, then you interpret their raw materials using the tools of your own culture, so it's not necessarily something I find to be completely problematic, but I prefer to be giving more credit to the Russian scholar community, knowing about what they're publishing, but that is a problem for me, getting that information, making personal contact there and being able to find out what's coming from Russian scholars, it's not the easiest actually."

Subject S48 explains that, despite Egyptology teams originating from a wide variety of countries, "French used to be the language of publication because the French were running the antiquities service. Nowadays, it's really transitioned to English."

The language issue of less common languages (see section 6.2.5, above) can strongly impact authorship and publishing culture in either a positive or negative way depending on the culture's commitment to language preservation. As explained in section 2.3.2.1, there is often a disconnect between the number of speakers of a language and the numbers of publications coming from that country or existing in that language. Some cultures that are the home to lesser spoken languages have realized this and are making concerted efforts to preserve and promote their languages. Subject S6 explains how this phenomena plays out related to the Catalan language:

> "Anything whatsoever that has anything to do with the Catalan side of the equation (because there were Beguines on the other side of the Pyrenees) is in Catalan because Catalans have a real political stance about publishing in their own language. Older stuff is sometimes in Spanish, but these days, no Catalan person would publish in anything other than Catalan."

While Portuguese is not a language that risks extinction by any means, it is still rarer worldwide

and in academia than Spanish, for example, and that feeling of being a minority can bring a pride

and promotion to a language as explained by subject S25:

> "Yes, I do think there's definitely that sense of 'we are the guardians of the Portuguese language.' In fact, the only museum dedicated to the Portuguese language is in Sao Paolo, Brazil and not in Portugal. So there's demographic reasons, economic reasons, why Brazil is at the forefront of Portuguese language publications."

When there is an active publishing culture that is aggressive in digitizing published works

and making them available on-line, barriers such as cost may intervene to make it more difficult

for scholars to access sources, as subject S18 explains, "The Japanese are somewhat funny

because they're really on the leading edge and innovative in a lot of ways with digitization

projects and things like that, but much of that is very expensive, subscription-only." Copyright

laws also contribute to the publishing culture of a country or culture. Subject S36 explains:

> "You can get (sources) through Yandex or Google - databases of people's memoirs, most published works in Russian that are considered literary class have been digitized, much more so than here, because of the very weak protection of copyright in Russia, so people are much more likely to full text digitize things there, maybe not recent books, but anything from the Soviet period."

Related to authorship and publishing is the issue of the scholarly community in the target

country, language, or culture. Subject S42 explains that scholarship on her topic is hard to find

because it is rare in the U.S. and culturally very different in the target country of France:

> "There's not a lot of Yourcenar scholars in this country, and I find most of the French scholarship pretty innocuous and useless because it's like a small group of people publishing the same thing over and over, and it's like a little club, and it's not very far reaching. It's really not what we would consider serious scholarship here. The styles are so different, and the academy is so, so different. There's not a community here in this country so much."

In order to have a vibrant culture of authorship, research, and creativity, the environment must

support and encourage those pursuits. Unfortunately, subject S43 indicates:

246

"Somehow the layers of structure and so forth sometime squashes the empowerment of those who want to bring about new things. South Africa is even worse than Europe in prohibitive kinds of mindsets and layers of bureaucracy to tell you 'Go away. Succeed somewhere else. This is not the place to do it,' and that's the sad thing."

6.3.3. Internet culture

The factor of internet culture is significant enough to distinguish from publishing culture. Some countries and languages have an older and more significant internet history and presence. The prime example of this is Russia, as subject S36 notes, "Because Russia's a pretty internetted place, and Russian's enough of a major world language that just using the basics of my Mac, loading the alphabet and all that, it's actually very seamless to use internet things there." To have a mature and vibrant internet culture requires several factors to come together in a perfect storm. First, the infrastructure must exist for citizens in general and scholars in particular to access and contribute to internet-based resources. Second, the language in which these resources are presented must be spoken widely enough – either by a large population in the same country or by populations in other countries – so as to appeal to a significant audience. Third, the environment must encourage and promote scholarship and the on-line publication with minimal if any censorship. Relatively few non-English speaking countries or cultures might fit into this category, so countries like Russia, France, and perhaps Germany or Spain are unique in this regard.

6.3.4. Conditions in libraries and archives

Unfavorable and challenging conditions are not unique to overseas libraries and archives. However, in some cases, as the participants explain challenges they experienced in archives, they identified these as being more prevalent in the target culture. Various descriptions of conditions in libraries and archives are discussed below in what are often fairly long and colorful stories.

6.3.4.1. Facilities challenges

Physical conditions in foreign archives can be surprising. Scholars in my pilot studies reported documents in bags on dirt floors, oppressively hot work rooms, and insect-eaten documents, among others (Sabbar, 2013; Sabbar, 2016). The most severe of these are when archives are completely closed; subject S10 provides two examples from Germany:

> "The archive in Cologne collapsed because of construction of a subway that was poorly planned, and the building literally fell down. They're recovering the documents now; they will be able to recover most of them, but I would say Augsburg and Cologne are the two largest collections for this period of manuscripts, and right after Cologne collapsed, Augsburg was attacked by beetles that were eating the manuscripts, so they had to close the archive and take out all the documents to be fumigated. Then they reopened and eventually built a new archive in another part of town to put them in, but they're still fumigating those records and moving the more modern records. As a result, they were closed for two years, and now they are only open one day a week. This happened in 2009 when they closed. They closed at the end of 2009. They only gave us three months warning. In those three months, everyone was packed into the reading room trying desperately to get everything before they closed. Then they closed for about a year, and then they reopened for one day a week, and then they closed again for about a year and a half, and now they're open again for one day of the week."

Other simpler examples include the "many stories of enduring the physical challenges of the under-heated Russian archives in the winter" (S36); an archive in Turkey with unavailable sources "because the condition was too bad and it was under restoration." (S35); or old, broken microfilm readers in Brazil so "you couldn't print anything." (S15).

6.3.4.2. Conditions adverse to research

Even when physical conditions of the archival or library facilities are acceptable, the overall ambiance and "vibe" of the library or archives may not be research-friendly. Subjects in prior pilot studies talked about poor lighting or other uninviting qualities (Sabbar, 2013; Sabbar, 2016). Subject S43 tells of conditions in some South African libraries that make them inhospitable to researchers and scholars:

"A library is a scarce resource. It's overcrowded; they stand 60 to 80 deep by the Xerox machine. It's like a voting booth; they go around the block waiting to print. Everyone is in the library. You go to the library on a Saturday afternoon. The stores close at noon… any Saturday afternoon are their dead hours. The library is packed. It's the study hall for the community because they're so poor, and it's so noisy. There's no ambiance to study at home; it's the only place you can do it. It's so crowded that for them; it's just too much."

6.3.4.3. Facility hours

A prevalent theme among users of foreign archives are shortened, less convenient, or less predictable hours. Subject S24 gives an example of Italian archives:

"One of the things about Italy is that a lot of the archives are only open very limited hours, like three days a week for two hours each day and that kind of thing, so everybody who works intensively in Italy then has an entire patchwork schedule. Venice has one private library that's open until 11:00 or 11:30 at night, but all the rest are closed. The very latest that any of them go is like 7:00 or 7:30 P.M. It's sort of like here where you can't go to the library at 10:00 P.M., you've got to go when they're open and not on holidays."

Subject S15 finds similar issues in Brazil:

"I have one archive I'm going to try to visit, and you have to be very cautious about their hours, but what they don't tell you… not the National Archive but smaller archives tend to close for long lunches, so be prepared to get kicked out at around 12, and they might open by 2:30 or 3."

Conversely, subject S36 finds that library and archives hours in Russia are longer and more generous than in some U.S. archives, "They often have different hours. Their hours go until 8 o'clock at night; the library is open until 9 o'clock at night; it's open on weekends, so you can do a whole lot of work in a lot of different ways, very efficiently there. You go to the U.S. archives, and they're only open nine to five."

6.3.4.4. Closed stacks

Scholars who began to do research in the United States and now conduct research in other countries admit to being somewhat spoiled by open stacks in U.S. libraries and frustrated by closed stacks elsewhere. Subject S6 gives an example of France:

"I was in Paris, and the Bibliothèque National supposedly had copies of this journal. But they're the kind of library where you have to submit a request, and they will bring you out the volume. You can't go into the stacks. So, I made my request, and despite several conversations with the circulation desk and kind of going back to a superior or something, they were unable to locate the volume. And I thought that's ridiculous. If you could let me into the stacks for 5 seconds, I could find it. So, in Paris, I had to give up."

Subject S7 discussed her experience with closed stacks in Spain, as portrayed in section 5.2.2.1.

6.3.4.5. Closed records and unavailable documents

Most archives – whether overseas or in the U.S. – have policies regarding closed records. These records and documents may be unavailable for a specific number of years based on their official nature or their sensitivity. Which records are closed and for how long are determined by archival and copyright policies whether at the local or national level. In addition, archivists in some special collections may hesitate to make rare and fragile documents available – especially to outsiders – ostensibly to preserve them and prevent them from harm.

Subject S31, who conducts research in Czech and Slovak, knows that she is fortunate not to be affected by closed records related to her topic: "There are less issues in periods that I'm studying than more contemporary history, so I have that advantage in terms of road blocks. A lot of my stuff is declassified or available because the people died, essentially, so, there's less concern about protection of the legacy of some of these individuals."

Subject S6 understands the motivation of the archivists to protect ancient documents, but she has also learned how to convince the archivist to grant access:

"All of the books that they have in their manuscripts collection, for them, they are holy relics because they are found there. The guy who owned these books is a Saint – John of Capistrano, the same guy with the swallows… There was only one friar left in the entire convent… and (one of the visiting researchers) didn't lock the cupboard or something, and he just said 'I'm sorry, but you can't see them anymore.' So we had to just sweet talk him for a day so he would let us back and look at the precious relic books."

250

6.3.4.6. Items "gone missing"

Items in any archive can be misplaced or deaccessioned, and several subjects in my study

related stories that could likely happen in the United States as well. Subject S23 explains how he

would request access to an archival resource and a common response:

> "You go to Western Turkey and you spend a couple of hours having chai or tea
> with the kasheesha (local priest) of the local church, right, and because you think
> something's there because it said so in the catalog and you're trying to get into his
> good graces and at the end he says, 'Well that… I think, that was given away by
> so-and-so, as a gift, way back when.' So nobody has a clue sometimes where
> something has gone or where it has ended up, and this is particularly true in the
> Middle East. Sometimes very difficult to trace where things are."

In one of the most amusing stories of this study, subject S10 cites a number of phenomena that

have led to the loss of records in archives and libraries that she uses in Germany:

> "If something was produced in the 16th or 17th century, it was either printed in a
> print shop or written by hand and there's only one. There's multiple copies of
> books, but every court record, there is just one hand written thing and whether
> they survived or not is pure chance. Aside from collapsing archives and beetle
> infestations in this century, there have been wars that wiped them out, many
> archives were bombed out in World War II. A lot of records on Jews were
> intentionally destroyed prior to WWII. There were other wars and fires; there
> were pigeon droppings. I had one case where the reading room was on the top
> floor of a seven-story city gate tower that was the archive, but at one time it had
> been the storage room for some of the records, and I was working through budget
> records, and I found that there were always six in a row and then one or two
> missing and then six or seven in a row and then one or two missing, and I asked
> the archivist why every six or seven years there were one or two missing, and he
> said because they were all stacked in piles on that top floor, and someone left that
> window open for like two years. They were completely covered with pigeon
> guano, and the pigeon guano completely ate down through the paper and went
> down through the top one or two in each stack. There's all kind of things that can
> happen to archival records, so anything that has survived as long as it has is kind
> of like a little gem."

Subject S30 recounts additional, more political reasons for documents gone missing related to

her research in South Africa:

> "My major challenge with my first project was that at the time in the 70s, the
> police in the South African government had confiscated a lot of documents. This

is why oral interviews became so important in my research, because they either shredded or lost those documents because they are nowhere to be found. There were also other missing records because of fires. There was a fire that destroyed some church records. There was a printing press that had printed some papers for an organization that I was studying, and when I called to ask if I could come look at their records, they said 'We don't have those. They don't exist. There was a fire in 1980-something. They're gone."

With ancient documents, their destruction and loss may mean that a piece of history has also gone missing, as subject S20 explains:

"In the stuff that I work on, like Roman comedy and Roman erotic energy, there's a lot fewer copies and sometimes there might only be one copy of this play, and so when a worm got through the book or when someone spilled coffee on it, or wine on it in the copying process, then those words that were eaten away or stained away, are gone forever and scholars are doing their best to fill in the blanks."

Sometimes it isn't clear if the items are truly gone or if the archivists are not willing to make them available, as subject S41 recounts related to his experience in China:

"For foreigners to get access, you know I always bring a letter of introduction with me, a lot of the times it meant nothing to these people, and they would make up stories and say those documents don't exist anymore. I would say I saw these documents in these footnotes; they say they're in these archives, and when I went there, they said they had to get rid of those old documents because they didn't have room for the new stuff. My guess is that they still existed somewhere, they just didn't want me to see them for whatever reason."

6.3.4.7. Collections poorly indexed or organized

Even when documents exist, they may be difficult or very time consuming to find in the archives due to poor or non-existent indexing or description. Subject S8 finds challenges and provides suggestions regarding his research in France because much of the relevant collections are not indexed or described, "I've also used the archives of the Paris Mission Society, and these archives are not well-indexed. They're not online, so basically it's a matter of going there and just going through materials, just looking for what's appropriate, just going through hundreds and hundreds of pages of documents."

Subject S35 attributes poor indexing to several different factors that make the sources she needs harder to identify, and she credits a recent indexing project for making some of her research much easier:

> "There's a really important collection of documents in the main Ottoman archive that… which order it has been done in is a mystery to me; they're not consistent. There is a section that has been transliterated into modern Turkish and indexed, and that has made for my current project really a big difference. Sometimes the person I'm interested in appears in the document just as a person in the document. It's not to him, it's not about him, but nevertheless, the fact that he's somewhere mentioned in this document is really important, and if there wasn't an index, I wouldn't know because where in this document he could be could be very variable because he's not a really important person… Because the handwriting is so terrible, there's no way you could look through every document because there are thousands, and they're hard to read. When something like this happens, when they decide someone is going to go through to put this into modern Turkish and have an index with all the names, it's just wonderful. That's like one hundredth of a percent of all documents. Every now and again something like that happens, and it's a great resource."

A somewhat common reason for items being either unfindable or going missing is that the archive has been reorganized with no updates to the finding aids, or the organization is so unique as to make finding materials nearly impossible. Regarding archival re-organization related to her research on Venice, Italy, subject S24 says:

> "(This author) discussed at length the material in these documents, but shortly after he did so the archives were re-organized and he gave them a kind of odd name of his own. He gave them a title of his own which didn't correspond to the archival fonds. And so no one really took it all that seriously because nobody could use or nobody could find the documents again."

Subject S11 provides an example of problems brought on by unique and inconsistent archival organization: "And the other is just the problem of the French archives and their organization, trying to do a search from afar because each archive organizes itself differently and their search engines are different." This phenomenon is understandable given the history of libraries and archives in France whereby the treasures of the people were distributed amongst the

253

people, and archives were established as autonomous entities to be organized as determined by local constituencies (Manross & Sabbar, 2010).

6.3.5. Political considerations

Political factors can be a major subset of cultural factors and can contribute to the information-seeking context, most often in an adverse way. Fay (1994) discusses the interaction of politics, culture, and human behavior, and the scholars in this study very often make note of such interactions, considering them a formidable aspect of culture. Scholars conducting research in the U.S. take for granted many of the phenomena discussed by the participants in this study. While most are not common worldwide, when they do arise, they present formidable barriers to research both at a broader macro level and at a more targeted local, micro level. Several specific political considerations are presented below.

6.3.5.1. Dangerous political climate

As governments change, the political climate of a country can make it more difficult or even dangerous for a U.S. scholar to travel to the target country to conduct research. Subject S17 conducted research in Iran earlier in his career, but now he says "going to Iran for me is not a good option, but I have for instance made friends with a colleague there who is the editor of the academic journal on music, and …he keeps me abreast of publications." Similarly, subject S49 conducted her dissertation research in 2011 – just after the revolution in Egypt – "as a single female travelling around the country" and taking photographs. She reports how the environment changed by 2013:

> "About two years later, in 2013, I went to the major Egyptological conference and ran into the people from the Cairo office and said 'How are things in Cairo?' and they said 'Well, it's hard to be female and walk down the street. The government is cracking down on journalists, so anyone with a camera is immediately suspect, and being foreign is not easy.' And I had a moment standing in the hallway of the

conference hotel going 'Oh, so my research fellowship wouldn't be possible now.'"

She continues to explain that government permission is needed to access archaeological sites in Egypt, which can create barriers in normal situations but has recently become even more problematic.

> "In some situations, access to countries and resources has improved. Subject S44 studies Vietnam and tells how access has recently changed and how some aspects remain difficult:

> "Beginning in the late 1980s, people began to be able to go to Vietnam for research, where before it was completely closed, at least to most everybody. It's still challenging to do research in Vietnamese archives. The libraries are a pretty easy place to research; they have a much more open policy; you can bring a digital camera; you can photograph materials. You cannot use a digital camera in the archives typically, and basically every document is considered classified, so it can take two to three weeks between requesting and when you can look at it. If you want to photocopy something it can take another two to three weeks to get the photocopies."

Not only are civil wars and revolutions dangerous for scholars, they may also be dangerous to the sources and documents themselves. Subject S23 expresses his concern:

> "There are still a couple manuscripts that I've been looking for for years now, nearly a decade, that are in Damascus and it used to be that the Patriarchate there, the Syrian Orthodox Patriarchate, it used to be that they wouldn't let people look at it because of fears of Westerners, again, nowadays it's the civil war going on, so I have no clue of what's happened to these very, very important manuscripts."

6.3.5.2. Government and regime change

A common cultural phenomena in our post-colonial world is that governmental and regime change have created changes in the language, the structure of libraries, processes related to access, and more. A common effect of these changes is the change of alphabet used by a language. While language grammar and structure are very resistant to change, the written representation of the language can be relatively quickly changed base on the occupier. This has

happened in older times with Egyptian and Turkish and more recently with languages like

Vietnamese. Subject S44 give some examples of issues related to research in Vietnamese:

> "Basically before Western involvement in Vietnam, Vietnamese was written
> using Chinese characters, the literary texts were written in classical Chinese but
> pronounced used Vietnamese pronunciation and then there was a hybrid script
> that developed out of modifying Chinese characters which was called 'Nom',
> which was used for also for literary purposes but more kind of a demotic
> character. I guess sort of high culture and philosophy and poetry, were modeled
> after Chinese literature was written using classical Chinese and then more kind of
> popular or low or rivaled humor – texts and plays - were written using this
> demotic script which is kind of modified, quite complex set of characters built out
> of different Chinese characters."

Wars and national reorganizations, whether related to colonization or not, also have an

effect. Scholars who have worked with East German sources, for example, have found that the

unification of East German and West German collections was a time-consuming and perhaps

imprecise process that often made sources more difficult to identify and obtain. The French

revolution and subsequent Napoleonic rule led to a dispersed model of libraries and archives in

France (Manross & Sabbar, 2010). Subject S33 gives an example related to postwar Japanese

language reforms, again related to alphabet:

> "As a result of the postwar reforms during the occupation period of Japan, one of
> the many results of those reforms was a sort of renovation, if you will, of the
> Japanese writing system to simplify some of these characters that were originally
> adopted from Chinese and had been progressively simplified in both countries
> over time so that they now look quite different from one another."

Similarly, subject S31 explains the challenges she encounters when searching for place names in

Slovakia that can be attributed to governmental and subsequent language change over time:

> "I do the work on Slovakia also, they were part of the same state at the time, and
> Slovakia for a long time wasn't a country, it wasn't even Czechoslovakia, it was
> Hungary, so we were talking about the Hungarian roots that were all written in
> Magyar, which I have no familiarity with, so anything pre-1918 in terms of
> Slovakia is an entirely different language. The Slovaks have some differences
> from the Czechs, their own language, so when it comes to writing about Slovakia,
> the Slovaks kind of renamed what they see as Slovak or Slavic names, but then it

also gets renamed after the changing to the republic after that divorce, so it's incredibly hard to do Slovakia, as well."

Russian scholar S36 tells how "there was alphabet reform after the revolution and because all my work has been after that, I've never run into any of that." Subject S45 elaborates, "The Russians got rid of some letters in 1918, and that really hasn't been an issue. You can probably spell it the way it would be spelled now and still get to it. Usually these were just letters that they didn't use too much as far as I can tell. They were kind of obsolete."

6.3.5.3. Limitations based on policy or practice

Participants enumerated various limitations and barriers that they encounter based on policies or common practice within a specific country, culture, or institution. Many of these cause scholars to be less productive than they intend to be, but most have learned to take these factors in stride and work through them. Subject S40 provides this example:

> "Norwegians move at glacial speeds… The American efficiency model is just not present, partially due to the oil money that is there, partially due to the fact that people work about six and a half to seven hours a day, and that's it. And so, if you're going to get your library card, it will probably take a day. For you to get an appointment with a librarian will probably take a week or two."

Subject S8 compares another type of policy between two countries (France and China) where he conducts research, "At the Paris archives, I'm not really able to make photocopies, so I have to rely on copying material myself. And in Chengdu, I was able to do several hundred pages of photocopies, so bringing back material that I could work through at a more leisurely pace."

A phenomenon mentioned by several scholars – primarily those working in Europe – is a limitation to the number of requests that can be made in an archives on a given day. Subject S24 says, "I mean I was there for a year and you're only allowed three requests per day in this archive because it's rather badly understaffed and there are a lot of people there, and so you're limited to three requests a day." She explained that she experienced the phenomenon in Italy because of

negotiated labor laws, and "this is the case in France because it's very much a kind of Latin-based Mediterranean mentality, is that with all the unions being so strong and there often being very strict limits on things you can ask for like numbers (of items) and hours in the day and that kind of thing."

Though not technically related to policy, subject S15 mentions a situation that arose in Brazil that is considerably less likely to happen in the U.S., "I was in Brazil last summer with the intention of doing research, but the archivists are public sector employees, and they had all gone on strike."

Despite policies and procedures in place, subject S36 explains a possible exception:

"Generally, Russia is very good at having a way to work around rules effectively. So a lot of the time things are more efficient there because you don't really have to listen to the rules. There's a limit of ten things per order, but they're willing to work with you, and it's usually because ten things will keep you busy for a while.... It's less about government refusal to do anything for you."

6.3.5.4. Hesitance to show documents to foreign researchers

In some specific cultures, researchers pretty uniformly find resistance to showing documents – especially governmental documents – to foreign researchers. While subject S36 finds gaining access to documents in Russian archives to be normally quite simple, she found that scientific documents related to Russian space projects were more protected:

"I emailed them and made an appointment and met with them and talked to them, and because it has to do with satellites, which seem to have to do with the space race and to be important and may involve secrets and spies, none of which I know anything about, and God knows I could see satellite documents all day long and never help anyone build one. So I'm thinking about how to convince them not to be threatened right in this particular geopolitical context, by my interest in them, that my interest is purely scholarly."

Subject S41 tells a story that illustrates how the Chinese limit access to documents for their own scholars and even more for foreign scholars. His observations were echoed by other scholars of Chinese as well:

> "I was on a Fulbright scholarship a few years ago in China and I was giving a talk in Beijing and afterwards a few historians from Beijing got together and we were just talking about research topics and it kind of went to the idea of truth, is there truth in history? And one of the Chinese guys said yeah, there is truth in history but sometimes it's not the right type of truth and I think that really did encapsulate one of the big problems in China. If you're Chinese historian working in China and you find some very interesting materials and you want to publish that material, it has to go through a censor and that censor can say, 'No, that's too sensitive.' There's a couple of different levels of censorship. If you happened to get past that and you published it and it was too sensitive, the censor would be in trouble and so would you, and so would the person who gave you the materials in the first place. Because anything can be labeled state secrets… in fact just this year they put out a new policy about what's a state secret, more or less. And that's the fear of archivists. Working in China, they have some sort of checks and balances, because a historian isn't going to ask for something, or the censors aren't going to mark it down so you won't get in trouble, but working with a foreigner who can take that material and then go abroad and publish his or her opinion on that material, could get that archivist in trouble."

Subject S39 speaks about access to materials by foreign scholars in India as it relates to gaining the needed credentials to gain entrée and conduct research in country.

> "It says nothing of the issue of gaining time and legal access to the National Archives, which is a frustratingly bureaucratized process because India has a particular class of visa for research. You don't apply for a business or student visa; you apply for a research visa, and the research visa requires a substantial description of your project – a description of what kinds of sources and archives you expect to use. So this allows the government to basically refuse to issue a visa to somebody who will carry out ethnographic fieldwork in an area where there is a lot of, say, politicized violence, so that the scholar is not doing work that makes the government look bad."

Subject S44 tells the type of nightmarish story that all scholars are afraid of that illustrates how political situations can pose a very real threat to success in research:

> "They keep a close watch on the kinds of documents people are looking at and making sure they are keeping with the research plan they have already submitted, so in certain locations I was told I couldn't look at a certain file because it wasn't

related to my research, so it can be difficult. I had a colleague doing research in Vietnam, and when she got to the airport, all her documents were confiscated by the airport police because they said while she was allowed to have the photocopies and read them, she wasn't allowed to bring them out of the country, so she ended up losing a whole year's worth of research. I think it's mainly the government that is making guidelines to the people that work in the archives, and they seem to be zealous in enforcing them just in the fear of doing something wrong and things getting out. There are people who end up get visas denied for writing things that upset the Ministry of the Interior. So you can apply, and I have colleagues who have not had visas granted to them to do research in Vietnam after having publishing something that had to do with something the authorities considered inflammatory. This has very often to do with ethnic minorities or an unsavory aspect of the history of the communist party. Aside from the distance and needing funding in order to make archival visits, the real impediment is the cultural surveillance in Vietnam because they don't really like people looking into history. It can be difficult to get permission. When you first arrive, first you get a visa sponsored by a Vietnamese institution. When you get there they have to write a letter for you to take to the archive, and then it can take up to a week until you can look at the finding aids. Once you've located materials, another week to three weeks until you can get the materials, and very often you will get denied. I have gotten denied for things that I thought were perfectly anodyne, like a report on an orphanage from the 1920s. I was told I was not able to look at it. That kind of paranoia by the authorities I would say is the biggest impediment."

In addition to a general mistrust of foreign scholars, some countries and cultures have more specific anti-American feelings which can lead to increased reticence to make sources accessible. Subject S25 says, "There is a resentment toward any American inquiry into Argentina at this point. My French Canadian last name is hurting me in terms of the Anglo-American bureaucracy, but it has done me a few favors in Argentina because I'm immediately perceived to be French or French Canadian."

6.3.5.5. Limitations by IP and filtering

For researchers who need to access overseas resources from their home base in the U.S., they may encounter technological barriers. Organizations and governments can filter content by the IP addresses trying to access them, or they can filter content outright so that it is not available at all. Subject S47 gives an example, "It could be that because it's coming from a foreign IP, the

Chinese don't want foreigners accessing this kind of thing. Historical matters in China can be classified as only within the party. When you do archival research on China, that happens quite a bit."

On a related note, subject S48 tells of working with a fellow scholar from Cuba and being unable to contact her using common electronic communication tools. "The means that I set up for the 17 collaborators to talk to one another is via a Google site, but she doesn't have access to the internet at all. All she had access to is email, and all of the email she gets is censored."

Sometimes the limitation is even more severe than by country. Subject S18 provides a specific example related to digitized records in Japan:

> "So I really have to go to Japan. This is kind of a… constant issue. So, to give you a perfect example: the National Diet Library has a major project to digitize historical materials, and particularly in the two time periods that I'm working on right now, but some significant number of those materials, you actually can only view on-site in the Tokyo Main Library."

6.3.6. The role of culture in gaining entrée

In addition to the role played by the political constructs of a country or culture, there are other more purely cultural factors that come into play especially when scholars travel abroad and request access to libraries, archives, and materials. In the recruiting process of this study, subjects were often initially skeptical that they could provide insight into the role that language, culture or geography might play in the research process. Once I would mention the word "bureaucracy," the usual response was one of "aha." Very often a chuckle followed and sometimes an immediate anecdote. Human beings are often very cognizant of the differences between the bureaucracies they are accustomed to in their own native culture as compared to those they encounter in other cultures. Most researchers have learned to take these cultural difference in stride and even find them amusing, but they are nonetheless very aware of the

barriers caused by these cultural factors. This section includes both access to library and archival collections as a whole as well as access to specific materials once inside the library or archive.

The term adopted by this study to describe the accomplishment of being allowed into a library or archive and being given access to the materials is "gaining entrée." It would be only a slight exaggeration to say that each country or culture has their own distinct set of cultural differences that must be navigated in order to gain entrée to the resources being sought by the scholar. Subject S20 explains:

> "I work on these texts that are ancient curse magic texts, and they are physical things that survive, and they're located in different museums in the U.S., but the vast majority are in Athens, in Rome, in Berlin, and even if I had the ability to get over there without difficulty, there are within those bureaucratic and culture barriers. The Berlin team does not seem to be sharing their work with anyone besides the Germans and people already associated with the University of Berlin, and the Greeks have very specific procedures for even getting to look at those sorts of things."

6.3.6.1. Gaining entrée by credentials

Getting access to archives overseas most often involves pre-determined processes that require requesting permission. It is very unpredictable, however, which archives might make the process simpler or arbitrarily more difficult. In some countries and cultures, and in some specific archives, scholars have found their credentials as a university professor with a doctoral degree to be very helpful. Subject S30 explains, regarding South Africa, that "culturally, archivists do like PhD or professor researchers, maybe because they don't get as many and are excited about what people are doing." There are, however, protocols to be followed, especially to access government documents. Her account also provides an example of hesitance to provide access to foreign scholars:

> "I have definitely done the run around with the East Cape Department of Health. That's a pretty different game because dealing with governments that don't have archives… I wanted to access the hospital archives, so I had to submit a letter of

262

request. I knew I had to be really respectful of authority and pay deference and make sure that in my communication, I was using the right titles, but I knew I really had to follow up. You're talking to someone who's the secretary of the office. They have way more things to think about than your letter of permission to go look in these archives. I did recently receive a letter that it is ok and as long as you followed this protocol. But they sent me through the legal department, and I think a lot has to do with the health department nowadays. It has been targeted in the news for corruption and also not doing well delivering services. So here's this American coming in asking to look at their archives. Even though it's way in the past, they're still very sensitive and suspicious. I just had to be aware of that and make sure I'm covering my bases so I'm not going to be accused of doing something I'm not really doing."

In Russia, one's status as a foreign scholar is most often sufficient to gain access to

relevant documents, as subject S36 explains, below. She finds that showing up unannounced can

also be advantageous:

"That was another one of my major Russian research strategies… just to physically go there, because it's awkward for them to decline to help you when you just show up in their waiting room as a foreigner, so they kind of let me in and showed me around and they have staff to help you, and it was really lovely… It's very specific to the post-1991 relationship to the West in the sense that Westerners come, and it's quite imperial; it's definitely using a kind of privilege that foreign scholars do not have in the U.S. As a foreign scholar is Russia you have that privilege that you could use and throw around and they take you seriously much more easily and can't be rude to you in a way that definitely is not the case here, and I'm sure in other places."

At the opposite end of the spectrum, some countries such as India have multiple layers of

bureaucracy that scholars must overcome in order to conduct research in their archives. Subject

S39 explains:

"In addition, to get a research visa, one must be affiliated with an Indian institution of some kind. Normally it's a university. The problem, of course, particularly if you want to work at the National Archives in New Delhi… The two big universities there and because of their bigness, the time frame before getting these institutional affiliations arranged is glacial. It's not quite the Indian-legal-system glacial, but it's still slower than one might want. There is a group called the American Institute of Indian studies which is relatively old and is based in Chicago at the University, and what they do, among other things, is facilitate the institutional affiliation among scholars in the United States to do work there and

even then, AIIS's offices are slow to process paperwork if one is not being funded by that organization."

Another tale of bureaucratic procedures and red tape is retold by subject S44 about Vietnam:

"It's very bureaucratic. It can take anything up to a week to ten days, sometimes longer once you submit a request for a file to be given access to it because it goes through multiple levels of approval. The person who runs the reading room has to approve it; the director of the archive has to approve it; and then somebody at the basically the ministry – the archives in Hanoi has to approve it. The form that you submit gets signed by these officials and then faxed to Hanoi and things can take days and days to get a response and it takes further time for them to retrieve the documents from the storage facility which is not in the same building as the reading room."

Subject S49 explained the multiple institutes in Egypt set up by various countries, each housing their own non-circulating archival collection and each having differing protocols for gaining entrée. She recalls that she was allowed access to the Austrian Institute because she knew someone who worked there. For the German institute, she explains that she "had to apply for a reading card in order to get in."

6.3.6.2. Culture-specific standards of politeness and approach

Politeness may seem like a basic skill that one could take for granted, but specific cultures have higher expectations and more complex codes. These cultural differences can cause barriers to access if the librarian or archivist feels disrespected or offended when approached by the scholar for access to materials or for assistance. The most common and prominent example is that of *keigo* – the language of politeness (Sabbar, 2013; Sabbar, 2016). Every Japanese scholar in both prior pilots and this study talked about the art of keigo; learners of Japanese as a second language are often required to take a course to specifically learn to use the requisite forms of address in order to be considered polite and respectful. Subject S18 explains it very well:

"Japan is also hugely bureaucratic, and that is further complicated by their language, which has varying levels of politeness, and so there are whole grammatical structures and words that you use in polite Japanese, and there are

264

actually different levels of polite Japanese… the kinds of language that you'd use for important people, people who are a little bit more important than you, people who are maybe not as important as you but are guarding the access to the archive. There's a lot of subtleties and nuances to be aware of. Most foreigners - even people who are quite fluent – often are not that great with the super-polite language part. As a listener trying to negotiate maybe with an archivist or librarian or something, it is very typical with them to use quite polite language; it's a really basic interaction point. So, I might be prepared to have a conversation using plain language, so I ask my question, and their response to me is just polite language. And I just don't quite get the meaning."

In most other cultures, politeness is less regimented, but it is still worthwhile to learn how to best approach the archivist or librarian in order to get access to the needed materials. Subject S24 presents what she has learned from her experience working in archives in Italy and how using the right approach allows her to get the assistance she needs:

"There are two really important strategies that I learned in big archives … First of all, say 'May I ask for your advice? I'm really not sure how to do this.' Oh, they love that… The second one is, if there's something that you need and you're not making any headway, just kind of stand there with this kind of puzzled… slightly fearful (look), like 'how am I ever going to finish this project if I can't get a person to help me?'… Then with the staff at the desk, they actually give you things that you've called up."

In another story, subject S24 tells how she learned to be deferential show respect for the archivist even though they had made a mistake and how that consideration was rewarded:

"I filled out a request list at the archive and the staff member brought me back something that had nothing to do with what I was looking for. And then I was kind of stuck because I thought I'm sure I'm on the right track here. There was a more knowledgeable staff member, and you don't want to embarrass one guy by asking his colleague in front of him, 'This guy brought me the wrong one. What's the right one?' because you never get any (resources or help) after that. So I waited until the more knowledgeable guy went out for his coffee break and I caught him in the hall where it was just him and me so the other guy wouldn't be embarrassed. I said, 'This is what how I filled out the slip. Did I fill it out wrong? Do you recognize what I'm looking for?' He said, 'I know exactly what you want.' It was a separate sub-designation within the main archival fonds. And he went and he brought the correct one right back. Then that gave me these five very large buste full of documents to work with."

265

In some cultures, offering a gift is a useful way of showing proper respect for the archivist. Subject S4 said simply, "In Germany, what you do is take a bottle of wine every now and then." When researching in France, subject S38 says:

> "Being able to get along with people… my director when I was doing research in France said 'Now make sure you bring flowers to the archives.' So I would do that and next thing I knew I had access to collections which I didn't even imagine. Just by showing courtesy and recognizing the dignity of a librarian, that's important too. I happened to mention I was studying Bennabi, I don't remember where he was from, but he kind of said out loud 'An American studying Bennabi?' and that meant something to me.

Subject S39 provides similar counsel related to archives in India and how to approach the archivists to successfully get access to materials:

> "I would say that being nice to archivists is a good idea, not like in the bribery kind of way, but to be personable, treat them like people and being understanding that they have lives and things they need to do. If you interact with them on a personal level, they will interact with you on a personal level. How one works in the archives makes a difference."

6.3.6.3. Lack of assistance from library or archival staff

As presented in section 5.2.3.3, scholars using archives and libraries overseas are dependent on the assistance of librarians and archivists, and generally, the participants in this study were complimentary and appreciative of the expertise and helpfulness of the archivists. However, it was a recurring theme that both incompetent and uncooperative librarians and archivists exist and that there may or may not be ways to work with them or work around them. Subject S6 tells one of the most comical stories of the study:

> "I worked in Montpelier for a long time, and in the municipal archives there, they had one competent guy, and all the rest of the staff was incredibly incompetent. One woman – the story in the archives was that she was actually demoted from the Lost and Found of the city of Montpelier. Like she wasn't competent enough to do the lost and found, so they put her in the archives. She couldn't find anything. And when the good guy went on vacation for a week, I was just beside myself. I would have the exact call number, and she wouldn't be able to get anything, and so I just had to do work someplace else for the rest of that week and come back when my friend was there."

Subject S10's story of an uncooperative archivist is more tragic than comical and demonstrates

the value of creating a personal network in order to gain access to material that would otherwise

be unavailable:

> "That monastery… the archivist there is very unpleasant, and a friend of mine
> said there is no way to get into that arching because you call up and make an
> appointment, and then you show up and he's not there. Or you might not be able
> to make an appointment; he will make up an excuse not to. Once you get in there
> he won't bring you the records. There will be some reason why he can't, he's just
> an uncooperative person. The only reason she was able to get in and get records
> was because she knows somebody at this other archive… just an older local
> person who just happened to know this archivist at the monastery and could get
> you in. So it was sort of a personal connection thing."

Subject S11 tells another story about archives in France that combines the factor of

request limits and the factor of an archivist whose service ethic leaves room for improvement:

> "The Bibliothèque Historique de la Ville de Paris, I think it's still this way, only
> four (items can be requested) in the morning and four in the afternoon. Even if
> it's just a newspaper clipping, that still counts. And even if it turns out to be no
> help at all, it still counts. They still have card catalog, and some of the cards were
> hand written in fountain pen from the 1920s, and one time I misread a numeral,
> and the guy ripped me up one side and down the other in front of the entire
> reading room for giving him the wrong number. Half of my requests were rejected
> after that with no explanation given."

Subject S36 has learned how to understand and work with the archivists in Russia and what

constitutes usual behavior. She tells this story:

> "There was this archivist who I had ordered a bunch of stuff, and she couldn't
> give it to me, and she started yelling at me for having ordered it. And it occurred
> to me that she was yelling at me because she was embarrassed. She actually
> wanted to give me this stuff herself, but it really wasn't her call. The people who
> were making those decisions were above her, but she was the one who had to
> deliver the bad news. So I always try to be very accepting of the bad news and say
> 'Oh, I totally understand,' It's really unpleasant to have to be the one to tell
> people who are there for a very short period of time and really want to get their
> work done and who are prone to stereotypes about how Russian bureaucracy
> doesn't work well."

6.3.6.4. Gaining entrée by developing a human network

Based on the various factors related to access to archives and libraries overseas, most scholars have found it crucial to success to build a human "kinship network" that they could call upon when more straightforward methods of gaining entrée fail. Subject S8 finds the Chinese concept of "guanxi" expresses well what he needed to build to be successful in his research in China:

> "I used the Sichuan Provincial Archives in about 1997. By that time, many archives in China were theoretically open… In Chinese, there's a word, guanxi, which means 'connection, relationship.' It helps to have this kind of guanxi. And I wrote to the archives, saying, 'I would like to come and use your archive. Can I do that?' And I never heard back, and in the meantime I had gotten a grant from the American Philosophical Society to pay my expenses, so I arranged to travel to China, and I went there not knowing whether I'd actually have access to the archives. In the meantime, I had made connections with Sichuan University, where they were much more open. The person who I connected with there was a former graduate student of someone I knew in graduate school, so here too there was guanxi involved. I was a visiting scholar of Sichuan University, and they asked me to give me a talk. Afterwards, the head of the Institute of History of the Sichuan Academy of Social Sciences came up and said, 'I understand you'd like to use the Sichuan Provincial Archives. I'll drive you there tomorrow morning.' So he came and introduced me, and from then on, everything worked."

Subject S41, who also conducts research in China, echoes the importance of knowing people:

> "I think that's one of the cultural differences with doing work in China. You have to gain the trust of someone in China who can help you, or an archivist who knows you enough that they can feel confident in giving you the information that is in the archives but is not necessarily open, so I would say that's one of the greatest problems. It's just one of those things you have to work around or take your time and cultivate a trust."

Subject S8 also developed his human network to conduct research in France. Here, he discusses the similarities between the normally closed natures of two very different societies, comparing his experience in France to his experiences in China, explained above:

> "Both China and the Roman Catholic Church are not exactly open societies. So the first archives that I used were the ones in Paris, and I wrote ahead of time, basically explaining my academic interest. I did say that I wasn't Catholic myself, but I was approaching the subject with respect. I wanted to assure them that I didn't have any kind of unscholarly agenda. And I was very pleased with how

graciously I was received. When I got to the archives, the archivist also introduced me to the librarian of the society, and the society had an excellent library of printed materials, and she in turn introduced me to a retired French bishop who had served as a missionary in this area. They ended up being very gracious."

He adds a specific insight into how he was able to gain favor with the archivists in France:

"When I first used the Archives des Missions Étrangères in 1989, I wrote in advance to explain my research and give my bona fides as a scholar. The archivist was gracious and welcoming, and I invited him and the librarian of the Bibliothèque to lunch in a nearby restaurant. The gesture seems to have been well appreciated. I made return visits in 1994, 2003, and 2012."

Subject S30 gives advice for working with archivists in South Africa, "Most archivists I know are pretty nice, but some people it is helpful to know people and say oh, so-and-so referred me to you." She also points out the importance of cultivating relationships over time in a reciprocal way:

"The local Eastern Cape archives, I know the people really well because they don't have many people like me going there, and they are smaller. Culturally South Africans are pretty welcoming anyway, but I think because I know them and chat with them and because we have similar interests, now there are a couple of archives where I can just show up and say, 'Hey, how's it going? I want to look at this.' Anytime I publish something, if I send it back to them, they really appreciate that because then they can show how their archive is being used."

6.3.7. Summary of culture

Culture is complex. As shown in the various sections above, many phenomena are unique to a particular culture, and many are unpredictable. Especially for scholars using informal resource strategies and interactive human strategies, the role of culture cannot be underestimated. Moreover, understanding cultural factors and preparing to overcome cultural barriers is crucial to their success.

6.4. Geography

Geography is the third aspect being considered, and its effects are simpler and more direct. All of the subjects in this study have a home base in the United States, yet the topic of their study is about another location. In most cases, travel is both time consuming and costly. Before travelling to overseas collections, scholars often calculate what might be accessible electronically, from U.S.-based archives, through inter-library loan, from colleagues and other scholars, and so on. They then prioritize the most crucial research activities that can only be done in libraries, archive, or bookstores overseas. Geography factors have greatest effect on informal resource strategies. Several distinct aspects of geography are discussed below along with explanations of their role in the information-seeking context.

6.4.1. The need to travel to collections

Before discussing specific geographic factors that may impact the information-seeking strategies, we look at the importance of and need for travel to access sources. For subjects who use the informal resource strategies of visiting libraries and archives, a common theme was the need to travel to access and browse collections. Subject S44 sums it up well:

> "There's really little substitute for spending time in the field. Vietnamese history and indeed the history of much of Southeast Asia is so based on local connections with local Vietnamese or Southeast Asia scholars to write letters to give you access to libraries and those kinds of things, that forming a network of scholars in country and also with other people, I think it is indispensable to being successful."

Subject S7 explains that, for her topic, there are materials that only exist in libraries in Spain:

> "So that gets tricky at times. For that, you have to be in the library. So at one point, I was in Spain one summer; I had to go to the archive in Spain. There are unique material that are nowhere- anywhere else. So then you have to apply for grants to make it possible."

Subject S31 gives a very specific example of organizational records that require overseas travel to access:

"And there are some sources that are just being sorted and being made available abroad. So the work of the YWCA in Prague, between 1918 and 1948 has just been in the past 5 years, sorted, classified and been made available to the public. So that is primarily… I have to go overseas to get that."

A common theme among humanities scholars, especially historians, is the need to see the original sources, and this in turn, most often necessitates travel. While digitized versions may be available on-line, and sometimes copies or other surrogates may be available in more convenient locations such as libraries in the U.S., being able to see the original document often provides enough additional benefit that it merits a trip to an overseas archive, such as the time when subject S4 "traveled halfway across Germany just to hold a particular page in the light just to see what the watermark was. It dated the document." Subject S23 explains that it isn't always necessary any more to access the originals, but there is often value in doing so. It is at the discretion of the scholar whether surrogates are sufficient:

"You sometimes can't make out very intricate details on copies of the manuscripts. If you wanted to distinguish between the first and second hands or somebody who's written later on in the manuscript, you really have to be there first-hand to see it. What I do more frequently is order manuscripts online from libraries, and I am able to receive them in good order. Places like the British Library will give you really high quality color copies of manuscripts you can just order over the internet, so it really makes it a lot easier to be able to access."

Not all archives allow visitors to access originals, so scholars must investigate and carefully plan their archival research to make sure that the trip is worth the time and expense. Subject S23 elaborates:

"When I go to the British Library or the Vatican, I'm always touching the sides of the manuscripts. You're able to hold them; you're able to put them on a scan and read them through. So yes, definitely you're looking at the original. And there's no substitute, I think, for looking at the original. I would not travel all the way to the British Library if I didn't have a chance to see the original. You can distinguish things that maybe are unclear in the copy. Even the best and highest-quality copies I receive, there are still ambiguities there compared to when you look at a manuscript; you can see it in different shades of light and different angles. It's particularly true looking at vellum manuscripts – things that are

271

written on skin. Often with dirt and water damage on it, you need to be able to see the original."

Even when the scholar isn't insistent on seeing **the** original document, they tell about the issues created by poor reproductions, especially of ancient and hand-written documents. Subject S35 encounters various issues with digital reproductions:

> "A lot of the things aren't actually the originals. Sometimes they are, but sometimes they're digitized, and sometimes you'd really much rather have the original than the digital copy because you're trying to use it on their computer screen, and it does weird things and puts it at a weird angle. A lot of the things I'm working with are hundreds of years old, and they could very well be in restoration, but it is frustrating when you finally found something you think is going to be absolutely wonderful, but you can't see it."

Many, many additional examples of visiting overseas libraries and archives are provided in section 5.2.2.2.

6.4.2. Location and distribution of archives

As stated earlier, one of the most fascinating questions asked of the subjects is how they know which archives contain the materials they need to advance their research. While, for some, they may be a central national archive that holds their best sources, there are numerous examples of archives, libraries and museums spread over a wide geographical area. The factors that determine the locations of these archives are complex and often related to historical events. Subject S10 frequents an archive in Augsburg, Germany, where she consults documents related to the history of the Jewish people in Germany. When I asked her why she uses an archives in Augsburg and if there would be an archive in Berlin that would be helpful, her answer explains that the importance of the city at the time of the period being studied is more significant than the current importance of that city, "Well, Berlin would be less useful because it was basically so much less significant in the 16th and 17th century. It came up in the modern period." Moreover, she explains the scattering of resources across Germany:

"In Germany, there is no central repository. It was in the early modern period, a very un-unified group of the Holy Roman Empire and in each of the cities was an individual city state basically, the territories. If I want to know about village life, I might have to go to the village archives, as well. It's a very decentralized system of record keeping."

Similarly, subject S15 tells about the distribution of archives in Brazil and the lack of central libraries in universities there. She says, "Part of the problem is that there's a National Library in Brazil, but the university system and the university infrastructure in Brazil is very different, so most universities don't have like a general library."

In a post-colonial world, resources for some countries and cultures may exist in multiple locations, possibly halfway across the globe. Subject S44 conducts his research on post-colonial Vietnam, and subject S38 studies post-colonial Algeria. Because both of these countries were colonized by France, and both use archives in their target country as well as the Archives Nationales d'Outre Mer in France. Subject S44 explains how history has determined the distribution of these archives:

"For me, the bulk of important documents for my research are in Vietnam because when the Geneva Accords was signed in 1954, the documents that were considered aspects of French sovereignty were repatriated to France, and things that were considered to do with local administration were left in Vietnam, and because they were from southern Vietnam, which was a colony of Cochin, China, all of the documents for Cochin China were left in place, so none of those were repatriated to France. I've had to use both places. The documents in France are helpful for looking for higher level policy decisions, reports on major goings-on like rebellions and unrest and economic data, but the documents that remain in Vietnam are useful for local detail and having a sense of what was happening in the village or who was writing to this or that provincial chief seeking some kind of assistance."

Some scholars have an easier time because the archives are all located near each other in a major city. This is the case for subject S36 who studies Russian topics:

"It's all in Moscow, for my work at least. So you could just be taking the subway around town and be using ten different resources in a short period of time. It takes time to do your orders and to get new stuff. Once you get a rhythm you can make

sure you get an archive every day, but those archives sometimes have off days, some of them are only open three days a week, so you can fit in a million different tasks in your free time."

6.4.3. Need to plan ahead for overseas travel

Scholars who use archives overseas have learned how to prioritize their research and plan

for their trips abroad. Since trips to the target culture or country entail time and expense, the

time there must be used wisely for activities that could not be done from the home office or a

nearby research library. Subject S33 conducts research in Japan and explains:

> "There are some things that you actually do have to physically travel to Japan to get. Occasionally you can put in an order from here, have them make a photocopy, and then send it to you. They'll charge you, but it's possible. But I wind up going to Japan once a year usually anyway, so I save up my requests for stuff I can't get here and do that in two or three very intensive days of searching all at one time and bring it back."

Subject S31 further explains how she leverages her time abroad, including making copies of

materials to study later:

> "When I do the work overseas, I tend to try to copy and gather as much as I can, because it is expensive to be over there, so I do such a broad net and then bring things back. I know I will want to dive deeper, and I have to plan another trip over there and start tapping into those other access pathways to get back and deeper into that issue or that person's history or that organization's history. Being over there, the advantage is I could just go back the next day and would know scholars on the ground that I could reach out to."

One type of division of work is to do research in secondary sources from the U.S. and use time

abroad to obtain primary source only available there. S36 gives an example:

> "I'm also going to go in May to work on the same project, but those obviously require a lot more planning and preparation. You also need to be there longer in Russia, you can't photograph things, so you go long enough so you have time to take notes and get access. Things take a lot longer there, so I try to go for like three weeks if I'm going to work in Moscow. I do certain kinds of secondary source work here but a lot of my Russian language research does take place outside of the U.S., in person at the archives."

274

Subject S10 explains how it is sometimes possible to find out what sources exist in a foreign archive by searching on-line OPACs or finding aids so that the scholar can decide if a trip is warranted at all. "You know what you're looking for and what words to search under in archival catalogs and stuff like that. A lot of archives (in Germany) have catalogs online so you can search online before you make the decision to go on the trip." Subject S32 makes a similar recommendation regarding research in France, "In the French archives and national archives, the main inventory is online. Figure out what that is and spend time before you go and familiarize yourself with the structure of the archives, how it is organized."

Subject S30 tells how she uses the interactive human strategies of conference attendance to begin to identify archival materials and asking an archivist to determine whether the archives she is considering will have materials she can use:

> "It's hard to know what's in South Africa sometimes, and it takes quite a bit to plan and get over there for these sources, so it's important to have somebody on the ground to tell you what they have. At the conference, I was talking to somebody about a particular archive in South Africa that has newspapers, and I had never been there, and I did a search online and figured there probably wasn't anything for me. This other, I think she was doing her PhD; she was telling me the archivist there is really good at knowing what's in the archive and telling you what they have, so you should just email her and ask her what they have. So that's where having people on the ground becomes really important."

6.4.4. Summary of geography

Because the vast majority of participants in this study perform research overseas, the factor of geography is significant in their information-seeking contexts. Cost and time are two significant factors that determine how often, to what extent, and how effectively the participants can pursue research on their topic. It is clear that all agree on the importance of being in country and that they plan carefully so that they can maximize the value of the time they spend there.

6.5. Comparison with domestic research topics

275

On three occasions during this study, participants were able to compare their own experiences with research on domestic topics that would not require non-English resources or resources in locations outside of the United States. These examples provide powerful insight to confirm the various examples provided in this chapter in support of research question 3. Subject S33 compares her situation to that of colleagues:

> "Researchers who are doing U.S. political systems; their language is English; they're living in the U.S.; and they're researching something on the U.S. That's a different sort of problem from someone like me who's physically located in the United States and working within the U.S. academic system, but my frame of reference is Japan, so the number of human resources that I have immediate access to is fairly limited. I can always, of course, email colleagues, and I do some of that."

Similarly, subject S44 admits to being envious of his colleagues whose research takes them to more accessible and convenient locations:

> "I think for people who work on world areas like Southeast Asia, the strategies tend to be a lot different from my colleagues who work in American or British history. For us things are still relatively primitive, having to travel all the way to Vietnam or wherever and go through decaying card catalogs to get anything. I am always envious of people who do British history because every book that has been published in Britain, I think until 1800, has been digitized, which strikes me as amazing."

Subject S31 was able to compare her own experience since her intended field of research changed from that of a topic related to the United States to a topic based in Europe. She explains this in the context of discussing how places are named:

> "You don't realize about the politics of naming places. When I did U.S. history, Springfield has always been Springfield for the most part, so it was a new exposure to me getting into European history and learning about how political the name of a town could be. My first dissertation proposal was this project I wanted to do working with the YWCA in St. Louis, but someone else published a really great book on it a few months before, and I felt like it would be really similar. Then I stumbled into this project and everything changed. The timeline, the types of resources I had, your whole strategy changes because there's this whole layer you don't have to think about when you're studying U.S. history."

6.6. Chapter summary

This chapter provides evidence and examples of how language, culture, and geography each play roles in the information-seeking context. Subjects made it abundantly clear that all three factors are significant aspects – often barriers – to their research and can create problematic situations. One might, however, wonder if some of these factors might also affect research done by scholars who do use English or research topics related to the United States. Since this is a qualitative study, there is no control group and no statistical analysis that might prove, for example, that gaining entrée to archives overseas is more difficult than in the U.S. or that certain types of documents are more difficult to obtain.

In all, the experiences of scholars who use non-English sources strongly support the notion that language, culture, and geography play important roles in the information-seeking context. Moreover, the participants in this study indicate that they have learned what barriers might arise and how to conduct their information-seeking effectively in routine, disruptive, and problematic situations by choosing effective strategies and making shifts among strategies to advance their research.

Chapter 7 – Discussion

7.1. Overview

The previous two chapters presented the findings from this study related to the three research questions:

Q1: What information-seeking strategies are used by scholars conducting research in languages other than English?

Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?

Q3: In what ways do language, culture, and geography play a role in the information-seeking context; in particular, how do they create barriers that cause problematic situations?

This chapter discusses the implications and the significance of the findings in relationship to the research questions. The significant contributions of this study are: 1) the unique and groundbreaking nature of the topic, being the first user study to explore the role of language, culture, and geography in the information-seeking context and 2) the new conceptual model of the information triangle. Following the theory of transformational validity (Cho & Trent, 2006), the implications of this study to effect change will be discussed. Finally, limitations of the study are discussed.

7.2. Theoretical implications and significance of the research findings

Because these research questions have not previously been studied, this dissertation breaks new ground related to all three research questions. Theoretical implications and significance are discussed as related to each of the three research questions, below. Moreover, the theoretical significance of the information triangle model should not be understated. Prior theories most often utilize rigid models to visualize user behavior where the components and

arrows are static. My model seeks to provide a more flexible framework that is unconstrained, providing a structure that can be used in diverse situations with diverse user populations to portray any type of information seeking thread. In any new study that uses the information triangle, the list of information-seeking strategies can be built inductively, and each case of each study can be captured using this flexible tool.

7.2.1. Theoretical implications and significance related to strategies

The question of information-seeking strategies has been considered by prior scholars such as Bates (1979a), Marchionini (1988), Ellis (1989), Belkin (1993, 2010), and many others, each of whom developed their own taxonomies of strategies. This study, however, does not use any pre-existing list of strategies, nor does it try to specify strategies based on their specific tactical components. Instead, the list of strategies was compiled inductively based on terminology used by the participants combined with common terminology found in information studies literature. It is notable, however, that several of the informal resource strategies identified in this study resemble those enumerated by Bates (1989) in her earlier works, as shown in Table 7.1, below.

Table 7.1 – Strategies compared to Bates

Strategy from Bates' prior studies	Strategy from this study
Footnote chasing (backward chaining)	Citation tracing
Citation searching (forward chaining)	Chaining forward
Journal run	Journal run
Area scanning	(does not exist)
Subject searching in bibliographies and abstracting and indexing (A&I) services	Using a print index
Author searching	(does not exist)

However, Bates' work predates common use of online search tools that make formal system strategies possible, and interactive human strategies were not within the scope of her study, so the strategies named in her work cannot be closely compared to the strategies from this study.

Since the definition of strategies used in this study is a multi-dimensional one, strategies with varied and specific entities tend to be unique. For example, prior literature, such as the studies of Marchionini (1988) draw a distinction between search-based and browsing-based activities (a concept that is fundamental to the categorization of information-seeking strategies and the form of the information triangle,) but he draws no distinction between the browsing-based strategies of using libraries, bookstores, or private collections. Neither Bates nor Marchionini include activities that are characterized by human interaction within the scope of their studies, nor do hybrid strategies appear in their work.

As a result of both the inductive methods used to create the list of strategies and the multi-dimensional definitions, several of the strategies enumerated in this study are new and had not appeared in prior literature. In the category of formal system strategies, previously uncited and unexplored strategies include: Search Google Books or Google Scholar, search an online archival finding aid, search in a discovery tool (e.g. Summon), searching social media or streaming sites. In three of these four (Google Books, discovery tool, and social media,) the entities are relatively new phenomena yet distinct enough to be considered separate and discreet strategies by the participants and for the purposes of this study. In the category of informal resource strategies, several strategies are also unique, including using private collections, browsing in bookstores, and reading newspapers (on-line or in print.) Again, because strategies in this study are defined both by method and entity, and because the entities are identified somewhat narrowly, gathering the data inductively allows strategies such as browsing in archives

280

in the U.S., browsing in archives outside of the U.S., and browsing in bookstores to be considered as distinct from a previously identified strategy such as browsing a library collection. Prior studies of information seeking by historians identified nearly all of the interactive human strategies discussed in this study with the exception of attending conferences which is new to this study (Beattie, 1989; Delgadillo, 1999; Green, 2000; Graham, 2002). Finally, the hybrid strategies of using inter-library loan and purchasing a copy on-line are unprecedented in prior studies.

Much prior theory, including taxonomies of strategies, have been created by studying a wider variety of users, and they most often seek to generalize lists of strategies across a broad user population. It is perhaps counterintuitive that narrowing the group of users would result in a more extensive list of strategies. The more we focus on scholars who use non-English resources, the more we see that they embrace a wide variety of very specific information-seeking strategies of all four types discussed in this study: formal system strategies, informal resource strategies, interactive human strategies, and hybrid strategies. It is precisely because of the targeted nature of the studied population that we can define information-seeking strategies in such as granular way, as explained above.

Because the participants in this study are experienced researchers with doctoral degrees in their fields, they have considerable prior experience, and they have learned through that experience which strategies work best. In addition, they have learned what types of strategies are best suited to the information task they are pursuing, the topic of their current study, and the options they have available at the time. While it cannot be proven conclusively by a qualitative study of this nature, we could infer that the strategies mentioned most often are the most effective for the types of research pursued by these subjects. Various rationale were discussed

281

by participants as to why they choose strategies. Throughout their education and especially at the graduate school level, mentors and instructors play a major role in their choices of sources such as specific academic databases (e.g. MLA,) journal titles, or on-line repositories. Others credit more recent interactions with colleagues, librarians, and archivists as helping them to discover new sources and therefore new strategies. Some scholars, especially those using less common languages, note that their scholarly community is instrumental in providing support and direction for their research methods.

Overall, the answer to research question 1 is that scholars use a well-rounded arsenal of at least 30 different information-seeking strategies in four different categories, some of which are commonly used by users in other information-seeking contexts while others are unique to this group of scholars. Further exploration and comparison are warranted to determine to what extent the use of non-English resources influences the choices of strategies used.

7.2.2. Theoretical implications and significance related to shifts

The types of shifts were categorized based on the data gathered inductively during the study and resemble those in Xie's prior research (2000, 2007, 2008). This study also breaks new ground by mapping the shifts to the information triangle to show how shifts can and do connect specific strategies in various types of situations and for varying types of information needs. 14 triangle diagrams are provided in Chapter 5 to provide examples of planned shifts, opportunistic shifts, and alternative shifts made in routine, disruptive, and problematic situations, and to show how these become more evident and easily understood when presented visually using the triangle model.

The subjects of the study were previously unaware of the concepts of shifts between information-seeking strategies, yet the data they provided pointed out that they make frequent

shifts between strategies in order to successfully meet their own information-seeking needs. In

the initial interviews, the primary focus was on planned shifts in routine situations. Most

scholars could easily articulate "what I usually do" and describe their information-seeking

threads as a progression of strategies logically linked together. Over the years of their education

and scholarly experience, they have mastered common progressions involving planned shifts

such as:

- First, search a scholarly database to identify and access one or more relevant scholarly
  journal articles on their topic; second, use citation tracing to identify additional relevant
  sources from the footnotes and bibliography; third, use an OPAC to locate the sources
  from the footnotes and bibliographies; then fourth, use inter-library loan to obtain the
  articles that were not readily available from their own local library collection.

- First, browse a known web site such as that of a library or archives at a nearby research
  university; second, search through a finding aid on that web site to identify any relevant
  primary sources described therein; third, contact the archivist at the university to be sure
  that the documents in question really exist and are available; and fourth, travel to visit the
  university library or archive to access the documents.

At any point in their planned shifts such as those above, scholars are prepared to

encounter barriers or serendipitous discoveries that result in either problematic or disruptive

situations where they need to make either alternative or opportunistic shifts.   Examples of these

unexpected shifts were more often documented in the research diary or final interview. In

problematic situations, scholars learn from trial and error, and what they learn is either

incorporated into their repertoire of coping skills, or they are internalized to become part of their

planned research strategies for the future. In a few cases, scholars may also learn from

problematic situations that they encounter that some specific strategies are not very useful in specific types of situations. An example might be that they seldom consult their own local library collection or their own library staff because they are persuaded – whether true or not – that their local library does not have the materials or provide the services they need.

The data reported in the diaries portrays 44 different information-seeking threads undertaken to meet 44 different research needs. Of these, 25 used 3 or more strategies and were sufficiently clear and well documented to map to information triangles; 14 of those are included in Chapter 5 as examples of various types of shifts. The information-seeking threads in the diaries range in length from a single strategy to shifting between eight strategies. The most common lengths of threads used by participants were two strategies and only one shift (13), four strategies and three shifts (11), and three strategies and two shifts (10). In the 44 examples, participants documented 41 planned shifts, 19 opportunistic shifts, and 43 alternative shifts. These latter counts are in keeping with the observations that scholars who use non-English resources frequently encounter problematic situations and must make alternative shifts. If any number is surprising, it is the number of disruptive situations (19) that occur, perhaps because the scholars' expectations of finding relevant materials are often low.

The concept and discussion of shifts discussed in this dissertation extends beyond the prior work of Kuhlthau (1991) and Belkin, Marchetti & Cool (1993) who looked primarily at users' shifts from one formal IR system to another. Overall, this dissertation confirms that the categorization of planned shifts in routine situations, opportunistic shifts in disruptive situations, and alternative shifts in problematic situations (Schutz and Luckmann, 1973; Xie, 2007) is an effective way to represent information-seeking behavior. This study then builds upon Xie's model by using the information triangle as a visualization tool to document and map the shifts

according to the categories of the strategies that they connect. Certainly there is opportunity for further research into what strategies used in what contexts are most likely to cause disruptive or problematic situations and which perform as expected leading to a routine situation.

7.2.3. Theoretical implications and significance related to the role of language, culture, and geography

The most unique aspect and significant contribution of this dissertation occurs in relation to the third research question. Chapter 6 is dedicated to the aspects of language, culture, and geography that play a role in the information-seeking context, specifically in problematic situations. Unlike the discussion of shifts, no pre-existing model or theory exist to provide a classifications system for the study of language, culture, and geography aspects in information-seeking contexts. And unlike the discussion of strategies, no simple three-sided categorization tool can capture the complexity of these factors.

Previous literature related to language and information-seeking are limited primarily to implication for formal IR systems (Large et al, 1999; Moukdad & Large, 2001; Garcés, Olivas & Romero, 2004; Mustafa, 2005; Bar-Ilan & Gutman, 2005; Ahmed & Nürmberger, 2007; Vilares et al, 2008; Lazarinis et al, 2009; Dolamic & Savoy, 2009; Chang et al, 2010), including problems presented by other languages in the areas of language recognition, stemming, and term matching. While the mechanical aspects of language have been explored in the information-seeking context, the role of culture and geography have gotten very little attention, and user studies related to language and information-seeking are virtually non-existent. Meho and Tibbo (2003) conducted a qualitative study of 14 social science faculty who conduct research on stateless nations with the goal of revisiting the information-seeking process posited by Ellis (1987, 1989). The subjects of their study, like the participants in this dissertation, also discussed

trips requiring "long distances" as well as barriers related to political climate and the role of culture in gaining entrée to archives. Their study is, in many ways, similar to mine although much more limited in its scope and implications.

The gathering, coding, and analysis of data in this study show that language, culture, and geography are significant both in determining the choices of information-seeking strategies used by scholars who rely on sources in languages other than English, and that they also contribute to routine, disruptive, and problematic situations that cause researchers to make shifts between strategies. In addition to factors traditionally studied in the information-seeking context such as user characteristics (e.g. prior experience, domain knowledge) or task characteristics (e.g. task type), for non-English researchers, language can play a formidable role when using formal system strategies, while culture and geography come to bear more heavily on informal resource strategies.

7.2.4. Modeling information-seeking strategies and shifts and the role of language, culture, and geography in the information-seeking context

Based on Glaser's (1965) constant comparative method, data were analyzed in the pursuit of developing grounded theory and developing a model that illustrates the role of language, culture, and geography as they relate to strategies and shifts in the information seeking context. At the axial coding stage, categories were developed for strategies and shifts as well as for the roles of language, culture and geography. From the subsequent analysis of the categorized data during the selective coding process, a model emerged that explores how the three factors of language, culture, and geography impact the information-seeking process of scholars who use sources that are not in English, including the strategies used and the shifts made between them. At this preliminary stage in the exploration of language and information-seeking, generalizable

connections cannot be made between individual factors (e.g. diacriticals or political climate) and specific strategies (e.g. searching scholarly databases or browsing in overseas archives) without further study and analysis. Nor can this one study completely isolate language, culture, and geography from intervening conditions within the information-seeking context. Nonetheless, anecdotal evidence of connections between language, culture, and geography and choice of strategies is abundant.

Most previous grounded theory models of information-seeking take into account various individual elements, including: elements associated with the problem (e.g. subject area, specific requirements); elements associated with the system (e.g. nature of returned results) (Hert, 1997); elements associated with users (e.g. domain knowledge, attitudes, expertise with information-seeking) (Hert, 1997; Xie, 2007). While some aspects of culture could be considered part of the social-organizational context, most aspects of language, culture, and geography are clearly outliers that have not been considered in previous studies. When studying the population of scholars who use non-English resources, the roles of language, culture, and geography quickly eclipse more traditional factors like system response, subject area, or user's domain knowledge. The information-seeking model shown in Figure 7.1, below, is largely based on the paradigm model and coding theories of Flick (2009, p. 310-312). In the model of language, culture, geography, strategies, and shifts in the information-seeking context proposed below, the central phenomenon is the information-seeking process as represented by a cloud; this indicates that its boundaries are somewhat permeable, fluid, and susceptible to influence by various factors. In the center of the information-seeking context is the information triangle with the three main categories of strategies placed around it; these represent what Flick also terms "strategies" (not to be confused with information-seeking strategies) which are used to address the central

phenomenon. The yellow circle in the center is made up of bi-directional arrows that represent

the shifts made among information-seeking strategies; these are also strategies that address the

central phenomenon. The double-ended shape of each arrow indicates that the shift can occur in

any direction across the triangle. The causal factors of interest in the model are represented as

rectangles around the edge of the diagram; these include language, culture, and geography,

shown in blue with red arrows. The arrows represent the influence that the causal factors have

on the central phenomenon. The more traditional elements of task characteristics and user

characteristics, shown in gray with blue arrows, are part of the context, and more broadly the

interviewing conditions that influence the strategies in the central phenomenon.

Figure 7.1 – Language, culture, geography, strategies, and shifts in the information-seeking context: a model
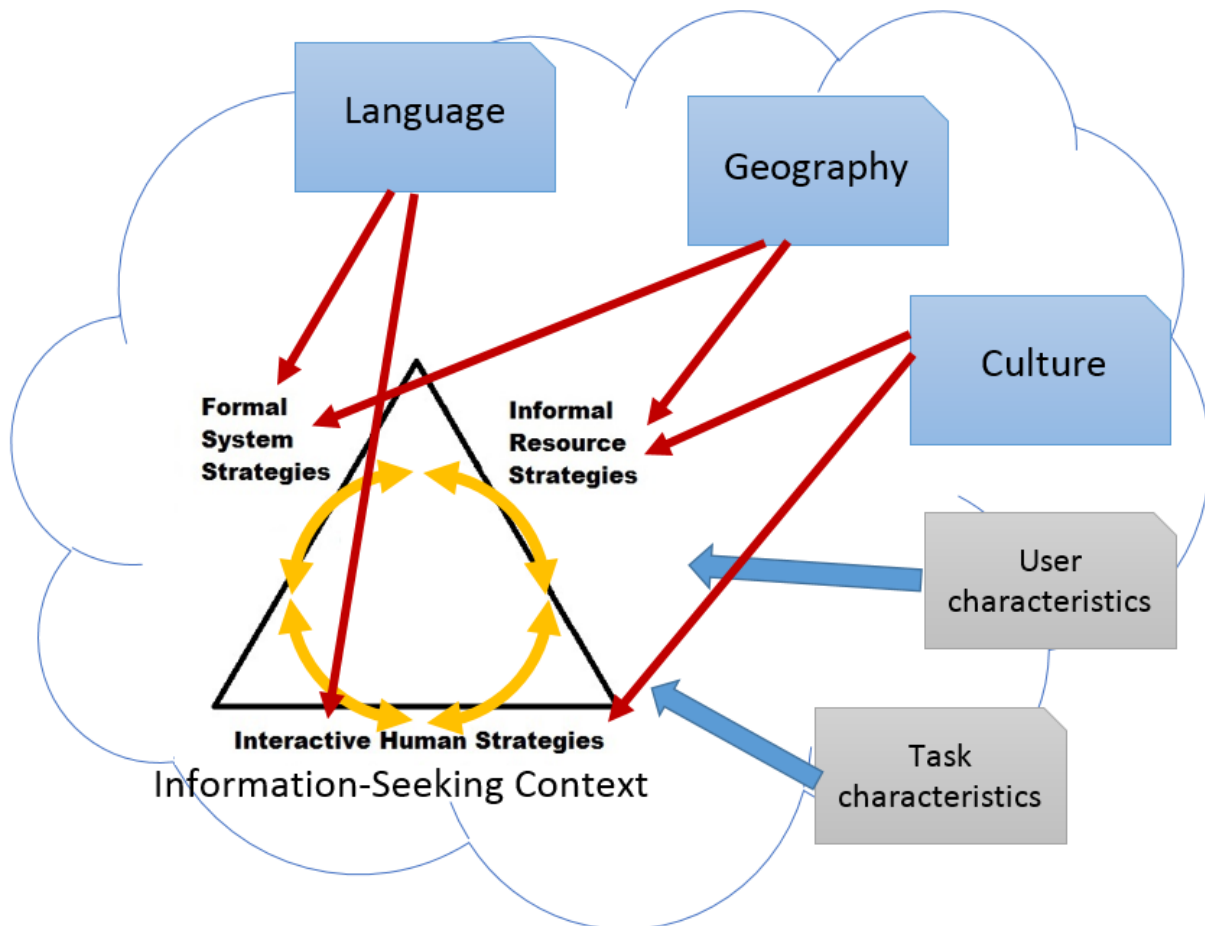
At this early, exploratory stage of research on this topic, it is too early to prove actual causality, and because the study is based on qualitative methods, further investigation with confirmatory statistics needs to be made to prove the level of correlation between the specific causal factors, the strategies, and the central phenomenon.

The purpose of the model is to illustrate the developed theory and portray how the various elements impact the information-seeking context and what aspects of the information-seeking process they most directly impact. The major emphasis is the impact on choices of information-seeking strategies, including the shifts made between strategies.

As data were analyzed from this study, it became unexpectedly evident that there are more observable links between some specific factors and specific categories of strategies. The blue arrows that emanate from the gray boxes representing user and task characteristics indicate a general and untargeted influence on the information-seeking context. For example, a user's expertise in information-seeking will certainly have an impact on their choices of strategy but not more strongly for one type of strategy than another. Similarly, a task characteristic such as the task type of a known fact search may well point the scholar to a formal system strategy such as searching the web with a search engine, but the task type is not likely to impact the outcome of the strategy itself more than it would impact the outcome of another strategy.

The red arrows that emanate from the factors of language, culture, and geography represent a role and influence that are more targeted toward a specific category of strategies. Some examples of how these causal factors influence the central phenomenon of the information-seeking process include the following:

- The language factor of diacriticals has a strong influence on the usage and success of formal resource strategies, especially searching scholarly databases, OPACs and the

web. The example given by subject 25 in section 6.2.1.3 indicated that for searches to work properly, some tools required the diacriticals to be removed, some required them to be correctly represented, and some were more forgiving. The factor of diacriticals was not shown to have any documented effect on either informal resource strategies or interactive human strategies.

- The culture factors related to gaining entrée have a strong influence on the usage and success of informal resource strategies, especially visiting or browsing in archives outside of the U.S., as well as visiting or browsing libraries, while it has no documented effect on formal system strategies and may less frequently influence interactive human strategies such as asking a librarian or archivist.

- The geography factor of the location and distribution of archives has a strong influence on the usage and success of informal resource strategies, especially visiting or browsing in archives outside of the U.S., while it has no documented effect on formal system strategies or interactive human strategies.

This model is in its nascent stage in terms of development since this line of inquiry is so new in the study of information-seeking. This is the first study that has explored language, culture, and geography in the information-seeking context, so it can be documented, for example, that when issues with proper names arise, they create a problematic situation, and an example of an alternative shift (or more likely several alternative shifts) that one scholar made in response to that particular problematic situation can also be documented. However, causality cannot yet be conclusively determined between specific language, culture or geography factors and specific shifts between specific strategies. There are too many intervening factors (the scholar's prior experience, resources available, domain knowledge, system knowledge, etc. plus additional

language, culture, and geography factors) to determine which caused the shift to the subsequent strategy. Additional study is needed to isolate factors and refine the model and thereby document more targeted relationships between specific aspects of these factors and specific shifts, strategies or strategy types.

7.3. Practical implications

Several observations can be made from this study that would inform development of systems, resources, and services provided by publishers, system developers, libraries, archives, and other information providers. One of the final questions asked of the participants was "How could either systems, libraries, or any other resources help you better?" While it was evident that scholars do not often think in these terms, they had comments and suggestions to make. Some of the implications discussed herein are based on their suggestions, and some of the more technical suggestions are conclusions drawn from connecting their observations with literature in the field.

7.3.1. Design implications

As shown above in the model of language, culture, and geography in the information-seeking context, aspects of language have the strongest impact on formal system strategies. Many of these are related to the written representation of the language, especially for languages using non-Latin alphabets. Based on prior studies of IR algorithms and stemming, such as those of Large, Tedd & Hartley (1999) or Lazarinis et al. (2009), information science professionals and system designers are well aware of the aspects of language that cause problems for search algorithms and stemming routines. For morphologically complex languages, the challenges are significant, and in order for a system to accommodate multiple languages, it would need to be able to determine the language(s) of both the documents and the search terms and then apply the

291

appropriate stemmers and search algorithms for those languages. Documents that contain

multiple languages would exacerbate this situation even further.

Xie's model of shifts (2007) includes a fourth type of shifts: assisted shifts made in

problematic situations. Assisted shifts differ from alternative shifts in that, in assisted shifts, the

IR system in use or an outside force may suggest or initiate the shift. Throughout my two pilot

studies and this study, scholars consistently do not credit systems or related resources in

suggesting or helping them to make shifts during problematic situations. One recent exception,

though still not common, is the implementation, within some library OPAC systems, of a direct

link to request an item through inter-library loan. By using the results of this study, system

developers could take into account the types of strategies that scholars use and build in more

direct links to facilitate shifts between strategies. Some examples would include clickable

footnotes and in-line citations embedded in electronic documents in scholarly databases; this

would facilitate the shift from the formal system strategy of searching scholarly databases to the

informal resource strategy of citation tracing. Likewise a link from an article to the broader set

of articles from that same title would facilitate a shift to the informal resource strategy of a

journal run. In a world of systems that runs on hypertext, the possibilities for linking to authors

web sites or email addresses, linking to citing sources (chaining forward,) etc., are endless.

Some of the issues related to formal system strategies are not caused by the systems or

search algorithms themselves but rather by the way the data are represented within these

systems. If documents and artifacts are indexed and catalogued by machine, these algorithms

must be robust and savvy enough to deal with varying alphabets, non-conforming characters,

language reform and evolution over time, regional variations, transliteration errors, and much

more. If the documents are indexed and catalogued by an information professional such as a

librarian, that person must be aware of and adhere to any standards for the aforementioned issues. But before all that can happen, standards must exist and be agreed upon beyond national or even regional boundaries. If multiple standards exist, such as mentioned by subject S20 in section 6.2.4, systems could make use of relevant authorities tables in order to make searching more effective.

7.3.2. Library service implications

Of the practical implications directly discussed by participants in this study, most were related to library services provided at their own institutions. Many paid compliments to their local librarians, speaking positively about the services they provided. While they knew that there was no reasonable expectation that librarians in the U.S. would have other language skills, subjects did indicate that language expertise from their local library staff would be beneficial to their research.

A second type of observation was the understanding that their local library – primarily at smaller colleges and universities – rarely had funding for resources that would advance their specific research. Cost factors were cited by a number of scholars when talking about on-line scholarly databases that would provide relevant secondary sources. They also understand that, most often, the primary sources they need are not available for loan from the libraries that own them, even when those libraries and archives are in the U.S.

A primary audience for this dissertation is the information profession; librarians and archivists have much to gain from recognizing the needs and habits of this group of scholars. With a clearer understanding of the strategies most often used by non-English researchers, services can be devised and implemented to provide improved support and instruction to scholars regarding best practices in non-English information-seeking as well as information-seeking in

general. Librarians can learn domain-specific skills and concepts from the scholars, and scholars can benefit from the information-seeking expertise of librarians and archivists who are more familiar with this special information-seeking context.

Globally, scholars would benefit greatly from progress in digitization and indexing of foreign collections, but they understand that many countries – especially developing countries – do not have funding for such projects. Indeed, in some cases, funding is inadequate to organize and describe their collections and then provide any type of finding aid. On their wish list would be national union catalogs that could be accessible worldwide to at least identify sources and their location. A worldwide inter-library loan service would make that circle complete.

In terms of physical access, scholars in the study would advise libraries worldwide to create a vibrant authorship and publishing culture and provide a process of gaining entrée that is predictable and free from unnecessary bureaucracy and delay, all while respecting local norms and the larger cultural context.

7.3.3. Collegial implications

In addition to the usual implications discussed in research papers, I believe that the results of this dissertation offer great promise for positive change if they are shared collegially in the academic community of non-English researchers. As stated earlier, scholars are one of the two major audiences for this research. By learning from the information behaviors of their peers, humanities scholars using non-English sources can gain insight into the aspects of language, culture, and geography that contribute to the information-seeking context, and they can learn about best practices and effective strategies used by colleagues in situations much like their own. Opening up the conversation about the unique needs and challenges of conducting research using sources in languages other than English would provide scholars with an opportunity to share

294

their hard-won expertise and learn from each other. In the course of the study, it became evident that scholars would have insight that would be helpful to others. Bringing these scholars together collegially would provide a great benefit to the academic community.

7.4. Methodological implications

In general, the qualitative methods chosen and used for this dissertation were well suited to the topic and effective in answering the research questions. The recruiting methods used with purposive sampling were effective though time consuming. The semi-structured interviews gathered excellent data and were a joy to conduct. The success of the research diary is perhaps worthy of note in that its format worked well in gathering the needed information and facilitated data coding in initial and axial coding. This study therefore confirms the efficacy of a research diary as a method for gathering qualitative data about information-seeking behaviors. The questionnaire helped to gather specific information on resources used that would have been difficult to gather verbally. The various methods complemented each other: the bulk of the data was gathered through the interviews, especially regarding usual information-seeking processes in planned situations; the diaries helped to provide information about specific strategies and shifts that were more specific; the questionnaires then filled in the missing specifics. The combination of the three methods allows for detailed documentation and exploration of the information-seeking strategies of these scholars.

The most noteworthy observations made from a methodology perspective occurred at the data coding and analysis stage. At the axial and selective coding stages, the gathered and categorized data were used to create the information triangle and to then use that framework to document and visualize the information-seeking strategies and shifts. One of the most important contributions of this dissertation is the creation of the information triangle as a conceptual model

and way of documenting and visualizing strategies and shifts used by researchers in user studies. While it was developed specifically for this dissertation, it would provide a framework for qualitative research in any study of information-seeking behaviors. Throughout the course of the study, subjects reacted positively to the concept of the three major categories of information-seeking strategies represented by the information triangle. Their comments point to the triangle as also providing a useful way for them to think about their own research habits and possibly identify additional avenues for research and unexplored strategies that they may have previously overlooked.

7.5. Limitations

The results of the study were confined by several different types of limitations, including research design, data collection, and data analysis. In terms of the subjects studied, some limitations were purposeful and others not. As explained in section 4.3, purposive sampling was used, and most of the participants were from smaller private colleges. Of the larger or public universities represented, nearly all were from the Midwest region. Due to previous experience in recruiting, participants were not solicited from either Ivy League institutions or major research universities nationwide since scholars from those types of institutions are rarely willing to participate in such studies. In order to have more generalizable results, humanities scholars from a wider variety of institution types and geographical locations in the U.S. should be studied as well. In terms of languages or disciplines represented within the humanities, it is unlikely that expanding the scope of subjects would add any truly new observations, but frequencies and examples would likely be affected if, for example, more participants were to be selected from disciplines such as religion or even outside the humanities in political science.

At the data collection stage, only data that was directly relevant to the research questions was collected. Moreover, information about strategies used was collected using open ended questions which may not lead to an exhaustive or complete representation of all strategies used since unprompted self-reporting likely favors recent or common experiences. Future research might include a checklist of possible strategies along with some scale for participants to indicate if they used a particular strategy often, rarely or never. Additionally, there are a number of demographic factors that could be gathered in order to either study or isolate them; this might include number of years of study, type of degree, and so on.

Documenting the exact steps in a scholar's information-seeking process is a challenging pursuit. Think-aloud protocols or direct observation could be used in some subsequent studies in place of or in addition to the research diaries, especially where the focus is the use of formal system strategies. However, many scholars from this group conduct their research over hours, weeks and even months, working during extended periods in relative silence in faraway places. Additionally, as the data gathered in this study showed, informal resource and interactive human strategies are very common, and these types of behaviors are more difficult to document, even with think-aloud or direct observation methods.

At the data analysis stage, many more types of analyses could be undertaken on the gathered data. For example, more in-depth analysis could be performed on the strategies themselves, especially in regard to their dimensions of method and entity. Much could be gained by characterizing and understanding what specific methods and entities are used to categorize strategies and to distinguish one strategy from another.

Specific language, culture, and geography factors could be coded from the diaries and interview narratives as they occur in specific situations, and then those factors that cause

problematic situations could be linked with the strategy in use at the time the barrier occurred. By using more extensive data gathering and more specific data analysis, more targeted connections could be made between factors, strategies and shifts, including causality. Additional data analysis might allow for participant statements of cause and effect to be gleaned from the collected interviews; in this type of data analysis, care would need to be taken to account for self-reporting inaccuracies and bias.

Subsequent studies should use quantitative and mixed methods in order to isolate causal factors from intervening factors, creating a more controlled environment, and pursuing data analysis of statistical significance. Using quantitative analysis would allow for the further development and refinement of the model of language, culture, and geography in the information-seeking context as developed in this dissertation. With additional data points analyzed quantitatively, the influence of specific factors on specific information-seeking strategies could be assessed more closely, allowing the arrows in the model to be weighted according to that influence.

Finally, in a qualitative study such as this, it is difficult to isolate, prevent or control for the intervening factor of the researcher's influence in the study. As explained in Section 4.2, above, qualitative researchers must build the confidence of and rapport with the subjects in their study (Glesne, 1992). In that process, the risk is present and unavoidable that the researcher will, in some way, influence the behaviors and perceptions of her subjects. This is certainly the case in this study because explaining the structure of the information triangle becomes necessary in order for the participants to fill out the diary form. In turn, that knowledge may influence the use and reporting of information-seeking strategies. It is most appropriate not to attempt to prevent

this but rather engage in a process of self-reflexivity and identify and comment on such influences.

7.6. Chapter summary

This chapter presents the theoretical, practical, and methodological implications of this dissertation including presenting a model for the roles of language, culture, and geography in the information-seeking context. Connections are indicated in the model between specific categories of factors (e.g. language factors) and the types of strategies (e.g. formal system strategies) on which they have the greatest impact. Unique aspects of this study, including the innovative aspects of the topic and the conceptual model of the information triangle, are discussed. Specifically, the suitability of the information triangle for further study is explained. Limitations of the study that impact generalizability, along with some suggestions for future adjustments, are also explored.

Chapter 8: Conclusion

8.1. Overview

This dissertation explores the strategies used by scholars in the humanities that use sources in languages other than English; the shifts made between strategies; and the role of language, culture, and geography in the information-seeking context. A new conceptual model for visualization of the information-seeking process – the information triangle – is presented in Chapter 4 and provides a significant framework for the analysis of strategies and shifts used by the subjects of this study.

8.2. Key findings

The study answers each of the three research questions using qualitative methods to gather data from 40 subjects representing at least 30 different languages in 8 different disciplines within the humanities.

Q1: What information-seeking strategies are used by scholars conducting research in languages other than English?

A list of 30 distinct information-seeking strategies was built inductively based on the data provided by the subjects through semi-structured interviews and research diaries. These were categorized based on the dimensions of method and entity into four types: formal system strategies, informal resource strategies, interactive human strategies, and hybrid strategies. Seven different formal system strategies were mentioned 92 times; 12 distinct informal resource strategies were mentioned 196 times; nine different interactive human strategies were mentioned 88 times; and two hybrid strategies were mentioned 37 times. The frequencies of individual strategies, along with definitions and the list of which subjects used them are provided in Table 5.2.

Q2: What shifts do scholars make among strategies in routine, disruptive, and/or problematic situations?

Shifts are organized and discussed by type and the situation in which they are used: planned shifts in routine situations, opportunistic shifts in disruptive situations, and alternative shifts in problematic situations. A discussion of the shifts made by scholars in various types of situations is presented in Section 5.3. Data from the research diaries were coded into information threads, 25 of which were mapped to information triangle diagrams, and 14 of which were subsequently chosen as examples in this dissertation. The triangle diagrams and accompanying narrative provide valuable insight into the usual, serendipitous and challenging aspects of the information-seeking processes undertaken by researchers who use non-English sources. A noteworthy observation is that no two triangles are identical in their progression of strategies used and the intervening shifts.

Q3: In what ways do language, culture, and geography play a role in the information-seeking context, in particular, how do they create barriers that cause problematic situations?

This is certainly the most fascinating, complex, and understudied of the research questions. Participants in this study provide hundreds of data points primarily from the interviews but also through diaries; these data are related to aspects of language, culture, and geography. Six factors related to language were discussed in Section 6.2 based on themes cited by participants in the study. These show significant influence on the choice of strategies chosen by scholars who use non-English sources, and they contribute to problematic situations that cause alternative shifts, especially when formal system strategies are being used. Six themes related to culture are discussed in Section 6.3. These have a more mixed impact on the information-seeking process, sometimes contributing to routine situations or even disruptive

301

situations that lead to fortuitous discoveries, but more often contributing to problematic situations. Factors related to culture are more likely to impact the information-seeking context when informal resource strategies are being used, especially visiting or browsing archives outside of the U.S. Section 6.4 discusses four themes related to geography, most of which also have the strongest impact on informal resource strategies, especially visiting or browsing archives outside of the U.S. Throughout these findings, we see that a few strategies could be considered more resistant to negative influence by language, culture, or geography. These more "language-proof" strategies include citation tracing, using private collections, and journal run.

8.3. Opportunities for future research

Since this dissertation breaks new ground in exploring the population of humanities scholars who use sources in languages other than English, it opens up many avenues for future research. Some specific areas of study that would benefit from further investigation include:

- Expansion of the research to a larger group of subjects, pursuing confirmation and triangulation using quantitative and mixed methods

- Targeted study to isolate factors of language, culture, and geography from more traditional factors such as user characteristics, task characteristics, etc.

- Further development and refinement of the research model to investigate connections between specific aspects of language, culture or geography and specific strategies

- Exploration of the dimensions of strategies and how they help determine which strategies are distinct from others and how they can be categorized

- Focused study on the information-seeking strategies used in the pursuit of primary sources

- An analysis of "identifying" versus "obtaining" sources by scholars who use non-English sources

- Further investigation of which strategies could be considered most "language-proof"

- Investigation of narrower segments of non-English researchers as defined by specific discipline, specific language group, older vs. younger scholars, etc.

The hope and ambition of this dissertation is to open the discussion of this nascent and fascinating field of investigation where the study of information-seeking and the study of language intersect.

References

Ager, Simon. *Omniglot - writing systems and languages of the world*. Accessed September 21, 2013, from www.omniglot.com.

Ahmed, F. & Nürnberger, A. (2007). N-Grams Conflation Approach for Arabic Text. In F. Lazarinis, J. V. Ferro, & J. Tait (Eds.), Proceedings of the Workshop on Improving Web Retrieval for Non-English Queries (iNEWS), SIGIR 2007, Amsterdam, The Netherlands, 23–27 July, 2007.

Altmann, G., & Fan, F. (2008). *Analyses of script: Properties of characters and writing systems.* Berlin: Mouton de Gruyter.

Anderson, D. L. (1998). Academic Historians, Electronic Information Access Technologies, and the World Wide Web: A longitudinal Study of Factors Affecting Use and Barriers to That Use. *Journal of the Association for History and Computing, 1*(1).

Ardila, R. (1999). Scientific publishing in Latin America: the case of journals in the behavioural and social sciences. *in: Revistas científicas en América Latina – Scientific Journals in Latin America; edited by* A.M. CETTO and O. ALONSO. Paris: International Council of Scientific Unions; Mexico: UNAM, CONACYT, and Fondo de Cultura Económica. p. 209.

Bailey, C. A. (2007). *A guide to qualitative field research*. Thousand Oaks, Calif: Pine Forge Press.

Bar-Ilan, J., & Gutman, T. (2005). How do search engines respond to some non-English queries? *Journal of Information Science, 31*(1), 13-28.

Bates, M. J. (1979a) Information search tactics. *Journal of the American Society for Information Science, 30*(4), 205-214.

Bates, M. J. (1979b) Idea tactics. *Journal of the American Society for Information Science, 30*(5), 280-289.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*, 407-424.

Beattie, D. L. (1989-90). An Archival User Study: Researchers in the Field of Women's History. *Archivaria*, *29*, 33-50.

Belkin, N. J. (1980). Anomalous States of Knowledge as a Basis for Information Retrieval. *The Canadian Journal of Information Science, 5*, 133-143.

Belkin, N. J. (1982). Models of Dialogue for Information Retrieval. In: Friberg, I. (ed.): *Proceedings of the4th International Research Forum in Information Science.* Boras, Sweden: Skrifter fran Högskolan i Boras, pp. 15-36.

Belkin, N. J. (1993). Interaction with texts: Information retrieval as information-seeking behavior. In G. Knorz, J. Krause, & C. Womser-Hacker (Eds), *Information Retrieval '93: Von der Modellierung zur Andwendung* (pp. 55-66). Kostanz: Univertaetsverlag Konstanz.

Belkin, N. J. (1996). Intelligent information retrieval: whose intelligence? In J. Krause, M. Herfurth, and J. Marx (Eds). *Harausforderungen an die Informationswirtschaft. Informationsverdichtung, Informationsbewertung und Datenvisualisierung. Proceedings of the 5th International Symposium for Information Science* (ISI '96) (pp. 25-31). Konstanz: Universitätsverlag Konstanz.

Belkin, N. J., Cool, C., Koennman, J., Ng, K.B., & Park, S. (1996) Using relevance feedback and ranking in interactive searching. In E. M. Voorhees & D. M. Harman (Eds.), TREC-4, Proceedings of the Fourth Text Retrieval Conference (pp. 181-209). Washington, DC: NIST.

Belkin, N. J., Cool, C., Stein, A., & Thiel, U. (1995). Cases, scripts, and information-seeking strategies: On the design of interactive information retrieval systems. *Expert Systems with Applications, 9*(3), 379-395.

Belkin, N. J., Marchetti, P. G., & Cool, C. (1993). BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing and Management, 29*(3), 325-344.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982a). ASK for information retrieval: Part I. Background and theory. *Journal of Documentation, 38*(2), 61-71.

Bennett- Kapusniak, Renee K. (2015). "Baby Boomers and Technology: Factors and Challenges in Utilizing Mobile Devices" *Theses and Dissertations.* Paper 1040. http://dc.uwm.edu/etd/1040

Berg, B. L. (2001). *Qualitative research methods for the social sciences*. Needham Heights, MA: Allyn & Bacon.

Blanco, R., & Lioma, C. (2009). Mixed monolingual homepage finding in 34 languages: the role of language script and search domain. *Information Retrieval, 12*(3), 324-351.

Bogdan, R., & Taylor, S. J. (1975). *Introduction to qualitative research methods: A phenomenological approach to the social sciences*. New York: Wiley.

Brahme, M., & Walters, L. (2010). While Technology Poses as the Great Equalizer, Distance Still Rules the Experience. *Journal of Library Administration, 50*(5/6), 484-514.

Brunn, S. D., & Dodge, M. (2001). Mapping the "worlds" of the world wide web: (re)structuring global commerce through hyperlinks. *The American Behavioral Scientist, 44*(10), 1717-1739.

Busha, C. H. and Harter, S. P. (1980). *Research methods in librarianship,* New York: Academic Press.

Cairncross, F. (1997). *The death of distance: How the communications revolution will change our lives*. Cambridge, MA: Harvard Business School Press.

Callahan, E. (2005). Interface design and culture. *Annual Review of Information Science and Technology. 39*, 257-310.

Carroll, T. (2005). Beyond keigo: smooth communication and the expression of respect in Japanese as a Foreign Language. *International Journal of the Sociology of Language, 2005*, 233-247.

Case, D. (1991). Conceptual Organization and Retrieval of Text by Historians: The Role of Memory and Metaphor. *Journal of the American Society for Information Science. 42*(9), 657-668.

Catalano, A. (2013). Patterns of graduate students' information-seeking behavior: a meta-synthesis of the literature. *Journal of Documentation, 69*(2), 243-274.

Chang, S-J. & Rice, R. E. (1993). Browsing: A Multidimensional Framework. In M.E. Williams (Ed.), *Annual Review of Information Science and Technology.* Medford, NJ: American Society for Information Science by Learned Information, Inc.

Chang, Y. K., Morales-Arroyo, M. A., & Spink, Amanda H. (2010). Multimedia search capabilities of Chinese language search engines. *Information Processing and Management, 46*, 308-319.

Chau, M., Qin, J., Zhou, Y., Tseng, C., & Chen, H. (2008). SpidersRUs: Creating specialized search engines in multiple languages. *Decision Support Systems, 45*(3), 621-640.

Chen, H., & Dhar, V. (1991). Cognitive process as a basis for intelligent retrieval systems design. *Information Process & Management, 27*(5), 405–432.

Chi, M. T. H. (1997). Quantifying Qualitative Analyses of Verbal data: A Practical Guide. *The Journal of the Learning Sciences, 6*(3), 271-315.

Chickering, F. W., & Yang, S. Q. (2014). Evaluation and Comparison of Discovery Tools: An Update. *Information Technology and Libraries, 33*(2), 5-30.

Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research, 6*(3), 319-340.

Choo, C., Detlor, B., & Turnbull, D. (1999). Information-seeking on the Web – An integrated model of browsing and searching. *First Monday*, Retrieved September 12, 2013, from http://www.firstmonday.org/issues/issue5_2/choo/

Chowdhury, G. G. (2004). *Introduction to modern information retrieval*. London: Facet.

Clavel, P. (2006). How localization challenges international portals: character sets and
        international access. *IFLA Conference Proceedings*, 1-6.

Cook, T. (1997). What is Past is Prologue: A History of Archival Ideas since 1898, and the
        Future Paradigm Shift. *Archivaria, 43*.

Creswell, J. and Miller, D. (2000). Determining Validity in Qualitative Inquiry. *Theory into
        Practice, 39*(3). 124-130.

Crystal, D. (2001). *English as a Global Language*, Cambridge University Press: Cambridge,
        U.K.

Crystal, D. (2006). *Language and the Internet.* Cambridge University Press: New York.

Davis, M. W., & Ogden, W. C. (2000). Towards universal text retrieval: Tipster text retrieval
        research at New Mexico State University. *Information Retrieval, 3*(4), 339-356.

Delgadillo, R., & Lynch, B. P. (1999) Future historians: Their quest for information. *College &
        Research Libraries, 60*(3), 245.

Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods.*
        Englewood Cliffs, N.J: Prentice Hall.

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. Thousand Oaks: Sage
        Publications.

Dervin, B. (1983). An overview of sense-making research: Concepts, methods and results. Paper
        presented at the annual meeting of the International Communication Association, Dallas,
        TX, May. [On-line]. Available: http://communication.sbs.ohio-state.edu/sense-
        making/art/artdervin83.html

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative
        methodology. In J. D. Glazier & R. R. Powell (Eds), *Qualitative research in information
        management* (pp. 61-84). Englewood, CO: Libraries Unlimited.

Dervin, B. & Dewdney, P. (1986). Neutral Questioning: A New Approach to the Reference
        Interview. *Research Quarterly, 25*(4), 506-513.

Dervin, B., Foreman-Wernet, L., & Lauterbach, E. (2003). *Sense-making methodology reader:
        Selected writings of Brenda Dervin*. Cresskill, N.J: Hampton Press.

Dolamic, L., & Savoy, J. (2009). Indexing and stemming approaches for the Czech language.
        Information Processing and Management, 45(6), 714-720.

Donmoyer, R. (2001). Paradigm Talk Reconsidered. In V. Richardson (Ed.), *Handbook of
        Research on Teaching*. Washington, DC: American Educational Research Association.

Duff, W., Craig, B., & Cherry, J. (2004). Historians' Use of Archival Sources: Promises and Pitfalls of the Digital Age. *The Public Historian, 26*(2), 7-22.

Duff, W., & Johnson, C. (2002). Accidentally found on purpose: Information-seeking behavior of historians in archives. *The Library Quarterly, 72*(4), 472-496.

Efthimiadis, E., Malevris, N., Kousaridas, A., Lepeniotou, A., & Loutas, N. (2007). How do search engines handle Greek queries? In F. Lazarinis, J. V. Ferro, & J. Tait (Eds.), Proceedings of the Workshop on Improving Web Retrieval for Non-English Queries (iNEWS), SIGIR 2007, Amsterdam, The Netherlands, 23–27 July, 2007.

Egbokhare, F. O. (2003). *Breaking Barriers: ICT, Language Policy and Development*. Ibadan, Nigeria: The Postgraduate School, University of Ibadan.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation, 45*(3), 171-212.

Ellis, D. (1992). The Physical and Cognitive Paradigms in Information Retrieval Research. Journal of Documentation, *48*(1), 45-64.

Ellis, D. (1996). *Progress Problems in Information Retrieval.* London: Library Association Publishing.

Ellis, D. (1993). Modeling the information-seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly, 63*(4), 356-369.

Esterberg, K. G. (2002). *Qualitative methods in social research.* Boston: McGraw-Hill.

Fattah, M. A., Ren, F., & Kuroiwa, S. (2006). Stemming to improve translation lexicon creation form bitexts. *Information Processing and Management, 42*(4), 1003-1016.

Fay, B. (1994). General Laws and Explaining Human Behavior. In Martin, M., & McIntyre, L. C., *Readings in the philosophy of social science*. 91-110. Cambridge, Mass: MIT Press.

Fernandez-Quijada, D. (2011). Appraising Internationality in Spanish Communication Journals. *Journal of Scholarly Publishing, 43*(1), 90-109.

Fidel, R. (1985). Moves in online searching. *Online Review, 9*(1), 61-74.

Fidel, R. (1992). The Case Study Method: A Case Study. In J. D. Glazier, & R. R. Powell, (Eds.) *Qualitative Research in Information Management*. Englewood, CO: Libraries Unlimited, Inc.

Fidel, R. & Pejtersen, A. M. (2005). *Cognitive Work Analysis*. In K. E. Fisher, S. Erdelez, & L. McKechnie. Theories of information behavior. Medford, N.J: Published for the American Society for Information Science and Technology by Information Today.

Fidel, R., Pejtersen, A.M., Cleal, B., & Bruce, H. (2004). A multidimensional approach to the study of human-information interaction: A case study of collaborative information retrieval. *Journal of the American Society for Information Science, 55*(11), 939-953.

Flick, U. (2009). *An introduction to qualitative research*. Los Angeles: Sage.

Foster, A. & Ford, N. (2003). Serendipity and information-seeking: an empirical study. *Journal of Documentation, 59*(3), 321-340.

Fox, B., & Fox, C. J. (2002). Efficient stemmer generation. *Information Processing and Management, 38*(4), 547-558.

Friedman, T.L. (2005). *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.

Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report, 20*(9), 1408-1416.

Garcés, P. J., Olivas, J. A., & Romero, F. P. (2006). Concept-matching IR systems versus word-matching information retrieval systems: Considering fuzzy interrelations for indexing Web pages. *Journal of the American Society for Information Science & Technology, 57*(4), 564-576.

Ghemawat, P. (2007). Why the World Isn't Flat. *Foreign Policy, 159*, 54-60.

Glaser, B. G. (1965). The Constant Comparative Method of Qualitative Analysis. *Social Problems, 12*(4). 436-445.

Glaser, B. G. & Strauss, A.L. (1967). *The discovery of Grounded Theory.* Chicago: Aldine Publishing, Company.

Glaser, B. G., & Strauss, A. L. (1970). In N. K. Denzin (Ed.), *Sociological methods: A sourcebook*. Chicago: Aldine Pub. Co.

Glazier, J. D. & Powell, R. R. (1992). *Qualitative Research in Information Management*. Englewood, CO: Libraries Unlimited, Inc.

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. White Plains, N.Y: Longman.

Global Reach, *Global Internet Statistics (by Language)* (2004). Available at: www.global-reach.biz/globstats (accessed 25 July 2004).

Gorman, G. E., Clayton, P., Shep, S. J., & Clayton, A. (2005). *Qualitative research for the information professional: A practical handbook.* London: Facet.

Gould, S. (2000). Interlending and document supply: A review of recent literature XXXVIII. *Interlending and Document Supply, 38*(3), 143-150.

Graham, S. R. (2000.). *Historians and Electronic Resources: A Citation Analysis*. Ann Arbor, MI: Scholarly Publishing Office, University of Michigan Library.

Graham, S. R. (2002). Historians and Electronic Resources: Patterns and Use. *Journal of the Association for History and Computing*, 5(2).

Green, R. (2000). Locating Sources in Humanities Scholarship: The Efficacy of Following Bibliographic References. *Library Quarterly, 70*(2), 201.

Grover, R., & Glazier, J. (1985). Implications for Application of Qualitative Methods to Library and Information Science Research. *Library and Information Science Research, 7*(3), 247-60.

Guba, E. G. (1978). *Toward a Methodology of Naturalistic Inquiry in Educational Evaluation.* CSE Monograph Series in Evaluation No. 8. Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles.

Guba, E. G., & Lincoln, Y. S. (1988). Do Inquiry Paradigms Imply Inquiry Methodologies? In D. Fetterman (Ed.). *Qualitative Approaches to Evaluation in Education: The Silent Scientific Revolution.* New York: Praeger.

Ha, W. (2014). A Proposal to Study of Cross Language Information Retrieval (CLIR) System Users' Information-seeking Behavior. In M. N. Al-Suqri, L.L. Lillard & N. E. Al-Saleem (Eds.), *Information Access and Library User Needs in Developing Countries*. Hershey, PA: IGI Global.

Hammarstrom, H. (2007). A fine-Grained Model for Language Identification. In F. Lazarinis, J. V. Ferro, & J. Tait (Eds.), Proceedings of the Workshop on Improving Web Retrieval for Non-English Queries (iNEWS), SIGIR 2007, Amsterdam, The Netherlands, 23–27 July, 2007.

Hansen, P., & Karlgren, J. (2005). Effects of foreign language and task scenario on relevance assessment. *Journal of Documentation, 61*(5), 623-639.

Harter, S. P. (1986). *Online information retrieval: Concepts, principles, and techniques.* Orlando, FL: Academic Press.

Harter, S. P., & Rogers-Peters, A. (1985). Heuristics for online information retrieval: A typology and preliminary listing. *Online Review, 9*(5), 407-424.

Hawk, W. B., & Wang, P. (1999). Users' interaction with the World Wide Web: Problems and problem solving. In M. M. K. Hlava & L. Woods (Eds.), *Proceedings of the 62nd ASIS Annual Meeting,* (Vol. 36, pp. 256-270). Medford, NJ: Information Today.

Hawkins, D. T., & Wagers, R. (1982). Online bibliographic search strategy development. *Online, 6*(3), 12-19.

Hert, C. A. (1997). *Understanding information retrieval interaction: Theoretical and practical implications.* Greenwich, CO: Ablex Pub.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations.* Thousand Oaks, CA: Sage Publications.

Hover, P. L., & Lu, J. (2009). "Sentences like these:" Multicultural information dynamics and international diversity of thought. *International Information and Library Review, 41*(3), 196-218.

Huang, C.-Y., & Chang, S.-S. (2009). Commonality of web site visiting among countries. *Journal of the American Society for Information Science and Technology, 60*(6), 1168-1177.

Ingwersen, P. (1992). *Information Retrieval Interaction.* London: Taylor Graham.

Ingwersen, P. (1996). Cognitive Perspectives of IR Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation, 51*(1), 3–50.

Ingwersen, P. & Jarvelin, K. (2005). *The Turn: Integration of Information-seeking and Retrieval in Context.* Heidelberg: Springer.

Internet World Stats, *Internet World Users by Language* (2010). Available at: http://www.internetworldstats.com/stats7.htm (accessed 26 September 2013).

Janelle, D. (1991). Global interdependence and its consequences. In S. D. Brunn & T. R. Leinbach, *Collapsing space and time: Geographic aspects of communications and information.* London: Harper Collins Academic.

Johnson, C. A., & Duff, W. M. (2005). Chatting up the Archivist: Social Capital and the Archival Researcher. *The American Archivist, 68*(1), 113-129.

Johnson, I. M. (2006). Breaking out of the linguistic ghetto? The role of new media in facilitating international research. *IFLA Conference Proceedings*, 1-14.

Karnes, D., Lajba, F., & Shorb, S. (2007) Technology for providing materials in 94 world languages: a partnership for user-driven learning, and improved awareness of the people and events that shape our world. *IFLA Conference Proceedings*, 1-6.

Kellsey, C. (2003). Crisis in Foreign Language Expertise in Research Libraries: How do we fill this gap? *C&RL News 64*(6), 391-397.

Kellsey, C., & Knievel, J. (2004). Global English in the Humanities? A Longitudinal Citation Study of Foreign-Language Use by Humanities Scholars. *College and Research Libraries, 65*, 194-204.

Krefting, L. (1991). Rigor in qualitative research: the assessment of trustworthiness. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association, 45*(3) 214-22.

Kuhlthau, C. C. (1991). Inside the search process: Information-seeking from the user's perspective. *Journal of the American Society for Information Science, 42*(5), 361-371.

Large, A., Tedd, L., & Hartley, R. (1999). *Information-seeking in the Online Age: Principles and Practice*. London: Bowker-Saur. Print.

Lazarinis, F., Vilares, J., Tait, J., & Efthimiadis, E. N. (2009). Current research issues and trends in non-English Web searching. *Information Retrieval, 12*(3), 230-250.

LeCompte, M. D. (2000). Analyzing Qualitative Data. *Theory into Practice. 39*(3), 146-154.

Liebscher, P. (1998). Quantity with quality? Teaching quantitative and qualitative methods in an LIS masters program. *Library Trends, 46*(4), 668.

Liebscher, P., & Marchionini, G. (1988). Browse and Analytical Search Strategies in a Full-Text CD-ROM Encyclopedia. *School Library Media Quarterly, 16*(4), 223-33.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications.

Loria, A., & Arroyo, P. (2005). Language and country preponderance trends in MEDLINE and its causes. *Journal of the Medical Library Association, 93*(3), 381-5.

Losee, R. M., & Worley, K. A. (1993). *Research and evaluation for information professionals.* London: Academic Press Limited.

Lu, W. H., Lin, R. S., Chan, Y. C., & Chen, K. H. (2006). Overcoming terminology barrier using Web resources for cross-language medical information retrieval. *AMIA Annual Symposium Proceedings.* 519-523.

Madkour, M.A. (1975). Linguistic impediments on information transfer for and from the Arab states: towards an automated solution. in: Proceedings of the symposium on scientific and engineering secondary information transfer for the developing countries: sponsored by ICSU Abstracting Board and the World Federation of Engineering Organisations Committee on Engineering Information in association with UNISIST; held at the Bibliothèque Royale Albert Ier Brussels, 23-24 June 1975. Paris: International Council of Scientific Unions Abstracting Board, pp. 47-59.

Manross, M., & Sabbar, C. (2010), *Libraries in France* (Unpublished paper). UW-Milwaukee, Milwaukee, Wisconsin, USA. https://dspace.carthage.edu/handle/123456789/1825

Marchionini, G. (1995). *Information-seeking in electronic environments.* Cambridge, U.K.: Cambridge University Press.

Marchionini, G. (2006). Exploratory Search: From Finding to Understanding. *Communications of the ACM, 49*(4), 41-46.

Marchionini, G. & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 21*(1), 70-80.

Markey, K., & Cochrane, P. (1981). *Online training and practice manual for ERIC database searchers* (2nd ed.). Syracuse, NY: ERIC Clearinghouse on Information Resources.

Matthiessen, C. M. I. M., Teruya, K., & Lam, M. (2010). *Key terms in systemic functional linguistics*. London: Continuum.

McGarry, K. J. (1993). *The changing context of information: An introductory analysis*. London: Library Association Pub.

McGlamery, S., & Ulrich, P. (2007). Serving Patrons in Their Language, Not Just Ours. *IFLA Conference Proceedings*, 1-12.

Meadow, C.T., Boyce, R.R., & Kraft, D.H. (1999). *Text information retrieval systems* (2nd ed.). San Diego, CA: Academic Press.

Meadow, C. T., & Cochrane, P. (1981). *Basics of online searching.* New York: Wiley.

Meho, L. I., & Tibbo, H. R. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology, 54*(6), 570.

Mellon, C. A. (1986). Library Anxiety: A Grounded Theory and Its Development. *College and Research Libraries, 47*(2), 160-65.

Meltzer, B. N., Petras, J. W, and Reynolds, L. T. (1975). *Symbolic Interactionism: Genesis, Varieties and Criticism.* London: Routledge & Kegan Paul.

Meneghini, R., & Packer, A. L. (2007). Is there science beyond English? Initiatives to increase the quality and visibility of non-English publications might help to break down language barriers in scientific communication. *EMBO Reports, 8*(2), 112-116.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis.* Thousand Oaks, CA: Sage.

Moukdad, H., & Large, A. (2001). Information Retrieval from Full-Text Arabic Databases: Can Search Engines Designed for English Do the Job? *Libri, 51*(2), 63-74.

Mustafa, S. H. (2005). Character contiguity in N-gram-based word matching: the case for Arabic text searching. *Information Processing and Management, 41*(4), 819-827.

Nederhof, A. J. (2011). A bibliometric study of productivity and impact of modern language and literature research. *Research Evaluation, 20*(2), 117-129.

Neuman, W. L. (1991). *Social research methods: Qualitative and quantitative approaches.* Boston: Allyn & Bacon.

Nobelprize.org. Nobel Media AB 2013. (21 Sep 2013). *Facts on the Nobel Prize in Literature*. Retrieved from http://www.nobelprize.org/nobel_prizes/facts/literature/

Notess, G. R. (2008). Multilingual searching: Search engine language tools. *Online, 32*(3), 40-42.

Olszewski, B., Macey, D. & Lindstrom, L. (2006). The Practical Work of <Coding>: An Ethnomethodological Inquiry. *Human Studies, 29*(3), 363-380.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage Publications.

Pejtersen, A. M., & Fidel, R. (1998). *A Framework for Work-Centered Evaluation and Design: A Case Study of IR on the Web*. Report for MIRA, Grenoble, France, March, 1998.

Pennanen, M., & Vakkari, P. (2003). Students' conceptual structure, search process, and outcome while preparing a research proposal: A longitudinal case study. *Journal of the American Society for Information Science and Technology, 54*(8), 759.

Pharo, N. (2004). A new model of information behavior based on the search situation transition schema. *Information Research, 10*(1). Retrieved September 12, 2013, from http://InformationR.net/ir/10-1/paper203.html.

Pierce, J. E. (2006). Things I Wish I Knew about Documents, and Things I'm Glad I Know: Dysfunction Junction: The Intersection of Historians & Government Documents. *DttP: A Quarterly Journal of Government Information Practice & Perspective, 34*(3), 35-36.

Rasmussen, J., Pejtersen, A. M., & Goldstein, L. P. (1994). *Connitive systems engineering.* New York: Wiley.

Riley, C. (1992). The foreign language barrier in information transfer at the University of Tasmania. *Aslib Proceedings, 44*(10) October, 331-40.

Rimmer, J., Warwick, C., Blandford, A., Gow, J., & Buchanan, G. (2008). An examination of the physical and the digital qualities of humanities research. *Information Processing and Management, 44*(3), 1374-1392.

Ruthven, I. (2008). Interactive information retrieval. *Annual Review of Information Science and Technology*, *42*, 43-91.

Sabbar, C. (2010), *Information Retrieval Behaviors of Historians* (Unpublished paper). UW-Milwaukee, Milwaukee, Wisconsin, USA. https://dspace.carthage.edu/handle/123456789/1787

Sabbar, C. (2011), *What Difference Does Language Make? A comparison of search results based on the language of the user query and the browser* (Unpublished paper). UW-Milwaukee, Milwaukee, Wisconsin, USA. https://dspace.carthage.edu/handle/123456789/1823

Sabbar, C. (2013), *How Scholars Identify and Obtain Primary Sources from Archives: A Qualitative Study* (Unpublished paper). UW-Milwaukee, Milwaukee, Wisconsin, USA. https://dspace.carthage.edu/handle/123456789/1786

Sabbar, C. & Xie, I. (2016), Language in the information-seeking context. *Journal of Documentation, 72*(1), 103 – 126. http://dx.doi.org/10.1108/JD-07-2014-0094

Sandler, M. (1992). Qualitative Research Methods in Library Decision-Making. In J. D. Glazier & R. R. Powell. *Qualitative research in information management*. Englewood, CO: Libraries Unlimited.

Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. In *Proceedings of the 60th Annual Meeting of the American Society for Information Science, 34*, 313-327.

Scheurich, J. (1996). The masks of Validity: A Deconstructive Investigation. *Qualitative Studies in Education, 9*(1). 49-60.

Schutz, A., & Luckmann, T. (1973). *The structure of the life world.* Evanston, IL: Northwestern University Press.

Schwartz, C. (1998). Web Search Engines. *Journal of the American Society for Information Science, 49*(11), 973-982.

Seale, C. (1999). *The Quality of Qualitative Research.* London: Sage.

Seale, C. (2007). Quality in Qualitative Research. In C. Seale (Ed.) *Qualitative research practice*. London: Sage.

Sexton, M. (2008). Preferably in English: Surfing the Pacific in Search of Law in Translation. *Syracuse Journal of International Law and Commerce, 35*(2), 275.

Shiri, A. A., & Revie, C. (2003). The effects of topic complexity and familiarity on cognitive and physical moves in a thesaurus-enhanced search environment. *Journal of Information Science, 29*(6), 517-526.

Shoki, G. E., & Oyelude, A. A. (2006). Cultural and linguistic barriers to information retrieval and dissemination. *IFLA Conference Proceedings*, 1-15.

Shute, S. J., & Smith, P. J. (1993). Knowledge-based search tactics. *Information Processing and Management, 29*(1), 29-45.

Spink, A. (1997). Study of Interactive Feedback during Mediated Information Retrieval. *Journal of the American Society for Information Science. 48*(5), 282-394.

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage Publications.

Taylor, R. (1968). Question-negotiation and information-seeking. *College and Research Libraries, 29*(3), 178-194.

Vakkari, P. (2001). A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation, 57*(1), 44-60.

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes of search terms and tactics while writing a research proposal:A longitudinal case study. *Information Processing and Management, 39*(3), 445-463.

Valge, J. & Kibahl, B. (2007). Restriction on Access to Archives and Records in Europe: A History and the Current Situation. *Journal of the Society of Archivists, 28*(2), 193-214.

Vilares, J., Alonso, M. A., & Vilares, M. (2008). Extraction of complex index terms in non-English IR: A shallow parsing based approach. *Information Processing & Management, 44*(4), 1517-1537.

von Ungern-Sternberg, S. (2003). Cultural and linguistic differences in information retrieval: the impact of indexing (English). *Tidskrift for Dokumentation, 58*(2), 55-60.

Wang, J., & Oard, D. W. (2012). Matching meaning for cross-language information retrieval. *Information Processing and Management, 48*(4), 631-653.

Wersig, G. (1979). The problematic situation as basic concept of information science in the framework of the social sciences. In *Theoretical problems for Informatics: New trends in informatics and its terminology* (pp. 48-57). Moscow: International Federation for Documentation.

Wersig, G. (1992). Information science and theory: A Weaver Bird's Perspective. In: B. Cronin & P. Vakkari (Eds.), *Conceptions of Library and Information Science. Proc. of the First CoLIS Conf.*, Tampere, August, 1991, London: Taylor Graham.

Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation, 37*(1), 3-15.

Wilson, T. D. (1999). Models in information behavior research. *Journal of Documentation, 55*(3), 249-270.

Wilson, T. D. (2000). Human information behavior. *Information Science Research, 3*(2), 49-54.

Wolcott, H. F. (1982). Differing styles of on-site research, or, "If it isn't ethnography, what is it?" *The Review Journal of Philosophy and Social Science, 7*(1 & 2), 154-169.

Xie, H. (2000). Shifts of Interactive Intentions and Information-Seeking Strategies in Interactive Information Retrieval. *Journal of the American Society for Information Science, 51*(9), 841-857.

Xie, H. (2002). Patterns between interactive intentions and information-seeking strategies. *Information Processing & Management, 38*(1), 55-77.

Xie, H. (2006). Understanding human-work domain interaction: Implications for the design of a corporate digital library. *Journal of the American Society for Information Science and Technology. 57*(1), 128-143.

Xie, H. (2007). Shifts in information-seeking strategies in information retrieval in the digital age. A planned-situational model. *Information Research*, *12*(4), 1-14.

Xie, I. (2008). *Interactive information retrieval in digital environments.* Hershey, PA: IGI Global Inc.

Xie, I. (2009). Dimensions of tasks: influences on information-seeking and retrieving process. *Journal of Documentation, 65*(3), 339-366.

Xie, I. (2010). Information searching and search models. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of Library and Information Sciences* (3rd ed., pp. 2592-2604). New York, NY: Taylor and Francis.

Xie, I. & Joo, S. (2011). Factors affecting the selection of search tactics; Tasks, knowledge, process, and systems. *Information Processing and Management, 48*(2), 254-270.

Yin, R. K. (2009). *Case study research: Design and methods.* Los Angeles, CA: Sage Publications.

Yu, K.-A. (2011). Culture-specific concepts of politeness: indirectness and politeness in English, Hebrew and Korean requests. *Intercultural Pragmatics, 8*(3), 385-409.

Yuan, X., & Belkin, N. J. (2010a). Evaluating an integrated system supporting multiple information-seeking strategies. *Journal of the American Society for Information Science & Technology, 61*(10), 1987-2010.

Yuan, X., & Belkin, N. J. (2010b). Investigating information retrieval support techniques for different information-seeking strategies. *Journal of the American Society for Information Science & Technology, 61*(8), 1543-1563.

Appendix A: Recruiting message text

The following message was sent to most potential participants in my study. The language was modified somewhat in cases of participants who already knew me or at the point when I began asking for participants who spoke specific languages.

Hello. This is an invitation to humanities scholars to participate in my doctoral dissertation study about scholars who use resources in languages other than English. This is a largely unexplored topic, and your participation could help provide valuable insight into the unique needs of scholars whose work requires sources that may pose unique challenges due to language, culture, or geography. The working title is "Information-seeking strategies of scholars using sources in languages other than English." So I would really like to hear about the research you do and learn how you find and obtain sources.

I am seeking 30 participants from various institutions, languages, and disciplines who are or will be engaged in research between June and December 2015. Each participant will be asked to participate in two interviews (the first about 45 minutes long and the second about 20 minutes) and also use a structured diary (a journal with specific questions) to log their research activities during a 2 to 3 week period between the two interviews. Each participant who completes the study will receive $20 and help to advance my work significantly. Participants will also receive an electronic copy of my dissertation and may have their names included as sources in the finished thesis or may remain unidentified. Perhaps most importantly, subjects in my two pilot studies found the topic and the conversation to be quite interesting and rewarding in and of itself. For a few, it even provided a bit of extra motivation to advance their research during the study.

Specifically, the activity I'm interested in is how you identify sources in the language of your research, how you access and acquire them, and what your best keys to success are. If you are engaged in any of these stages of your research during the remainder of 2015, then I hope you will agree to participate in my study. Your work is especially of interest to me if you use primary sources or items obtained from archives such as manuscripts, memoirs, government documents, newspapers, and so on, but any resource in various languages may qualify. If you know of colleagues at your institution or other institutions that might be qualified or interested, please do let me know as well. An RSVP with a description of your research and what language(s) you use is most helpful.

Thanks in advance for your help and participation. I would be very grateful for the help.

Thanks again,
Carol Sabbar

Appendix B: Initial interview questions

Initial interview

The topic of today's interview is about research and the strategies that you use to find information on your topic.
Can you please indicate if you give your permission to record today's interview?

1. Please tell me about the subject of your research.
2. What about this topic led you to pursue it?
3. What types of resources do you use? Do you use both *primary* sources (*archival materials*) as well as secondary resource? Tell me a little about it, including some examples if you like.
4. HOW do you begin to identify the resources you need? What are your usual strategies?
5. Once you have identified a resource, how do you go about acquiring it? Are there times when you can't obtain a resource that you know does exist?
6. Tell me about a recent time in your research when you identified and obtained a helpful source?   What did you do, use, or find?
7. Can you tell me about a recent time that you had difficulties obtaining resources? What are the barriers caused the difficulties? If you eventually overcame the barriers, how did you do so?
8. If you use primary resources, how did you first find out about the resource and/or its location?
9. What kinds of challenges do you face to find and obtain appropriate sources for your topic? How do you overcome these challenges?
10. What would you like to tell me about your research process and strategies that I haven't asked you about?

    In the next month, I would like you to please document at least TWO specific information-seeking tasks that you accomplish (or possibly that you don't). What were you looking for? How did you look for it? Did you find it? Tell me about the steps you took, barriers you encountered, etc. Use the diary form.

Appendix C: Demographic and source identification form

Scholars' strategies for finding and obtaining non-English sources

Thank you for participating in my study about how scholars identify and obtain sources that are not in English. Please fill out this questionnaire about the sources you use, including what libraries and archives. I do need to correlate your answers with your interview, so please do include your e-mail address. Some questions are optional as they may take some added time to fill out. Please just provide what information you can. Thanks so much in advance.

Please indicate your e-mail address below:

About your use of Archives and Primary Sources
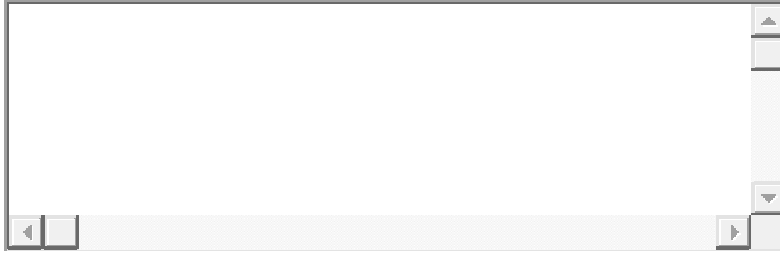
1. What types of source documents do you look for and use? Check all that apply.

☐ Newspapers

☐ Government or organizational documents/records

☐ Letters, diaries, manuscripts or memoirs

☐ Photographs, films, audio, or other multimedia

☐ Journal articles and/or books

☐ Conference proceedings (live or in print)

☐ Popular web sites (not peer-reviewed)

☐ Human subjects (interviews or observation)

☐ Other:

2. In what countries do you conduct research? (e.g. visit libraries or archives, etc.)

3. Please provide the specific names and locations of the sources you have used (e.g. Bibliotheque Nationale de France, Paris, France)
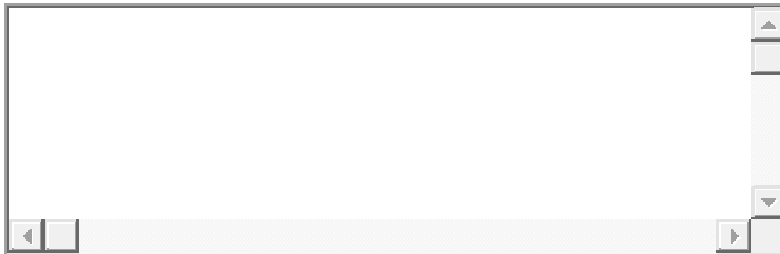
3. Please list the languages of the documents that you use. Include "English" if appropriate.

[text box]

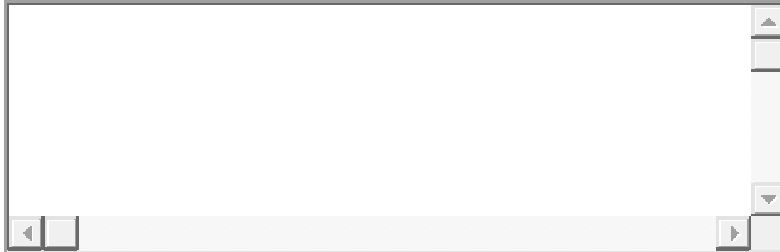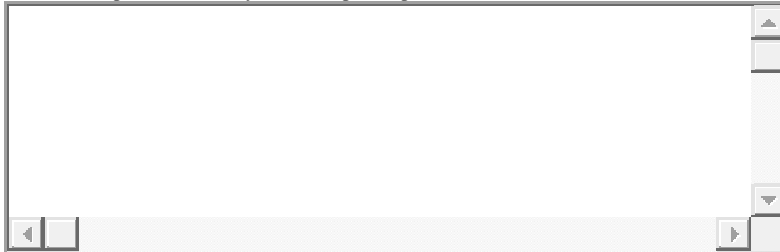4. If there are specific publications, web sites, search tools, etc. that help you get access to the materials you use or help you know what is available in what locations, please provide their titles, names, or web addresses below

[text box]

5. Cultural differences in accessing or obtaining sources. If you have noticed differences between gaining access to materials in one country or culture vs. another, please comment on that, below. This may include tips on how to get access to libraries, archives, or institutions; how to best get assistance from librarians, etc. If you have not noticed any difference or have not used such sources in other cultures, indicate "No difference" or "NA."

[text box]

6. General observations and comments. If you have any other observations or comments about how to identify or obtain sources that you think would be helpful to me in my research, please provide those below.

[text box]

Demographic Information

12. Gender - Please indicate your gender

○ Male

○ Female

13. Age - Please indicate your age

◉ 20-35

○ 36-45

○ 46-54

○ 55+

14. Native Language - please indicate your native language

○   English

○   The same as the language of my research

○   Other: [   ]

15. Degrees - Please indicate your various degrees or at least your highest degree. Please include the field and institution.

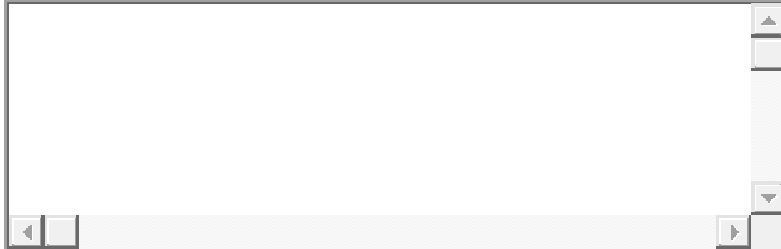**Appendix D: Diary instructions and form**

Please complete this form for at least two information-seeking tasks.

Subject ID:

Resource/archive/library/web site used: _____

_____

Describe your information-seeking task: _____

_____

Are you looking for: ❑ A specific piece of information (a name, a date, a description, etc.)
      ❑ A specific item or type of item (a specific document, video on a topic, etc.)
      ❑ Information on a topic (information from sources of any kind)

Please log the steps you take and resources you use to complete your information-seeking task. (If you were unable to complete the task and gave up, that's significant, too. If that is the case, check this box - ❑) Record in as much detail as you can what methods and sources you used in the column at the left. Use the column on the right to classify your activities as closely as you can.) Add as many more steps as you need. Each time you shift to a new strategy, enter that strategy as a new step, even if you used the same strategy earlier in the process. If you change keywords in a search or look at multiple items using the same strategy, please record that as part of the same step. Leave any unnecessary spaces blank.
Explain (if you can) why you used this resource or strategy

Step 1: Resource used: Why did you start with this strategy and/or resource?	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies ❑ Other _____

Shift A: What caused you to change strategy?	Cause of shift: (check 1) ❑ Normal change that I planned to do ❑ Good change based on a positive situation ❑ Hit an obstacle and had to change strategy
Step 2: Resource used:	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies ❑ Other _____
Shift B: What caused you to change strategy or resource?	Cause of shift: (check 1) ❑ Normal change that I planned to do ❑ Good change based on a positive situation ❑ Hit an obstacle and had to change strategy
Step 3: Resource used:	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies ❑ Other _____
Shift C: What caused you to change strategy?	Cause of shift: (check 1) ❑ Normal change that I planned to do ❑ Good change based on a positive situation ❑ Hit an obstacle and had to change strategy

Step 4: Resources used:	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies ❑ Other _____

Appendix E: Sample completed diary page

Diary instructions and form – Please complete this form for at least two information-seeking tasks. (Make two or more copies of the form and fill them out separately for each task.) You may not fill out the entire form for each task or you may need to add additional rows, depending on how many steps you take. When you have satisfied your information need, you are done and do not need to fill out any more rows.

Subject ID: S16

Resource/archive/library/web site used: __ _____

_____

Describe your information-seeking task: __Children's lit in early 20th Century Germany

_____

Are you looking for: ❑ A specific piece of information (a name, a date, a description, etc.)
  ❑ A specific item or type of item (a specific document, video on a topic, etc.)
  X Information on a topic (information from sources of any kind)

Please log the steps you take and resources you use to complete your information-seeking task. (If you were unable to complete the task and gave up, that's significant, too. If that is the case, check this box - ❑) Record in as much detail as you can what methods and sources you used in the column at the left. Use the column on the right to classify your activities as closely as you can.) Add as many more steps as you need. Each time you shift to a new strategy, enter that strategy as a new step, even if you used the same strategy earlier in the process. If you change keywords in a search or look at multiple items using the same strategy, please record that as part of the same step. Leave any unnecessary spaces blank.
Explain (if you can) why you used this resource or strategy

Step 1: Search the internet for institutes or archives in Germany that deal with kids' lit and young adult literature Resource used: Google Why did you start with this strategy and/or resource? Usual starting place	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases X Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies

Shift A: What caused you to change strategy? I can use the web sites I found to get faculty names	Cause of shift: (check 1) x Normal change that I planned to do ❑ Good change based on a positive situation ❑ Hit an obstacle and had to change strategy
Step 2: Browse the web sites I found in Step 1 Resource used: (depends on the search in step 10	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery x Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies ❑ Other _____
Shift B: What caused you to change strategy or resource?	Cause of shift: (check 1) x Normal change that I planned to do ❑ Good change based on a positive situation ❑ Hit an obstacle and had to change strategy
Step 3: Search for the faculty names in WorldCat Resource used: WorldCat	Strategy/method: (check 1) ❑ Browsing in libraries or archives ❑ Tracing footnotes, citations, or bibliography ❑ Consulting with an archivist or other person ❑ Searching scholarly databases ❑ Searching the Worldwide Web ❑ Serendipitous discovery ❑ Using known internet source or web site ❑ Using my own private collection ❑ Using known indexes and bibliographies X Other _Searching library OPACs (WorldCat)_

Shift C: What caused you to change strategy? To check to see if there are OTHER unpublished works by the people from Step 2	Cause of shift: (check 1) x Normal change that I planned to do ❏ Good change based on a positive situation ❏ Hit an obstacle and had to change strategy
Step 4: Search on Google for works by the authors Resources used: Google (The settings on my computer give me German results)	Strategy/method: (check 1) ❏ Browsing in libraries or archives ❏ Tracing footnotes, citations, or bibliography ❏ Consulting with an archivist or other person ❏ Searching scholarly databases X Searching the Worldwide Web ❏ Serendipitous discovery ❏ Using known internet source or web site ❏ Using my own private collection ❏ Using known indexes and bibliographies ❏ Other _____
Shift D: What caused you to change strategy? To check to see if there are OTHER unpublished works by the people from Step 2	Cause of shift: (check 1) x Normal change that I planned to do ❏ Good change based on a positive situation ❏ Hit an obstacle and had to change strategy
Step 5: Check forums and listservs for publications and calls for papers on this topic Resources used: Google (The settings on my computer give me German results)	Strategy/method: (check 1) ❏ Browsing in libraries or archives ❏ Tracing footnotes, citations, or bibliography ❏ Consulting with an archivist or other person ❏ Searching scholarly databases ❏ Searching the Worldwide Web ❏ Serendipitous discovery X Using known internet source or web site ❏ Using my own private collection ❏ Using known indexes and bibliographies ❏ Other _____

Appendix F: Final interview questions

Final interview

Collect their research records and offer to go through them.

This time, I want to ask you about some of the specifics of the information-seeking that you did during the study period.

1. Please walk me through your information-seeking need:

    a. What were you trying to find? (document, fact, subject info)

2. What strategy did you start with? Did you change to a different strategy at any time in your process? If so, were those planned or unplanned changes?

3. Did you encounter any barriers that led you to change strategies or methods? If so, what were they? Were any of them related to language, culture or geography? How did they change your strategies?

4. How could either systems, libraries, or any other resources help you better?

5. If you had advice for a junior colleague on how to conduct research on a topic similar to yours, what would that be?

(Highlight any barriers that are related to the use of primary sources or resources in other languages.)

Appendix G: Strategies used by discipline

S#	Discipline	FS-OP	FS-WB	FS-SD	FS-GB	FS-FA	FS-DT	FS-SM	IR-BL	IR-BA	IR-PC	IR-CT	IR-KW	IR-BU	IR-NW	IR-BK	IR-FA	IR-PI	IR-JR	IR-CF	IH-CC	IH-AC	IH-AL	IH-SM	IH-AU	IH-MN	IH-BL	IH-RA	IH-HU	HY-IL	HY-PC
S49	Archaeology	X		X					X	X	X	X		X				X			X										
S18	Art History	X	X	X					X					X	X	X			X		X				X	X	X			X	X
S19	Classics	X		X	X				X		X	X	X												X		X			X	
S20	Classics	X	X	X	X							X	X					X					X			X				X	
S23	Classics		X	X					X	X	X		X					X			X										
S17	Dance History								X	X	X	X		X		X					X	X									
S2	History		X		X			X						X	X			X				X			X						
S4	History				X				X	X				X	X								X		X					X	
S6	History	X			X				X	X	X	X	X	X	X		X	X			X		X	X						X	X
S8	History	X	X	X					X	X	X	X		X								X								X	
S10	History	X				X			X	X	X	X	X				X						X			X				X	
S15	History									X	X				X						X	X								X	
S29	History	X				X			X		X	X	X		X		X				X		X		X			X		X	X
S30	History	X		X						X	X		X	X	X						X	X				X				X	X
S31	History	X	X		X	X			X		X	X		X	X		X				X	X				X					
S32	History	X	X		X				X	X	X	X		X	X							X	X		X					X	
S33	History	X	X						X	X				X	X						X									X	
S35	History									X	X	X	X	X		X	X				X		X				X	X			
S36	History	X	X	X					X	X	X			X	X	X	X	X						X							X
S38	History	X	X	X					X	X	X	X	X	X	X	X					X	X	X		X		X			X	
S39	History	X							X	X		X		X				X				X									
S41	History		X	X	X				X	X		X	X								X	X								X	
S44	History								X	X	X	X			X	X	X				X									X	X

S#	Discipline	FS-OP	FS-WB	FS-SD	FS-GB	FS-FA	FS-DT	FS-SM	IR-BL	IR-BA	IR-PC	IR-CT	IR-KW	IR-BU	IR-NW	IR-BK	IR-FA	IR-PI	IR-JR	IR-CF	IH-CC	IH-AC	IH-AL	IH-SM	IH-AU	IH-MN	IH-BL	IH-RA	IH-HU	HY-IL	HY-PC
S45	History	X	X							X		X	X	X			X		X		X		X							X	
S46	History	X	X			X				X	X	X	X	X			X				X		X							X	
S47	History								X				X	X																X	
S48	History	X	X							X	X	X	X									X		X						X	
S43	Information Studies		X		X				X													X	X	X			X			X	
S3	Modern Languages	X	X	X					X	X	X		X									X	X							X	
S7	Modern Languages	X	X	X	X				X	X	X				X							X									
S11	Modern Languages	X	X	X			X		X	X		X		X		X						X							X	X	X
S13	Modern Languages	X			X						X		X	X		X		X				X	X							X	X
S14	Modern Languages		X			X		X							X										X						
S16	Modern Languages	X	X	X					X	X			X			X		X	X			X				X	X	X		X	X
S24	Modern Languages									X	X	X					X								X			X			
S25	Modern Languages		X	X	X		X											X	X		X				X					X	X
S34	Modern Languages	X	X	X	X				X		X	X	X		X	X					X	X	X							X	X
S40	Modern Languages	X	X	X					X	X	X	X			X	X			X		X		X							X	
S42	Modern Languages		X	X									X	X	X															X	
S37	Religion	X	X	X					X	X			X	X				X			X				X					X	

Key to abbreviations in the table above

Abbreviation	Strategy
Formal System Strategies:	
FS-OP	Search library OPACs
FS-WB	Search the web using a search engine
FS-SD	Search scholarly databases
FS-GB	Search Google Books or Google Scholar

Abbreviation	Strategy
FS-FA	Search an online archival finding aid
FS-DT	Search in a discovery tool (e.g. Summon)
FS-SM	Searching social media or streaming sites
Informal Resource Strategies	
IR-BL	Visit and/or browse libraries
IR-BA	Visit and/or browse archives outside the U.S.
IR-PC	Consult private collections
IR-CT	Citation tracing
IR-KW	Consult known web sites
IR-BU	Visit and/or browse archives in the U.S.
IR-NW	Read newspapers (on-line or in print)
IR-BK	Visit and/or browse bookstores
IR-FA	Consult archival finding aids
IR-PI	Use a print index
IR-JR	Journal run
IR-CF	Chaining forward
Interactive Human Strategies	
IH-CC	Consult a colleague or other expert
IH-AC	Attend a conference
IH-AL	Ask a librarian or archivist
IH-SM	Use social media to contact other people
IH-AU	Contact an author or family member
IH-MN	Learn from a mentor or senior scholar
IH-BL	Use forums, listservs, or blogs
IH-RA	Use a research assistant or other surrogate
IH-HU	Interview human subjects
Hybrid Strategies	
HY-IL	Use interlibrary loan (ILL)
HY-PC	Purchase a copy online

Appendix H: Participant names and institutions

The following participants gave permission for their names and/or institutions to be shared in this dissertation. They are listed in alphabetical order by last name without subject numbers to preserve confidentiality.

Dr. Ana Adams, Gustavus Adolphus College

Dr. Lennie Coleman Amores, Marian University (Indiana)

Dr. Julia Bullock, Emory University

Dr. Louisa Burnham, Middlebury College

Dr. Linda Carroll, Tulane University

Dr. Brian Caton, Luther College

Dr. Siegfried Christoph, University of Wisconsin-Parkside

Dr. Scott Denham, Davidson College

Dr. Jacques du Plessis, University of Wisconsin-Milwaukee

Dr. Virginia L. Emery, Carthage College

Dr. Robert Entenman, St. Olaf College

Dr. Christine Evans, University of Wisconsin-Milwaukee

Dr. Jennifer Foray, Purdue University

Dr. Karen Fraser, Santa Clara University

Dr. T.H.M. Gellar-Goad, Wake Forest University

Dr. Melissa Gormley, University of Wisconsin-Platteville

Dr. Leslie Hadfield, Brigham Young University

Dr. Christine Isom-Verhaaren, Brigham Young University

Dr. Maren Johnson, Luther College

Dr. Rebecca Kennedy, Denison University

Dr. Jonathan Loopstra, Capital University

Dr. Stephanie Mitchell, Carthage College

Dr. Daniel Meissner, Marquette University

Dr. Phillip Naylor, Marquette University

Dr. Lynn Palermo, Susquehanna University

Dr. Elisheva Perelman, College of St. Benedict/St. John's University

Dr. Brent Peterson, Lawrence University

Dr. Steven Rowan, University of Missouri-St. Louis

Dr. Anthony Shay, Pomona College

Dr. Erika Smith, Nichols College

Dr. Ronen Steinberg, Michigan State University

Dr. Stephen Udry, Carthage College

Dr. Robert Weinberg, Swarthmore College

Dr. Timothy Wilkerson, Wittenberg University

Dr. Nicholas Wolf, New York University

CURRICULUM VITAE
Carol Sabbar

Education

Doctor of Philosophy in Information Studies, University of Wisconsin-Milwaukee. Major: Information Retrieval. Minor: Language and International Studies. Begun 2010; projected completion in December, 2016.

Master of Arts in French, Middlebury College. Coursework at Vermont campus and year abroad at Université de Paris X-Nanterre and l'Institut d'Etudes Politiques in Paris. Certificate of Proficiency in Phonetics. Thesis Title: L'Evangelisation parmi la Jeunesse Parisienne. 1983.

Bachelor of Arts in French and Religion, Carthage College. Summa cum Laude. 1982.

Positions

Director of Library and Instructional Technology, Carthage College, 2014-Present. Oversee all aspects of Hedberg Library operations, including collection development, electronic systems and resources, reference, circulation, library liaison and instruction programs, as well as all aspects of instructional technology support for the College. Planning, promotion, and assessment of all services.

Director of Information Services, Carthage College, 2009-2014. Oversaw library, media and computer services for the College, including combined reference and technology helpdesk, library media, and technology user support. Assisted with Jenzabar ERP rollout and support, including related systems: eLearning LMS, Cognos reporting software, and Ad Astra room scheduling software.

Instructor, Carthage College, 1992-Present. Taught Introduction to Linguistics, Introduction to Organizational Technology, Understandings of Religion, Beginning and Intermediate French, Introduction to Business, Introduction to Management Information Systems, Applications in Management.

Director of Computer Services, Carthage College, 1987-2009. Oversaw the operation of IT services and systems for the College, including networking, student computing, faculty/staff computing, educational resources, and helpdesk. Established the first campus network, internet connection, computer labs, electronic classrooms, faculty laptop distribution, and all related support services.

Instructor, University of Wisconsin-Parkside, 1990-1992. Taught short courses in word processing, spreadsheets, desktop publishing.

Director of Administrative Computing, Carthage College, 1984-1987. Oversaw the development of administrative software for the College.

Instructor, Carthage College. 1983-1986. Taught Elementary French I, Elementary French II, French Conversation and Composition, 20th Century French Theatre, Introduction to Computing.

Operations Manager, Carthage College, 1983-1984. Provided support service for the Computer Center, including budget supervision, user support, and routine operations.

Grants and Awards

ASIS&T. 2016. Winner, Student Doctoral Research Videos.
https://www.asist.org/doctoral-student-research-videos/

University of Wisconsin-Milwaukee. 2015. Doctoral Research Award Grant Opportunity. $1,000 to study the information-seeking strategies of humanities scholars conducting research in languages other than English. In support of my dissertation, the grant allowed me to conduct research with 40 research subjects whose research uses non-English sources.

University of Wisconsin-Milwaukee. 2011. Doctoral Research Award Grant Opportunity. $500 to study the information-seeking strategies of scholars conducting research in languages other than English. The grant provided support for a pilot of my dissertation study.

Ameritech. 2001. $10,000 to conduct a "laptop training camp" for faculty receiving laptop computers. The week-long camp provided training on using the Blackboard system, e-mail features, and Microsoft Office software, and data organization for 15+ faculty members.

Ameritech. 2000. $10,000 to study electronic classroom design. Funds were used to travel to 18 different colleges to tour their classrooms and technology installations and to bring in a classroom design consultant.

Wisconsin Advanced Telecommunications Fund, 1999. $50,000 to create the Portable Wireless Electronic Classroom, consisting of WiFi access points and wireless capable laptop computers.

Publications

Sabbar, C., & Xie, I. (January 11, 2016). Language in the information-seeking context: A study of US scholars using non-English sources. *Journal of Documentation, 72,* 1, 103-126.

Weekly columnist for the <u>Kenosha News</u>. 2004 – Present. Column topics include basic computer troubleshooting, internet use, digital photography, viruses, newly released versions of Windows, and Macintosh/iPad.

Sabbar, Carol. "Surprises Happen… When Kids Read the Bible." *The Lutheran*. August, 2007.

Sabbar, Carol (1991). *Voyelles françaises*. Library for the Macintosh. Santa Barbara, CA: Intellimation. An interactive tutorial in French phonetics.

Conference Presentations and Participation

Mathematica Conference, (August, 2016) Organized conference including recruiting 9 presenters and 30+ participants from Wisconsin and Illinois. Responsible for all logistical arrangements. Carthage College, Kenosha, WI.

Sabbar, C. (2015) *Beyond Research Guides: Innovative Uses of LibGuides,* Presentation at WiLSWorld library conference, Madison, WI.

Kelley, T., Hoare, W., Garrigan, D., Sabbar, C. (2014) *Using Values and Culture to Calibrate Planning Trajectories*, Presentation at SCUP (Society for College and University Planning) Annual, International Conference, Pittsburgh, PA.

Forecasting Next Generation Libraries: A Course-ference. (July – August, 2013). Organized and managed a virtual on-line course and conference hybrid. It presented two keynote speakers with live and virtual audiences originating from Carthage College; four panel discussions conducted on Google Hangouts with four presenters each, including nationally known library, education, and technology experts. Attendees participated in synchronous and asynchronous activities over a 2-month period.

Sabbar, C. (2012) *What Difference does Language Make?* Poster presented at Carthage College Celebration of Scholars, Kenosha, WI.

Sabbar, C. (2012) *When Language Causes Shifts*. Poster presented at UW-Milwaukee School of Information Studies Poster Day, Milwaukee, WI.

Sabbar, C. and Zitron, E. (2012) *Understanding SOPA and PIPA Proposed Legislation*. Radio interview on WGTD Morning Show, Kenosha, Wisconsin.

Sabbar, C., Hanna, A., and Robinson, D. (2011) *It Takes a Village: Collaboration Strategies to Introduce your Portal.* Presentation at JAM Jenzabar user conference, Nashville, TN.

Kelley, T., Huenink, R., and Sabbar, C. (2011) *Where is your Campus in the ERP Development Cycle?* Presentation at JAM - Jenzabar user conference, Nashville, TN.

Kelley, T., Huenink, R., Sabbar, C., and Faraca, D. (2011) *Leveraging Organizational Technology Tools and Expertise to Maximize ROI.* Presentation at JAM - Jenzabar user conference, Nashville, TN.

Sabbar, C., Bannigan, K., Jenkins, D., Meyer, J., Tanner, J. (2010) *E-mail in the Cloud: Outsourcing your E-mail Service*. Panel discussion presentation at the WiscNet Future Technologies Conference, Madison, WI.

Huenink, R., Myers, K., Robinson, D., Sabbar, C. (2009) *Next Generation Information Services at Carthage.* Poster session at the NITLE national conference, Philadelphia, PA.

Sabbar, Carol (2009) *Using Social Media Web Sites*. Radio interview on WGTD Morning Show, Kenosha, WI.

McGrath, M., Sabbar, C., Walsh, J. (2006) *Putting Wireless Access into Residence Halls.* Presentation at EDUCAUSE Midwest Regional Conference, Chicago, IL.

Sabbar, Carol. (2005) *Old Paradigm, New Paradigm: Does Anything Ever Really Change?* Lunch speaker presentation at WFIC Best Practices Conference, Milwaukee, WI.

Sabbar, C., Walsh, J. (2005) *Towards a Better September: Controlling Residence Hall Computing.* Presentation at Educause Midwest Conference, Chicago, IL.

Sabbar, Carol. (2005) *Creating an IT Disaster Recovery Plan*. Presentation at WAICU IT Directors Meeting, Sheboygan, WI.

Heffner, D., Sabbar, C., Stack, D. (2004) *How Collaboration with Neighboring Institutions can Enhance Professional Development and Cut Costs*. Presentation at EDUCAUSE Midwest Regional Conference, Chicago, IL.

Sabbar, C., Walsh, J. (2004) *Moving your Network Operations Center in 12 Hours or Less.* Presentation at Educause Midwest Conference, Chicago, IL.

Ananou, S., Love, M., Sabbar, C., Wisniewski, B. (2004) *How Many Support Desks Does a College Need?* Presentation at EDUCAUSE Midwest Regional Conference, Chicago, IL.

Sabbar, C. (2003) *Designing Technology-Friendly Classrooms.* Presentation at the Classrooms of the Future Conference, Minneapolis, MN.

Sabbar, C. (2003) *Calculating and Controlling the Cost of a Course Management System.* Presentation at WFIC Best Practices Conference, Milwaukee, WI.

Sabbar, C. (2002) *Choosing the Right Training Models*. Presentation at the WFIC Mini-Institute, Milwaukee, WI.

Sabbar, C. (2001) *From Pushing to Running Behind: Getting Faculty to Take the Lead in Technology on Campus.* Presentation at the Governor's Wisconsin Educational Technology Conference, Green Bay, WI.

Engeldinger, E., Love, M., Quick, A., Sabbar, C. (2000) *Wireless Laptops for the Library and Portable Classroom.* Presentation at the EDUCAUSE National Conference, Nashville, TN.

Sabbar, C. (2000) *Using Wireless Technology to Create a Portable Electronic Classroom*. Presentation at MiCTA Conference, Fontana, WI.

Love, M., Sabbar, C., Unterholzner, D., (1998) *Staffing the HelpDesk in a Merged Academic Environment.* Presentation at WILSWorld, Madison, WI.

Sabbar, C. (1999) *Setting up a HelpDesk that Works*. Presentation at WFIC IT Conference, DePere, WI.

Sabbar, C. (1998) *How Not to Spend Money on Network Stuff you don't Need*. Presentation at Governor's Wisconsin Educational Technology Conference, Madison, WI.

Sabbar, C., Underholzner, D. (1997) *Cross-training Library and Technology Staff*.
    Presentation at Wisconsin Library Association Conference, Madison, WI.
Sabbar, C., Underholzner, D. (1997) *Cross-training Library and Technology Staff*.
    Presentation at the Wisconsin Educational Technology Conference, Green Bay, WI.
Sabbar, C. (1990) *Using HyperCard to Create an Educational Tutorial.* Presentation at
    Apple Educational Conference, Milwaukee, WI.

Professional Organization Service

Member, ASIS&T
Member, Wisconsin Library Association and American Library Association
Member, WAICU Library Directors Group
Member, WAICU IT Directors Group. 1996 – 2014.
At-large representative, Board of Directors, WiscNet Educational Research Cooperative.
    2006 - 2011.
Member, WiscNet Critical Data Backup and Recovery Workgroup. 2005-2008.
Member, WiscNet Security Workgroup. 2005-2007.
Chair and Co-founder, Packeteer User's Group. 2003 – 2011.
Participant, Midwest Higher Education Compact. Novell Executive briefing in Provo,
    Utah, in 2003. Learning Management System comparison project, 2005.
Planning Committee Member, WFIC/Ameritech Faculty Development Program, 1999 –
    2004.
President and member, Wisconsin Banyan User's Group. 1993 - 1997.

College Committees and Service

Learning Commons Steering Group, 2016-Present.
Member and moderator, Enterprise User's group, 2012-2015.
Chair and founder, Disaster Recovery Task Force. 2004-2011.
Chair, founder, and member. Classroom Design Team. 2000-2003, 2006-2014.
Organizer, Annual Faculty/Staff Technology Conference. 1995-Present.
Member, Academic Information Services Advisory Group. 1996-2001.
Member, Web Oversight Committee. 1994-1998.

Community Service

Instructor, Kenosha Senior Citizen's Center. 2005–2008. Conducted workshops for senior
    citizens, including topics on internet use, e-mail, and using Ebay.
Sunday School teacher, Sunday School Superintendant, Children's Music Leader,
    Christmas Program Organizer, Lay Reader. Gethsemane Lutheran Church, Racine,
    Wisconsin. 1992-2005.