BUILDING PUZZLES AND GROWING PEARLS
A QUALITATIVE EXPLORATION OF SUBJECT DETERMINATION

Daniel N. JOUDREY, djoudrey@stargate.net

Simmons College. Graduate School of Library and Information Science (Boston, USA)

Abstract

Conceptual analysis, the essential first step in the subject analysis process, is an attempt by a cataloguer or indexer to determine the subject matter, or the «aboutness», of a document. Despite centuries of organizing information in libraries, little is known about how documents are analyzed to determine their subject matter. This paper, an attempt to better understand the processes involved in determining aboutness, is based on the author's ongoing dissertation research. It examines how interested, yet untrained, participants perform the tasks of conceptual analysis when no process is suggested or imposed. The study used observation, think-aloud methods, and semi-structured interviews to examine the participants' subject determination processes. Transcripts of the analysis sessions and the interviews, as well as the participants' aboutness statements, were examined for underlying patterns. This paper focuses on one of the five activities identified in this research: how the understanding of a document's aboutness is formed. Five models for this process are presented, illustrated, and discussed.

1 BACKGROUND

Subject access to information is, has been, and, most likely, will remain one of the most difficult aspects of information organization and retrieval. For more than a century, some of the greatest minds of library and information science (LIS) have struggled with the difficulties of subject access. Shera states that systems of content-accessibility are without question the least satisfactory of bibliographic instruments (SHERA 1996). Jolley states that making subject catalogs has received little attention, and little «success has been obtained in drawing up authoritative and detailed rule. The reasons for this comparative neglect are to be found not in the lesser importance of the subject catalogue, but in the intractability of the difficulties it presents» (JOLLEY 1961). Despite the difficulties, providing subject access to information is still a core function of our bibliographic retrieval tools. For over 125 years, Cutter's «objects» of the catalog have been among the guiding principles of information organization,

reminding catalogers of their obligation to provide an assortment of means with which to access information, including subject-based access (CUTTER 1904).

In order to provide subject access, a process known as subject analysis is performed. The subject analysis process is «one of the most complex and least understood aspects of bibliographic control» (WILLIAMSON 1996). It is performed to identify the topical contents of documents so that a work can be «retrieved uniquely according to its particular aspects,» and can be «related to other materials and retrieved in conjunction with them» (HICKEY 1976). The subject analysis process begins with a procedure referred to as conceptual analysis. Conceptual analysis is the attempt to uncover the subject of a document, i.e., to determine the document's «aboutness.» That aboutness is then translated into one or more indexing languages. But how is conceptual analysis performed? Despite centuries of organizing information, it is still a mystery; little LIS research has explored the nature of the subject determination process. While researchers have been prolific in addressing translation-related subject access issues, a basic understanding of the conceptual analysis process is lacking. This paper, based on the author's dissertation research, is an exploration of a fundamental (yet little understood) aspect of subject access to information; it is an investigation into the conceptual analysis process. This paper addresses the techniques used by research participants to determine aboutness. It looks briefly at the LIS literature of aboutness, describes the study's design and methodology, discusses approaches to understanding aboutness, and illustrates them with qualitative data.

2 LITERATURE REVIEW

While the determination of aboutness is occasionally addressed in the English language LIS literature, it is rarely explored in-depth. In discussions of subject access, little information is provided about conceptual analysis. It is often assumed that catalogers and indexers can determine the aboutness of documents, and authors have focused, instead, on how to translate aboutness into controlled vocabulary or classification notation. Even in cataloging and indexing textbooks, discussions of subject analysis often assume a starting point at which the aboutness of a document is already understood, and move directly to the translation step. «The literature seems to ignore or skate over the mental processes that take place during the subject analysis phase. There is little about how people actually decide what the subject of a document is, what they actually do to achieve this» (TODD 1992).

There have been relatively few attempts in LIS literature to examine the conceptual analysis process. Notable exceptions include Wilson, Langridge, Ranganathan, and Taylor. Wilson proposes four possible approaches to aboutness, including identifying: 1) the author's intent or purpose in writing the work; 2) what in the text stands out from the background; 3) the concepts most frequently mentioned in the item; and 4)

what holds the work together and what has been left out (WILSON 1968). Langridge believes that in addition to asking «What is it about?» two other questions must be asked. He asks, «What is it?» and «What is it for?» He identifies disciplines, topics, and the forms of knowledge, thought, and writing as concepts to be addressed in the analysis as well (LANGRIDGE 1989). Ranganathan, in his tracts on classification, describes how to analyze complex subjects into facets, but does not explain how the subjects of individual works are to be determined. His process begins with analyzing the title, but he acknowledges that a title can be unhelpful or incomplete; he states that relevant concepts not represented in the title words are to be included. He does not, however, address how these concepts are to be identified (RANGANATHAN 1967). Taylor's approach to subject determination focuses on examining bibliographic features. She provides a list of items to consider, including tables of content, illustrations, introductions, etc. Her method is similar to that of other authors, but Taylor also offers a list of concepts to note during the item's examination. These include topical elements, names (personal, corporate, and geographic), chronological periods, and form/genre (TAYLOR 2004). While these approaches are useful, none are considered standard practice, and not one has been evaluated for its effectiveness in determining aboutness.

Rather than exploring the subject determination process, some authors have instead addressed the nature and meaning of «subject,» including Cutter (1904), Ranganathan (1967), Svenonius (1999), and Lancaster (2003). Others have described varying approaches to the concept without actually addressing aboutness determination methods. Fairthorne, who is credited with coining the term «aboutness», makes a distinction between «extensional aboutness» and «intensional aboutness». Extensional aboutness addresses what he considers to be the relatively stable, recognizable, inherent subject properties of the work. Intensional aboutness addresses properties associated with the users, their requests, or the reasons for which the document has been acquired; this is a meaning-based, changing, interpretive aboutness (FAIRTHORNE 1969). Beghtol (1986), Svenonius (1999), and others make similar distinctions. Some authors offer textual or linguistic approaches to understanding aboutness (HUTCHINS 1978) (BEGHTOL 1986) (SVENONIUS 1999), while others emphasize a use-based approach to the concept (MARON 1977) (SOERGEL 1985) (WELLISCH 1996). Again, few of these approaches to the concept have been examined empirically. More exploration is needed.

3 RESEARCH DESIGN AND METHODOLOGY

This paper examines the basic nature of the conceptual analysis process, how it is performed, and what processes are involved. Little is understood about conceptual analysis, so, the research is exploratory in nature. Qualitative methods are exceptionally well suited for exploration (GAY 2000). They are appropriate for exploring complexities, processes, little-

known phenomena, and circumstances in which relevant variables have yet to be identified (MARSHALL 1995). All of these apply to this inquiry, so qualitative methods were deemed the most appropriate for the study.

In order to investigate the conceptual analysis process, 12 participants were recruited for a case study in which the participants were asked to: analyze three books while being observed and tape-recorded; determine the aboutness of the items; write a statement describing the aboutness of each item; and discuss the process once the analyses were complete. During the study, the participants received no training or instructions on how to determine aboutness, and no time-restrictions were placed on the participants' analyses. A two-hour time frame was suggested, but not enforced.

3.1 Data Collection Methods

Data collection (September / October 2004) incorporated the use of think-aloud methods, observation, and semi-structured interviews. Multiple data collection techniques were used for the purpose of triangulation. As aboutness determination is primarily an internal process, the participants were asked to verbalize their thoughts during their observed analysis sessions. Once the participants had determined the aboutness of an item, they were asked to write a statement describing their understanding of that aboutness. After observing each session, the researcher then discussed the process with the participants. An interview guide was used, but the semi-structured nature of the interviews allowed the researcher to ask questions based on observations and on participants' responses to previous questions. All of the sessions were tape-recorded and the recordings were transcribed by the researcher for analysis.

3.2 Participants

All of the participants in the study are MLIS students from the School of Information Sciences at the University of Pittsburgh. Persons without LIS backgrounds were not recruited. At the time of the study, participants had not taken coursework in information organization, indexing, or cataloging, nor had they had any practical experience in these areas. The population of concern is new and future catalogers and indexers, so LIS students were ideal since they are interested in and committed to LIS values and principles, but are inexperienced in information organization. Potential participants with practical experience or coursework in information organization were eliminated from the study.

A recent study by Sauperl illustrates why experienced catalogers and indexers were not sought as participants for this research (SAUPERL 2002). Exploring similar interests in subject determination, Sauperl found that working professionals frequently relied on purely practical approaches, shortcuts, and their extensive knowledge of indexing systems. Instead of gaining insight into subject determination, her participants demonstra-

ted processes of *subject indication*. Instead of observing thoughtful examinations of the items to carefully determine aboutness, Sauperl observed catalogers often hurrying to find workable subject headings and classification numbers. Many relied heavily on catalog records for items they considered to be similar to the items in hand. So, working professionals and translation-related activities were purposely excluded from the research design to avoid shortcuts and the commingling of the conceptual analysis and translation processes that occurred in Sauperl's study. The focus of this research, instead, is on the inner subject-determination processes of untrained LIS students.

3.3 Materials

The researcher selected three books for the participants to analyze. All are non-fiction, English language monographs (non-juvenile and text-focused). Two of the books are popular literature; one item is scholarly in nature. Item 1 is from the humanities dealing with the evolution of the concept of evil in America as illustrated in American literature and history. Item 2 is from the sciences relating diet and nutrition to brain health. Item 3 is a social sciences monograph informing members of Generation X about important political issues in the 2000 U.S. presidential election. The items have disparate approaches to the organization and structure of their content. The social sciences text contains 17 chapters, each beginning with a chapter abstract and containing numerous section breaks with descriptive section headings. It has a pop culture orientation, vivid cover art, and an introduction and a conclusion. Item 2 contains 8 chapters. Each begins with a relevant quote, but one that does *not* summarize the content. The text contains descriptive section headings, charts, lists, textboxes, and tables. It does not have a conclusion; the final chapter contains recipes and meal plans. The humanities text comprises 7 chapters, which do not contain descriptive section headings using, instead, section numbering. There are few breaks in the continuous, scholarly text. The diversity of the items gave the participants a range of content with which to work despite the small number of items to analyze. A range of items was also desired to avoid giving some participants an unfair advantage based on subject expertise.

3.4 Data Analysis

Data analysis began in November 2004. The transcripts and the participants' aboutness statements were examined using content analysis techniques. QSR's N6 (NUD*IST) software for qualitative analysis was used to analyze and code the data. The initial coding scheme was developed from LIS literature and from a June 2004 pilot study. Three rounds of data analysis have been conducted thus far. The first round was a rough coding of the data to explore the transcripts and expand, develop, and refine the categories in the coding scheme using an inductive approach. Af-

ter the first round of analysis, the coding scheme underwent an extensive examination during which all text assigned to each of the categories was reviewed. This allowed for greater understanding of the nature of each category, for refining the definitions of the categories, and for developing an initial understanding of the relationships among the categories. During this examination of the coding scheme, many synonymous categories were merged, and others were separated into discrete classes. The revised scheme contained 225 active categories. As data analysis is on-going, the coding scheme continues to evolve. A second round of detailed, line-by-line coding began in December 2004 and was completed in January 2005. In this second round of coding more explicit patterns in the transcripts began to appear. A third round of data analysis occurred during March and April 2005. It focused on refining the models of aboutness determination that are discussed below.

4 Preliminary Research Findings

The 209 categories in the current coding scheme are hierarchically nested under five major areas of investigation. These areas are: 1) the physical item; 2) the intellectual content; 3) aboutness models; 4) supporting processes; and 5) human factors related to the analyses. This paper addresses some of the preliminary findings in two of these top-level categories: aboutness models and supporting processes.

4.1 Aboutness Models

In this study, transcripts were analyzed to find patterns in the participants' approaches to determining aboutness. Preliminary data support the identification of two subject determination models and several key supporting processes. Both models (Puzzle Building and Pearl Growing) and their supporting processes can be used to determine the aboutness of an entire work (macro-aboutness), a particular chapter (chapter-level aboutness), or even a single section or paragraph (micro-aboutness). Each participant in the study employed at least one of the two models, fortifying it with supporting processes, and sometimes with the other model. The participants showed considerable variation in their use of the models and the supporting processes; thus, no single approach to subject determination can sufficiently represent all of the participants. The configurations of models and supporting processes reflect the needs of the analysts and the materials being analyzed.

4.1.1 Puzzle Building

Puzzle Building is a model in which the participants gather individual details from the item's content. The process is an attempt to fit discrete pieces of the puzzle together to construct a complete picture of the item's

aboutness. Puzzle construction may occur at various points during the process. Puzzle Building can occur on both the macro- and chapter-levels. These levels may or may not appear in the same analysis. This model can best be summarized by the phrase «the whole is the sum of its parts.»

4.1.2 Pearl Growing

Pearl Growing begins with a core notion of the item's aboutness that evolves into a more complete understanding. In this process, the core idea is a grain of sand that develops into a fully grown pearl of understanding as layers of complexity are added. Pearls may require further growth if the participants' understandings are inadequate. Pearl Growing may be used to determine either chapter- or macro-level aboutness.

4.2 Supporting Processes

The supporting processes identified thus far in this study include: data input, sense-making, ignoring/skipping, note taking, categorization, summarization, extraction, assumption making, reinforcing, refuting, refining, making associations, and text reduction. These processes support both: 1) aboutness determination and 2) the formulation of verbal expressions of the understood aboutness. The following are the most commonly observed supporting processes; this list is in no way comprehensive.

4.2.1 Categorization

Categorization is the process in which ideas and objects are recognized and understood through their placement in either formal or ad hoc personal groupings. Library classifications were not often used by the participants, most likely due to their lack of familiarity with those systems at the time of the study. Categorization may be an early step in determining aboutness.

4.2.2 Summarization

Summarization is a process in which participants present the content of a document, chapter, section, or paragraph in a condensed form. Summarization occurs not only at the end of the process, but throughout the analysis.

4.2.3 Extraction

Extraction is a process of excavating words from the text. This process is primarily used with smaller units of text, such as extracting key words from section headings or tables of contents to represent a broad notion of aboutness. It is often used in connection with summarization activities.

4.2.4 Assumption Making

Assumption making is essential to aboutness determination. It may take the form of a simple guess about the content or a generalization based on textual or visual cues. Each participant in the study made numerous assumptions about the nature of the content. Assumptions can be made about all levels of aboutness (macro-, chapter-, and micro-levels).

4.2.5 The R^3 (R-Cubed) Processes

The R-Cubed processes include: refining, refuting, and reinforcing. These processes are closely tied to assumption-making and categorization. Refining is a progression from a broad, imprecise impression to a more sharpened notion. Refuting is recognizing information that proves an initial assumption to be false or erroneous, forcing the participant to rethink their assumptions. Reinforcing involves information that strengthens and supports an initial assumption.

4.2.6 Making Associations

Making associations is a process that finds connections between what is read or observed and other documents, personal experiences, or knowledge. Each of the twelve participants made associations to other texts they felt were similar or associations with their own lives.

4.2.7 Text Reduction

Text reduction begins with the individual details of a text. It is a process of distillation or condensation. It identifies the commonalities amongst the details and reduces them into inclusive, broader concepts or macro-statements. This process may be used to summarize content and communicate an understanding of aboutness.

5 Illustrations and Discussion

The following excerpts are from a transcript of one participant's analysis of Item 1 (Bagby 2000). This transcript was chosen because the participant used many of the processes in her examination of a single item. The excerpts illustrate how these approaches are used, and how these processes can interact. The excerpts contain the codes assigned to each text passage and are followed by the researcher's analysis.

5.1 Overview

This participant determines aboutness by examining both macro- and micro-levels of aboutness. She uses Puzzle Building along with various sup-

porting processes to determine the macro-aboutness of the entire item. Pearl Growing is used to determine the aboutness of each chapter. Each chapter's pearl is a piece of the macro-level puzzle being constructed.

5.2 Coding

The following codes are nested under the top-level categories «Aboutness Models» and «Supporting Processes». Nested under the primary processes of Puzzle Building and Pearl Growing are their component parts. Most supporting processes do not at this time contain additional levels of complexity.

```
1 Models
        1.1 Puzzle Building
                A The Edges/Frame
                B Puzzle Pieces
                C Completed Puzzle
        1.2 Pearl Growing
                A The Sand
                B Layers
                C The Pearl
2 Supporting Processes
        2.1 Categorization
        2.2 Summarization
        2.3 Extraction
        2.4 Assumption Making
        2.5 The R³ (R-Cubed) Processes
                A Refining
                B Refuting
                C Reinforcing
        2.6 Making Associations
        2.7 Text Reduction
```

5.3 · Example 1

5.3.1 Excerpt

So the title of the book is: *We've Got Issues: the Get Real, No B.S., Guilt-free Guide to What Really Matters* [by] Meredith Bagby. (1.1.A)

[I am] looking through the contents and chapters: seventeen of them. [The chapter titles are:] Politicians, Paramours… America's Report Card, Trick or Treat, The Tooth Fairy, Mysterious Machinations of the Reverse Robin Hood… hmm, confusing… Social Insecurity, Mediscare, Poor Fred, What Happened to George?, Ward and June, Who?, Public babysitting, The Graduate, You Want Fries with That, The Green Generation… NOT, Civil Fights, Yale or Jail, Golden Arches, Under the Covers. (1.1.A) (2.3) (2.4) (2.7)

So that's the basic rundown. [It] sounds cool. (2.4)
[I will] keep my finger in there because I know I will go back to it. [I am] jotting down 17 chapters. (1.1.A)
Introduction – No, I will make up my own mind.

5.3.2 Analysis

The conceptual analysis process generally begins with an examination of the preliminary pages in the item (cover, title page, table of contents, etc.); it is no different for this participant. She makes the unusual choice of not looking at the introduction, preferring to «make up [her] own mind» instead. The participant begins by attempting to build a frame of reference for the macro-level aboutness puzzle. In order to build the puzzle's border, extraction is used. Text reduction may play a role, but it is not explicitly mentioned. The participant collects the title and the author's name, but the initial framework is not complete until she examines the table of contents. This framework is a vague understanding of the nature of the work; it is not a complete understanding of the aboutness. The existence of the initial framework is inferred from the statement that the book «sounds cool,» i.e., the participant has made a macro-level assumption about what the work is about, and she believes that she will find the book interesting. In this example, three or four processes begin the analysis. Sharpening and categorization may also be occurring, but because the initial framework is never explicitly described and think-aloud methods cannot capture unspoken processes, it is unclear if the initial notion of aboutness is being sharpened during the examination of the table of contents or if the participant is placing this work in either formal or ad hoc categories. The participant does not exhibit the use of summarization; she never verbalizes her macro-level assumptions of the item's aboutness.

5.4 Example 2

5.4.1 Excerpt

[Chapter 1's title is] Politicians, Paramours, and Peccadilloes. «Forget drugs. This generation is «just saying no» to politics. Lies, scandal, bickering, and partisanship has [sic] turned us off. How can younger Americans take Washington seriously? And what will happen to our democracy—and our issues—if we don't?» [Opening quote]

NOTE: Political figures (1.2.A) (2.2) (2.3) (2.4)

Stereotypes: yeah, yeah, we know about that. (1.2.B)
Dirty Rotten Scoundrels… yeah, this is confirming everything I feel. (1.2.B) (2.5.C)
Forget the picture, it's not the text.

WEB OF DECEIT: It jumps out at everybody. Who's the Enemy Here? (1.2.B)

Reading this part here that jumped out at me, because I am real interested in young adults. «The good news is that despite our current lack of interest, Generation X is not totally devoid of political potential.» I can testify to that. (1.2.B) (2.6) «Only 1 percent of 17-24 years olds identified themselves as apolitical.»

No Partying Down. (1.2.B)
Impotence in Washington—Political Impotence That Is. (1.2.B)…
NOTE: Political figures—opinion—trust—the public—vote. (1.1.B) (1.2.C) (2.2) (2.4) (2.5.A)…
This is geared toward young adults.

NOTE: young adults. (2.1) (2.2) (2.4) So, the first one's voting. (1.1.B) (2.2) (2.4)

5.4.2 Analysis

This excerpt illustrates how the aboutness of a chapter can be used as a piece of the macro-aboutness of the item. The participant begins by examining the chapter title and abstract. Using extraction and chapter-level assumption making, these features provide the initial grain of sand that will eventually be grown into a pearl of understanding. Taking notes, the participant jots down «political figures.» This is a verbal description of her initial understanding of chapter. It is neither complex nor complete. Having discarded numerous concepts from the text (apathy of younger Americans, scandals, lack of credibility, etc.) she favors a broad general description, which will need to be sharpened if it is to be useful. As the participant encounters more of the chapter content, she adds layers of complexity to the sand to grow it into a pearl. As she encounters more of the content, she begins to make associations between the information and her own experiences. Some information helps to reinforce her chapter-level assumptions, while others may help her to refine or refute her assumptions.

After developing a more complete pearl, the participant describes her understanding again. The revised verbal representation is more focused, going from the broad «political figures» to the more specific «political figures – opinion – trust – the public – vote». The participant's pearl (her understanding) has expanded and her summary, in turn, has been sharpened. Her verbal description does not reflect her total comprehension of the chapter's content, but it does contain enough information to represent the chapter's puzzle piece. This piece is then added to the overall picture of aboutness to help complete the puzzle. Another puzzle piece is the category «young adults», which reflects the item's audience.

5.5 Example 3

5.5.1 Excerpt

Third chapter: Trick or Treat, the Tooth Fairy, and a Tea Party. «Did you know that more than half our budget goes toward just three programs… ? Where does the rest go? How does the government spend your tax money? Do we really have a budget surplus or might it be just be another sneaky accounting trick?» (1.2.A) (2.4)

Federal Budget (1.2.B) Fast Facts-Facts, see they're highlighted, or they're bolded. So it makes me jump towards them naturally, which is the purpose of course. [I] don't need to read them all, but… well, yes and no, I don't agree with [her] facts, some general opinions she has. I did not read them aloud because I didn't read everything in them. I just looked at 2-10 words in each Fast Fact there. (1.2.B)

The Tooth-Fairy Surplus (1.2.B) Just looking through it but not really reading anything. Democrats, Republicans, Social Security, Clinton… I wonder if this guy is a Republican. «It's good that we are paying down the debt, it is a little sketchy that we are using Social Security to do it.» (1.2.B) (2.5.C)

I wish I knew when this was written too: 2000, well, copyrighted in 2000… Talking about the debt. Tooth fairy. (1.2.B)
NOTE: National debt. (1.1.B) (1.2.C) (2.2) (2.3) (2.4) (2.5.C)

Cassandra Speaks. We know who Cassandra is. (1.2.B) After getting the gist of two and a half chapters, I think I know where he is coming from… «So when you hear a politician say "I want to use the surplus to make tiny ceramic tea party sets for all fat men over forty in Wichita, Kansas," and so on, just know you're being duped. Ask the hard questions of the candidates: What happens to your plan if there is no surplus? Where do you get the money then? If the answer is the Tooth Fairy, jump ship.» [Last paragraph in chapter] (1.2.B) (2.5.C)

Alright so, the third chapter's about «Fiscal Responsibility» (1.1.B) (1.2.C) (2.2) (2.4) (2.5.A)

5.5.2 Analysis

In this excerpt, many of the same processes that appeared in Example 2 are again present, but it also illustrates a two-stage Pearl Growing process. The title of chapter 3 is not particularly informative, thus the initial grain of sand is not meaningful. The participant conducts an extended examination before an understanding is developed and verbalized. That summarization is extracted directly from the text. Her description, «national debt», is an incomplete pearl; it reflects only a part of the chapter. Her understanding needs to be developed further, so Pearl Growing continues. Throughout the examination, the participant looks for mate-

rial to reinforce her understanding. Once the participant's understanding is more complete, she again summarizes her understanding in a verbal description. This description, «National debt—fiscal responsibility», is somewhat more focused than her earlier concept. This pearl becomes another piece to be placed into the puzzle's frame.

5.6 Example 4

5.6.1 Excerpt

> I'm going to stop doing this [extended analysis of each chapter] because I feel likes it's taking too long. I am just going to write down the chapter headings.
>
> NOTE: Green Gen, Civil Rights, Yale/Jail, McDonalds. (1.1.B) (2.2) (2.3) I don't care what «Under the Covers» is because I know what this book is about.
>
> ABOUTNESS STATEMENT: This book is about one individual's opinion on Generation X's political and societal state. It examines over a dozen issues that young adults should be conscious of and proposes some remedial actions for those young adults reading the book. What struck me as the most important issue is that young adults must be informed and vote as soon as they are able. (1.1.C) (2.2) (2.5.A) (2.7)

5.6.2 Analysis

This is an excerpt from the final step in this participant's Puzzle Building process. After examining 12 chapters and identifying 12 pearl-grown puzzle pieces, the participant feels that she has enough information to complete the puzzle and to describe her understanding of the item's aboutness. She uses extraction to identify the aboutness of the final chapters from their titles and then summarizes her understanding of the macro-level aboutness in a written statement. Her final aboutness statement is constructed primarily from the participant's notes. Each note represents a piece of the puzzle. The completed puzzle (the participant's complex understanding of the document) is not in itself communicable. It must be translated into words (bounded by the limits of language) and reduced to a macro-proposition that addresses the subject content of the entire document. The macro-proposition/aboutness statement is the minimal verbalization of the participant's fuller understanding.

6 Conclusion

In this ongoing dissertation research, two key approaches to determining aboutness and numerous supporting processes have been observed. The approaches, Puzzle Building and Pearl Growing, are construction-based

models. Both approaches were observed in the participants' analyses, but only in conjunction with one or more of the supporting processes, or with the assistance of the other model. Each participant used multiple techniques across the three items and within each item. It appears that the models and processes are interdependent, and that the primary approach cannot be fully successful without the assistance of at least one other process. The supporting processes, in addition to assisting the primary models, also appear to be helpful in articulating the participants' understanding of aboutness, i.e., text reduction, categorization, and others were used when the participant wanted to express complex concepts in smaller, manageable, communicable units.

Of the models, the most frequently observed is Pearl Growing. All of the participants used this model to determine either macro- or chapter-level aboutness. In Pearl Growing, the grains of sand may come from one or more of the other processes. The grains grow into pearls as more content is examined. While the pearl grows, the description of that pearl must be narrowed or refined in order to express the concepts concisely. The final explanation of the pearl is a more acutely focused representation of the initial concept. Pearl Growing is supported by the supporting processes, but it can also be supported by Puzzle Building. Some participants use Puzzle Building to reinforce or verify their understanding once their pearl is sufficiently complex, i.e., they collect additional discrete details from the text to validate their notions of aboutness. Puzzle Building might also be used to develop an initial grain of sand through the collection of individual details.

The data suggest that Puzzle Building is more prevalent among participants using strictly linear approaches. It begins with the construction of an overall framework for the item, i.e., building the puzzle's border from preliminary details. One or more of the other methods may help to develop the initial framework for the puzzle. With an initial framework in place, subsequent pieces identified from the details can be used to complete the overall picture of the item's aboutness. Text reduction is used at the end of the process to describe the participant's grasp of the macro-aboutness. Some participants use Pearl Growing to create individual pieces of the puzzle, which may or may not represent chapter-level aboutness. It appears that dual-level Puzzle Building can also be performed, in which Puzzle Building can be used to complete a chapter-level puzzle. This chapter puzzle, in turn, can be used as a puzzle piece in the macro-level puzzle for the entire item.

Text reduction, identified in this study as a supporting process, was described in 1986 by Beghtol as a model for aboutness determination. She states, «Our understanding of a document's aboutness results from our ability to reduce the information in the text to manageable and therefore memorable portions». The text reduction process hierarchically organizes «all the detailed textual propositions [into] the most general macro-proposition that meaningfully expresses the aboutness of the whole text» (Beghtol 1986). Beghtol may be correct that text reduction can be a model for understanding macro-aboutness, but that has not

been observed in this study. The current data suggest that text reduction is primarily used as a supporting process related to expressing an understanding of aboutness. Further investigation is warranted.

This paper describes only the beginnings of a larger research project. As it continues, other models of subject determination may develop and current models may evolve or vanish. Additional supporting processes may also be identified. What has been observed thus far has encouraged the author to continue to explore the rich ground of subject determination. It is hoped that others will embrace this area of study so that additional research regarding the possible connections between aboutness determination and cultural background, text comprehension, reading skills, cognitive psychology, personality type, and information-seeking models will be explored. These and other topics are in need of investigation from many different perspectives and approaches.

References

(BAGBY 2000) BAGBY, Meredith. *We've Got Issues: The Get-Real, No B.S., Guilt-Free Guide to What Really Matters*. New York: Publicaffairs, 2000.

(BEGHTOL 1986) BEGHTOL, Clare. «Bibliographic Classification Theory and Text Linguistics: Aboutness Analysis, Intertextuality, and the Cognitive Act of Classifying Documents». *Journal of Documentation.* v. 42, n. 2 (1986), p. 84-113.

(CUTTER 1904) CUTTER, Charles A. *Rules for a Dictionary Catalog*. 4th ed. Washington, DC: Government Printing Office, 1904.

(FAIRTHORNE 1969) FAIRTHORNE, Robert A. «Content Analysis, Specification, and Control». In: CUADRA, Carlos A. (ed.); LUKE, Ann W. (ed.). *Annual Review of Information Science and Technology,* v. 4 (Chicago, Ill.: Encyclopaedia Britannica, 1969), p. 73-109.

(GAY 2000) GAY, Lorraine; AIRASIAN, Peter. *Educational Research: Competencies for Analysis and Application*. Upper Saddle River, NJ: Merrill/Prentice-Hall, 2000.

(HICKEY 1976) HICKEY, Doralyn J. «Subject Analysis: An Interpretative Survey». *Library Trends.* v. 25, n. 1 (1976), p. 273-291.

(HUTCHINS 1978) HUTCHINS, W. J. «The Concept of «Aboutness» in Subject Indexing». *Aslib Proceedings.* v. 30, n. 5 (1978), p. 172-181.

(JOLLEY 1961) JOLLEY, Leonard. *The Principles of Cataloguing.* New York: Philosophical Library, 1961.

(LANCASTER 2003) LANCASTER, F.W. *Indexing and Abstracting in Theory and Practice.* 3rd ed. Champaign, Ill.: University of Illinois, 2003.

(LANGRIDGE 1989) LANGRIDGE, D.W. *Subject Analysis: Principles and Procedures*. London: Bowker-Saur, 1989.

(MARON 1977) MARON, M.E. «On Indexing, Retrieval and the Meaning of About». *Journal of the American Society for Information Science.* v. 28, n. 1 (1977), p. 38-43.

(MARSHALL 1995) MARSHALL, Catherine; ROSSMAN, Gretchen B. *Designing Qualitative Research*. Thousand Oaks, Ca.: Sage, 1995.

(RANGANATHAN 1967) RANGANATHAN, S.R. *Prolegomena to Library Classification*. Bombay: Asia Publishing, 1967.

(SAUPERL 2002) SAUPERL, Alenka. *Subject Determination During the Cataloging Process*. Lanham, Md.: Scarecrow Press, 2002.

(SHERA 1966) SHERA, Jesse H. *Documentation and the Organization of Knowledge.* Hamden, Conn.: Archon Books, 1966.

(SOERGEL 1985) SOERGEL, Dagobert. *Organizing Information: Principles of Data Base and Retrieval Systems*. San Diego: Academic Press, 1985.

(SVENONIUS 1999) SVENONIUS, Elaine. *The Intellectual Foundation of Information Organization*, Cambridge, Mass.: MIT Press, 1999.

(TAYLOR 2004) TAYLOR, Arlene G. *The Organization of Information*. 2nd ed. Westport, Conn.: Libraries Unlimited, 2004.

(TODD 1992) TODD Ross J. «Subject Access—What's It All About?» *Catalogue Australia.* v. 19, n. 3/4 (1993), p. 259-262.

(WELLISCH 1996) WELLISCH, Hans H. «Aboutness and Selection of Topics». *Keywords.* v. 4, n. 2 (1996), p. 7-9.

(WILLIAMSON 1996) WILLIAMSON, Nancy J. «Standards and Rules for Subject Analysis». *Cataloging & Classification Quarterly.* v. 21, n. 3/4 (1996), p. 155-176.

(WILSON 1968) WILSON, Patrick. *Two Kinds of Power: An Essay on Bibliographic Control.* Berkeley, Ca.: University of California Press, 1968.