

Lei Zhang

University of British Columbia, Vancouver, Canada

Approaching Navigational Cues of Web Documents from Relevance Theory and Genre Theory

Abstract: Based on Relevance Theory proposed by Sperber and Wilson, this research explores genres at multiple levels and its influence on expectations of relevance in interpreting the space particular details occur and in comprehending particular details, and navigation patterns directed by expectations of relevance. It proposes a conceptual model of genre-assisted navigation through web documents for particular details and describes the empirical study for verifying the model.

1 Introduction

With the increasing availability and accessibility of web documents enabled by the information technology, online reading has become a common practice for various activities. A web document loses discrete boundaries as indicated by paper documents while incorporates the properties of hypertext: the content is likely to be broken into pieces on several pages, and one document piece is likely to be linked to other pieces within or across documents. Thus navigation has been acknowledged as “the single greatest difficulty for readers of electronic text” (Dillon, 2004, p. 50), as additional work dealing with the structure which takes away effort and concentration from the content.

Navigation is defined as “moving oneself sequentially around an environment, deciding at each step where to go” (Jul & Furnas, 1997). Two major means to improve hypertext comprehension are using metaphors like cities, maps, etc. for spatial display and creating coherent structures to explicate relations (Vora & Helander, 1997). A recent trend arises to assist navigation by means of structural genre conventions. Genre refers to the relatively stable and expectable form and content for communication within a particular community (Breure, 2001). Web genres may be migrated or adapted from traditional genres like journal article, or emerged unique to the web like home page. Genre-assisted navigation is based on the idea that the perceived cues of a document enables one to be aware of its form, which in turn enables one to be aware of its type of content (Toms & Campbell, 1999). It is agreed that the identification of structural regularities can facilitate navigation through either web sites (Symonenko, 2007) or web pages (Vaughan, 1999). Certain components of journal articles have been identified functions to play such as orientation, overview, etc. (Bishop, 1999). However, all of these studies neglect to address that readers seldom employ one strategy (e.g., one type of structural knowledge only) and stick to that single strategy during navigation in online reading.

Coherent hypertext structure is built on discourse comprehension; web genre is derived from rhetorical or linguistic genre. Both are based on linguistics-related areas. However, linguistics has not affected hypertext studies as it ought to (Esperet, 1996; Tosca, 2000). As a theory of everyday speech utterances, Relevance Theory proposed by Sperber and Wilson (1995) is also

applicable to the thought processes in the general sense, and what's more the interface has been likened to a kind of communication. Relevance Theory claims that the expectations of relevance raised by the utterance guides the hearer toward the speaker's meaning (Wilson & Sperber, 2002). With the communicative media as the web and the type of input as a document genre, how is human mind geared to obtaining the most relevant information from web documents in the interaction?

In realistic and complex reading situations, readers often need to use more than one text, and need to use a particular detail of text (Rouet, 2006). It remains unclear cognitive processes and communicative behavior involved in dealing with particular details of web documents. The current limited understandings are due to a lack of robust underlying framework, which is just this research intends to tackle equipped with the explanatory and predictive power of Relevance Theory.

This paper reports part of ongoing research on exploring cognitive processes and communicative behavior in locating and consuming particular details within and across web documents from the perspective of Relevance Theory. Section two discusses Relevance Theory and related studies; section three proposes a conceptual model of genre-assisted navigation based on relevance-theoretic comprehension procedure; section four describes the upcoming empirical study.

2 Relevance Theory

2.1 Linguistic Theory of Relevance

The publication of the book *Relevance: Communication and Cognition* co-authored by Sperber and Wilson in 1986 marks the establishment of Relevance Theory. In their second edition, Sperber and Wilson (1995) differentiate between two principles of relevance (p. 260):

Cognitive Principle of Relevance: *Human cognition tends to be geared to the maximization of relevance.*

Communicative Principle of Relevance: *Every act of ostensive communication communicates a presumption of its own optimal relevance.*

Cognitive Principle of Relevance states that human cognition tends to maximize relevance in processing information - to gain the greatest cognitive effects at the least processing effort. Applied to cognitive processes in verbal communication, Communicative Principle of Relevance states that an intentional communication conveys the presumption of optimal relevance - it is at least relevant enough to be worth the addressee's attention and it is the most relevant one the addresser could have made with her abilities and preferences.

The communication Sperber and Wilson advocate is an ostensive-inferential one, in which the communicator makes manifest both his intention and that he has this intention, while the audience tries to recognize the communicator's intention, first by decoding the literal meaning, then by inferring the intended meaning. It follows Relevance-theoretic Comprehension Procedure (Wilson & Sperber, 2002, p. 613):

- a. *Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.*
- b. *Stop when your expectations of relevance are satisfied (or abandoned).*

2.2 Related Studies

As noted by Saracevic (2007), relevance in communication, that is, Relevance Theory proposed by Sperber and Wilson, has some impact on thinking about relevance in information science than thoughts on relevance from other fields. Harter (1992) made the first attempt to apply Relevance Theory to information studies by proposing psychological relevance, which is based on the essence of relevance of Relevance Theory, thus “references on the topic may be less important than relevant references not on the topic - references that allow the making of new intellectual connections or cause other cognitive change” (p. 612). A recent research by White (2007) uses cognitive effects and processing effort within Relevance Theory to indicate term frequencies and inverse document frequencies in a two dimensional pennant diagram of bibliometric retrieval.

Tosca (2000) applies cognitive effects and processing effort to guidelines of writing relevant hypertexts depending on what kind of interpretive movement to provoke: *Maximal (informational) cognitive effects at Minimum processing effort* for presenting concrete information to enable the reader know where she is and where she can go at all times, or *Maximal (lyrical) cognitive effects at Increased processing effort* for taking advantage of hypertext’s power of suggesting implicatures for the readers to explore. Laine (2003) extends Tosca’s pragmatics of links to all clickable textual labels on hyperlinks and buttons, namely, interactive texts. Laine uses explicitness within Relevance Theory to show the explicitness of interactive text can enhance the strength of interactivity in the setting of online shopping.

Unger (2002, 2006) points out the influence of genre on comprehension because genre information enters into the comprehension procedure to the fine-tuning expectations of relevance. Nevertheless, Unger’s statement is limited to linguistic discourses. Yus (2007) extends the ideas of genre and relevance to weblog template in stabilizing weblog genre, “genre identification is bound to save mental effort and direct the addressee towards particular interpretive paths and lead to specific expectations of weblog information” (p. 124).

3 A Conceptual Approach

3.1 Theoretical Framework

Relevance in information studies is inspired by its borrowed theory Relevance Theory. However, if examine Tosca’s and Laine’s work which start from Relevance Theory itself, we may find the notion of relevance in information studies does not include all aspects of Relevance Theory. This linguistic theory of relevance, particularly relevance-theoretic comprehension procedure, is just the standpoint of this research.

According to Relevance Theory, humans pay attention to information seemingly relevant, and communication claims the audience’s attention, implying the information communicated should be relevant. Applied to web environment, the web document aims to communicate meaning by the writer, and thus the way it is presented is to be inferred as relevant by the reader. In this regard Relevance Theory meets with genre theory - a web document is a genre in its own right with anticipated content and form for supporting the communicative purpose. Genre theory

approaches communication and cognition from the artifact aspect involving human communicative behavior while Relevance Theory, from the user aspect involving the cognitive processes of artifact. The two are connected by the mediated artifact, utilized by the technology.

In discussing Relevance Theory in the linguistics context, Unger (2006) argues that relevance can explain what falls within the scope of global coherence and beyond, and the role of genre can be integrated with the inferential process of comprehension. The combination of Relevance Theory and the concept of genre are expected to provide a holistic view and in-depth interpretation toward the role genre plays in assisting navigation through web documents for particular details.

3.2 Genres at Multiple Levels

The unit of genre can be a document component like scientific abstract (Cross & Oppenheim, 2005), a document like journal article (Vaughan & Dillon, 1998), a web page like web newspaper (Vaughan, 1999), or a web site like university site (Symonenko, 2007). Content structure of an individual document component has its regularities; so does a document the component belongs to or a web page the document is situated in. Take online journals for example, a subdivision of an article, such as abstract, tables, figures, etc., has distinct characteristics in its structure and content. A journal article has multiple components organized in a regular manner, i.e., which components are normally included, which component often comes before which (e.g., methodology and results appear in sequential order). A web page which contains the document has also evolved its regularities with what and where other elements are likely to appear on that page (e.g., global, local, and contextual navigations).

When a particular detail is presented on the web, it is tied with those wrapped around and infused within the content-bearing page. When dealing with a particular detail, it should be placed in a macro hierarchical structure of genres, each of which contains its navigational cues. Thus it needs to examine, for a particular document genre, distinct characteristics recognized of genre at different levels - individual document components, the document the components belong to, and the web page the document is situated in, and navigational cues conveyed at each level, to support the typified situation and activity.

We are more interested in text types in three distinct groups identified by Dillon (2004): work-related texts, 'news'-type texts, novel. These documents in print format are traditional genres frequently used, and they are prevalent on the web as adapted genres, thus providing an examination on what differences will make in exploiting a genre with the incorporation of characteristics of web media. Also we are interested in some novel genres like weblogs, which have grown as a cited source. The newly emerged genres provide an examination on what differences will make in exploiting a genre in terms of structuredness compared with adapted genres. These differences, if have, are revealed by variations in expectations of relevance and navigation patterns.

3.3 Cognitive Processes - Expectations of Relevance

Genre information comes as an input to the inferential process of comprehension and specifies that inferential process. Genre information may exist at multiple levels of a space in which a particular detail occurs. For a particular document genre, expectations of relevance can be more or less precise dependent on which level of genre information the reader applies to. For example, expectations of relevance vary with web document genres such as a journal article, and of components included in that document such as an abstract, and of the web page holding that document such as the presentation of an electronic journal system. Each may, to a certain extent, indicate its physical document space, which in turn suggests the semantic information space. Readers are likely to employ genre information from more than one level, with navigational cues at multiple levels playing complementary roles.

Of genre information at different levels, not all of them enter into readers' inferential process: which comes first into inference, what is the order of accessibility for inference, during the process which is heavily exploited, and in what condition is the closure of inference. According to Relevance Theory, from among the competing stimuli available, the one selected is more relevant than any alternative stimulus in that context. Readers consider interpretive hypotheses in order of accessibility and stop when it arrives at an interpretation that satisfies the expectations of relevance raised by the stimulus, that is, a balance between the cognitive effects provided and the processing effort demanded. It is tasks and texts that largely decide which genre information on which level readers first attend to, which they mostly rely on, and which motivates them to move a level up or down or exit. That is, it is determined by whether the particular details are used within a document component, or across components within a document, or across components of multiple documents, and which web document genres are employed.

Thus it needs to examine what different roles genres at multiple levels play in generating the expectations of relevance in interpreting the space particular details occur and in comprehending particular details within a document component, across components within a document, and across components of multiple documents. It also needs to examine further whether the involvement of task goals and individual characteristics would alter and how alter the genre-driven expectations of relevance.

3.4 Communicative Behavior - Navigation Patterns

Communication via web media, on the part of writer, is to convey intention by genre information; on the part of reader, is to recognize intention by processing genre information cognitively, represented by communicative behavior as navigation patterns. Expectations of relevance may direct navigating and reading, such as motivating readers to turn to another piece of text assumed more interesting within or across documents, or intensifying readers' identification of the details before going through a component or a whole document.

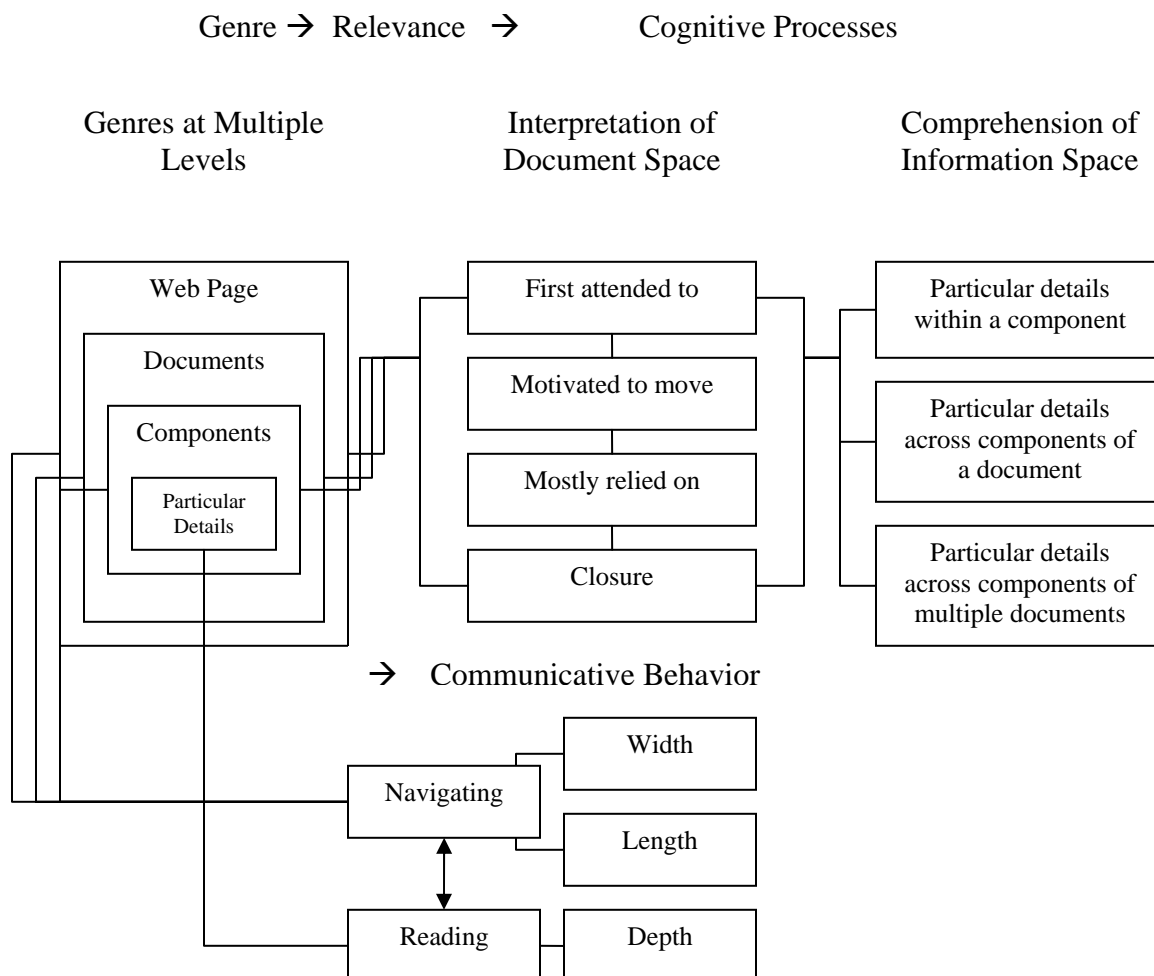
When one refers to a web document, one often needs to locate a particular detail of the document efficiently, and to consume the information related to that detail effectively. Navigating can be viewed as communicative action in responding to navigational cues conveyed at different levels: width refers to the levels of genre information are employed, counted as the frequency readers move away from the current level to another level up or down; length refers to the genre

information on a level are employed, counted as the constancy readers stay within a single level; depth refers to the long attention lingering on a block of text and can be considered as reading.

Thus it needs to examine patterns of navigating and its relation with reading in employing genre information across levels as width of navigating path, across blocks of text on a level as length of navigating path, and within a block of text as depth of navigating path, a shift to reading.

3.5 An Integrated Conceptual Model

What elaborated above can be illustrated as an integrated conceptual model:



The proposed conceptual model needs to be verified to see in what respects it needs to be adjusted in light of empirical study.

4 An Empirical Approach

4.1 Sample

The scope of this research is the situation of using web documents for academic research activity. Online journals and weblogs are chosen as corpus for they represent characteristics of two major types of web genres: maintain most characteristics of paper documents, or obtain new characteristics unique to the web. The implications and applications drawn from these two are to be generalized to other web genres. Samples are 30 journals articles by Google Scholar and 30 weblogs by Google Blog randomly selected on the topic of “information architecture”. Multiple methods for data collection are used including transaction logs, retrospective interview, comprehension writings, drawings, observation, and pre-test questionnaire. The collected data is to be analyzed by content analysis, which is most commonly used to interpret rich qualitative data, and has been widely employed in empirical analyses of web genres.

4.2 Data Collection and Analysis

Phase I:

To identify genres at multiple levels, 5 domain experts with experiences of online reading are recruited in a focus group. Tasks are to identify typified situation and activity engaged in the genre of online journals vs. that of weblogs, and identify and categorize distinct characteristics recognized of genre at different levels - individual document components, the document the components belong to, and the web page the document is situated in, from 10 randomly selected journal articles and 10 randomly selected blogs. Thus generating a list of genre information dependent on levels and the exploitation on each level.

Phase II:

To identify expectations of relevance generated by genres at different levels, 30 LIS graduate students are recruited as participants to conduct three reading tasks; they can read in any way they like within 5 minutes for task a, 5 minutes for task b, and 30 minutes for task c. Before each reading, they are given instructions on the comprehension writings and write answers within limited time afterwards:

- a. Answering true-false questions referred to information contained in a component within a journal article or a blog assigned. This is to examine readers' comprehension of particular details within a document component.
- b. Answering true-false questions referred to information contained in at least two components within a journal article or a blog assigned. This is to examine readers' comprehension of particular details across components within a document.
- c. Writing a review on journal articles or blogs assigned. This is to examine readers' comprehension of particular details across components of multiple documents.

For each of tasks a and b readers are assigned one journal article and one blog; for task c, 28 journal articles and 28 blogs, including the ones used in phase I, but excluding the ones used for tasks a and b. We use Pre-test questionnaire to understand individual characteristics (e.g., demographic information, domain and system expertise); Scores on above three comprehension tasks, and Retrospective interview to elicit readers' comprehension of particular details; Transaction logs, Retrospective interview, Drawings, and Unobtrusive observation to elicit readers' interpretation of the space particular details occur.

To identify navigation patterns directed by expectations of relevance, the sample is from RQ 2 in participants, tasks, and materials. Data from Transaction logs and Retrospective interview are used to characterize the way navigated across levels of genre information, the way navigated across blocks of text on a level, the way content read within a block of text.

We take genres at multiple levels, together with task goals and individual characteristics as independent variables, and effects and effort as dependent variables; effects can be measured by structure interpretation and content comprehension, while effort measured by width, length and depth of navigating paths. From the above experiments, by varying tasks and texts, effort and effect factors may be manipulated to produce observable changes in the inferential process toward relevance. This may result in the readers' corresponding navigation paths, their ensuing comprehension outcomes, information usage and feelings.

From the data collected, coding categories are identified in terms of genres at multiple levels; exploitation of genre information in interpreting structure and in comprehending content; navigation patterns. Then relationships are identified between:

- a. genres at multiple levels and the expectations of relevance in interpreting the space particular details occur;
- b. genres at multiple levels and the expectations of relevance in comprehending particular details;
- c. expectations of relevance and navigation patterns.

5 Conclusion

As an approach of a linguistic theory of relevance, combined with the concept of genre, this research will help better understand the cognitive processes and communicative behavior involved in interacting with web documents in location and consumption of particular details. This research intends to offer a conceptual model for genre-assisted navigation in online reading, from both document and user aspects, and from both social and cognitive perspectives.

The results from this research can be used to inform practical web document design. On the basis of readers' tendency to the pursuit of relevance, with predicting and manipulating the mental states of readers, improved design can be executed by making use of the interplay between cognitive effects and processing effort in the inferential process, with the influencing factors like the inputs from genre information at multiple levels, and task goals and individual characteristics in a certain context.

Unlike prior works, this research intends to enrich the current studies on genre-assisted navigation by identifying genres at multiple levels; to derive convincing evidence by incorporating its influence on expectations of relevance as cognitive processes in interpreting structure and in comprehending content for particular details; and to formulate patterns of navigating and its relation with reading as communicative behavior.

References

- Andersen, J. 2008. The concept of genre in information studies. *Annual Review of Information Science and Technology* 42: 339-367.
- Askehave, I., & Nielsen, A.E. 2005. Digital genres: A challenge to traditional genre theory. *Information Technology & People* 18, no. 2: 120-141.
- Bishop, A.P. 1999. Document structure and digital libraries: How researchers mobilize information in journal articles. *Information Processing and Management* 35: 255-279.
- Breure, L. 2001. Development of the genre concept.
<<http://people.cs.uu.nl/leen/GenreDev/GenreDevelopment.htm>>
- Cross, C., & Oppenheim, C. 2006. A genre analysis of scientific abstracts. *Journal of Documentation* 62, no. 4: 428-446.
- Dillon, A. 2004. *Designing usable electronic text*. 2nd ed. Boca Raton, FL: CRC Press.
- Esperet, E. 1996. Notes on hypertext, cognition, and language. In *Hypertext and cognition*, 149-155. Eds. J.F. Rouet, J.J. Levonen, A. Dillon & R.J. Spiro. Mahwah, NJ: Lawrence Erlbaum Associates.
- Harter, S.P. 1992. Psychological relevance and information science. *Journal of the American Society for Information Science* 43, no. 9: 602-615.
- Jul, S., & Furnas, G.W. 1997. Navigation in electronic worlds: A CHI 97 workshop. *SIGCHI Bulletin* 29, no. 4. <<http://www.sigchi.org/bulletin/1997.4/jul.html>>
- Laine, P. 2003. Explicitness and interactivity. *ACM International Conference Proceeding Series* 49: 421-426.
- Rouet, J.F. 2006. *The skills of document use: From text comprehension to Web-based learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Saracevic, T. 2007. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology* 58, no. 13: 1915-1933.
- Sperber, D., & Wilson, D. 1995. *Relevance: Communication and cognition*. 2nd ed. Oxford: Blackwell.
- Symonenko, S. 2007. *Websites through genre lenses: Recognizing emergent regularities in websites content structure*. Doctoral dissertation. Syracuse University.

- Toms, E.G., & Campbell, D.G. 1999. Genre as interface metaphor: Exploiting form and function in digital environments. *Proceedings of the 32nd Hawaii International Conference on System Sciences*.
- Tosca, S.P. 2000. A pragmatics of links. *Journal of Digital Information* 1, no. 6.
<<http://jodi.tamu.edu/Articles/v01/i06/Pajares/>>
- Unger, C. 2002. Cognitive-pragmatic explanations of socio-pragmatic phenomena: The case of genre. *EPICS I Symposium*.
- Unger, C. 2006. *Genre, relevance, and global coherence: The pragmatics of discourse type*. New York: Palgrave Macmillan.
- Vaughan, M.W., & Dillon, A. 1998. The role of genre in shaping our understanding of digital documents. *Proceedings of 61st Annual Meeting of the American Society for Information Science*, 559-566.
- Vaughan, M.W. 1999. *Identifying regularities in users' conceptions of information spaces: Designing for structural genre conventions and mental representations of structure for web-based newspapers*. Doctoral dissertation. Indiana University.
- Vora, P.R., & Helander, M.G. 1997. Hypertext and its implications for the Internet. In *Handbook of human-computer interaction*, 877-914. 2nd ed. Eds. M.G. Helander, T.K. Landauer & P.V. Prabhu. Amsterdam: Elsevier.
- White, H.D. 2007. Combining bibliometrics, information retrieval, and relevance theory. Part I: First examples of a synthesis. *Journal of the American Society for Information Science and Technology* 58, no. 4: 536-559.
- Wilson, D., & Sperber, D. 2002. Relevance theory. In *The handbook of pragmatics*, 607-632. Eds. L.R. Horn & G. Ward. Oxford: Blackwell.
- Yus, F. 2006. Relevance Theory. In *Encyclopedia of language and linguistics*, 512-519. 2nd ed. Ed. K. Brown. Amsterdam: Elsevier.
- Yus, F. 2007. Weblogs: Web pages in search of a genre? In *The texture of Internet: Netlinguistics in progress*, 118-142. Eds. S. Posteguillo, M.J. Esteve & M.L. Gea-Valor. Cambridge: Cambridge Scholars Publishing.