International Journal of Information Management 31 (2011) 21–29

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Making functional units functional: The role of rhetorical structure in use of scholarly journal articles☆

Lei Zhang*, Rick Kopak, Luanne Freund, Edie Rasmussen

School of Library, Archival and Information Studies, University of British Columbia, Vancouver, BC, Canada V6T 1Z1

ARTICLE INFO

Keywords:
Functional units
Information use tasks
Document components
Genre analysis
Relevance theory

ABSTRACT

Scholars are reading more journal articles than ever, so it is important that they focus on the relevant text within the articles they read. To support this goal, this study explores enhancements to a journal reading system by applying the idea of the functional unit, the smallest information unit with a distinct function within four major components of scholarly journal articles—Introduction, Methods, Results and Discussion. This study examined a set of functional units and their associations with scholarly journal article use tasks through literature analysis and validation surveys. Forty-one typical functional units were found in psychology journal articles, with varying relevance to five tasks requiring use of information in journal articles. The relationships among sets of functional units for particular tasks were also identified. A taxonomy was developed incorporating the relationships between functional units and information use tasks, which can be used to inform system design. Based on this taxonomy, a prototype journal reading environment signalling functional units was designed and implemented for testing.

1. Introduction

Within academia, scholars are reading an increasing number of scholarly articles. With no corresponding increase in the time available for this activity, the time spent on each item read has declined. For a university science faculty member in the United States, the average number of articles read per year increased from 150 in 1977 to 280 in 2005, while the average time spent per article read decreased from 48 min in 1977 to 31 min in 2005 (Tenopir, King, Edwards, & Wu, 2009; Tenopir, King, Spencer, & Wu, 2009). This suggests that there is a need for ways to support more effective and efficient information use within academia.

The concept of genre, referring to the relatively stable and expected form and content for communication within a particular community (Breure, 2001), provides a means of looking at information system design from a document-oriented perspective. Most genre-based information science research is focused at the document level, but some studies (Dillon, 2004; Vaughan & Dillon, 1998) have taken a more analytical approach by studying the genre of components within journal research articles—*Introduction*, *Methods*, *Results*, and *Discussion* (IMRD). The current study seeks to facilitate the use of journal research articles by utilizing the functional units within components. Here, a functional unit is defined as a chunk of information embedded in the Introduction, Methods, Results, or Discussion components of a journal article, which serves a distinct communicative function. The concept of functional units is based on Swales' CARS (Create a Research Space) model (1990) and Sperber and Wilson's Relevance-theoretic Comprehension Procedure (1995).

This study seeks to identify the functional units in the core components (IMRD) of journal articles and to map their relationships with typical information tasks that prompt use of journal articles. By exploiting the mapping between functional units and tasks, we may be able to help users complete a particular task more efficiently by presenting them with the text in the article that is most relevant to the task, rather than presenting the article in its entirety. "Journal article" in this paper specifically refers to scholarly journal articles that follow the conventional IMRD format for reporting research. This study focuses on the psychology domain because adherence to APA (American Psychological Association) style in this domain has resulted in a relatively mature research article genre. This study addresses three research questions:

(i) What functional units exist within scholarly journal articles in the field of psychology?
(ii) How are functional units related to different tasks requiring use of information in journal articles?
(iii) How are functional units related to each other for a particular task requiring use of information in journal articles?

☆ This article is an extended version of a paper (Zhang, Kopak, Freund, and Rasmussen, 2010) presented at 2010 Annual Meeting of American Society for Information Science & Technology.

* Corresponding author.
E-mail addresses: leizhang@interchange.ubc.ca, iamleizhang@gmail.com (L. Zhang), rkopak@interchange.ubc.ca (R. Kopak), luanne.freund@ubc.ca (L. Freund), edie.rasmussen@ubc.ca (E. Rasmussen).

0268-4012/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijinfomgt.2010.10.003

Sections 2 and 3 present the background for the study, Sections 4–6 present the study itself, and Section 7 applies the results to the design of a prototype reading environment, which makes use of functional units to support efficient use of scholarly journal articles.

2. Theoretical framework

The concept of functional units as used in this study is based on Swales' "Create a Research Space" (CARS) model, while the idea that functional units are inter-related is based on Sperber and Wilson's Relevance-theoretic Comprehension Procedure.

2.1. Swales' CARS model

According to Swales, "a 'move' in genre analysis is a discoursal or rhetorical unit that performs a coherent communicative function in a written or spoken discourse" (2004, p. 228). As further noted by Swales, the move structure of an individual article component consists of functionally distinct steps. Based on a move analysis of 48 articles in the "hard" sciences, social sciences, and life and health sciences, Swales (1990, p. 141) proposes the CARS model for writing academic introductions:

Move 1:	Establishing a territory
	Step 1: Claiming centrality *and/or*
	Step 2: Making topic generalization(s) *and/or*
	Step 3: Reviewing items of previous research
Move 2:	Establishing a niche
	Step 1A: Counter-claiming *or*
	Step 1B: Indicating a gap *or*
	Step 1C: Question-raising *or*
	Step 1D: Continuing a tradition
Move 3:	Occupying the niche
	Step 1A: Outlining purposes *or*
	Step 1B: Announcing present research
	Step 2: Announcing principal findings
	Step 3: Indicating RA structure

In this way, the overall meaning of "introduction" is realized through a sequence of moves, each of which is realized through several steps. The boundaries between moves are indicated by changes in the type of information communicated. A number of studies of IMRD components within different corpora in various disciplines have been based on Swales' genre model. These include work on the Results (Brett, 1994; Thompson, 1993) and Discussion (Hopkins & Dudley-Evans, 1988; Lewin, Fine, & Young, 2001) components.

2.2. Sperber and Wilson's Relevance-theoretic Comprehension Procedure

Relevance theory, proposed by Sperber and Wilson (1986), addresses everyday speech utterances from the cognitive perspective. Sperber and Wilson (1995) differentiate between two principles of relevance. The Cognitive Principle of Relevance states that human cognition tends to maximize relevance in processing information—to gain the greatest cognitive effects with the least processing effort. Applied to cognitive processes in verbal communication, the Communicative Principle of Relevance states that an intentional act of communication conveys the presumption of optimal relevance: that it is at least relevant enough to be worth the addressee's attention and is as relevant as the addresser could have made it given his or her abilities and preferences. Comprehension, then, starts with the recovery of linguistically encoded meaning, and continues with the recovery of the explicit meaning and the implicit meaning. The addressee follows a path of least effort and stops at the first interpretation that satisfies his or her expectations of relevance. This is the Relevance-theoretic Comprehension Procedure (Wilson & Sperber, 2004, 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).

Saracevic (2007) notes that relevance theory as proposed by Sperber and Wilson has had more impact on thinking about relevance in information studies than work on relevance from other fields. Some examples of this impact are Harter's (1992) work on psychological relevance and White's (2007) examination of cognitive effects and processing effort in bibliometric retrieval.

3. Previous research

Current genre research in information studies has focused on the genres of digital documents, such as web pages and weblogs. Nevertheless, a few studies have examined article components. According to Dillon (2004), a component is a part-genre of a journal article. Vaughan and Dillon (1998) recruited domain expert users to categorize a set of paragraphs according to where they belong in an academic journal article. The experts' verbal protocols were subjected to a "how, why, what" content analysis. IMRD components have well-established roles to play: how they are read, why they are read, what content they should contain. However, results of "how, why, what" in reading were identified from users' conceptions of article components rather than from the documents themselves. IMRD components are also discussed briefly in related work that considers the role domain expertise can have in helping users locate information in articles (Dillon, 2000; Dillon & Schaap, 1996).

In work by Bishop et al. (Bishop, 1998, 1999; Bishop et al., 2000), components refer to all the logical subdivisions of a journal article, including article titles, author names, external links, abstracts, and references. The functions of journal article components can be to support finding relevant documents, assessing document relevance before retrieval, reading articles, creating document surrogates, reaggregation and integration into new documents. They found that readers tend to extract individual components from journal articles and incorporate them into their own writing. This idea was applied by Sandusky and Tenopir (2008) to the components of tables and figures. However, the implementations of this idea, Bishop's DeLIVER testbed and Sandusky and Tenopir's ProQuest CSA prototype, were focused on extracting logically discrete components from their embedding articles for the sake of searching and viewing. The work has also raised questions of whether or not individual components can stand alone, and what the minimum necessary information required would be (Sandusky & Tenopir, 2008). Other studies on structured document retrieval, though addressing the importance of document parts in relation to document structure, do not consider genre conventions.

Unger (2002, 2006) was one of the first to bridge genre and relevance theory through his work on linguistic discourse. He suggests that genre information contributes to the comprehension procedure by providing contextual assumptions for the inferential process, thus fine-tuning expectations of relevance. Genre information can generate expectations of relevance that are more or less precise: more precise in terms of what utterances to expect in which sequence; less precise in terms of the expected form and content of the text, or the kind of cognitive effects or level of relevance to be expected. Unger argues that genre can be incorporated into the relevance-theoretic comprehension procedure because of its influence on comprehension.

Yus (2007) extends the idea of bridging genre and relevance to weblog templates with the aim of stabilizing the weblog genre. Rel-

evance theory differentiates between procedural meaning (words encoding the manipulation of conceptual representations) and conceptual meaning (words encoding concepts). Yus considers the weblog template to possess a procedural quality, for verbal or visual features of weblogs can trigger an instant identification of the weblog genre. He suggests that "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).

Previous studies have not fully investigated the genre of article components and the link between genre and relevance in using digital documents. Research is needed in this area.

4. Methods

Following Swales' CARS model, a functional unit, which is derived from "steps", is the smallest possible unit of information, and related functional units should contain the minimum information one desires in a certain context. Following Sperber and Wilson's Relevance Theory, the comprehension procedure proceeds as follows: expectations of relevance generated by the most relevant functional unit can be extended to other related functional units within the component, which can be further extended to more related functional units beyond the component.

Our approach to studying functional units in the context of information use was, first, to identify common information use tasks that involve scholarly journal articles and to create a taxonomy of the common functional units within psychology journal articles. We then conducted two surveys to validate the findings from the identification study with members of the user population, and to refine the taxonomy. Survey I was conducted to validate the information use tasks and the functional units within four components of psychology journal articles. The purpose of Survey II was to validate the relationships between functional units and information use tasks, and the relationships among functional units for particular tasks.

4.1. Information tasks and functional units

We identified a set of information use tasks associated with scholarly journal articles from the literature on scholarly journal use and from Taylor's (1991) information use model. Six tasks requiring use of information in journal articles were identified:

- Keeping up: to keep current with articles in the user's area of research;
- Refer to facts: to consult specific factual information, e.g., data, phenomena;
- Refer to arguments: to consult arguments, ideas or suggestions supporting a point made by the user;
- Learn about background: to get to know a new area on which the user is embarking;
- Learn about particular: to understand a particular problem with its details and associated interpretation, judgment, etc.
- Learn how to: to learn how to do something, e.g., operation, procedure.

We developed the functional unit taxonomy for psychology journal articles through the following steps. We first selected prototype models from existing move structures, drawing upon Swales' model (1990) of Introductions, Brett's model (1994) of Results, and Hopkins and Dudley-Evans' model (1988) of Discussions, with move analysis in other works (Dubois, 1997; Holmes, 1997; Kanoksilapatham, 2005; Lewin et al., 2001; Nwogu, 1997; Thompson, 1993) as supplements. Secondly, we derived and refined the set of functional units from these prototype models based on their descriptions and examples. Lastly, we applied the framework of functional units developed in the first two steps to Introduction, Results, and Discussion components in twelve sample articles. The results of this endeavor showed that our initial functional unit framework was able to cover almost all functional units in psychology journal articles. Because the Methods component is highly discipline specific, functional units for this component were identified directly from the corpus. The result was a set of 52 functional units, distributed among the four journal components.

4.2. Instruments

Survey I consisted of three parts: task validation, functional unit validation, and participant background information. First, the participants were asked to indicate how frequently they used journal articles for the six information use tasks listed, by rating on a seven-point Likert scale (1 = Never, 7 = Very Frequently) and by putting them in rank order by frequency. They were free to suggest tasks other than those provided. Second, the participants were asked to indicate how frequently they thought each functional unit typically occurred in the Introduction, Methods, Results, and Discussion components of a psychology journal article by rating them on a five-point Likert scale (Never–Rarely–Occasionally–Very Frequently–Always). They could also suggest other functional units they thought occurred frequently, but were not in the list. To avoid cognitive overload and to reduce misinterpretation of the meaning of the functional unit involved, a one-sentence definition was provided for each functional unit in place of a title. Each definition of a functional unit was listed as a separate item for rating.

Based on the responses from Survey I, 41 functional units were included in Survey II (Appendix A), including 11 functional units in the Introduction, 10 in Methods, 7 in Results, and 13 in Discussion. Survey II consisted of six task scenarios designed to validate relationships between functional units and information tasks. Given each scenario, i.e., refer to facts, learn about background, refer to arguments, learn about particular, keeping up, learn how to, the participants were asked to rate the usefulness of the functional units within each of the components on a five-point Likert scale (1 = Not Useful at All, 5 = Highly Useful). They also were asked to rank the six most useful functional units within each component.

4.3. Participants

From mid June to mid July in 2009 we sent email advertisements to the graduate student listservs of the Departments of Psychology at two major research universities in Vancouver. Participants received $10 for completing two online surveys, each of which took approximately 30 min. Thirteen people participated in Survey I. Nine participants from Survey I also participated in Survey II. The thirteen participants, eleven female and two male, included six PhD students, five Masters students, one postdoctoral fellow and one PhD graduate. Most were between the ages of 26 and 30. Experience using journals within the group ranged from 4 to 18 years, with half the participants citing 6 or 7 years. Six participants used journal articles daily, and the remainder used them one to three times per month.

5. Results

5.1. Survey I

Data on the six information use tasks showed a high level of consistency between the task frequency rating and ranking. The results, in decreasing order of mean frequency, are: Learn about background ($M = 6.23$), Refer to facts ($M = 6.00$), Learn about particular ($M = 5.62$), Refer to arguments ($M = 5.23$), Keeping up ($M = 4.77$), and Learn how to ($M = 4.23$).

Table 1
Common functional units within components.

Components	Functional units
Introduction	Claim importance of topic
	Narrow down topic
	Review previous research
	Indicate a gap in previous research
	Provide reason to conduct research
	Point out contribution of previous research
	Introduce present research
	Present hypotheses
	Clarify definition
	Summarize methods
	State value of present research
Methods	Relate to prior/next experiments
	Justify methods
	Preview methods
	Describe participants
	Describe materials
	Describe tasks
	Outline experimental procedures
	Present variables
	Outline data analysis procedures
	Present reliability/validity
Results	Describe analysis conducted
	Restate hypotheses
	State findings
	State additional findings
	State non-validated findings
	Evaluate hypotheses
	Summarize results
Discussion	Recapitulate present research
	Provide established knowledge of topic
	Highlight overall outcome
	Indicate (un)expected outcome
	Compare results with previous research
	Interpret outcome
	Support explanation of results
	Generalize results
	Recommend future research
	Indicate significance of outcome
	Ward off counterclaim
	Indicate limitations of outcome
	Evaluate methodology

The mean scores on the prevalence of the 52 functional units within the four components were used to divide the functional units into three categories: 4.0–5.0 (Very Frequently–Always), 3.0–3.9 (Occasionally–Very Frequently), and 2.0–2.9 (Rarely–Occasionally). Scores were quite consistent: only two functional units had a standard deviation higher than 1.0.

The fifteen functional units in the 2.0–2.9 category were considered to be rarely or occasionally used, and most were dropped from further study. The Results component had the most functional units in this category because these functional units derived from the literature included commentary statements which overlapped with those in Discussion component. Two functional units "restate hypotheses" and "non-validated findings" which had no counterparts in Discussion were kept for subsequent study. Also we kept one low-scoring functional unit from the Methods component, which was observed with a high frequency in the identification study. Additionally, an item suggested by a participant was added to the Methods component in Survey II. The 41 functional units remaining after this selection process are shown in Table 1.

5.2. Survey II

To examine how useful participants considered the functional units to be for different information use tasks, a multivariate analysis-of-variance was conducted on functional units within four components for five tasks. The sixth task, "Learn about particular", was removed from the analysis because the data indicated that participants had not been consistent in their interpretation of the task. Each component showed significant differences between the five tasks. Post hoc tests were conducted to see which functional units were significantly different in each component.

First, a few functional units were significantly more useful than others in the same component. For example, in the Introduction component, the functional units "review previous research" and "point out contribution of previous research" were shown to be significantly more useful for the task "learn about background", while another three "indicate a gap in previous research", "provide reason to conduct research" and "state value of present research" were significantly more useful for "refer to arguments". In the Results component, the functional unit "state findings" was shown to be significantly more useful for the task "refer to facts", while another, "describe analysis conducted", was significantly more useful for "learn how to". The functional units which were significantly more useful in the Discussion component were: "provide established knowledge of topic" and "compare results with previous research" for the task "learn about background".

Furthermore, the same functional unit can be more or less useful for different tasks: in the Methods component the functional unit "justify methods" rated significantly lower than "outline experimental procedures" and "describe tasks" for "refer to facts", whereas it was significantly higher than "preview methods" and "describe participants" for the task "refer to arguments". For some tasks a functional unit was not significantly different from others, such as functional units in the Introduction component for the task "keeping up", those in Methods for "learn how to", and those in Discussion for "refer to facts". All these findings show that the usefulness of a component or a functional unit within a component varies with information tasks.

The highest values of ranking scores were also considered. A ranking score was calculated for each functional unit based on the frequency with which it was assigned each rank, using the formula $\sum(7-n)\times \text{freq}(n)$ where n is the rank and $\text{freq}(n)$ is the number of times the unit was assigned rank n.

As shown in Table 2, the functional units were grouped in terms of how useful they were for a particular task: the primary functional unit, related functional units in primary component and additional related functional units in other components. This was determined by their rating and ranking scores in Survey II. In a process similar to that used in Survey I, we placed the functional units in one of three categories 2.0–2.9, 3.0–3.9, and 4.0–5.0, based on their mean rating scores. The three categories represented the degree of usefulness of functional units within IMRD components (1 = Not Useful at All, 5 = Highly Useful) for each of five tasks. Functional units with the highest rating score across the four components were categorized as "primary component and functional unit". Those categorized as "related functional units in primary component" were functional units that scored from 4.0 to 5.0 in the same component as the primary functional units. Those in "additional related functional units in other components" were functional units with the highest rating scores in the other three components. The functional units with the top ranking scores were added to "additional related functional units in other components" if they did not duplicate others in this category.

6. Discussion

In this research, a taxonomy of functional units was developed in which the units are classified by the level of usefulness they have for a particular information use task. This taxonomy was developed

Table 2
Key relationships between information tasks, component and functional units.

Information tasks	Primary component and functional unit	Related functional units in primary component	Additional related functional units in other components
Learn about background	Introduction Review previous research	Point out contribution of previous research Indicate a gap in previous research Narrow down topic Clarify definition	Methods Relate to prior/next experiments Justify methods Results Summarize results Discussion Provide established knowledge of topic Compare results with previous research
Refer to facts	Results State findings	Evaluate hypotheses Summarize results	Introduction Review previous research Methods Outline experimental procedures Describe tasks Discussion Highlight overall outcome
Refer to arguments	Discussion Support explanation of results	Compare results with previous research Highlight overall outcome Interpret outcome Provide established knowledge of topic Generalize results Indicate (un)expected outcome Indicate significance of outcome Ward off counterclaim	Introduction Indicate a gap in previous research Claim importance of topic Methods Justify methods Results State additional findings State non-validated findings State findings
Learn how to	Methods Describe materials Describe tasks Outline experimental procedures	Justify methods Present variables Outline data analysis procedures Preview methods Describe participants Present reliability/validity	Introduction Summarize methods Results Describe analysis conducted Discussion Evaluate methodology
Keeping up	Introduction Indicate a gap in previous research	Provide reason to conduct research Point out contribution of previous research Review previous research Claim importance of topic Introduce present research	Methods Justify methods Relate to prior/next experiments Outline experimental procedures Results State findings Discussion Recommend future research Highlight overall outcome

by identifying and validating the functional units in IMRD components and also their relationships with information use tasks in the case of psychology journal research articles. We use the term taxonomy, which normally implies hierarchical classification, because functional units are viewed at the component level, and can be further classified into three sub-categories: primary, related, and additional.

First, six common information tasks and 41 common functional units were identified for psychology journal articles. Second, it was found that to a component, the extent of usefulness is not the same for five information tasks, and even when a component proves useful for a task, not all functional units within the component are equally useful for that task. Furthermore, for a particular task, a functional unit is more closely associated with certain functional units than other functional units in the same component and in other components.

The taxonomy of task-related functional units indicates the relevance of functional units for each information use task, with functional units in three categories. There is clearly a relationship between functional units and information use tasks, and among a set of functional units for a particular task. This taxonomy of functional units can provide guidance for the design of journal reading systems that facilitate information use by emphasizing the functional units strongly related to particular tasks. This study indicates how functional units support information tasks associated with journal reading in the following ways:

a. A functional unit is the smallest information unit. By employing functional units, we can help users to focus on the highly relevant information within an article.
b. A functional unit is associated with other functional units in the same and different components for a particular task. By employing the associations between functional units, we can help users to connect pieces of relevant information across journal articles.
c. Functional units are classified into three categories according to how useful they are for a particular task. By employing functional units of varying relevance, we can help users to move from the most relevant to the least relevant information, and stop at the amount of information the user desires.

A focus on the smallest information unit can help the reader by enabling them to spend more time on reading rather than on locating information, while connecting these information units by function and by relevance to particular tasks can help the reader achieve the greatest possible effect with the least possible effort. The notion of “moves” has long been confined to the pedagogy of academic writing. This study extends the idea of functional units, originating from “moves”, to information use of digital documents. Additionally, this study incorporates the concept of genre into the cognitive processing of relevant information by following a relevance-theoretic comprehension procedure.

Journal usage is discipline-dependent. Even within journal publications, article genre may vary, e.g., theory pieces, review articles,

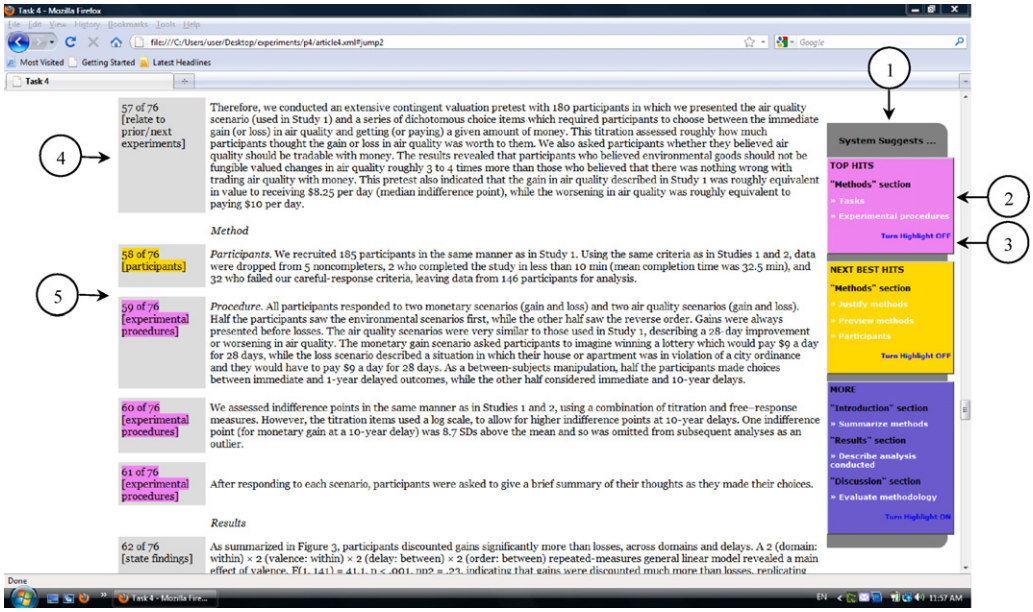

Fig. 1. Prototype reading environment. (1) Relevant functional units categorized into three boxes, (2) functional unit titles, (3) toggle on–off button, (4) paragraph numbers and labels (5) highlighted paragraphs.

data-based research articles, and shorter communications (Swales, 2004). Although this taxonomy is developed from one genre in a specific domain – the psychology journal research article – the research methods and results from this research can be used to inform investigations into other genres and disciplinary domains.

7. Implications for system design

Currently, there are a limited number of approaches to using genre in information system design. Genre information has been incorporated into document representations (Glover, Lawrence, Gordon, Birmingham, & Giles, 2001; Rosso, 2008), query formulation (Roussinov et al., 2001), and has been used to customize the ranking of search results for different task scenarios (Freund, 2008; Yeung, Freund, & Clarke, 2007). These approaches have all focused on using genre at the aggregate, document level to support information retrieval, rather than at the granular, within-document level to support reading and use.

Applying the results of our study to the design of a reading environment has the aim of enhancing reading effectiveness and efficiency by signalling functional units. Based on these two approaches, the signalling of functional units can be realized by labelling each paragraph by function or by highlighting the functional units relevant to a particular task. We designed a prototype reading environment that uses both these approaches. It incorporates a functional unit indicator to inform the reader of the function of each paragraph, and a functional unit selector to highlight functional units that are most likely to be relevant for particular tasks. This prototype was designed to investigate the extent to which this type of system would benefit users.

7.1. Interface

Fig. 1 shows the prototype reading environment. The functional unit indicator, located along the left margin, shows the label(s) and paragraph number for each paragraph. For the prototype, we manually identified functional units at the paragraph level within a corpus of psychology journal articles, drawing upon the taxonomy of 41 functional units developed in Survey I (Table 1). Each paragraph is labelled with at least one and at most two functions if both are equally important. Paragraphs are numbered sequentially and the number is shown as XX (current paragraph number) of YY (total paragraph number) to give the user a sense of the size of the article and his or her location within it.

The functional unit selector is located along the right margin. It displays functional units categorized into three boxes labelled as Top Hits, Next Best Hits, and More. The functional units in the three boxes are categorized by relevance: “primary functional units” are “Top Hits”, “related functional units in primary component” are “Next Best Hits”, and “additional related functional units in other components” are “More”. These relationships are based on the results of Survey II, as presented in Table 2. In each box is a toggle on–off button. When a user clicks on the button, those paragraphs whose functions are listed in that box are highlighted in the left margin with a corresponding colour, and the user is taken to the first highlighted paragraph. In addition to the toggle on–off button, each box contains a list of functional units and the sections in which they occur so the users may have an idea of what is available. People may combine use of interface functionalities to narrow down reading step by step: functional unit titles in three boxes help one decide where to go first, a click on toggle on–off button directs one to highlighted paragraphs, and the paragraph labels on the left margin enable one to narrow down reading further. People may also rely primarily on the functional unit selector or the functional unit indicator.

The prototype was created with XML, XSLT and JavaScript. We are currently conducting a user assessment in which the same group of participants read journal articles using the prototype interface incorporating functional units and using a baseline interface with no added functionality. The study is designed to test for an effect of signalling functional units on reading effectiveness and efficiency. This part of the research is currently underway.

8. Conclusion

This study explores the functions of types of information within the components (Introduction, Methods, Results, and Discussion) of a journal article. Based on the results of analysis of research articles in psychology, the conclusions related to the original research questions are: (i) a set of 41 functional units typically exists within psychology journal articles; (ii) different functional units are more or less useful for different tasks, and this variation is not consistent

with respect to all components of the article; and (iii) for a particular task, a functional unit is more or less closely connected with other functional units in the same component and in other components.

By modeling functional units and their relationships with information tasks, we have created a prototype reading environment that indicates the functions of each paragraph and highlights relevant paragraphs for a particular task. The preliminary results of the experimental user study are promising and suggest that users may derive real benefit from such systems. Further research is needed to develop methods for automatic classification and annotation of discourse, which are necessary pre-requisites for the real-world implementation of genre-enabled reading environments. Given the growing pressure of reading more in less time, the development of such systems is a priority. This work represents one step in that direction.

Acknowledgments

Our thanks to the participants who took part in the study.

Appendix A. Validation Survey II: Relationships between Functional Units and Information Tasks

Use Scenario 1: Refer to Facts

Imagine yourself in the following scenario: You are trying to refer to specific factual information, e.g., data, phenomena. With a

1.1 To refer to facts, how do you rate the usefulness of the following types of information within the *Introduction* section? Please select one of the five responses (1 = Not useful at all, 5 = Highly useful). Please also rank the **SIX** most useful types of information (e.g., put 1 next to most important, 2 next to second most important).

	Rating (1 = Not useful at all, 5 = Highly useful)					Rank
	1	2	3	4	5	
Showing that the research area is important/central/interesting/problematic/ relevant in some way	☐	☐	☐	☐	☐	
Increasingly narrowing down the topic of the research	☐	☐	☐	☐	☐	
Reviewing items of previous research in the area	☐	☐	☐	☐	☐	
Pointing out deficiencies in the present state of knowledge	☐	☐	☐	☐	☐	
Providing positive reasons for conducting the research reported	☐	☐	☐	☐	☐	
Pointing out contribution of previous research	☐	☐	☐	☐	☐	
Outlining purposes of the present research	☐	☐	☐	☐	☐	
Presenting hypotheses or research questions	☐	☐	☐	☐	☐	
Discussing the definitions of key terms	☐	☐	☐	☐	☐	
Summarizing the methods used	☐	☐	☐	☐	☐	
Stating the value of the present research	☐	☐	☐	☐	☐	

Any Comments

1.2 To refer to facts, how do you rate the usefulness of the following types of information within the *Methods* section? Please select one of the five responses (1 = Not useful at all, 5 = Highly useful). Please also rank the **SIX** most useful types of information (e.g., put 1 next to most important, 2 next to second most important).

	Rating (1 = Not useful at all, 5 = Highly useful)					Rank
	1	2	3	4	5	
For several experiments in a paper, describing the relations between the current experiment and prior/subsequent experiments	☐	☐	☐	☐	☐	
Stating the rationale for the decision to use particular experimental methods, procedures, or techniques	☐	☐	☐	☐	☐	
Previewing methods in the following experiment	☐	☐	☐	☐	☐	
Describing the participants in the study	☐	☐	☐	☐	☐	
Describing materials/stimuli in the study	☐	☐	☐	☐	☐	
Describing tasks in the study	☐	☐	☐	☐	☐	
Describing the procedures of an experiment	☐	☐	☐	☐	☐	
Describing variables in the study	☐	☐	☐	☐	☐	
Describing the procedures used in data analysis	☐	☐	☐	☐	☐	
Describing the reliability/validity of the measures used in the experiment	☐	☐	☐	☐	☐	

Any Comments

1.3 To refer to facts, how do you rate the usefulness of the following types of information within the *Results* section? Please select one of the five responses (1 = Not useful at all, 5 = Highly useful). Please also rank the **SIX** most useful types of information (e.g., put 1 next to most important, 2 next to second most important).

	Rating (1 = Not useful at all, 5 = Highly useful)					Rank
	1	2	3	4	5	
Providing a description of the analysis conducted	☐	☐	☐	☐	☐	
Restating the aims of the research, or creating further hypotheses from the findings	☐	☐	☐	☐	☐	
Providing the results of the data analysis	☐	☐	☐	☐	☐	
Stating the data that neither support nor conflict with the major finding	☐	☐	☐	☐	☐	
Accounting for the data that do not support the major finding	☐	☐	☐	☐	☐	
Evaluating findings with regard to the hypotheses	☐	☐	☐	☐	☐	
Summarizing a number of results and explanations	☐	☐	☐	☐	☐	

Any Comments

1.4 To refer to facts, how do you rate the usefulness of the following types of information within the *Discussion* section? Please select one of the five responses (1 = Not useful at all, 5 = Highly useful). Please also rank the **SIX** most useful types of information (e.g., put 1 next to most important, 2 next to second most important).

	Rating (1 = Not useful at all, 5 = Highly useful)					Rank
	1	2	3	4	5	
Strengthening the discussion by recapitulating main points such as research questions, aims and purposes, theoretical or methodological information	☐	☐	☐	☐	☐	
Describing established knowledge on the topic	☐	☐	☐	☐	☐	
Highlighting the overall research outcomes	☐	☐	☐	☐	☐	
Commenting on whether the results are expected or not	☐	☐	☐	☐	☐	
Referring to previous research for comparison	☐	☐	☐	☐	☐	
Explaining specific research outcomes	☐	☐	☐	☐	☐	
Claiming support for explanation by exemplifying or citing	☐	☐	☐	☐	☐	
Making a claim about the generalizability of the particular results	☐	☐	☐	☐	☐	
Making suggestions for future research	☐	☐	☐	☐	☐	
Indicating significance of the outcome	☐	☐	☐	☐	☐	
Providing arguments in response to the potential criticism raised by the reader	☐	☐	☐	☐	☐	
Indicating limitations of outcomes	☐	☐	☐	☐	☐	
Evaluating the effectiveness of the methodology in hindsight	☐	☐	☐	☐	☐	

Any Comments

journal research article at hand, you need to identify how useful various types of information are for referring to facts.

(This is repeated for each of the six scenarios.)

References

Bishop, A. P. (1998). Digital libraries and knowledge disaggregation: The use of journal article components. In *Digital libraries 98: Proceedings of the 3rd ACM conference on digital libraries* (pp. 29–39).

Bishop, A. P. (1999). Document structure and digital libraries: How researchers mobilize information in journal articles. *Information Processing and Management*, *35*(5), 255–279.

Bishop, A. P., Neumann, L. J., Star, S. L., Merkel, C., Ignacio, E., & Sandusky, R. J. (2000). Digital libraries: Situating use in changing information infrastructure. *Journal of the American Society for Information Science*, *51*(4), 394–413.

Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, *13*(1), 47–59.

Breure, L. (2001). *Development of the genre concept*. Retrieved from http://people.cs.uu.nl/leen/GenreDev/GenreDevelopment.htm.

Dillon, A. (2000). Spatial-semantics: How users derive shape from information space. *Journal of the American Society for Information Science*, *51*(6), 521–528.

Dillon, A. (2004). *Designing usable electronic text* (2nd ed.). Boca Raton, FL: CRC Press.

Dillon, A., & Schaap, D. (1996). Expertise and the perception of shape in information. *Journal of the American Society for Information Science*, *47*(10), 786–788.

Dubois, B. L. (1997). *The biomedical discussion section in context*. Greenwich, CT: Ablex.

Freund, L. (2008). *Exploiting task-document relations in support of information retrieval in the workplace*. Doctoral dissertation. Faculty of Information Studies, University of Toronto.

Glover, E. J., Lawrence, S., Gordon, M. D., Birmingham, W. P., & Giles, C. L. (2001). Web search—your way. *Communications of the ACM*, *44*(12), 97–102.

Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, *43*(9), 602–615.

Holmes, R. (1997). Genre analysis, and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, *16*(4), 321–337.

Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes*, *7*(2), 113–121.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(5), 269–292.
Lewin, B. A., Fine, J., & Young, L. (2001). *Expository discourse: A genre-based approach to social science research texts*. London: Continuum.
Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, *16*(2), 119–138.
Rosso, M. A. (2008). User-based identification of web genres. *Journal of the American Society for Information Science and Technology*, *59*(7), 1053–1072.
Roussinov, D., Crowston, K., Nilan, M., Kwasnik, B., Cai, J., & Liu, X. Y. (2001). Genre based navigation on the web. In *Proceedings of the 34th Hawaii international conference on system sciences*
Sandusky, R. J., & Tenopir, C. (2008). Finding and using journal-article components: Impacts of disaggregation on teaching and research practice. *Journal of the American Society for Information Science and Technology*, *59*(6), 970–982.
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*(13), 1915–1933.
Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.
Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.
Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge, UK: Cambridge University Press.
Taylor, R. S. (1991). Information use environments. In B. Dervin, & M. J. Voigt (Eds.), *Progress in communication sciences* (pp. 217–255). Norwood, NJ: Ablex.
Tenopir, C., King, D. W., Edwards, S., & Wu, L. (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings: New Information Perspectives*, *61*(1), 5–32.
Tenopir, C., King, D. W., Spencer, J., & Wu, L. (2009). Variations in article seeking and reading patterns of academics: What makes a difference? *Library & Information Science Research*, *31*(3), 139–148.
Thompson, D. K. (1993). Arguing for experimental "facts" in science: A study of research article results sections in biochemistry. *Written Communication*, *10*(1), 106–128.
Unger, C. (2002). Cognitive-pragmatic explanations of socio-pragmatic phenomena: The case of genre. In *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. In *Proceedings of 61st annual meeting of the American Society for Information Science* (pp. 559–566).
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*(4), 536–559.
Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607–632). Oxford: Blackwell.
Yeung, P. C. K., Freund, L., & Clarke, C. L. A. (2007). X-site: A workplace search tool for software engineers. In *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval* (p. 900). New York, NY: ACM.
Yus, F. (2007). Weblogs: Web pages in search of a genre? In S. Posteguillo, M. J. Esteve, & M. L. Gea-Valor (Eds.), *The texture of Internet: Netlinguistics in progress* (pp. 118–142). Cambridge: Cambridge Scholars Publishing.
Zhang, L., Kopak, R., Freund, L., & Rasmussen, E. (2010). A taxonomy of functional units for information use of scholarly journal articles. In *Proceedings of 2010 annual meeting of American Society for Information Science & Technology*.

Lei Zhang is PhD Candidate at the School of Library, Archival and Information Studies at the University of British Columbia. She is currently completing her dissertation, which studies the functions of types of information within journal articles and its application in supporting information use. Her main research interests include human computer interaction and information architecture.

Rick Kopak is Senior Instructor at the School of Library, Archival and Information Studies at the University of British Columbia, where he teaches courses in information design. He received his Ph.D. from the Faculty of Information at the University of Toronto. His research interests are in the area of human information interaction and information design focussing on the use of information structure in guiding user navigation.

Luanne Freund is an assistant professor at the School of Library, Archival and Information Studies at the University of British Columbia, where she teaches courses in information retrieval and information sources and services. She received her Ph.D. from the Faculty of Information at the University of Toronto. Her research interests include human information interaction and information retrieval, with a particular focus on the role of domains, tasks and digital genres.

Edie Rasmussen is Professor in the School of Library, Archival and Information Studies at the University of British Columbia, Vancouver, BC, Canada, where she served as Director for six years. Prior to joining UBC she was a Professor in the School of Information Sciences, University of Pittsburgh, Pittsburgh, USA. She has also held appointments at the School of Library and Information Studies at Dalhousie University, Nova Scotia, Canada, the School of Library Science at the Institiut Teknoloji MARA, Kuala Lumpur, Malaysia, and visiting positions at Nanyang Technological University, Singapore, Victoria University of Wellington, New Zealand, and Oslo University College, Norway. Her current research interests include indexing and information retrieval in text and multimedia databases and digital libraries.