June 2017

Linguistics and LIS: A Research Agenda

Mary K. Bolin
University of Nebraska–Lincoln, mbolin2@unl.edu

Follow this and additional works at: https://scholarworks.sjsu.edu/ischoolsrj

Part of the Library and Information Science Commons

Recommended Citation

This article is brought to you by the open access Journals at SJSU ScholarWorks. It has been accepted for inclusion in School of Information Student Research Journal by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.
Linguistics and LIS: A Research Agenda

Abstract
Linguistics and Library and Information Science (LIS) are both interdisciplinary fields that draw from areas such as languages, psychology, sociology, cognitive science, computer science, anthropology, education, and management. The theories and methods of linguistic research can have significant explanatory power for LIS. This article presents a research agenda for LIS that proposes the use of linguistic analysis methods, including discourse analysis, typology, and genre theory.

Keywords
linguistics, library and information science research, discourse analysis, genre theory, systemic-functional linguistics

About Author
Mary K. Bolin, PhD, is Professor and Catalog and Metadata Librarian at the University of Nebraska–Lincoln, as well as a Lecturer in the School of Information at San Jose State University. She is one of the founding editors of the open access, peer-reviewed e-journal Library Philosophy and Practice.

This invited contribution is available in School of Information Student Research Journal:
https://scholarworks.sjsu.edu/ischoolsrj/vol7/iss1/2
Linguistics is an interdisciplinary field that draws from study of languages, including English, and fields such as psychology, sociology, cognitive science, computer science, and anthropology. Library and Information Science (LIS) is also interdisciplinary, and can be studied using techniques from the humanities, social science, and science. The many theories and methods of linguistic research can be extremely useful and have significant explanatory power for LIS. This article presents a research agenda for LIS that proposes the use of linguistic analysis methods.

The elements of language are phonology, morphology, syntax, and semantics. The study of linguistics includes those areas, but also includes discourse analysis, linguistics universals and typology, sociolinguistics, psycholinguistics, language and cognition, language acquisition (including child language and second language acquisition), and many other topics and approaches. Language is a semiotic system, a system of signs. Halliday (1978) calls language a social semiotic. Written and spoken language are systems of signs that are used and understood by speakers. Languages and variants of languages are used in speech communities (e.g., speakers of Parisian French) and discourse communities (e.g., librarians) for purposes that include those of business and commerce, education, government, medicine, law, and every kind of human social and cultural event and occasion. We talk to each other, we read and write, and we carry out daily endeavors and long-term goals using language. As librarians, we already recognize the significance of the language that we use, in controlled vocabularies, in OPAC displays, in library signage and marketing, and in planning and problem-solving. As researchers, we can use the techniques of linguistic analysis to further unpack those plans and problems, and discover new theories and frameworks for helping library patrons discover and use information.

There is substantial and groundbreaking work being done in areas of library and information science such as search engine optimization, semantic web, natural language processing, and linked data. Those subjects are certainly linguistically oriented and often draw on the techniques of linguistic analysis, but this article does not focus on that area of the LIS and related literature. It provides information on frameworks, theories, and methods used in linguistics as they might be applied to many areas of LIS.

**Typology**

Typologies are used by many fields, but they are widely used in linguistic research, often as part of the search for linguistic universals: features or elements that are common to all, most, or many languages, and the contrast between the most and least common types in an area. Typology is used in research on semantic areas like kinship, color terms, and other culturally-salient phenomena, as well as syntax (the most and least common order of grammatical constituents in different languages). For example, the default order of constituents in English is subject-verb-object (SVO). SVO is one of the most common word order types, while OVS is the least common (“Word Order,” n.d.). Comrie (1989), Croft (1990), Greenberg (2005), and other scholars have compared the characteristics of different languages to identify universal phenomena. Prototype semantics can also be used to create typologies (Lakoff, 1986; Rosch, 1973, 1977). In prototype theory, there are
semantic categories with central and peripheral members. Members of a speech or discourse community may not agree on the boundaries of a category, but there is agreement about the center of the category or about its best representative, for example, a sparrow is a more typical bird than a penguin. Typology can be used in LIS research for studying things like librarian faculty status (Bolin 2007, 2008a, 2008b), library organizational patterns (Bolin, forthcoming 2017), and many other areas of library programs and services. It requires gathering data to answer a research question and then answering the question by dividing the data into types. The types are created using clusters of characteristics; for example, Bolin (2007) gathered data about librarian status at US land grant universities, and used characteristics such as eligibility for tenure, librarian rank system, and other things to create a typology of librarian status that including three faculty types and one staff type. There could be many other applications of linguistic typology to LIS. Those include:

- Models of liaison librarianship in academic libraries considering assignment of subject areas, services provided, types of instruction, and so on.
- A framework for collection evaluation based on format, age, use and other characteristics.
- Performance evaluation for librarians and staff, including frequency, depth, interactivity, rating scale, and areas of assessment.
- Access policies, including patron categories, loan periods, fines, licensing, and use of electronic resources.
- Cataloging and metadata workflows, including division of labor, MARC and non-MARC metadata, use of repositories such as CONTENTdm and Rosetta, et cetera.

The creation of a typology could be used to explore any of these areas (and many others), by posing a question and gathering data to categorize attributes. For example, information on performance evaluation at a group of 50 academic libraries might yield a typology such as:

- Department chair writes a letter of evaluation for librarians once a year.
- Librarian does self-evaluation and meets with department chair to come to agreement on strong and weak points.
- Department chair uses evaluation form with rating scale.
- Some mixture of these processes is used.

The creation of the typology is a qualitative activity that assesses which characteristics are salient (e.g., the use of a rating scale in performance evaluation), as well as lumping or splitting characteristics to create types. It is a lens for analysis that can help make sense of large amounts of data. Creating a typology often uses a kind of componential analysis, first used in research on phonology to describe how sounds are differentiated (e.g., \(t\) and \(d\) are distinguished by the voicing of the alveolar stop in the case of \(d\). \textit{Voice} is the component that is used to distinguish the two sounds. Trubetskoy, 1969). Componential analysis was adopted in other areas of linguistics and has been used in semantic analysis as well, for example, the difference between the cooking terms \textit{fry} and \textit{bake} includes the component \textit{oven}. Bake is described as +oven, while fry is -oven (Coseriu, 1973; Katz & Fodor, 1963).
Semantic Fields and Frames

Semantic fields are also called lexical fields, and they are groups of related words that might be synonyms from a domain (e.g., cooking), or words related in some other way. They are often used in contrastive linguistics, which compares one or more languages to see how concepts map in different languages. Bolin (1999) compared the semantic field grace in texts from the Bible in their original languages as well as in Latin, English, and German. The words in the field (English words include grace, mercy, kindness, compassion, and pity) did not have a one-to-one correspondence between languages. Semantic fields and frames deal with different types of meaning, which include referential, social, and encyclopedic meaning. Bolin describes these categories of meaning, saying that,

‘Referential’ meaning is the denotational, dictionary definition of the meaning of a word … ‘[s]ocial’ or emotive meaning includes ... connotations that include social or class markers, differences in register such as slang, a word’s pejorative connotation … ‘[e]ncyclopedic’ meaning … is all the baggage that any word carries, referential and social meaning, plus the combined weight of all the accumulated meanings, history, and cultural associations that the word carries. (1999, p. 8)

Semantic frames start with a domain or concept rather than with a group of words. They use the encyclopedic meaning of words and concepts to understand the social, cultural, historical, and any other aspects of meaning of words in a domain. The University of California, Berkeley maintains a site called FrameNet ([https://framenet.icsi.berkeley.edu/fndrupal/](https://framenet.icsi.berkeley.edu/fndrupal/)) that is a collection of semantic frames. An excerpt from the frame accuracy is shown in Figure 1.

**Accuracy**

**Definition:**

An **error** is involved in an activity whose degree of success is dependent on a parameter of the action matching a particular set of values of a continuous variable or variables (e.g. quantity, location, time). The **agent**, or metonymically the **instrument** or **means** action, is described in terms of the actual or expected **deviation** between the location, time, or quantity in the activity and the location, time, or quantity which is necessary for the intended event.

The **range** of the agent is **accurate** to within one foot per hundred yards.

His estimate was **very** by an order of magnitude.

The fork truck operator must be **very** accurate in his aim.

The **precision** of the daily measurements is dependent on a number of factors.

**FEs:**

Core:

<table>
<thead>
<tr>
<th>Agent (Ago)</th>
<th>The individual who tries to interact with the Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Type: Sentient</td>
<td>The maintainer is very <strong>exact</strong> about addresses -- he has to be.</td>
</tr>
</tbody>
</table>

**Figure 1.** The semantic frame **accuracy**.

This excerpt shows the referential definition of accuracy, examples of its use, and the social and grammatical participants in the concept. There are many uses for semantic fields and semantic frames in LIS research. Research using semantic fields could include:

- An analysis of the syndetic structure of Library of Congress Subject Headings (LCSH), Library of Congress Classification (LCC), or Dewey
Decimal Classification (DDC) using a semantic field for a domain, for example, food, war, industry, shelter, music, and so on.

- Semantic field analysis using an aspect of library terminology, for example, \textit{format} as perceived by librarians and library users.
- Semantic field analysis of library staff and librarian job descriptions, for example, what are the relationships among responsibilities assigned to one or more persons or positions?
- Attributes of information as one or more semantic fields as used in MARC, Dublin Core, and other metadata schemes.

Semantic frames are broader and not as based in lexical items. LIS research could use semantic frames in many ways, for example,

- As with semantic fields, semantic frames could be used to examine areas of LCSH, LCC, and DDC to see how relationships are expressed, and determine how much of the encyclopedic meaning of words and concepts can be expressed in a controlled vocabulary or thesaurus.
- Interviews with users could be used to create semantic frames for library services. Examples include instruction, collections, spaces, electronic resources, and so on. Cognitive framing by users may be quite different than the frames used by librarians. Reconciling those frames could improve library services.
- The organizational structure of libraries could be analyzed and re-engineered using semantic frames. Exploring frames such as service, employment, education and training, as well as the frames for library services such as cataloging, reference, circulation, and so on, can provide insight and help generate new ideas.

**Discourse Analysis**

Discourse is often defined as “language in use” or “language above the level of the sentence,” (“Discourse,” n.d.), for example, longer texts or utterances that have significant social and cultural meaning. Discourse analysis is used by many fields, sometimes using techniques that may not be considered linguistic analysis. Approaches to discourse analysis that may be useful in LIS including analysis of spoken discourse, for example, a reference interview, which is a communicative event that has meaning in the discourse community of librarians. The need for positive interactions makes it worthwhile to analyze the discourse of events such as these in which librarians and users interact.

Discourse analysis methods include examining the intersection of syntax and semantics, that is, how grammatical forms encode meaning, the study of dialects and \textit{registers} (language varieties used in social or professional situations, e.g., the language of medicine), critical discourse analysis (Fairclough 1989), which decodes and critiques power relationships, and many others. Conversation analysis, for example, studies the interactions of two or more speakers, including turn-taking, pragmatic meaning, and so on. The analysis of written texts can examine the cohesive devices link parts of a text together. Intertextuality — the relationship of one text with another — is a vital concept in the analysis of both written and spoken discourse (Kristeva, 1984). Both written and spoken discourse can follow scripts,
patterns, and schema that can be analyzed. (Halliday, 1978; Hoey, 2001; Hodge & Cress, 1988, 1993; Swales, 1990)

A useful place to start in considering discourse analysis techniques is Halliday’s (1978) systemic-functional linguistics (SFL). SFL approaches language by considering its functions; that is, linguistic elements as they are used to create meaning. SFL’s system networks, are systems which give choices to speakers. Those choices are determined by social identities and situations. The options and choices create a register, which Halliday calls “a recognizable language variety” (1978, p. 7). Examples include the language of medicine, education, or of a situation such as a reference interview in a library. SFL uses register variables to encode meaning. Field encodes ideational meaning (what a text or discourse is about). Tenor encodes interpersonal meaning (the participants and their roles and status). Mode encodes textual meaning (the devices that link the text together). Situational contexts of language are expressed by registers, and genre is the outermost layer, representing the cultural context and the genres used by a culture. Halliday (1978) describes language as a social semiotic, which is a system of signs that encode meaning. That social setting includes discourse communities (Nystrand 1982), which are professional or other social or cultural groups, who use language to mark themselves as members of their communities.


Discourse analysis may draw on the concept of a communicative event (Gumperz and Hymes, 1972). Communicative events (e.g., a job interview, a lecture, religious service) have rules and expectations that are familiar to discourse community members. Discourse analysis pays close attention to the concept of voice, that is, the people and communities implicitly present in a text. Texts with more than one voice represented are called heteroglossic or described as having voices in “heteroglossic opposition.” (Bakhtin, 1935). Bolin (2014), states that, “among academic librarians, there are the voices of reference, instruction, and collection development that were identified and discussed by Lemke (1999a) in his analysis of an academic library’s re-design of its website.” In Lemke’s view, “the Reference Orientation voice articulates a discourse formation in which primary positive valuations attach to servicing the user’s needs for information” (p. 30). The voice of the reference orientation advocated for a website that would give maximum access to users. In heteroglossic opposition was the instruction orientation voice that advocated the “teach a man to fish” approach, that is, to instruct users in how to find information rather than simply providing the information to them. This illustrates how contrasting voices and opposing discourses can still be based on the same ideology: the idea that librarians should use their expertise to provide services to users. Geertz (1973) introduced the idea of thick description, which examines a culture or community from the inside (as a
Like wise, Pike (1967) described *emic* and *etic* description, an allusion to the phonological concepts phonemic and phonetic. Lib rarians who do research on the discourse of their own community will produce an emic description, while an outsider would produce an etic one.

Examples of LIS research projects using discourse analysis include:
- Analysis of spoken or other interactive discourse in the library, including in-person, phone, and chat reference.
- Examination of internal communicative events such as evaluation conferences, interviews with job candidates, committee meetings, and so on.
- Research on the various discourse communities among library users including students (who come from different speech communities, socio-economic levels, and academic fields), faculty (who also vary demographically and have various information needs depending on their area of research), and other library users.
- Analysis of written texts and images such as letters and emails sent to library users, signage, press releases and announcements, and so on.

**Genres of Organizational Communication**

Genre analysis may be viewed as an aspect of discourse analysis. All organizations use both spoken and written genres to communicate. They may be unique to one organization or type of organization, but in practice there are genres that are shared by nearly all organizations and certainly by types of organizations. They may include something as generic as the memo, genres associated with employment such as vacancy announcements, letter of offer, contracts, job descriptions, and evaluations, as well as common but more specialized genres such as invoices, budget documents, annual reports, et cetera. Swales (1990, 2004) is a leading scholar on genres, and he describes genre sets and genre chains that are used in organizations, for example, the chain of documents used in hiring: vacancy announcement, letter of application, resume, interview questions, and letter of offer. Genres must meet expectations that are understood by the communities that use the genres. In hiring, for example, an organization judges a letter of application according to whether it meets the genre expectations, in terms of formal writing, appropriate content, and general characteristics of its appearance (e.g., not written on purple paper using Comic Sans). Bolin (2007, 2014, forthcoming 2017) examines genres used in academic libraries, including academic librarian appointment documents (e.g., promotion and tenure standards), academic library websites, and organizational charts. Genre analysis uses the techniques of discourse analysis, including determining authorship, uncovering the voices that are present in the text, the patterns the texts follow, who the participants are, what their relationship is, and how language encodes all these things. Possible research projects using genre analysis include:
- Examining a genre of organizational communication to gain understanding of how the use of that genre affects the library’s programs and services, for example, what is being communicated by the library website?
Examining internal genres to reveal how employees are being hired, retained, educated, and encouraged, and to see what organizational values are encoded in genres such as performance evaluation.

Looking at interactions with patrons as a genre and using data such as chat reference transcripts to improve service by understanding how this genre can be used.

Simmons (2005) discusses the application of genre theory to instruction in information literacy by librarians. She proposes using genre theory to introduce students to the discourse of various disciplines and move toward Critical Information Literacy, a version of Freire’s (1970) critical pedagogy.

Existing Studies

There is already interesting LIS research the uses linguistic approaches, including discourse analysis, various linguistic approaches to semantics, and the examination of documents, conversations, and other texts produced in and by libraries. The following is a selection of recent studies.


Conclusion

Librarians come to the profession with a master’s degree in library and information science (MLIS) that was preceded by an undergraduate degree that is virtually always in some other field: English, French, history, art history, music, biology, computer science, and so on. Depending on the nature and quality of their undergraduate program, librarians may be informed by the literature and research methods of those disciplines. They may also have other graduate degrees, a second master’s or a doctorate in a subject such as education, history, English, or any other discipline that will have provided formative experiences with professional literature and research methods and theoretical frameworks. Any of these can fruitfully inform LIS research and practice. Linguistics, with its focus on discourse, semantics, syntax, anthropology, and sociology, among other things, can be useful in any area of LIS. This article has briefly reviewed some prominent frameworks and methods in linguistic research, along with ideas for applying them to LIS research. These ideas may be more familiar and straightforward to librarians who have a background in linguistics, but there is a large body of interesting literature that is accessible to librarians and scholars who would like to learn more about linguistics and its methods.

References


