

# Peircean Semiotics and Subject Indexing: Contributions of Speculative Grammar and Pure Logic<sup>†</sup>

Carlos Cândido de Almeida\*, Mariângela Spotti Lopes Fujita\*\*,  
and Daniela Marjorie dos Reis\*\*\*

\*/\*\*/\*\*\*Information Science Department, São Paulo State University – UNESP, Av. Higino Muzzi Filho, 737, Marília, São Paulo, Brazil, CEP: 17525-900  
\*<carlosalmeida@marilia.unesp.br, \*\* <fujita@marilia.unesp.br>, \*\*\*<danielamajorie@yahoo.com.br>

Carlos Cândido de Almeida is Associate Professor, Information Science Department, São Paulo State University (UNESP). He holds a PhD in information science from São Paulo State University, a master's in information science from the Federal University of Santa Catarina, and graduated in library science from the State University of Londrina. He is the leader of the Theoretical Foundations of Information Research Group. His main research interests are epistemology of information science, Peircean semiotics, communication, and knowledge organization.

Mariângela Spotti Lopes Fujita is Professor, Information Science Department, São Paulo State University (UNESP). She is a researcher of the Brazilian National Research Council (CNPq). She holds a master's and PhD in communication sciences from São Paulo University and she graduated in library science from São Carlos School of Librarianship and Documentation. Her main research interests are indexing, documentary languages, documentary reading methodology, and knowledge organization.

Daniela Marjorie dos Reis graduated in library science and has a master's in information science from São Paulo State University (UNESP). Her main research interests are subject indexing and knowledge organization.

de Almeida, Carlos Cândido, Spotti Lopes Fujita, Mariângela, and dos Reis, Daniela Marjorie. **Peircean Semiotics and Subject Indexing: Contributions of Speculative Grammar and Pure Logic.** *Knowledge Organization.* 40(4), 225-241. 46 references.

**ABSTRACT:** The semiotics of C. S. Peirce presents fundamental concepts to discover aspects of the indexing process, including representation and classes of signs. However, we still know little of its theoretical potential for subject indexing. We believe that the main difficulty in the proposals to understand the process of subject indexing based on Peircean semiotics stems from an incomplete interpretation of his semiotic system. This paper attempts to describe the contributions of Peircean semiotics to subject indexing. First, we analyze some of the concepts of the branches of semiotics, after which, we discuss strategies for conceptual approximation. Secondly, and aiming to raise the level of interlocution between the areas, we intend to argue that subject indexing is an inferential process, as explained by the second branch of semiotics. Thus, we seek to go beyond the level of speculative grammar, the first branch of semiotics, to forge a closer link with pure or critical logic, the second branch. We conclude that the indexer's work does not produce a mere reflection of what already exists in documents, but involves an instigating action to discover, through the inferential matrix, the meaning of a text in order to find the subject and the most appropriate subject added entry to the information system.



† The authors would like to acknowledge the careful reading and contributions of Daniel Martínez-Ávila and José Augusto Chaves Guimarães as well as financial support from São Paulo State University Research Foundation (FUNDUNESP) and the Brazilian National Research Council (CNPq).

Received 21 May 2012; Revised 16 May 2013; Accepted 12 June 2013

“Among American thinkers, the most inventive and universal was probably Charles Sanders Peirce; so great was he that no university could find a position worthy of his merits.” (Jakobson 1999, 99). [our translation]

## 1.0 Introduction

Information is not an easily decipherable element. Documents, as records of signs that convey information, possess an aspect of content and another of form, and any knowledge organization system must observe these two facets. However, unlike Danish linguist Hjelmslev's concept of the sign (1975), form and content in the field of knowledge organization and representation are not equivalent to planes of expression and planes of content. These planes, which originate from the concepts of significance and meaning of Saussure (1970), are restricted to mental entities. Moreover, in the field of knowledge organization and representation, form and content materialize as, on the one hand, a physical entity and, on the other, a mental entity, i.e., content.

The first entity supposes written records on physical objects (monographic and iconographic materials) or those that can be materialized physically, as in the case of digital documents, which are also treated in terms of the physical characteristics of the cataloging process. The second entity suggests the existence of what we know generically as subjects or themes, i.e., what a document actually deals with in terms of semantic information, and what exists from at least two dominant philosophical perspectives. On the one hand, the subject is considered a reality, in fact an objective one, defined as realism in the concept of subject; on the other, it is a phenomenon present only in people's minds, like a kind of platonian idea. It is a concept defended by idealism. Realism and idealism are opposed when one seeks to establish a reasonable concept of subject in the area of knowledge organization and representation, and these same theoretical perspectives determine ways for extracting the subject of a document.

Extracting is a word that serves to describe the process of retrieving conceptual elements from texts to construct the subject; however, it does not reveal the essence of this process, which at other times is called indexing. Sometimes, indexing is one of the leading mediation processes between information available in documents and users. The indexing process is mistakenly regarded as technical or practical activity, although it should be understood as an intellectual process par excellence. The main issue surrounding indexing is how professional indexers extract, select, and translate the information available in the documents.

To do so, we must adopt an explanation of the indexing process that stresses the mental operations that generate signs for the representation of information. Thus, we need another way to understand the indexing process, a perspective that adds the production of signs as an activity that guides the indexing process. Accordingly, we assume that indexing is a representational activity, and that, at its core, it is a process that produces representations in general. This means that indexing is not only a representation action that establishes as goals words about concepts to control vocabulary more suitable for the information retrieval and organization system. Our argument is that subject indexing is a representational process and therefore requires theories of various disciplines, to wit: linguistics, semiology, philosophy of language, logic, and semiotics.

The efforts and contributions of linguistics to understanding this process are inestimable. In the last four decades, several linguistic theories such as textual linguistics, semantic linguistics, discourse analysis, etc., have competed to propose solutions for problems involved in the process of subject indexing, both manual and automated. Semiology, which studies the signs produced and communicated in the sociocultural sphere, also suggests ways of understanding nonverbal texts, such as codes or signaling systems. This has expanded the knowledge of indexing specialists on the concept of reading and text. Philosophy of language, in its most varied theories, roughly discusses the notion of referent and of how it determines the connection between concept and expression. Logic, in turn, seeks to suggest ways of understanding the structure of scientific discourse and the mechanisms of reasoning involved in the indexing process. Thus, subject indexing is the recipient of a series of disciplinary contributions, whose diagnosis would require extensive intellectual and practical work.

Nevertheless, we find that some disciplinary connections can still be clarified with respect to the relationship of subject indexing with semiotics, the science or study of signs in society and in nature. This definition of semiotics is not the same as the one espoused by Greimas, who proposed a textual semiotics, understanding it as a theory of the meaning of textual macrostructures. The results of Greimasian semiotics can be observed in the way that indexers understand a text and the strategies that are employed to grasp its principal meaning, evaluating the overall textual structure instead of an isolated word

or phrase. The research line of semiotics that is central to our discussion is that of Peircean philosophy, i.e., related to the philosophy and theory of signs propounded by Charles Sanders Peirce.

The semiotic perspective of Peirce offers essential contributions to the field of knowledge organization. Thellefsen (2002) uses concepts of Peircean semiotics to define and discuss what he called "semiotic organization of knowledge." The concepts of knowledge domain, significant effect (significance-effect), and fundamental sign are used by Thellefsen to propose an eight-step method for organizing knowledge in an area, which he illustrates with the field of occupational therapy in Denmark, whose sign is the fundamental concept of activity. Similarly, Thellefsen (2003) sought to examine and validate scientific terms that integrate knowledge domains based on the pragmatism of Peirce. According to the author, "Here, the pragmaticism of Peirce offers a way to analyze and develop the terminology of the knowledge domain" (Thellefsen 2003, 11).

Similarly, Thellefsen and Thellefsen (2004) proposed a theoretical framework for organizing knowledge based on what they called a sociopragmatic epistemology rooted in the pragmatic realism of Peirce. This sociopragmatic view explains the evolutionary dimension of knowledge, i.e., the analysis of the mechanisms of production, communication, and social interactions involved in scientific development. Thellefsen and Thellefsen (2004) allude to a method of organizing the knowledge of domains, arguing that, in knowledge, the concept or fundamental sign is the center of a radial structure composed of a series of related concepts. The birth of a sign is crucial to forming a habit of conduct, according to Peircean semiotics, and to identify it and estimate its value in a speech community is a required profile method called knowledge (knowledge profile). According to Thellefsen (2004), the profile of knowledge seeks to describe the epistemological foundations of any organization of knowledge. Returning to the premise of Peirce's pragmatism, Thellefsen (2004) argues that the meaning of a sign is in the test, i.e., the identification of the practical consequences of the sign.

These studies sought to apply Peirce's theories to understand knowledge in specific fields and to indicate ways to organize it. Another contribution to organizing knowledge was the work of Friedman and Thellefsen (2011), which sought to relate Dahlberg's theory of the concept and Peirce's semiotic theory to the organization of knowledge. Although both theories study representation, they do it differently: Semiotics are based in philosophy and logic to explain how signs evoke meanings, although generally Peirce's theory does not address knowledge organization systems, while Dahlberg's theory does not directly consider the representation of knowledge, but rather, how these

representations, in the form of concepts, can be related and classified. According to Friedman and Thellefsen (2011), the semiotic model is a general model of signs, while Dahlberg's conceptual model involves an instrumental perspective on how to make the best possible organization of a system of knowledge.

Despite the importance of these studies for the organization of knowledge in general, they do not specifically address the indexing process. An analysis of previous studies shows that we know which variables imply the semiotic organization of knowledge, but they tell us little about how to proceed semiotically to represent the essential information in a document. Thus, to gain a deeper view of subject indexing, we must consider the Peircean semiotic perspective. Peircean semiotics contains fundamental concepts for discovering aspects of the indexing process, including representation and classes of signs. However, we still know very little about its theoretical potential for subject indexing. The literature contains conceptual models of the indexing process (Mai 1997a, 1997b, 2001) suggesting concepts of semiotics, but the relationship of these models with the branches of semiotics requires clarification. We believe that the main difficulty in the proposals aimed at shedding light on the process of subject indexing based on Peircean semiotics stems from an incomplete reading of his semiotic system.

In general, semiotic proposals to explain the indexing process emphasize the concepts present in speculative grammar, the first branch of semiotics. However, a general reading of the Peircean semiotic system should integrate other branches of semiotics. For instance, semiotics is not a science that is limited to the elaboration of an extensive, detailed, and esoteric catalog of signs whose type-names are disconnected from the reality of people and society. Thus, the first step we must take is to gain an understanding of the branches of semiotics, after which we will focus on conceptual approximation strategies. Thereby we understand Peircean semiotics as a fundamental theoretical framework that explains the intricacies of the process of representation in general. So, in this paper, we adopt a theoretical approach to the issue of Peircean semiotics in indexing. What we want to do is to explain indexing from the inferential processes present in pure logic, the second branch of semiotics. For this, we must also identify the basic concepts of speculative grammar for indexing in order to achieve a more comprehensive view of the contribution of Peirce's general semiotics.

In view of these considerations, we intend to conceptualize indexing according to the knowledge organization and representation literature. Secondly, and aiming to raise the level of dialogue between the areas, we present subject indexing as an inferential process, as explained by the second branch of semiotics. Thirdly, we explore the

close links between speculative grammar, the first branch of semiotics, and pure or critical logic, which is the second branch.

The inference in Peircean semiotics is the guideline logic of indexing, and the types of inference are closely related to the phases of the indexing process. Thus, the inference or cognitive activity that leads to discovery of new information is a fundamental variable for the indexing process. In addition to the inference as a semiotic process, other factors contribute to the success of indexing. The collateral relationship to the experimental object accumulated by the index, as well as knowledge of an area's discursive universe of documents and users greatly influences the success of indexing. However, although essential to indexing semiotically, the concept of collateral experience will not be the object of analysis of this paper.

o

## 2.0 Indexing theory

Indexing, as a field of study and research, encompasses different views. However, from the perspective of the area of knowledge representation and organization represented by the systematization of Dahlberg (1993), it is seen as a process of knowledge organization and the result of its application as knowledge representation. This is undoubtedly the view of the theoretical area of information science that underpins indexing. This perspective is important, because it serves as reference for certain statements and viewpoints that will be made throughout this paper about indexing theory concepts and phases.

In referring to indexing and classification, Mai (2000, 1) wrote that "The representation of documents and the knowledge expressed by them is one of the central and unique areas of study within library and information science." This statement as well as Dahlberg's systematization (1993) lead us to the conclusion that indexing is a process of knowledge organization whose products accomplish knowledge representation. Moreover, indexing is a central and intermediate area that lies between an author and his readers by realizing the representation of knowledge that will be retrieved from a search interface.

If one looks back in time a little, one can see that the nomenclature of the investigative field in indexing is determined by the historical perspective of the specific purpose of creating indexes, an activity that the term "indexing" matches perfectly. Starting from the evolution that determined the importance of the context of documents in information retrieval, the area of indexing became part of the studies aimed at understanding the content of texts to be analyzed. These studies, however, are clearly embedded in theoretical currents, and, in the literature, it is easy to mistake the concept of indexing for its phases upon analyzing the contents of information

that have become diversified due to advances of information and communication technologies.

When one considers indexing as a field of investigation, one must understand that the activity of building indexes is necessary and that this product is the result of an indexing process whose procedure produces index entries (Cleveland and Cleveland 1983). Although from a similar point of view with which we do not agree entirely, the author Fidel (1994, 572) states that "indexes are needed to facilitate the retrieval of information." However, not only indexes facilitate information retrieval. When one goes to a search interface in a database on the web, be it with catalogs or an online journal portal, one does not need indexes. One uses a term or descriptor that is not part of an index. Stated simply, they are linked to a document or a subject field within a document or in metadata, without being associated in any way with an index. However, terms or descriptors are products of an indexing process.

Leiva (2008) considers that the diversity of these two products determines differences between what he calls *indexación* and *indización*. In Spanish, they differ even in spelling, but in Portuguese one has to use indexing for both meanings. The same holds true in English and French for the terms *indexing* and *indexation*. According to Leiva (2008, 66), in Spanish, the difference between *indexación* and *indización* is determined by the product resulting from the process. In *indexación*, different types of indexes are built (thematic, onomastic, of authors, etc.). Indexes "are lists that serve to connect a word or phrase taken from the text with the exact place it occupies in the text." Therefore, "to make an index is to extract words or phrases from its context, order them appropriately and indicate the place in the document where each word or phrase is located." In *indización*, "the resulting product (descriptor, subject heading or identifier) is not associated with an exact location in the document." Indexing, according to the meaning of *indización*, is performed by means of a set of operations that imply an effort to evaluate and condense both implicit and explicit content. One must consider that every document has implicit and explicit content and that the development of indexes does not require the inclusion of the implicit, only the explicit, since the location searches for exact reciprocity between the word of the index and the word of the text.

Fidel (1994) believed that indexing for the construction of indexes determines differences between indexes and subject indexes, and cites the example of the development of a name index, which differs from that of a subject index. For this reason, she considered that the most theoretical work on indexing involved subject indexing and that "indexing" was normally employed to signify "indexing by subject." Therefore, indexing aimed

at the construction of indexes is performed by extracting words that will be located through a remissive index. On the other hand, indexing aimed at representing the knowledge contained in documents for access and retrieval is accomplished through a process of analysis and representation.

The original purpose of index construction indexing coexists as one of its products for practical purposes, but the indexing process and its application have become a necessary investigative dimension due to the difference it causes in document representation for information retrieval. In this sense, one cannot but note that indexing, as a theoretical and methodological area of research, has been driven, to a large extent, by market needs for specialized information systems designed to meet the demands of their users searching for more accurate and specialized information without having to leave their research offices. The investigation of human and automated methods, followed by indexing assessment studies, has contributed to improve theoretical and methodological studies of indexing in the decades since the advent of scientific journals, and later, of specialized bibliographies.

From the standpoint of information systems, indexing is recognized as the most important part because it conditions the results of a search strategy for retrieving documents whose contents have a greater relevance in representing information needs. All the considerations about the performance of indexing and its importance for retrieval require an examination of how it is conducted in view of the existing analytical procedures. In the literature, most studies subdivide the indexing process into two phases: analysis, to determine the subject of the content of the document, and translation of this subject into the indexing language. According to the conception of Borko and Bernier (1978, 8), "indexing is the process of analyzing the information content of knowledge records and expressing this content in the language of the indexing system." However, in the literature of the area, one finds further developments of these two phases. Vickery (1980) believes that the indexing process comprises yet another stage, that of summarization, between the analytical and translation stages, dividing the analytical stage into two stages, analysis and summarization.

The normative literature published by the World Information System for Science and Technology (United Nations International Scientific Information System 1981) defines indexing based on the process and its purpose. Indexing, as a process, consists of describing and identifying a document with the help of representations of the concepts it contains, and, as for its purpose, it enables the search and retrieval of stored information. Although process and purpose are distinct points of view, they strongly influence each other, contributing jointly to

the existence of document-oriented approaches to indexing, for both user and domain.

In a study about document reading in indexing using a cognitive approach, Fujita (2007) presents the results of his observation of the reading of indexers during the first phase of the indexing process and finds, like Chu and O'Brien (1993), that subject analysis identifies and selects the main topics of the subject of a document, preceding the translation phase of these topics by the indexing language. In addition, he confirms that during the indexer's reading, subject analysis is subdivided into the stages of identification and selection of concepts (Fujita 2003), according to the Brazilian ABNT 12.676 standard, and that determining intrinsic and extrinsic aboutness is part of subject analysis (Fujita 2007).

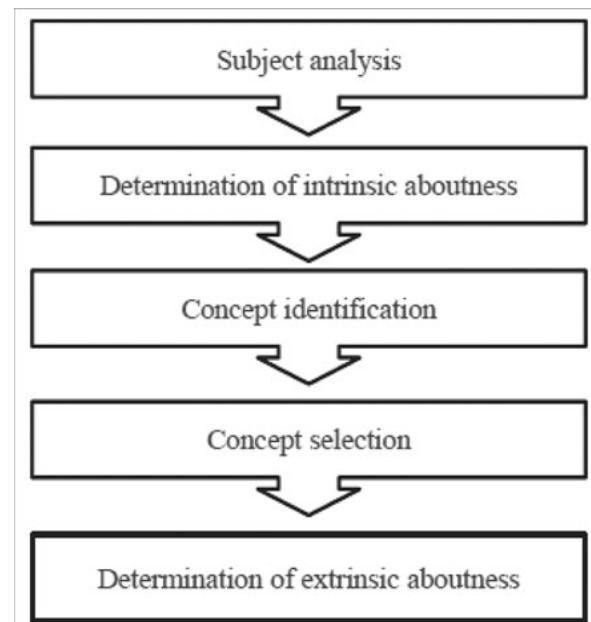


Figure 1. Stages of Subject Analysis (Fujita 2007, 44)

Thus, as Figure 1 illustrates, subject analysis consists of four phases, beginning with the determination of the intrinsic aboutness of the document and ending with the determination of its extrinsic aboutness. In addition, we believe that the determination of aboutness in subject analysis is influenced by the indexer's sociocognitive context, which, in turn, is linked to the interests and demands of information retrieval by the system and by the user (Fujita 2003). Another factor that influences this process is the semiotic competence of the indexer.

In 1974 *PRECIS, a Manual of Concept Analysis and Subject Indexing* appeared (Austin 1974). Albeit now outdated in terms of practical applicability, this manual offers an in-depth, interdisciplinary, and innovative theoretical and methodological foundation when considering the theoretical underpinnings of transformational generative and

case linguistics, as well as working within a cognitive and sociocognitive approach. The cognitive approach of PRECIS (Preserved Context Index System) involved the conceptual analysis of the indexing process performed by the human indexer to allocate function operators that automatically generated the subject added entries of the indexes of the British National Bibliography. The conceptual analysis applied to the document by the indexer proposed knowledge construction based on document context associated with constructive logic. The sociocognitive approach was present in the philosophy of preservation of the context of the content of the text and of the cognitive work of the indexer, since the conceptual analysis indicated the use of surface and deep structure operators which were equivalent, respectively to the explicit and implicit content. PRECIS went beyond the proposal of applying an indexing system in that it was innovative and was a relevant theoretical and methodological milestone in the area of indexing.

Other theoretical milestones emerged with innovative components in indexing:

Decade	Theoretical milestones
1970s	In the 1970s, addressing linguistic and cognitive aspects, authors such as Austin (1974), Jones (1976), Borko (1977), Cooper (1978), and Fugmann (1979) proposed theoretical foundations for later research.
1980s	In the 80s, Fugmann (1984) developed an indexing theory based on five axioms related to database search and retrieval strategies and started a lengthy discussion about an indexing policy that anticipated the results of information search and retrieval in order to "save" the user's time, while Soergel (1985) used the concept of document-oriented indexing that did not consider the context or user needs.
1990s	In the 90s, Fidel (1994) discussed indexing based on user-oriented conceptions and their demands for information retrieval, which differ from more strictly document-oriented conceptions. From the user-oriented perspective, studies that evaluate indexing use the characteristics or qualities of exhaustivity, specificity, accuracy and consistency, which can be measured by retrieval and, at the same time, can modify and enhance retrieval. Albrechtsen (1993) proposed the existence of three conceptions of subject analysis in the indexing process: simplistic, content-oriented, and demand-oriented. Farrow (1991), with his model of the cognitive process of document indexing, as well as Bertrand and Cellier (1995) emphasized the cognitive perspective of indexing. Also in the 90s, Lancaster (1993) and Fugmann (1993) combined theory and practice, and Frohmann (1990) proposed indexing rules without the perspective of what he called "mentalism."

Decade	Theoretical milestones
2000s	Starting in 2000, Mai (2001) introduced the perspective of semiotics to analyze the nature of the subject indexing process and the perspective of domain in domain-oriented indexing (Mai, 2004).

Table 1. Evolution of Indexing Theory

The content of this table summarizes the evolution of indexing theory without making a critical assessment. However, it should be pointed out, based on this summary, that there has been a theoretical evolution in indexing starting from approaches with an interdisciplinary focus, as in the cases of linguistics, cognitive psychology, and semiotics, because they contribute to the investigation of aspects relating to the nature of the indexing process, as well as of conceptions that focus on the purposes and elements of the indexing process, the user, the document, or the domain.

In a recent paper about the importance of knowledge theories for indexing and information retrieval (Hjørland 2011), one section is dedicated to theories of indexing starting from a systematization, which the author calls a classification of indexing approaches, based on four epistemological premises of indexing theories: rationalist; empiricist; historicist and hermeneutical; and pragmatic and critical. In the Rationalist approach, he cites Ranganathan's theory to explain that subjects are constructed logically starting from a set of categories, and thus considers all the other rules that are included in logical divisions. Therefore, he includes cognitive studies in the rationalist premise because he considers that, in the cognitive approach, "indexing rules are part of our cognitive structures and that, from this point of view (contrary to the historicist and pragmatic view), they are linked to universal biological structures" (Hjørland 2011, 74). He also says, however, that cognitive views of indexing are "theoretically unclear and problematic." Moreover, the author believes that literature of indexing itself is still unclear about the differences between cognitive approaches, user-oriented approaches, and other approaches, which therefore are difficult to test.

Although an assessment is controversial, because the development of the theory of indexing shows that the approaches are intrinsically linked to the nature of the indexing process and connected, indeed, with the mental processes still poorly understood by cognitive studies, linguistics, and semiotics. However, we believe that such approaches, although provisional and inconclusive about human thought, help to illuminate certain aspects and points that theoretically can not yet be explained. It should be kept in mind that the domain analysis suggested by Hjørland and Albrechtsen (1995) is an ap-

proach that needs clarification because it is still difficult to apply, although it has been helpful for information science with regard to knowledge organization from the perspective of domains and domain communities. It can also be inferred that interdisciplinary studies conducted by researchers in the area of indexing or information science depend, to a large extent, on a good understanding of the theories of these areas, whose applicability was not tested. Be that as it may, the theoretical breakthrough achieved by indexing can be credited to these approaches. However, there are recent contributions on exploratory levels in the field of knowledge representation and organization which suggest additional explanations about the semiotic nature of indexing.

### 3.0 Peircean semiotics and indexing

The science of signs, generally called semiotics (from the Greek *semeiotike*), studies the process of semiosis or the action of signs in the human and natural context. This means that, unlike semiology, which investigates signs in their social life, semiotics does not prioritize a sole significative manifestation of language. Before engaging in an analysis of the influence of Peircean semiotics on indexing, with respect to the semiotic process of indexing and its inferential character, we must circumscribe the theoretical dimension of the categories of the experience, concept and branches of semiotics, in addition to defining the entities related to the fundamental concept of signs.

Charles Peirce, the precursor of semiotics, contributed to several disciplines, including philosophy, logic, linguistics, biology, geodesy, and others. Although he did not become an acclaimed academician, he taught courses as a professor at Johns Hopkins University in Baltimore. Peirce's published work is estimated to be around 90,000 pages. The semiotic concepts presented below will be helpful to understand general Peircean semiotics in its three branches, as well as the ramifications of this for our analysis. Peirce acknowledges that the name semiotics (*semeiotike*) is attributed to John Locke (1632-1704), who also understood it as a discipline equivalent to logic. According to Liska (1996, 14) "semiotic, as a branch of philosophy, is a formal, normative science that is specifically concerned with the question of truth as it can be expressed and known through the medium of signs, and serves to establish leading principles for any other science which is concerned with sign in some capacity." In other words, semiotics seeks to study how things should be, not as they are currently in real life. The study of things as they are is a special objective of empirical sciences.

Semiotics is a broader logic in that it promotes the study of the various forms of thought representation. According to Peirce, semiotics—the "science of the gen-

eral necessary laws of Signs" (Peirce, 2000, 29, CP 2.93) or the formal doctrine of signs (Peirce, 2000, 29, CP 2.227)—comprises three branches known as speculative grammar, pure or critical logic, and speculative or methodeutic rhetoric. The overall understanding of these three branches prevents constant misinterpretations, such as the notion that semiotics is a solely classificatory science. In this paper, we will discuss speculative grammar and pure logic in greater detail, recognizing in advance that in order to go deeper into the issues of semiotics we must investigate the intricate relationship between indexing and speculative rhetoric. We recognize beforehand the importance of symbolic logic in Peirce's semiotic studies; however, we do not discuss this approach in our analysis.

According to Peirce (2000, 29, CP 2.93), speculative grammar "is the doctrine of the general conditions of symbols and other signs having the significant character;" pure logic "is the theory of the general conditions of the reference of Symbols and other Signs to their professed Objects, that is, it is the theory of the conditions of truth," and lastly, speculative or methodeutic rhetoric "is the doctrine of the general conditions of the reference of Symbols and other Signs to the Interpretants which they aim to determine." Although the three branches constitute the entirety of Peircean semiotics, few authors in the field of knowledge organization and representation have addressed anything beyond speculative grammar. And even at this level of semiotics, it is necessary to indicate the phenomenological categories.

The phenomenological categories gestated within the phenomenology of Peirce, the first of the philosophical sciences his view, seek to classify the broader aspects of phenomena (in Greek, *fanerons*). Such phenomena must be described, catalogued, and typified by phenomenology, which contains general categories of all the experiences: Firstness, secondness, and thirdness. In earlier moments of Peirce's thought, these categories were conceived as logical categories, namely: quality, relationship and representation. Perhaps the most general version of Peirce's understanding about these categories was expressed in a letter to Lady Welby dated October 12, 1904. In this letter, Peirce discussed the notion that all ideas could fit into three classes, and, although he did not like the notion, it ultimately dominated him completely. After this, he stated directly: "Firstness is the mode of being of that which is such as it is, positively and without reference to anything else" (Peirce, 1972, 136, CP 8.328). This logical mode is also present in the definition of the second and third categories: "Secondness is the mode of being of that which is such as it is, with respect to a second but regardless of any third. Thirdness is the mode of being of that which is such as it is, in bringing a second and third into relation to each other" (Peirce, 1972, 136, CP 8.328). Note that the catego-

ries of experience are applicable to any ideas, thoughts, feelings, concepts, etc. In the particular case of semiotics, they are essential to understand the phenomenological nature of the types and classes of signs.

In the category of firstness, we identify ideas referring to the first: original, independent, feeling, sense, quality, possibility, potentiality, diversity, multiplicity, uncertainty, and chaos. In this sense, none of these characters has a relation with a second. In a consciousness, the phenomenon of firstness arises simply because something is present, i.e., the presentness of the phenomenon or its ability to be present, without any relation to anything else, at the opportune moment of wonderment in a given instant of time. Hence, the feeling and sense of alterity are not possible in firstness.

The second category, secondness, is that which ontologically is such as it is according to a relationship that it establishes with another, and suggests what Peirce conceived of as experience. In every situation in the Peircean perspective, experience is felt alterity, in that it identifies the presence of another imposing on the first consciousness. It is what we find in the ideas of action and reaction, ego and non-ego, duality, relationship, experience, reality, factuality, and existence. Representative mediation is not present in secondness. The awareness of secondness precludes any possibility of a third or of a general law with an interpretive function connecting a first to a second.

Thirdness presupposes the notions of synthesis between a first and a second, mediation among phenomena, representation, law, regularity, generality, and thought. Whatever the complexity of a symbolic thought, it develops within thirdness. In other words, this means that all thought is triadic by nature, in that a quality could simply elicit wonderment without a real world to give it materiality, i.e., an individual occurrence in secondness. Such quality and factuality could not be represented in themselves at the instant in which they occur in direct perception, but would require a mediative phenomenon with representational power of the order of a third. Hence, thought as a third mediates the understanding of the phenomena of secondness and firstness.

In Peircean semiotics, the sign is the best example of a phenomenon in thirdness. For Peirce (1972, 143, CP 8.332) the sign is understood as “an object which is in relation to its object on the one hand and to an interpretant on the other, in such a way as to bring the interpretant into a relation to the object, corresponding to its own relation to the object.” Furthermore, according to Peirce (1972, 94, CP 2.228):

A sign, or *representamen*, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the

mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the *ground* of the representamen.

According to Peircean phenomenology, the representamen is a phenomenon of firstness, the object of secondness and the interpretant of thirdness. In addition, all the types and classes of signs in semiotics obey the categorical structure proposed by Peirce. However, the Peircean sign is not only one of the entities defined previously, but is the complete triadic relationship that exists among elements, such that there will be no flow in the interpretation without the participation of the sign, object and interpretant. This interpretive series, in Peirce's approach, is understood as the action of signs or the interpretation of a sign in the form of another sign (semiosis).

As has been observed, the foundation of the sign in a community of interpretants will be understood as the particular aspect or quality denoted by the sign. The object of the sign will be a second element, which differs from the sign because it is the element that instigates the representation of the sign. If there is a sign, it will be a substitute of an object. Peirce (1972, 96, CP 2.230) explains that a sign may contain more than one object; in this sense, this object may be complex and reveal a set of things or individuals that coexist as the object of the representation. In a letter to William James, dated March 14, 1909, Peirce (2000, 168, CP 8.314) argued that: “We must distinguish between the Immediate Object – i.e. the Object as represented in the sign—and the Real (no, because perhaps the Object is altogether fictive, I must choose a different term, therefore), say rather the Dynamic Object ...,” the latter, in turn, “sign *cannot* express, which it can only *indicate* and leave the interpreter to find out by *collateral experience*.” Understanding the immediate object, as it is present in the sign, is only the first stage to understanding the two facets of the object. The second facet deals with the dynamic object, which is the object as an existent reacting on other individuals and undergoing the wear of time. We can only approximate our immediate object to the dynamic when we value collateral experiences that support and adjust our representation of the object.

The interpretant, the effect of the sign or habit of action that suggests to the mind a mode of behaving to understand the object, is the third correlate of the sign. The nature of the interpretant is also signic, but unlike the first sign or representamen, this sign is more complex and developed. At some point, one might even feel the need to explain the interpretant with the word meaning

or sense; however, the allusion to structural linguistics may cause problems in understanding the dynamic and processual aspect of the sign in Peircean semiotics. A study of the influence of Saussure and Peirce on the field of knowledge organization and representation might well clear up these misapprehensions.

Like the object, the interpretant receives a division in order to deepen our understanding of it. However, this division is triadic and arranged in two series. In the first series, Peirce (2000, 168, CP 8.314) discusses the existence of three types: immediate interpretant, dynamic interpretant, and final interpretant. The immediate interpretant is the interpretive effect in the form of an idea, representation, or concept that is present in the first instant of the sign. The dynamic interpretant is the result, understanding, or concrete effect conveyed by the sign, i.e., the effective and real meaning of the sign. The final interpretant, in turn, is the interpretive effect the sign tends to produce at the end of the interpretive chain. In other words, the final interpretant can be understood as the total effect of a sign at the end of an interpretive series when the latter offers favorable conditions to reach the full meaning. However, there is a second series into which interpretants are divided: emotional interpretant, energetic interpretant, and logical interpretant. Following the categorical matrix, Silveira (1991, 49) states:

The first is a mere availability, feeling or affection, which, upon falling into a state of indifference, simply predisposes the behavior of the sign to move in search of the object. Having undone the psychological effects of the exposure, it becomes an interpretant of mere possibility. The second is an action that interprets the relationship of the sign to the object: the response to a command and the entire reaction are characteristic examples and forms of genuine approximation of energetic interpretants. Lastly, the third is a representation that interprets the relation of the *representamen* and the object.

It is possible that this second series of interpretants may contribute to an understanding of the cognitive dimension of the dynamic interpretant, as the sign's real effects on an individual mind. At this point, we can state that the relationship of the sign with itself, of the sign with the dynamic object, and of the sign with the interpretant, comprise the three principal trichotomies studied by Peirce in speculative grammar. In truth, Peirce's signs have ten trichotomies, although three of them are more fully developed. The most important text about this is an excerpt from approximately 1897, entitled "Division of Signs" in which Peirce (2000, 51-53, CP 2.247-253) pre-

sents the three trichotomies. In the first trichotomy, the relation of the sign with itself, the types of sign are: qualisign (a quality that works as a sign), sinsign (an existing thing or event that is a sign), and legisign (a law that is a sign). The second trichotomy, the sign in relation to the dynamic object, comprises the following types: icon (a sign resembling the object), index (a sign that refers to the object because it is affected by it), and symbol (a sign that refers to the object due to its ability to associate general ideas addressed to the object). The third trichotomy, the sign in relation to the interpretant, comprises the following types: rheme (a sign of qualitative possibility), dicisign (a sign of real existence to its interpretant), and argument (a sign of law to its interpretant).

The combination of these nine types of sign engenders several possible classes of signs, but only ten classes are considered valid. In the same text, "Division of Signs," Peirce (2000, 55-57, CP 2.254-264) presents the ten classes of signs—to wit: rhematic iconic qualisign (example: a sense of redness), rhematic iconic sinsign (an individual diagram), rhematic indexical sinsign (a spontaneous shout), dicent indexical sinsign (a weather vane), rhematic iconic legisign (a general diagram), rhematic indexical legisign (a demonstrative pronoun), dicent indexical legisign (the cry of a peddler), rhematic symbolic legisign (a common noun), dicent symbolic legisign (a proposition), and argumentative symbolic legisign (a syllogism). The classes of signs that are mentioned when the subject is Peircean semiotics only make sense when one states that, through the study of the last class—argument—one enters the field of the second branch of Peircean semiotics: pure logic. So having briefly discussed the main aspects of speculative grammar, we can now proceed by defining the types of inference present in Peirce's logic.

First, however, we must explain what inferences are. We know that we come to conclusions from inferences. The inference is that we have the intellectual apparatus to reach a new idea. The arguments are the linguistic expression that demonstrates the reasons that led us to conclude. For Salmon (1969, 22), an argument is a linguistic entity and is not an inference. An argument is a type of inference that is based on evidence and not on opinions. But not all the inferences can be linguistically expressed in the form of arguments. Accordingly, for Copi and Cohen (2002, 6), inference "refers to the process by which one proposition is arrived at and affirmed on the basis of one or more other propositions accepted as the starting point of the process," while the argument is considered "any group of propositions of which one is claimed to follow from the others, which are regarded as providing support or grounds for the truth of that one." Logic studies the argument already established and the

propositions already expressed, and is not subject to analyze the inference that, technically, is a process to formalize the previous argument.

Handbooks of logic commonly treat argument in the analysis of hypothetical deduction and analogical argument in a chapter about induction. Although there are distinctive aspects, what interests us in this paper is how the three types of inferences or arguments addressed by Peirce can illuminate our understanding of the indexing process. When it comes to this matter, Peirce uses the terms "inference" and "argument." Despite the distinction, he treats "inference" and "argument" as equivalent terms or synonyms, but we are aware of the distinctions they have.

There are other types of inference beyond the inductive and deductive arguments, like the analogy. Peirce emphasized the abductive, inductive, and deductive inferences, which are fundamental to the analysis of arguments. This analysis will highlight the types of inference studied by Peirce, recognizing that other types may guide further studies on the relation of inference to the indexing process. These types of inference or argument are examined in Peircean semiotics in the branch known as pure logic.

Pure logic studies the types of inference known as abduction, deduction, and induction. In 1903, Peirce (1980, 46, CP 5.171) defined the type of reasoning that engenders hypotheses as follows: "Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value, and deduction evolves the necessary consequences of a pure hypothesis." Peirce was aware that the surprising ideas of human ingenuity are not the products of deliberate reasoning, but of our ability to create and accept suppositions as possible representations of phenomena. Peirce anticipated our current understanding that scientific research originates with a hypothesis.

Deduction, as an analytical inference, presupposes the existence of a law which should be applied to all cases subjected to a norm. It is impossible for original knowledge to be manifested, since knowledge of the case lies in the law. The inventive power to suggest original ideas is extremely reduced in deduction. It is almost an automatic manifestation of human reasoning. In 1896, Peirce (2000, 5, CP 1.66) explained the following:

*Deduction* is that mode of reasoning which examines the state of things asserted proposed in the premises, forms a diagram of that state of things, perceives in the parts of that diagram, relations not explicitly mentioned in the premises, satisfies itself by mental experiments upon the diagram that these

relations would always subsist, or at least would do so in a certain proportion of cases, and concludes their necessary, or probable truth.

For Peirce, the purpose of deduction is to start from a presumed law and reach the ultimate consequences of this law in the application of the case. In this sense, "In deduction, or necessary reasoning, we set out from a hypothetical state of things which we define in certain abstracted respects .... Our inference is valid if and only if there really is such a relation between the state of things supposed in the premises and the state of things stated in the conclusion," (Peirce 2000, 215, CP 5.161). This hypothetical state of things reveals the integration of abduction in deduction. However, we must also mention the continuity of the inferential process, which materializes with the approximation to reality in experimental trials, i.e., the direct contact of the mind with reality.

Induction, on the other hand, does not stem from a hypothetical law accepted by the mind to know a case. On the contrary, it is the inferential ability to admit a rule starting from the case and the result (Peirce 1972, 149, CP 2.622). In other words, "Induction is where we generalize from a number of cases of which something is true, and infer that the same thing is true of a whole class. Or, where we find a certain thing to be true of a certain proportion of cases and infer that it is true of the same proportion of the whole class." (Peirce 1972, 150, CP 2.624). We find that induction, like an inferential process, contains a fundamental phase of the stage of scientific investigation, i.e., experimentation and testing with the individuals of reality. This meaning is present in the Lectures on Pragmatism given in 1903 at Harvard University, for induction, now as a stage of the process of investigation, requires a return to experience in the sense of adjusting the representation. Peirce (1980, 46, CP 5.170) wrote:

Induction consists in starting from a theory, deducing from it predictions of phenomena, and observing those phenomena in order to see *how nearly* they agree with the theory. The justification for believing that an experiential theory which has been subjected to a number of experimental tests will be in the near future sustained about as well by further such tests as it has hitherto been, is that by steadily pursuing that method we must in the long run find out how the matter really stands.

The purpose of an experimental procedure is to demonstrate the vitality of the germinal hypothesis, which has been generalized in the rational ambit through deduction. Indeed, speculative grammar and pure logic account for the typological characteristics of signs and their inferen-

tial workings. However, without knowledge of the interpretive potential of signs and of how they shift semiotically to a new stage of translation to generate new knowledge, these semiotic disciplines do not achieve Peirce's purpose, which is that of being a complete theory of the formal and normative conditions of all signs. The authority of a speculative rhetoric, which encompasses the investigation of belief fixation, is essential to understand Peirce's pragmatism, a method and, at the same time, a theory of meaning. Nevertheless, the mention of the first two branches of the doctrine of signs suffices for us to begin our analysis of how Peircean semiotics is related to subject indexing.

### 3.1 Speculative grammar in the semiotic indexing model

Subject indexing is a determining process in the organization of documents in libraries based on terms used in searches by subject and may also be presented in alphabetical subject indexes at the end of books. It is a fundamental process directly related to the individuality of each indexer, since each professional possesses previous knowledge of the world and references that differentiate him from others; these are factors that will be reflected in the end products of indexing. In addition, the process involves the creation of signs, the complexity of which may vary. Simpler types of signs lead to cognitions that are more prone to misinterpretation, which is the case of indexical signs. In Peircean semiotics, the disciplinary branch that deals with the diagnosis of general types of signs is speculative grammar. The taxonomy of signs provides specialists in the subject with core concepts to understand the semiotic nature of the indexing process. The phases of subject indexing can be seen as the types of signs of the trichotomy described by Peirce.

A contribution in this sense was offered by Mai in his studies on the relation of subject indexing with the classes of signs in Peircean semiotics. For Mai (2000), the subject indexing process is divided into four elements and three steps. He explains that little is known about subject indexing, and that it should be considered an interpretive

process. Mai (1997a) believes it is possible to understand how each of the elements and steps of the indexing process work based on some aspects of Peircean semiotics. These elements and steps of indexing are presented by Mai (1997a, 2001, 1997b) in an indexing model adapted from Francis Miksa (1983). Miksa's indexing model consists of three squares arranged sequentially, the last one ending in a point, representing each of the elements of the indexing process, namely: the document under analysis (first square), the subject (second square), the description of the subject (third square), and the subject entry (the point at the end). The three steps performed between the elements are presented between the squares: the document analysis process (between the document square and the subject square), the subject description process (between the subject square and the subject description square), and the subject analysis process (between the subject description square and the subject entry point). Shown below in Figure 2 is the model of the indexing process used by Mai.

Mai (1997a) adapted Miksa's indexing model simply by progressively reducing the size of the squares, from document (largest square) to subject entry (smallest square), indicating that the variety of referents is larger at the base and that they are filtered by the indexer as the steps are performed. This model containing the indexing elements and steps is used later as the basis for a semiotic indexing model. Based on Miksa's indexing model and some conceptual aspects of Peircean semiotics, such as semiosis or sign-activity, Mai believes one can explain how the interpretive process of the professional who performs indexing takes place. Mai (2001, 1997a, 1997b) uses the concept of semiosis during the indexing process based on Peirce. Each element of the indexing process is considered a sign, and each step works as an interpretive action connecting the signs in a sequential process (Mai 2001, 603). Mai's model is depicted in Figure 3 below.

In the semiotic indexing model (Mai 2001) presented in Figure 3, the process is initiated by a sign, the document (represented by the triangle M). Upon performing an action of interpretation in the first step of the process

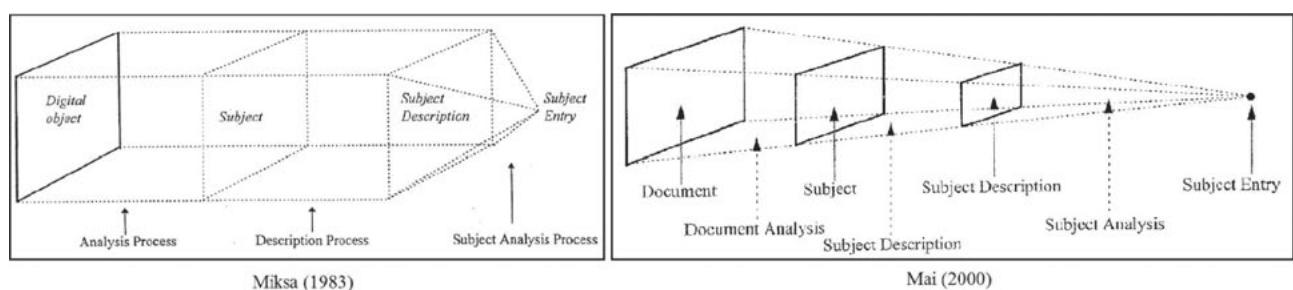


Figure 2. Indexing Models of Miksa and Mai (Mai 1997a, 2000)

(document analysis), the indexer will develop a new sign, the subject (triangle N). In the following step, the subject description process, when the subject that was in the indexer's mind becomes something more palpable, a new interpretive action occurs, resulting in a new sign, the subject description (triangle O). Upon performing the subject analysis, the subject description is transformed into an indexing language, giving rise to another sign, the subject entry (triangle P) (Mai 2001, 603).

Semiosis in the indexing process is thus understood as a sequence of interpretive sign-generating actions resulting from its steps and elements. At the level of possibility, we also consider the continuity of semiosis from the contact of users with the end products of indexing, how it interacts with the system and how the terms selected by the indexer interfere in his search, and hence, in the production of new signs. According to Mai (2001), the document can be classified as an argument, since it represents a group of ideas and knowledge. The subject, in turn, will be a mental sign of the dicent symbol type (Mai

2001, 615). The subject description acts as a dicent indexical legisign and, depending on the level of interpretation and on the interpretant, the entity called subject entry will be a rhematic indexical legisign (Mai 1997a, 62):

The subject entry is categorized as a *rhematic indexical legisign*. The subject points out its subject, and is therefore an index, it requires a minimum of interpretation, and is therefore a rheme. The subject entry therefore give very little information about its object but merely points it out.

At this point, we find a convergence between Peircean semiotic and subject indexing studies, insofar as the indexical and potential nature of the subject entry is highlighted by both approaches. However, Peircean semiotics suggests other explanations to understand indexicality, which may suggest a referenced object without presenting it readily, leaving it to the interpretant to produce the necessary contiguity between sign and object. Perhaps,

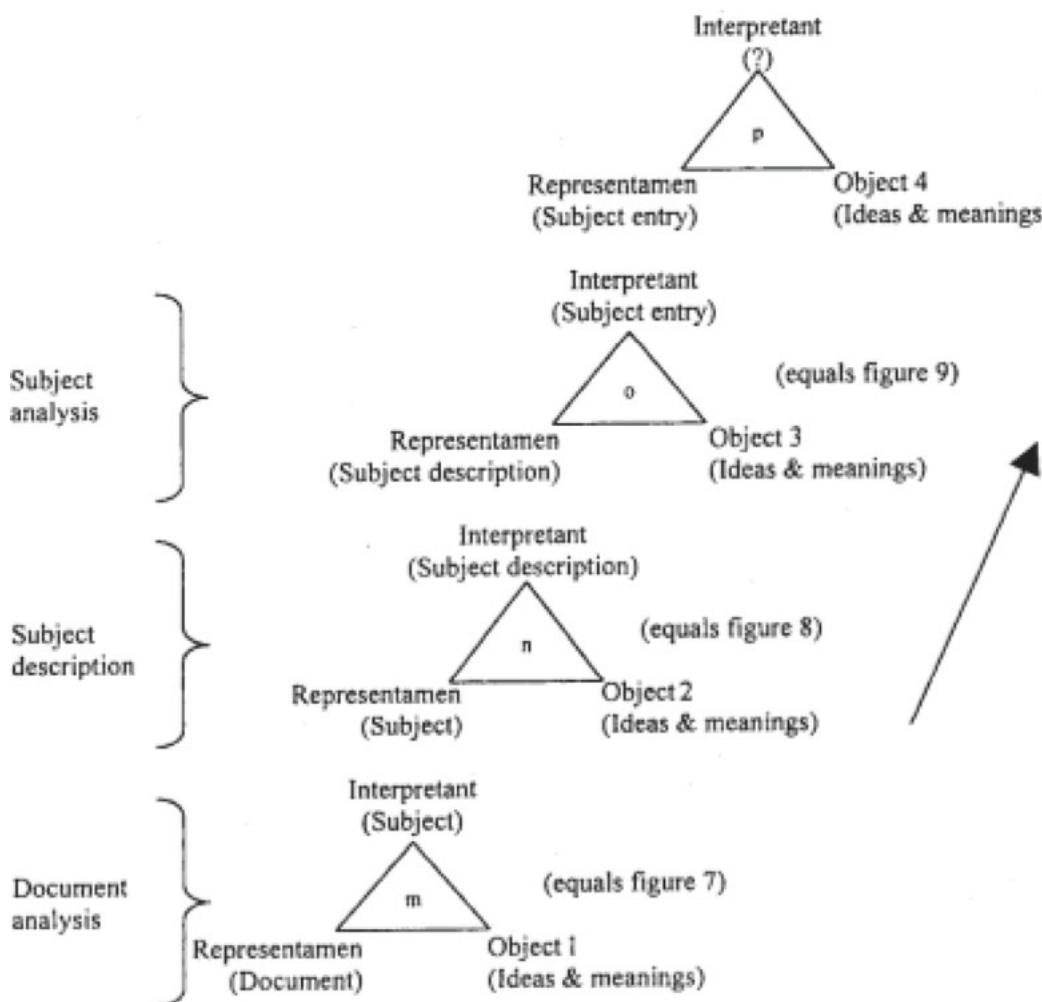


Figure 3. Semiotic Indexing Model (Mai 2001).

the ultimate goal of the indexing process lies in indexicality, for it is through this type of sign that we bring together signic realities by providing opportunities for connections between signs (subject entries) and objects (ideas and documents).

It should also be noted that the semiotic indexing model illustrated in Figure 3 is only part of a larger process. The triangles M, N, O, and P are used to represent the indexing elements presented by Mai to explain indexing, but when the steps of the process are put into practice, the indexer will perform them simultaneously without thinking about their limits (beginning and end of each step). This sequence of steps and elements of indexing helps shed light on how the entire process works, contributing directly to the practice itself, since the area of the thematic treatment of information lacks formal manuals with guidelines on how to perform subject indexing according to its signic processes. With the semiotic indexing model, Mai means that indexing is a process that depends not only on the technique itself, but on how the indexer thinks in signs and acts to perform certain steps, for the theory of indexing is fundamental to knowing how semiosis acts and classes of signs are generated in the indexing process. We argue that Mai's approach consisting of four elements and three processes is not enough to understand indexing from the viewpoint of Peircean semiotics, as it is necessary to consider the types of inference or argument. At first, we hold that the indexing process is inferential, i.e., the types of inference are the rationales of the process even before its linguistic manifestation.

### 3.2 Pure logic and indexing: indexing as an inferential process

In Mai's analysis (1997a, 1997b, 2000, 2001), we find the products generated in the indexing process (document, subject, subject description, and subject entry), which are considered the formal elements obtained in the process. However, the nature of this process is inferential. Pursuant to recent research in this field, particularly Almeida's study (2011), we are now able to gain a semiotic understanding of the indexing process and answer a subjacent question arising from the works of Mai. To wit—if the mental entities generated by the indexing process can be created without the help of some of Peirce's classes of signs, could the phases of the subject indexing process have the same fate? We believe that the answer to this question is positive, since classes of signs can only be generated by a combination of feasible signs through types of inference, as defined by Peirce. In this sense, to broaden the statements of Almeida (2011), we argue that indexing is not only a representative and interpretive activity but also an inferential process, and, as an enquiring action, it must approach the most demanding modes of

scientific reasoning. By inference, we understand the inclination of reasoning to generate new information from previously known elements or premises.

Pure logic seeks to address the relation of signs to objects, and to understand how knowledge can be created by ideal types of inference, namely abduction, deduction, and induction. Indexers, like scientists and ordinary people, are predisposed to reach conclusions through inference. This natural and deliberately improved process is not the privilege of the scientist in a laboratory or in a field of research, but is an existential condition, depending on semiotically available devices to know reality. The production of representations of the world lies at the core of the creative knowledge of indexers and scientists. The three types of inference discussed by Peirce in the second branch of semiotics may receive various sequential compositions, i.e., the direction of the inference is not prescriptive, from abduction to induction, but depends on the perspective we consider in Peirce's thinking in this respect. In this sense, if Peirce's logical categories of quality, relation, and representation are involved in our analysis, the reasonable inferential sequence would be abduction, induction, and deduction, since deductive inference would represent the highest degree of generality. If one were to emphasize the categories of experience firstness, secondness and thirdness, the appropriate metaphysical sequence would be abduction, deduction, and induction, because the latter inference manifests in the continuity of the process of representation. Given the breadth and emphasis on continuity of the process suggested by the latter categorical point of view, we decided to organize our explanation of the indexing process in the inferential sequence: abduction, deduction, and induction, in order to reflect phenomenologically firstness, secondness, and thirdness.

The problem of classes of signs put forward by Mai can be understood perfectly if one associates the inferential platform that supports the production sign modalities. In this sense, abduction is the creative stage of indexing, deduction is the generalizing stage, and induction is the phase of testing and continuity of subject representation. Thus, the abduction is the type of inference that creates suggestions of subject in the form of hypotheses. The Peircean perspective argues categorically that we only achieve advances in knowledge due to our ability to formulate hypotheses, and these are not as difficult to produce as scientists believe. We do not control the issuance of original ideas; it is only after they have been put forth that we work on them in a self-controlled and deliberate way. The absence of hypothetical inferences or their complete control on the part of reasoning comprises the target of this process, which, in the Peircean point of view, supports the variety and diversity of ideas in the life of thought.

In evolutionary terms, there will always be documents whose subjects are new, and there will be no representative terms to substitute them in the indexing languages. These circumstances will require that the indexer generate hypotheses through abductive inference, the only way to ensure the feedback of systems in the absence of available subject representations. Notwithstanding the beneficial strategies of terminological control and professional qualifications, we must realize that the fundamental semiotic characteristic of indexing lies in the ability of indexers to generate suppositions or creative ideas as the focal points of subject representation. In this sense, the indexing process resembles the intellectual activity of investigation, in that the creative power of the scientist is also required of an indexer in the context of the identification of potential subjects of a document. In Peirce's approach, abduction stems from a process of perception, for perceptual judgments, irrecusable premises, are like the initial premises of our reasoning, and "the first judgment of a person as to what is before his senses, bears no more resemblance to the percept" (Peirce 1980, 34, CP 5.115).

Considering that abduction is a fundamental inference for the indexer's activity and that this process begins with perceptual judgments, we should value perception as a determining instance to understand the indexing process. In this regard, we can state that the production of perceptual judgment lies outside the possibility of rational control (Peirce 1980, 37, CP 5.115; 43, CP 5.157; 51, CP 5.181). Perceptual judgment is an uncontrollable element, and it arises in a relatively chaotic manner in the mind. It is reasonable to assume that, due to the uncontrollable nature of perceptual judgment, one cannot refuse to consider it when it emerges in the mind. In other words, it would be immediate perception, which we do not have the power to discern, but can only accept due to its power to produce constraints on the senses. In short, the indexer is an unpredictable condition with regard to the subject and enters into a collateral relationship with the document in the hope of being able to suggest creative suppositions, after which he will proceed to evaluate the suitability of the mental product thus generated.

The deduction, in turn, is the second inference that supports the indexing process. Being the type of inference that intends to extract consequences and relate them to a case, it assumes a preestablished law. As in analytical or decompositional reasoning, the intention of the deduction is to refer to a quality in the conclusion, but this attribute is subjacent in the premises. There is no innovation in deduction; it is a simple operational application to clarify the intellectual procedure. In this study, without discussing the nature of the interpretant in deduction, we will simply state that it has an iconic and diagrammatic character, since it establishes relationships found in one part of thought and

transfers them to another, as yet unknown part. Deduction, relying on its general and relatively true nature, compares the hypothetical suggestion in indexing to other possible subjects, reaching an understanding of the relationship among subordinate subjects. In other words, if one assumes that the document deals with paleontology, one should deduce that the document also deals with fossils (symbolically: S is P/ P is D/ S is D). If the conclusion accepted deductively proves to be credible at the end of the subject retrieval process, the pertinence of the original hypothesis is justified. Therefore, we believe that the indexer knows the subject of a document not only because he decomposes the textual structure into central contents, but because this knowledge derives from the indexer's semiotic competence in inferring it starting from hypotheses. Thus, we find that deduction is the platform that underpins rational operations with hypotheses, giving them a more coherent and plausible form.

The third type of inference studied within pure logic in semiotics is induction, which plays a central role in the particular context of indexing. It is through induction, in the Peircean concept of the term, that we ensure the quality of indexing. In the inductive procedure, we make use of schemes or models of the world of knowledge in a given field, represented by bibliographic terminologies and classification systems. The indexer's daily activities of experimentation or testing require inductive inference, which involves checking and comparisons against reality in the context of the information system. The adequate subject representation of a document requires observing the quality of the hypotheses and of the supporting instruments available to the indexer. Thus, the terms used to explain the indexing process, such as translation and representation, hinder the understanding of the act of indexing as a semiotic activity. In other words, previews are decisive in indexing and, as a semiotic and creative act, indexing must involve much more than the indexer's technical training, experience, and knowledge of the document's language. Indexing is a process open to creativity, a necessary condition for reaching a reasonable offer on the subject.

The products generated in the indexing phases, some of them classified according to classes of signs, are only concretized in inferential processes. Inference is the mechanism that works as a setting or backdrop of the indexing process so that it is possible to question the semiotic nature of the phases of the process, since subject analysis—the last phase according to Mai—is not a procedure that segments the content, but a phase that tries, tests, and compares it with the choices of representation suggested hypothetically. If we think of induction as a common procedure that conditions the indexing process, we will transfer the indexer's activity to the space that it should in fact occupy, i.e., that of professional actions

aimed at solving problems of representation, which, like those of a scientist, follow an inferential sequence.

In short, by abduction, we mean the process of reading (perception) and the creation of hypothetical suggestions of subject representation. Deduction, in our view, deals with the selection of representative terms according to the intellectual generalization of the consequences. And lastly, we understand that induction involves the functions of testing and comparing potential subjects to the indexing language, the user's language, and the language of the information system, aiming to continually evaluate and improve subject representations. These concepts are illustrated in Table 2, which shows how the types of inference in the indexing process are integrated.

#### 4.0 Final Remarks

We understand the Peircean semiotic approach to the process of indexing the subject as a field of research for knowledge organization and representation. We pointed out that the semiotics of Charles Peirce cannot be examined only in its taxonomic bias recommended by speculative grammar. Semiotics should be valued by the general theory of signs to understand the primary intention of Peirce. Finally, it is essential to understand Peircean semiotics as a logical array of the indexing process. Indexing is an intellectual activity, thus inferential.

In addition, we must emphasize another semiotic contribution to the indexing process: speculative rhetoric. In other words, subject indexing presupposes its own representational dynamics, which must suggest the creation of interpretants that are mutually related in a continuous process of signification. Any subjects, as mental objects, would be correlated by the change in signs, converging to the object. In this sense, we should also note the role of the referent as a condition of subject representation. When a referent is known to the indexer, its interpretation is obviously somewhat more effective, as is the pro-

duction of the subject entries. This can be illustrated by the discrepancy between the indexing performed by an expert and that produced by a non-specialized technician. One may speculate that differences in the interpretation of indexers correlate to their knowledge of and collateral contact with the referents dealt with in documents.

To continue this study would inevitably involve a review of the impact of speculative rhetoric on subject indexing, since the indexer is not outside of the phenomenal world when he engages in indexing. Therefore, the next step in the research will be to examine the connection of speculative rhetoric with indexing, i.e., how the process of semiosis, communication, and continuity of meaning occurs in the form of a chain of semiotic phenomena. It is precisely the indexer's interaction with the world, with objects, and with signs that determines his actions of representation. The understanding of these actions has an inevitable consequence on the indexer, who is guided by his perception of reality and of the subjects of documents, which also determine the success of his work and should be the object of analysis.

We know that the subject of the document often becomes an objective entity (a reality independent of the indexer), and, at other times, a subjective entity (a reality dependent of the indexer), and that these two entities do not always converge to the same point. Accordingly, we suggest, based on the proposal of Peirce's semiotics, a pragmatist reading of the subject, bringing together the intricate and complex encounter between the inner reality and external reality to the subject of the document. Thus, the subject of the document would not be an objective entity outside the indexer or a subjective entity dependent on the indexer's awareness, but a pragmatic reading combining elements of both in order to semiotically serve a reality of representation ultimately aimed at guiding people in information systems. A pragmatic view of the document subject overcomes the illusory dichotomy of objectivist and idealistic

INFERENCES	ABDUCTIVE	DEDUCTIVE	INDUCTIVE
DEFINITION	Creation of hypotheses or explanatory suggestions about the document's contents	Analysis of the consequences of attributing a subject to the document	Testing and experimentation with the presumed language of the system and user
CATEGORIES	Firstness	Secondness	Thirdness
NATURE	Possibility	Generality	Continuity
STAGES	Creation	Analysis	Comparison

Table 2. Inferential Indexing Process

Another logical development of this analysis of Peircean semiotics in subject indexing will be to analyze the usefulness of his version of pragmatism as a method for discovering the process of producing meanings associated with the set of habits involved in conception. In the words of Peirce, "To develop its meaning, we have, therefore, simply to determine what habits it produces, for what a thing means is simply what habits it involves" (Peirce 1972, 58, CP 5.400). This epistemological proposal for indexing is not the same as Peirce's, which emphasizes the practical effects of individual action to obtain practical benefit. Peirce's pragmatism concentrates on the rules and intellectual content that govern human action. Finally, the work of the indexer does not produce a mere reflection of what is already contained in documents, but involves the authoritative and challenging action of discovering, through the inferential matrix, to determine the subject and the subject entry that are most suitable for the information system.

## References

Albrechtsen, Hanne. 1993. Subject analysis and indexing: From automated indexing to domain analysis. *The indexer* 18: 219-24.

Almeida, Carlos Cândido de. 2011. Sobre o pensamento de Peirce e a organização da informação e do conhecimento. *Liinc em revista* 7: 104-20. Available: <http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/405/265>

Austin, Derek William. 1974. *PRECIS: A manual of concept analysis and subject indexing*. London: Council of the British National Bibliography.

Bertrand, Annick and Cellier, Jean-Marie. 1995. Psychological approach to indexing: Effects of the operator's expertise upon indexing behavior. *Journal of information science* 21: 459-72.

Borko, Harold. 1977. Toward a theory of indexing. *Information processing and management* 13: 355-65.

Borko, Harold and Bernier, Charles L. 1978. *Indexing concepts and methods*. New York: Academic Press.

Chu, Clara M. and O'Brien, Ann. 1993. Subject analysis: The critical first stage in indexing. *Journal of information science* 19: 439-54.

Cleveland, Donald B. and Cleveland, Ana D. 1983. *Introduction to indexing and abstracting*. Littleton, CO: Libraries Unlimited.

Cooper, William S. 1978. Indexing documents by Gedanken experimentation. *Journal of the American Society for Information Science* 29: 107-19.

Copi, Irving M. and Cohen, Carl. 2002. *Introduction to logic*. 11th ed. Upper Saddle River: Pearson Education.

Dahlberg, Ingetraut. 1993. Knowledge organization: Its scope and possibilities. *Knowledge organization* 20: 211-22.

Farrow, John F. 1991. A cognitive process model of document indexing. *Journal of documentation* 47: 149-66.

Fidel, Raya. 1994. User-centered indexing. *Journal of the American Society for Information Science* 45: 572-6.

Friedman, Allon and Thellefsen, Martin. 2011. Concept theory and semiotics in knowledge organization. *Journal of documentation* 67: 644-74.

Frohmann, Bernd. 1990. Rules of indexing: A critique of mentalism in information retrieval theory. *Journal of documentation* 46: 81-101.

Fugmann, Robert. 1979. Toward a theory of information supply and indexing. *International classification* 6: 3-15.

Fugmann, Robert. 1984. The five-axiom theory of indexing of indexing and information supply. *Journal of the American Society for Information Science* 36: 116-29.

Fugmann, Robert. 1993. *Subject analysis and indexing: Theoretical foundation and practical advice*. Frankfurt: Verlag.

Fujita, Mariângela Spotti Lopes. 2003. A identificação de conceitos no processo de análise de assunto para indexação. *Revista digital de biblioteconomia e ciência da informação* 1 no.1: 60-90.

Fujita, Mariângela Spotti Lopes. 2007. *Leitura em análise documentária: Uma contribuição à formação de indexadores* (Relatório final de pesquisa, Bolsa PQ/CNPq). Marília: Faculdade de Filosofia e Ciências da UNESP.

Gil Leiva, Isidoro. 2008. *Manual de indización: Teoría y práctica*. Gijón: Trea.

Hjelmslev, Louis. 1975. *Prolegómenos a uma teoria da linguagem*. São Paulo: Abril Cultural.

Hjørland, Birger. 2011. The importance of theories of knowledge: Indexing and information retrieval as an example. *Journal of the American Society for Information Science and Technology* 62: 72-7.

Hjørland, Birger and Albrechtsen, Hanne. 1995. Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science* 46: 400-25.

Jakobson, Roman. 1999. *Linguística e comunicação*. São Paulo: Cultrix.

Lancaster, Frederick Wilfrid. 1993. *Indexação e resumos: Teoria e prática*. Brasília: Biquet de Lemos/Livros.

Liska, James Jakób. 1996. *A general introduction to the semiotics of Charles Sanders Peirce*. Bloomington: Indiana University Press.

Mai, Jens-Erik. 1997a. The concept of subject in a semiotic light. In Schwartz, Candy and Rorvig, Mark, eds., *Digital collections, implications for users, funders, developers and maintainers: Proceedings of the 60<sup>th</sup> ASIS Annual Meeting*. Medford, NJ: Information Today, pp. 54-64.

Mai, Jens-Erik. 1997b. The concept of subject: on problems in indexing. In I. C. McIlwaine ed., *Knowledge or*

ganization for information retrieval: 6<sup>th</sup> International Study Conference on Classification Research. The Hague: FID, pp. 60-67.

Mai, Jens-Erik. 2000. *The subject indexing process: An investigation of problems in knowledge representation*. Ph.D. dissertation. Austin: The University of Texas at Austin. Available [http://jenserikmai.info/Papers/2000\\_PhD\\_diss.pdf](http://jenserikmai.info/Papers/2000_PhD_diss.pdf).

Mai, Jens-Erik. 2001. Semiotics and indexing: An analysis of the subject indexing process. *Journal of documentation* 57: 591-622.

Mai, Jens-Erik. 2004. The role of domains, documents, and decisions in indexing. In McIlwaine, Ia C. ed., *Knowledge organization and the global information society: Proceedings of the Eighth International ISKO Conference 13–16 July 2004 London, UK*. Advances in knowledge organization 9. Würzburg: Ergon Verlag, pp. 207-13.

Miksa, Francis. 1983. *The subject in the dictionary catalog from Cutter to the present*. Chicago: American Library Association.

Peirce, Charles Sanders. 1972. *Semiótica e filosofia*. São Paulo: Cultrix.

Peirce, Charles Sanders. 1980. *Escritos coligidos*. São Paulo: Abril Cultural.

Peirce, Charles Sanders. 2000. *Semiótica*. São Paulo: Perspectiva.

Salmon, Wesley C. 1969. *Lógica*. Rio de Janeiro: Zahar Editores.

Saussure, Ferdinand de. 1970. *Curso de lingüística geral*. São Paulo: Cultrix.

Silveira, Lauro Frederico Barbosa da. 1991. Na origem está o signo. *Trans/form/ação* 14: 45-52.

Soergel, Dagobert. 1985. *Organizing information: principles of data base and retrieval systems*. Orlando: Academic Press.

Thellefsen, Torkild L. 2002. Semiotic knowledge organization: theory and method development. *Semiotica* 142: 71-90.

Thellefsen, Torkild L. 2003. Pragmaticism and the role of terminology. *Impact: an electronic journal on formalisation in text, media and language*. Available <http://www.impact.hum.auc.dk>.

Thellefsen, Torkild L. 2004. Knowledge profiling: The basis for knowledge organization. *Library trends* 52 no.3: 507-14.

Thellefsen, Torkild L. and Thellefsen, Martin M. 2004. Pragmatic semiotics and knowledge organization. *Knowledge organization* 31: 177-87.

Vickery, Brian Campbell. 1980. *Classificação e indexação nas ciências*. Rio de Janeiro: BNG/Brasilart.