

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

## SEMIOTICS AND INDEXING: AN ANALYSIS OF THE SUBJECT INDEXING PROCESS

JENS-ERIK MAI

*jemai@u.washington.edu*

*The Information School, University of Washington, Seattle  
Washington 98195-2840*

This paper explains at least some of the major problems related to the subject indexing process and proposes a new approach to understanding the process, which is ordinarily described as a process that takes a number of steps. The subject is first determined, then it is described in a few sentences and, lastly, the description of the subject is converted into the indexing language. It is argued that this typical approach characteristically lacks an understanding of what the central nature of the process is. Indexing is not a neutral and objective representation of a document's subject matter but the representation of an interpretation of a document for future use. Semiotics is offered here as a framework for understanding the 'interpretative' nature of the subject indexing process. By placing this process within Peirce's semiotic framework of ideas and terminology, a more detailed description of the process is offered which shows that the uncertainty generally associated with this process is created by the fact that the indexer goes through a number of steps and creates the subject matter of the document during this process. The creation of the subject matter is based on the indexer's social and cultural context. The paper offers an explanation of what occurs in the indexing process and suggests that there is only little certainty to its result.

### 1. INTRODUCTION

In the literature, the indexing process is often described as a process of multiple steps. However, discussions have not been concerned with the nature of the indexing process, but mostly with the last step, that of producing an appropriate subject entry. The aim of this paper is to present a theoretical framework for understanding the nature of the indexing process that explains why a predictable result cannot be expected. The attempt is to explain at least some of the major problems related to representing the subject matter of documents; more specifically, to explain the nature of the subject indexing process in a new way. This study is based on the assumption that it is not possible to make a general prescription of how to index and explores the indexing process from the perspective that the process is one of interpretation.

The paper provides an understanding of the subject indexing process that views the process as a number of interpretations that to some degree depend on

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

the specific cultural and social context of the indexer. The aim is *not* to provide a new and improved method for indexing. The investigation is held at a level independent of specific indexing languages and indexing practices.

The main problems of representing the subject matter of documents for retrieval are concerned with meaning and language, more specifically how a statement can be represented using a few words or symbols. Philosophy of language is concerned with how meaning is determined and established and how language can represent reality. There seems to be an overlap of interest between understanding the subject indexing process and philosophy of language; the subject indexing process is, therefore, explored here from a philosophy of language perspective. Others have begun with similar assumptions. Fairthorne (1969), for instance, noted that 'special topics can be treated as isolated topics only at the risk of sterility; therefore some acquaintance with the general problems of language and meaning is essential'. Blair (1990, pp. vii–viii) notes that: 'The central task of information retrieval research is to understand how documents should be represented for effective retrieval. This is primarily a problem of language and meaning. Any theory of document representation ... must be based on a clear theory of language and meaning'. In this respect, this study argues that the subject indexing process consists of a number of steps that should be viewed as interpretations. Benediktsson (1989, p. 218) has noted the interpretative nature of the indexing process and the need for guidelines that recognise the significance of interpretation: 'Any sort of bibliographical description ... can be considered descriptive. When it comes to interpretation, the question is: ought not the description to follow a method or standard as any canon, which makes interpretation possible?'

The present study will explore the approach to studies of indexing and library and information science (LIS) suggested by Fairthorne, Blair, Benediktsson and others.

### *1.1 Steps in the indexing process*

In the literature, the indexing process is often portrayed as involving two, three, or sometimes even four steps. The two-step approach (cf. e.g. Benediktsson, 1989; Frohmann, 1990) consists of one step in which the subject matter is determined and a second step in which the subject is translated into and expressed in an indexing language, i.e.:

1. determine the subject matter of the document;
2. translate the subject matter into the indexing language.

The three-step approach (cf. e.g. Miksa, 1983; ISO, 1985; Farrow, 1991; Taylor, 1994; Petersen, 1994) adds one more step to the process. The subject is still determined first. However, a second step is then included in which the subject matter found in step one is reformulated in more formal language. Thereafter, in a third step, the more formally-stated subject is further translated into the explicit terminology of an indexing language, i.e.:

1. determine the subject matter of the document;
2. reformulate the subject matter in a natural language statement;
3. translate the subject matter into the indexing language.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

The four-step approach (cf. e.g. Langridge, 1989; Chu & O'Brien, 1993) is similar to the three-step approach in the first two points. The first step determines the document's subject matter more or less informally. In the second step, the indexer then summarises the subject matter of the document more formally, usually in his or her own vocabulary and in the form of a more compressed statement.

From this point forward, this approach differs from the three-step approach. Here the translation of the subject matter into an indexing language consists of two steps rather than a single step. In a third step the indexer translates the sentences into the vocabulary used in the indexing language. And in a fourth step the indexer constructs one or more subject entries in the indexing language – in the form of index terms, class marks or subject headings – with respect to their syntax and relationships, i.e.:

1. determine the subject matter of the document;
2. reformulate the subject matter in a natural language statement;
3. reformulate the statement into the vocabulary of the indexing language;
4. translate the subject matter into the indexing language.

It should be noted that the idea of 'steps' as recounted here has to do chiefly with the logic of the indexing process, not necessarily with the actual sequence of mental and physical operations. It may well be that some indexers, particularly those who are beginners in such work, may accomplish their indexing 'by the numbers', ticking off the steps as they go. However, this is less likely as experience is gained. In reality, experienced indexers and cataloguers may not be conscious of the various steps at all, and all steps, regardless of how many one supposes are most accurate, may well take place almost simultaneously. In short, an experienced indexer will perform the indexing process in just one complex action<sup>1</sup>. It is useful, however, to operate with the idea of steps when analysing the process, because breaking down the process into its individual parts will allow one to examine it in greater detail.

The three-step approach is chosen here for several reasons. The two-step model is too simplified in its conception of the subject indexing process. In fact, the two-step approach appears to be used chiefly as a device to separate two distinct activities in the subject indexing process: determining the subject of a document and converting that subject to the terminology of an indexing language. It is seldom used to discuss the details of the process itself. In contrast, the four-step approach appears to add an unnecessary complication to the final part of the process which consists of the activity of translating the subject of a document into the terminology of an indexing language. The four-step approach breaks that final part of the process into two parts which is not useful as there is no essential difference between these two steps but only a difference of general versus specific activity. In the first of these two final steps, the subject of a document is said to be translated into the language of a given subject access vocabulary, whereas the next step only translates the results into indexing terms or strings of terms (i.e. the syntax) in the system.

---

<sup>1</sup> Mai (1999) has explored this development of indexers from being novice indexers to becoming experts.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

However, the focus here is not merely on the steps themselves, but rather an enhanced rendition of the steps is presented. More accurately, the view of the indexing process presented here consists of four elements and three steps, where an element consists of an object that is acted upon and a step that is the action taken upon the object. The sequence of the elements and steps is as follows: Element 1 – Step 1 – Element 2 – Step 2 – Element 3 – Step 3 – Element 4.

The first element consists of the document under examination. As an object upon which action is focused, this document is a given. Its presence causes the indexing process to swing into action.

The first step, called the document analysis process, occurs in response to the presence of the document. It consists of the act of examining the document<sup>2</sup> (i.e. the title, the table of contents, the abstract, if there is one, the back of the book index, reviews of the item, and so on) in order to identify its subject.

The second element is the product of the first step. It consists of some mental sense of the subject of the document on the part of the indexer. It could be called the subject of the document as it exists initially in the mind of the indexer and includes a relatively unordered mass of mental impressions, phrases, terms etc. which have been collected in the process. These ideas have been generated from the sources that form the basis of the examination process in the first step.

The second step is the indexer's response to the second element and is named the subject description process; it consists of the act of attempting to create a cohesive formulation of the subject of the document in language. In short, the product of the first step is a relatively unordered mass of mental impressions, phrases, terms etc. which have been collected in the document examination process. The product of the second step is the result of a concerted effort to give those various impressions, phrases, terms, etc. some sort of order and structure.

The third element is the product of the second step. It consists of the more or less cohesive formulation of the subject of the document in language – a subject description. The second element consists primarily of a mental product, a sort of running mental tab of the various candidate terms, ideas, concepts and so on that one collected in examining a document. This third element represents an attempt to compress all of these into something that in a relatively cohesive way summarises the subject of the document.

The third step is prompted by the presence of the third element, that is, of a relatively cohesive summary of the subject of the document in language. This step is named here the subject analysis process and consists of translating the product of the third element into a formal statement of the same thing, only this time in terms of the language of the appropriate subject access system. In short, it means converting one's language statement into, for example, class numbers, subject headings or descriptors. In this activity, one must, of course, be aware of all of the various rules, conventions, proscriptions and so on that any system uses.

The fourth element, the terminus of the process, is simply the product of the third step. It consists of the completed subject entry from a given system that the indexer has finally chosen to represent the subject of the document.

---

<sup>2</sup> Although the ideas presented here could be generalised to other media (such as web pages, film, sound, images, etc.) the concept of documents is here limited to books and journal articles because most work in indexing has been concerned with these media.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

To rephrase, the sequence of the elements and steps in the indexing process is: the document (element) – the document analysis process (step) – the subject (element) – the subject description process (step) – the subject description (element) – the subject analysis process (step) – the subject entry (element).

### *1.2 Semiotics in information science*

From the above description of the indexing process, it should be clear that it actually consists of multiple interpretations. If this process is a series of interpretations, then a theory that can explain the nature of the process from this perspective is needed. The study of signs and semiotics, as discussed in the writings of Charles Sanders Peirce (1839–1914), is suggested here as a useful theoretical framework for studying and understanding the interpretative nature of the subject indexing process. Peirce's semiotics is useful for this because it includes an explanation of how the meaning of signs is generated, interpreted and represented.

Some scholars within the LIS field have found semiotics a useful theoretical framework. Cronin (2000) suggested semiotics as a framework for understanding citations and bibliometrics. Smiraglia (2000) used semiotics in his analysis of the concept of work and Buckland (Buckland & Day, 1997; Buckland, 1997) in his analysis of the concept of document. Brier (1996) argued that semiotics together with second order cybernetics and Wittgenstein's pragmatic philosophy of language could form the theoretical foundation for the field. Karamüftüoğlu (1996) has used semiotics to analyse the information retrieval process, Wagner (1991) to analyse the communication processes in public libraries, and Warner (1990) has noticed that there is a conceptual overlap between semiotics and LIS, which has not yet been investigated thoroughly. In a National Science Foundation (NSF) funded research project, Pearson and Slamecka (Pearson, 1980; Pearson & Slamecka, 1977) used Peirce's semiotics to form the foundation of a pragmatic approach to programming and understanding information systems.

Perhaps the best known discussion of semiotics in LIS, and the most important for the present study, is Blair's analysis of language and representation problems in information retrieval. In his book, *Language and representation in information retrieval*, Blair (1990) argued that theories of indexing and retrieval have to include explicit theories of language and meaning in their foundation. Blair especially used Wittgenstein's pragmatic philosophy of language for understanding information retrieval.

The major part of Blair's book is an analysis of the importance of language in indexing and representation. Blair argues that Wittgenstein's philosophy of language has significant bearings on the understanding of indexing and representation of documents. However, Blair rejects semiotics as a possible foundation for understanding indexing and information retrieval. He argues that 'semiotics begin from the perspective that certain words/expressions exist and that they need explanation' (Blair, 1990, p. 145). This may be true of Saussure's semiology, but not of Peirce's semiotics. Semiotics, in Peirce's understanding, can be defined as the study of meaning as represented by signs, what meaning is, how and where meaning comes into existence, and how meaning is transformed and combined. Semiotics does not focus on what a specific phenomenon means, but rather on why and how meaning exists. Instead of semiotics, Blair argues that the later

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

Wittgenstein's (1958) theories are useful as a foundation for understanding how to represent documents for retrieval. However, Peirce's semiotics and Wittgenstein's pragmatic philosophy of language are quite alike.

## 2. SEMIOTICS

Semiotics is generally defined as the study of signs. Two traditions of the study of signs can be identified, a European and an American. The European tradition is based on the work of the French linguist Ferdinand de Saussure (1857–1913) (Saussure, 1966). This school is usually named semiology. The American tradition is based on the work of the American scientist and philosopher Charles Sanders Peirce (1839–1914) and is called semiotics (or semeiotic, as Peirce preferred to spell it). Although there have been attempts to define a unified theory of semiotics, most notably by Eco (1984), the two traditions are distinct. Saussure's theory is a theory of how to derive meaning from words. Peirce's theory, on the other hand, is about how signs in general, and not only words, are attributed meaning. Johansen (1985, p. 225–226) has discussed the distinction between the two traditions: 'As a contra distinction to the concept of sign of continental structuralism (Saussure, Hjelmslev), defining the sign as an immanent solidarity between two formal entities (an element of expression and one of content), Peirce conceives the sign as an element in a signifying process'. In short, Saussure operated with a dual concept of the sign. He suggested that words are not merely names that represent things, but are expressions that stand for some content. By this, he separated words and their content. Saussure argued against the notion that words have an inherent quality, as earlier linguistics had suggested. Instead, he argued that the connection between a word and its content is arbitrary and his theory is centred on how to derive meaning from words.

Peirce (1955; 1958) defined a sign as a relation among three entities, the sign itself, the referent of the sign, and the meaning that is derived from the sign. Peirce's concern was how meaning is derived from a sign and transformed into another sign. He operated with a three-sided, or a triadic concept of sign, which he (Peirce, 1955, p. 99) defined as:

A sign, or *representamen*, is something that 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.

He distinguished between the physical entity, for example words, the ideas that these words refer to, and the meaning one derives from the words. Peirce's concept of a sign is represented as a triangle, as shown in Figure 1, based on a figure by Johansen (1993). The triangle is sometimes referred to as the Ogden Triangle, although it is evident that Ogden & Richards (1923) got their inspiration from Peirce (Fisch, 1986, p. 344).

The representamen is that which represents the sign, often in the form of a physical entity or at least manifested in some form. The representamen is, in other

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

Figure 1. *The semiotic triangle*

words, the entity of the sign relation that is perceived and therefore often denoted the 'sign'.

The representamen represents an object. However, there is not a one-to-one relationship between the representamen and the object. The object is not some identifiable entity that exists independent of the sign. Peirce (1955, p. 101) states about the object that,

The Objects – for a Sign may have any number of them – may each be a single known existing thing or thing believed formerly to have existed or expected to exist, or a collection of such things, or a known quality or relation or fact, which single Object may be a collection, or whole of parts, or it may have some mode of being, such as some act permitted whose being does not prevent its negation from being equally permitted, or something of a general nature desired, required, or invariably found under certain general circumstances.

The sign can only represent the object and tell about it, it cannot furnish acquaintance with or recognition of the object. The object, therefore, is not some objective entity that exists and which can be known or realised through the sign. The object is 'that with which ... [the sign] presupposes an acquaintance in order to convey some further information concerning it' (Peirce, 1955, p. 100). The object should be understood as the background knowledge that one needs to understand the sign, or the range of possible meaningful statements that could be made about the sign. The representamen could be any item that represents or stands for something else – Peirce's notion of signs is not limited to words or language. As will be shown later a document can therefore be regarded as a sign.

The connection between the representamen and its object is made by the interpretant, which is the third entity in the sign relation. The interpretant is *not* a person who interprets the sign, but rather the sign that is produced from the representamen. In other words, when the representamen is perceived as a sign, a new and more developed sign is created on the basis of the representamen. The person who interprets the sign makes a connection between what he or she sees (which is the representamen) and his or her background knowledge (which is the object) and thereby creates an understanding or meaning of the sign (which is the interpretant). This process is called semiosis, the act of interpreting signs.

The connection of the representamen and the object to create the interpretant as a process of semiosis is emphasised in the Y-leg model of the sign in Figure 2, based on a figure by Larsen (1993). The bold line from representamen to object stresses the connection between the primary sign (the representamen) and its referent (the object). The connection between these two entities is the meaning of the



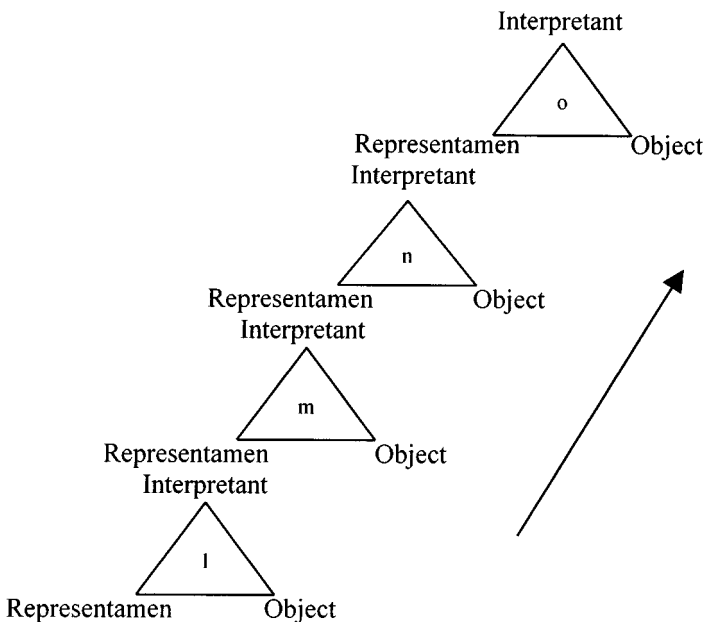
All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

Figure 2. *The Y-leg model*

representamen which is represented as the interpretant. This idea is stressed in the Y-leg model, but is less clear in the semiotic triangle. Throughout this paper, both models will be used, however.

A key element in Peirce's theory of semiotics is the notion of 'unlimited semiosis' which could be seen as the connecting of sign or the process of one sign producing another sign. Unlimited semiosis is based on the fundamental idea of semiosis; that a sign (b) is generated on the basis of another sign (a). When a new sign (c) is generated on the basis of the second sign (b), still another semiosis process occurs. Because new signs will always generate still more signs, this process can continue indefinitely and is, therefore, unlimited, hence the term, 'unlimited semiosis'.

The unlimited semiosis process is represented in Figure 3, based on a model by Johansen (1993, p. 80). The interpretant of the first sign in unlimited semiosis changes to become the representamen in the second sign. There is a relation between these, but the object in each case remains independent of both the representamen and the interpretant. The object will change throughout the process. Each object relation in the unlimited semiosis process will be unique to that sign

Figure 3. *Unlimited semiosis*



All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

relation. The single objects in the unlimited semiosis process are independent of each other. The latter is crucial for Peirce's theory of semiotics.

In Figure 3, the triangles in the figure will continue to generate new triangles. It should be understood that there have been triangles before them, and there will be triangles after them. What this means is that an understanding of something is always based on an understanding of something else; and it will always generate still another understanding. Two important ideas are illustrated here. First, the different sign relations have different objects, each of which is dependent on the person for whom the interpretant is created. There are no necessary relations between the different objects. Second, each representamen is based in turn on an interpretant, which again is based on a representamen.

### 2.1 Categories of signs

Peirce divides signs into a number of categories to illustrate their different kinds. One set of sign categories commonly associated with his work consisted of icon, index and symbol. This approach to categorisation grouped signs on the basis of their relation to their referent and object. In this respect, an icon sign is based on resemblance (like the sign on a bathroom door), an index sign points to what the sign refers to (like smoke to a fire) and a symbol sign refers to a convention (like language).

The categorisation into icon, index and symbol is a simple representation of Peirce's full categorisation of signs. To reach this, Peirce defined three modes of each entity (interpretant, representamen and object) of the sign. These are based on Peirce's phenomenology, in which he argued for a division of the world into three modes of phenomena – or three modes of being. Before this categorisation of signs is more fully explored, however, Peirce's phenomenology must be introduced.

*2.1.1 Three modes of being* Peirce argues that everything that exists in the world, including feelings, ideas and thoughts, belongs to one of three fundamental modes of being. These are the modes of being of positive qualitative possibility, of being of actual fact and of being of law (or conventions). Peirce named these respectively firstness, secondness and thirdness.

Firstness is the mode of monadic being that consists of the category of qualities of phenomena, such as red, bitter and hard. This existence is neither dependent on its being in the mind of some person, whether in the form of sense or in thought, nor on its being in the form of some material thing possessing the quality (Peirce, 1955, p. 85).

Secondness is the dyadic mode of being that tells something about other objects. Secondness is the relations between things (Hoopes, 1991, p. 10); Peirce furthermore describes secondness as facts. It is the direct relation between things, for instance, between the whistling locomotive and the perception of the whistle.

Thirdness is the triadic relation between something first and something second, which reveals information about something third. This can generally be defined as meaning. Meaning is not inherent in signs, but something one *makes* from signs. Peirce speaks of thirdness as the category of law (e.g. Peirce, 1955, p.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

90); by this he means that thirdness is a relation between two things that is established by humans.

*2.1.2 Trichotomies* Each sign consists of three entities: representamen, object and interpretant, which all must be present to make a sign. The three entities of the sign have three elements, which reflect the three modes of being; firstness, secondness and thirdness. This trisection of the sign, and the further trisection of its components are essential to Peirce's semiotics. The trisection of the trisection is represented graphically in Figure 4, based on a figure by Christiansen (1988).

The three inner categories – rheme, icon and qualisign – represent firstness. The middle categories – dicent sign, index and sinsign – represent secondness. And the outer categories – argument, symbol and legisign – represent thirdness.

Any sign consists of an element from each of the three legs in Figure 4 and this combination of sign elements makes up the individual categories of all signs. In other words, not only did Peirce divide the sign into three elements but these elements were furthermore divided into three elements each.

The representamen is divided according to whether the sign itself is a mere quality (qualisign), an actual existent (sinsign), or a convention (legisign):

- a qualisign is a quality, which is a sign;
- a sinsign (the syllable 'sin' is derived from singularity as in single, simple etc.) is an actual existent thing, an individual object, an act or an event. In other words, the sinsign is thisness in the sense it represents specific objects, acts or events;
- a legisign is a general type, law, habit or convention, which is established by humans.

For instance the sign 'A' could be considered (1) black lines or the quality of black ink on paper (i.e. a qualisign), (2) a good example of the class of letter 'A', which would be an actual existent (i.e. a sinsign), or (3) an expression of satisfaction with a term paper, that is, a convention (i.e. a legisign).

The object is divided according to the sign's relation to the object it represents. The sign could either have some character in common with its object (icon), some existential relation to that object (index), or only have a representational relation to its object (symbol).

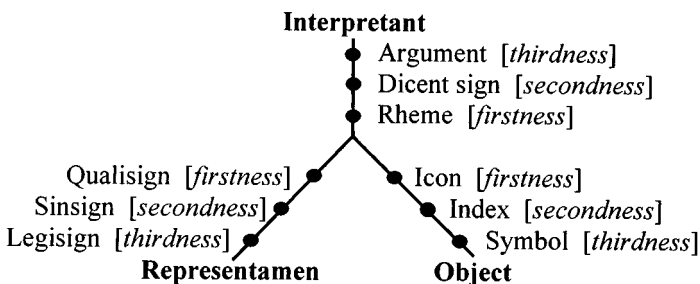


Figure 4. *Triadic classification of signs*

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

- An icon is a sign that shares some kind of likeness with that which the icon represents;
- an index is a sign, which refers to its object by being affected by that object and as such points out its object;
- a symbol is a sign, which refers to its object through law, habit or convention. This usually takes place through an association of ideas by which the symbol is interpreted as referring to its object.

An example of an icon is a pictogram where the sign resembles the object. An example of an index is a footprint, which points to a person. And an example of a symbol would be a sign based on the context in which it occurs, such as a street sign with the letter P, which means 'parking allowed'.

The interpretant represents the sign as a sign of possibility (rheme), a sign of fact (dicent sign) or a sign of reason (argument).

- A rheme is understood as representing a certain kind of possible object. The meaning of a rheme is easily understood;
- a dicent sign is more complex than the rheme, which means that it requires more knowledge to interpret it than to interpret a rheme;
- an argument is a sign of reason or law and is understood to represent its object in its character as sign. The argument should be 'contemplated as a sign capable of being asserted or denied' (Peirce, 1955, p. 104).

Examples of these three are: rhemes are nouns (e.g. 'house', 'car'), sinsigns are propositions (e.g. 'the house is green', 'the car is fast') and legisigns are arguments, i.e. meaningful links of propositions (e.g. 'Jones has a green house and a fast car. Smith on the other hand does not like to drive and therefore prefers to bike...').

The above examples are all rather weak since any sign is defined as a combination of all the elements of the sign, such that each sign consists of one element from each of the three trichotomies. The examples for each aspect of the sign are therefore incomplete since two elements of the sign are missing. A sign will always consist of three elements and each of the examples depends on the two missing elements.

By combining the above categories of sign elements Peirce defined ten categories of signs. Although a total of  $3^3$ , or 27, different categories of signs could be enumerated<sup>3</sup>, Peirce only enumerated ten, since some possible signs are logically excluded. A qualisign will, for instance, always be a rhematic icon (because a mere quality cannot be a convention). A symbol will always be a legisign (a symbol is a representation of its object based on context, and a legisign is a sign based on convention). An argument will always be a symbolic legisign (since an argument always is thirdness to the interpretant and requires a high degree of interpretation).

---

<sup>3</sup> Many other numbers have been considered; Seboek (1994) for instance, expanded the three basic signs – icon, index and symbol – into six signs. Marty (1982) enumerated twenty-six, and Weiss & Burks (1945) sixty-six. Since each sign possesses its own triads, Peirce argues that a total of  $3^{10}$ , or 59,049, signs could be enumerated (Merrell, 1997).

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

2.1.3 *The individual categories of signs* Each of the ten categories of signs is loosely defined. Though they are not clearly distinguished, they can be viewed as ten points on a continuum from the mere sense of a feeling to a complex statement:

- I. a qualisign is 'a feeling, a sensation, for example, the sense of "blueness" upon one's being subjected to a blue object' (Merrell, 1997, p. 193);
- II. an iconic sinsign is any 'object of experience in so far as some quality of it makes it determine the idea of an object' (Peirce, 1955, p. 115);
- III. a rhematic indexical sinsign is any 'object of direct experience so far as it directs attention to an object by which its presence is caused' (Peirce, 1955, p. 115);
- IV. a dicent sinsign is 'any object of direct experience, in so far as it is a sign, and, as such, affords information concerning its object' (Peirce, 1955, p. 115);
- V. an iconic legisign is 'any general law or type of sign, insofar as it manifests some likeness with something other than itself' (Merrell, 1997, p. 194);
- VI. a rhematic indexical legisign is 'any general type or law of sign, however established, which requires each instance of it to be really affected by its [semiotic] object' (Peirce, 1955, p. 116);
- VII. a dicent indexical legisign is 'any general type or law, however established, which requires each instance of it to be really affected by its object in such a manner as to furnish definite information concerning that object' (Peirce, 1955, p. 116);
- VIII. a rhematic symbol is a 'sign connected with its object by an association of ... ideas' (Peirce, 1955, p. 116);
- IX. a dicent symbol, 'or ordinary proposition, is a sign connected with its object by an association of ... ideas, and acting like a Rhematic Symbol, except that its intended interpretant represents the Dicent Symbol as being, in respect to what it signifies, really affected by its object' (Peirce, 1955, p. 117);
- X. an argument is a sign 'whose interpretant represents its object as being an ulterior sign through a law, namely, the law that the passage from all such premises to such conclusions tends to the truth' (Peirce, 1955, pp. 117–118).

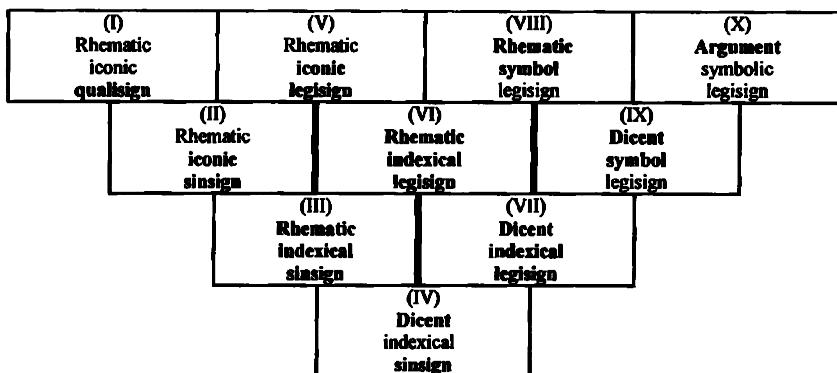
These ten categories of sign are regarded as basic and provide a framework for discussing different kinds of interpretation, in the sense that different kinds of signs require different kinds of interpretation (see Figure 5).

After this introduction to Peirce's categorisation of signs it should be clear that the everyday use of the concept of sign is rather limited in scope. There are in fact many different kinds of signs, a difference that can be ascribed to the way signs are attributed meaning, and to the way they are interpreted. It should also be clear that there are different kinds of interpretation. Interpretations are sometimes a mere translation of a sign into an action, and at other times an interpretation requires an involved understanding of the social context in which the sign is used.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

Figure 5. *Ten categories of signs*

### 3. A SEMIOTIC ANALYSIS OF THE INDEXING PROCESS

As was outlined in the introduction, the subject indexing process consists of four elements (document, subject, subject description and subject entry) and three steps (document analysis, subject description and subject analysis). These elements and steps are interconnected in such a way as to be explainable in terms of Peirce's ideas of unlimited semiosis and signs. The next section (Section 3.1) will show briefly how unlimited semiosis fits the case. It will be followed by detailed explanations of how the individual steps of the subject indexing process are to be viewed (Section 3.2) and how Peirce's categories of signs lend insight into the nature of the interpretation that occurs (Section 3.3).

#### 3.1 *The subject indexing process as unlimited semiosis*

The subject indexing process can be expressed in terms of Peirce's idea of unlimited semiosis: each element of the subject indexing process is to be regarded as a sign, with each step functioning as an act of interpretation linking the signs in a sequential process.

The process begins with an initial sign, the document. The indexer initially makes an act of interpretation (the first step) in order initially to determine what the first sign, the document, is about. The product of this act is a new (or second) sign, the subject. A new act of interpretation (the second step) is then made in order to convert what the indexer has come up with as a subject to something more manageable and concise for indexing. The product of this act is still another new (a third) sign, the subject description. Finally, still another act of interpretation (the third step) is made in order to fit the subject description into a given subject indexing system's vocabulary. This act in turn develops still another new (the fourth) sign, the subject entry. One could extend this process further, of course. For example, the user will come to the index and view the subject entry (a sign) and in an act of interpretation view it as a statement of aboutness for the document, though in this case, the aboutness will likely be related in some fashion to the reason for which the user is searching out information in the first place. The user's conclusions about what the subject entry means will constitute still another sign. And so on.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

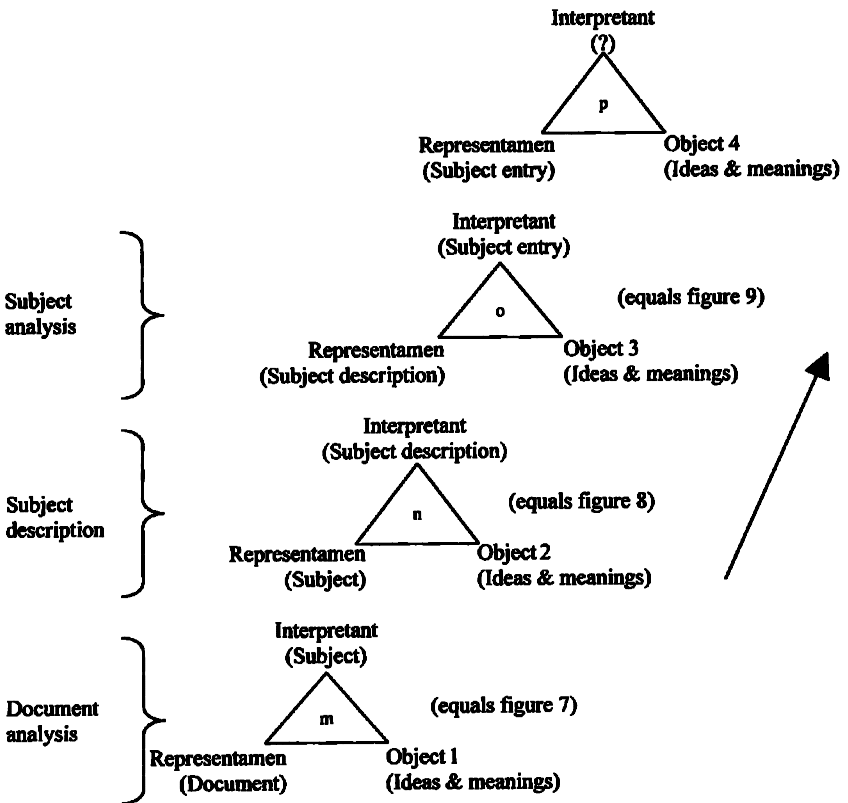


Figure 6. *The semiotic model of indexing*

The entire process is presented in Figure 6. It should first be noted that the triangles in the figure are called *m*, *n*, *o* and *p* (rather than, say, *a*, *b*, *c* and *d*) to emphasise the fact that in reality subject indexing is part of a much larger process of interpretation.

Before an indexer begins the subject indexing process, the document will have been created in some sort of a discourse community, perhaps a scientific discourse community. Its very creation is the result of many acts of interpretation on the part of the document's author and on the parts of those to whom the author refers. Once completed and published and after the subject indexing process has made the document accessible, the document will be retrieved and used by a number of information users, some within that discourse community, and others outside it. The activities of those information users in consulting a catalogue and focusing on the subject entry terms that represent the document in an indexing system (among other documents) and subsequent uses of the document as a whole or in part are likewise acts of interpretation. In short, the process of unlimited semiosis, confined here chiefly to the subject indexing process, started before the subject indexing process began and will continue after it is completed. Figure 6 simply represents an intercepted portion of the larger process.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

The second thing to be noted in Figure 6 is the layout of its triangles. Each triangle is a sign that constitutes an element in the process of unlimited semiosis. It should be remembered that the sign is defined as a relation between three entities: the representamen, the interpretant and the object. This relation constitutes the sign and is in itself a process of semiosis, or interpretation. In other words, each element is a sign and the interpretation of the sign is a process of semiosis. As such, each triangle shows a process of semiosis with the beginning sign that is interpreted (representamen) in its lower left corner, the newly created sign from the act of interpretation (interpretant) at its apex, and the range of ideas and meanings associated with the representamen (object) in its lower right corner.

The third thing to be said of the figure is that it should be noted that the clear distinction between the elements and the steps of the subject indexing process, which was outlined in the Introduction, collapses here. It was argued in the Introduction that an element consists of an object that is acted upon and a step is the action taken upon the object. This argument was put forth in order to take the elements of the indexing process into consideration. Earlier explanations had merely focused on the steps and ignored the position of the elements in the process. However, in view of the above explanation of Figure 6 it should be clear that no precise lines of demarcation exist between the elements and the steps. Rather, the elements and steps collapse into one single act of interpretation, semiosis. When the indexer acts upon an element, he or she is in fact already thrown into the step leading to the next element. For instance, when the indexer views and acts upon the document, that act is in fact the first step, the document analysis, of the subject indexing process. The indexer cannot view or act upon the document and then afterwards go into the first step. The elements and steps cannot be separated into two different kinds of phenomena. However, in order to reach a better understanding of the subject indexing process, the following discussion will continue to analyse elements and steps distinctively, but it should be clear that this distinction in reality cannot be maintained.

The final thing to be said of the diagram is that references to Figures 7, 8 and 9 are placed in the diagram for the purposes of correlating it with the discussions that are to follow which are expansions of the description of this diagram. The nature of the various acts of interpretation in the continuous semiotic process will be presented in the next section; and the nature of the particular signs in the process will be presented in Section 3.3.

### 3.2 *The steps of the subject indexing process*

In order to provide a greater degree of understanding of how Peirce's process of unlimited semiosis can be used as a basis for understanding how the subject indexing process works, the individual acts of interpretation of that process will be discussed in greater detail next. One aspect of this discussion will be the inclusion of an example of subject indexing, in this case, determining the subject, subject description and subject entries for *The organization of information*, a book of nearly 300 pages by Arlene Taylor (1999).



All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

*3.2.1 Step 1: document analysis* The first step in the subject indexing process is to analyse a document in order to determine its subject matter. In this step, the document, as a sign that is being interpreted, is the representamen, and the product of the step is a new sign. The sign consists of its point of departure, the document (representamen), the subject (interpretant) and the range of ideas and meanings associated with the document (object). Figure 7 illustrates this in the form of a diagram, but it should be noted that this diagram merely represents the lowermost triangle in Figure 6 extracted as a separate diagram.

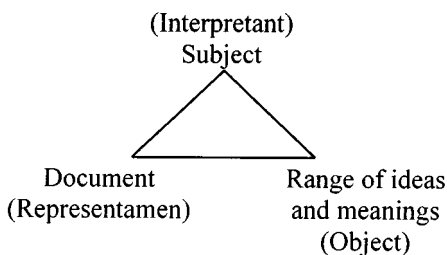


Figure 7. *Document analysis*

It would be nearly impossible, of course, for any single person or, in this case, any single indexer, to determine *all* of the ideas and meanings which might be associated with any particular document, since there might always be potential ideas and meanings which different people at different times and places might find in the document. Furthermore, it would be well nigh impossible to predict precisely which of the many possible ideas and meanings that could be associated with the document would be specifically valuable to the users or would have some sort of lasting value for the document. To recognise and accept this fundamental openness is of utmost importance. The indexer must realise from the start that he or she will never discover all the ideas and meanings that could be associated with the document and that, therefore, it is not possible to describe all these ideas and meanings.

How might an indexer discover 'the subject' of Arlene Taylor's book, the first step in the subject indexing process?<sup>4</sup> For this purpose, the indexer would look at different places in the book – for example, the title, the tables of contents, the preface, etc. This would provide ideas about the topical content of the book. By doing this, a general impression of the document would begin to accumulate. By just looking at the title it might be supposed that the book is simply about knowledge organisation. However, from reading the preface the indexer is informed that Taylor intends the book to be used as a textbook and introduces to students a work

<sup>4</sup> It should be noted that what is referred to as determining the subject in this semiotic process is only taking an initial step in a more involved process. It should not be confused with the more common way of stating the process, which involves going all the way to the subject entry. In short, normally when an indexer or cataloguer hears the statement, 'discover the subject of the book', it is not truncated to a single step. The subject, subject description, and subject entry are, in common parlance, all the same thing. Here they are stages in a series of interpretative steps.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

on library cataloguing and classification before they encounter Wynar's *Introduction to cataloging and classification* (Wynar, 2000). This gives the indexer some idea about how the author herself viewed the content of the document, intending it to provide information on matters of information organisation that precedes library cataloguing and classification.

The titles of the individual chapters in the book at first glance suggest that the range of topics discussed is much broader than simply library cataloguing and classification. There are chapters on 'Organization in human endeavors', 'Retrieval tools', 'Development of the organization of recorded information in Western civilization,' 'Encoding standards,' 'Metadata,' 'Verbal subject analysis,' 'Classification,' 'Arrangement and display' and 'System design.' These titles suggest that the book deals with topics ranging from philosophical issues to historical issues to technical issues. At the same time, however, by looking at the titles of the sub-chapters the indexer will learn that the book not only covers a wide range of problems, standards and issues in knowledge organisation in general, but it also does so with a special orientation to library cataloguing and classification. For example, the sources of some of the discussions are clearly from library cataloguing and classification, rather than from some more general level. In short, while the book at first appears to be about knowledge organisation in general, it also appears to treat that topic at least some of the time from the narrower standpoint of library cataloguing and classification.

The sources that supported the foregoing ideas were found within the book. However, as the document analysis proceeds, external sources will also inevitably play a role which could well include the indexer's general knowledge of library and information science, a possible knowledge of Arlene Taylor's other works, a knowledge of the users of the information system, and ultimately, the indexer's personal situation and experience. With respect to the latter, were I indexing this work for a given system, I would consider, for instance, whether the book could be used in courses I teach in knowledge organisation, how the book supplements other standard works on cataloguing and classification, and how Taylor talks about the subject indexing process.

By the above process, the indexer ultimately collects what earlier was called the 'range of ideas and meanings associated with a document' – the 'object' of the sign in the document analysis step. The accumulated ideas and meanings are at this point, however, more like a collage of impressions of the book rather than some systematically organised statement about it. To arrive at a point of more formal organisation will require the second step in the ongoing process of unlimited semiosis as applied to the subject indexing process.

*3.2.2 Step 2: subject description* The second step (Figure 8), creating the subject description, begins with the subject that was reached in the first step. The representamen of the sign relation in the second step is now the subject of the document that the indexer reached in the first step, rather than the document itself. And the interpretant of the sign relation, which is the product of the second step, is the subject description, more formalised and condensed than the subject matter that resulted from the first step.

To say that the subject description is more formalised and condensed than the

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

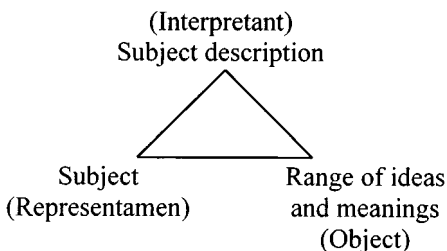


Figure 8. *Subject description*

subject is not a reflection about committing it to writing, although the indexer may in fact write the subject description down at this point. It is rather a reflection of picking and choosing from among the range of ideas of meanings encountered in assembling the subject, or of combining elements of the collage assembled during step 1 in order to produce a sensible assertion or set of assertions about the document's subject.

At this point, it is important to remember that the subject reached in the first step, the document analysis process, was primarily a mental matter. As such, it contained a great many associations and couplings the indexer found in the text and in other sources. By way of contrast, in the subject description process the indexer summarises the information compiled in step 1 in a more or less formalised subject description. Such a description will not likely contain all the associations the indexer made during the first step but rather only those that for various reasons he or she concludes should eventually become statements of the document's subject matter within the system for which he or she is working. The reasons why some might be used and not others will include things like limitations on how many indexing entries may be prepared per item, a sense that some of the ideas encountered in step one are better representatives of the document than others, and so on.

In order to provide a more realistic illustration of this process, it will be best to return to the process of indexing the Taylor book.

The initial point for this step, when applied to Taylor's book, is the subject collage that was accumulated as the product of the first step. A subject description of the subject collage accumulated for Arlene Taylor's book in step 1 might be something like the following:

This book gives a broad introduction to the fundamentals of knowledge organisation. It introduces and discusses the most important issues, concepts and problems in knowledge organisation and shows how the novice information scientist designs and implements information systems.

This description focuses on the most obvious and broadest of the themes accumulated in the document analysis step. At the same time, it includes only a part of the collage of ideas that were accumulated in the document analysis step. Another theme noticed at that stage was the fact that this book has a special relationship to the narrower knowledge organisation practices known as library cataloguing and classification. Were that to be acknowledged in a subject description, it might appear something like this:

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

This book is somewhat of an introduction to the fundamentals of library cataloguing and classification.

This subject description is not nearly as accurate as the previous. However, the major theme of the book appears to be captured by this description as well. A number of other ideas were represented in the accumulation of ideas and meanings found in the subject collage, but many of those were simply sub-elements of one or the other of these two descriptions. And, if the goal is to find subject descriptions that cover or summarise or match the main scope of the content of the work, then the two illustrated above are about as close as one might get to the goal. Still, choosing these two over any others, or even between these two are the kinds of decisions the subject indexer must make, for that is the heart of subject description – creating more formalised statements that will be amenable to conversion to subject entries, the final step in the process. It is important to note that this too was a process of interpretation, for just as one interprets in going from document to subject collage, so also one interprets anew in going from subject collage to subject description.

3.2.3 *Step 3: subject analysis* In the third and final step (Figure 9), the subject analysis, the indexer moves from the subject description to a subject entry which is itself based on an understanding of the subject description, and not directly on dealing with the initial subject estimation of the document. In this step, the representamen of the sign is the subject description already arrived at in step 2, and the product of the subject analysis is the subject entry, i.e. the interpretant of the sign. Again, the object is a range of ideas and meanings, although this time that range is associated with and limited to the subject description.

As the indexer moves from the subject description to the subject entry he or she should convert the subject description directly into a subject entry according to the requirements of the system being used. The subject entry will result in expressing the subject description reached at the apex of the second step in terms of the particular indexing language the indexer works with. The third step may thus be characterised as being similar to translating a text from one language to another. In the latter, a translation is a mixture of the syntax, semantics and pragmatics of the two languages involved. Likewise, a similar thing occurs as the indexer proceeds from the subject description to the subject entry. These too represent two languages: first in the indexer's own language and second in the language of the indexing system.

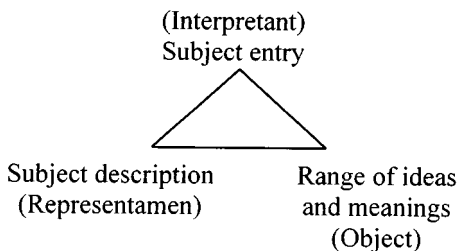


Figure 9. *Subject analysis*

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

It should be noted at this point that some have argued that the first step should be conducted without any focus on the indexing language. Langridge (1989, p. 7), for example, argues that because the first step 'relates to the document and not to the system' the indexer should not be concerned about the system in this initial step. It is the opinion here, however, that no matter how exemplary Langridge's ideas may sound, in practice it will be almost impossible to do as he says in his warning, and in reality doing so may lead to unnecessary work. Avoiding the indexing system's language will be impossible because the indexer will know which indexing language the subject indexing process will eventually be expressed in, and the indexer will, if not consciously, then unconsciously, think in the terms of the ontology of that system. By using the system repeatedly the indexer will have learned to represent documents according to this particular view.

Returning once again to Taylor's book, step three takes place as the subject description is analysed and translated into the indexing language. The basis for this analysis is the written subject description reached in the subject description process.

The two themes of the subject description of Arlene Taylor's book that were produced in step two are converted into an indexing language. Here for the sake of the illustration, the indexing language to be used is that of the Library of Congress Subject Headings (LCSH).

The most obvious theme in the subject description is that the book is about organising knowledge. However, LCSH has no term that specifically matches that phrase and the translation of the subject description into the subject entry will be distorted. If one uses only the terms already extant in the LCSH and does not create a new term, then one is hard put to find a term that exactly matches the emphasis here. About the closest one might come is the term 'information retrieval', but this is deceptive because all things being equal this term is ordinarily used to cover many subtopics that are not included in the Taylor work. That this is the case can be seen in the list of narrower terms associated with the term 'information retrieval' in the LCSH system. Thus, to choose it to represent the theme of knowledge organisation would be to list this book in a system using LCSH among works that cover a much broader range of subtopics than Taylor's work covers.

Other terms are even less representative. One might consider, for example, 'information storage and retrieval systems', but this work is not specifically about particular systems. Or, again, one might use 'information science', but this term is manifestly too broad for the theme of knowledge organisation. This is the case because it is not limited to the process of knowledge organisation but rather includes a wide range of subtopics that have to do with the field of information science.

The second theme that arose in determining the subject description is library cataloguing and classification. Here too, LCSH offers only modest help as there is no single term for this entire theme. Rather, two separate terms, 'cataloguing', on the one hand, and 'classification – books', on the other hand represent this theme. To choose either of these terms would be to place Taylor's work among others

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

that are fully about both topics when, in fact, Taylor's work is only partially about each topic.<sup>5</sup>

The important thing here, however, is not the particular terms that one has to contend with, but rather the fact that when the indexer moves from the subject description to subject entries, he or she is again engaging in an act of interpretation. This time, however, the sign being interpreted is neither the document nor even the subject as an interpretation of the document, but rather only the results encapsulated in the form of the subject description. And the result of the interpretation has a strong possibility of being even farther removed from the reality of the document as a sign than even the more formalised statements that one arrives at in step two when a subject description is produced.

### *3.3 Peirce's categories of signs and the four elements*

To elaborate on the differences between the elements in the subject indexing process each element is now categorised in terms of Peirce's categorisation of signs. The reason for doing this is because Peirce's categorisation highlights the potentially different kinds of interpretation that can be involved for any particular sign, and these differences in interpretation are in reality related to the fact that signs are of different kinds. In other words, although each of the elements in the subject indexing process – document, subject, subject description and subject entry – are signs, that each is a different kind of sign will make a considerable difference in how one approaches them and interprets them. However, by matching the elements of the subject indexing process to the most appropriate of Peirce's ten categories of signs, the kind of interpretation required at each step in the subject indexing process will become more evident.

Although the task of matching the elements with Peirce's ten categories of signs might seem like a straightforward task, it is actually a complex matter. This is the case because each of Peirce's ten categories of signs appears as a trichotomy with three modes of involvement, one for each of Peirce's ideas of firstness, secondness and thirdness. Thus, the task here has been to determine which of the ten possible categories of signs fits each of the four elements in the subject indexing process: document, subject, subject description and subject entry. A brief list of the matches is found in Figure 10 and the justification for the matching is found in the following.

In each case, the mode of meaning in the sign will be described as to how it best fits the case and as to its effect on interpretation. When the choice of category for each element in the subject indexing process is made in the following, each element will first be placed in the categorisation of signs, then afterwards this choice will be justified by looking at the trichotomies for each entity of the sign. In other words, after a category is chosen, then the relations between the sign and each entity in the sign will be discussed. This involves discussing and justifying the choice of relation between the sign and the representamen as a qualisign, a sinsign and a legisign; the object as an icon, an index and a symbol; and the interpretant as a rheme, a dicent sign and an argument. This order of argumentation is chosen because otherwise it would be possible to determine a set of trichotomies,

<sup>5</sup> The Library of Congress Cataloging-in-Publication Data classifies the book under, 1. Information Science. 2. Information Science – United States.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

| <i>Sign</i>         | <i>Best fit among Peirce's sign categories</i> |
|---------------------|--|
| Document            | Sign category X (Argument)                     |
| Subject             | Sign category IX (Dicent symbol)               |
| Subject description | Sign category VII (Dicent indexical legisign)  |

Figure 10. *Matches between categories of signs and elements*

for instance a sinsign, an index and an argument, for which Peirce had defined no category of signs.

3.3.1. *Peirce's categories of signs and the first element – the document* The document is the initial element in the subject indexing process and is therefore the first element that is analysed. A document represents a range of ideas and meanings, and it is the indexer's task to select from among these and capture them first in a more formalised subject description and ultimately in a subject entry. The document is a complex entity that involves many different statements. It is for that reason to be regarded as equal to Peirce's definition of the class of signs called argument<sup>6</sup> (category number X, in Figure 5).

The argument is the most complex of the ten categories of signs in the sense that it embeds all the other kinds of signs in the other nine categories. It is the most developed of the ten categories of signs and it involves the greatest amount of interpretation and uncertainty in its result. The reason for this is that the interpretation of the sign is based on being part of a certain social and cultural context. Peirce (1955, pp. 117–118) defined an argument as a kind of law in the following way: 'An Argument is a sign whose interpretant represents its object as being an ulterior sign through a law<sup>7</sup>, namely, the law that the passage from all such premises to such conclusions tends to the truth'. The interpretation of the arguments is therefore based on what Peirce called conventions. By this, he meant the agreements of practice or customs based on a general consent within specific social and cultural contexts. The document is therefore seen as a product of certain conventions, and its meanings are to be understood in a social and cultural context. The determination of the document's aboutness cannot be separated from this context, although there may be overlaps in determination of subject matter between different social and cultural contexts. However, this overlap is as much based on an overlap in conventions as determined by the particular document itself.

To justify this choice of category for the document, the document's relation to each of the entities of the sign will be elaborated in the following. A summary of the argumentation can be found in Figure 11 in which the choices of trichotomies for each of the three entities are underlined and connected with dotted lines.

<sup>6</sup> Peirce named the category 'Argument.' He did not use all three terms in the category – argument, symbolic, legisign – to identify the category; although this may cause confusion, Peirce's use of names for the categories will be followed here.

<sup>7</sup> The word 'law' here means customs, habits or practices.



All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

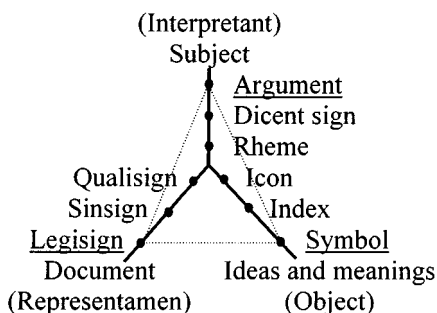
SEMIOTICS AND INDEXING

The tenth category of signs, arguments, is defined as those kinds of signs that are thirdness to all three entities of the sign. That is, the representamen is a legisign, the object is a symbol, and the interpretant is an argument; each of these represents thirdness. That category of signs best matches the document element in the subject indexing process. By investigating each of the three entities of the sign this match will become clear.

The representamen, or in this case the document itself, can either be a qualisign, a sinsign or a legisign depending on what kind of phenomenon the sign itself is. Qualisigns are mere qualities like colours and smell and sinsigns are specific objects like events or acts that only occur once. The legisign is a convention that gives meaning through laws, or habits, which are 'determined' culturally by humans. Accordingly, the document is a legisign, for it will be senseless to claim that the document is a mere quality or a once-occurring event.

The object, or in this case the range of ideas and meanings associated with the document, can either be an icon, an index or a symbol. The icon represents through likeness in the sense that there is some kind of likeness between the sign and that which the sign represents. An index refers to its object by a direct connection between the sign and that which it represents. The relation between the document and the range of ideas and meanings associated with it is not a relation of likeness or direct connection. It is therefore neither an icon nor an index. The symbol, on the other hand, refers to its object through a law, habit or convention and is based on the social and cultural context of the document. The document is therefore a symbol. In practice, this means that a document can have a different range of ideas and meanings, depending on social and cultural context.

The interpretant, or the understanding of the document's subject matter, can either be a rheme, a dicent sign or an argument. Rhemes are signs that are easily understood and understood equally by all humans, or at least by all humans within a certain social and cultural context. It is understood as representing a certain kind of possible object. A rheme could therefore be compared to a single noun, like 'cat'. The word 'cat' will give nearly the same information to all who speak English. A dicent sign is more complex than the rheme and as such requires more

Figure 11. *The document*<sup>8</sup>

<sup>8</sup> It could be confusing that the term 'argument' also appears in this figure. However, it should be noted that 'argument' in this figure is distinct from 'argument' in Figure 5 where it denotes a category of signs; in this figure it denotes a relation to the interpretant.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

knowledge to interpret and yield different understandings for different people at different times. Dicient signs could be compared to sentences, like 'the cat lies on the mat'. Here there is greater variability in the interpretation of the sign. However, the argument is the most complex relation between sign and interpretant. It is a sign of reason or law and can only be understood in relation to the interpreter's actions. It cannot be understood by itself and is as such highly dependent on the individual interpreter. The argument is thus to be compared to text where interpretation varies highly depending on the individual reader. The document element in the subject indexing process therefore is an argument.

In short, other sign categories could not be seriously considered because they used (1) rheme or dicient sign instead of argument, (2) qualisign or sinsign instead of legisign or (3) icon or index instead of symbol, and as already shown, none of these alternatives is appropriate.

The categorisation of the document in category X, argument, can be further elaborated and justified by referring back to the previous example of the document analysis of Arlene Taylor's book in Section 3.2.1. It was seen here how the indexer developed what was referred to as a collage of impressions of the book. This collage was developed by investigating different parts of the book and consulting external resources like the indexer's knowledge about library and information science, knowledge about the users of the information system etc. The book itself, the representamen, is a convention that is only understood through certain laws and habits related to the understanding of library and information science and the users of an information system etc. The book refers to a range of ideas and meanings associated with the book, the object of the sign; however, this reference is made neither through resemblance or likeness nor through direct connection with the object. The book refers to its object through laws, habits or social conventions and is therefore a symbol. In other words, Taylor's book belongs to sign category X, argument, because it will only be understood by referring to a certain set of laws, habits or conventions. The collage of impressions developed by the indexer depends as much on his or her knowledge as on the book itself.

*3.3.2 Peirce's categories of signs and the second element – the subject* The subject is generally only present in the mind of the indexer as something like a collage of impressions and ideas taken from the document. The indexer has not at this point distinguished between what should be represented and what should not. The subject is whatever the indexer associates with the document, the range of which could include, besides data from the document itself, such disparate things as other documents the indexer might think of, the fact that the indexer's brother could possibly make use of the document, or a discussion the indexer had with some friends on a previous day. In this respect, the subject is not solely something that has to do with the document and the text in the document. The subject has much to do with the person who reads, looks at and evaluates the document as well as the document itself. In other words, at this point it makes less sense to say that a document *has* a particular subject than to say that a subject is something a document is *given*. In sum, the subject of a document could be almost anything. In most cases, however, there are only a limited number of accepted and useful

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

interpretations of a document's subject matter, since the social practice in which the document will be used determines the use, meaning and subject matter of the document.

Social practice refers to the particular social context (including the domain, the users, the organisation etc.) of the indexing situation. This practice will limit the range of ideas and meanings the indexer will often associate with the document. That the indexer is aware of this serves in turn to limit the range of ideas and meanings to a smaller number than the entire total possible. One element of professionalism in indexing is an awareness of the particular setting in which the document should be used. The subject is categorised in Peirce's sign category number IX (Figure 5), which is the category of dicent symbols.<sup>9</sup>

A summary of the following argumentation for the choice of sign category IX for the subject can be found in Figure 12; the choices of trichotomies for each of the three entities are underlined and connected with dotted lines. As is seen in Figure 12 sign category IX, dicent symbol, is thirdness to the representamen (legisign) and the object (symbol), and secondness to the interpretant (dicent sign). This means that the dicent symbol is somewhat less complex than sign category X, argument, and it therefore involves a less complex kind of interpretation. This in turns means that interpreting the subject is less involved than interpreting the document. The choice is justified next by investigating each of the three entities of the sign.

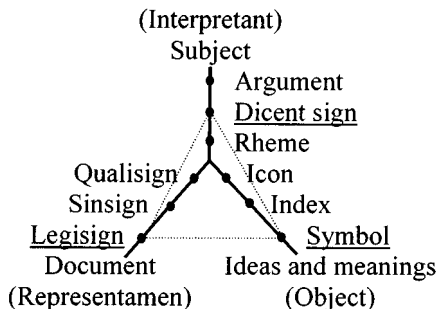


Figure 12. *The subject*

The representamen, or the subject, is a legisign. As already discussed in the previous section, where the document was categorised, legisigns are signs that represent through conventions. At this point the subject exists in a purely mental mode. However, the ideas and meanings associated with the subject are represented, through laws and conventions. It could be argued that the subject is a sinsign, since sinsigns are defined as actual existent things or events. However, one further implication of sinsigns is that they are embodiments of qualities of actual existences. Sinsigns, therefore, are signs of external, or physical, existence and the subject is therefore not a sinsign. As mentioned earlier, qualisigns are mere qualities, wherefore the subject cannot be a qualisign.

<sup>9</sup> It should be noted that the category is referred to as a 'dicent symbol', which is Peirce's denotation of the category. The category name 'dicent symbol' should not be confused with the dicent sign. The term dicent sign is used to denote a relation between the sign and the interpretant.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

The object, or the range of ideas and meanings associated with the subject matter, is a symbol. If the subject were considered an index, this would mean that the subject should point out its object or that it would be directly connected with its object. However, this does not happen; the object's ideas and meanings are represented through associations. The subject cannot be an icon, because icons represent through likeness, as mentioned in the previous section.

The interpretant, which in this case is the subject description, is a dicent sign. Peirce's definition of the dicent sign states that the dicent sign is a sentence or proposition. As was described in the previous section, an argument equals text and a rheme merely a noun. A subject description, being the interpretant of the subject, is therefore best matched with the dicent sign. For the indexer, the subject will only refer to a fraction of all the possible ideas and meanings, something that can be captured in a few sentences.

In short, other sign categories could not be seriously considered because they used (1) rheme or argument instead of dicent sign, (2) qualisign or sinsign instead of legisign, or (3) icon or index instead of symbol, and as already shown, none of these alternatives is appropriate.

This choice of category could further be illustrated by referring back to the indexing of Taylor's book the subject of which is the collage of impressions and ideas that the indexer associates with the book after the document analysis. This was discussed in Section 3.2.1, and in Section 3.2.2 how the indexer formulates a more condensed description of the book's subject matter based on that collage of impressions and ideas was discussed. The nature of the interpretation that is required for the subject is that which equals the ninth category of signs, dicent symbols.

### *3.3.3 Peirce's categories of signs and the third element – the subject description*

The subject description is a formalisation of the subject collage of impressions that resulted from the second step in the subject indexing process. It may be a written statement of the subject or simply a condensed implicit formulation of the subject matter. The essence of the formalisation resides in the conscious effort to reduce the various elements of the subject collage to a sensible statement. Thus, in the second step, in which the movement is from the subject to the subject description, the indexer decides which of the many topics and other elements associated with the document in the subject collage should be represented. Hence, the subject description will almost assuredly be considerably narrower in focus than the subject. Whereas the subject includes everything the indexer could possibly infer from and associate with the document, the subject description is limited to the information that the indexer concludes is worthwhile or important to represent in the indexing language. It is often recommended here that the indexer produce a series of concise statements that describe the subject. Limiting the description in this way will help the indexer to distinguish the different kinds of relations that comprise the subject and pick the most appropriate in the given situation.

The subject description, therefore, consists of a few sentences or statements that are made to represent the content of the document more formally. More

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

specifically, it represents the document as viewed by the indexer, and reflects those topics of the document the indexer has chosen to emphasise. In terms of Peirce's sign categories, the subject description is a dicent indexical legisign (sign category VII, Figure 5). A summary of the following argumentation for the choice of sign category VII for the subject description can be found in Figure 13; the choices of trichotomies for each of the three entities are underlined and connected with dotted lines.

As it can be seen in Figure 13, the subject description is only thirdness in its relation to the representamen (legisign), whereas it is secondness to the object (index) and the interpretant (dicent sign). This means that the interpretation involved with the subject description is less complex than with the document and the subject. The production of the interpretant, the subject entry, of the subject description will therefore be less dependent on the individual indexer than the interpretation of the previous two elements.

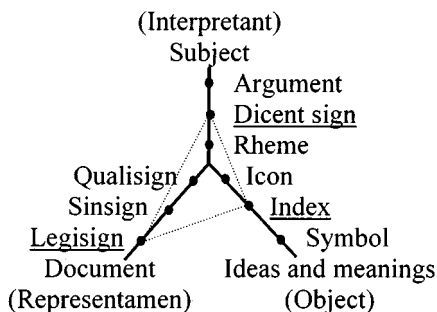


Figure 13. *The subject description*

The representamen, which in this case is the subject description of the subject matter, is, like the previous two elements, a legisign. Legisigns represent through conventions, and the subject description must be regarded as such. The subject representation is not a sinsign because it is not a sign of actual existence and it is not a quality and therefore not a qualisign. Rather the subject description is a sign that represents through conventions.

The object, which in this case is the range of ideas and meanings associated with the subject description, is an index. The collage of impressions and ideas of the subject produced in the first step of the subject indexing process limits the range of ideas and meanings associated with the subject description. In other words, the subject collage that the indexer produced directly affects the interpretation of the subject description. Even if the indexer actually wrote down a few sentences that condensed the subject collage to a subject description, the range of ideas and meaning that the indexer associates with the subject collage directly affects the interpretation of these sentences of the subject description. The definition of an index includes that it points to its object and is affected by it. The distinction between a symbol and an index is important here. Symbols are interpreted as referring to their objects and thereby operate through laws and conventions. Indexes, on the other hand, are closely associated with their objects, since they are affected by them. Indexes point out their objects, whereas symbols

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

are interpreted as merely referring to their objects. Even though it could be argued that the subject description under some circumstances might be regarded as a symbol – for example, when one indexer has analysed the document for its subject matter and expressed this in a subject description and another indexer converts the subject description into the subject entry. In such a case the subject description is not affected by its object and is therefore a symbol belonging to sign category IX, the dicent symbol. However, it is maintained here that the subject description is an index. The subject description is not an icon, since icons are defined as likeness, and the subject and subject description are not alike, but different in the sense that the subject description is a condensed version of the subject collage.

The interpretant of the subject description, which is the subject entry, is a dicent sign. The three possible trichotomies for the interpretant are the rheme, which is a sign that is easily understood and interpreted by most humans; the dicent sign, which is a more complex sign and involves more knowledge of the interpreter; and the argument, which is the most complex sign or a sign of reason or law. The three trichotomies can be exemplified by a noun (rheme), a sentence (dicent sign) and a text (argument). The subject description, formalised in a few sentences, best matches the dicent sign.

In short, other sign categories could not be seriously considered because they used (1) rheme or argument instead of dicent sign, (2) qualisign or sinsign instead of legisign, or (3) icon or symbol instead of index, and as already shown, none of these alternatives is appropriate.

The choice of sign category for the subject description could further be illustrated by the indexing of Taylor's book. It was shown in Section 3.2.2 how the subject description could vary according to the focus or point of view the indexer chose to adopt. Two examples of subject descriptions were given, one that focused on the broadest theme in the book, namely the introduction of the fundamentals of knowledge organisation for the novice information scientist. The other description focused narrowly on the book as an introduction to library cataloguing and classification. These two themes were converted to Library of Congress Subject Headings in Section 3.2.3. It was shown in Section 3.2.3 how the indexer interprets the subject description to create a subject entry. This interpretation was constrained by the object of the sign, the range of ideas and meanings associated with the subject description. Among these ideas and meaning were the indexer's initial subject collage of impressions of the document, but also knowledge about the users and about the indexing system itself. The indexer's final interpretation was therefore based as much on his or her personal knowledge as on the subject description itself. In other words, the interpretation of the subject description for Taylor's book as being about, for example, 'information retrieval' or 'classification – books' is equally good and equally correct. However, whether one or the other interpretation is found best suited for the representation of Taylor's book depends much more on the social and cultural context of the indexing than on the book itself.

*3.3.4 Peirce's categories of signs and the fourth element – the subject entry* The fourth, and last, element of the subject process is the subject entry. The subject

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

entry could be a verbal indexing term or it could be a notation in a classification system. The exact form of the subject entry will depend on the system for which it has been created. The categorisation of the subject entry in Peirce's categories of signs depends on the exact situation in which the subject entry is interpreted and on who interprets the subject entry. The exact categorisation of the subject entry in Peirce's categories of signs is therefore beyond the scope of this paper which is limited to the subject indexing process in which the subject entry is the end product. An analysis of the subject entry, therefore, belongs to a study of information seeking.

However, a brief discussion of the categorisation of the subject entry will be valuable here to show the range of possible categorisations.

The interpretation of the subject entry will depend on various things, such as the interpreter's knowledge of the indexing system and of its users, and the domain of the system and documents etc. If the user of the indexing system is very familiar with it then the interpretation will be less involved and the interpretant will be a rheme, because the subject entry will be interpreted as easily as a noun and almost alike for all users in the same situation. The object will in this case be an index that simply points out the object of the subject entry or range of ideas or meanings associated with it. This situation would be the case, for instance, for professional librarians who have worked with a specific indexing system for a number of years. On the other hand, a novice user of an indexing system will interpret the subject entry in a different way. In such a case the interpretation of the subject entry will be more complex. The interpretant will be a dicent sign, because the user had to interpret the subject entry as statement or sentence of the content of a class or indexing term. The object, the range of ideas and meanings associated with the subject entry will be a symbol, because the subject entry will be a convention based on the social and cultural context of the subject entry.

The possible subject entries for Taylor's book might serve to illustrate this point. A professional user of the LCSH system would know that subject headings such as 'information retrieval' and 'classification – books' cover many other topics than those that the headings imply. The professional user would therefore look under both of these and a number of other headings in a search for a book such as Taylor's. The novice user of the LCSH, on the other hand, would interpret the subject headings as a statement of what could be found under these headings and would not off hand know that the LCSH lacks a heading that covers the focus of Taylor's book and would therefore, perhaps, not look under these headings for the book. He or she might look for the topic under a broader heading such as 'information science', which he or she might interpret to cover books with a broad focus on knowledge organisation and system design.

One obvious conclusion that may be derived from applying Peirce's semiotics to the subject indexing process is to demonstrate how fundamentally interpretative and, therefore, variable, the entire process is. To portray the process in this way and to make a point of saying that it is a useful conclusion should not be seen as an attempt to demean the process as something that will not yield itself to precision and exactness. Rather it is a way of showing how inexplicably profound and human the process is. Indeed, it is the profoundly human nature of the subject



All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

JOURNAL OF DOCUMENTATION

vol. 57, no. 5

indexing process that makes it so impervious to analysis solely by quantitative empirical methods on the one hand, and so demanding of the need for qualitative and humanistic approaches to understanding it on the other.

#### 4. CONCLUSIONS

One of the central problems in the construction and use of systems for representation and retrieval of documents is how the subject matter of the documents is captured and represented. Although technological advances might provide means for fast access to documents, the problem of representing the subject matter of documents is first and foremost a problem of how people interpret and understand documents. The core of the problem is to understand how the interpretation of a document's subject matter can be represented to make it possible for others to gain access to the document. Indexing, therefore, is concerned with meaning and language – and any theory of indexing explicitly or implicitly includes a theory of language and meaning.

The understanding of the nature of the subject indexing process is more critical now than ever. Indexing of documents is being employed in settings far beyond the traditional library; it is, for instance, now being discussed in organisations that implement intranets and knowledge management. In reality, however, the documents are often made accessible both through human-selected and automatically generated index terms. A clear understanding and a sound theory of the human subject indexing process is, therefore, as important as ever.

The semiotic analysis of the subject indexing process demonstrates that the first step, the analysis of the document, involves a kind of interpretation that is highly dependent upon the social and cultural context of the indexer and the indexing process. Although this dependency has been discussed previously in the literature, this study demonstrates how inescapable the interpretation of the document is. Not only is the initial interpretation inescapable, the entire indexing process is made up of multiple inescapable interpretations. In this sense, the study of indexing is the study of documents and how documents are used.

Furthermore, since the representation of documents is the first in a series of activities that provide the user with requested documents, a clear understanding of the nature of the process is central to many activities and studies within library and information science. In other words, any study of information seeking, information retrieval, evaluation of information systems and so on should take the fundamental and inescapable interpretative nature of the subject indexing process into account.

#### ACKNOWLEDGEMENTS

I would like to thank Francis Miksa, University of Texas at Austin, Richard Smiraglia, Long Island University and Hanne Albrechtsen, Risø National Laboratory, Denmark for comments and suggestions to earlier versions of this paper. Thanks also to Gorm Larsen and Peder Voetmann Christiansen for permission to reuse their figures.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

September 2001

SEMIOTICS AND INDEXING

## REFERENCES

- Benediktsson, D. (1989). Hermeneutics: dimensions toward LIS thinking. *Library and Information Science Research*, 11, 201–234.
- Blair, D. (1990). *Language and representation in information retrieval*. New York: Elsevier Science.
- Brier, S. (1996). Cybersemiotics: a new interdisciplinary development applied to the problems of knowledge organisation and document retrieval in information science. *Journal of Documentation*, 52(3), 296–344.
- Buckland, M.K. (1997). What is a 'document'? *JASIS*, 48(9), 804–809.
- Buckland, M.K. & Day, R. (1997). The semiotics of 'document' and the antelope of Suzanne Briet. In: *Semiotics around the world: synthesis in diversity. Proceedings of the Fifth International Congress of the International Association for Semiotics Studies, Berkeley, June 12–18, 1994*. Mouton de Gruyter, 1203–1206.
- Christiansen, P.V. (1988). *Charles S. Peirce: Musten og Mørtel til en Metafysik*. Roskilde, Denmark: IMFUFA.
- Chu, C.M. & O'Brien, A. (1993). Subject analysis: the first critical stages in indexing. *Journal of Information Science*, 19, 439–454.
- Cronin, B. (2000). Semiotics and evaluative bibliometrics. *Journal of Documentation*, 56(4), 440–453.
- Eco, U. (1984). *Semiotics and the philosophy of language*. London: Macmillan.
- Fairthorne, R.A. (1969). Content analysis, specification, and control. *ARIST*, 4, 73–109.
- Farrow, J.F. (1991). A cognitive process model of document indexing. *Journal of Documentation*, 47(2), 149–166.
- Fisch, M.H. (1986). *Peirce, semeiotic, and pragmatism: essays*. Bloomington: Indiana University Press.
- Frohmann, B. (1990). Rules of indexing: a critique of mentalism in information retrieval theory. *Journal of Documentation*, 46(2), 81–101.
- Hoopes, J. (1991). Introduction. In: *Peirce on signs: writings on semiotics by Charles Sanders Peirce*. Chapel Hill: University of North Carolina Press, 1–13.
- ISO. (1985). *Documentation – methods for examining documents, determining their subjects and selecting indexing terms*. Geneva: International Organization for Standardization.
- Johansen, J.D. (1985). Prolegomena to a semiotic theory of text interpretation. *Semiotica*, 57, 3/4, 225–288.
- Johansen, J.D. (1993). *Dialogic semiosis: an essay on signs and meaning*. Bloomington: Indiana University Press.
- Karamüftüoğlu, M. (1996). Semiotics of documentary information retrieval systems. In: Ingwersen, P. & Pors, N.O., eds. *Information science: integration in perspective. Proceedings [of] CoLIS 2, Second International Conference on Conceptions of Library and Information Science, October 13 – 16, 1996*. Copenhagen: The Royal School of Librarianship, 85–97.
- Langridge, D.W. (1989). *Subject analysis: principles and procedures*. London: Bowker-Saur.
- Larsen, G. (1993). System og semiosis – eller forholdet mellem semiologi og semiotik. In: *Tegn og Data. En semiotisk tilgang til humanistisk datalogi*. Aalborg, Denmark: Aalborg Universitetsforlag, 139–160.

All rights reserved. Except as otherwise permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise without the prior written permission of the publisher.

JOURNAL OF DOCUMENTATION

vol. 57, no. 5

- Mai, J-E. (1999). Deconstructing the indexing process. *Advances in Librarianship*, 23, 269–298.
- Marty, R. (1982). C.S. Peirce's phaneroscopy and semiotics. *Semiotica*, 41(1/4), 169–181.
- Merrell, F. (1997). Do we really need Peirce's whole decalogue of signs? *Semiotica*, 114(3/4), 193–286.
- Miksa, F. (1983). *The subject in the dictionary catalog from Cutter to the present*. Chicago: American Library Association.
- Ogden, C.K. & Richards, I.A. (1923). *The meaning of meaning*. London: Kegan Paul.
- Pearson, C. (1980). The basic concept of the sign. In: *Proceedings of the 43rd ASIS Annual Meeting*, 17, 367–369.
- Pearson, C. & Slamecka, V. (1977). *Semiotic foundation for information science. Final Project Report*. Atlanta, Ga: School of Library and Information Science, Georgia Institute of Technology. (National Science Foundation Grant GN-40952)
- Peirce, C.S. (1955). *Philosophical writings of Peirce*. New York: Dover Publications.
- Peirce, C.S. (1958). *Charles S. Peirce. Selected writings (values in a universe of chance)*. New York: Dover Publications.
- Petersen, T. (1994). Introduction. In: Petersen, T., ed. *Guide to indexing and cataloging with the Art and Architecture Thesaurus*. New York: Oxford University Press, xiii–xvi.
- Saussure, F. de. (1966). *Course in general linguistics*. New York: McGraw-Hill.
- Seboek, T.A. (1994). *Signs: an introduction to semiotics*. Toronto: University of Toronto Press.
- Smiraglia, R.P. (2000). Works as signs and canons: toward an epistemology of the work. In: Beghtol, C., Howarth, L.C. & Williamson, N.J., eds. *Dynamism and stability in knowledge organization*. Würzburg, Germany: Ergon Verlag, 295–300. (Advances in Knowledge Organization, vol. 7)
- Taylor, A.G. (1994). Books and other bibliographic material. In: Petersen, T., ed. *Guide to indexing and cataloging with the Art and Architecture Thesaurus*. New York: Oxford University Press, 101–119.
- Taylor, A.G. (1999). *The organization of information*. Englewood, Colo: Libraries Unlimited.
- Wagner, G. (1991). Semiotics: a tool of communication research in public libraries. *Libri*, 41(3), 207–215.
- Warner, J. (1990). Semiotics, information science, documents and computers. *Journal of Documentation*, 46(1), 16–32.
- Weiss, P. & Burks, A.W. (1945). Peirce's sixty-six signs. *Journal of Philosophy*, 42, 383–388.
- Wittgenstein, L. (1958). *Philosophical investigations*. New York: Macmillan Publishing.
- Wynar, B.S. (2000). *Introduction to cataloging and classification*. 9th ed. by Arlene G. Taylor. Englewood, Colo: Libraries Unlimited.

(Revised version received 17 April, 2001)