



The concept of "subject" in Information Science

Item Type	Journal Article (Paginated)
Authors	Hjørland, Birger
Citation	The concept of "subject" in Information Science 1992, 48(2):172-200 Journal of Documentation
Journal	Journal of Documentation
Download date	01/04/2020 06:19:23
Link to Item	http://hdl.handle.net/10150/105962

THE CONCEPT OF 'SUBJECT' IN INFORMATION SCIENCE

BIRGER HJØRLAND

*Royal School of Librarianship, Danmarks Biblioteksskole
Birketinget 6, DK-2300 Copenhagen S*

This article presents a theoretical investigation of the concept of 'subject' or 'subject matter' in library and information science. Most conceptions of 'subject' in the literature are not explicit but implicit. Various indexing and classification theories, including automatic indexing and citation indexing, have their own more or less implicit concepts of subject. This fact puts the emphasis on making the implicit theories of 'subject matter' explicit as the first step.

A very close connection exists between what subjects are, and how we are to know them. Those researchers who place the subjects in the minds of the users have a conception of 'subject' different to that possessed by those who regard the subject as a fixed property of the documents. The key to the definition of the concept of 'subject' lies in the epistemological investigation of how we are going to know what we need to know about documents in order to describe them in a way which facilitates information retrieval. The second step therefore is an analysis of the implicit epistemological conceptions in the major existing conceptions of 'subject'. The different conceptions of 'subject' can therefore be classified into epistemological positions, e.g. 'subjective idealism' (or the empiric/positivistic viewpoint), 'objective idealism' (the rationalistic viewpoint), 'pragmatism' and 'materialism/realism'. The third and final step is to propose a new theory of subject matter based on an explicit theory of knowledge. In this article this is done from the point of view of a realistic/materialistic epistemology. From this standpoint the subject of a document is defined as the epistemological potentials of that document.

1. THE NAIVE CONCEPTION OF SUBJECT

FROM A NAIVE POINT OF VIEW the concept of 'subject' or 'subject matter' poses no problem: it is rather obvious what subjects are. The book *General psychology* has quite naturally the subject 'psychology', and the *Cambridge history of England* has 'history' as its subject, which can be further subdivided if one wishes to do so into 'world history' and the 'history of England'.

A slightly less naive viewpoint would recognise that there need not be a correspondence between, for example, the title of a book and its actual 'subject'. Not all handbooks (for example 'Handbook of psychology') use this term in their titles, nor do all such titles necessarily correspond to the user's view of the content of the book. Authors with a background in one particular

discipline (for example psychology, psychiatry or sociology) may have a tendency to give their works titles which name their own discipline, even though the contents of the works might just as easily justify mention of another field. 'A history of dynamic psychiatry' could also properly be titled 'A history of dynamic psychology', and what is its actual subject? The naive viewpoint has run into difficulties!

The naive point of view corresponds in part to a child's lack of differentiation between linguistic forms and meanings. It is apparently typical of a primitive perception of language that a word and its phonetic construction are viewed as an attribute of the thing itself which cannot be separated from its other characteristics (cf. Vygotsky [1, 358-359].) The naive person typically views a subject as a part of, for example, a book's attributes, a concentration as it were of what is stated in its title and which cannot be separated from the other attributes of the book. This attitude is in a way related to the philosophical concept *naive realism* (according to which the experience of the senses provides direct access to reality: the naive realist, for example, sees that the stars are smaller than the moon, and therefore assumes that they are smaller).

A more detailed characterisation, scrutiny or investigation of the naive conception of the concept of subject requires that we ourselves have attained a solid conception of subject, which is the purpose of this work.

2. SUBJECTIVE IDEALISM

Idealism is a fundamental concept in philosophy, of which the main characteristic is that the mental process or consciousness is viewed as primary, or determining, in relation to reality or the material world. In opposition to idealism are the different varieties of realistic or materialistic philosophy, in which the mental is conceived of as something secondary, or derived, in relation to reality or the material world. Some researchers and philosophers are proclaimed idealists, but it is far more common that researchers do not consider themselves to be idealists, nor do they assume a consciously idealistic point of departure (and, for example, view the clash between idealism and materialism as an irrelevant issue), but in their thinking inadvertently fall into idealistic modes of thought. In the field of library and information science this is indeed the case, for example, as regards the concept of 'subject matter'. A worthwhile critique of mentalistic (and thereby idealistic) tendencies in 'information retrieval' theory has recently been published by Frohmann [2]. My own attempts at clarification of information science are in definitive ways identical to Frohmann's point of departure.

An idealistic concept of subject matter encompasses that a 'subject' is an 'idea', either in an objective (i.e. Platonic) sense, or in a more subjective sense. In this section we will look more closely at subjective-idealistic concepts of 'subject'; in the next section, the objective-idealistic will be considered.

Subjective idealism takes concepts and subjects to be the expression of the perceptions or views of one or more individuals (subjects). Concepts and

subjects are that which is subjectively comprehended or understood by them. The key to the concept of subject therefore lies in the study of the minds of some people, for example, the authors or users of documents. From the point of view of epistemology, subjective idealism is characterised by making perception and thinking independent in a subjectivistic manner. Positivism is the most common representative of subjective idealism.

If the issue is the subject matter of a book, there are many possibilities: the author's version (often as expressed in the title or the text, either implicitly or explicitly), the reader's version (great variation is possible here), the publisher's version, as often indicated in a series title (for example 'European Monographs in Social Psychology'), and the librarian's version, which may well be expressed in terms of the library's classification.

Bente Ahlers Møller [3] has published a brief paper in which she compares classification of the same books by the system used at the State and University Library in Aarhus, Denmark, with the Dewey Decimal classification. This demonstrates that there can be amazing differences between subjective perceptions of what the subjects of the books are. But this subjectivity may well be extremely well-founded: *subjectivity is not noise or error, it is a consistent and thoroughly underpinned analytical tendency*. We are not merely speaking of the different structures which different classification systems give to subjects (i.e. more or less subdivision), but unequivocal differences in the conception of the subject of a book, where one view places a book under the subject 'books', and another view places the same book under the subject 'trade'.

In connection with subjective idealism special consideration is given to the intentions of the author, his view of its subject, and what new things he has to relate. This has given rise to the concept of 'aboutness' in library and information science literature, an interest which in my view represents a blind alley, an attempt to escape from the difficulties in the concept of subject (Note 1). Devotees of the concept 'aboutness' assign to it special clarity and significance in the analysis of subjects, but are evidently unaware of its epistemological position as subjective-idealistic.

With regard to the subjective-idealistic theory of 'subject matter' I will demonstrate that neither the author's, the reader's, librarian's/information specialist's nor any other person's (for example the publisher's) points of view or subjective understanding can have any certain or objective knowledge about the subject of a document, nor define the concept of 'subject'. Each of these viewpoints can contribute something to a determination of the subject, but the subjective-idealistic conception of subject over-emphasises certain aspects of the document either from the author's, the reader's or an interpreter's point of view.

1. A book can – but need not – contain an assertion of what its subject is. The *author* can explicitly discuss the subject of his work, for example in the introduction, and he may note its relation to other subjects. If a book is called 'general psychology' it may contain a discussion of 'what is general psychology?'. Since the basis of psychology is a complex theoretical problem, the author's views need naturally not be true, merely the

expression of his more or less well-founded (subjective) ideas. That which is psychology for some may – after theoretical considerations – prove rather to be sociology or physiology. The book may not deal at all with that which the author thinks it does, nor with what the title indicates.

Just as often, however, a work does not contain any explicit discussion of its subject. 'The history of dynamic psychiatry' assumes implicitly that psychoanalysis is part of medical science (psychiatry) and not of psychology. Much can be said about this, but the given label of a given book need not be correct. A book need not treat the subject of psychiatry because it says it does.

A truly scientific analysis of the subjects of documents for databases would have to assume certain consistent definitions, which would sometimes, but by no means always, be in agreement with the version of the subject given in the document itself.

2. With regard to the *user*, a document can be ordered with the user's conceptual structures and subject perceptions in mind. The user may well have his subjective grasp of what the subject of the book is.

Some information retrieval theorists appear to work from the premise that an information retrieval system ought to order subjects according to each user's subjective reading. They are inclined to build on psychological investigations of the users' perceptions of the subject, their 'knowledge structures'. There are also examples of investigations carried out on such a basis (Mark Pejtersen [4, 5] is clearly an example of this). A related mode of consideration is, for example, Belkin's ASK model [6–8]. Though J.E. Farradane [9, 10] assumes an explicit psychological approach within library and information science literature, a closer interpretation of his work seems to imply more an objective than a subjective-idealistic model.

We claim that there are types of information systems which clearly ought to aim to tailor the description of the subjects to the user's subjective perceptions. Examples of this are library systems for children or pedagogical systems in which a point of departure and a goal can be described for both a learning process and for advising students. Both types express a certain *paternalism*, i.e. someone assumes the responsibility for the direction of others' information searches. This is done by presuming to create the connections between given documents and the user's subject universe, i.e. undertaking to interpret the subjects or information content of the documents from a psychological or pedagogical evaluation of needs and goals.

Aside from such paternalistic approaches, should subject descriptions then take the psychology of the user into account? Yes, in a certain way this is indeed desirable. Information retrieval systems should be made user-friendly, and this can be done by having knowledge of the user's language and subjective perceptions, and use this knowledge, for example in *see* references to the preferred terms. So perhaps it is even the ideal, that all systems in a certain way relate to the users. But this does not mean that one interprets the subject content of documents based on knowledge of the users' subjective perceptions, but that these perceptions are employed to create the necessary references and instructions, i.e. to make the system user-friendly. In my opinion the question of user-friendliness is not *the* central theoretical issue in information retrieval. The central issue is knowledge-representation, how to represent the knowledge in documents. The question of user-friendliness is a cognitive-ergonomic question that must be implemented in a system, but is of secondary interest compared to the adequate representation of knowledge in databases.

Scientific information systems must in my opinion presuppose that the user acquires the categories, terminology and classifications of science, scholarship and information systems, rather than the reverse. The adoption of the user's categories and terminology by science and its information systems is a job for popularisation, not primarily for information science. Reference is often made to using the principles of psychology and linguistics for system design, but such principles often present dilemmas or contradictions in contrast to purely disciplinary considerations. Our conclusion here is

that he who seeks the key to the concept of 'subject' in the mind of the user commits an error of psychologism.

3. A third subjective conception can be expressed by the *librarian* or *information specialist* in a subject description of documents in a database. In the best instances a system is used (of classification, a thesaurus or something else) which makes possible a high degree of explicit and consistent basis for analysis. As demonstrated (for example in Møller [3]), different systems employ different (subjective) principles of analysis and thereby determinations of subjects. This situation will not be further documented here, since it makes up a significant part of the argument in the section on materialistic theory of subject matter. I will here merely establish that both the individual information worker and the different IR systems display considerable variations in their descriptions of the subjects of given documents. To the extent that this subjectivity is made a quality of the subject concept itself, I am talking about a subjective-idealistic conception.

Thus it is typical of the subjective-idealistic conception of subject that it over-emphasises certain aspects of the document either from the author's, the reader's or an interpreter's point of view. Insofar as no subjective instance in its role relative to the document can guarantee a correct analysis of subject matter, that analysis always is subjective, this can lead to an *agnostic conception of 'subject'*: it is impossible to say what a subject is, and how it is to be determined. Such a view has been expressed by Patrick Wilson [11].

Patrick Wilson investigates – especially via thought experiments – the suitability of different methods of determining the subject of a document. Among these methods are 1. to identify the author's purpose in writing the document, 2. to weigh the relative dominance and subordination of different elements in the picture given by reading the document, 3. to group or count the document's use of concepts and references and 4. to invent a set of rules of selection for what are the 'essential' elements (in contrast to the inessential) of the document in its entirety. Patrick Wilson demonstrates convincingly that each of these methods by itself is insufficient to determine the subject of a document, and concludes: 'the notion of the subject of a writing is indeterminate ...' (p. 89); or (on what a user can expect to find under a particular position in a library's classification system): 'for nothing definite *can* be expected of the things found at any given position' (p. 92). In connection with this last remark Wilson includes an interesting footnote, in which he directs attention to the often imprecise use made of concepts by the authors of documents ('hostility' is mentioned as an example). Even though the librarian personally might attain a very precise grasp of a concept, he will be unable to make use of it in his classification since none of the documents use the concept in the same precise way. Therefore Wilson concludes: 'if people write on what are for them ill-defined phenomena, a correct description of their subjects must reflect the ill-definedness'.

Renouncing an exact determination of one of the basic concepts of library and information science is a questionable matter. We do not think that such agnosticism as Patrick Wilson expresses in the above citations is an acceptable solution. As we shall see later, it *is* possible to define subjects. But it is not possible to determine subjects by examining the minds of authors, users or any other specific group of people. To do this would be a kind of 'mentalism'.

Attempts to move beyond this raise the question: what are the objective criteria for the subject of a document? If subjects are not perceptions or 'ideas' in some people's minds, what else can they be? What is to be understood by the statement 'document A belongs to subject category X'?

3. OBJECTIVE IDEALISM

The subjective-idealistic subject theory viewed subjects as subjective categories, for which person X and person Y each had his own subjective grasp of the subject of a given document. (These subjective categories may be more or less identical – this is another issue; the principle is that they are individual, dependent on a subjective conception.)

Objective idealism does not consider a subject as subjective in this way: persons X and Y will – if they perform correct analysis – arrive at the same subject for a given document, the subject of which can then be termed objective (at least in a particular meaning of the word). Whereas subjective idealism in general is characterised by over-emphasis on the perceptions of the senses, objective idealism tends to over-emphasise certain aspects of theoretical analysis and make them absolute.

The idealistic conception indicates that a subject is a designation of an *idea*. In Ranganathan's system this is made explicit, as cited by one of his students, Gopinath: 'Subject – an organised body of ideas, whose extension and intension are likely to fall coherently within the field of interests and comfortably within the intellectual competence and the field of inevitable specialisation of a normal person'; and: 'A subject is an organised and systematised body of ideas. It may consist of one idea or a combination of several ...' [12]. This comes very close to Ranganathan's own conception, even though he often avoids the problem, as in *Documentation and its facets* [13, p. 27], where he declares the subject to be an 'assumed term'.

To elucidate more closely the view which objective idealism takes of the concept of the subject, we will start by looking at its view of concepts in general. Objective idealism (as represented, for example, by Plato or scholastic realism) considers a concept to be an abstract psychic or mental entity (an *idea*), which exists in and of itself, and the relationship of this to concrete things is such that these things share in the mental entities which represent them via the concept. Realism (in the above meaning) considers, in other words, that general concepts represent something universal, which exists outside and independent of the human consciousness, and which at the same time exists prior to separate things (originally with reference to God, today rather a form of *a priori* cognition in a Kantian sense).

Translated into the terms of the problem of 'subject', this means that the concrete documents share in the 'ideas' expressed in a given subject. These ideas exist outside the human consciousness (or within it as *a priori* perceptions) and are also prior to the individual concepts expressed in the individual documents. These ideas or subjects have universal or fixed properties; they can once and for all be analysed in a universal system, or separated into individual parts.

This theoretical point of departure still has a far-reaching influence in today's theories about subjects which can be traced from the views of Ranganathan [12], Tranekjær Rasmussen [14, p. 26] following the Danish philosopher Harald Høffding, Thomas Johansen [15–19] and others on subject as an idea which can be analysed in its individual parts.

Ranganathan's 'Colon Classification' is discussed in an article by Gopinath, in which he states [12, p. 60]:

2.7 Absolute syntax of ideas

a subject is largely the product of human thinking. It presents an organized pattern of *ideas* created by the specialists in any field of inquiry. Working at the near-seminal level and postulating about helpful sequence among the facets and isolates *has led to the conjecture that there may be an 'absolute syntax' among the constituents of the subjects within a basic subject, perhaps parallel to the sequence of thought process itself, irrespective of the language in which the ideas may be expressed, irrespective of the cultural background or other differences in the environments in which the specialists, as creators as well as the users of the subject, may be placed . . .* (emphasis added).

This view, that human thought, human language, human consciousness, the human subject universe has an 'absolute syntax', i.e. that it is fundamentally independent of the functional context of the mental processes, is a pattern of the idealistic conception, a direct contrast to the view that the mental processes are tools, formed by and suited to the tasks and conditions in which they function. Since there is no question of person X and person Y having different 'syntax', this is an objective, not a subjective, idealism.

Objective idealism expresses itself in its classification process with the view that classification of documents can be done independently of the context in which classification is being used. The 'syntax' in Ranganathan's system is the PMEST formula (Personality, Matter, Energy, Space, Time). Gopinath [12, p. 60] gives an example of the analysis of a document. The subject 'exercise of franchise by the Indian citizen in 1960s' is analysed as follows in the Colon system:

History (basic subject)
Indian community [Personality round 1, level 1]
Citizen [Personality round 1, level 2]
Franchise [Matter round 1, level 2]
Exercise [Energy round 1]
1960s [Time level 1]

It is my claim that this type of analysis, which determines the priorities of the viewpoints to be taken on a document, is not optimal in every situation. One can imagine researchers working on technical aspects of the election process who wish to compare them in several countries. For such a person the election would be the central subject, and it would be inconvenient if this were a sub-topic of History and India. (Computer searching has to a large degree made fixed sequences among facets superfluous; the problem only remains for printed catalogues and other one-dimensional ordering systems, but that is another issue.)

It is indeed our claim that an objective idealist concept of subject matter tends towards subject descriptions which only have an abstract relationship to

the needs for subject description and the contexts in which they are used, because such descriptions are based on the *a priori* given properties of ideas. One can also express this as that subjects are viewed as 'innate properties' in things or documents. This is a consequence of the theory's concept of objective ideas, separated from the individual items of reality. In other words, this is also an expression of objective idealism's special conception of the relationship between the general and the particular: that the general exists outside and independent of the particular. This is in contrast to the concept that a subject only exists in specific documents, and that every subject description contains an analysis with its point of departure within the very contexts of its use, which is to be examined more closely below. *The idealistic concept of 'subject' has furthermore the consequence that neither the world views nor the academic disciplinary and political priorities expressed in information systems are recognised*, which has been criticised by Steiger [20], among others.

To sum up: the objective-idealistic point of view does not – as did the subjective-idealistic viewpoint – match the concept of subject in the minds of some people. Instead it presumes that some kind of abstract analysis or fixed procedure could be used to penetrate the surface of documents, thereby revealing their true subjects. As we shall see later, no such fixed procedure can guarantee a correct subject analysis. Among other things, this approach lacks consideration of the pragmatic aspects of subjects: the potential use of the documents.

4. PRAGMATIC CONCEPT OF SUBJECT MATTER

A user has a particular (specific) need for information, a problem to be solved for which information is required. This information is searched for in libraries or databases in which documents (carriers/conveyors of information) are registered by *subject*.

The registration of subjects by librarians or information specialists must – for the process to be meaningful – anticipate the needs of the user: it must make it possible for the user to find that for which he is searching. Subject data in libraries and information systems have an instrumental or pragmatic function. As Bookstein and Swanson [21] write: 'documents are indexed for the purpose of retrieval, and one can arrive at a theoretically well-founded procedure for indexing by being true to that purpose'.

Dagobert Soergel [22] has introduced a distinction between 'content oriented indexing' and 'request oriented indexing' which has proved most stimulating in my philosophising on the concept of subject. Whether Soergel really invented 'request oriented indexing' or just the name has not been investigated here. He points out that it is only the first of these which is described in library and information science literature, and that the second is hardly known in theory, though examples do exist in practice (for example the database Ringdok, which describes chemical literature in a different way from Chemical Abstracts, because Ringdok pays special attention to the needs of the pharmaceutical industry).

Content-oriented indexing is a description of subjects which has to be conceived as purely a function of the attributes of the document: as in the observation that 'this document contains the chemical formula for sulphuric acid' (and the consequent categorisation such as 'inorganic chemistry').

User-oriented or need-oriented indexing is a description of a subject which must be perceived as the relation between the properties of a document and a real or anticipated user need. 'This document deals with sulphuric acid. Sulphuric acid corrodes. Signmakers need corrosive agents' – thus follows a categorisation, for example, 'Literature on chemicals for use in signmaking'. *Need-oriented indexing is an instrumental (means-goals) relation between a document and a user need.*

Within information science aids such as *Science Citation Index*, *Social Science Citation Index* and *Atlas of Science* (all published by the Institute of Scientific Information in Philadelphia) provide links between subjects or the categorising of documents on the basis of a prior purely instrumental or means-goal relationship: the documents cited by the same document are assumed to be related in subject, since they have all contributed to the results of the document in question. In other words, these atlases (or the concept of bibliometric linking and co-citation) are implicit expressions of a concept of 'subject' in which a prior factual instrumental relationship (as reflected in citation practice) provides the basis of definition.

Bibliometric linking etc. is one method of searching for literature which has taken its place in the system, and which has its advantages and disadvantages. It occupies a niche: *it is not a question of merely mapping such prior instrumental connections and thereby producing a patent medicine for literature searching, nor reducing the concept of the subject to these empirical relationships.*

Several reasons play a role in this. First, a potential instrumental relationship cannot be extracted from a prior instrumental relation. In information science, the literature about 'telecommunications' may be linked (co-cited) with the literature on 'information retrieval', because telecommunications at a certain stage of development was a crucial problem for information retrieval. But at a later time, the problems of telecommunications may be regarded as trivial, and this bibliographic linking may be a bad expression of 'subject-relatedness'. Second, certain conditions, cultural or sociological within the research environment, skew the picture, insofar as the most epistemologically fertile documents are often not cited as much as those documents which easily lead to concrete investigations (that is to say, there is overemphasis on empiricism). A third and final reason is that a particular document most often contains essentially different types of information which it is useful to categorise in other ways from that to which a purely use-oriented practice would lead. For example, many psychological investigations cite statistical and methodological literature as well as literature of psychological substance. It would be expedient to operate with these as different subjects, even though they appear together (through bibliometric links) within the psychological literature of a given period.

Pragmatic subject theory runs into other difficulties: if it is assumed that a given document is to be included in relation to all its possible uses, then this would give rise to all too numerous repetitions or multiple classifications. In the above example with sulphuric acid it would be impossible for a universal library to classify sulphuric acid under all its potential uses. Therefore Soergel's concept of request oriented indexing is indeed significant, and for specialised information services it is important to classify according to the need of the target group.

Of course the problem with a pragmatic concept of the subject lies in the most basic sense in the condition it shares with pragmatic philosophy: even though the goal is to develop human practice, a narrow practice-orientation is too short-sighted and superficial in its truth criteria. Pragmatism contains no profound criteria for significance that can give direction to indicating the priority of the properties of a document.

A cow can be described both zoologically as a mammal and pragmatically as a domestic animal or livestock. Dalhberg [23, p. 194] designates the last relation as the relation between man and object, but assigns to the first another type, namely 'ontological'. We are not in agreement in this absolute distinction: all cognition is fundamentally instrumental for man. The concept of 'domestic animal' has a more *immediate* connection to human practice, whereas the concept 'mammal' is an abstraction with a *less immediate* relation to human practice. Classification of a book on cows in the subject category 'mammals' or in 'domestic animals' is not dependent on the most significant property of the book (the central object is a cow in both cases). It depends basically on the evaluation of whether the book is of most use to people looking for literature under zoology or agriculture, i.e. whether the book is of most use to a biologist or a farmer. This is a judgement based on the properties of the book in relation to perception of interests in an epistemological sense. This judgement is perhaps made primarily on the basis of the content of the book, but when subject description is intended for another target group, other decisions would be made (cf. this example with Chemical Abstracts and Ringdok).

Abstract and general knowledge of biology and the other sciences have clearly demonstrated their significance for man, even though their designation of useful functions is less immediate than 'domestic animal'. Scientific systematisation and terminology provide a topical organisation of knowledge which on a superior level assures the most effective communication in the development of human knowledge. Such an organisation of knowledge is difficult to justify from a pragmatic philosophy, in the usual understanding of this concept in philosophy.

Even though pragmatic subject theory has its limitations, it makes an important contribution to perception of central properties of the concept of the subject by pointing out its means-goal nature (and thus repudiating the view of subjects as 'inherent qualities'; subjects are no more inherent qualities than is the *value* of a thing).

This is supported by the etymology of 'subject' (especially in the Scandinavian languages, but also in English and German, see Note 2). 'Subject' (Scandinavian: 'emne') means 'raw material', among other things. Iron is a subject for the smith. A cow is a subject for the zoologist and the farmer. Epistemology is a subject for the philosopher and the information researcher. A subject thus is always a subject for someone or for something.

5. A REALIST/MATERIALIST SUBJECT THEORY

According to the realistic and the materialistic viewpoint things exist objectively and encompass objective properties. This is a crucial point of departure which is to be taken for granted in this article (see Note 3). In this paper, no efforts will be made to illuminate the differences between 'scientific realism' and 'materialism'.

Documents are (in this context) a theoretical problem. On the one hand, naturally, documents reflect the author's subjective view of the subjects treated. On the other hand the document has objective properties. If a document states that 'a person's intelligence is correlated with the size of his brain', this is a subjective (and false) judgement. But it is an objective fact that this document contains this (false) judgement. We are interested in the objective properties of documents. The objective properties are not the subjective judgements or evaluations contained in the documents; objective properties have cognitive or (informative) *potential* (provided that the reader can differentiate between false and true statements). Our conception of the objective properties of documents is reminiscent of Karl Popper's concept of 'World III' [24], in which he refers to books as 'objective knowledge', and operates with thought experiments very similar to mine. However, my concept of the objectivity of documents is not borrowed from Popper, and there are great differences between them because Popper's theoretical base is dualism and mine is monism. There is no space here to evaluate Popper's theory in relation to mine. It is controversial and has been seriously criticised both in philosophy and in information science (as to the latter, see Rudd [25]).

What is to be understood by the properties of a document?

In the broadest sense, the properties of a document are every true statement that can be said about that document.

A document can describe the achievements of Christian the Fourth, state the melting points of metals, present information on the composition of food additives and their consequences for human health, investigate the unicorn as a psychoanalytical symbol etc. The properties mentioned here can be said to deal with a document's reflection, representation or treatment of a part of reality (or of human consciousness and imagination). Which aspect of reality it reflects (its 'aboutness') is one of a document's central properties. It is also significant how it treats or reflects reality, for example whether its claims are true or false, representative, superficial or fundamental etc. A category of properties may be termed relational: how is this document related to other documents? Does it elaborate, overlap, correct or make other documents superfluous?

Documents can be characterised by their language, form, type, etc., which often represent lesser properties (cf. Hjørland [26]). And finally, documents can be characterised by type of paper, binding, typography etc., which in most cases would be insignificant, but for special purposes (the history of the book) may be central properties. The properties of a document emerge especially in the use of a document, for example by reading the document in connection with a particular activity (research, education or other). The frequency and structure of the words used, i.e. the language expressed in the document, also belong among the properties of the document. These latter properties do not normally appear directly through reading a document, but, for example, through processing it for automated functions, searching or automated indexing, classification etc. I will end discussion of these latter properties here, even though they do naturally play a large role in information science literature. The language in which the document is expressed plays a large practical role in information searching, because these elements often are accessible for searching, either in full text bases (still the exception), or in the form of representation of parts of the text in databases, usually the titles and abstracts. I will pass over this problem here. I am in agreement with Spang-Hanssen [27, p. 20] that a document's content cannot be described in depth merely by a formalisation of its language.

I have now provided a brief definition of the properties of a document. Now we must consider to what extent the properties of a document can be described objectively.

Curiously enough, objectivity means two different things in relation to judging the properties of a book (described here according to the realistic epistemology): 1. independent of the subject who apprehends; 2. in agreement with reality. In the first of these senses, the more readers who identify these same properties with the book, the

higher the objectivity. In the sense of 'in agreement with reality', the relationship is inversely proportional. Since special qualifications are necessary to be able to identify the significant properties in a scientific book, perhaps only a limited group can grasp the full potential of a work. In other words, the properties easily identified by the many will often be the less significant (or the more indiscriminate), and thus less objective in the second sense of this word. (This situation is especially the case in basic research, where theoretical re-orientations take place. In more everyday contexts, the 'normal process of research' (in the Kuhnian sense), this expressed contrast between the two objectivity requirements need not obtain).

To repeat: there is a direct contrast between the two concepts of objectivity in the evaluation of a book's most significant properties and thereby its subjects. The solution of this problem is not a decision by majority. The solution is an explicit argumentation and, if not a provision of proof, at least an establishing of a probability. We have seen that the description itself of the properties of a document is not a simple thing, susceptible to automation, but that it is highly dependent on particular conditions (which often are of a theoretical nature). When we maintain that the properties of a document are objective, even though the description of them requires special subjective prerequisites, this implies that reality, the testing of the document in practice, will in the final analysis decide its informative potential, no matter how many earlier misconceptions have been made. History becomes the final judge of the objectivity of statements about the properties of a document. (And even though history will rarely finally decide this, we retain the concept of objective properties in documents which are the basis of our attempts to analyse them.)

Different properties of documents can have different meanings for different purposes or scientific disciplines. Scientific disciplines or theories can have different foci or different epistemological interests. Therefore there can be marked differences in identifying central properties of documents. An identification of properties from a narrow theoretical point of view is more pragmatic than a more general perspective. Identification of the properties of documents from a superior or general point of view presupposes an ability to evaluate the potentials of different theories, that is it presupposes more a philosophical perspective. Library and information science personnel with a deep degree of subject knowledge and with expertise in searching databases and evaluating searches done for professionals, do often have important prerequisites for identifying such general properties.

Subjects and the properties of documents.

In philosophical usage the documents represent the individual variable and their properties and relations the predicates (together the properties and relations are termed the logical attributes of the document).

The examples mentioned of the properties of a document (the part of reality with which it deals, its truth value, its method etc.) make up predicates of the first degree (or first order predicates), just as does its lexical structure, etc.

When a librarian or information specialist categorises documents with a subject description, it is these predicates of the first degree with which he interacts: either by reading the book, or by inspecting its lexical structure (and in the extreme case he can construct a computer program that categorises documents from this structure). On the basis of this analysis of the first degree predicates of the document, he assigns it a predicate of the second degree, a predicate predicate (see Note 4). *An assignment of a subject is thus a function of*

the properties of a document and is in itself an attribute of a document (see Note 5).

Viewing the subject as a function of the properties of a document in this way does not in itself say what the subject is. Despite this, the predicate concept clarifies the relationship between the subject of a document and its other attributes (see Note 6).

To determine the concept of subject we must concern ourselves with *which* properties of documents enter into a subject description, and *in what way* they play this part. In practice it is often an extremely simple thing to say what the subject is (cf. the naive concept of the subject): designation of a subject often merely requires pointing out one or a few significant properties in a document, in particular the conditions in the real world that the document reflects. If the document has the property that it treats the building style of Christian the Fourth, then the document can be assigned the subject predicate 'Christian the Fourth's building style'. In this example there is an apparent identity between what we have defined as a property of the document and its subject, but since a choice has been made among the theoretically infinitely many properties, the subject description is in principle not identical with the predicates of the first order of the document. An explanation is lacking for why just this property, in just this case, has been selected as the subject. In other words, we must look more closely at this subject function (see Note 7).

Which properties of the document enter into the subject description?

As emphasised above, very often in practice rather simple and hard properties form the basis of subject analysis. Theoretically, however, this becomes extremely complicated, and as soon as an attempt is made to exclude a property, a hypothetical example pops up in which just that property would be part of determining a subject. The authorship of a document is hardly part of analysing the subject? Yes, in the case of autobiographies (and as Boserup [28] indicates, also hypothetically in other situations). I will not attempt to demonstrate here that all properties of documents enter into the subject function or to eliminate those which do not. My point of departure is that there is not a well defined or definable portion of the properties of documents which enter into the analysis of the subject (and that just exactly this situation leads to Patrick Wilson's agnostic concept of subject matter).

In the same way I would make the claim that the subject function cannot be a previously fixed procedure at analysing properties, such as Ranganathan's PMEST formula attempts to set up. It is my opinion that exactly the choice of specific properties of documents or specific functions of these properties inevitably leads down the idealistic path. Since librarians and information specialists would very much like to have clear and firm directives and procedures, an idealistic tendency continually lurks in the wings within the conception of the subject itself. (But of course in the concrete development of information systems procedures must be described, for example in the use of classification systems and thesauri, and I myself in another connection have been a spokesman for definite and explicit procedure (checklists) in the description of subjects [29]).

My point of departure for a materialistic theory of the subject lies in the pragmatic conception of the subject presented earlier. Subjects build on an evaluation of the documents' properties with regard to optimising the potential perception of the document. Which properties of documents are relevant, and which analytical functions are to be instituted with regard to these properties is not given *a priori*, but is, *inter alia*, dependent on context (see also Note 10).

Subjects in themselves must thus be defined as the epistemological potentials of documents. A potential is a rather intangible property – hence the problem with defining subjects. But the potential of a thing is not a subjective or objective 'idea'. A potential is an objective possibility. Uranium held its potential as an atomic fuel before science was aware of these possibilities, and many authors have been buried before the significant potential of their work has been recognised. Which things and works have which potentials is determined by the current stage of society's development. At one stage uranium is a not particularly valuable metal without special potential. At another stage it is an important energy source, and at a third stage it is perhaps something else again. *This is to say that it is the level of development of human society, the human practice, that constitutes a subject* (see Note 8).

Thus a subject description of a document is in one or another way an expression of the epistemological potentials of the document, such as these appear to one who describes the subject. The better the description predicts the potentials of the document, the more correct, more objective, the description of the subject is. The understanding of this should become more clear by reading the concrete example analysed in the appendix to this article. However, an interpretation of a given description of a subject must involve the qualifications (and interests) of the person who has carried out the subject description. When Patrick Wilson [11, p. 92] wrote (with regard to what the user can expect to find within a particular location in a library's classification system): 'for nothing definite *can* be expected of the things found at any given position', this is only correct from this subjective prerequisite. We can affirm with the adherents of hermeneutics that perception of the potential of documents depends on the *pre-understanding* of the person who carries out the determination of the subject. In contrast to many adherents to hermeneutics, I, however, wish to retain the concept of the objective potential or subjects of documents.

A subject description is thus a prognosis of future potentials. This prognosis can be based on positive as well as negative judgements. The subject description can be seen as both a kind of vision and as an evaluation in relation to current research. The most important prerequisite in subject description is not a special kind of method, but is maturity in judgement.

The use of subject systems thus also assumes interpretation. The user must enter into the universe of the system and its devising. This is hardly exceptional. In some cases documents are ordered by the so-called 'principle of provenance', which requires that documents remain in the collections and the order in which they were originally organised. This requires an insight into

the organisation that existed when the collection was established. Ordering of documents and knowledge is always based on particular premises, world views, assumptions. Knowledge of these premises is often necessary to obtain a satisfactory return from descriptions of subjects. The necessary degree of interpretation depends upon the extent to which the subject description has anticipated and met the needs of the user. In the principle of provenance only a low degree of anticipation is attempted, because the principle makes no attempt to consider the current user's context. In contrast, the above-mentioned pharmacology Ringdok database exhibits a high degree of accommodation of user needs. Information systems which take into account the needs of the users are more expensive to establish and maintain, but are in return economical of resources in use.

A description of a subject is rarely presented as a direct statement about the potential of a document; more often it appears in the form of a reference to an academic discipline ('the subject is psychology'), i.e. a socially defined problem area, within which the document particularly contributes to problem-solving. As previously mentioned, subjects can also be expressed indirectly by merely emphasising special qualities ('treats the architecture of Christian IV'), which can also be located in a discipline (history, art history) or which serve directly as the base from which the user himself evaluates the subject of the document (for example 'tourist attractions').

The issues of the expression of subjects, of 'information retrieval languages' and of representation in text go beyond the framework of this article. But since these issues presuppose a knowledge of what subjects are, the proposed theory of 'subject' presented here is a prerequisite for more profound theories on these questions.

We can now return to Patrick Wilson's problem regarding the 'ill-defined phenomena' of authors. A designation of a subject reflects the clarity or imprecision of a document, but not in the way mentioned by Wilson. The purpose of analysing the subject is to determine whether a document has an epistemological potential in relation to future users of a given category or a given concept, for example 'hostility'. If it does, it is classified under that concept, if not, then it is not put there. (If it is put there in order to illuminate the unclear terminology in the field, this could also be regarded as a kind of informative potential, even of a more indirect kind.) Assignment of a subject to a document is indeed a clear judgement that 'this document has epistemological potential within "hostility"', even though this clear judgement is based on many deliberations of whether or not the document actually contributed to this subject, because it was imprecise in its use of concepts. In actual practice other possibilities do often exist, preferable from an ideal point of view, for example characterization of the methods or theoretical approach of the work, which will may give the work a higher profile in a database depending on its structure; in other words decisions on the subject of a document are not typically an 'all or nothing' judgement (see Note 10).

Subjects and epistemology

Documents are sources for the cognitive process just as people, things, processes, statements etc. also are sources of human cognition. How man achieves knowledge preoccupies epistemologists. Part of human cognitive activity (an important special case) is scientific cognition, which in addition to

epistemology also builds on the theory of science and the methodologies of academic disciplines.

Various types of epistemology exist, for example idealism (positivism), scientific realism and materialism. It lies outside the scope of this article to sketch or treat epistemology itself. The purpose of this article is to clarify the concept of subject, and with that objective in mind it is necessary to view the determination of subject from an epistemological point of view. This follows in particular from the conclusion of the previous section: that a determination of a subject is an evaluation of and assignment of priorities to the properties of a document with regard to a categorisation and a subject description of that document. How this categorisation and description go on is decisive for the 'visibility' of the document in libraries and databases, and thereby for its potential role in future development of knowledge.

The most generalised knowledge about how a person, for instance a researcher, or a whole discipline, ought to examine the world in order to expand human knowledge is lodged in philosophical epistemology. Therefore I conclude that insofar as such a theory is at all able to produce useful results, this theory also is the foundation for the determination of the subjects of documents.

If a researcher poses a particular question, for instance about apes, or the origins of life, it is the hypotheses and formulation of the question which are primary. What methods may be used to investigate the question, 'empirical', or 'theoretical analysis' or 'library investigations' (i.e. a search for literature), are secondary. The clarifying of the question and the central concepts involved would be the same at a certain level. The question *determines* which things, processes, documents etc. are relevant to the study, and how they are relevant. Another issue is, to what extent the relevant documents can be identified. It is my claim that it is extremely difficult to identify the most relevant documents in modern science (cf. Hjørland [29] for an analysis of this problem with a case study). The effect of this identification being so difficult is that the theoretical basis of information systems assumes the status of an important scientific problem. The description of the subject of a document (i.e. the evaluation, assignment of priorities and the consequent categorisation of the potentials of the document) assumes an insight or understanding of which future problems can give rise to use of the document in question. The reason for this lies in two assertions: 1. any document possesses an infinite number of properties (so that it is not possible to count them all); 2. the properties which are central to one context need not be so in another (thus a fixed set of priorities cannot be established once and for all, as the example from Ranganathan's system illustrated).

Epistemology has something relevant to say on what it means 'to describe'. What does it mean to describe, for instance, the content of a book? We will touch lightly on the epistemological aspects of this, based on Kröber and Segeth [30]. The concept of description is most commonly used about the perceptions of the senses, which are presented in a systematic and ordered way through deliberation and language. A successful description can achieve quite a precise picture of the item described, but it

can only state how this object is constituted, not why it is constituted as it is. *For the same reason a description keeps to the superficial aspects of an object, and does not pursue its essence*, including the reasons for its existence. A description is thus a first step in cognition, which is later replaced by other modes of cognition that delve deeper into the essence of things. Positivistic epistemology's programme of restricting scientific method to mere description of facts is too narrow in relation to the above. Positivism's requirement of a *complete* description of a phenomenon is both impossible and unnecessary. A complete description is impossible, because the infinite number of properties of a phenomenon would require an infinitely extensive description. A complete description is unnecessary, because both for scientific knowledge as well as for practical human purposes, an equally detailed description of all significant and insignificant, necessary and random, general and particular properties and relations is pointless. What is needed is knowledge of the significant, the general among the particular, the necessary and the typical. *Description can therefore only fulfil its function in the knowledge-gathering process insofar as it is not made absolute and discrete from other means of cognition, such as explanation, hypothesis, prognosis etc.* Description must, indeed, be viewed in the context of other such modes of cognition.

We see no reason to doubt that the very same situation holds regarding the description of the subjects of documents: a 'pure' description of documents without connection to other modes of cognition such as hypothesis, prognosis etc. can only extract the more trivial and superficial properties of the document. Comparison of the subject descriptions made by librarians and sociologists of sociological literature, for instance, gives some insight into this situation [31]: because the documents are not merely 'described', but evaluated in relation to their sociological value, the sociologists' judgements on subject were the most precise and useful. It is banal to discover that the better the qualifications one has in an academic discipline, the better the judgements made on the significant properties of a book from that field; and the reverse is also true: the poorer the qualifications, the more random and superficial the assessment and the properties that are emphasised.

We have in this section seen an example of how two epistemological theories (positivism and materialism) view the role of description in the development of knowledge, and we have from this example seen the fundamental role epistemology plays in the evaluation of subjects, and how the same theoretical problems that occur in regard to material objects also occur with regard to the role of documents in the development of knowledge.

It is naturally decisive for a theory of subject matter to recognise how to distinguish between the superficial and accidental properties on the one hand, and the significant properties on the other. Once again this is a basic problem of epistemology (as well as a problem of scientific method). Just as it is pointless to describe flora by superficial characteristics (such as colour) instead of meaningful characteristics (for example categorisation in plants with seeds or with spores), it is naturally just as necessary to describe documents according to meaningful rather than superficial characteristics. Thus an epistemological theory which facilitates the development of knowledge in the direction of the substance of things is what is needed. *Such a theory stands out in sharp contrast to conceptions which are based on research and analysis of subjects as an algorithm, a 'trick' or an a priori method.* It is rather the method

which should be a reflection of the essence of the object.

Materialistic theory, in contrast to pragmatic theory, is characterised by a broader and more far-sighted interest in epistemology. Realistic and materialistic theory of the concept of subject does not merely attempt to solve the limited problems of here and now, but hopes to contribute the greatest possible consciousness of long-term consequences. Subjects are not merely to be structured in a narrowly instrumental way, but the attempt must be made, for instance, to contribute to a deeper penetration of the sciences into the innermost essence of reality. Subject categories should exhibit this in such a way that they reflect significant and general aspects of reality. *In practice it will often be the concepts of the sciences with which materialistic subject theory operates, because the sciences are the cognitive organs of society* (see Note 10). Of course, the sciences are naturally neither uncontroversial, objective nor infallible, but, at least as an ideal, the debate about the objectivity of scientific research is a part of science (see Note 11). *Thus an analysis of a subject is itself, at its most profound, a part of the scientific process of knowledge gathering.* This analysis is dependent on contextual factors, including the existing volume of literature and the system of its access points (see Note 12).

ACKNOWLEDGEMENTS

This paper was partly written during my affiliation to the Royal Library in Copenhagen, partly in my present position at the Royal School of Librarianship. I should like to thank both places for giving me good working conditions to do this research.

NOTES

Note 1

Occurrence of the concept 'about?' in the *Library and Information Science Abstracts (LISA)* database:

S1	5504	ABOUT?	= total occurrences in July 1989
S2	560	1/DE, TI	= occurrences in titles* and descriptors
S3	74	2/DE	= occurrences as descriptor
S4	68	PY = 1989	
S5	2865	PY = 1988	
S6	5744	PY = 1987	(*manual check shows that title-occurrences make up noise)
S7	5872	PY = 1986	
S8	5392	PY = 1985	
S9	5933	PY = 1984	
S10	5986	PY = 1983	
S11	5963	PY = 1982	
S12	5651	PY = 1981	LISA's total number of references sorted by printing year
S13	5469	PY = 1980	
S14	5388	PY = 1979	
S15	4506	PY = 1978	
S16	4171	PY = 1977	
S17	3790	PY = 1976	
S18	3681	PY = 1975	

S19 2695 PY = 1974
 S20 2978 PY = 1973
 S21 2985 PY = 1972
 S22 2576 PY = 1971
 S23 0 S4 AND S3 1989 The descriptor 'about?'
 S24 1 S5 AND S3 1988 distributed by printing year
 S25 1 S6 AND S3 1987
 S26 1 S7 AND S3 1986
 S27 0 S8 AND S3 1985
 S28 0 S9 AND S3 1984
 S29 0 S10 AND S3 1983
 S30 1 S11 AND S3 1982
 S31 0 S12 AND S3 1981
 S32 1 S13 AND S3 1980
 S33 1 S14 AND S3 1979 It shows, that the use of the descriptor
 S34 2 S15 AND S3 1978 concentrates on the years
 S35 2 S16 AND S3 1977 around 1975, which we
 S36 11 S17 AND S3 1976 interpret as a fad
 S37 30 S18 AND S3 1975 which has not
 S38 15 S19 AND S3 1974 caught on.
 S39 3 S20 AND S3 1973
 S40 0 S21 AND S3 1972
 S41 0 S22 AND S3 1971
 S42 0 S3 AND PY = 1970
 S43 0 S3 AND PY = 1969

Note 2

The etymology of the concept 'subject' (Scandinavian 'emne').

Nudansk ordbog (13. udgave) maintains that the word 'emne' was borrowed about 1760 from Norwegian 'emne' or Swedish 'ämne'; same word as 'evne'. It mentions three meanings of which only the first two are of interest in this connection: 1. material for treatment in speech or writing; theme; motive; 2. material ('raw material'), which is partly worked up, e.g. about keys before the final filing. *Nusvensk ordbok* mentions four meanings of which the first is 'raw material', 'something to produce out of'.

'Emne' can be translated into 'subject' in English. The concept of 'subject' has in the *Oxford English dictionary*, second edition, eighteen main meanings. It is complicated that the English 'subject' has many meanings, among them the Danish 'subjekt' (i.e. grammatical 'subject'). Of the eighteen meanings in *OED* the following should be mentioned:

5. The substance of which a thing consists or from which it is made.
7. Logic. a. That which has attributes; the thing about which a judgement is made. b. The term or part of a proposition of which the predicate is affirmed or denied.
8. Gram. The member or part of a sentence denoting that concerning which something is predicated (i.e. of which a statement is made, a question asked, or a desire expressed); a word or group of words setting forth that which is spoken about and constituting the 'nominative' to a finite verb.
9. Modern Philos. More fully conscious or thinking subject: The Mind, as the 'subject' in which ideas inhere; that to which all mental representations or operations are attributed; the thinking or cognitive agent; the self or ego (correlative to *object* sb.6).

(The meanings 5, 7, 8 and 9 are derived through the Latin 'subjectum' from Aristotle's use of *τὸ ὑποκείμενον*, with the meanings 1. the material of which the things consist; 2. subject for attributes (qualities); 3. subject for predicates (names)).

- 10. The subject-matter of an art or science.
- 12a. That which is or may be acted or operated upon; a person or thing towards which action or influence is directed, or that is the recipient of some treatment.
- 13a. In a specialised sense: that which forms or is chosen as the matter of thought, consideration, or inquiry; a topic, theme.
- 14a. The theme of a literary composition; what a book, poem, etc is about.
- 18. attrib. and Comb . . . (sense 14, chiefly with reference to cataloguing books according to their subjects) subject card, catalogue, cataloguing, entry, heading, index, list, reference; . . .

The meaning of special concern to us, is of course especially #14 (and the combinations #18), that is 'subject' in the meaning 'what a book is about'. This definition does not, however, solve the problem. What does it mean that 'a book is about' the subject *x*? According to the meanings 12a and 13a – and the above mentioned definitions from Danish and Swedish – we find evidence of our conception of the concept of 'subject' or 'emne' as being a 'raw material' for humans to act upon.

In German terminology, you will see that subject indexes and the like in libraries, books etc. are often called 'Sach-' or 'Fachregister'. 'Fach' is a reference to professions or scientific disciplines. That means that in German there is a direct connection between the terminology used for our 'subject' and the social groups which may be using those documents. That is, the concept of 'subject' has no precise equivalent in German, but the corresponding concepts underline the function to refer documents to categories of users.

The etymological meaning of 'raw material' underlines the fact that it is not the innate properties in the things themselves, but their functions for the human user, which make up 'subjects'.

(In the article I have compared the concept of 'subject' with the concept of 'value'. This gives a better grasp on the meaning of 'subject': gold has its value not from the chemical properties in themselves (they are necessary: that gold is 'precious' is partly due to the fact that it is not easily corroded by chemical influences), but because of special cultural conditions. The 'value' is not a built-in property in things but is still a function of the properties of the things and of the human culture.)

Thus we have seen that our conception of the concept of 'subject' in library and information science is not in contrast to important meanings in general language. If such a contrast had existed, our position would have been weakened because we then had to argue for a special usage of the word. Of course we are not claiming that the general concept 'emne' or 'subject' cannot have other meanings too, as seen in *OED*, but we emphasise one side of the concept which supports our theoretical points.

Note 3

Not all modern researchers are of the opinion that things exist objectively and have objective properties. For example the influential book *Understanding computers and cognition: a new foundation for design* by Terry Winograd and Fernando Flores [32, p. 73 ff.] takes the opposite position.

Note 4

An example of a predicate predicate is 'F is symmetric', where the property of symmetry is a predicate for parts of a body which have a particular relationship to each other [33].

Note 5

There are other predicates of the second degree than the assignment of subject. If, for example, a document is said to be characterised by belonging to the structuralist school, (and this judgement is indeed made directly from the properties of the document), this is a meta-description which is not identical with a subject description, but might sometimes be a part of a description of subject. (If an assignment of subject is

based on such a secondary attribute it may itself become an attribute of the third degree, but that does not have to be dealt with here).

Note 6

Another concept significant for the concept of subject is the very concept of 'concept'. In the last twenty years significant changes have taken place regarding *concepts* in research within psychology, philosophy and linguistics. These developments cannot be summarised here, but they are of great importance for the meaning of the concept of the subject. One result is that some concepts today must be viewed as the result of an inductive argument. Smith [34, p. 518] gives the following example:

The animal originally had typical bird properties.
The animal accidentally acquired typical insect properties.
The animal produced offspring with typical bird properties.

This animal is probably a bird.

That is, human beings, when faced with a problem of categorisation, are capable of rising above similarities and employing deductions, which require a facility for further assumptions. This is in direct conflict with the view expressed by Beghtol [35, p. 95–96] that the classifier judges the class relationship on *the basis of similarities between documents*. *Here we are proposing the view that just as modern research in concepts has gone beyond similarity as the only valid criterion in the view of concept, it is correspondingly necessary to move beyond the similarity of documents as the only criterion for subject relationships.*

Note 7

It is my experience that many people view this discussion as unnecessarily complicated. Why is it not possible to grasp subjects as more tangible properties of documents? This does of course work in most cases. But it is my view that in particular work on the concept of subject in psychology and the social sciences necessitates a far more abstract and complicated conception of subject than that which previously has been discussed in LIS literature. Examples are given in the appendix to deepen understanding of the problems of subject analysis in psychology and the social sciences. It is noteworthy that the criticisms of other conceptions of subject (for example 'aboutness') often come from persons with a background in the social sciences. (cf. Swift *et al.* [36]). This does not, of course, mean that the concept of subject proposed here only has validity for the social sciences. Rather, the needs of the social sciences contribute to a generalisation of the concept of subject in such a way that will be fruitful in other areas. A general information science theory has to be based on such a generalisation of experience and theories within specific disciplines (as opposed to the opposite: that a finished theory be forced on specific fields).

Note 8

I owe the expression 'it is the human practice that constitutes a subject' to my colleague Anders Orom, who coined it in response to an oral presentation of my theory of subject.

Note 9

This relationship leads us to a new question: are there documents without subjects? In theory one has to answer no to this question; we cannot imagine documents without any cognitive potential. And it is a rare experience to consider in practice not assigning any subject designation. In specific cases the lack of clear possibilities for classification usually reflects that the document in question was inappropriate for acquisition or inclusion in the particular database. Thus the lack of a 'subject' usually expresses an inconsistency between policies of accession and indexing.

Unfortunately a contradiction in subject descriptions can occur. Documents which correspond to a classification system (or IR language) receive single, or few,

classifications, which correspond to the respective category in the system. Vague or cross-cutting documents often receive far more classifications and thus achieve an unintended visibility. This phenomenon ought to be contained. Information systems ought to provide optimal use of the knowledge in the collected mass of documents. In the above case a document achieves visibility at the expense of other documents: if all documents were placed in all categories all value of categorisation would be null and void. Rare situations can thus also occur where a subject description of a document does more harm than benefit, and such descriptions should be avoided.

Note 10

In addition to subject analysis for scientific/scholarly purposes, subject analysis of a more pragmatic nature also exists. Subject analysis of documents does not always have to be viewed as a scientific process of cognition, even though scientific perception/cognition often naturally extends into, and becomes relevant to, more ordinary perception. This view on the role of scientific disciplines is in opposition to many information scientists, who try to avoid scientific disciplines and instead – as, for example, the Classification Research Group – describe documents according to 'more fundamental semantic categories'.

Note 11

This emphasis on disciplines rather than 'forms of knowledge' or 'topics' represents an alternative to a widespread viewpoint represented in library science, for example in Langridge's recent book *Subject analysis* [37]. Because this book represents a different theory about subject analysis, I shall give a short comment on it.

Langridge analyses the concept of subject in two major components:

(a) Central to his book is the thesis that fundamental categories of knowledge exist. These are the philosophical categories, which go back to Plato and Aristotle, introduced to LIS especially by S.R. Ranganathan. Langridge prefers the expression 'forms of knowledge' to these fundamental categories.

There are relatively few 'forms of knowledge'; Langridge lists twelve, for example Philosophy, Natural science, Technology, Human (behavioural or social) science, History, Religion, Art, Criticism and Personal experience.

(b) Beside these 'forms of knowledge' Langridge operates with 'topics', which are 'the phenomena that we perceive'. Where 'human science' is a 'form of knowledge', 'human behaviour' is a topic.

Besides the two fundamental components, a third one exists:

(c) the concept of discipline (or 'field of learning') (p. 31): unfortunately, this extremely important distinction has been blurred in many people's minds by the existence of a third kind of term which combines both form of knowledge and topic. For example, ethics is the philosophy (form) of morals (topic); zoology is the science (form) of animals (topic); psychology is the science (form) of human behaviour (topic).

Langridge does not like the concept of scientific disciplines as a concept in subject analysis. They are unstable: '... the disciplines that constitute specialisations may be unstable, but the fundamental disciplines, or forms of knowledge, are not. The specialisations are a practical convenience for sharing the world's intellectual labours: the forms are permanent, inherent characteristics of knowledge' (p. 32).

Langridge's concept of subject takes the above mentioned 'fundamental components' as point of departure for subject analysis. It does *not* make reference to the user context, to the 'pragmatic viewpoint' of subject analysis.

In my classification of conceptions of 'subject', Langridge's theory – in the tradition of Ranganathan – must be labelled 'objective idealistic'.

My own viewpoint differs in more ways:

First, in my theory, 'disciplines' are the central point of departure. They are often unclear and unstable – admitted, but they are the best we have. It is the job of the disciplines themselves – together with philosophy (and hopefully LIS) – to make them more clear, more well defined.

Second, the fundamental philosophical categories are important, but epistemologically they should be understood as generalisations of scientific research. Scientific research by the way is not only empirical research but also theoretical research. No sharp boundary exists between science and philosophy. These philosophical categories are *relatively* stable, but they are *not* 'permanent, inherent characteristics of knowledge' (I read such a statement as a clear idealistic position).

Third, Langridge's concept of 'topics' as 'perceived phenomena' represents the positivistic, empirical and 'subjective-idealistic' position as the fundamental point of departure. From the 'realist' (in the Platonic and scholastic sense) or 'rationalist' position, the opposite is the case: the perceived phenomena are subsumed by the 'immortal ideas'.

Langridge seems to follow the 'rationalistic' or 'objective-idealistic' position, where 'perceived phenomena' are subsumed by the 'immortal ideas'.

Both the rationalistic and the empirical points of view contain part of the truth: it is the one-sided emphasis on one of these viewpoints at the expense of the other which leads to either 'subjective idealism' or 'objective idealism'. Science starts with such perceived phenomena as flowers (botany), stones (geology), stars (astronomy), chemicals (chemistry) etc., but in the development of science, the perceivable objects are turned to more unperceivable objects. Plants, for example, are defined as living organisms with chlorophyll granule and microbiology recognises living organisms which are both plants and animals (having both mouths and chlorophyll granule). That is: perceived things influence sciences and 'forms of knowledge' (empiricism) and the theoretical knowledge thus obtained changes our perceptions and lets us see new things (rationalism).

From a materialist and modern realist position ('qualified realism' as opposed to 'naive realism'), the scientific disciplines represent or reflect the world, the same world as we perceive. But these questions are difficult, and many sciences are having difficulties in saying what their objects are. This ought to be clarified, but it is not sensible for library and information science to pursue its own way, to try to go it alone and to avoid this unclear status by choosing an idealist instead of a materialist theory of knowledge, to base its subject analysis on either 'immortal ideas' or 'perceived phenomena'.

'Topics' or 'perceived phenomena' must be part of the same reality as the study of science. The scientific and the non-scientific perception should both be subsumed by one theoretical dimension of analysis, where 'the theory of integrative levels' is a good starting point.

Langridge follows one tradition in library and information science, a more library-oriented line, with S.R. Ranganathan and the British Classification Research Group as leading figures.

This tradition seems to be separated from another avenue of research, represented, for example, by Blair's *Language and representation in information retrieval*, one could say a more database-oriented line. Both lines are very preoccupied with epistemological questions, and their main difference can be seen as different epistemological positions, where the school of Ranganathan and followers represent a rationalistic, or 'objective idealistic' line, while Blair, following the late Wittgenstein, represents a pragmatic point of view.

In my own research, I try to use the best of both traditions (and others as well) and integrate it into another epistemological tradition – that of materialism/realism. The selection of the epistemological position is not a 'free choice'. A wrong position is scientifically infertile and research on such a line will be contradicted by reality and the research will not flourish, but represent a blind alley. Epistemological positions are

therefore not chosen, but worked out in fundamental research in order to solve theoretical problems. A materialistic or realistic position does not – contrary to commonly held belief – represent ready-made solutions. It paves the way for concrete theoretical and empirical work.

Note 12

Developmental principles for subject description

In practice, of course, there will often be several subject descriptions of a given document. In addition to the subject descriptions there are the properties of the document, for example in the form of concepts in databases (from title, full text or other sources). The function of subject descriptions has naturally to be viewed in relation to such a system of possibilities. This belongs to the technical questions ('information retrieval languages') which will not be treated here. That which is of significance in this connection is that the information explosion (i.e. the growth in the number of documents among which discrimination must take place) has had consequences for qualitative aspects within subject description. The user of the documents becomes acquainted, of course, to a greater or lesser degree, with their properties. On the basis of this, the user himself makes a subject evaluation. The fewer the documents involved, the more carefully the properties of the document can be described and analysed, and the more certain is the subject description. Through an implicit understanding of this situation, many librarians and information specialists will, of course, provide access to as many of the document's properties as possible, and account for as many of the properties as the practical possibilities of their subject system allows. The greater the size of the document mass in which one searches, the more difficult it becomes to locate the truly relevant documents. Thus it would be best if the more the mass of documents grows, the more selective the subject descriptions become. In other words: the greater the mass of documents, the greater the necessity for an actual subject description rather than a mere registration of the properties of a document.

Insofar as a predicate is a more indirect product than a predicate, there will always be a greater uncertainty in relying on a subject description rather than investigating the primary properties personally. In contrast, reliance on others' subject evaluation exploits value-added service and saves time. Information systems must strive for an optimal solution to this dilemma.

The hypothesis can be formulated more precisely: *the greater the mass of documents, the more necessary it is to describe their subjects on the basis of user needs (rather than properties of documents)*. The multiplicity of properties and relationships between them creates a surfeit that leaves the user unable to determine relevance by means of analysis of properties. The burden is simply too great.

An example which supports this is the development of the subject index in *Kompas Danmark*, an index of products on the Danish market. The more products that appear within a field, the more the descriptions are based on the needs of the users. Thirty years ago, chemicals were primarily described by chemical properties, today they are more commonly described by their types of use (for example, fertilizers, photographers' chemicals etc.). An exception to this general tendency is the computer field, where it was previously common to describe hardware according to specific purpose, but today the tendency is to emphasise universality and describe properties.

Another example is a proposal to introduce the concept of the 'political civil servant' into the Danish central administration. (Cf. *Weekendavisen*, 27.7.1990.) This is in accordance with our view that different political parties will need different 'subject descriptions' of existing information, and this need will be all the greater, the more extensive the amounts of information become. The safeguarding of the principles of democracy lies perhaps not so much in having officially 'neutral' civil servants and neutral information systems, as in having consistent analysis and information systems that can provide real alternatives.

The above comments are included here in order to demonstrate that the subject does not consist of an *a priori* function of the properties of documents, but that the whole context in which subject description is carried out determines this function, and that regularities can apparently be described for the dependence of the subject description on contextual factors.

APPENDIX

Subject analysis: a concrete example

What is the subject of the book by Robert A. Wicklund with the title: *Zero-variable theories and the psychology of the explainer* [38]?

According to the title of the book, it is about certain kinds of theory ('Zero-variable theories') and about 'the psychology of the explainer'. The latter subject is related to the 'psychology of science'.

If you look at the book, you will see that 'zero-variable theories' are not favourably evaluated; they are described as simplistic theories and the book tries to explain why these kinds of theories are produced so much in modern psychology. Why do so many psychologists (or why do so many explainers in general) tend to use these kinds of simplistic theories on behalf of more varied theories?

In the preface to the book, you can read the following sentences: 'The reader should not suppose that this is a book about the philosophy of social science, or about moral pronouncement on what is good or bad in ancient and current psychological theorising. Instead, the reader is invited to consider the psychological side of the *explainer*'.

Before I present my own subject analysis of the book, we will have a look at the Library of Congress (LC) analysis. In the Cataloging-in-Publication Data the following subject terms appear: '1. Psychology – Philosophy. 2. Psychologists – Psychology. 3. Explanation'.

This means that LC, in its first selection of subject terms, is disposed *not to follow Wicklund's statement in the preface*, while the two following subject statements can be said to be in accordance with self-understanding of the book. This applies especially to the last subject expression.

My own subject analysis is the following: I regard the book as important, because it deals with a neglected issue in psychological research, or psychology as a science: the apparent decay in the theoretical level in psychology. This condition is illustrated by a number of concrete analyses of psychological theories, which in the succeeding psychological research has become substantially reduced. One such example is the almost classical theory of personality by H.A. Murray from 1938.

In my opinion, the most essential thing about Wicklund's book is in particular the concrete documentation of the apparent decline in psychological theory. There are lots of books about the philosophy and methodology of psychology, giving direction to the science of psychology, but there are relatively few books documenting the apparent decline in theory. It seems as if psychology does not exploit the best of its own theory and knowledge from philosophy and other sciences. How can this be explained?

Wicklund's explanation of this apparent condition is in my opinion not correct. Wicklund's explanation is different to the way I see things. Wicklund sees the documentation of theoretical decline as something less important in his book. His main interest is to use this material to give an explanation not only about the condition of psychology, but about the psychology of explainers in general. The material which I consider having the most potential value is, for the author of the book, only a minor thing.

This means there is a marked difference between the author's and my own judgement about what the potential value, the epistemological potential, of the book is. And therefore what its subject is. The book has – as any book – an unlimited number of

properties. To analyse a book's subject is to choose the properties which have the greatest potential for human knowledge. Therefore my subject analysis is other than that of the author as indicated by the title and the quoted sentences from the preface. The reason that Wicklund's and my own analyses of the central subject of the book differ so much lies in my professional evaluation of Wicklund's explanation, which I will characterise as being individualistic: Wicklund seeks an explanation of the decline in psychological theory in psychological mechanisms in the persons producing those theories.

Certainly Wicklund, in connection with his explanation, writes about interesting and relevant psychological phenomena (such as rumours and competition) which should be a part of the pattern of explanation, but in my opinion, a broader cultural and social description is needed as a background for the understanding of these mechanisms.

I believe the documented examples of decline in psychological theory can in part be traced to the market for psychological books (and the market for psychologists!). In a long period after World War II, the market for psychological books (and for psychologists) was the 'seller's market', and it was all too easy to sell even very poorly written psychology books (and to do poor research). This phenomenon is described in an article by Jürgen Kagelmann, psychological consultant for Psychologie Verlags Union, Munich, in the magazine *Psychologie Heute*, October 1988. Kagelmann's main point is that the (far too) easy sales possibilities in the 1970s made an overwhelming production of psychological books of a very doubtful quality. All that could be printed between two covers was thrown on the market, and the market was insatiable. This is an example of a non-individualistic explanation, which in my opinion comes closer to the truth than Wicklund's explanation, even if this is not a full explanation.

Therefore I think that Wicklund has a tendency to individualise and psychologise a social problem, and his book contains a contradiction in a way. Wicklund acts in this book also in the role of 'explainer', and he too has a tendency toward a very simplistic, positivistic theory, which the book is actually meant to fight against.

The epistemological potential of Wicklund's book lies in my opinion especially in its documentation of certain conditions in psychological science which it is important to set right. Therefore *the subject of the book is the epistemology of psychology, methodology, theory of science and philosophy*. In my opinion, LC was right in its first selection of subject terms (Psychology – philosophy), which, as mentioned, was in contradiction to Wicklund's statement in the preface.

I would *not* consider 'zero-variable theories' the subject of the book. It is hardly a concept with a future, not even as an explanation of the decline in theorising. It is an open question, whether what has been called 'variable psychology' [39, p. 522] is a valuable concept or not.

As regards the proposed subject 'psychology of the explainer' it is for me a theoretical question whether the behaviour of different explainers can be explained by the same psychological mechanisms disregarding what they are trying to explain. The question is whether 'a theory of explainers' can exist. Such a theory ought to include not only explanations of human behaviour (that is psychological explainers, professionals as well as laymen), but all other kinds of explanation too. Such a book would in reality come close to the discipline named 'decision theory', and that is not what Wicklund's book is about. My conclusion is that I tend to doubt the value of the proposed subject 'psychology of the explainer'. This doubt also includes LC's subject term 'Explanation'. Wicklund's book is hardly a contribution to the concept of explanation in general.

The last proposed subject which I want to discuss is 'psychology of psychologists' (LC: 'Psychologists–Psychology'). Such a subject does exist, and books are written about it. They can describe, for example, the recruitment of psychologists, the motivation for choosing the profession, professional socialisation and many other things. Wicklund's book is in my opinion not of this kind.

In my judgement – as already noted – the subject of Wicklund's book is 'philosophy

and epistemology of psychology'. My judgement is of course subjective, and could be wrong, in general or in part. The only way to decide this is to analyse the arguments. The arguments about the subject of a book are fundamentally the same as arguments about the advancement of knowledge.

The subject of a book is its (objective) epistemological potentials. The subject description which comes closest to the prediction of the role of a document in the advancement of knowledge is its most correct subject description. The evidence of the truth of the subject-statement lies in the argumentation. If my argumentation above cannot be rejected it constitutes a better suggestion about what the subject of Wicklund's book is than both Wicklund and LC have provided. If it can be rejected my subject description of that particular book is wrong, but this does not change my theory about what subjects are: the potentials of documents for the advancement of knowledge.

REFERENCES

1. VYGOTSKY, LEV SEMENOVICH. *Tænkning og sprog*. Bind 1-2. København: Hans Reitzel, 1982.
2. FROHMANN, BERND. Rules of indexing: critique of mentalism in information retrieval theory. *Journal of Documentation*, 46 (2), 1990, 81-101.
3. MØLLER, BENTE AHLERS. *Vidensklassifikation. En komparativ analyse af Statsbibliotekets systematiske katalog*. Århus: Statsbiblioteket, 1979.
4. MARK PEJTERSEN, ANNEISE. The meaning of 'about' in fiction indexing and retrieval. *Aslib Proceedings*, 31, 1979, 251-257.
5. MARK PEJTERSEN, ANNEISE. Design of a classification scheme for fiction based on an analysis of actual user-librarian communication, and use of the scheme for control of librarians' search strategies. In: HARBO, O. and KAJBERG, L., eds. *Theory and application of information research: Proceedings of the Second International Research Forum on Information Science, 3-6 August 1977, Royal School of Librarianship, Copenhagen*. London: Mansell, 1980, 146-159.
6. BELKIN, NICHOLAS J. The problem of 'matching' in information retrieval. In: HARBO, O. and KAJBERG, L., eds. *Theory and application of information research: Proceedings of the Second International Research Forum on Information Science, 3-6 August, Royal School of Librarianship, Copenhagen*. London: Mansell, 1980, 187-197.
7. BELKIN, N.J., ODDY, R.N. and BROOKS, H.M. ASK for information retrieval: part I. Background and theory. *Journal of Documentation*, 38 (2), 1982, 61-71.
8. BELKIN, N.J., ODDY, R.N. and BROOKS, H.M. ASK for information retrieval: part II. Results of a design study. *Journal of Documentation*, 38 (3), 1982, 145-164.
9. FARRADANE, J.E. Fundamental fallacies and new needs in classification. In: *The Sayers Memorial Volume*. London: Library Association, 1961, 120-135.
10. FARRADANE, J.E. Concept organization for information retrieval. *Information Storage and Retrieval*, 3, 1967, 297-314.
11. WILSON, PATRICK. *Two kinds of power: an essay on bibliographical control*. Berkeley: University of California Press, 1968.
12. GOPINATH, M.A. Colon Classification. In: MALTBY, A., ed. *Classification in the 1970s: a second look*. Revised edition. London: Clive Bingley, 1976, 51-80.
13. RANGANATHAN, S.R. *Documentation and its facets*. London: Asia Publishing House, 1963.
14. TRANEKJÆR RASMUSSEN, EDGAR. *Bevidsthedsliv og erkendelse. Nogle psykologisk-erkendelsesteoretiske betragtninger. Festskrift udgivet af Københavns Universitet i anledning af Hans Majestæt Kongens Fødselsdag 11. marts 1956*. København: Munksgaard, 1956.

15. JOHANSEN, THOMAS. *Indledende betragtninger over emners beslægtethed*. København: Danmarks Biblioteksskole, 1975.
16. JOHANSEN, THOMAS. An outline of a non-linguistic approach to subject-relations. *International Classification*, 12(2), 1985, 73–79.
17. JOHANSEN, THOMAS. Elements of the non-linguistic approach to subject-relationships. *International Classification*, 14(1), 1987, 11–18.
18. JOHANSEN, THOMAS. On the relationships of material subjects. *International Classification*, 14(3), 1987, 138–144.
19. JOHANSEN, THOMAS. Om sammensatte emners struktur. In: *Orden i papirerne – en hilsen til J.B. Friis Hansen*. Redigeret af Ole Harbo og Leif Kajberg. København: Danmarks Biblioteksskole, 1989, 157–165.
20. STEIGER, ROLF. Zu philosophisch-weltanschaulichen Fragen der Informationssprachen. *Informatik*, 20, 1973, 52–55.
21. BOOKSTEIN, ABRAHAM and SWANSON, DON R. A decision theoretic foundation for indexing. *Journal of the American Society for Information Science*, 26(1), 1975, 45–50.
22. SOERGEL, DAGOBERT. *Organizing information: principles of database and retrieval systems*. London: Academic Press, 1985.
23. DAHLBERG, INGETRAUT. *Grundlagen universaler Wissensordnung. Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens*. München: Verlag Dokumentation, 1974.
24. POPPER, KARL R. *Objective knowledge: an evolutionary approach*. Oxford: Clarendon Press, 1972.
25. RUDD, DAVID. Do we really need World III? Information science with or without Popper. *Journal of Information Science Principles & Practice*, 7, 1983, 99–105.
26. HJØRLAND, BIRGER. Indledende betragtninger over faglitteraturens typologi og udtryksformer. *Biblioteksarbejde*, 29, 1990, 35–50.
27. SPANG-HANSSEN, HENNING. Kunnskapsorganisasjon, informationsgjennfinning, automatisering og språk. In: *Kunnskapsorganisasjon og informationsgjennfinning. Seminar arrangert 3.–7. desember 1973 i samarbeid mellom Norsk hovedkomite for klassifikasjon, Statens Biblioteksskole og Norsk Dokumentasjonsgruppe*. Oslo: Riksbibliotekstjenesten, 1974, 11–61. (Skrifter fra Riksbibliotekstjenesten, Nr. 2)
28. BOSERUP, IVAN. Hvad er emnedata? In: *Emnedata i online-alderen*. Under redaktion af Niels-Henrik Gylstorff, Niels C. Nielsen og Morten Laursen Vig. Danmarks Forskningsbiblioteksforenings Internatmøde Nyborg Strand 7.–8. februar 1984. København: Bibliotekscentralens Forlag, 1984, 31–42.
29. HJØRLAND, BIRGER. Information retrieval in psychology: implications of a case study. *Behavioral & Social Sciences Librarian*, 6(3/4), 1988, 39–64.
30. KRÖBER, GÜNTER and SEGETH, WOLFGANG. Beschreibung. In: KLAUS, GEORG and BUHR, MANFRED. *Marxistisch-Leninistisches Wörterbuch der Philosophie I–III*. Reinbeck bei Hamburg: Rowohlt, 1983, Band I, 214.
31. KRARUP, KARL and BOSERUP, IVAN. *Reader oriented indexing: an investigation into the extent to which subject specialists should be used for the indexing of documents by and for professional readers, based on a sample of sociological documents indexed with the help of the PRECIS indexing system*. Copenhagen: The Royal Library, 1982.
32. WINOGRAD, TERRY and FLORES, FERNANDO. *Understanding computers and cognition: a new foundation for design*. New York: Addison-Wesley, 1987.
33. SEGETH, WOLFGANG. Prädikat. In: KLAUS, GEORG and BUHR, MANFRED. *Marxistisch-Leninistisches Wörterbuch der Philosophie I–III*. Reinbeck bei Hamburg: Rowohlt, 1983, Band III, 961–962.

34. SMITH, EDWARD E. Concepts and induction. *In*: POSNER, MICHAEL I., ed. *Foundations of cognitive science*. Cambridge, Mass., London: MIT, 1989, 501–526.
35. BEGHTOL, CLARE. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, 42, 1986, 84–113.
36. SWIFT, D.F., WINN, V. and BRAMER, D. 'Aboutness' as a strategy for retrieval in the social sciences. *Aslib Proceedings*, 30, 1978, 182–187.
37. LANGRIDGE, D.W. *Subject analysis: principles and procedures*. London: Bowker-Saur, 1989.
38. WICKLUND, ROBERT A. *Zero-variable theories and the psychology of the explainer*. Berlin: Springer, 1990.
39. HOLZKAMP, KLAUS. *Grundlegung der Psychologie*. Frankfurt, 1983.

(Revised version received 4 November 1991)