

The Usefulness of Standardized Definitions in Thesauri: An Assessment Through Interindexer Consistency Measurements

Michèle Hudon
École de bibliothéconomie et des
sciences de l'information
Université de Montréal
hudonmi@ere.umontreal.ca

A study was conducted to determine whether the availability of standardized definitions in a thesaurus of descriptors used as indexing aid could raise levels of terminological consistency among novice non specialist indexers. This paper recalls the project's objectives, briefly describes the methodology, and discusses some of the most interesting findings.

Brève présentation d'une recherche ayant pour but de déterminer si la disponibilité de définitions normalisées dans un thésaurus de descripteurs utilisé comme outil d'aide à l'indexation pourrait augmenter la cohérence terminologique entre indexeurs novices et non spécialistes. Rappel des objectifs du projet, description de la méthodologie et discussion de certains résultats intéressants.

Introduction

Indexers working with a thesaurus are expected to consistently select the same descriptor(s) to represent a subject, thus creating predictable clusters of documents on a topic. They cannot do so, however, if they do not perceive clearly the meaning of each descriptor. The meaning of thesaurus descriptors is to be found in their immediate environment. It is believed that, in a well-constructed thesaurus, a term is self-defined by its explicit relationships with other terms in the vocabulary, and that thesaurus users will be efficiently guided by these semantic links to appropriate indexing terms.

Given that low levels of terminological consistency among indexers¹ have been observed repeatedly over the past thirty years, it would appear

that the defining information provided with thesaurus terms might be insufficient or inappropriate, causing inconsistency in descriptor selection, and leading to vagueness in information and knowledge organization. The inconsistency problem is more apparent among novice indexers unfamiliar with the complex structure of controlled indexing languages.

Since the thesaurus is, and is likely to remain for years to come, an essential indexing aid in many types of information transfer environments, it seems important to consider at this time alternate and more direct ways of providing indexers with much-needed semantic information about individual subject descriptors.

In termbanks, a product of terminology² work similar in content and structure to the indexing and retrieval thesaurus, one or more standardized definitions accompany each term to prevent misunderstanding, misinterpretation and misuse, and to ultimately facilitate scientific communication. We believe that there might also be a place and a role for standardized definitions in the thesaurus. For the indexer, an access to standardized definitions in the thesaurus itself may make it easier to be accurate and consistent in term selection and assignment. The availability of standardized definitions may also make it possible for individuals who do not relate well to complex networks of semantic relationships (novice indexers, for example) to make better use of a controlled indexing language.

The focus of our study (Hudon 1997) was thus on the potential usefulness of definitions in the thesaurus of descriptors used as indexing aid. Looking more closely at the behaviour of novice non specialist indexers, our research project was designed and conducted with the aim of providing an answer to the following two research questions:

1. Does the availability of standardized definitions in a thesaurus of descriptors, as an addition to the conventional display of semantic relationships generally provided in such a tool, lead to an increase in consistency among novice non specialist indexers?
2. Does the availability of standardized definitions in a thesaurus of descriptors lead to acceptable levels of consistency among novice non specialist indexers?

who do not have access to the conventional display of semantic relationships generally provided in such a tool?

Methodology

The project's general objectives were:

1. to create and to integrate into a traditional thesaurus structure a different type of semantic information, namely a set of standardized definitions;
2. to assess the usefulness to the indexer of this new type of semantic information, by way of calculating and comparing interindexer consistency measurements.

A two-phase project was designed to reach these objectives. The first phase of the project involved the development of a prototype thesaurus including definitions written specifically for our purpose with the help of a defining model borrowed from the field of terminology. In the second phase of the project, the prototype thesaurus was used as indexing aid in a controlled indexing experiment.

Developing the Core Literacy Thesaurus

The *Canadian Literacy Thesaurus (CLT)*⁴, which describes a field characterized by ill-defined and shifting conceptual boundaries, and by fluctuating terminology, was our source for the creation of a prototype thesaurus in the field of adult literacy theory and practice. Despite an obvious need for clarification of meaning and specification of differences between terms as similar as Literacy consultants, Literacy Coordinators, Literacy facilitators, Literacy instructors, Literacy practitioners, Literacy specialists, and Literacy workers, 398 definitions only are provided in the current edition of *CLT* which includes a total of 1,890 terms.

367 descriptors and the corresponding 243 non-descriptors representing core concepts in the field of literacy were extracted from the main thesaurus file and transferred to the *Core Literacy Thesaurus* file. All semantic relationships linking terms were transferred with the descriptor from the source to the prototype thesaurus file, but existing definitions were stripped from the descriptor records.

Standardized definitions were then written for all descriptors selected to appear in the *Core Literacy Thesaurus*, regardless of the amount of semantic information provided in the rest of their records through relationships. Following practices recommended by terminologists, existing definitions for the terms were first collected in a wide variety of appropriate specialized sources (dictionaries, textbooks, thesauri, etc.) Those definitions were then rewritten to conform to a model used in terminology to produce a particular type of analytical definition. In the analytical definition, the concept being described (i.e. the *definiendum*) is referred to the broad class of concepts to which it belongs, and some of its essential characteristics (i.e. the *differentia*) are presented. A defining model proposed by Sager and L'Homme (1994) provided most of what was necessary to create the flexible template needed to write definitions for the 367 descriptors in our corpus⁵. A few examples of definitions created are given below. Italics are used in the definitions to identify terms defined elsewhere within the same corpus.

After school literacy programs = *literacy programs* designed for and offered to adolescents after regular school hours.

Functional literacy = *ability* of an individual to read and write at the level required to cope with the demands of everyday life in his or her family, community, and workplace.

Reading across the curriculum = educational movement advocating the incorporation of *reading activities* into *instructional strategies* for all subjects and disciplines, to help *learners* develop, improve, and/or retain their *reading skills*.

Using the Core Literacy Thesaurus for indexing

Three versions of the prototype thesaurus were created. The standard version (C) of the thesaurus contained no definitions. The augmented version (A) of the thesaurus provided at least one standardized definition for each descriptor, in addition to the conventional display of semantic relationships of equivalence, hierarchy, and association generally found in such a tool. The stripped version (S) of the thesaurus provided the same definitions, but did not display hierarchical and associative relationships among terms.

The controlled indexing experiment set up to obtain consistency measurements for comparison and analysis was conducted in conditions similar to those of previous consistency studies⁶. Twenty-five novice indexers, randomly assigned to one of three test groups, were asked to index a collection of twelve informative abstracts using one of the three versions of the prototype thesaurus, in a period of two hours. Indexers were asked specifically to assign a minimum of four and a maximum of eight descriptors to represent the subject content of each one of the abstracts, and to identify a "main" descriptor for each source document.

Hooper's indexer-pair consistency formula (i.e. the ratio of the number of term matches to the total number of unique terms assigned by two indexers) was used to calculate consistency ratios; group consistency figures were obtained by averaging indexer-pair results for each document indexed by all participants in the study. One-tailed paired *t* tests, based on differences in group consistency in descriptor assignment, were run to test for statistical significance.

Results

In this study, the test groups, as well as the sample collection, were quite small. The pairing of indexers within each group, however, allowed us to gather a considerable number of results representing as many separate observations. A total of 2,152 pair-results were used in the analysis.

Consistency in assignment of complete sets of descriptors ("all")

Our first research hypothesis relating to consistency in assignment of a complete set of descriptors to a document was:

On average, novice non specialist indexers working with an augmented thesaurus will achieve higher levels of interindexer terminological consistency than novice non specialist indexers working with a standard thesaurus.

This hypothesis was not supported by the data obtained through the indexing experiment. In fact, indexers in group A were more consistent among themselves than indexers in group C in less than half of the test cases (i.e. five out of twelve).

The second research hypothesis relating to consistency in the assignment of a complete set of descriptors to a document was:

On average, novice non specialist indexers working with a stripped thesaurus will achieve levels of interindexer terminological consistency at least equal to those of novice non specialist indexers working with a standard thesaurus.

This hypothesis was not supported either by the data obtained through the indexing experiment. Indexers in group S were equally or more consistent among themselves than indexers in group C in three cases only. Statistical testing provided evidence that indexers in group S were less consistent on average than indexers in group C.

Consistency in main descriptor selection ("main")

The first research hypothesis relating to consistency in the selection of a main descriptor for a document was:

On average, novice non specialist indexers working with an augmented thesaurus will achieve higher levels of interindexer terminological consistency in their selection of a main descriptor for a document than novice non specialist indexers working with a standard thesaurus.

This hypothesis was only weakly supported by the data obtained through the indexing experiment. Indexers in group A were more consistent than indexers in group C in their selection of main descriptors for seven documents. Statistical testing revealed that the null hypothesis could not be rejected at the 5% significance level, but would be rejected at the 10% level.

The second research hypothesis relating to consistency in the selection of a main descriptor for a document was:

On average, novice non specialist indexers working with a stripped thesaurus will achieve levels of interindexer terminological consistency in their selection of a main descriptor for a document at least

equal to those of novice non specialist indexers working with a standard thesaurus.

This last research hypothesis was supported by the data obtained through the indexing experiment. Indexers in group S were more consistent in their selection of a main descriptor for a document than indexers in group C in eight cases out of twelve. Statistical testing provided no evidence that indexers working with the stripped thesaurus were less consistent in their selection of main descriptors than indexers working with a standard version of the same thesaurus.

Discussion and Conclusions

Our data analysis, based on an assessment of consistency which did not involve a judgment on the accuracy, appropriateness, and relative importance of descriptors assigned, suggests the following interpretation of the results:

1. the availability of standardized definitions in a thesaurus does not lead to an increase in overall terminological consistency among novice non specialist indexers;
2. the availability of standardized definitions in a thesaurus which does not display the conventional associative relationships among terms is not sufficient to prevent a decrease in overall terminological consistency among novice non specialist indexers;
3. the availability of standardized definitions in a thesaurus may lead novice non specialist indexers to acceptable levels of consistency in main descriptor selection, even if the thesaurus does not display the conventional associative relationships among terms.

It would thus appear that the writing of standardized terminological definitions for integration into a thesaurus used as indexing aid may not be worth the effort and the cost, at this time and in most environments, where novice indexers would be the main users.

This preliminary conclusion is justified and valid in the framework of this study, given the data collected. We could not affirm without a doubt, however, that the standardized definitions integrated into the *Core*

Literacy Thesaurus were not at all helpful to the participants in this experiment. This investigation focussed on outcomes rather than on process. Although we found that the availability of standardized definitions at the time of descriptor selection did not make a significant positive difference in indexing outcomes, our research design did not allow us, unfortunately, to determine if and how the indexing process itself had been influenced by the availability of a new type of semantic information in a traditional indexing language structure. Posttest interviews with the participants would have been useful to gather more information on the behaviour of the indexers in each of the three test groups, and on their reaction to the content of the indexing tool they had used to index the sample collection. Any further research in this area should include this supplemental tool for more "qualitative" data collection.

Given the scope and limitations of our project, it is evident that much more research is needed before we can reach a definitive conclusion on the usefulness of standardized definitions in the thesaurus as indexing aid. It would be interesting, for example, to verify whether terminological consistency among experienced indexers would increase if they were given access to standardized definitions. In a study involving trained indexers, both specialists and non specialists should be recruited, making it possible also to compare the reactions of each group to the availability of clear defining information. An investigation conducted in a "natural" rather than experimental setting might also lead to interesting discoveries on the relationship between standardized definitions and indexing consistency. In a natural setting, indexers would be allowed to work at their own rhythm, to perform postindexing editing, to access and review previously assigned descriptor sets, etc.

Important changes have taken place in the information transfer environment since the thesaurus of descriptors' first appearance on the scene forty years ago. Technological innovations, and more recently economic restrictions, have contributed to a revision of general views and opinions on the value and necessity of human-based indexing. Human-based subject indexing is still considered at this time value-added to the information transfer process, but the increasing capabilities of search engines for accessing full texts or document surrogates have modified considerably the tasks and responsibilities of the indexer. In many systems, the number

of descriptors an indexer is required to assign has decreased; the one or two descriptors assigned have now acquired greater importance, and they must be accurate, precise, and highly efficient. The amount of information available makes it impractical to continue to use generic terms to group related documents and information. Indexers must use specific descriptors, and they must be provided with tools that will help them do so consistently.

Further research should concentrate on the link between definitions and consistency in main descriptor selection and assignment. In this study, we did a separate calculation on a single main term; in any further study, it would be appropriate to require the identification of the two (or even three) most important descriptors since that many terms are often needed to express a subject.

Particular emphasis should be put on the use of the stripped version of the thesaurus for main descriptor selection. The current practice in many organizations of making documents widely available on the Internet, and of building complex intranets for their own internal operations, has exacerbated the access problem that existed before but was not as visible and probably not as critical. These organizations get quickly to the point of needing a controlled list of terms and/or categories to structure this mass of information. Most organizations are reluctant, and rightly so, to get into thesaurus construction, but they can rarely find a list which is totally appropriate to their needs. A solution might be found in a new type of tool, such as our experimental stripped thesaurus, which offers basic control of synonyms and meanings, without involving huge costs in development and maintenance. Much of the defining work could be shared with terminologists and translators who are already pursuing the goals of tracking concepts and standardizing term use in these same organizations.

End Notes

1. Interindexer terminological consistency is a quantitative measure of the degree to which two or more indexers agree in their selection and assignment of index terms to represent indexable concepts in a document..
2. Terminology is this field of knowledge and practice which deals with concepts and their verbal representations.

3. Hudon, M. 1997. An Assessment of the Usefulness of Standardized Definitions in a Thesaurus Through Interindexer Terminological Consistency Measurements. Ph.D. Diss., Faculty of Information Studies, University of Toronto.
4. Toronto : Canadian Literacy Thesaurus Coalition, 1996. 2nd ed..
5. The original model is described in: Sager, J.C., and M.C. L'Homme. 1994. A model for the definition of concepts: rules for analytical definitions in terminological databases. *Terminology* 1:351-373. A description of the template used in my own study is found in: Hudon, M. 1996. Preparing terminological definitions for indexing and retrieval thesauri: a model. In: *Knowledge Organization and Change: Proceedings of the Fourth International ISKO Conference, 15-18 July 1996, Washington, DC, USA*, 363-369. Frankfurt: Indeks.
6. Previous interindexer consistency studies are described in Leonard, L.E. 1977. *Inter-indexer Consistency Studies, 1954-1975: A Review of the Literature and Summary of Study Results*. Urbana, IL: University of Illinois, Graduate School of Library Science.