Effects of Subject Specificity: Part I: Specificity of LC Subject Headings and Depth of Subject Analysis in Monographic Records

A. Khosh-Khui

ABSTRACT. Subject headings in 7834 LC monographic MARC records were analyzed to determine relationship between the specificity of LC Subject Headings and the number of subject headings per bibliographic record. The findings indicated that a weak inverse correlation existed between the specificity of LC Subject Headings and the number of subject headings assigned to each document. The study concluded that using more specific subject headings does not reduce the mean of headings per bibliographic record. The number of headings assigned to each document most likely depends on the variation of topics discussed in a document, the availability of specific headings to cover multiple topics, and local management policy.

Specificity is the extent to which a system allows precision in explicitly stating the subject contents of a document, while the depth of subject analysis or exhaustivity refers to the extent to which a document is being analyzed to completely identify the contents of a document.[1] Specificity is a function of the system and is responsible for the accuracy of document representation and thus controls the precision capabilities. Exhaustivity is the result of management decision and has to do with completeness of document representation and controls recall potentialities.[2]

Specificity is a relative term, and the concept is not precisely defined in the literature. Subject specificity may be discussed from

59

different points of view. It may refer to the basic concept as distinguished from many aspects of the subject of a document. This view has been seen implicitly in the library subject headings.[3] It may also be extended to cover a variety of aspects of the subject. It may be applied to hierarchical subordination of one term with respect to another, the extent of precision to which a term describes a document, or, it may also be applied to the guidelines for assigning the most specific headings to cover the content of multi-subject documents.

The purpose of this study is to determine if there is a significant correlation between the extent of subject specificity and indexing exhaustivity in monographic records. Does the average number of subject headings per bibliographic record significantly reduce when more specific subject headings are used to represent the topics discussed in a document? This paper attempts to analyze the association between specificity of subject headings and the total number of subject headings assigned to a document.

LITERATURE REVIEW

The probable association between subject specificity and number of subject headings per bibliographic record was first mentioned by Black.[4] He pointed out that there is a relationship between subject specificity and the number of headings assigned to documents. Black suggested that the use of specific headings reduces the average number of headings per title, but he added that this is a question to be tested. Since Black's suggestion for such a study, several papers have appeared on subject specificity, but this particular question is not addressed in the literature. This may be due to the fact that the concept of specificity is a difficult attribute to define and to measure.

The principle of unity requires that a subject catalog must bring together under one heading all materials which deal principally with one subject. The principle of specific and direct entry, on the other hand, requires that each subject in the catalog be entered under one name and in one form. This implies that each work has to be entered under a heading which expresses the topic of that work precisely rather than entering the work under a broad subject. As Wilson pointed out, the idea of specific entry does not mean that a work is to be assigned only a single subject heading.[5] If no single heading

is available that accurately and specifically describes the topic of a work, then, rather than using a broader heading, two or more specific headings which will approximately cover the contents of the document should be used. If a work does not have a single topic and a single subject heading is not specific enough to describe it, then additional subject headings are assigned.

The Library of Congress Subject Heading has often been criticized for its lack of in-depth analysis and its lack of representing the contents of library collections in research libraries. Chen believes that, in the course of the development of the Library of Congress Subject Headings, views on the functions of the subject catalog varied according to differing demands of users. This caused the absence of a consensus on functions, which in turn caused difficulties in defining basic principles, such as specificity and usage, in both the construction and the application of subject headings.[6]

The principle of uniform headings as applied to the construction of the Library of Congress Subject Headings, suggests essentially a single-heading for each subject as opposed to multiple-headings for the same subject.[7] However, if two parts of a heading are equally important, two subject headings are created: "When two components of the heading are equally significant and a reasonable choice between the two cannot be made, each component is listed as an entry element in LCSH."[8]

Haykin identified availability of appropriate terms as the determinant factor of subject specificity.[9] "It is language which is the determining factor in the question of specificity." He said that the specificity of ideas, facts, and relationships may be beyond the limit of availability of the subject headings.

Angell identified two types of specificity: (1) the specificity that is possible as determined by the document in hand and (2) the specificity that is desirable, which is not determined by the document but by the characteristics of the demands which are made upon an information system in a particular application or installation.[10]

Lilley described the relative nature of subject specificity, and pointed out that the concept of specificity is not constant.[11] Specificity may vary, he said, because (1) specificity is in part a function of a subject area, that is, the same word can vary from the extremely general to the extremely specific depending on the context, and (2) specificity is a function of the library because the same book can be described at various levels of specificity depending on the users of the library. Part of Lilley's dissertation research

dealt with the question of specificity of subject headings in the field of English literature. He found that "the headings at the same level of specificity may be employed for such a wide variety of purposes that their exact uses are highly unpredictable on the basis of terminology and form alone."[12]

Steinweg believed that the "specificity of the subject covered by the work should be matched by the specificity of subject headings to cover it."[13] She examined the methods by which specificity is achieved in the Library of Congress Subject Headings. Specificity is achieved by using modifiers in parentheses, adjectival and phrase modifiers, dates, subdivisions by place, time, form, and topic; but the most obvious way to achieve specificity is to use the appropriate term for any given level of specificity.

Wilson supported the idea of specific entry, that is, an entry under a heading which expresses the topic of a work precisely rather than a broader heading in general, and he disapproved of the policy of the Library of Congress in providing duplicate entries at both specific and general levels.[14]

Greenberg examined characteristics of 1,892 scope notes in the eighth edition of the Library of Congress Subject Headings.[15] He identified three functions of scope notes which can ensure the consistency of usage: specifying the range of subject matter to which a subject heading can be applied; drawing necessary distinctions between related subject headings; and stating which of several meanings of a term is the one that its usage in the library catalog is limited.

The above review implies that the number of subject headings assigned to a record is a factor of: (1) the availability of specific subject headings that precisely describe the content of a document, (2) the number of topics discussed in a document, and (3) the degree of precision demanded in a particular system. This study examines the effect of the first factor on the number of subject headings per bibliographic record.

METHODOLOGY

Measures of Specificity

The combination of the following factors was used as the measure of specificity in this study: the level of subject subdivisions; the number of words per subject heading; and the number of

characters per heading. Many examples from the Library of Congress Subject Heading List can be cited where the specificity of subject headings increases as more topical or form subdivisions are added to the main subject headings:

Roads
Roads—Riding qualities
Roads—Riding qualities—Testing
Roads—Riding qualities—Testing—History

The use of modifers is another way to achieve specificity. Thus, in many cases, an increase in the number of words in the subject headings increases the degree of subject specificity:

Libraries
Libraries and the handicapped
Libraries and mentally handicapped children

There are other cases where an increase in the characters per subject heading implies greater specificity:

Pressure
Pressure-gage
Man
Man-machine systems
Man-made fiber industry

Measures of specificity in this study have certain limitations. These measures are limited to the specificity of subject headings as recognizable by the format of the representation. There are other ways that specificity may be achieved. The measures used in this study identify neither the cases where specificity is recognizable through the meaning of the term, nor the cases where scope notes are used to define the degree of specificity.

Sample

A ten percent sample of the MARC records produced by the Library of Congress during the year 1980–81 was selected through systematic sampling. There were 7,834 bibliographic records in the sample which had a total of 15,072 LC Subject Headings. All of the

subject headings for each selected record were retrieved. Each subject heading in the sample was coded according to the degree of specificity. This was done by counting the number of subdivisions in each subject heading, the total number of words per subject heading, and the number of characters per heading.

RESULTS

All subject headings were analyzed for their degree of specificity in relationship to the number of subject headings per bibliographic record to determine if there was a significant association between these two variables. The results of the analysis showed that a weak inverse relationship existed between measures of specificity and the number of subject headings per bibliographic record. As Table 1 shows, there was −.04 correlation between the number of subdivisions per subject and the number of subject headings per bibliographic record. Similarly, the correlation between the number of words per subject heading and the number of subject headings per bibliographic record was −.04. There was only −.02 correlation between the length of each subject heading and the number of subject headings per bibliographic record.

All selected LC Subject Headings were further analyzed to compute the probability of the number of subject headings per

TABLE 1

Correlation of Subject Specificity with the Number of

Subject Headings Per Bibliographic Record

Library of Congress Subject Specificity	Mean	Standard Deviations	Correlation	Significance of r
Subdivision Per Subject	1.11	0.88	−.04	.001
Words Per Subject	2.69	1.37	−.04	.001
Characters Per Subject	28.41	13.26	−.02	.216

bibliographic record for subject headings with different levels of specificity. The mean number of subject headings per bibliographic record for LC Subject Headings with different levels of specificity decreased as the level of specificity increased. That is, whenever the number of words per subject heading increased, there was a slight reduction in the number of the Library of Congress Subject Headings per bibliographic record. The chi square test of the difference between the average Library of Congress Subject headings per bibliographic records for subject headings with different levels of specificity was significant at the .05 level.

The correlation measures between the number of subject headings per bibliographic record, the subject headings subdivision level, and the length of subject headings were computed to determine if the pattern of variation could be identified. In addition, the measures of specificity were combined to determine if the combination of measures would reveal any new information. The results showed that no recognizable pattern of variation existed between the variables. The correlation between the increment in the number of subject headings per bibliographic record and the variation in the LC Subject specificity was not statistically meaningful. The combination of measures of specificity resulted in the disappearance of the negative signs in correlation values. This was an indication that when the measures of specificity were combined, each cancelled out the specificity effect of the other.

Finally, the degree of association was separately computed for the four groups of bibliographic records, that is, for the groups of records with one, two, three, and four subject headings. The results revealed a slight difference between measures of association in each group. As shown in Table 2, the chi square test of difference between measures of association in each group was not statistically significant.

CONCLUSIONS

The study showed that there is no statistically significant correlation between the number of subject headings per bibliographic record and the level of subject specificity. This leads to the conclusion that most likely the number of subject headings per bibliographic record is a factor depending upon the number of topics discussed in a document and upon the availability of specific subject

TABLE 2

Means and Standard Deviations of Subject Specificity

Measures in Four Groups of Bibliographic Records

No of LCSH Per Record	No of Records	LCSH Subdivision Level		LCSH Lenght		No of Words Per LCSH	
		X	S	X	S	X	S
1	3340	1.19	.89	28.99	13.68	2.77	1.44
2	2510	1.12	.87	28.85	13.15	2.77	1.37
3	1224	1.12	.88	29.22	13.26	2.79	1.37
4	760	1.08	.81	28.92	13.15	2.76	1.39
Total	7834	1.11	.88	28.41	13.26	2.69	1.41
Chi Square		.782		.022		.782	
Significance		.10		.20		.50	

headings to cover such topics when the work is being cataloged. It is thus basically a pragmatic managerial decision. The specificity of the subject headings per se does not affect the number of subject headings per record.

REFERENCES

1. A. C. Foskett, *The Subject Approach to Information*, 4th ed. (Hamden, CN: Linnet Books, 1982), p. 25–26.

2. John Balnaves, "Specificity," *The Variety of Librarianship: Essays in Honour of John Wallace Metcalfe*, W. Boyd Rayward ed. (Sydney, Australia: Library Association of Australia, 1976), pp. 47–56.

3. Jessica L. Milstead. *Subject Access Systems: Alternative in Design* (Orlando: Academic Press, 1984), p. 143.

4. Henry Black. "An Approach to a Theory of Subject Headings," *College and Research Libraries* 7:244–48 + (July 1946).

5. Patrick Wilson. "The End of Specificity," *Library Resources & Technical Services* 23:116 (Spring 1979).

6. Lois Mai Chan, *Library of Congress Subject Headings: Principles and Application* (Littleton, CO: Libraries Unlimited, 1978), p. 20

7. Ibid. p. 25.

8. Lois Mai Chan, "The Principle of Uniform Heading in Library of Congress Subject Headings," *Library Resources & Technical Services* 22:134 (Spring 1978).

9. David Judson Haykin, "Subject Headings: Principles and Development," In *The Subject Analysis of Library Materials* ed. Maurice F. Tauber Papers Presented at an Institute. June 24–28, 1952, Under the Sponsorship of the School of Library Service. Columbia University, and the A.L.A. Division of Cataloging and Classification. (New York: School of Library Service, Columbia University, 1953), p. 48.

10. Richard S. Angell, "Standards for Subject Headings: A National Program," *Journal of Cataloging & Classification* 10:193 (October 1954).

11. Oliver Linton Lilley, "How Specific is Specific?," *Journal of Cataloging & Classification* 11:3 (January 1955)

12. Oliver Linton Lilley, "Terminology, Form, Specificity and the Syndetic Structure of Subject Headings for English Literature." (DLS dissertation, School of Library Service, Columbia University, 1959), p. 400.

13. Hilda Steinweg, "Specificity in Subject Headings," *Library Resources & Technical Services* 23:55 (Winter 1979).

14. Wilson, "The End of Specificity," (see ref. # 5), p. 116.

15. Alan M. Greenberg, "Scope Notes in Library of Congress Subject Headings." *Cataloging & Classification Quarterly* 1:95 (1982).