City University of New York (CUNY)

CUNY Academic Works

Publications and Research

College of Staten Island

2014

Controlling Our Vocabulary: Language Consistency in a Library Context

Mark Aaron Polger

CUNY College of Staten Island

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/si_pubs/72

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

Controlling our vocabulary: language consistency in a library context

Mark Aaron Polger

As a result of his experience as an interim academic web/systems librarian, Mark Aaron Polger embarked on a study of consistency of terminology in a library context, looking at usage across three media – promotional material, signage and websites. In this article, after reviewing the literature, he reports the results of his study, points out its limitations and suggests ways in which the work could be taken forward.

Introduction

This study originated in my experience as a web/systems librarian at an academic institution in Toronto, Canada, where I was hired to maintain the library website while the permanent librarian was seconded to another position. I found a great deal of inconsistency in the website terminology, and I volunteered to undertake an audit of inconsistent terms and to create a controlled vocabulary for future use. This work led me to reflect further on the importance of consistency of vocabulary, and the extent to which this consistency was lacking, a point borne out by my review of the literature with which I begin this article.

Many people visit a library's website to obtain basic information such as library hours and policies, or to access the library catalogue. They also, particularly when they visit the library itself, rely on promotional material and signage to guide, inform and educate themselves on the library services and resources available to them. The library's website, promotional material and signage must be informative, inviting and user-friendly, and if they are to be fully effective in promoting the library as a valuable service, they need to relay a message that is consistent across the three media.

This study compares the language consistency of five library terms across the three media: websites, signage and library promotional materials (brochures, posters, flyers, newsletters, bookmarks, and business cards sometimes known as 'swag'). My sample, for reasons I explain, was smaller than I would have wished, and there were other limitations which I point out in my concluding section. It is a first step towards my aim of demonstrating my premise that consistency of language would help librarians promote the value of their services, and that to achieve this they should collaborate on developing a controlled vocabulary.

Literature review

There is an extensive literature on consistency in a library-related context, on consistency of controlled headings, on inter-indexer consistency (IIC), on thesaurus design and on the use of controlled headings in library catalogues, which I review here. I have omitted literature relating to visual consistency of user interfaces, as this study focuses on language-related consistency.

J. I. Iwe (2001) analyzed the semantic consistency of two

Library of Congress Subject Headings (LCSH), 'Literature' and 'Librarianship'. She compared the syntactic structures of the two terms, their word structure and prepositional phrasing from the 19th edition of the LCSH. She found real shortfalls in the level of consistency between the two LCSH. Some headings in the two lists such as 'Literature and society' and 'Libraries and society' are essentially identical. 'Literature and mental illness' is similar to 'Libraries and the mentally ill' but not entirely consistent. 'Library education' and 'Literature-study and teaching' are not consistent. Some terms such as 'Library fund raising' or 'Library science teachers' cannot be matched. There is no equivalent in 'Literature' to match these terms in 'Librarianship'. She concluded that for LCSH the application of the headings is more important than their consistency (Iwe, 2001).

Li (2004) explored the inconsistencies in Chinese language cataloguing, identifying the following problems:

- Inconsistency in romanizing: Li found four different ways of romanizing the same Chinese word.
- Incorrect spacing which affected the meaning of the word.
- Lack of cultural knowledge of China. Li identified errors in the catalogue record such as Shanghai University Press being shown as publishing in Beijing. It is clear that the cataloguer did not have even a basic knowledge of Chinese geography.
- Use of incorrect currency. Li gives examples of catalogue records where the currency should be renminbi but is coded as US dollars.
- Incorrect language codes in the MARC record. Li points to technical inconsistency, with materials being catalogued using the incorrect language code.
- Inconsistent use of traditional or simplified Chinese characters. Li recommends the use of simplified ones.

Hearn (2008) limited his evaluation to the author and subject fields. He argues that for users to be able to locate all items under a single heading, consistent headings are necessary. He compared two source lists, the LCSH 'Weekly Lists' and the OCLC lists of 'Closed Dates in Authority Records', selecting 50 sample headings from these lists and comparing them with headings from the library catalogues of a group of 13 universities, all members of the Committee on Institutional Cooperation (CIC). Hearn found that most of the 13 catalogues were not consistent with Library of Congress (LOC) subject or author headings. For author headings, only three

catalogues matched. For subject headings, only four catalogues matched. The author suggests that his method serves as a useful way for measuring data quality, acknowledging that maintaining catalogue headings presents challenges, and currency and consistency can be difficult.

Sievert and Andrews (1991) addressed consistency in *Information Science Abstracts (ISA)*, reviewing duplicate records in *ISA*. They took 496 records from four publications from *ISA*: 65 records from *Online Review*, 128 records from *Database*, 106 records from the *Journal of Documentation* and 197 records from the *Journal of Information Science*. Out of the 496 records, the authors found 72 duplicate records, with an average consistency of 52.25 per cent for the main headings and 32.89 per cent for the identifiers. They also compared consistency percentages for *ISA* and *Medline*, finding that *ISA* has a lower indexing consistency percentage for all headings than *Medline*. They also noted the inverse relationship between indexing depth and indexing consistency. As indexing depth increased, indexing consistency decreased.

Monreal and Gil-Leiva (2011) took three LIS journal article abstracts, and using three different thesauri for analysis, looked for consistency between the analyses produced by four groups of people – three groups of students (second, fourth, and fifth-year information science students) and a group of professional indexers. Each student group comprised 18 people who were subdivided into three subgroups, each using one of the three thesauri. The remaining group of nine professional indexers was similarly subdivided into groups of three. The authors found that the mean consistency across three thesauri between the student indexers and the expert indexers was 39.5 percent, 29.8 percent, and 27.5 percent. Their findings suggest that inconsistencies in indexing are going to be present.

Symonds (2008) discussed the development of the Wabash Valley Visions & Voices: A Digital Memory Project (WV3), a collaborative digital collection involving Indiana State University departments and other cultural and historical organizations. One of the biggest challenges, she found, was maintaining consistency across collections, and she pointed to software limitations and human error as contributory factors. She also drew attention to a lack of standards in the metadata, the main consistency issues relating to the 'title,' 'subject,' 'repository,' and 'material type' fields, and argued that this was because of a failure to adopt the Dublin Core standard across the entire collection.

Agee (2008) discussed the process of maintaining controlled vocabulary in the cataloguing process, and pointed out that controlled vocabulary can come from a list, a synonym ring, a taxonomy, or a thesaurus. She also discussed the more recent use of folksonomies, giving the examples of photo-sharing sites such as Flickr to illustrate how folksonomies make it easier for users to access information. However they represent the antithesis of controlled vocabulary.

Fenton (2010) analyzed the use of controlled vocabularies and thesauri in UK archive finding aids, distinguishing between controlled vocabulary (i.e. subject headings) and thesauri (hierarchical listing of those subjects). She contacted 454 repositories, classifying them into academic,

local government, national, and 'other'. She received 298 responses, a 66 percent response rate. Fifty-five percent of respondents made use of a controlled vocabulary. Of those using a controlled vocabulary 32 percent developed an in-house thesaurus, 14 percent developed a specialized thesaurus, 31 percent used LCSH, 25 percent used the UNESCO thesaurus, and 25 percent used a UK thesaurus.

Falahati Qadimi Fumani (2010) looked at the inter-indexer consistency (IIC) from a Persian-language perspective. Three indexers were assigned to index 111 Persian scientific articles, each indexer selecting terms from the same thesaurus. Using the Hooper, Rolling, and Cosine evaluation measures, he found the IIC to be very low between the three indexers. He argued that collaboration between the indexers would help increase the consistency, and suggested that collaboration on other projects, such as the development of a controlled vocabulary for all library publications, might help prevent confusion. This suggestion is relevant to my own study because the language on signage, promotional material and websites may be developed differently by people who are not in communication with each other.

Portaluppi (2007) analyzed the consistency of 10 medical subject headings (MeSH) in chronobiology journal articles. The author downloaded 228 citations from Pubmed from 2005 to 2007, of which 70 articles related to animals, 122 involved humans, and four (including two editorials and a historical paper) presented experiments on both humans and animals. The remainder of the articles explored (32) were reviews. Portaluppi found 99 medical subject headings that were variants of the initial 10 headings. She found great variability in the consistency and accuracy of the MeSH headings.

Svenonius focused in her two articles (1986 and 2003) on the creation of controlled vocabularies, arguing (1986) that a collection of library 'publications' (web, signage, and promotional material) best serves its users if it uses a controlled vocabulary. In her 2003 article, she described the process of designing a controlled vocabulary, outlining the following steps: selection of terms with scope notes, context, literary warrant, use warrant, scholarly usage, and consistency. She maintained her thesis that controlled vocabularies should improve access, decrease confusion and help improve the user experience, but acknowledged that there may be situations where natural language is more appropriate.

Methodology

For my own investigation of consistency of terminology in a library context, I selected 24 library terms based on library literature, reference desk transactions, and classroom interactions. Table 1 shows the preliminary list of terms, divided into three categories: resource, service, and information. From these I selected five terms that appeared most frequently in promotional material, signage and on websites, and compared their use in these three media.

The five selected terms were:

- library catalogue
- databases
- guides

- interlibrary loan/document delivery
- job title.

I then compared these library terms with each other using two media at a time. This means, for example, that the term ‘library catalogue’ can yield three consistency counts:

- website vocabulary with promotional material
- promotional material with signage
- signage with website vocabulary.

To obtain consistency counts from other libraries, I ran an electronic questionnaire using SurveyMonkey in February 2012 (see the Appendix), sending a recruitment email containing a consent form, cleared by my Institutional Review Board, to six library listservs, using the ‘Bcc’ header to mask email addresses. The targeted listservs (ili-l (ALA), infolit (ALA), academicpr (ALA), librarymarketing (MLA), CANMEDLIB (CHLA), and MEDLIB-L (MLA)), had a total of approximately 11,000 subscribers. I chose these listservs because I wanted to recruit respondents from different types of libraries in the United States and Canada, including both large and small public, academic, and special libraries that develop signage, promotional materials, and websites. IP addresses were blocked and respondent data was anonymous.

This drew 450 responses, a low rate of return (5 percent). This is partly because many of those who received the questionnaire subscribed to more than one of the listservs so they received the questionnaire several times but naturally only completed it once. Moreover, since this was not a random sample, the data from the questionnaire is not representative of all libraries in Canada and the United States. Respondents chose to complete the questionnaire and only listserv subscribers knew the URL to complete the questionnaire.

Survey results and discussion

Most respondents were from the United States (89.5 percent). This is not surprising since the listservs targeted more respondents in the United States and the US population is ten times greater than Canada’s. This is likely to lead to greater representation from academic institutions from the United States.¹

Sixty-six percent of respondents were from academic institutions, which also included academic health institutions (medical libraries affiliated with universities). Nearly 20 percent of respondents identified themselves as hospital librarians/library staff who worked in community or non-teaching hospital library settings. Very few respondents came from public libraries (4.7 percent) or school libraries (5.6 percent). Only 4 percent of respondents were from special libraries. Since the six listservs targeted were primarily academic, it is not surprising that most respondents were from academic institutions. Perhaps the data would have been different had there been more respondents from public libraries.

Tables 2 to 4 examine the consistency count (N=450) by using nominal data (yes/no responses) using two media

Table 1 List of 24 library terms

Resource	Service	Information
Library catalog	Reference services	Library hours
Databases	Circulation services	URL of website
Guides	Interlibrary loan / Document delivery	Job title
Print periodicals	Ask a librarian	Cell phone policy
Microform readers	Course reserves/ Reserves	Food / drink policy
Photocopiers	Chat /Virtual/ Email Reference	Computer use policy
Scanners	Media services	Noise policy
Study carrels	Technical services	Gift/donation policy

at a time for comparison. Questions concerning business cards and email signatures were collapsed into the category of promotional material (which was defined on the questionnaire). Signage was assigned its own category.

It is important to emphasize that the sample (N=450), limited as it was to subscribers to the six selected listservs, is not representative of the population of librarians in the United States. The data distribution suggests that there may be a relationship between the consistency count and the publication type. For example, website vocabulary paired with either signage or promotional literature has a higher consistency count than signage paired with either publication type. This may suggest that signage has an overall lower internal consistency count than website vocabulary.

The consistency count relates to how respondents report on the consistency of the library term compared with two publication types. Respondents were asked whether the publication types, website and signage, were written consistently. The highest consistency count was for job title. Librarians’ job titles were the most consistent when comparing promotional material with signage (39.5 percent), as opposed to signage with website vocabulary (33.3 percent). Job titles in signage and promotional material were the least consistent (25.8 percent). Since library website vocabulary is typically more current and up to date, it may be more consistent with promotional material. It appears that when the library term is compared with the signage, the overall consistency count decreases. This might be because signs are less current than promotional material, or website content is updated more frequently.

The lowest consistency count was for the library term ‘guides’, with a mean consistency of 20.9 percent. Guides on the library website may be written as ‘research guides’, ‘web guides’, ‘subject guides’, or branded ‘LibGuides’. Further, promotional material may discuss the presence of web-based guides and list them as LibGuides, guides, or research guides. There is little or no signage on guides since they represent an electronic resource and many guides are not printed in hard copy.

For Interlibrary Loan (ILL), all three consistency counts appear to be relatively similar (mean consistency of

Table 2 Consistency counts, promotional material versus website vocabulary (N=450)

Library term	Promotional material versus website	
Job title	178	39.5%
Library catalog	159	35.3%
Databases	144	32.0%
ILL	124	27.5%
Guides	114	25.3%

Table 3 Consistency counts, website vocabulary versus signage (N=450)

Library term	Website versus signage	
Job title	150	33.3%
Library catalog	128	28.4%
Databases	100	22.2%
ILL	106	23.5%
Guides	99	22.0%

Table 4 Consistency counts, promotional material versus signage (N=450)

Library term	Signage versus promotional material	
Job title	116	25.8%
Library catalog	96	21.3%
Databases	86	19.1%
ILL	90	20.0%
Guides	70	15.6%

23.7 percent). This might be because ILL appears equally across all three media. On signage, ILL may be abbreviated while on the library website it may be spelled out. On promotional material, there may be a variety of ways in which it is written: 'ILL', 'Inter-library Loan and Document Delivery', 'Document Delivery and Interlibrary Loan', or simply 'ILL/DD'.

The consistency count of 'library catalogue' is significantly decreased when signage is compared with promotional material (21.3 percent) or websites (28.4 percent). When signage is removed and 'library catalogue' is compared for websites and promotional material, the consistency increases (35.3 percent). Signage for 'library catalogue' can be written in two ways, 'library catalogue' or 'find books', so this may affect its consistency (Polger, 2011). Promotional material may contain a branded version of the library catalogue, and its name may not match signs promoting the library catalogue workstations.

'Databases' consistency on signage is also low because the term does not typically appear on signs but mostly on promotional material and on the home pages of websites. This may be why 'databases' has a higher consistency count (32 percent) when comparing websites with promotional material.

Library terms that were compared using signage and promotional material appear to have the lowest consistency pattern (25.8 percent for job title, 21.3 percent for library catalogue, 19.1 percent for databases, 20 percent for ILL,

and 15.6 percent for guides). This might be because library terms on signage are less up-to-date than those used in promotional material, which tends to be time-sensitive.

Nearly 78 percent of respondents did not have a controlled vocabulary to manage language across the different media. While approximately 10 percent of respondents were thinking about developing a controlled vocabulary, only 4.5 percent had already created an in-house thesaurus. Developing a committee or working group to help create a controlled vocabulary for signage, promotional material, and website vocabulary might be useful. An in-house thesaurus might help identify which library terms are used the most often. A controlled vocabulary might help develop a consistent message to library users. Preparing a controlled vocabulary might also help in the development of a cohesive and consistent library marketing plan. Having a consistent set of controlled vocabulary terms might be helpful in the overall marketing message the library is trying to disseminate.

Limitations of the study

The study I have described had a number of obvious limitations which would need to be addressed in any future investigation:

- This study only looked at language-related consistency and did not address visual consistency issues such as font type, font colour, logos, graphics and overall design.
- Language consistency can be a relative matter unless the parameters are very clearly set. For example, interlibrary loan and inter-library loan are written identically with the exception of the hyphen. Some of the respondents to my questionnaire may have overlooked such small inconsistencies, which will have affected the result.
- The study did not address the question of whether the same people were responsible for developing the three media under consideration (promotional material, signage, website). It is, for example, possible that the library website may have been developed by IT professionals with very little consultation with library staff.
- The five library terms are not equally distributed across the three publications. 'Library catalogue', for example, may be listed more frequently on the library website and promotional material than on signage. This unequal occurrence of the library terms might affect the consistency count.
- Although the five library terms were selected on the basis of their frequency in the literature and in reference desk transactions, I acknowledge that there are other terms that could have been explored for consistency.

Future study projects

Taking into account the results of this initial study and its limitations, I suggest there would be advantage in future studies focusing on:

- Institutions that have developed controlled vocabulary

for their library, exploring the extent to which this affects consistency in practice in the areas I have discussed here, and how much use is made of the controlled vocabulary in other areas such as information literacy classes and on the reference desk. Are academic departments outside the library concerned with language consistency or is this only a concern to library and information science professionals?

- Responsibility for creating promotional material, signage and websites.
- Visual consistency.
- Language consistency in public libraries. Are public libraries more (or less) consistent than the academic libraries considered in this study?
- Language consistency on social networks such as Facebook, Twitter, Foursquare, or YouTube (a rather more difficult task given the lack of defining boundaries offered by closed institutional environments).

Appendix: Questionnaire

- 1 Where is your library located?
 - Canada
 - United States
 - Other
- 2 Type of library
 - Academic
 - Hospital (Community, Teaching, Geriatric, etc.)
 - Public
 - School
 - Special/Other (Government, Museum, Corporate, etc.)
- 3 Is your job title (on business cards) written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 4 Is your job title (on signage) written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 5 Is your job title (on business cards) written consistently with signage?
 - Yes
 - No
 - Not applicable
- 6 Is 'Interlibrary Loan and Document Delivery' in your signage written consistently with your promotional materials?
 - Yes
 - No
 - Not applicable
- 7 Is 'Interlibrary Loan and Document Delivery' in your signage written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 8 Is 'Interlibrary Loan and Document Delivery' in your promotional material written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 9 Is 'Databases' in your promotional materials written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 10 Is 'Databases' in your signage written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 11 Is 'Databases' in your promotional materials written consistently with signage?
 - Yes
 - No
 - Not applicable
- 12 Is 'Library Catalog' in your promotional materials written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 13 Is 'Library Catalog' in your signage written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 14 Is 'Library Catalog' in your signage written consistently with your promotional materials?
 - Yes
 - No
 - Not applicable
- 15 Is 'Guides' in your promotional materials written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 16 Is 'Guides' in your signage written consistently with the library website?
 - Yes
 - No
 - Not applicable
- 17 Is 'Guides' in your signage written consistently with your promotional material?
 - Yes
 - No
 - Not applicable
- 18 Are there any plans to develop a thesaurus (i.e. authority list) of consistent library terms used in library signage, library promotion, and the library website?
 - Yes (great idea)
 - No
 - Not applicable
 - We're thinking about it
 - We've already developed this list
 - We can't manage a list

Notes

- 1 See www.aacu.org/membership/list.cfm vs. www.aucc.ca/canadian-universities/our-universities/

References

- Agee, V. (2008) 'Controlling our own vocabulary: a primer for indexers working in the world of taxonomy.' *Key Words* **16**(1), 30–1.
- Falahati Qadimi Fumani, M. R. (2010) 'Inter-indexer consistency (IIC) in a Persian context.' *The Indexer* **28**(1), 12–17.
- Fenton, C. (2010) 'Use of controlled vocabulary and thesauri in UK online finding aids' *Journal of the Society Of Archivists* **31**(2), 187–205.
- Hearn, S. S. (2008) 'Comparing catalogues: currency and consistency of controlled headings.' *Library Resources and Technical Services* **53**(1), 25–40.
- Iwe, J. I. (2001) 'Linguistics and information processing: provision of syntactic and semantic consistency in the language of Library of Congress Subject Headings (LCSH) pertaining to literature and librarianship: a comparative analysis.' *Cataloging and Classification Quarterly* **32**(2), 107–26.
- Li, L. (2004) 'Consistency versus inconsistency: issues in Chinese cataloging in OCLC.' *Cataloging and Classification Quarterly* **38**(2), 17–31.
- Monreal, C. S. and Gil-Leiva, I. (2011) 'Evaluation of controlled vocabularies by inter-indexer consistency.' *Information Research* **16**(4), 8.
- Polger, M. A. (2011) 'Student preferences in library web site vocabulary.' *Library Philosophy and Practice*, 69–84.
- Portaluppi, F. (2007) 'Consistency and Accuracy of the Medical Subject Headings® thesaurus for electronic indexing and retrieval of chronobiologic references.' *Chronobiology International* **24**(6), 1213–29.
- Sievert, M. C. and Andrews, M. J. (1991) 'Indexing consistency in information science abstracts.' *Journal of the American Society for Information Science* **42**(1), 1–6.
- Svenonius, E. (1986) 'Unanswered questions in the design of controlled vocabularies.' *Journal of the American Society for Information Science* **37**(5), 331–40.
- Svenonius, E. (2003) 'Design of controlled vocabularies.' *Encyclopedia of Library and Information Science*, 822–38.
- Symonds, E. (2008) 'Controlling the vocabulary: maintaining consistency across collections.' *Visual Resources Association Bulletin* **35**(2), 106–11.

Since completing his Master's in library and information science at the University of Western Ontario in 2000, Mark Aaron Polger has had a variety of library-related jobs including cataloguing, web management and teaching. He has written widely on a number of topics of particular interest to him and is a frequent conference speaker (details available on request to him at MarkAaron.Polger@csi.cuny.edu).

ANZSI conferences: last time and next time

ANZSI takes stock of the 2013 conference and looks forward to the 2015 conference scheduled for 6–9 May in Canberra

ANZSI conference, 'Intrepid indexing without borders', Wellington, NZ, 13–15 March 2013

The conference attracted 71 attendees, most from Australia, and 13 workshop participants. The conference theme was carried through by many of our speakers. Our keynote speaker, Jan Wright, gave an inspiring address entitled 'Intrepid indexing: from the sea to the stars'.

An important part of this conference was looking at indexing for languages and cultures other than for English-speaking communities. Sessions covered Māori names and terms, East Asian (CJK) indexing, understanding Asian names, indexing practice in Japan, archiving and indexing history in the Pacific Islands, and database indexing for indigenous collections in Australia.

Panel discussions and roundtable discussions proved popular, with ideas and comments freely flowing! Topics covered were Māori names and terms, typesetting dilemmas; Asian names; publishers, editors and indexers; indexing techniques and EPUB; running an indexing business; future electronic indexing; Pacific Islands archives and indexing history; indexing numbers; military history indexing, and

reports from overseas societies about developments in Canada, the United States, China and the United Kingdom. Ebook publishing and the EPUB3 Standard and Charter were also covered.

The conference also included workshops on advice on tips and traps for indexers wishing to become accredited, advanced SkyIndex, and intermediate CINDEK.

Jan Wright held an InDesign workshop on Tuesday morning, 12 March, attended by 13 people including two non-members. The ICRIS meeting was held at the conference venue on Tuesday afternoon. The cocktail function on Wednesday evening was well attended and useful as a 'mixer' for attendees.

There were many requests pre-conference for a *Lord of the Rings* tour, and Hammonds Tours created a tour for ANZSI on Thursday, collected the 17-strong group near the conference venue, made sure they were provided with lunch, snacks and drinks, and showed them the city according to the Hobbits.

The conference dinner on Thursday was attended by 52 people (including partners). Diners were transported by charter coach to the art-deco style Roxy. On arrival, attendees were presented with pre-dinner drinks and

Copyright of Indexer is the property of Society of Indexers and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.