

---

## Ghettoes and diaspora in classification: Communicating across the limits

---

Hope A. Olson and Dennis B. Ward  
School of Library and Information Studies  
University of Alberta  
<hope.olson@ualberta.ca>  
<dennis.ward@ualberta.ca>

---

*The Dewey Decimal Classification (DDC) has limitations that reflect the biases of its cultural origins, resulting in problems representing marginalized knowledge domains such as women's studies and feminism. To locate the limits and make them permeable, we adopt theoretical conceptions of space from: Drucilla Cornell's philosophy of the limit, Lorraine Code's rhetorical space, Henri Lefebvre's transparent space, and Gillian Rose's paradoxical spaces. Our project operationalizes this theoretical approach in a relational database which provides term-by-term linkages from A Women's Thesaurus to DDC.*

### Introduction

Addressing the biased representation of marginalized knowledge domains in traditional classification, this pilot study explores a concrete technique for ameliorative change. The research problem is that the *Dewey Decimal Classification (DDC)* marginalizes or excludes groups and topics relating to women's issues, women's studies and feminism. *DDC*, like other subject representation standards, is inhospitable to groups and topics outside of our social and cultural mainstream. Criticism of *DDC*'s treatment of marginalized groups and topics dates from at least 1971. It has included theoretical critiques, general critiques regarding racism and sexism, and specific critiques relating to Native North Americans, Chicanos, people with disabilities, gay men and lesbians, and women, and critiques of *DDC*'s treatment of African, Melanesian, and Arabic materials, and Hebraica.

Since these critiques of *DDC* relate to specific discourses, recent theory on domain analysis offers an environment for this research. Instead of pretending to a "value-free ethics of dissemination of knowledge" in which we adopt a "prudent objectivity," Hanne Albrechtsen suggests that indexers and classificationists need to recognize that social and cultural realities call for subjectivity in taking

responsibility for information access for different groups (1993, 223). Domain analysis has been used successfully to map the features of two very different knowledge domains: fiction studies (Beghtol 1995) and mathematics (Iyer and Giguere 1995). Our proposed research follows from the premise that the landscapes of different knowledge domains need what David Horner calls a "metascientific and epistemological ground cleaning" to develop "methodological frameworks for a more cogent construction of classification and retrieval systems" (1993, 291). It questions the "metanarrative" of existing classification, by developing a "micronarrative" in one knowledge domain.

In a deconstruction of Melvil Dewey's work, *DDC*, and the application of *DDC* in catalogue records, Olson (1996) concludes that the biases apparent in *DDC* reflect and perpetuate Dewey's underlying presumption that the classification must and can be a universal language. Dewey asserted that *DDC* can accommodate all topics for all time at the same time that he recognized that this supposed universal language would not accommodate diversity. *DDC* fails as a universal language because it has limits. It marginalizes and excludes some knowledge domains, particularly domains marginalized or excluded by the society in which *DDC* exists. The worldwide use of *DDC* gives urgency to adapting it for use by diverse publics (see for example, Mitchell 1995). Developing a means of addressing *DDC*'s inherited biases has an enormous potential impact. This will be particularly so if our theoretical approach and its implementation can be used as a prototype for ameliorating biases in other marginalized knowledge domains.

### **Theoretical approach**

As a language with limits or boundaries rather than a universal language, *DDC* is a hierarchical construction that marginalizes some disadvantaged groups and topics by locating them in ghettos, diasporized across the system. Other disadvantaged groups and topics are totally excluded from the system — outside of its territorial limits. To open this bounded system we follow Drucilla Cornell's *The Philosophy of the Limit* which focuses on the constructed limits of the system, suggesting that they be made permeable to allow the voices of those excluded to be heard within. In relation to *DDC*, this means that instead of trying to perfect the system to overcome its biases, the system be made permeable.

Because classification is a locational system, we will use spatial analyses borrowed from various disciplines to identify and address its problems. The philosophical basis for these analyses comes from Lorraine Code's conception of rhetorical spaces as:

fictive but not fanciful or fixed locations, whose (tacit, rarely spoken) territorial imperatives structure and limit the kinds of utterances that can be voiced within them with a reasonable expectation of uptake and "choral support": an expectation of being heard, understood, taken seriously. They are the sites where the very possibility of an utterance counting as "true-or-false" or of a discussion yielding insight is made manifest (1995, ix).

In *DDC* there is rhetorical space for most mainstream social and scholarly discourses, but not always for marginalized discourses. This project breaches the boundaries or limits of closed rhetorical space by adopting Gillian Rose's technique of "paradoxical spaces" (1993). By occupying both poles of a multiplicity of one-dimensional spectra, Rose sees potential for overcoming the inside/outside either/or of boundaries. Being simultaneously or alternately in the centre and at the margin, same and other, inside and outside, women and other marginalized groups can develop a more complex and meaningful geography. Mapping a language from a marginalized knowledge domain, *A Women's Thesaurus*, to a language with pretensions to universality, *DDC*, creates a paradoxical space that occupies both poles of a spectrum. Creating this paradoxical space and a system to accommodate it is the mode of our research.

## Method

Our research questions are:

- To what extent can *DDC* represent women's studies/feminist topics?
- Is a spatial analysis applicable for mapping marginalized knowledge domains?
- Can an information technology be devised to accommodate the needs of marginalized knowledge domains in *DDC*?

This research is a pilot study to develop a specific application of our theoretical cartography. The application is a term-by-term linkage of a subset of *A Women's Thesaurus* to *DDC*. The methodology for developing guidelines for application of this theory is a reflexive one. "Classification Guidelines" operationalizing our theoretical approach are used by research assistants to derive classification numbers from *DDC* for each thesaurus term. It is not sufficient to find a classification number that fully and specifically represents a given term. The number must also be in a viable rhetorical space. Therefore, both coextensiveness and rhetorical space are considered as variables.

## 22 CAIS / ACSI 97

*A Women's Thesaurus* includes lists of its approximately 5,000 terms organized into eleven general subject areas. Four of the eleven will make up the sample for this pilot study:

Communication	224 terms
History and social change	
<i>a subset</i>	250 terms
International women	258 terms
Science and technology	185 terms

The research assistants and investigator making the links between these terms and *DDC* meet at intervals to critique and revise the "Classification Guidelines" and to consider problem terms. All linkages are reviewed on an ongoing basis and the guidelines further revised. Final analysis of this data will assess the linkages for patterns of coextensiveness and rhetorical space.

To successfully apply this approach, a software system is needed to support the construction of linkages and the provision of access to other researchers and eventually to information managers and patrons. A system is being developed in which the thesaurus and linkages to *DDC* are imported into a commercial thesaurus program that provides a convenient environment for editing and updating. The final product will then be exported to a relational database which can be searched using standard and inexpensive tools. Access to the database will be provided from central sites via the World Wide Web or through direct distribution of the database and ancillary programs. The software used in this project is envisioned as a prototype for a more general system which, like the theoretical approach, could be applied to other subject domains and linked to larger systems.

### **Classification guidelines**

The result of linking terms from *A Women's Thesaurus* to *DDC* is, in essence, a feminist index to *DDC*. It charts a feminist naming of subjects on a mainstream map of knowledge. Each link is assessed for:

- coextensiveness — how well the term in general can be linked to a number;
- gender coextensiveness — how well the term as it relates to women and/or feminism can be linked to a number;
- rhetorical space — whether or not that number is in an appropriate context.

Jessica Milstead defines coextensiveness as "the extent to which the index term reflects the precise content of the item of information . . ." (1984, 143), defining coextensiveness to take in the whole shape of the subject, not just its dimensions up or down a hierarchy as for specificity. Most classification constructs

pigeonholes which are preformed. Therefore, documents are put into the pigeonhole "closest in size to the subject" (1984, 144). Coextensiveness is achieved when the size and shape of the topic to be classified are the same as the size and shape of the classification number.

Gender coextensiveness as applied in this project is similar to the concept of race indicators used by Doris Hargrett Clack who studied the coextensiveness of *LCSH* for Black literature resources (1978). Gender may be indicated by the *DDC* base number itself (such as 331.4 for women workers) or it may be appended to the base number. The most common gender indicator to be appended is 082 for women, females and feminist perspectives from *DDC*'s Table 1: Standard Subdivisions (T1-082), because the standard subdivisions may be used relatively freely throughout *DDC*.

Five categories of coextensiveness have been defined for this study: match, limited match, broad match, partial match, and no match. They incorporate general coextensiveness and gender coextensiveness.

**match** — The thesaurus entry and the classification number can be readily interpreted as representing the same topic and the number includes a gender indicator or is hospitable to a gender indicator. No match is expected to be exact due to the subjectivity of understanding topics.

For example, *balance of payments* is coextensive with 382.17 and T1-082 could be appended to indicate the relation of balance of payments issues to women or a feminist perspective on balance of payments.

**limited** — The thesaurus entry and the classification number can be readily interpreted as representing the same topic, but the number *does not include a gender indicator and a gender indicator cannot be appended*. Typically, the number for the topic will already include a standard subdivision so T1-082 cannot be appended, because only one standard subdivision may be used at a time.

For example, *agricultural extension* is reasonably coextensive with 630.715, but because "0715" is a standard subdivision, T1-082 cannot be appended to this number.

**broad** — The classification number is broader than the thesaurus entry, but does include the topic. Because of the *DDC* practice of "approximation of the whole," standard subdivisions may not be appended to numbers broader than the topic under consideration. Therefore, *DDC* numbers that are broad will typically also be limited and both designations should be added. Some broad numbers will, however, indicate gender.

For example, *African studies* is classified in 960 along with general history and civilization of Africa. Not only is the number far broader than

the topic, but because *African studies* does not approximate the whole, standard subdivisions, including T1–082, cannot be appended for this topic. However, specific feminisms such as *liberal feminism* and *radical feminism* are classified under the general number for feminism, 305.42, which is broad, but does indicate gender.

**partial** — The classification number represents only one aspect or perspective on the topic represented by the thesaurus entry. Numbers that are partial will typically allow the addition of a gender indicator.

For example, *appropriate technology* is represented by 338.927, but that classification number is located in the hierarchical context of "Economics. Production." The scope note for *appropriate technology* reads:

*Simple, low-cost techniques and tools that help increase productivity in agriculture and household tasks. (A Women's Thesaurus, 26)*

The number under "Production" does not include the actual "techniques and tools" which would be located somewhere in the 600s and it does not indicate the context of "agriculture and household tasks." However, it is a specific number for this topic and does reflect the aspect of "productivity."

**no match** — A classification number that represents the topic or a significant aspect of it cannot be located by the researchers.

For example, there is no *DDC* number for *basic human needs*.

Coextensiveness, then, relates to how well the *DDC* number represents the topic of the thesaurus term. Rhetorical space, on the other hand, assesses the friendliness of the *DDC* number's context. This project uses five categories of rhetorical space: open, closed, transparent, ghettoized and diasporized.

**open** — The classification number provides a rhetorical space in which the topic represented by the thesaurus entry is meaningful, can be taken seriously and can be discussed openly.

For example, *appropriate technology* in 338.927 is only partially coextensive, but its rhetorical space is not inhibiting or judgmental. Its hierarchical context gives it credibility:

- 300 Social sciences
- 330 Economics
- 338 Production
- 338.9 Economic development and growth
- 338.92 Specific policies
- > 338.927 Appropriate technology

And the topics between which it sits on the shelves are similar types of issues:

- 338.922 Subsidies and grants
- 338.924 Nationalization
- 338.925 Privatization
- 338.926 Information policy
- 338.927 Appropriate technology
- 338.93-338.99 Economic development and growth in specific continents, countries, localities

**closed** — The classification number does not offer a hospitable environment to the topic represented by the thesaurus entry.

For example, *concubinage* in 392.6 is in the hierarchical context of customs, etiquette and folklore with no link to sex roles or exploitation:

- 300 Social sciences
- 390 Customs, etiquette, folklore
- 392 Customs of life cycle and domestic life
- > 392.6 Customs of the relations between sexes

And its entry trivializes this topic by grouping it with "chaperonage" and topics from the index such as "dating."

Class Number: 392.6

Caption: Customs of the relations between sexes

Notes: Including chaperonage

DDC Index Terms:

- Celibacy—customs
- Chaperonage—customs
- Cohabiting—customs
- Concubinage—customs
- Dating (Social practice)—customs
- Premarital sexual relations—customs
- Sex—customs

**transparent** — "Transparent space" according to Henri Lefebvre (1991) denies the existence of anything excluded from its mapping because it appears to be all there is. It is the illusion that location is neutral — that mapping of territory can be a true representation of some essence of reality. Making this space visible means identifying its boundaries to allow recognition of what is outside those boundaries.

For example, *colonialism* is classified in 325.3, "Colonization." This number views colonies from the perspective of the colonizers rather than the colonized. Its context confirms the focus on the colonizers, the

population that moves, as opposed to the colonized, the previously resident population:

325 International migration and colonization

325.1 Immigration

325.2 Emigration

325.3 Colonization

**ghettoized** — Ghettoization consists of isolating marginalized groups by concentrating them in one area. The ghetto is formed by a forced concentration of people sharing a particular characteristic. Jewish ghettos, African American ghettos, Latino ghettos are all familiar images. "Women's" occupations are sometimes referred to as ghettoized. Numerous forces create and maintain ghettos: military, social, political, economic, psychological. They become institutionalized in various forms — even libraries' organizations of information. An aspect of ghettoization is confinement. Gillian Rose draws together evidence of women's spatial confinement in the private/domestic (as opposed to the public) sphere and in the gaze of the subject regarding the object. She suggests that the objectification of women by a masculine subject confines women by making us conspicuous, just as visibility enables racism (1993, 144-149). Visibility, then, enforces ghettoization — like the wearing of a yellow star. The visibility of marginalized groups makes space an important concept to consider in our representation. Members of the mainstream are allowed to live in transparent space — without the barriers of the ghetto (Rose 1993, 147).

For example, *African studies* in 960 and *Asian studies* in 950 are ghettoized by the technique of making them historically and geographically defined. Such area studies deal with social, economic, political, cultural and other aspects of a group of people. By classifying them in the 900s with history of a geographical area, they are taken out of the mainstream of everyday life in the world at large.

Another example is the classification of women workers in specific industries. Workers in specific occupations are classified in 331.7 "Labor by industry and occupation." However, women workers have a separate section in 331.4 which is then subdivided by industry in 331.48 "Women workers by industry and occupation." The result is that women workers are taken out of the mainstream and located in a feminized ghetto. There is no parallel number for men workers.

**diasporized** — Diasporas disperse marginalized groups so that they have no concentrated existence, risking assimilation and loss of identity or creation of a different identity. Diasporization is the opposite pole from ghettoization,



implying the availability of only two options of which only one must be chosen.

For example, material on *Gays* in general is classified in "Occupational and miscellaneous groups" at 305.90664. However, material only on *Lesbians* is classified under women first with additional notation for sexuality: 305.489664. And material only on *Gay men* is classified under men first with additional notation for sexuality: 305.389664. The general number is separated from the specific numbers by all groups identified according to class, religion, language and race. Works on lesbians that specify other characteristics such as race or ethnicity will be even further dispersed resulting in the exclusion of sexuality for purposes of browsing. The result is that people browsing for information on lesbians and gays will find it dispersed — if they find it. Lesbians and gays are diasporized throughout this section of *DDC*.

Assessing the linkages between thesaurus terms and *DDC* numbers is intended to give a conceptual understanding of both acute and systemic problems so that in a later stage of this research we can develop global and local techniques for change.

### **Software to support development and access**

A software system is being developed to support this project. Its immediate purpose is to provide a development environment and an access method for researchers. In the longer term an interface for information managers and users is envisioned, as well as generalization into a toolkit that could be used in other domains.

The major requirements for the software system are:

- a convenient means for recording linkage information, ideally allowing students and other research personnel to work at a convenient location (e.g. at home);
- a thesaurus maintenance system which will allow thesaurus and linkage information to be imported, displayed, edited and exported in a variety of formats;
- a means for providing access to researchers and other interested parties.

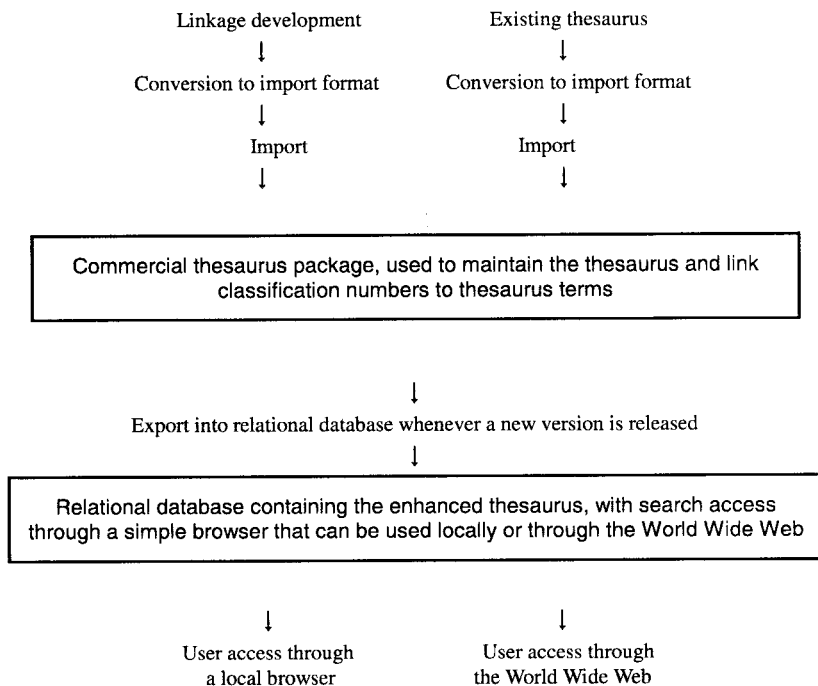
It is also vital that the components of the system be widely available, inexpensive and simple to use. This is particularly important for the access tools, since we hope that this work will find an application in the real world.

Figure 1 summarizes how the system will function. A commercial thesaurus maintenance program (MultiTes) provides the capabilities identified in the second requirement above. A digital version of *A Women's Thesaurus* has been successfully imported into MultiTes, and word processors and spreadsheets are being used to prepare tables defining the linkages between thesaurus terms and

*DDC*. These tables will be written out as comma delimited files, and then converted into a term list which can be imported into MultiTes. The linkages are stored as a special kind of "narrower term", comprising the Dewey number plus notes and the other information that has been described in the previous section.

A relational database management system will be used to provide access to the thesaurus and linkages. This approach has been investigated in the past (Jones, 1993), but usually with large systems like Oracle. We plan to use Microsoft Access and Visual Basic, so that the database can be used on standard PC's. Initial tests have been promising, *A Women's Thesaurus* having been easily exported into a database both reasonable in size (several megabytes) and providing quick response on a low-end Pentium machine. The database search and World Wide Web interfaces are currently being developed. We plan to implement the Web interface using the tools that are provided with the Microsoft Internet Information Server under Windows NT 4.0.

**Figure 1 - Software System for Development and Distribution of Results**



### Preliminary results of sample

Preliminary results suggest methodological considerations, the possibility of quick amelioration of acute problems, and serious systemic concerns with *DDC* that will demand major and/or very creative solutions.

A methodological concern in assessing the efficacy of this approach is that statistical analysis will be of questionable validity, because of the different relations of terms to the domain. For example, several terms for which there is no match are core concepts in women's/feminist studies. *Objectification*, "the process of viewing women as sex objects that exist only in relation to men and male desires," is not classifiable in *DDC*. Neither is *private sphere*, the "assumption that society is divided into male and female spheres, women's sphere private and domestic, men's sphere public"; or *race, class, and gender studies*, "interdisciplinary area of study that attempts to integrate understanding of the complex current and historical forces affecting economic, political, and social status of women and men." A *Women's Thesaurus* includes many terms that could relate to women and women's issues, or could be viewed from a feminist perspective, but are not core, such as *copyright, lobbying, nonviolence, and information*. That these concepts can be represented in *DDC* is of less importance in this context than that *objectification* and *private sphere* cannot, so statistical analysis would only be useful if terms could be weighted for their significance to the domain — a task of unworkable subjectivity, particularly in an interdisciplinary, quickly-changing knowledge domain.

A second methodological concern is that, as anticipated, the categories of coextensiveness and rhetorical space have the same problems as any classification: "Despite our desperate, eternal attempt to separate, contain, and mend, categories always leak" (Trinh 1989, 94). At present separate categories for gender indicators are evolving and the coextensiveness category *limited* is becoming redundant as a result. Also, several linkages fall into more than one category, but that is probably not a problem.

Some patterns are beginning to emerge that could be remedied by minimal changes to *DDC*. An example is the classification of occupations described above. An option to subdivide workers in a given occupation by gender would allow works on women in an occupation to be classified with that occupation.

Another optimistic indication is that there are relatively few headings with no match or with closed rhetorical space and that the ones that do exist represent acute problems that can probably be addressed through a reasonable number of specific changes to the classification.

Of more concern is the high level of diasporization. It seems to reflect two fundamental tenets of *DDC*: that gathering is by discipline rather than topic (for example, cats may be classified as animals in 599.752, as pets in 636.8, and as

images in each of art and literature separately, using a subdivision from T3C); and that topics are always subdivided in an order of preference (for example, first by age, then by sex, then by race, then by language, then by . . .). Classification by discipline necessarily diasporizes interdisciplinary topics. Preferred order always privileges some differences over others and those least privileged are most diasporized (see example of lesbians and gays above).

Another serious concern is that while the ghettoization and diasporization of topics is fairly easy to identify, transparent space is both elusive and systemic. Because *DDC*, like other classifications, reflects its society of origin, it can appear neutral when it is actually inherently biased. Transparent rhetorical space often coincides with partial coextensiveness. That is, entire aspects of a topic may be unrepresentable while the classification wears a cloak of objectivity and completeness — a mantle of invisibility in relation to bias. For example, *DDC* numbers for *oppression* highlight the victims, not the perpetrators. The ubiquity and subtlety of transparent space also has the methodological ramification that it will take several revisions of the links to see through this apparent neutrality and make the biases of rhetorical space visible.

### Tentative conclusion

Thus far our approach is producing helpful results and shows potential for a better understanding of *DDC*'s representation of women's/feminist studies topics and techniques for its improvement. We are addressing a knowledge domain that is interdisciplinary and dynamic as well as marginalized so final results are unpredictable and, in fact, unlikely to ever be really final. However, we are optimistic that we can create a paradoxical space and a system to accommodate it where marginalized knowledge domains can simultaneously occupy the margins and the centre.

### References

- Albrechtsen, Hanne. 1993. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer* 18(4): 219-224.
- Beghtol, Clare. 1995. Domain analysis, literary warrant, and consensus: The case of fiction studies. *Journal of the American Society for Information Science* 46(1): 30-44.
- Capek, Mary Ellen S., ed. 1987. *A Women's Thesaurus*. New York: Harper & Row.
- Clack, Doris Hargrett. 1978. The adequacy of Library of Congress Subject Headings for Black literature resources. *Library Resources and Technical Services*, 22: 137-44.
- Code, Lorraine. 1995. *Rhetorical spaces: Essays on gendered locations*. New York: Routledge.
- Cornell, Drucilla. 1992. *The philosophy of the limit*. New York: Routledge.
- Dewey *Decimal Classification and Relative Index*. 1996. Devised by Melvil Dewey; ed., Joan S. Mitchell. Edition 21. Albany, NY: Forest Press.

- Horner, David Sanford. 1993. Paradigms, discourses and language games: Categorical frameworks and signs of the times. In *ASIS '93: Proceedings of the 56th annual meeting*. Medford, NJ: Learned Information for the American Society for Information Science.
- Iyer, Hemalata, and Mark D. Giguere. 1995. Towards designing an expert system to map mathematics classificatory structures. *Knowledge Organization* 22: 141-147.
- Jones, Susan. 1993. A thesaurus data model for an intelligent retrieval system.. *Journal of Information Science* 19: 167-178.
- Lefebvre, Henri. 1991. *The production of space*. Donald Nicholson-Smith, trans. Oxford: Basil Blackwell.
- Mitchell, Joan. 1995. Options in the Dewey Decimal Classification system: The current perspective. *Cataloging & Classification Quarterly*: 19(3/4), 89-103.
- Olson, Hope A. 1996. *The power to name: Marginalizations and exclusions of subject representation in library catalogues*. Ph.D. diss., University of Wisconsin-Madison.
- Rose, Gillian. 1993. *Feminism & geography: The limits of geographical knowledge*. Minneapolis: University of Minnesota Press.
- Trinh T. Minh-Ha. 1989. *Woman, native, other: Writing postcoloniality and feminism*. Bloomington: Indiana University Press.