Multiple Entry Points
Variants and Cross-References in Indexes and Thesauri

American Society for Indexing Annual Conference
Chicago, Illinois, June 17, 2016

Heather Hedden, Senior Vocabulary Editor
John Magee, Director,
Indexing and Vocabulary Services
Gale | Cengage Learning
About Heather and John

- Each has worked at Information Access Company/Gale Research/Thomson Learning/Cengage Learning since the 1990s (Heather with a break of 10 years)
- Each has experience with back-of-the-book indexing, periodical database indexing, and thesaurus development for database indexing
Outline

Introduction
- Exercise
- Book indexing vs. periodical/database indexing
- Thesauri
- Multiple entry points overview

Back-of-the-book index multiple entry points
- Double posts
- See references
- Issues specific to book indexes

Thesaurus multiple entry points
- Nonpreferred terms
- Issues specific to thesauri

Comparison summary
Outline

Introduction
- Exercise
- Book indexing vs. periodical/database indexing
- Thesauri
- Multiple entry points overview

Back-of-the-book index multiple entry points
- Double posts
- See references
- Issues specific to book indexes

Thesaurus multiple entry points
- Nonpreferred terms
- Issues specific to thesauri

Comparison summary
Exercise
Propose multiple entry points for these entries:

- Campaign finance
- Automobile industry
- Aviation
- French Navy
Three related functional/skill areas:

   - Identifying the concepts and names mentioned in the book and organizing them into an index

2. Periodical/database indexing
   - Identifying the main ideas of an article or content item and assigning the most appropriate index terms available from a controlled vocabulary, which results in creating an index

3. Controlled vocabulary (thesaurus) creation
   - Creating and editing a structured list of terms used for database indexing (and sometimes multi-volume book indexing) for supporting end-user retrieval
Back-of-the-book indexing vs. periodical/database indexing


- Also called “closed indexing”: the index is created for the single work, then is done (closed)
- Indexing subsequent editions may involve referring to previous edition’s index, but usually are indexed from scratch again
- Embedded indexing (linking to text location in the electronic file) enables index reuse and revision in subsequent editions
2. Database indexing

- Also called “open indexing”: indexing is an ongoing process as additional periodical issues or content is added, and the index is used yet never “finished” (open)

- A controlled vocabulary is necessary to provide consistent indexing to the same concepts from different sources indexed by different indexers over time.

- Originally was mostly for periodical articles. Now for any content in a content management system of digital asset management system: HTML files, PDFs, PPTs, brochures and ads, test questions and learning activities, images, audio, video, etc.
1. Similarities between the two kinds of indexing

- Read/examine and analyze content for what the main concepts are
- Consider different ways the concepts might be named
- Consider to how much detail to index

2. Differences between the two kinds of indexing

A. Tasks

- Back-of-the-book indexing requires the indexer to additionally come up with (invent) all of the index terms and their variants and arrange them into an index
- Database indexers utilize the existing controlled vocabulary (and may suggest terms subject to approval)
2. Differences between the two kinds of indexing (continued)

B. Differences in the resulting indexes

- Back-of-the-book indexing results in a fully displayed browsable alphabetical index.
- Database indexes may or may not be displayed to end-users. Maybe just portions (such as terms in a type-ahead scrollbox)

C. Differences in the indexers

- A book is indexed by a single indexer.
- Database indexing projects are shared by multiple indexers.
Controlled vocabulary

- Each term stands for an unambiguous concept.
- There is control over the addition of terms to the vocabulary.
- Indexers should not add terms already in the controlled vocabulary.
- A very small controlled vocabulary might be a simple list without multiple entry points.

Thesaurus - a kind of controlled vocabulary that has multiple entry points and structure

- **Multiple entry points** are “equivalent” terms, with a nonpreferred term pointing to a preferred term.
- **Structure** is through relationships between terms
  - Hierarchical (broader term/narrower term: BT/NT)
  - Associative (related terms: RT)
Thesaurus excerpt example

Alphabetical browse:

- Corporate trust services (Subjects)
- Corporate turnarounds (Subjects) (NPT)
- Corporate videos (Subjects) (NPT)
- Corporate welfare (Subjects)
- Corporate wellness programs (Subjects) (NPT)
- Corporation directors (Subjects) (NPT)
- Corporation executives (Subjects) (NPT)
- Corporation law (Subjects)
- Corporation reports (Subjects) (NPT)
- Corporation secretaries (Subjects)
- Corporations (Subjects)
- Corporatism (Subjects) (NPT)
- Corporative state (Subjects) (NPT)
- Corporativism (Subjects) (NPT)

Selected term details:

Descriptor Corporation law

Relationships

- UF Company law (Subjects)
- UF Corporate law (Subjects)
- BT Commercial law (Subjects)
- NT Antitrust law (Subjects)
- NT Business judgment rule (Subjects)
- NT Disregarding corporate entity (Subjects)
- NT Incorporation (Subjects)
- NT Railroad law (Subjects)
- RT Articles of incorporation (Subjects)
- RT Business enterprises (Subjects)
- RT Business trusts (Law) (Subjects)
- RT Bylaws (Subjects)
- RT Corporate counsel (Subjects)
- RT Corporate domicile (Subjects)
Introduction: Thesauri

Thesaurus excerpt example

Hierarchical view excerpt
Multiple Entry Points

- **Defined:** Synonyms or roughly equivalent concepts (not just words), for the context.

- **Purpose:** To capture different wordings of how different people might describe or look up the same concept or idea.
  - Differences between that of the author and the user/reader
  - Differences among different users/readers

- A concept may any number of (multiple) entry points, or it may have only a single entry name.

- Multiple entry points can point to the preferred entry/term, or they can point directly to the content.
### Multiple Entry Points Overview

<table>
<thead>
<tr>
<th>Have various designations:</th>
<th>Found in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aliases</td>
<td>SKOS vocabularies</td>
</tr>
<tr>
<td>Alternate labels</td>
<td></td>
</tr>
<tr>
<td>Alternate terms</td>
<td></td>
</tr>
<tr>
<td>Cross-references</td>
<td>BOB indexes</td>
</tr>
<tr>
<td>Entry terms</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Equivalency relationship</td>
<td>Thesaurus creation</td>
</tr>
<tr>
<td>Equivalent terms</td>
<td>Thesaurus creation</td>
</tr>
<tr>
<td>Keywords</td>
<td></td>
</tr>
<tr>
<td>Nondescriptors</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Non-postable terms</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Nonpreferred terms</td>
<td>Thesaurus creation</td>
</tr>
<tr>
<td>NPTs</td>
<td>Thesaurus creation</td>
</tr>
<tr>
<td>See references</td>
<td>BOB indexes</td>
</tr>
<tr>
<td>Synonyms</td>
<td>Taxonomies</td>
</tr>
<tr>
<td>Use for terms</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Use references</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Used for terms</td>
<td>Thesaurus end-use</td>
</tr>
<tr>
<td>Variants</td>
<td></td>
</tr>
</tbody>
</table>
Types include:
- synonyms
- near-synonyms
- variant spellings
- lexical variants
- foreign language names
- acronyms/spelled out forms
- common/scientific names
- phrase variations (in print)
- antonyms
- more specific concepts (depending on circumstances)
Exercise: Create each kind of multiple entry point

- synonyms: Cars
- near-synonyms: Junior high schools
- variant spellings: Defense
- lexical variants: Hair loss
- foreign language names: German Air Force
- acronyms/spelled out forms: United Nations
- common/scientific names: Cancer
- phrase variations (in print): School buses
- antonyms: Behavior
- more specific concepts: Computers
Exercise answers:
- synonyms: Cars / Automobiles
- near-synonyms: Junior high schools / Middle schools
- variant spellings: Defense / Defence
- lexical variants: Hair loss / Baldness
- foreign language names: German Air Force / Luftwaffe
- acronyms/spelled out forms: United Nations / UN
- common/scientific names: Cancer / Neoplasms
- phrase variations (in print): School buses / Buses, school
- antonyms: Behavior / Misbehavior
- more specific concepts: Computers / Laptops
Outline

Introduction
- Exercise
- Book indexing vs. periodical/database indexing
- Thesauri
- Multiple entry points overview

Back-of-the-book index multiple entry points
- Double posts
- See references
- Issues specific to book indexes

Thesaurus multiple entry points
- Nonpreferred terms
- Issues specific to thesauri

Comparison summary
Double posts

- Multiple entries that refer to the same concept/name/topic/idea with the same locators.
- Desirable for many entries, but not all.
- Although called “double” posts, can be for three or more.
- Use double posts instead of See reference, for entries with no subentries.
Double post examples:

Film reviews, 162–166, 173
Movie reviews, 162–166, 173

Ethics of communication, 113–114
Communication ethics, 113–114

Search, web, 41–42
Web-based research, 41–42

GDP (gross domestic product) 233, 251, 265, 295
gross domestic product 233, 251, 265, 295
See references

- Entries that point to another entry, to use instead
- Locators are at the referred entry only.
- Used instead of double posts when entries have subentries, and it is undesirable to repeat all subentries. So it saves space.
See reference examples:

arms purchases. See weapons purchases
labor unions. See unions, labor
parties, political. See political parties
war on drugs. See drug war
Handicapped persons. see Persons with disabilities
Daoism. see Taoism
MAS. See Movement toward Socialism (MAS)
Movimiento Nacionalista Revolucionario (MNR). See Nationalist Revolutionary Movement (MNR)
Issues specific to book index multiple entry points:

**Alphabetical display**

Multiple entry points next to each other or very near (a couple entries away) are not needed, and are better omitted.

Do not have:

- Ethnic groups
- Ethnic minorities

Close lying entries may be permitted for educational purposes:

- Net profit. See Net operating profit
“Flipping” main entries and subentries can serve as double posts

APA documentation style, 649, 665–677
parenthetical references, 666

Parenthetical references
in APA documentation style, 666

African Americans
business ownership, 247–248

Business ownership
by African Americans, 247–248

Active voice, 213–216
tone and, 235

Tone
active voice and, 235
More common than full flipping:
Double-post of main entry and main entry/subentry

Design
  clothing. see Fashion design
  furniture, 223
  industrial, 60–61
  interior. see Interiors
  product, 224–226

Furniture design, 223
Double post variation:
Slightly broader concept with added locator(s)

Body language, 85–87, 118

Gestures, 85–87

Hand gestures, 85–87
Outline

Introduction
- Exercise
- Book indexing vs. periodical/database indexing
- Thesauri
- Multiple entry points overview

Back-of-the-book index multiple entry points
- Double posts
- See references
- Issues specific to book indexes

Thesaurus multiple entry points
- Nonpreferred terms
- Issues specific to thesauri

Comparison summary
Thesauri, by themselves, do not have locators.

- So, true (re-directing) cross-references (in addition to double posts) are not needed for the purposes of saving space.

Thesauri are rarely in print or print-only.

- So, multiple entry points can link directly to content.
- For the end-user, there is no need to re-direct/point to another term in the thesaurus first with a cross-reference.

Thesauri are structured hierarchically.

- So, a single “preferred” term is needed for each concept in order to display a simple hierarchy of main entry terms. Double post terms cannot be accommodated.
Thesauri generally do not have traditional cross-references that redirect within the thesaurus prior to pointing/linking to the content.

Thesauri do not have double post versions of concepts, where multiple entries are on equal standing.

Each concept is described by a single “preferred term” which has hierarchical and related relationships with other preferred terms.

Additional entry points for the concepts are “nonpreferred terms” (NPTs).

Nonpreferred terms (NPTs) point to the preferred term, and have no relationships to other preferred terms.
Standard thesaurus notation: USE / UF (Used for or Used from)

Nonpreferred term USE Preferred term
Preferred term UF Nonpreferred term

- Public procurement USE Government purchasing
  Government purchasing UF Public procurement

- Dress design USE Costume design
  Costume design UF Dress design

- Eskimos USE Inuit
  Inuit UF Eskimos

- Inundations USE Floods
  Floods UF Inundations
Thesaurus nonpreferred terms:

- Considered a kind of “relationship” of the Equivalency type.
- Reciprocity of relationship, pointing in both directions.
- Both preferred terms and nonpreferred terms are “terms.”

In contrast with SKOS vocabularies:

- Instead of terms, there are concepts.
- Concepts have multiple labels.
- Concepts have a preferred label (for each language).
- Concepts have any number of alternate labels and hidden labels (for each language).
- Alternate and hidden labels are part of a concept’s attributes, not equivalent terms and not connected by “relationships.”
Many-to-one, not one-to-many

- Typically a nonpreferred term may point to only one preferred term.
- But a preferred term, may have multiple nonpreferred terms pointing to it.
Thesauri, typically in electronic form, can be searched in addition to being browsed.

- For users who search, nonpreferred terms for entries that are alphabetically next to each other should still be. For example, *do* have both:
  - Ethnic groups
  - Ethnic minorities

- If the search system is “smart” to search on words within terms, *do not* include simple inversions as nonpreferred terms. For example, *do not* add the nonpreferred term:
  - Political fundraising
  - UF Fundraising, political
Narrower concept as a nonpreferred term

Examples: Laptops USE Computers

Hand gestures USE Body language

As the preferred term is used for the narrower concept and fully encompasses the nonpreferred term.

Can be problematic if:

- The nonpreferred/preferred term relationship is not displayed to the end-users, and
- There are multiple narrower concepts as nonpreferred terms

Computers
  NT Laptops
  NT Desktops
  NT Servers
  NT Supercomputers
Narrower concept as a nonpreferred term problem

Scenario:

- Indexer indexes article on *Supercomputers* with *Computers*.
- End-user looks up term *Laptops*, and is taken directly to result set of articles indexed with *Computers*.
- Result set includes articles on supercomputers and other computers that are not laptops, in addition to articles on laptops.
- End-user thinks the indexing is wrong by retrieving articles on other computers besides the selected laptops.
Know your user interface for thesauri

- Thesauri don’t stand alone like a book index. They are used within systems and applications.

- Decisions about additional entry points for a thesaurus are best made when you have a full understanding of your end users and the user interface through which they will use the thesaurus.

- Always ask yourself, “How are these NPTs being used?” The answer to that question will be the answer to many, many questions about specific NPTs.

Example on the following screenshot slides:

Education standards USE/SEE Educational standards
User interface of the thesaurus editor – Begins search:

**Search Form**
- **Either enter a search phrase**
  - education standards
- **Smart**
  - Begins
  - Contains
  - Exact
- **or select an alphabetical range**
  - Range:
    - a
    - thru:
    - z
  
- **or enter a specific item uid**
  - Item UID:

**Select search criteria**
- **Obj**: Subjects
- **Cat**: All Categories
- **Act**: Active

**Display batching**
- 1000 Items per batch

**Search Results**
- Elapsed Time for Query: 0.019 seconds
- 1 Items Found

- Education standards (Subjects) (NPT)
User interface of the thesaurus editor – Smart search:

**Search Form**
- Enter a search phrase
- Education standards
- Smart box
- Begins, Contains, Exact
- Select an alphabetical range
  - Range: a through Z
- Enter a specific item UID
- Item UID
- Select search criteria
  - Obj: Subjects
  - Cat: All Categories
  - Act: Active
- Display batching
  - 1000 Items per batch
- Start Search

**Search Results**
- Elapsed Time for Query: 0.068 seconds
- 5 Items Found
- Education standards (Subjects) (NPT)
- Educational standards (Subjects)
- State education standards (Subjects) (NPT)
- State educational standards (Subjects) (NPT)
- State standards (Education) (Subjects)
User interface of the indexer – Alphabetical browse:
User interface of the indexer – Smart search:

Thesaurus Multiple Entry Points: Specific issues – user interface
User interface of the end-user - search on Subjects:

- Education standards
  - See Educational standards
- Australia. Tertiary Education Quality and Standards Agency Act 2011
User interface of the end-user – autosuggest enabled:

- Subject Guide Search
- Publication Search

Search for: education stan

- Education Statutes and Regulations of Ontario 1998 Consolidation (Nonfiction work)
- Educational standards
- United States. National Center for Education Statistics
Introduction

- Exercise
- Book indexing vs. periodical/database indexing
- Thesauri
- Multiple entry points overview

Back-of-the-book index multiple entry points

- Double posts
- See references
- Issues specific to book indexes

Thesaurus multiple entry points

- Nonpreferred terms
- Issues specific to thesauri

Comparison summary
Multiple Points of Entry Comparison: Similarities

Same goal:
- To direct various users, who / various terms that mean the same thing, to the same content location

Same approach:
- Utilizes synonyms, near synonyms, sometimes antonyms (e.g. behavior/misbehavior), slang or jargon, abbreviations or acronyms and spelled out forms, former and current names, pseudonyms, phrase variations and inversions, etc.
## Multiple Points of Entry Comparison: Differences

<table>
<thead>
<tr>
<th><strong>Book Indexes</strong></th>
<th><strong>Thesauri</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Two different methods:</strong></td>
<td>One method only:</td>
</tr>
<tr>
<td>1. Double-posts</td>
<td>Nonpreferred terms / Equivalency relationship: Use</td>
</tr>
<tr>
<td>Both or all of equivalent-meaning entry terms have equal standing</td>
<td>- Points the user from an entry term <em>not</em> used in the thesaurus to one that <em>is</em> used in the thesaurus</td>
</tr>
<tr>
<td>2. <em>See</em> references</td>
<td>Thesaurus editor decisions:</td>
</tr>
<tr>
<td>- Points the user from an entry term <em>not</em> used in the index to one that <em>is</em> used in the index</td>
<td>- In all cases, what the preferred term will be</td>
</tr>
<tr>
<td><strong>Indexer decisions:</strong></td>
<td>Equivalency relationships are bi-directional and reciprocal: / and Used from (USE/UF)</td>
</tr>
<tr>
<td>- When to create double-posts versus <em>See</em> references (usually based on presence of subentries)</td>
<td></td>
</tr>
<tr>
<td>- If using a <em>See</em> reference, then what the preferred term will be</td>
<td></td>
</tr>
<tr>
<td><strong>See reference are one-directional:</strong> <em>See</em> (no corresponding “Seen from”)</td>
<td></td>
</tr>
</tbody>
</table>
Multiple Points of Entry Comparison: Differences

**Book Index**
The user will skimming the printed index.

Don’t create cross-references that fall close to each other alphabetically (starting with the same word or with the same first 3-4 letters).

*Do not create:*

- biological sciences. *See biology*

*Create adjective-noun inversions, as double-posts or cross-references to provide a different word to start on:*

- business zoning
  zoning, business

**Thesaurus**
The user might *search* the thesaurus instead of browsing it.

*Do create nonpreferred terms that would fall close to each other (starting with the same word or with the same first 3-4 letters).*

*Do create:*

- Biological sciences
  *Use Biology*

*If the thesaurus can be searched, do *not* create inverted nonpreferred terms.*

/ natural language only.

- Business zoning
## Multiple Points of Entry Comparison: Differences

<table>
<thead>
<tr>
<th><strong>Book Index</strong></th>
<th><strong>Thesaurus</strong></th>
</tr>
</thead>
</table>
| Create multiple entry points in anticipation of different ways:  
- various users may look up the concept. | Create multiple entry points in anticipation of different ways:  
- various users may look up the concept,  
- various indexers may look up the concept,  
- and various sources to be indexed may describe the concept |

A well-designed thesaurus generally has more nonpreferred terms proportionally, than a well-designed BOB index has of double posts of main entries and See references.
Book Index

With double (or triple) posts:

- computers in typography, 99–100, 145–146, 181
- digital typography, 99–100, 145–146, 181
- typography, digital, 99–100, 145–146, 181

Thesaurus

- Computers in typography
  USE Digital typography

- Digital typography
  UF Computers in typography
### Book Index

With *See* references:

AIGA. *see* American Institute of Graphic Arts

American Institute of Graphic Arts
- awards, 6, 55–56, 63, 96, 100
- founding of, 38
- Nash, Ray, involvement in, 96
- publications, 56
- SP meetings with, 8

### Thesaurus

- *AIGA*
  - USE American Institute of Graphic Arts
- American Institute of Graphic Arts
  - UF AIGA
Further information

Book indexing


Thesaurus creation

- “Practical Taxonomy Creation” ASI Online Learning 3-part webinar course [http://www.asindexing.org/online-learning/taxonomy-hedden](http://www.asindexing.org/online-learning/taxonomy-hedden)
- Construction of Controlled Vocabularies: A Primer [http://marciazeng.slis.kent.edu/Z3919/index.htm](http://marciazeng.slis.kent.edu/Z3919/index.htm)
- Thesaurus Construction tutorial by Tim Craven [http://publish.uwo.ca/~craven/677/thesaur/main00.htm](http://publish.uwo.ca/~craven/677/thesaur/main00.htm)
Questions?

Heather Hedden  
Senior Vocabulary Editor  
Cengage Learning  
20 Channel Center St., Boston, MA 02210  
(o) 617-757-8211 | (m) 978-467-5195  
Heather.Hedden@cengage.com | Heather@Hedden.net

John Magee  
Director, Indexing & Vocabulary Services  
Cengage Learning  
27500 Drake Rd.  
Farmington Hills, MI 48331  
248-699-8091  
john.magee@cengage.com

www.cengage.com