Thesauri

1 Thesauri: Introduction

A thesaurus is particular kind of **controlled vocabulary**

- **Controlled vocabulary:** anytime an input system reduces you to a fixed list of options
- There could be many reasons to limit the number of choices people have—usually it's about reducing synonymy, and reducing cognitive load for people (i.e., it's easier to browse a short list of options rather than think of the correct term)
- Example of controlled vocabulary: choose language from a dropdown list

Whereas a controlled vocabulary is usually just a flat list, in a **thesaurus** there is a hierarchical and network structure at play

- Typically a thesaurus is used for library purposes, where we need to describe in-depth many resources to surface their conceptual content, and make it more likely people will find all the resources about a particular concept, even if the authors didn't use the same term
- Example of a thesaurus: used in journal databases; "Subjects" (terms applied by professional indexer) (not author supplied keywords)

2 Syndetic Structure of Thesauri

3 principles for thesaurus organization:

1. Equivalence (synonymy; which terms are roughly the same for this purpose)
2. Hierarchy (some terms are more specific versions of each other)
3. Other relationsihps

Place where you can browse through a thesaurus: ERIC Thesaurus (Education Resources Information Center): on the topic of education

The default display of a thesaurus is its alphabetical list: e.g.,

> Art History
> > BT Intellectual History
> > RT Artists
>
> Artists
> > RT Art History
>
> Intellectual History
> > NT Art History

- Note: every tie is listed at both ends of the relationship (relationships are reciprocal; e.g., because "Artists" is RT of "Art History", "Art History" is RT of "Artists")

Can be helpful when making a thesaurus to make network view

3 Relationships Used in Thesauri

3.1 Equivalence

All these terms have inclusive relationships with other terms; users would be well served by considering a term closely related to the term they're using

Table 1: Equivalence Relationships

Type	Sub-type	Example	Notes
Equivalence	Equivalent Phrases	*Athabaskan* and *Athapaskan*	
	Inverted Forms	*Bilingual education* and *Education, bilingual*	Important for alphabetical order; no matter which term someone started with, they will be directed back
	Acronyms and Abbreviations	*UNDRIP* and *UN Declaration on the Rights of Indigenous Peoples*	
	Antonyms	*Student Retention* and *Dropouts*	Context-specific; whenever you are talking about one, you are talking about the other (the concept is the same, but you can express it in a positive or negative sense)

Lead-in Terms are terms that lead a person to an authorized heading

> Athabaskan USE **Athapaskan**

- "Athabaskan" is lead-in term; the only time it will appear in the system is to direct you to use another term

Upward Posting: e.g., if there is so little content about specific term, thesaurus directs user to use more general term

> Cats USE **Animals**

- obviously these are not equivalent concepts, but in the context of the collection's scope it might be appropriate

3.2 Hierarchical Relationships

Table 2: Hierarchical Relationships

Type	Sub-type	Example	Notes
Hierarchical Relationships	Generic Relationships	*Thinking* NT *Reasoning*	A link to a more specific type (e.g., genus)
	Instance Relationships	*English* NT *Germanic Languages*	
	Partitive Relationships	*Finger* NT *Hand*	

In hierarchies you can have lots of depth; long strings of narrow term / broad term relationships
You would never see any more together than just strict neighbours

Also might find lots of breadth: e.g., amusements: carnivals, casinos, charades, dance, etc.

3.3 Associative Relationships

The main difference between associative and hierarchical relationships is that the relationships are not transitive:

- Hierarchical: if bumblebees are type of bees, and bees are type of insects, bumblebees must be type of insects (transitive relationship)

- Associative: bees are related to honey, honey is related to toast, but bees not related to toast (non-transitive)

Table 3: Associative Relationships

Type	Sub-type	Example
Associative Relationships	Operations and Instruments	*Hairdressing* RT *Hair Driers*
	Actions and Products	*Speaking* RT *Speech*
	Causal Relationships	*Accidents* RT *Injury*
	Field of Study and Objects of Study	*Linguistics* RT *Languages*

If you don't stick to this list, you'll get a lot of bloat in thesauri

One way to cut down on bloat in a thesaurus: not to make related term relationships between things that have a hierarchical relationship (e.g., Flies and Bees don't need RT, if they both lead back to Insects)

3.4 Explanations and Clarifications

Scope Notes: a brief note that describes how the heading is to be used

> Art
> SN: used to refer to the general processes and results of aesthetic expression, as opposed to more precise headings such as "art products," "visual arts," "fine arts," "Painting (Visual Arts)," etc.