

Bibliotekarstudentens nettleksikon om litteratur og medier

Av Helge Ridderstrøm (førsteamanuensis ved OsloMet – storbyuniversitetet)

Sist oppdatert 05.02.24

Om leksikonet: https://www.litteraturogmedieleksikon.no/gallery/om_leksikonet.pdf

Tesaurus

(_sjanger) Fra et gresk ord for “skattkammer”. Et strukturert system av emneord der relasjonene mellom ordene (termene, begrepene) kommer tydelig fram. Dette systemet av ord er nyttig blant annet for søking etter informasjon om et emne. Flertallsformen er tesauri. Det finnes tesauri for medisinsk terminologi, pedagogisk terminologi og innen mange andre (fag-)områder og temaer.

Ordet tesaurus brukes noen ganger også om en vanlig ordbok og andre vokabularer: “However, the term ‘thesaurus’ is often used loosely for any subject-related vocabulary.” (<https://www.anzsi.org/member-area/recommended-reading-and-guidelines/thesauri/>; lesedato 11.12.23)

En tesaurus – i den mest presise betydningen av ordet – er en omfattende ordliste med emneord (substantiver, begreper) hvor relasjonene mellom innførlene vises, blant annet med over- og underordning, synonymer (eller sideordnete termer), relaterte termer og eventuelt antonymer (dvs. motsatte ord). Det kan være likeverdige (ekvivalente) ord. En tesaurus inneholder også ofte definisjoner og synonymkontroll (dvs. skillet mellom foretrukne og ikke foretrukne termer).

En tesaurus kan være en database eller et trykt verk som inneholder ord/termer med bestemte relasjoner seg imellom: innholdsrelasjoner, hierarkirelasjoner, abstraksjonsrelasjoner, assosiasjonsrelasjoner m.m. Det vil si at det for bestemte ord står oppført f.eks. synonymer og ord med under- eller overordnet betydning. Tesaurusen viser altså begrepsstrukturer i språket.

En tesaurus er en ordoversikt der ordenes hierarkiske forhold til hverandre blir forklart, med synonymer, homonymer (like ord med forskjellig betydning) m.m. (Rehm 1991 s. 265). Noen av relasjonene blir forkortet slik på norsk:

OT: overordnet term

UT: underordnet term

ST: sideordnet term

BF: brukt for (ekvivalensrelasjon)

I store engelskspråklige tesauri, blir disse forkortelsene brukt:

“BT = broader term
 BTG = broader term (generic) [f.eks. forholdet hund – rottweiler]
 BTI = broader term (instance) [f.eks. fjell – Alpene]
 BTP = broader term (partitive) [f.eks. hær – bataljon]
 GS = generic structure
 HN = history note
 NT = narrower term
 NTG = narrower term (generic)
 NTI = narrower term (instance)
 NTP = narrower term (partitive)
 RT = related term
 SEE = equivalent to U (USE)
 SN = scope note [forklaring, inklusiv definisjon]
 TT = top term
 U = use
 UF = used for
 UF+ = used for ... and ...
 USE+ = use ... and... ”
 (https://webstore.ansi.org/preview-pages/NISO/preview_ANSI+NISO+Z39.19-1993.pdf; lesedato 01.02.23)

“A ‘thesaurus’ in indexing and information retrieval contexts is a tool for subject indexing and searching which:

- identifies relevant terminology for a subject area
- maintains vocabulary control
- presents a structure for the vocabulary
- provides a hierarchical display of Broader and Narrower terms within the subject field
- indicates relationships between terms (Related terms), and
- maps non-preferred synonyms to the preferred thesaurus term.” (<https://www.anssi.org/member-area/recommended-reading-and-guidelines/thesauri/>; lesedato 11.12.23)

“A critical tool intended to promote retrieval of information through the use of a controlled vocabulary is the thesaurus. A thesaurus is “a controlled vocabulary of terms in natural language that are designed for postcoordination” (NISO 1993, p. 1). Indexers and catalogers use such a vocabulary to determine precisely what terms to assign as access points to a record or document. Users of indexes, catalogs, and databases can use thesauri to locate the proper terms for searching or to discover related terms and subjects. According to the National Information Standards Organization, thesauri accomplish four purposes: they (1) translate the language of authors and indexers into a fixed vocabulary that can be used for indexing and information retrieval; (2) help ensure consistency when assigning

headings and index terms; (3) reveal relationships between terms, and (4) assist in the information retrieval of information and documents (NISO 1993). As can be seen from these purposes, thesauri play critical roles in structuring access as well as in assisting in retrieval itself.” (Rubin 2000 s. 180-181)

“A structured list of expressions intended to represent in unambiguous way the conceptual content of the document in a documentary system and of the queries addressed to the system. [...] A controlled vocabulary in which concepts are represented by descriptors, formally organized so that paradigmatic relationships between the concepts are made explicit, and the descriptors are accompanied by lead-in entries for synonyms and quasi-synonyms. [...] The purpose of a thesaurus is to guide both the indexer and the searcher to select the same descriptor or combination of descriptors to represent a given subject.” (<http://www.parliaments.info/>; lesedato 10.06.13)

“Sometimes used interchangeably, taxonomies and controlled vocabularies are controlled lists of terms used to organize information. A controlled vocabulary is “an organized arrangement of words and phrases used to index content and/or to retrieve content through browsing and searching.” (Harpring). Controlled vocabulary can be used in any industry that collects and uses information, such as academic research, libraries, corporations, governmental organizations, etc. Common types of controlled vocabularies include term lists, authority files, and thesauri. Because controlled vocabularies require the use of predefined terms, they can be challenging to adopt and apply, but using controlled vocabularies during data or metadata creation supports consistency and accuracy. There are well-established vocabularies for personal and corporate names, geographic names, topics, concepts, resource types and genres, and languages. [...] A thesaurus is a kind of dictionary represents all the concepts for a specific domain in a consistent manner and labels each concept with a preferred term. Like the previously described examples of controlled vocabularies, thesauri contain preferred terms, variant terms, and broader and narrower terms. Additionally, the thesaurus also includes related terms, which may or may not be part of the same hierarchical structure of the term. A commonly used thesaurus for describing art, architecture, and material culture objects is the Getty Art & Architecture Thesaurus. For instance, the Getty Art & Architecture Thesaurus entry for “cellular telephones” gives alternate terms such as mobile phone in English, Handy in German, and shǒu jī in Pinyin Chinese. The entry also provides a definition and a hierarchical categorization of the term.” (<https://pitt.libguides.com/metadatadiscovery/controlled-vocabularies>; lesedato 11.12.23)

“A thesaurus combines the characteristics of synonym ring lists and taxonomies, together with additional features. A thesaurus is a semantic network of unique concepts, including relationships between synonyms, broader and narrower (parent/child) contexts, and other related concepts. Thesauri may be monolingual or

multilingual. Thesauri may contain three types of relationships: equivalence (synonym), hierarchical (whole/part, genus/species, or instance), and associative. Thesauri may also include additional peripheral or explanatory information about a concept, including a definition (or scope note), bibliographic citations, and so on.” (https://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/what.pdf; lesedato 25.01.24)

“Also referred to as in information thesaurus or information-retrieval thesaurus (to distinguish it from the synonym dictionary type) [...] a thesaurus: “A controlled vocabulary arranged in a known order and structured so that the various relationships among terms are displayed clearly and identified by standardized relationship indicators.” So, a thesaurus is a kind of controlled vocabulary or a kind of knowledge organization system which is quite structured and has certain standard features: terms that are noun phrases, hierarchical relationships between terms, associative (related, but not hierarchically) relationships between terms, “synonym” or variants, which are called nonpreferred terms, and scope notes on terms. Other metadata on terms is possible, and variations of hierarchical and associative relationships may also be possible. [...] Even if most relationships are hierarchical, there may be times when creating an associative relationship between related subjects seems logical and would be helpful to users, such as relating between a process and agent, action and property, cause and effect, object and origins, discipline and practitioner, etc.” (<https://www.hedden-information.com/what-is-a-thesaurus-and-what-is-it-good-for/>; lesedato 30.06.23)

“Thesauri usually consist of a core list of index terms, usually single words but sometimes combinations of words, phrases, or names. Terms in a thesaurus are sometimes referred to as descriptors. The terms comprising the core list are the ones determined to be acceptable for use to gain access to the information system (e.g., catalog, database, index). Associated with these terms are additional terms that are not used for direct access to the information system, but which, if consulted first, will direct a user to the appropriate term for use. These are sometimes referred to as cross-references or lead-in terms. Such an expansion is critical to a thesaurus because individuals trying to gain access to a particular information system are usually not aware of the terms that were selected. Rather they come with their own vocabulary and their own ideas as to which terms reflect the subject matter. As with controlled vocabularies, the terms selected are also enriched by notations about other associated terms and their relationships. Although thesauri have been in existence for many years, they have become especially important in the use of automated information-retrieval systems. Such thesauri provide a wide variety of available terms for those assigning terms to documents for access and for searchers seeking information. Examples of such thesauri include the *ERIC Thesaurus* and *INSPEC Thesaurus*.” (Rubin 2000 s. 180-181)

“Ivar Aasens *Norsk Maalbud* blir rekna som den første norske tesaurus, ei ordbok som ikkje er organisert etter alfabetet, men etter tyding, gruppert i ein

struktur av overordna og underordna omgrep. Aasen hadde *Norsk Maalbunad* (heretter NMB) ferdig i manuskript i 1876, tre år etter at *Norsk Ordbog med dansk Forklaring* kom ut. Manuset var berre førebels; Aasen vurderte å bearbeide det for trykking, men kom aldri så langt. NMB vart derfor liggande i papira hans, og først gitt ut etter Aasens etterlatne manuskript av Sigurd Kolsrud i 1925. Pionerverket og den internasjonale klassikaren på dette feltet er Peter Mark Roget sin *Thesaurus of English Words and Phrases* frå 1852, eit verk Aasen kjente til og brukte i arbeidet sitt.” (Lars S. Vikør i *Språknytt* nr. 2 i 2013 s. 32)

Johan Hammond Rosbachs *Ord og begreper: Norsk thesaurus* (2001) “er ingen vanlig ordbok, du kan ikke finne en definisjon eller direkte oversettelse, men du finner synonymer, kanskje også antonymer, og ord som kan gi assosiasjoner til det samme emnet [...] Han mener det er helt unødvendig å bruke de engelske ordene som har overstrømmet landet vårt de senere årene: det finnes da gode norske ord for skate-board, know-how, og hvorfor skal vi ha et uttrykk som heter fifty-fifty, når vi like gjerne kan si femti-femti?” (<https://www.nrk.no/kultur/sprakets-skattkammer-1.527296>; lesedato 27.02.13)

“The Ei Thesaurus was developed in response to requests from users of Engineering Information databases for an access tool that would make searching easier and more reliable. It contains nearly 9,400 descriptors (preferred terms) and over 9,000 entry (non-preferred) terms, for a total of over 18,400 access points.” (Milstead 1992 s. v)

“Four kinds of relationships may be provided with a descriptor. BT (Broader Term), NT (Narrower Term), RT (Related Term), and UF (Used For). [...] RTs may indicate any non-hierarchical type of relationship. [...] An example of a descriptor record is shown below:

Acoustic imaging

SN: Production of real-time images of the internal structure of opaque objects

DT: Predates 1975 [DT = “entry date”, gjelder første gang termen ble brukt i en index]

UF: Acoustic lenses

BT: Imaging techniques

NT: Ultrasonic imaging

RT: Acoustic holography

Acoustics

Diagnosis

Lenses” (Milstead 1992 s. vi)

Et annet eksempel fra Milstead:

“Waveguides

DT: Predates 1975

BT: Electronic equipment
NT: Circular waveguides
 Dielectric waveguides
 Helical waveguides
 Plasma filled waveguides
 Rectangular waveguides
 Waveguide components
RT: Antenna feeders
 Antennas
 Coaxial cables
 Guided electromagnetic wave propagation
 Gyrators
 Microwave devices
 Telecommunication links
 Telephone lines” (Milstead 1992 s. 748)

Franskmannen François Garniers *Ikonografisk tesaurus* (1984) er et system av deskriptive representasjoner (<http://bbf.enssib.fr/consulter/bbf-1985-02-0180-002>; lesedato 05.09.14). Verket er lagd for å kunne beskrive og identifisere gamle kunstverk, f.eks. malerier, tegninger, skulpturer og bygninger, uansett hvilke materialer og teknikker som er brukt til å skape verkene. Tesaurusen opererer med 3200 beskrivende kjennetegn som brukes til å karakterisere kunst, enten det kunsten viser er kropper, måltider, slåssing eller noe annet. Garnier har lagd lange lister over fenomener i naturen som kan forekomme i bilder (Garnier 1984 s. 51-61), samt fenomener knyttet til menneskets kropp og fysiske liv (63-73), moral og samfunn, politikk og administrasjon, våpen og militærvesen, jordbruk, jakt og fiske, industri, transport, vitenskaper m.m. Deretter kommer fantasiskikkelser og mytiske dyr, allegoriske skikkelser m.m. Under bibelske temaer navngir han en stor mengde katolske helgener (Garnier 1984 s. 153-163).

“The *Thesaurus for Graphic Materials I: Subject Terms (TGM I)* provides a substantial body of terms for subject indexing of pictorial materials, particularly the large general collections of historical images which are found in many libraries, historical societies, archives, and museums. Developed to support the cataloging and retrieval needs of the Library of Congress Prints and Photographs Division, *TGM I* is offered to other institutions in the hope that it will fill similar needs and will promote standardization in image cataloging. *TGM I* is primarily designed for automated cataloging and indexing systems and is authorized for use in MARC (Machine-Readable Cataloging) records. The vocabulary can also be used, however, in manual indexing systems, whether card catalogs or vertical files of pictures. *TGM I* may be used to index subjects represented in a variety of still image media and formats, including prints, photographs, drawings, posters, architectural drawings, cartoons, and pictorial ephemera. The thesaurus provides a controlled vocabulary for describing a broad range of subjects depicted in such materials, including activities, objects, types of people, events, and places. *TGM I*

includes subject categories only. Thus, proper noun names of people, organizations, events, and geographic places are excluded. Art historical and iconographical concepts are not included either, but *TGMI* does supply terms for abstract ideas represented in certain types of images, such as allegorical prints and editorial cartoons. *TGMI* is designed as a tool both for those who create catalog records and for those who search for them. Entry vocabulary and scope notes are designed to serve a dual purpose: assisting indexers in selecting terms for indexing and helping researchers find appropriate terms with which to search for pictures. At present, *TGMI* contains 5,504 authorized (“postable”) terms, and 4,324 cross-reference (“non-postable”) terms, giving it a very high ratio of entry vocabulary and increasing the ease with which searchers can translate their own terminology into the language of *TGMI*. *TGMI* adheres to the American National Standards Institute (ANSI) guidelines for thesaurus construction. In developing *TGMI*, emphasis has been placed on providing easily readable notations, logical hierarchical relationships, terms structured in natural language order, clearly distinguishable public notes and cataloger’s notes, and modern American spelling conventions.” (<https://www.loc.gov/rr/print/tgm1/ia.html>; lesedato 11.12.23)

I ANSIs retningslinjer for tesauri stod det i 2010: “There are four important principles of vocabulary control that guide their design and development. These are:

- eliminating ambiguity
- controlling synonyms
- establishing relationships among terms where appropriate
- testing and validation of terms

A major goal of vocabulary control is to ensure that each distinct concept is described by a single linguistic form. If multiple forms exist, these *should* be controlled or regularized so that information or content that is provided to a user is not spread across the system under multiple access points, but is gathered together in one place. Eliminating ambiguity and compensating for synonymy through vocabulary control assures that each term has only one meaning and that only one term can be used to represent a given concept or entity. [...] Ambiguity occurs in natural language when a word or phrase (a homograph or polyseme) has more than one meaning. [...] a single word can be used to represent multiple, very different concepts. [...] A controlled vocabulary *must* compensate for the problems caused by ambiguity by ensuring that each term has one and only one meaning.” (<https://groups.niso.org/higherlogic/ws/public/download/12591/z39-19-2005r2010.pdf>; lesedato 01.02.24)

Flertydighet gjelder f.eks. ordet Mars, som kan være både en måned, en krigsgud, en planet og et sjokolademerke. Flertydighet gjelder også mange substantiver, blant annet fordi de brukes på nye områder og metaforisk. En “plattform” kan f.eks. være laget av betong, være et politisk grunnlag eller en digital kommunikasjonskanal.

“The IEEE Thesaurus is a controlled vocabulary of about 11,730 descriptive engineering, technical, and scientific terms as well as IEEE-specific society terms. The thesaurus represents terminology and vocabulary usage drawn from IEEE [Institute of Electrical and Electronics Engineers] transactions and journal articles, conference papers, and standards. The vocabulary uses American-based spellings with cross references to British variant spellings. The scope and structure of the IEEE Thesaurus reflect the engineering and scientific disciplines that comprise the societies, councils, and communities of IEEE in addition to the technologies IEEE serves. The IEEE Thesaurus is now updated on a bi-annual basis. [...] Each descriptor included in the thesaurus represents a single concept or unit of thought. The scope of the descriptors is based on the material as presented in IEEE periodical articles, conference papers, standards, and/or IEEE organizational material. The IEEE Thesaurus is a tool to improve and enhance the indexing and retrieval of articles and other material from IEEE periodicals, conferences, standards, and other publications [...] It provides a common and consistent language for authors, researchers, and online discovery.” (<https://www.ieee.org/publications/services/thesaurus.html>; lesedato 11.12.23)

Det norske Nasjonalbibliotekets *Norsk tesaurus for sjanger og form* (<https://www.nb.no/nbvok/ntsf/nb/>) “dekker både barne- og voksenalderlitteratur. Termene finnes på bokmål, nynorsk og delvis engelsk. Referanse til tilsvarende begrep i Library of Congress Genre/Form Terms er tatt med der slike finnes.” Dette var innførselene på A i januar 2024:

ABC → ABC Bruk termen ABC-bøker

ABC-bøker

Abstrakt → Abstrakt Bruk termen Sammendrag

Abstrakter → Abstrakter Bruk termen Sammendrag

Action

Actionfilmer → Actionfilmer Bruk termen Action

Adapsjoner → Adapsjoner Bruk termen Adaptasjoner

Adapsjoner (film) → Adapsjoner (film) Bruk termen Adaptasjoner

Adaptasjoner

Adaptasjoner (film)

Adressebok → Adressebok Bruk termen Adressebøker

Adressebøker

Adressekataloger → Adressekataloger Bruk termen Adressebøker

Adresselister → Adresselister Bruk termen Adressebøker

Adresseregister → Adresseregister Bruk termen Adressebøker

Adresseregistre → Adresseregistre Bruk termen Adressebøker

Adventskalenderbøker → Adventskalenderbøker Bruk termen Julekalenderbøker

Aforismer

Agentfortellinger → Agentfortellinger Bruk termen Spionfortellinger

Aktivitetsbøker

Aktstykker
Aktualiteter (film)
Allmensoge → Allmensoge Bruk termen Bygdehistorie
Almanakker
Alterbok → Alterbok Bruk termen Alterbøker
Alterbøker
Alternativ historie → Alternativ historie Bruk termen Kontrafaktisk historie
Amatørfilmer
Amerikabrev
Anatomiske atlas
Andaktsbøker
Anekdoter
Animasjon (film) → Animasjon (film) Bruk termen Animasjonsfilmer
Animasjonsfilmer
Animasjonsinnslag
Anime (film)
Anmeldelser
Annaler
Antologier
Arbeidersanger
Arbeidsbøker → Arbeidsbøker Bruk termen Oppgavebøker
Arkeologiske utgravingsrapporter → Arkeologiske utgravingsrapporter Bruk
termen Utgravningsrapporter
Arkeologiske utgravningsrapporter → Arkeologiske utgravningsrapporter Bruk
termen Utgravningsrapporter
Artslister
Astronomisk årbok → Astronomisk årbok Bruk termen Efemerider
Astronomiske årbøker → Astronomiske årbøker Bruk termen Efemerider
Atlas
Auksjonskataloger
Autobiografier → Autobiografier Bruk termen Selvbiografier
Avhandlinger
Avisartikkel → Avisartikkel Bruk termen Avisartikler
Avisartikler
Aviser
Avisfortegnelser
Avisindekser
Aviskataloger → Aviskataloger Bruk termen Avisfortegnelser
Avisklipp → Avisklipp Bruk termen Avisartikler
Aviskronikker → Aviskronikker Bruk termen Kronikker
Avislister → Avislister Bruk termen Avisfortegnelser
Avlede verk

Niki Segnits *The Flavour Thesaurus* (2010) bruker sjangerbetegnelsen om et verk med 980 ingredienspar fra agurk og dill til østers og vannmelon. “The book takes

99 basic flavors and examines how each pairs with the remaining 98. The author acknowledges that the list of 99 and the pairings were only semi-scientific – instead it's more a way of exploring pairings, categorizing recipes, and sharing personal anecdotes. It's really about the emotions that different foods generate.” (<http://www.inventinginteractive.com/2013/10/16/empowering-discovery/>; lesedato 13.02.14)

Litteraturliste (for hele leksikonet): <https://www.litteraturogmedieleksikon.no/gallery/litteraturliste.pdf>

Alle artiklene i leksikonet er tilgjengelig på <https://www.litteraturogmedieleksikon.no>