

Libraries at University of Nebraska-Lincoln
Library Philosophy and Practice (e-journal)

University of Nebraska - Lincoln

Year 2008

Dialectic Schemes in Thesaurus Creation

Maziar Amirhosseini
AJT, Tehran, Iran, mazi.lib@yahoo.com

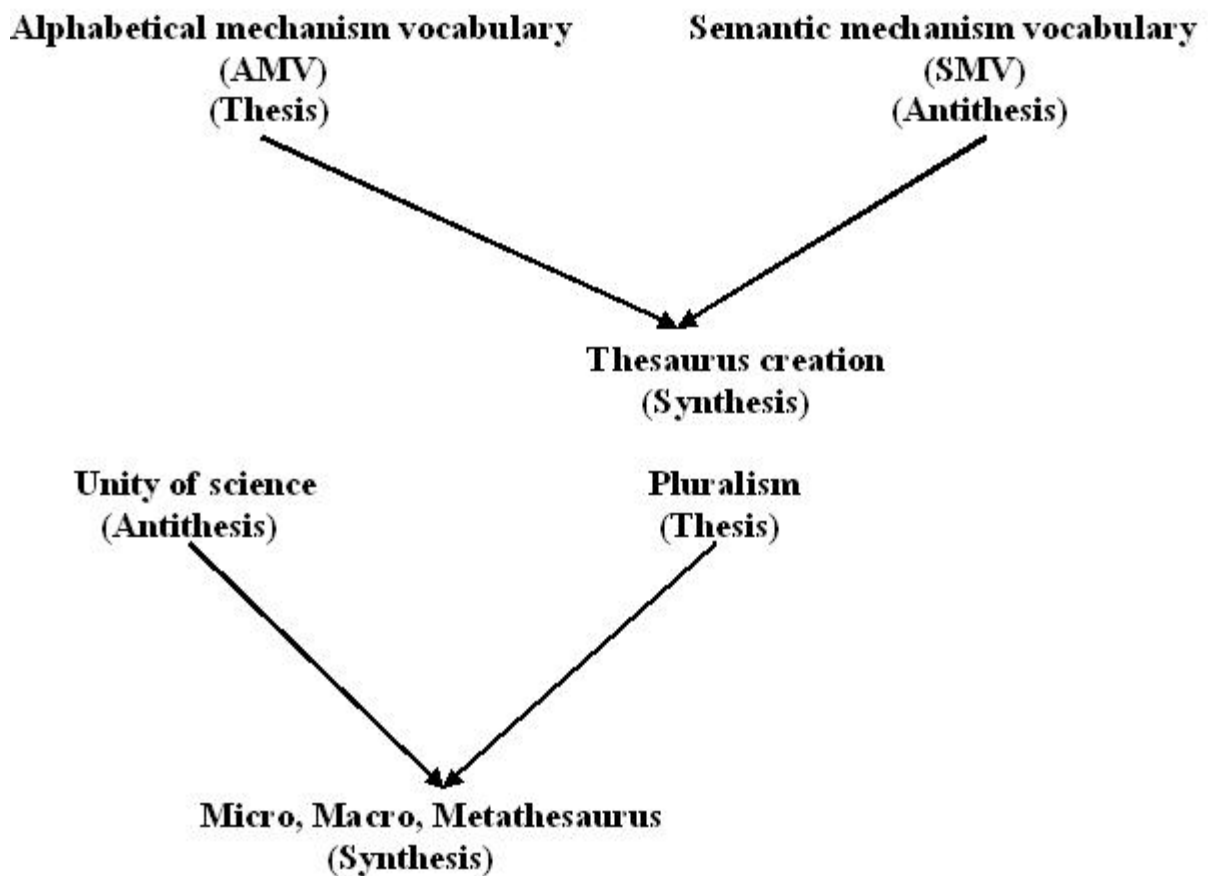
Dialectic Schemes in Thesaurus Creation

Maziar Amirhosseini
AJTKeshavarz Blvd.
Valiasr Compus
Tehran, Iran

Introduction

Dialectics is a form of reasoning and argument that aims to resolve and synthesize opposing views or ideas. Contemporary dialectic thought in the West is associated with Hegel, the German philosopher whose dialectical scheme described the progress of history and ideas from thesis to antithesis and then to synthesis. Hegel's work followed on that of Immanuel Kant, Johann Gottlieb Fichte, and others ("Hegel, Georg Wilhelm Friedrich"). Hegel's dialectics (the resolution of opposites) include triads that represent thesis, antithesis, and synthesis. In Hegel's view, we do not create dialectic categories. The categories already have a logical relationship that we must discover. The process of deduction shows the existence of an objective world which is separate from our minds. Our task is not to invent a method for deducing categories but to distinguish the categories that are deduced from each other (Stace 1955). That process includes This article looks at the logical relationships between the categories that have led to the development of the thesaurus, including the historical origins of thesaurus construction. The research uses the dialectical, logical relationships of thesis, antithesis, and synthesis. Each step in the dialectic of thesaurus creation includes subjective and objective thought as part of the thesis and antithesis, resolved by the more absolute thinking of the synthesis. [Table 1](#) summarizes the research results, which are explained in greater detail after the table.

Figure 1: Dialectical movement in creation of thesaurus



The first level of the figure has three steps, which are

1. Alphabetical mechanism vocabulary (AMV) as the thesis
2. Semantic mechanism vocabulary (SMV) as the antithesis
3. Thesaurus creation as the synthesis

The second level of the figure has three steps, which are

1. Pluralism as the thesis
2. Unity of science as the antithesis
3. Micro, Macro, and Metathesaurus as the synthesis.

The synthesis in the first level is the thesis in the second level.

Definition of Terms

Tradition: Tradition is contrasted with innovation. Tradition is a mixture of beliefs and opinions, which are in a historical process and are accepted in their own societies at the same time.

Modern thought: Modernism or modern thought uses reason to understand and navigate the world.

Modern tradition: This is a set of beliefs and attitudes associated with modern thought, the rational approach of the post-medieval world.

Unity of science: This concept states that all the sciences together are a unified whole. The unity of science thesis is associated with Paul Oppenheim and Hilary Putnam. Jerry Fodor presented the best-known counter-argument (20)

Philosophy of mind: Putnam's contribution to the philosophy of mind argue for functionalism, the idea that the mind is like a computer, although he later recanted this view ("Philosophy of Mind")

Functionalism: Functionalism accounts for mental states by focusing on manifestations such as beliefs, desires, and emotions, looking at consciousness as the interaction of functional processes, rather than seeking explanation by referring to physical media such as neurons ("Functionalism [Philosophy of Mind]"))

Thesaurus: Terms or vocabulary arranged to show their semantic relationships.

Microthesaurus: The convergence of several specialized thesauri (Yancey and Carson 2004)

Macrothesaurus: The convergence of several microthesauri. The interrelationship of microthesauri is created by broad classification codes, while the relationship between microthesaurus with macrothesaurus is created with bridge words. (Foskett 1980, p. 458-459)

Metathesaurus: A "thesaurus of thesauri," which serves as a way to harmonize different controlled vocabularies to achieve cross-file searching (Subcommittee on Metadata and Subject Analysis, 1999)

Level 1

Thesis: Alphabetical Mechanism Vocabulary (AMV)

Dictionaries are alphabetically organized. The advantage of alphabetical order is simplicity of access to vocabulary and simplicity in organization. It was an innovation that arose when literacy was becoming more universal. AMV reflects a one-dimensional way of thinking. The subjective aspect of this approach is the AMV hypothesis, i.e., the theory that AMV is a useful way to arrange material. The objective aspect is the resulting simplicity created by alphabetical organization. AMV reflects modernism, and the subjectivity of the AMV hypothesis is "crystallized" in the objective realization of alphabetical dictionaries.

Antithesis: Semantic Mechanism Vocabulary (SMV)

While alphabetical order is widespread and gives easy access to material, it also scatters related vocabulary. This need for integration of related material created the antithesis of AMV, which is organization by semantic concepts, or SMV. SMV is the spirit behind the approach of Peter Mark Roget (Roget 1995, 2000). In 1852, Roget introduced a multi-dimensional approach to vocabulary organization in his *Thesaurus of English Words and Phrases*. Roget provided not simple alphabetical lists, but a way of finding specific words, phrases, and idioms that express a concept or semantic domain. Roget's profound influence is still seen, both in the continued publication of his thesaurus, in the expansion of the concept of a semantically-arranged thesaurus to languages other than English, and in approaches to information retrieval. SMV is a modern, functionalist, and multi-dimensional way of thinking. The subjective aspect of this approach is the SMV hypothesis, i.e., the theory that SMV is a useful way to retrieve appropriate expressions and idioms. The objective aspect is the resulting collocation and retrieval of

synonyms created by semantic organization. The subjectivity of the SMV hypothesis is crystallized in the objective realization of thematically-organized thesauri.

Synthesis: Creation of the Thesaurus

A thesaurus is a set of terms that are semantically related. There was a turning point in thesaurus construction in 1951, when Hans Peter Luhn created a thesaurus that was influenced by Roget. Luhn's thesaurus model is a synthesis of AMV and SMV. Its structure creates a semantic map of the sciences, including broader and narrower terms, and a bridge between different scientific fields (related terms). The semantic pluralism of the sciences was introduced into information storage and retrieval. Luhn's thesaurus structure has had a deep and lasting influence. The one-dimensional thinking of AMV (the thesis, the dictionary) was in contradiction to the multi-dimensional thinking of SMV (the antithesis), which resulted in the creation of the thesauri used in modern information retrieval (the synthesis).

Level 2

Thesis: Pluralism

The thesis of level two is "pluralism," the idea that there are numerous disciplines in the sciences and elsewhere, which require separate, specialized thesauri. One purpose of a thesaurus is, "to provide a map of a given field of knowledge, indicating how concepts or ideas about concepts are related to one another, which helps an indexer or a searcher to understand the structure of the field" (Fosket 1980, p. 419) Only those terms that are closely related to a specific field of science are included in the thesaurus. The subjective aspect of the thesis is the pluralism hypothesis and the objective aspect is the expression of specific scientific fields. This thesis contrasts with the idea of the "unity of science," which is a concept of the second half of the 19th century. The Pluralism hypothesis helps crystallize the subjectivity in the objectivity of accounting for the vocabulary of a specific field of science. Pluralism is one-dimensional thinking, based in functionalism, and a reflection of modern thought.

Antithesis: Unity of Science

The unity of science is the idea that all the sciences form a unified whole. Physics and psychology, for example, are distinct disciplines, but the unity of science theory states that they must be part of a unified intellectual endeavor, which is science. Unity of science is usually associated with a hierarchical view of nature, where physics is the most basic, chemistry above physics, biology above chemistry, psychology above biology, and social sciences above psychology. The relationship between physics and psychology is mediated by the levels between them ("Unity of Science"). Unified intellectual endeavor is an important driver of thesaurus creation, and of the Micro, Macro, and Metathesaurus. The unity of science as the antithesis of pluralism. The subjective aspect of this subject is the unity hypothesis and the objective aspect is the integration of the sciences. The unity of science represents multidimensional thinking, and is a functionalist way of thinking.

Synthesis: Creation of Micro, Macro and Metathesaurus

With so many fields and organizations having found the need for a thesaurus, and so many thesauri having been compiled, it is surprising that more attention has not been given to the question of compatibility between the various schemes. In particular, one might expect this to apply to publishers of thesauri which cover to same or similar ground (Foskett 1980, p.456). The idea of consistency between thesauri is the first challenge to pluralistic thought. The goal of compatibility among thesauri is based in the concept of the unity of science.

The unity of science is the idea that all the sciences form a unified whole. Physics and psychology, for example, are distinct disciplines, but the unity of science theory states that they are part of a single, unified intellectual endeavor, which is science. Unity of science is usually associated with a hierarchical view of nature, where physics is the most basic, chemistry above physics, biology above chemistry, psychology above biology, and social sciences above psychology. The relationship between physics and psychology is mediated by the levels between them ("Unity of Science").

Unified intellectual endeavor is an important driver of thesaurus creation, and of the Micro, Macro, and Metathesaurus. The unity of science as the antithesis of pluralism. The subjective aspect of this subject is the unity hypothesis and the objective aspect is the integration of the sciences. The subjective aspect of unity of science is crystalized through the objective aspect: the creation of micro, macro, and metathesauri. The unity of science represents multi-dimensional thinking, comprising the integration of the sciences, and their interaction with each other. The basis of unity of science thinking is functionalism.

Compatibility and unity in thesaurus construction have not yet been achieved, for a number of reasons. Only recently has any attempt to create standard methods of construction begun. Many organizations appear unaware of either standard subject headings lists or existing thesauri in their own or related fields. The trend toward cooperation will doubtless continue, because there is a growing need, and because progress in technology makes it easier. Awareness of the value of international cooperation in intellectual matters is also growing, particularly in science and technology (Foskett 1980, p.457).

When the method of construction is the same, one classification scheme can draw on another, for example the faceted classification used by the *London Education Classification* thesaurus, *Thesaurofacet*, and the *London Classification of Business Studies* (LCBS). Cooperation among international organizations such as the United Nations, the International Labor Organization (ILO), and the Organization for Economic Cooperation and development (OECD), led to the creation of the OECD Macrothesaurus <http://bibliotecavirtual.clacso.org.ar/ar/oecd-macroth/en/index.htm>, which includes an "Aligned List of Descriptors," which is now "an authority list of terms on economic and social development, for use in the international context." The Macrothesaurus was created using a number of lists, and tried to compile all terms into a single scheme that would be useful to all constituents (9, p.457-458).

A metathesaurus is a "thesaurus of thesauri," a way to reconcile a number of controlled vocabulary schemes for cross-searching (Subcommittee on Metadata and Subject Analysis, 1999). A metathesaurus includes key concepts of the field or domain, semantic relationships between concepts such as synonym or antonym, as well as mapping from other thesauri to the key concepts. Due to the difficulty involved in constructing a metathesaurus, there are not many examples (Agogino 2004). Metathesauri are for use by system developers, but can also be a used by librarians and other information professionals (UMLS Metathesaurus 2004)

Figure 2 is a model for a complex set of controlled vocabulary components. The model has three macrothesauri, one of which is subdivided into three microthesauri. The three macrothesauri are mapped together to create a metathesaurus. Finally, one of the macrothesauri has been used to index an authority file of names and an authority file of documents (Yancey and Carson 2004)

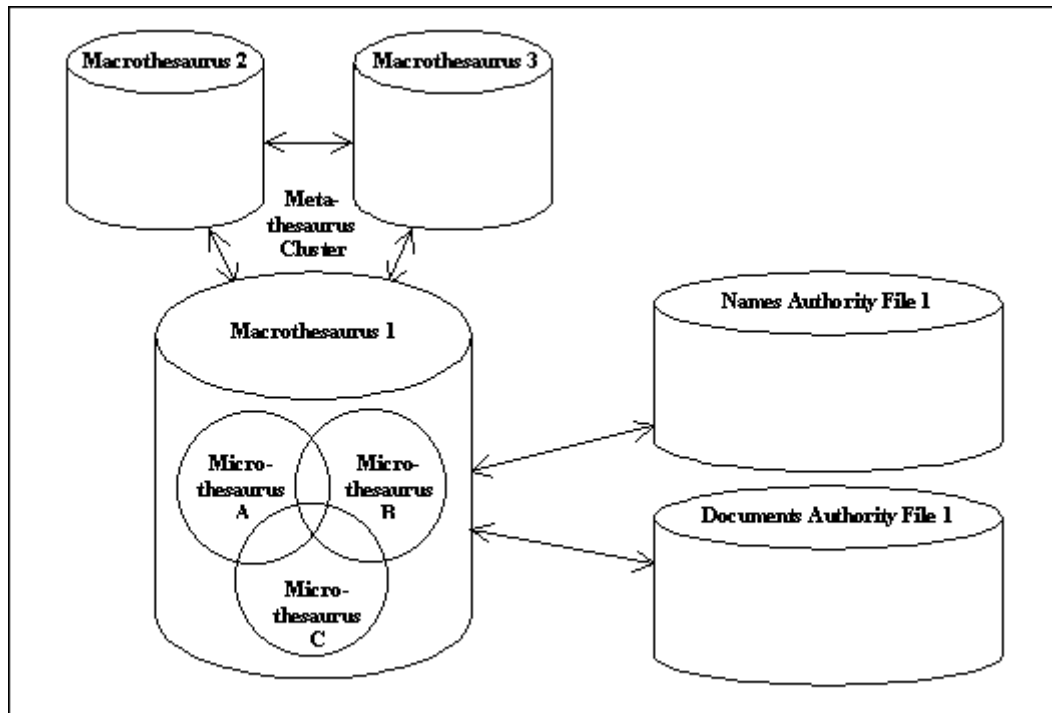


Figure 2: Controlled-Vocabulary Component Model

The convergence of several specialized thesauri produces a need for a microthesaurus. The relationship between specialized thesauri and microthesauri, and between micro- and macrothesaurus, is created through bridge words or exchange codes (broad classification categories). Microthesauri are mapped together to create a metathesaurus. The one-dimensional thinking of pluralism (thesis) was in contradiction with the multi-dimensional thinking of unity of science (antithesis), and resulted in the creation of micro-, macro- and metathesaurus (synthesis). The subjective thinking of the concept of the unity of science is manifested in the objective thought that produced these modern thesauri to achieve simplicity in reaching integrated knowledge. There is a kind of evolution to be seen in the creation of the new generation of thesaurus. The evolution and contradiction in these thesauri is reflected in the concepts and practices of information storage and retrieval.

Conclusion

In the evolution of thesaurus creation, we see the dialectic framework of thesis, antithesis, and synthesis, as well as the subjective and objective thinking that moves from concept to tangible outcome. The development of thesauri mirrors the dialectic seen in thinking about science: pluralism, unity, and the resolution of these contradictory views. The thesaurus and its creators demonstrate a psychological functionalism.

References

- Agogino, A. (2004). Developing a Learner-Centered Metathesaurus for Science, Mathematics, Engineering, and Technology Education. http://best.me.berkeley.edu/pubs/04_1202_P.pdf
- Foskett, D. J. (1980). Thesaurus. *Encyclopedia of library and information science*. New York: Marcel Dekker

Functionalism (Philosophy of Mind). *Wikipedia*.
http://en.wikipedia.org/wiki/Functionalism_%28philosophy_of_mind%29

Hegel, Georg Wilhelm Friedrich. *Encyclopedia Britannica*. <http://www.britannica.com/eb/article-41234>

Philosophy of Mind. *Wikipedia*. http://en.wikipedia.org/wiki/Philosophy_of_mind

Roget, P. (1995.) *Roget's II: The new Thesaurus*. <http://www.bartleby.com/62/>

Roget, P. (2000). *Roget's International Thesaurus of English Words and Phrases*. C. O. Sylvester Mawson. <http://www.Bartleby.com>

Stace, W. T. (Ed.) (1955). *The philosophy of Hegel: A systematic exposition*. New York: Dover Publications. Persian translation by Hamid Enayat. Tehran: Gibi.

Subcommittee on Metadata and Subject Analysis (1999). *Subject data in the metadata record, recommendations and rationale: A Report from the Subcommittee on Metadata and Subject Analysis..* http://www.ala.org/ALCTSTemplate.cfm/ALA_Subject_Data.htm

UMLS® Metathesaurus (2004). <http://www.nlm.nih.gov/pubs/factsheets/factsbj.html>

Unity of Science. *Wikipedia*. <http://en.wikipedia.org/wiki/Dictionary#searchInput>

Yancey, T., & Carson, J. (2004). *Lexicography without limits: A Web-Based solution*. <http://www.udc.es/grupos/lexicografia/x-zw.htm>

Table 1: The dialectic of thesaurus creation

Subject Characteristics	AMV	Dictionary	SMV	Roget	Thesaurus	Pluralism	Unity of science	Micro, Macro & Metathesaurus
Subjective aspect	AMV	AMV	SMV	SMV	SMV	pluralism	unity	unity
Objective aspect	reaching vocabulary	reaching vocabulary	expressing thought	expressing thought	reaching related vocabularies	expressing specific fields of science	integration between sciences	reaching integrated knowledge
Hegel's categories	Thesis	Synthesis	Antithesis	Synthesis	Synthesis	Thesis	Antithesis	Synthesis
Dimension	One	One	Multi	Multi	Multi	Multi	Multi	Multi
Modern							+	+
Modern tradition	+	+	+	+	+	+		
Functionalism	+	+	+	+	+	+	+	+
Monument	+	+	+	+	+	+	+	+