**Annette Béguin**  
UFR IDIST, University of Lille 3, France

**Thesaurus Usage and Mental Development**

**Abstract:** Linguistics cannot be the only criteria to assess the relevance of a thesaurus. Psycholinguistics, cognitive psychology and social psychology can also help you understand how users learn to master this tool. Indeed, using a thesaurus requires important abstracting abilities that may vary depending on the way the tool is structured. In a school environment with young users, difficulties in manipulating concepts through the application of a controlled language thesaurus become apparent. This finding will lead us to think about what training in the use of thesauri to offer users, and how thesauri can help in the sharing of knowledge.

1. **Introduction**

   How should we assess the cost effectiveness of a thesaurus? When comparisons are made of information retrieval in natural language with information retrieval in standardized language, the criteria that are taken into account are mostly collective. On the one hand, thesaurus indexing requires time and high skill; on the other hand, using a thesaurus enables the members of a community to query with more accuracy, as it limits the noise and the silence that retrieval in natural language would definitely entail.

   Now I would like to point out that the comparison would be more complete if it also took into consideration the cognitive cost for users in using thesauri. Indeed, the human mind will try to achieve its aim in the most economical way it can. In other words, if users have more problems in thesaurus use and expend more effort in using this tool efficiently than they feel is justified by the results, there is no doubt that they will tend to do without it.

   If average users think that such a thesaurus, even if deemed perfectly adapted by specialists, is too complicated a tool, they will not use it. They may not use it at all, if they have the opportunity to use natural language, or only in part where they think it is most convenient. For instance, they will only use the permuted alphabetic section and never refer to the tree structures or graphic presentation (terminogramme); or they will stop trying to manage by themselves and will become dependent on the librarian in order to get through their searches. Then the money invested in building the thesaurus and indexing documents with it for a community will be wasted.

   This is why I think that in order to make progress in designing and studying thesauri, we must give up a strictly linguistic point of view and observe a thesaurus as a mediation tool, in the light of research in psycholinguistics, cognitive psychology and social psychology. The cognitive cost when using a thesaurus could be assessed in proportion to the cognitive effort made by the users and the relevance of the information collected.

   I chose to turn my attention to the use of thesauri in a school environment for two reasons. Firstly, if we listen to what librarians or officials involved in education say, it seems to have become a current practice that cannot be ignored. Secondly, the difficulties and the benefits of learning to use thesauri are easier to apprehend in a school environment, where the cognitive improvement of individuals is easier to observe. Nevertheless, this is only a possible example. In a similar way, we should consider the various purposes and uses of thesauri in all the organizations which resort to them, depending on the cognitive and socio-cultural...
characteristics of their users.

I would add that my topic rests on research that has just begun. So, I cannot examine its outcome properly right now. I will only deal here with the main thoughts underlying the investigation we are setting up.

2. The Cognitive Difficulties Caused by the Use of Thesauri

2.1. Abstracting Processes such as Naming, Decontextualizing and Decentering

The major difficulty when using a thesaurus is the significant effort of abstraction which is necessary.

First of all, a thesaurus is a structured list of terms referring to the concepts of a definite field; now using a word to express a concept is a form of abstraction. The child who learns to speak first associates the word with a sample belonging to his/her experience. "Cat" stands for the white and black cat of the house. Little by little, the word will refer to a category rather than a specimen. "Cat" will mean any cat. The term will gradually emerge from the particular contexts in which it may be used and will become generally available in the vocabulary of the child who uses it (Ehrlich, Bramaud du Boucheron, Florin, 1978), (Bramaud du Boucheron, 1981).

The access to written language is seen as a decisive step in the development of abstracting ability, because it allows you to postpone an announcement in space and time. It is also possible to perceive which stage of this ability has been reached in the way the individual is capable of defining something. Children tend to use definitions based on examples (coal is used to warm yourself) before having access to logical and sectional definitions (coal; black solid fuel, the origin of which is organic).

If the gradual access to abstraction is obvious with young children, it actually concerns all learning situations, even those encountered by adults. As Lev Vygotski points it out in Pensée et Langage, a word cannot so easily be linked with a clearly defined concept. Or else, to know the word would mean to own the concept. But the role of the word is rather to lead towards the concept. Our empirical experience provides us with preconcepts that are efficient but unconscious and not formalized. It is thanks to verbal exchanges with someone else that we learn which word corresponds to which mental preformation. Then the formation of the concept, strictly speaking, starts up and it will never stop as long as we keep the word in mind. Any new piece of knowledge will complete the previous one and refine our representation. Thus, by referring to the concept, the word is a catalyst for all the information in keeping with it, as well as steering the person's attention through the various linguistic activities.

Naming is naturally associated with the abstracting process. Thesauri are very abstract tools too, since they organize names in a certain field of knowledge. As concepts are represented by keywords, they could be defined as potential associations of ideas. Keywords are cut out from their contexts, but they mean possible virtual contexts. When indexing has been carried out by a human mind, users have to make efforts to rise from the empirical context concerned by their searching in order to give a name to the object of this research. At the same time, they have to imagine the way the thesaurus has been used by the indexer, who also had to abstract concepts from their varied contexts in order to insert them in a structure. Therefore, users have to stop focusing on their own intellectual activity, they must understand the reversibility of the documentary mediation and consider their own searching process from a more general standpoint.

When one uses a thesaurus, the abstracting approach is thus undertaken at different levels: It can already be perceptible in the lemmatizing process, which consists in bringing the
inflected forms of a term back into its full form. This is the only way to introduce a term into a thesaurus. The abstracting approach is also crucial in the decontextualizing of concepts and in their possible insertions elsewhere. It finally affects the communication itself, as the use of thesauri implies that users will conduct efficient searching only if they can imagine the other side of their research, that is to say the indexing.

2.2. Sectional Hierarchies

Using a thesaurus also requires an important lexical flexibility. Users should know how to leave the syntagmatic order of experience to consider the paradigmatic order of a constructed knowledge. They should be able to give up the term that they spontaneously refer to as the object of their searching, and mentally move towards its synonyms or antonyms. They must show their ability to scan the hierarchy in which the terms are displayed in related categories.

But the ability to categorize and the mental flexibility we have just mentioned are not exclusively linked to the knowledge of a field. These are biologically constructed abilities, depending on the growth of one's sensible and formal way of thinking. According to Jean Piaget (Piaget, 1964), the classifying ability as well as the ability to focus on something outside the self emerge during the second stage of childhood. Sectional relations can only be controlled around the age of twelve. It is all the more difficult as classifying concerns concepts, the latter being expressed with words from the written language, which Vygotski calls the "algebra of language".

Therefore, as thesauri put order into concepts and relations between concepts, and because they constantly involve resorting to the hypothetico-deductive thought, they appear as very abstract tools. So, their users need a completed form of development in formal thought.

A master's degree student examined the use of the database Floris by adults who attended a training course in gardening. Her work clearly shows an underuse of the database, because the users do their searching with only the name of a plant, when they are left to themselves. Though the search could be more thorough and useful if it referred to the families of plants or their specific characteristics, they miss the information because of the level of abstraction it requires.

Under these conditions, what should we think about an earlier and earlier awareness to the use of thesauri in the education system, in keeping with the development of documentary software in schools and libraries?

2.3. A Few Implications for the Development of Thesauri

We must first pay attention to the way thesauri are designed. Terminologists should take account not only of the linguistic structuring of fields by specialists, but also of the cognitive level of users and their abilities to master terminologies.

Thesauri could be designed with several levels of searching. In Conceptions et connaissances, André Giordan and Gérard De Vecchi (Giordan, Girault, Clément, 1994) suggest supporting conceptual training by using different levels of formulation for concepts. Concepts belong to explanatory models that are more and more elaborate as the training develops. These levels are defined depending on the age or on the mastery in the field concerned. The formulations suggested by André Giordan provide examples where several concepts in biology rely on a terminology, which he carefully displays in bold type. Maybe we could instead imagine a thesaurus that evolves and is flexible thanks to modules fitting into each other, and the terms of which are in keeping with the users' knowledge.

Besides, classical thesauri only use substantives. The relevance between a name and an object is a form of abstraction. The research in cognitive representation through language
shows that the development of this ability is gradual. For instance, in every day life, the categorizing of objects more commonly results from typical examples of a class than from strictly logical relations. Here I am referring to the notion of typicality defined by Rosch and further studied by Françoise Cordier (Cordier, 1994). As far as thesauri are concerned, we could take account of the significant character of some typical examples. Logically enough, I can define a square as a parallelogram with four equal sides and four equal angles, as a lozenge with four right angles, or as a rectangle with four equal sides. And yet, the relationship between a rectangle and a square is usually easier to grasp than the relationship between a square and a lozenge or the definition by its including generic term, in other words a parallelogram.

Moreover, Françoise Cordier wonders about the ability to categorize actions. In this case, the naming necessary to change verbs into substantives introduces another difficulty. In order to nominalize a verb, you need to cut it away from both its agent and its object. The numerous uses of nominalizations in a text correspond to a high degree of abstraction. Once more, this observation could concern thesauri. The use of infinitives as substantive forms of verbs would often be a plainer means of access to information, especially when the nominalized form of a verb is related to a form different from its verbal stem: if digestion does not differ much from digest, it gets more difficult to pass from nourish to nutriment, and even more so to go from move to locomotion.

The choice of terms contributes to help users handle concepts, but other tools, added to the lists of thesauri, could also provide assistance, such as list of synonyms, and possible access to a dictionary or examples where the terms are shown in context. Computers increase the possibility of multiplying such aids, like understanding a wrong spelling, a shortened word, or automatic up or down tracing in a hierarchy of words. Such an impact, depending on the type of users, could then be an important subject of assessment. When observing this influence, the difficulty lies in the need to differentiate between what results from the inherent cognitive difficulties of language and what results from the use of computers.

3. Thesauri and Cognitive Structure

3.1. Terminograms and Mental Representation (Relations and Proximity between Concepts)

Up to now we have focused on terms. Now I would like to present their organization and recall, the often quoted analogy between the structuring of a semantic field through a diagram and the patterns of mental representation as elaborated by cogniticians. Concepts form the nodes and their relations are represented by arcs. Cogniticians generally agree to use this pattern to represent mental activities like understanding or memorizing (whether this pattern is considered as emerging from this mental activity or preexisting to it). As the cognitive structure of individuals is flexible, moveable and constantly developing, it can admit vague concepts that may also be located in two different places. On the other hand, ambiguity is undoubtedly out of place in a thesaurus which is set by a functional agreement inside a community. It is precisely the reason why individual knowledge and collective knowledge could meet in the patterns established for thesauri.

As far as that particular point is concerned, a school environment is literally a laboratory. There have been interesting attempts by Sciences didacticians to try and agree with the specialists of each subject about the knowledge that should be taught. The patterns of concepts are then the favored tools in discussions (Novak, 1990). By giving a representation in space of the short or long distance between concepts, one can clearly visualize a field of knowledge.
The diagrams that are consequently created are called maps, because they enable users to be directed. The metaphors for space (field, way, area...), that symbolically refer to human knowledge, are as old as graphic representations. Mental visualization refers us to the loci used during antiquity. Diagrams are both a support for memory and a means to agree about the things to invest within a community. They make metacognition easier, the latter having also a deciding role to play in the learning approach.

It is surprising to notice that up to now, in school environment, the research in didactics has not joined the practical experiences of terminologists. Indeed, using a thesaurus is a source of difficulties that are not to be left aside, while it can also be a support to training and to the structuring of concepts. The thesaurus Motbus, which is used in every French school, has been built by a team composed of librarians, who received no help from didacticians or psycholinguists.

3.2. Training through the Visualizing of a Field; Lexical Flexibility, Inference, and the Assimilation of Professional and/or Encyclopedic Languages

Some may think of the graphs of thesauri as tools of heuristic value. They enable users to have a general view of a field. In this field, they enable them to identify the relationship between the areas they have previously identified and the areas they have not explored yet. The relational set advises users to imagine intuitively the elements that could define terms, then to turn to a dictionary to check. They can perceive what separates them from the built-up collective knowledge.

The users of a thesaurus must show lexical flexibility, and we might think that the more they get used to thesauri, the more they will improve this lexical flexibility. If the socializing of knowledge requires mastering vocabularies, then giving access to a graph is an educational step.

Under these conditions, when designing the training for users, we should take this data into account. Therefore it is no more a matter of teaching users to use thesauri but to teach them, by getting them to use thesauri. It means that we no longer need to consider terminograms as tools exclusively made for librarians, but rather as temporary collective reference. If one learns how to move within a thesaurus, one is to get used to the mental working of the environment in which one develops.

In the teaching profession, it means that thesauri should be conceived by both terminologists and didacticians. When a standardized encyclopedic language is designed for use in teaching, it can hardly be done without basically analyzing not only the general state of knowledge in each subject, but also the knowledge in various subjects that a community chooses to pass on to the next generations depending on their stages of development.

Likewise in organizations, a collective agreement on the knowledge that one wishes to share is necessary. This is why Anne Condamines notes how interesting it would be for a linguist-terminologist to "get fully acquainted with the culture of the company" for which he/she works (Condamines, 1994).

For a global knowledge that is represented in a diagram to mean something for an individual, the interferences with his/her mental and linguistic structure must be good enough. As a consequence, a thesaurus, a linguistic mediation tool, may be seen as more or less user-friendly. It means that thesauri, if they are to be learning tools, should not be made according to the specialists' vocabulary only. They should also take account of the logic involved when exchanging knowledge within a social circle, such as it appears through the language that is used.
4. A Few Ideas for Training

4.1. Indexing one's own Work; Building Conceptual Maps with Mastered Fields

In this sense, we can consider training for the use of thesauri from two angles. First users must learn what a thesaurus is, then they must use the thesaurus with efficiency and judgment, in order to link their own knowledge to the collective knowledge.

The librarian usually swiftly explains what a thesaurus is and how one may use it, then he will let the users grapple with the tool. It is hardly efficient. In order to understand the mediation that a thesaurus entails, the ideal thing to do would be to explore the two poles of documentary exchange, by becoming an indexer before being a user.

Moreover, the training stages should be progressive, so as to single out difficulties. First, training should consist of creating key-words and organizing them in natural language based on one's own writing. Then, one should turn to other people's writing. Access to standardized language would come later, because it represents an additional step on the abstraction ladder.

At every stage, the librarian would be there to help, to comment, and to promote discussions. Working together with other people has the remarkable advantage of forcing users to put into words the reasons for their choice. It gives voice to the act of communication, which is implied by indexing.

4.2. A Real Interactivity; Participating in Designing and Developing the Tool

The above suggestions concern the school environment. Can we take this any further? If we consider that a thesaurus gives the image of a collective knowledge, it would appear necessary to involve users, or at least their representatives, as well as specialists, when designing, discussing and planning the thesaurus. Otherwise, the thesaurus is likely to be perceived as a magicians wand that enables users to find information "as if by magic". This behavior can discourage a user early especially if he/she is confronted with silence or noise. Then he will turn to the librarian... or give up! In this case, it becomes obvious that librarians unwittingly exert some undue power.

Notes

1 Besides, it would be necessary likewise to analyze the cognitive cost of the questioning in natural language.
2 Fiske and Taylor speak of "cognitive miser".
3 See also the n°5 review Didaskalia (1994) called "Les Cartes conceptuelles" [Conceptual Maps].
4 Interesting graphic representations in some fields of knowledge can be found in;


About the metaphor of the mental map, see SERRES, Michel (dir.). Éléments d'histoire des sciences. Paris; Bordas, 1989. Introduction.

References