Dmitrij Dobrovol'skij
Russian Academy of Sciences, Institute of Russian Language

Idioms in a Semantic Network: Towards a New Dictionary–Type

Abstract

In this paper a conception of a new dictionary of idioms is put forward. The proposed model is based on the thesaural principle and intends to reflect the language relevant "naive" picture of the world. To achieve this aim a nonconventional lexicographic approach is necessary, main priorities of which are psychologically acceptable solutions, and not the logical elegance of classification. An idiom–thesaurus of this kind, apart from its theoretical relevance as an instrument of description of the mental lexicon, can be used for purposes of foreign language acquisition and translation.

1. Introduction

The structure of most known dictionaries is based upon the principle "from sign to concept" (for monolingual dictionaries) or "from sign to sign" (for bilingual dictionaries). What we still often miss are dictionaries which are organized "from concept to sign" and can be used for text production. Such a dictionary containing idioms (i.e. an idiom–thesaurus) would make it possible to find all the idiomatic word combinations of a certain language which express the given concept.

The problems discussed in this paper should be situated within the field of both idiom–semantics and onomasiological lexicography. The originality of the proposed approach results not only from the fact that the traditional way of representing idioms in dictionaries ignores, as a rule, the thesaural relations, but also from the advisability of finding nonconventional cognitive based models which are able to reflect the language relevant "natural" picture of the world. The latter seems to be crucial to describing idioms because idioms by nature fix a "naive" pre–scientific view of reality.

Such models have to be predicated not upon a strictly hierarchical taxonomy (cf. traditional thesauri, e.g. Roget 1987; Sanders 1985; Hallig and von Wartburg 1952; Wehrle and Eggers 1954), but rather upon a kind of semantic network. This would allow to explicate not only "vertical", but also "horizontal" relations between lexical units under consideration. Besides, the proposed model tends to reflect some important features of the discourse structure, so the discussed questions can also be seen within the framework of general problems concerning the relationship between lexicon and discourse.

2. Thesaurus as a dictionary–type

Thesaurus is a dictionary presenting linguistic information in the direction "from concept to sign" and explicating relations between the entry–concepts (cf. Marello 1990). The basic entry of a thesaurus, traditionally called taxon, consists therefore of a semantic resp. conceptual marker (descriptor) used as a label of the taxon, and a group of linguistic signs (words and/or idioms) expressing the corresponding concept. In addition, a thesaurus should include a system of auxiliary devices for denoting relations between the concepts.

According to this view, there are two main questions which must be answered before compiling a thesaurus:

(i) At what level of categorization the descriptors have to be looked for?
(ii) How the taxa constructed on the basis of these descriptors have to be linked up with each other?

The answers to these questions can be very different. It depends on linguistic material that is the subject of lexicographic description, and on purposes of this description including a dictionary user's angle as well.

3. What is a basic unit of an idiom–thesaurus?

In every case the thesaurus–taxa must consist of lexical units providing certain semantic resp. conceptual similarity. Its degree and the structure of taxa as a whole depend, however, on the choice of the conceptual scheme fixed by descriptors.

At a first glance it seems that for the compiling idiom–thesauri it is sufficient to borrow the system of descriptors from some well–known thesaurus which contains the whole lexicon of a language with its both components: word–stock and phraseology, e.g. from (Roget 1987). But the analysis of this dictionary and all other reference books of this type shows that it is not as easy as that, because the idioms reveal important special features as compared with the word–stock. So, an adequate conceptual system which can be used as a basis for compiling idiom–thesauri must be developed in an inductive way, i.e. "from bottom to top". At the first stage a certain descriptor is attached to each of the idioms under consideration. The correctness of this conceptual marking is a major prerequisite for the application of the thesaural principle of regrouping the idioms at the second stage.

Looking for adequate descriptors we generally face three possibilities:

(a) One of these possibilities is to take for each idiom one descriptor only (let us call it "the main descriptor") in order to indicate its core meaning. This way proved successful by describing referential lexemes (i.e. words denoting

entities of the "visible world"), but it does not seem to be adequate in our case.

For example, I can indicate the word *oak* as a 'tree', but trying to indicate the German idiom *bei jmdm. ins Fettnäpfchen treten* as a "Verärgerung" ('annoyance') (so in Görner 1979) would be a failure because this label is not sufficient for covering even the core meaning of this idiom: from the viewpoint of the hearer it is 'annoyance', but from the viewpoint of the speaker it is 'offence'.

(b) The second possibility is to look for means of semantic decomposition in the strict, atomistic sense, i.e. to describe the meaning of every idiom as a set of very simple semantic components. But first, since there are no formalizable principles of semantic decomposition, this approach is rather subjective, and second, if we tried to decompose the meaning in really atomic parts, semantic primitives in the sense of Wierzbicka (1972), we would create a dictionary including very general taxa which would have no explanation power in this case, such as 'thing', 'say', 'want', 'good', 'bad', etc.

(c) The third possibility is the orientation towards the basic level of categorization (cf. Rosch 1978). In my opinion, these principles can be applied to non–referential entities as well. The orientation towards the basic level using mostly holistic heuristics allows to find an intuitively acceptable way of semantic labeling that is not as narrow as (a) and not as wide and abstract as (b).

In other words, trying to find adequate markers for units like e.g. the German idiom *jmdm. nach dem Leben trachten*, I neither postulate a set of atomic descriptors such as 'think', 'do' 'something', 'want', 'cause', 'somebody', 'dead', 'bad', nor do I take one specific label such as 'planning a murder' in order to cover the core meaning at once. In this case I rather prefer to put the idiom into two different taxa labeled as (i) 'wishes, intentions, hopes' and (ii) 'killing, murder'. Both of them are "natural" in the sense that they reflect some basic categories of the ordinary language.

Generally, indicating idioms I am against artificial restrictions such as "one idiom – one descriptor" or "one descriptor – one semantic primitive". Wherever it is impossible to find a descriptor covering the core meaning of the idiom in question at once without suggesting ad hoc solutions, to a certain extent the means of semantic decomposition can be used. However, the resulting components have to correlate with concepts of the basic level.

An important feature of idioms (as well as one–word–metaphors and many semantic predicates of other types) is a certain conceptual syncretizm. For example, the German idioms *einen Dachschaden haben, nicht alle Tassen im Schrank haben, nicht (recht) bei Trost sein,* etc. mean not only 'to be mad, crazy', but also depending on context 'to be stupid' and 'to act in an inadequate way'. Since these tendencies have made themselves clearly felt in many cases it seems to be practical to postulate a kind of conceptual field as a basic unit of an idiom–thesaurus. Whenever necessary, the field has to be labeled by a cluster of descriptors, e.g. 'to be mad, crazy, stupid, to act in

an inadequate way' the sum of which constitutes the corresponding syncretic concept (Dobrovol'skij 1992b:190–194).

As far as idioms are concerned, it seems appropriate to put together not only synonyms proper[1], but all idioms belonging to the same conceptual domain without taking into account even their word class identity, so e.g. within the domain 'obstacles, difficulties' such idioms as *der Stein des Anstoßes* (noun), *jmdm. im Weg(e) stehen/sein* (verb), and *mit Ach und Krach* (adverb) can be considered together.[2] The reason for this decision is the semantic vagueness of many idioms, their ability to change the formal parameters depending on the context, and the resulting ineffectiveness of postulating artificially precise distinctions.

A taxon including idioms of the same conceptual field labeled by the relevant (cluster of) descriptor(s) is the basic unit of the thesaural representation of idioms and therefore the main entry–form of the corresponding dictionary. Cf. two German example–taxa: 'success' and 'help' with their conceptual variations.

> *Erfolg, Realisation des geplanten Vorhabens, Erreichen des Ziels mit Geschick und Können (ohne jegliche Verluste)*
> *es zu etw. bringen (im Leben); das Rennen machen; ins Schwarze treffen; auf der Höhe sein; die Kurve kriegen; zum Zuge kommen; etw. auf die Reihe kriegen/bringen; etw. mit dem kleinen Finger machen; etw. mit der linken Hand machen; etw. mit links machen/erledigen; über die Runden kommen; den Nagel auf den Kopf treffen; den Laden schmeißen; zwei Fliegen mit einer Klappe schlagen; etw. in den Griff bekommen/kriegen; etw. im Griff haben; etw. auf die Beine stellen/bringen; etw. über die Bühne bringen; etw. unter Dach und Fach bringen; etw. unter einen Hut bringen; etw. auf einen (gemeinsamen) Nenner bringen; etw. auf den Punkt bringen; etw. in die Finger kriegen; etw. in die Hand/in die Hände bekommen; etw. (ein Hindernis, einen unangenehmen Zustand u.ä.) aus der Welt schaffen; mit jmdm./etw. ins reine kommen; sich (D) keine Blöße geben; eine gute Figur (bei etw.) machen; Fuß fassen.*
> *Hilfe, Eintreten für den anderen, Beschützung des anderen, Gewährung freier Aktionsmöglichkeiten, Sympathiebekundung, Zusammenarbeit*
> *jmdm. zur Hand gehen; jmdm. unter die Arme greifen; jmdm. aus der Patsche helfen; für jmdn. Partei ergreifen; auf jmds. Seite stehen/sein; jmdm. zur Seite stehen; für jmdn. durchs Feuer gehen; für jmdn./etw. die Hand ins Feuer legen; jmdn. unter seine Fittiche nehmen; jmdm./einer Sache nichts in den Weg legen; jmdm. freie Hand lassen; ein offenes Ohr für jmdn./etw. haben; jmdn. auf dem laufenden halten; jmdm. die Daumen drücken/halten; Hand in Hand arbeiten; an einem/am gleichen/am selben Strang ziehen; jmds. rechte Hand; mit Rat und Tat.*

4. Trees vs. networks: semantic networks as a structural basis of idiom–thesauri

The next problem we face is creating a coherent system of concepts in the thesaurus. The real associative links between the idioms reveal such a great diversity that it is practically impossible to explicate all of them in a dictionary. The conventional thesauri are normally based upon hierarchical taxonomies with a tree structure.

Following this principle we can bring forward a certain systematization of the relationship among the taxa of an idiom–thesaurus. Such conceptual fields as, for example, 'blame' (*come down on sb. like a ton of bricks; give/read sb. a lesson; tear sb. to bits; have sb. on the carpet*), 'deceive' (*blow dust in sb.'s eyes; pull the wool over sb.'s eyes; drag a red herring across sb.'s path; tell sb. a story; turn gees into swans*) can be grouped under a concept of a higher rank, i.e. the concept 'have an active non–physical verbal negative influence upon sb.', and so on, until they come under a very abstract notion such as 'act' (cf. Dobrovol'skij 1992a:285–288).

From the cognitive viewpoint this model is rather artificial because it ignores the psychologically real "horizontal" links such as e.g. action–result or beginning–continuation–finishing relations, and concentrates exclusively on "vertical" relations between superordinate and subordinate concepts. A possible alternative is a semantic network.

A dictionary designed according to the network–principle is better suited for exposing the underlying model of the world coded by the idioms. For realization of this project it is not only necessary to implement knowledge–based structures, but first of all to find out the most acceptable mapping from the psychological point of view. It is important to stress that the lexicon as a whole does not have a hierarchical structure.

> "A *rose* is, semantically, a subkind of *flower*, an *oak* a subkind of *tree* [...]; but *inform* is not a subkind of *tell*, *direct* is not a subkind of *order*, and *plead* is not a subkind of *ask (for)*" (Wierzbicka 1987:29).

Many lexical units – among them most idioms – are classified in our minds not only according to the relevant umbrella terms, but also on the basis of the knowledge about their potential co–occurrence in typical situations and corresponding texts, i.e. on the basis of a relevant cognitive structure, e.g. a prototypical scenario.

In general, it can be claimed that our lexical knowledge is organized as a kind of dynamic network which is governed by both the language system and situational interaction reflected in the discourse structure. The mapping of the mental lexicon depends therefore not only on semantic relations such as synonymy, antonymy, hyponymy, etc., but also on our knowledge of how to produce certain texts in relevant situations, and how certain lexical items

have to be embedded in relevant frames. How framelike cognitive structures work by compiling idiom–thesauri can be shown with the help of an example.

Taking e.g. the prototypical scenario ACHIEVING GOALS as a starting point we can connect some thesaurus–taxa with each other in a natural way based on the knowledge of everyday situations. Cf. the following taxa of my thesaurus of German idioms (Dobrovol'skij in press):

(i)   wishes, intentions, hopes: *die Weichen stellen; etw. ins Auge fassen; etw. im Schilde führen; drauf und dran sein, etw. zu tun*, etc.

(ii)   difficult tasks: *eine harte Nuß; ein schwerer Brocken; ein heißes Eisen; ein Buch mit sieben Siegeln*, etc.

(iii)   easy tasks: *eine der leichtesten Übungen; mit Händen zu greifen sein*, etc.

(iv)   good, favorable conditions; luck: *gute Karten haben; Schwein haben; (gut) in Form sein, gut drauf sein; auf der Höhe sein; (wieder) auf dem Damm sein*, etc.

(v)   bad, unfavorable conditions; bad luck: *schlechte Karten haben; Pech haben; schwache Stellen haben; nicht in Form sein; nicht auf der Höhe sein; nicht auf dem Damm sein*, etc.

(vi)   starting activities: *der erste Schritt; die Ärmel hochkrempeln; etw. in die Wege leiten; einer Sache zu Leibe rücken*, etc.

(vii)   obstacles, difficulties: *der Stein des Anstoßes; ein toter Punkt; jmdm. im Weg(e) stehen/sein*, etc.

(viii) overcoming obstacles: *etw. (ein Hindernis, einen unangenehmen Zustand u.ä.) aus der Welt schaffen; über den Berg sein; über die Runden kommen*, etc.

(ix)   giving up, resignation: *die Flinte ins Korn werfen; etw. an den Nagel hängen; Abstand von etw. nehmen; den Kopf hängen lassen*, etc.

These are only a few idiom–groups being relevant for ACHIEVING GOALS, but they are sufficient to illustrate the discussed idea. Between these taxa resp. the idioms belonging to them a sort of socket–plug connection can be postulated which is based on a corresponding prototypical scenario, i.e. on our knowledge that certain wishes, intentions and hopes (i) demand an evaluation of concrete tasks that can be difficult (ii) or easy (iii), and concrete conditions that can be good (iv) or bad (v). If the results of this evaluation are positive, certain activities can be started (vi). While being involved in these activities one can face certain obstacles (vii) which one either overcomes (viii) or does not and finally gives up (ix), and so on.

On the basis of this prototypical scenario a correlating fragment of the network of German idioms can be singled out and exemplified in the following way:

Wenn ich für etwas *Weichen stelle* (i), muß ich mir überlegen, ob diese Aufgabe *eine harte Nuß, ein schwerer Brocken* (ii) oder *eine der leichtesten Übungen* ist, also eine Sache, die ich *mit links* erledigen kann (iii). Ich muß mir auch überlegen, ob ich *gute* (iv) oder *schlechte Karten habe* (v), ob ich *gut drauf* (iv) oder *nicht in Form bin* (v). Wenn ich sehe, daß meine Chancen gut sind, *kremple ich die Ärmel hoch* (vi) und mache mich an die Arbeit. Wenn meine Chancen schlecht sind, wenn ich mit Hindernissen konfrontiert bin (*der Stein des Anstoßes* – vii), nehme ich das Ziel gar nicht in Angriff oder *werfe die Flinte ins Korn* (ix). Wenn ich aber trotz der für mich ungünstigen Situation *die Ohren steifhalte* (viii), *mache* ich doch *den ersten Schritt* (vi).

5. Conclusion

The discussed approach can be used as a model of idiom–representation for compiling dictionaries of different types for needs of foreign language acquisition and translation. The principle "from concept to sign" allows to put into every taxon idioms of different languages, so that multilingual thesauri are possible as well (Dobrovol'skij 1992a). The best practical application of this model is, in my opinion, a kind of computer–assisted dictionary that would allow to search for:

(a) idioms expressing the given concept,
(b) conceptual vicinity of the given concept and the corresponding idioms,
(c) prototypical scenarios describing situations in which these idioms normally appear,
(d) illustration of contextual interaction of idioms based on their "horizontal", non–hierarchical links in the mental lexicon.

Notes

1   Cf. in this connection dictionaries of phraseological synonyms, e.g. (Zhukov et al. (1987); Schemann 1989).
2   For another version of semantic mapping of the lexicon with regard to idioms cf. (Schindler 1993).

References

Dobrovol'skij, D. 1992a. "Phraseological universals: theoretical and applied aspects" in M. Kefer and J. van der Auwera (eds.), *Meaning and grammar. Cross–linguistic perspectives.* Berlin/New York: Mouton de Gruyter.
Dobrovol'skij, D. 1992b. "Phraseologie und sprachliches Weltbild. Vorarbeiten zum Thesaurus der deutschen Idiomatik" in C. Földes (ed.), *Deutsche Phraseologie in Sprachsystem und Sprachverwendung.* Wien: Praesens.
Dobrovol'skij, D. in press. *Kognitive Aspekte der Idiom–Semantik. Studien zum Thesaurus deutscher Idiome.* Tübingen: Narr.
Dornseiff, F. 1970. *Der deutsche Wortschatz nach Sachgruppen.* 7. Aufl. Berlin: de Gruyter.

Görner, H. 1979. *Redensarten. Kleine Idiomatik der deutschen Sprache.* Leipzig: Bibliographisches Institut.

Hallig, R. and von Wartburg, W. 1952. *Begriffssystem als Grundlage für die Lexikographie. Versuch eines Ordnungsschemas.* Berlin: Akademie Verlag.

Marello, C. 1990. "The thesaurus" in F.J. Hausmann, O. Reichmann, H.E. Wiegand and L. Zgusta (eds.), *Wörterbücher, Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie.* Teilbd. 2. Berlin/New York: de Gruyter.

Roget, P.M. 1987. *Thesaurus of English words and phrases.* New edition. London: Longman.

Rosch, E. 1978. "Principles of categorization" in E. Rosch and B.B. Lloyd (eds.), *Cognition and categorization.* Hillsdale NJ: Erlbaum.

Sanders, D. 1985. *Deutscher Sprachschatz.* Nachdrucke der Ausg. Hamburg 1873–1877, 2 Bde. (=Lexicographica. Series Maior. Suppl., 6/7). Tübingen: Niemeyer.

Schemann, H. 1989. *Synonymwörterbuch der deutschen Redensarten.* Unter Mitarb. von Renate Birkenhauer, Straelen: Straelener Manuskripte Verl.

Schindler, W. 1993. "Phraseologismen und Wortfeldtheorie" in P.R. Lutzeier (ed.), *Studien zur Wortfeldtheorie.* Tübingen: Niemeyer.

Wehrle, H. and Eggers, H. 1954. *Deutscher Wortschatz. Ein Wegweiser zum treffenden Ausdruck.* Stuttgart: Klett.

Wierzbicka, A. 1972. *Semantic primitives.* Frankfurt/M.: Athenäum.

Wierzbicka, A. 1987. *English speech act verbs. A semantic dictionary.* Sydney etc.: Academic Press.

Zhukov, V.P., Sidorenko M.I. and Shkljarov V.T. 1987. *Slovar' frazeologicheskich sinonimov russkogo jazyka.* Moskva: Russkij jazyk.