

Ivan Kassabov (Sofia)

## ON THE PROBLEM OF DEFINING THE CORE OF THE VOCABULARY OF THE BULGARIAN LANGUAGE

Ever since the introduction of the notion of lexical system the problem of establishing the basis, or the core of the vocabulary of a given language has been one of the main and still unsolved tasks of lexicology. This is a problem of considerable interest not only for lexicography and foreign language teaching, but from a theoretical point of view as well. For, and it is even more important, no adequate presentation of the systematic character of the lexical inventory of a given language is possible without a preliminary solution of the problem of its core.

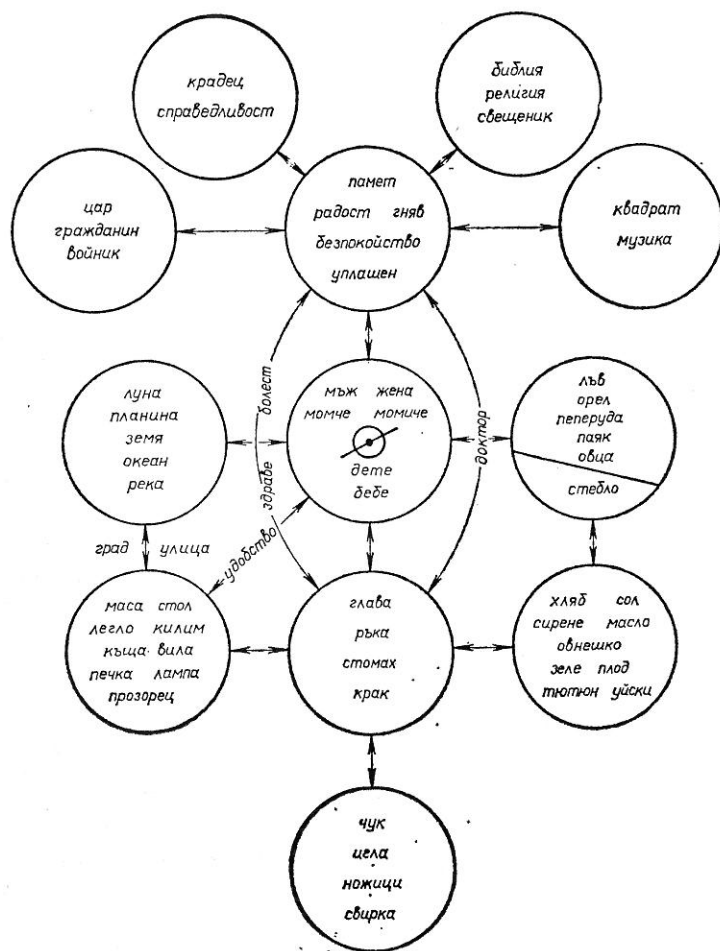
There exist different approaches in the attempts to define this core: a) statistical — used in the frequency dictionaries; b) hierarchial and taxonomic — in the thesaurus type dictionaries like that of Roget, in which thematic and notional fields are used and; c) associative — in the associative dictionaries of the different languages.

It is necessary to point out, however, that these approaches suffer from a grave methodological shortcoming: they all attempt to prove the systematic character of vocabulary not by establishing the inherent principles of its inner organization, but by forcing upon the lexical items the networks of pre-formulated systems. This insufficiency brings about the onesided restrictedness and inadequacy of these approaches, the results of which are often inferior to those achieved in a number of minimum-dictionaries, created intuitively for various practical purposes.

Much bigger adequacy in the presentation of the lexical core characterizes systems such as „Basic English“ and „The List of One Hundred Words“ of Kent and Roussanov. The reliability of the former has been demonstrated by the results of its application in foreign language teaching, while on the basis of the one hundred words, by applying different tests, there have been created associative dictionaries of several languages, Bulgarian included.

Our aim in this paper is to test and prove theoretically the possibility by means of a natural radial enlargement of the 100 words corpus of Kent and Roussanov (in translation) to determine a nucleus of about 800 basic Bulgarian words which, after one more step of enlargement, will amount to about 2000 lexical units, necessary and sufficient to be the core of the Bulgarian lexical system, representing it in its integrity.

In the dictionary of the Bulgarian associative norms these one hundred words seem as if chosen quite arbitrary (particularly in an alphabet order), but nevertheless they could be grouped into several associatived fields. Verbs, pure qualifiers and colour terms, excluded, we could distribute the remaining notional words into 12 associative fields, in accordance with the principle of the anthropocentric character of vocabulary (see Scheme 1).

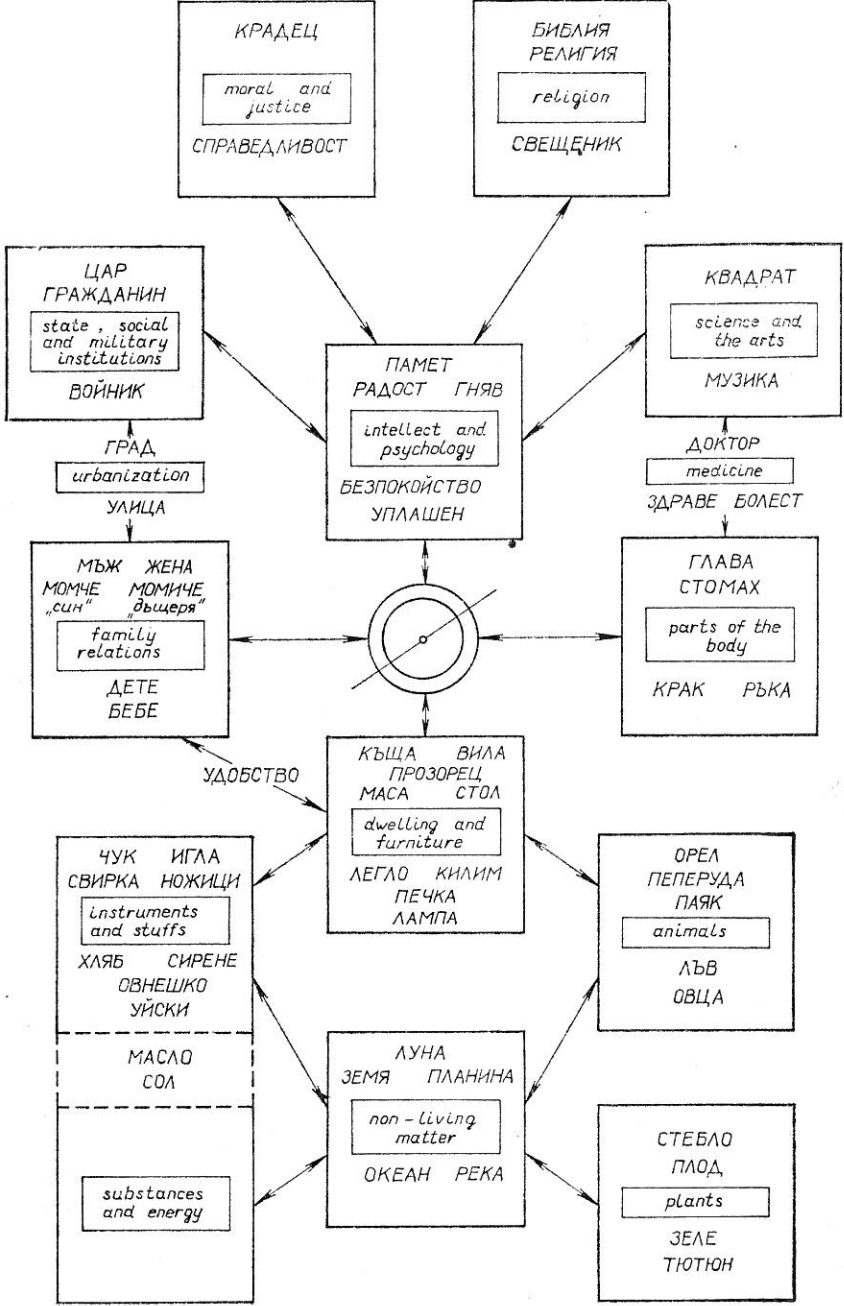


Scheme 1

It is interesting to note that each of these fields in its turn can be supplemented by as much as 50 words-associations, obtained as a reaction to the use of words-stimuli. In this way we get a basic corpus of about 800 words which after one more step of supplementation by applying the same procedure will come up to about 2000 units. This bigger volume will already include verbs, qualifiers and colour terms as well. A schematic presentation of three different fields of the type described above can be seen in Scheme 2.

The preliminary material of the associative fields obtained in this way could (and should) be tested and accordingly corrected by means of the hierarchial taxonomies of the thematic fields of some thematic dictionaries. In this particular case, we have used the Bulgarian part of the Thematic Russian-Bulgarian Dictionary (1961). As a result of its specificity, the organization of the thematic fields, presented in Scheme 3, acquired a configuration somewhat different from that on Scheme 2.

влак — кола — автобус							
УЛИЦА — път							
площад							
столица — ГРАД — село			дом				
квартал	парк						
	алея		жилище				
		блок	КЪЩА	ВИЛА	колиба		
		покрив	под	стая	МАСА	ЛАМПА	ПЕЧКА
		етаж	таван	хол	СТОЛ	светло	топло
		стълба	стена	антре	КИЛИМ		огън
		мазе	ПРОЗОРЕЦ	кабинет	бюро		дим
		двор	врата	спалня	библиотека		пожар
			ключ	кухня	ЛЕГЛО		
				баня	СПЯ (сън)		
				вода	чешма	ЯМ	
				КЪПАНЕ			
добър — лош							
честност		подозрение — вина				вина	
СПРАВЕДЛИВОСТ		престъпник	КРАДЕЦ			убийство	
съд		право				лъжец	
прокурор — адвокат	кодекс		закон	милиция		измама	
свидетел	подсъдим	устав	параграф	следовател		обида	
	присъда		член	ограбен			
	затвор		алинея	арест			
	милост			обиск			
	амнистия			разпит			
			свобода				
дЪЛГ							
	власт						
	държава	знаме					
		герб					
	конституция						
монархия —	република	—	учреждение				
придворен	президент		канцелария	отбрана			
боляр	министър		чиновник	армия			
ЦАР —	ГРАЖДАНИН	—		казарма			
крал	роб	народ		ВОЙНИК — дълг			
император	слуга	права	командир	униформа	война		
фараон		задължения	офицер	шинел	бой		
шах			генерал	мундир	победа		
султан			оръжие		герой		
вожд			пушка		убит		
			пистолет				
			сабя				
			танк				



Scheme III

These two aspects, which require two different types of organization of a given lexical corpus, are not only mutually complementary, but mutually restrictive as well. That is why the lexico-semantic field obtained by following the described procedures contains only such basic words which are necessary and sufficient both for its independent existence and for the integrity of its specificness.

In order to be confident in the reliability of our conclusions, we have considered it necessary to compare the corpus thus obtained with those of different types of minimum-dictionaries. In this way we acquired yet another proof that its scope is both necessary and sufficient.

The procedures described so far, by means of which we have achieved our results, have been theoretically substantiated elsewhere. Their validity is worth noting, since it proves that using a formal corpus (a core), and following a set of strict procedures for enlargement and corrections, it is possible to obtain a stable systematic basis of a vocabulary, which might be used as a point of departure in the creation of a theoretically reliable and adequate minimum-dictionary of the Bulgarian language.