

Svetlana Ter-Minasova, Moscow State University

The Freedom of Word-Combinations and the Compilation of Learners' Combinatory Dictionaries

ABSTRACT: The linguistic problem of freedom and productivity of word combination is closely connected with the lexicographical problem of compiling a Learners' Combinatory Dictionary meant to help students generate speech in its oral and written forms. Recent research in linguistics has convincingly shown that speech formation depends largely on word-combinations: multiword units used as prefabricated blocks. The dictionary presents words in their natural combinations and settings. Free word-combinations, which should be excluded from this type of dictionary turn out to be not so free.

The theory and method of compiling Learners' Combinatory Dictionaries has become a topical and urgent concern of Russian linguistics and pedagogics. The interest is deep-rooted and stems from several social, historical and cultural factors.

The USSR was a multi-lingual state with Russian as lingua-franca and Soviet linguists were committed to prodigious lexicographic activity and to the cultivation of a standard Russian language. They published very many dictionaries for the numerous language groups of the USSR and always paid attention to the development and maintenance of norms of Russian usage.

Another socio-historical reason for the intense interest in "speech-culture", of normalised, standard language can be found in the history of the USSR. After the 1917 revolution, millions of semi-literate peasants and workers suddenly had access to cultural values and the question of "How to speak properly", i.e. according to the norm, came to the forefront: the Elisa Doolittle problem, Russian style.

Today, one of the characteristic features of social life in the former USSR is an unprecedented and ever increasing desire to learn foreign languages, especially English. In Russia and elsewhere people now have many opportunities to communicate with foreigners: partnerships in business and trade on a much greater scale, an ever increasing number of cultural, scientific and sports exchanges, individual travel to and from Russia by private invitation, and so on. It is no wonder that, under such completely changed circumstances, the question: "How can I learn a foreign language quickly and effectively?" has become very urgent.

This question inevitably leads the would-be student to another: "Where can I obtain a good bilingual dictionary?" The average learner does not know enough to ask the really relevant question: "Where can I obtain a Learner's Bilingual Combinatory Dictionary?" – without which it is almost impossible to put two words together in the foreign language outside the context of the course. Indeed, for generations of Russians learning a foreign language, bilingual dictionaries were an inadequate source of information about the active use of the foreign language. They gave little or no help on productive skills, on speaking and writing. They were compiled for a different – recognition – purposes. It was... is... not their "fault" that they were/are used to satisfy all the needs of the foreign learner.

Bilingual dictionaries are meant to "translate" the word from one language into another, to give the word's meaning, to relate the word to a bit of reality in the extra-linguistic world reflected by the word. Users find very little information about the actual usage of the word, about the word's "life" in speech: about which words it goes with, likes, attracts and which words it does not go with, dislikes, repulses. Consequently, learners of a foreign language combine words – orally and in writing – following their mother tongue's collocational patterns, which results in all sorts of errors, ludicrous mistakes, linguistic monstrosities, and so on. Lexical-phraseological combinability or collocability of words was neglected. Using the bilingual dictionary alone, the learner is not made aware of habitual combinations of words in speech in the target language and mastery of the active skills of speaking and writing it is not achieved. That is why the problems of compiling a production-oriented bilingual combinatory dictionary are now at the focus of both scholarly and public attention. It is seen here as a panacea, an all-powerful medicine to cure the ineffectiveness of foreign language teaching methods, a key to the closed doors of foreign countries, those Gardens of Eden full of forbidden fruit.

I have no doubt that a really effective solution to the various problems connected with producing a dictionary of this kind must be sought with the help of linguists, that the practice of lexicography must be based on theoretical studies of language, that theory and practice must go hand in hand. It will be mutually beneficial.

The aim of this paper is to show linguistic grounds for some urgent problems of production oriented lexicography. In order to help language learners to master the active, production skills, it is necessary first of all to understand the "mechanism" of speech-formation, to penetrate the actual process of producing speech.

After many centuries of analysing speech and describing various facts of language possessed by a speaker of the language, 20th Century linguistics has shifted its interest to the question of synthesis: how are these facts put into operation? How are language units brought together in speech? The idea has always seemed to be clear: there are meaningful units of language – the ultimate and most important of them the word – and there are rules and regulations according to which these units are combined when speech is produced.

But, as linguistic research has shown quite convincingly (and the practice of foreign language learning has confirmed it), in actual utterances words do not appear as independent units. They do not form speech events like beads strung upon a thread. They always come up in combination with one another. Thus, in the flow of speech, side by side with monolexemic global units – words – there exist units which in Russian are called *slovosocetanie*, in English translation "word-combination", which, although com-

plex in structure and consisting of formally separable elements, are equivalent to separate words. To put it another way, word-combinations are complex equivalents of words functioning within a sentence in very much the same way as the separate word. To use the well known metaphor, if the output of speech production may be compared to a house, a building or a construction, then the ultimate units of speech – words and word-combinations – may be regarded as building material. Hence, a linguist is concerned with the two aspects, two sides of the problem: how to build and what to build out of; on the one hand, what are the rules and skills of building, and on the other what are the building materials? Linguistics deals with speech construction along two lines: what are the rules for combining ultimate units of speech, and what are these ultimate units of speech, especially those which have been built?

It follows from what has just been said that these complex units, articulated in form but functionally equivalent to a separate word, are bits of “building material” which have already been used in a construction. They are like prefabricated blocks. Being a functional equivalent of a separate word implies, inevitably, a certain degree of globality of nomination. Those units where globality of nomination is expressed in the highest degree, where it actually reigns supreme over the formal separability of elements to such an extent that in their meaning it is impossible to discern a connection with the meanings of the component elements, are called “phraseological” or “idioms”. Phraseological units function in speech in the same way as words. In the process of speaking, one does not merely bring separate words together in linear succession; one also uses “ready made” units, prefabricated blocks (i.e. phraseological units) that already exist in language as a global whole and function in speech as one word.

What, then are the relations between words, phraseological units and word-combinations? Their triple dependence was formulated by academician V. V. Vinogradov (1950, 36ff) in the following way: a word-combination (*slovosočétanie*) should be viewed as a free equivalent of a phraseological unit.

Then the building material out of which speech is produced comprises: 1. the word, 2. the phraseological unit, 3. the word combination as a free equivalent of the phraseological unit and, further – the word.

To sum up, a word-combination as a compound nominative unit, a free equivalent of a phraseological unit, takes part in speech production just like a word.

It goes without saying that the idea of word-combination is based on “gradience”: word-combinations display this or that quality, this or that characteristic to a lesser or greater degree. Some of them are less global, others are idiomatic units, repeated again and again in much the same way as are separate words. Between idiomatic phraseological units and “free” word combinations there come all those innumerable very complex transitional cases which, however different they may be from one separate global single word, are nevertheless functionally its obvious equivalents. (The word “free” is in inverted commas because, as will be shown below, the “freedom” of such word-combinations is relative.) So, if what we call *slovosočétanie* (word-combination, collocation) is viewed as a free equivalent of a phraseological unit, the latter in its turn being regarded as a complex word-equivalent, we no longer think of separate words as combined in speech according to certain syntactic rules.

We concentrate on the all-important problem of correct phraseological usage. This conclusion is part of the answer to the questions about the actual process of speech

production. What does the actual reality of speech consist of? Does it consist of those "surface structures" which the speaker freely generates according to certain rules of grammar and syntax, or does the speaker mainly use ready-made units, word-combinations that exist in language as prefabricated blocks? For a couple of decades, these questions have been at the centre of attention of Moscow University's school of linguistics. The research has been based on the following assumptions: speech-formation, speech-functioning and speech development are determined by the opposition, the dialectical unity of two contradictory factors. On the one hand: creativity, productivity, freedom, unlimited realisation of language potentialities and, on the other hand: non-creativity, non-productivity, fixedness, regular reproducibility, the use of prefabricated, closely bound complex units.

These two main dialectical tendencies of speech-formation are expressed by two different ways of word combination, which results in the dichotomy of complex units, forming two polar cases: productive, free word combinations which are obviously produced by the speaker for this particular speech event *versus* non-productive, tightly bound complexes that are reproduced as global ready-made units. Their correlation, their proportion and their role in speech-formation vary in different spheres of usage, and seem to be closely connected with the main functions of language: **message** and **impact**.

The most difficult problem we have been confronted with is that of finding some regularities and principles which enable us to distinguish between the two above-mentioned kinds of complex formations. The extreme case of "non-freedom" presents, comparatively speaking, not so much difficulty. It comprises all those complexes which may be called, in a broad sense, phraseological units, i.e. those compound word-equivalents which are used in speech as set entities: monemes, idioms proper, traditional set expressions, polylexemic terms, analytical forms, etc.

The notion of a free (or "productive") word-combination is based on the idea of productivity and presupposes the practically unlimited possibility of collocating words. Productivity can be defined as the freedom with which speakers coin new forms. It is very important to clarify the extremely complicated and controversial problem of the freedom of word-combinations, to try to answer the questions: how free and to what extent free are "free" word-combinations (i.e. those which are commonly regarded as "free")? What kind of word-combination may be considered to be free? What are the restrictions imposed on the freedom with which the speaker combines words in the process of speech-formation? These theoretical questions are extremely important for the practical purposes of a bilingual production oriented dictionary.

The authors of the very popular BBI COMBINATORY DICTIONARY OF ENGLISH specify the kind of language material which must be included in a dictionary of this type, drawing a line between collocations (recurrent combinations, fixed combinations) and free combinations which "consist of elements that are joined with the general rules of English syntax and freely allow substitution". (Introduction IX) This distinction is of diagnostic value because "Collocations should be included in dictionaries; free combinations on the other hand, should generally not be included. The inclusion of free combinations is sometimes essential to illustrate a sense of a polysemous entry in a general purpose dictionary". (Introduction IX) Consequently, the problem of "free" versus "non-free" word-combinations is of the utmost importance in the compilation of a combinatory

dictionary, where the word should be given in its most natural settings and contexts - in regular combinations with words that go together habitually and conventionally in normal, standard speech. This kind of dictionary is meant to help the user of the language to produce speech which, as has been shown by linguistic research, consists to a great extent not of separate words joined together in accordance with certain rules, but of prefabricated blocks, whose role is dominant in message-oriented areas of language usage. The legions of LSP learners must be helped to produce speech out of these non-idiomatic, regularly reproduced, firmly established, conventional, safe-to-be-used-by-foreigners word-combinations (collocations).

Investigations into the problem of "freedom" of word-combinations were held along the lines of analysing the restrictions imposed on the speaker in the process of speech formulation. The speaker's freedom is restricted intra-linguistically, i.e. by requirements of the language itself: its syntactic structure, grammatical rules and regularities, which can be presented in the dichotomy of colligation, or morpho-syntactically conditioned combinability of words, and collocation, or lexical-phraseological combinability of words. In the process of combining words in speech, conditioned by these two dialectically united factors: morphosyntactic (colligational) and lexical-phraseological (collocational), it is the latter that is dominant. The general colligational patterns, being more general, more abstract, a kind of formula separated from concrete lexical meanings of words, form as it were the "skeleton" of a speech event, while the lexical-phraseological or collocational factor, being concrete lexical filling of the formula, shapes its "body", makes it real, individual, alive - and difficult.

The main difficulty about the difference in lexical-phraseological combinability of words of different languages - the difficulty which a good combinatory dictionary is supposed to overcome, lies in the fact that, in different languages, words that reflect the same bits of reality, that "mean" the same, live different speech lives. They go together with different words - due to various intra- and extra-linguistic factors. Thus, the English word *book* and its Russian equivalent *kniga* mean the same thing on the level of reality and even on the level of concept (which is not always the case). However, on the level of speech these words are different because they are used in different word-combinations; they live, as it were, different speech lives.

For instance:

telefonnaja kniga - telephone directory

domovaja kniga - house register

kniga otzyvov - record of impressions, etc.

Quite often the most natural, regularly reproduced contexts of the Russian word *kniga* sound quite unnatural for the word *book*: *kniga lučšij podarok* - a book is the best present, *ljubite knigi* - istočnik znanija - love books, they are a source of knowledge, and so on.

The question of "free" word combinations becomes even more problematic if extralinguistic restrictions imposed on the speaker are taken into account.

These restrictions may be heuristically presented by the conceptual and/or sociolinguistic aspects of word-combination. The conceptual basis of word-combination implies the possibility - and naturalness - of bringing together certain notions of extralinguistic reality which underlie the combination of words in speech. The socio-linguistic aspect is

very closely intertwined with the conceptual one. Social structures underlie linguistic structures, and this shatters the very idea of a "free word-combination", which can be conceived only in the abstract, as every normal utterance is always coordinated in time and in space, i.e. restricted by the conditions of communication: who? when? why? where? (Not forgetting peculiarities of the culture, the customs and traditions, etc. of the speech community in question. Thus, theoretically speaking, there are no absolutely free word combinations. Practically speaking those that are relatively free must be sought in fiction, in poetry, where occasionally all the norms and standards, including the natural combinability of things meant are violated for stylistic purposes, for example to evoke some emotional-evaluative reaction from the readers. Word-combinations of this kind are highly connotative, individual, occasional. They are the private property of the author, not the common property of the speech community.

For the practical purposes of a bilingual combinatory dictionary, of all the restrictions imposed on the user of a language it is the lexical-phraseological combinability of words which is the most crucial. The main parameter for recognising word-combinations which should generally not be included is that of variability: "Free combinations ... freely allow substitution" (The BBI Combinatory Dictionary of English, 1986, IX). "A free collocation is one which allows substitution of either or any of its elements without semantic change in the other element or elements. By this criterion "open the gate" is free, as it allows the substitution of "fasten" or "lock" for "open", and "door" and "briefcase" for "gate" (Cowie, 3).

All this is very true. It is true when you look at these problems through the prism of one language. It is true on the level of a monolingual combinatory dictionary. However, the same problem viewed through two languages, on the level of a bilingual dictionary looks different. Many word combinations look deceptively free within their own language and their non-freedom becomes obvious only when they have to be translated into another language. For example, the freest of all combinations of words are those which are combined by the conjunction *and*, because any two nouns or pronouns may be combined in this way. So a collocation like "My wife and I" is free. Substitutes for "my", "wife" and "I" are easily found. But the Russian equivalent, "*My s ženoj*", translates into English as the monstrous looking "We with wife". An English learner of Russian will translate his or her free combination as "*Moja žena i ja*", which immediately shouts "foreigner" to a Russian. On the other hand, a Russian learner of English should be taught to give up all attempts to produce something like "we with wife" and to use "my wife and I" in this very word order, because it is more polite and more English than "I and my wife" – which would be quite normal for a Russian speaker. Another example: "to wash one's hair" allows substitutions. But Russians do not wash their hair! Let me hasten to add that they do wash their heads. The corresponding collocation to "to wash one's hair" is, in Russian: "*pomyt' golovu*" ("to wash head"). These examples could be multiplied indefinitely, but the idea is sufficiently clear. It can be summed up as follows. A combinatory dictionary meant for speech production is most urgently needed by foreign language learners. This kind of dictionary should exclude free, idiosyncratic word-combinations and give only non-free, regularly produced, set, settled collocationally-bound word-combinations. There are far fewer free word-combinations and many more non-free ones than lexicographers who are not well acquainted with linguistic research tend to believe. Speech production looks deceptively free on the collocational

level, and its deceptiveness greatly increases when the language is contrasted with (translated into) another language. It is this particular problem which is a stumbling block, a most difficult obstacle for foreign language learners – all the more difficult because it is hidden, hidden very deeply, not only from foreigners but from native speakers as well. It is this problem which is a weak point of those who teach their mother tongue as a foreign language, and, consequently and surprisingly, a strong point of foreign teachers of a foreign language. The native speakers are not aware of these hidden difficulties. It is *natural* for them to collocate this word with that word because that is the normal, conventional, regular-way of collocating in this particular language. It is taken for granted.

These hidden problems come to the fore and must be taken into consideration when a bilingual combinatory dictionary is being compiled.

Bibliography

The BBI Combinatory Dictionary of English (1986). John Benjamins Publishing Co., Amsterdam/Philadelphia.

COWIE, A.P. Collocational Dictionaries – A Comparative View. University of Leeds. Working Paper.

VINOGRADOV V.V. (1950) Voprosy sintaksisa sovremennogo russkogo jazyka. Moscow.