
Tatar Co-compounds as a Special Type of Classifiers

Alfiya Galieva, Dzhavdet Suleymanov

Research Institute of Applied Semiotics of Tatarstan Academy of Sciences,
Kazan Federal University
e-mail: amgalieva@gmail.com, dvdt.slt@gmail.com

Abstract

The so called co-compounds, or dvandva constructions, exist in many languages as specific type of lexical items, nevertheless they are poorly described in lexicology. The paper analyses the semantics of Tatar co-compounds formed on the basis of names of natural phenomena. The main hypothesis of the study is that co-compounds are a special type of nomination used to designate provisional, non-rigid classes of objects in different situations.

We study co-compounds proceeding from the assumption that there is no strict dichotomy between linguistic and encyclopedic knowledge and that there are no stable word meanings.

Co-compounds as signs which express concepts imply comparing properties of objects and phenomena by means of segregating some common features of these objects. Simultaneously the objects are differentiated by their properties and classified according to their similarities and differences, so co-compounds serve as a special type of classifiers. The analysis of the Tatar language data allows us to conclude that generalization of meanings, distinctive for co-compounds, can help overcome the discreteness of linguistic units and may lead to the overlap of different classes of phenomena with correlated or adjacent concepts.

All the examples are taken from the Tatar National Corpus.

Keywords: co-compounds; the Tatar language; semantics; lexical items

1. Introduction

The constructs formed by concatenating two components are one of the specific features of word-forming and functioning of lexical systems of the Turkic languages. Co-compounds (or dvandva constructions, or pair words) in the Tatar language express collective meaning and they are hyphenated in writing. In this paper we analyse semantic characteristics of co-compounds formed by composing two independent nouns denoting natural phenomena and objects of natural origin.

Co-compounds are studied from typological viewpoint (Wälchli 2005), and special literature is devoted to such composites in particular languages (Ditrich 2006, Bauer 2008, Kiparsky 2009). Main ways of formation, the structure and some aspects of semantics of co-compounds are studied and described in Tatar grammars (Tatar Grammar, 1993) and in special works (Mirgaleyev 2011, Safiullina 2009). Nevertheless many issues concerning the meaning and functioning of such constructs have not been studied yet. Co-compounds are frequently used in Tatar spoken language, in folklore and fiction texts, however, being “friable” in structure, compounds of this type are irregularly fixed in dictionaries, so we are to examine the corpus data that reflect actual types, functioning and distribution of linguistic items, including compounds, in their natural surroundings. All the examples presented here (compound words and text fragments) are taken from the “Tugan Tel” Tatar National Corpus (<http://corpus.antat.ru>). The Tatar National Corpus includes all sorts of writings from literary novels and popular scientific literature and educational texts to daily newspapers and magazines, texts of Internet publications on informative, social and political topics and official documents (Suleymanov et al. 2013).

We study the semantics of Tatar co-compounds taking into account the approaches in contemporary cognitive linguistics. The principles of cognitive linguistics are based on the belief in the contextual, pragmatic flexibility of meaning, the conviction that meaning is a cognitive phenomenon that exceeds the boundaries of the word, and the idea that meaning involves perspectivization (Geeraerts

2010: 182).

We also study the semantics of co-compounds, proceeding from the assumption that there is no strict dichotomy between linguistic and encyclopaedic meanings and that there are no stable word meanings. Rather, they are dynamic, context-sensitive and construed on-line (Paradis 2003, 2012), and the meaning is not a function of language per se, but arises from its use (Evans 2006).

The paper is organized as follows: Section 2 presents the relations of components of co-compounds from the semantic viewpoint, Section 3 examines co-compounds as special type of classifiers, Section 4 attempts to describe co-compounds as a means to overcome discreteness of lexical items.

2. The Relation of Component Meanings of Co-compounds

This paper focuses on co-compounds formed according to the model *noun + noun*, where nouns designate natural phenomena and objects of natural origin. The lexical class of nouns that denote phenomena of the nature (including names of objects of natural origin, atmospheric phenomena, names of landscape parts, animals and plants, etc.) is characterised by multidimensionality of semantics due to ontological (reference of nouns) and cognitive (various parameters of categorization and discretization of meaning) causes; in addition, names of natural phenomena have certain strongly marked language specific features.

Co-compounds usually contain constituents that have meaning components which are common on several grounds and may be described using different axes of structural organization:

1) synonymous relationship:

qır-basu (field + field) – ‘treeless plain’;

2) hypernymic relationship:

yılga-sulıq (river + water reservoir) - ‘water reservoirs, water bodies’;

3) equonimic relationship (components of the pair word are co-hyponyms to the same hypernym, in which case the co-compound designate the generic notion):

yılga-ineş (river + brook) - ‘water reservoirs of running water’,

yılga-kül (river + lake) – ‘water reservoirs’,

qar-yañgır (snow + rain) - ‘atmospheric precipitations’,

agaç-quaq (tree + bush) - ‘arboreous vegetation’, *ay-yoldız* (moon + star), *ay-qoyaş* (moon + sun) – ‘luminaries’.

Components may be combined by contiguity of observed phenomena or by associative links in the speaker’s mind, thus coining co-compounds is based upon knowledge learned from the external world:

a) cooccurrence of a natural phenomenon (object) and substance: *qar-su* (snow + water), *su-yañgır* (water + rain), *qiya-taş* (rock + stone); the word denoting some substance may be the first or the second component of the compound;

b) spacial and temporal adjacency of phenomena and objects (sometimes including causal relationships between phenomena): *yäşen-yañgır* (lightning + rain), *yañgır-boz* (rain + hail), *yañgır-bolut* (rain + cloud);

c) presence of a common semantic component in the parts of the paired word: *yañgır-saz* (rain + swamp) - the shared seme ‘moisture’.

Substantial amount of contexts extracted from the Tatar National Corpus contain co-compounds in plural, so the outstanding characteristic of co-compounds - collective meaning, is often supplemented by the idea of articulated multiplicity (the affix of plurality *-lar/-lär*):

(1) *ineş-yılgalar* (brook + rivers),

(2) *yılga-küllär* (river + lakes).

Sometimes the affix expressing plurality does not join the pair word but is used in the context:

(3) *yılga-ineş buy-lar-inda* river-brook length -PL, POSS_3, LOC

'along rivers and brooks'.

The meaning of a co-compound refers to the collection of its constituent members conceived as a whole. The constituents of co-compounds express semantically closely associated concepts which may be on the same or on different hierarchical levels, and the meaning of the whole is more general than the meaning of the parts.

3. Co-compounds as Classifiers

An ordinary (one-word) collective noun denotes a group composed of multiple members, identical or similar in any way, and taken as a whole (*foliage* 'the leaves of plant', *gentry* 'minor landed aristocrats'). Languages may have special indices for denoting collectivity meaning, for example Russian suffixes *-y-* (*zver-y-o* 'wild beasts'), *-stv-* (*yunoshe-stv-o* 'young people'). The Tatar language has no special affix to represent collective meaning, and the idea of collectivity can be expressed by means of some polysemous affixes, for example, *-lıq/-lek*:

- (4) *naratlıq* 'pinery', derived from *narat* 'pine',
- (5) *qayınlıq* 'birch wood', derived from *qayın* 'birch',
- (6) *agaçlıq* 'grove, small wood', from *agaç* 'tree'.

However we should specify that in (4-6) collective nouns denote no mere collection of similar objects (trees) but also the place which is characterized by a large number of objects named by the word stem.

Any common name has a notional meaning component representing in human mind essential properties of objects and connections between objects or phenomena, so empirical objects are selected and combined in a certain class according to a set of common characteristics and those which are specific for them. The selection of a verbal sign is based on such mental processes as comparison, analysis and synthesis, abstraction, and generalization.

The meaning of the Turkic paired word is more complex and wider than the meaning of the collective noun formed by adding an affix to one-component stem (examples like (4-6)), when the concept with collective meaning denotes a clearly defined set of identical objects.

Co-compounds as signs for representing concepts suppose comparing properties of the objects and phenomena named by their constituents, as well as defining certain common properties, thus distinguishing objects by different criteria and classifying them by similarities and differences; so co-compounds form various "temporary" or provisional (i.e. necessary in every given case and concerning demanded referents) classes.

Let us illustrate it with examples of co-compounds of different types.

A pair word may serve as a sign for denoting a group of different types of leaves for which the Tatar language has special words:

- (7) *yafraq-ilis* (leaf + conifer needles);
- (8) *yafraq-qiyak* (leaf +sharp long leaf, specific for some plants such as leek, cane or rush).

In cases (7-8) a pair word denotes a new class of objects having leaves of one kind or another.

The *yafraq* component often forms co-compounds with another component denoting parts of plants:

- (9) *yafraq-tamır* (leaf + root):

<i>üş</i>	<i>agaçı,</i>	<i>töp</i>	<i>käüsäse,</i>	<i>yafraq-tamırları</i>	<i>bula</i>
his	tree-POSS_3	main	trunk-POSS_3	leaf+root -PL	be-3
'has his own tree, main trunk, leaves and roots'.					

In (9) the pair word *yafraq-tamır* occurs in the context of the word *käüsä* (trunk of tree), so the tree is conceptually divided into two parts: trunk as the main, vertical hard part, formed of wood, and leaves and roots, as smaller and multiple objects.

- (10) *yafraq-botaq* (leaf + branch):

<i>agaçlar</i>	<i>tup-turi,</i>	<i>şäp-şärä</i>	<i>käüsäle,</i>	<i>bernindi</i>	<i>yafraq-botaglari</i>	<i>ıok</i>
tree-PL	very	very bare	trunk-MUN	any	leaf	branch-PL, no POSS_3
	straight					

'trees are very straight, bare, without any leaves and branches'.

(11) *botaq-yafraq* (branch + leaf):

<i>işelep</i>	<i>töşkän</i>	<i>botaq-yafraqlar</i>	<i>arasında</i>
wreck-CONV	fall- PARTC_PAST	leaf+branch-PL	among

'among fallen leaves and branches'.

In these phrases co-compounds with components *leaf* and *branch* designate different types of objects: In (10) small multiple overground parts of the tree in contrast to trunk as the single, big, hard, non-bracing and relatively smooth part; in (11) overground parts of the tree (leaves and branches) that may easily be separated from the rest of the tree.

(12) *yafraq-ülän* (leaf + grass):

<i>yafraq-ülänndärän</i>	<i>tözelgrän</i>	<i>öy</i>
leaf+grass-PL, ABL	build- PASS, PARTC_PAST	house

'house made of leaves and grass'.

In (12) combined into one class are the external organs of a plant as soft green plate and plants with narrow leaves growing from the base. Such a combination can be easily explained by the fact that leaves and grasses have significant similarities in their appearance and in use by people.

Semantics of many co-compounds is not inflexibly outlined; maintaining a general meaning of collectivity, they allow the speaker to vary the scope of class of named objects.

The flexibility of meaning of co-compounds is particularly evident in those cases where the second component does not represent any definite concept, but is only phonetically consonant to the first component (the so called "echo" word):

(13) *yafraq-çapraq* (leaf + unmeaning "echo" word):

<i>yulındagi</i>	<i>kori</i>	<i>yafraq-çapraqni</i>	<i>hawaga</i>	<i>kütärä</i>
road- POSS_3, LOC	dry	leaf+"echo" word-ACC	air-DIR	raise-3

'[Tempest] raises dry leaves [and other things] from the road into air'.

In cases like (13) the speaker or hearer are to fill the pair words content up according to their understanding of the world, and interpretation depends on the situation described and the background knowledge of the speakers.

As we mentioned above, in many cases co-compounds are not fixed in dictionaries. Easy formation of such constructs in speech enables words (concepts) to denote various classes containing diverse units depending on components. So co-compounds as a result of transformation of existing lexical concepts can form new, more abstract concepts, without inventing new words, i.e. using existing words in the most efficient way.

It should be noted that in the contemporary Tatar language there are many abstract (higher level) terms, nevertheless a large number of such terms are a product of development of scientific knowledge and of the language of science, and speakers always sense some artificiality of such stylistically marked bookish words as *sulıq* 'water body', *ciganaq* 'natural fountain, source' and so on. Co-compounds do not refer to the language of science, and they have been used for centuries for denoting generic concepts of various types regardless of the educational level of language users.

So the capacity for constructing co-compounds enables one to create new classes of objects by using linguistic resources economically, without contriving new terms, and thus to form new classes of objects required in different situations, as well as non-rigid classes with flexible borders for different types of objects, at the same time keeping the openness and versatility of the lexical and

semantic system of the language.

4. Co-compounds as a Means to Achieve Discreteness of Lexical Items

As it is generally known, one of basic characteristics of human speech is an ability to create discrete units at all levels, dissection and articulation of signifier as signified. V. Zvegintsev, a well-known Russian linguist, defines language as “a system of discrete units whose content is determined by the fact that the process of sampling is realised in relation to the phenomena of the ‘outside world’, in the broad sense of the word” (Zvegintsev 1973: 217).

The analysis of the Tatar language data allows us to conclude that generalisation of meanings, distinctive for co-compounds, can help overcome the discreteness of linguistic units and may lead to the overlap of different classes of phenomena with correlated or adjacent concepts.

In formation of co-compounds their components are acting as pre-constructs - units that are ready to use, but at the same time capable of combining with other words corresponding them by some parameters. In the process of speech, some components of the meaning or associative links are emphasised. This resembles the technique of *bricolage* in the works of C. Levi-Strauss (Levi-Strauss 1966), a French anthropologist. The term means construction using whatever was available at the time, and C. Levi-Strauss uses it for describing the characteristic patterns of mythical thought. In his description, *bricolage* is opposed to the contemporary engineers' creative thinking, which proceeds from goals to means. Mythical thought, according to Levi-Strauss, attempts to re-use available implements in order to solve new problems. The *bricoleur* is adept at performing a large number of diverse tasks and at putting preexisting things together in new ways, adapting his or her project to a finite set of materials and tools, but this universe of implements is limited. The set of the 'bricoleur's' means cannot therefore be defined in terms of a certain project but it may be defined by its potential use, because the elements are collected or retained on the principle that 'they may always come in handy' (Levi-Strauss 1966).

The Tatar co-compounds function in speech the same way, they represent a set of actual and possible relations; so they are 'operators' but they can be used for any operation of a certain type. A component of the pair word serves as an intermediary between a low level concept (the constituent of the pair word itself) and a higher level concept (the whole compound). This component maintains its specific meaning, but is also capable of expressing more generalized ideas, and the basis for classification, selecting the head characteristic is different every time; the capacity of components of co-compounds to join words of diverse thematic classes allows one to "remove", "dissolve" or neutralize the boundaries of these classes.

Here is a list of co-compounds containing constituents designating atmospheric phenomena and precipitations; the whole co-compound enables to distinguish different types of atmospheric phenomena and precipitations depending on the situation:

- (14) *qar-su* (snow+water);
- (15) *qoyaş-yaygır* (sun+rain);
- (16) *su-yaygır* (water+rain);
- (17) *yäşen-yaygır* (lightning+rain);
- (18) *yaygır-boz* (rain+hail);
- (19) *yaygır-bolit* (rain+cloud);
- (20) *yaygır-buran* (rain+snowstorm);
- (21) *yaygır-suiq* (rain+frost);
- (22) *yaygır-yawim* (rain+precipitation);
- (23) *yaygır-cil* (rain+wind);
- (24) *cil-buran* (wind+snowstorm);
- (25) *cil-dawıl* (wind+ storm);
- (26) *cil-qoyaş* (wind+sun);

(27) *cil-yangır* (wind+rain).

Here are some examples of co-compounds containing component *yilga* 'river', and their meanings (description of the meanings is approximate):

- (28) *yilga-ineş* (river+ brook) - 'water bodies with running water';
- (29) *yilga-kül* (river+lake) - 'water bodies';
- (30) *yilga-taw* (river+ mountain) - 'the most important elements of the landscape';
- (31) *yilga-çoqır* (river+ pit) - 'land forms';
- (32) *yilga-ärämä* (river+forest of thick bushes) - 'a type of landscape';
- (33) *yilga-öyänke* (river+ white willow) - 'a special type of landscape beside the river'
- (34) *yilga-kiçü* (river+ forwards) - 'crossing a river (by water and wading)';
- (35) *yilga-yul* (river+ road) - 'landscape elements having a great extension';
- (36) *yilga-kanal* (river+ canal) - 'water bodies having a great extension, of natural or artificial origin'.

These examples show that the word *yilga* easily joins other words denoting water bodies (28, 29), characteristics of locality (30, 31), certain plant species (32-33) and artefacts (35, 36). They form a compound, and its whole meaning is determined by the area of overlap of its component meanings.

Let us also consider the co-compounds containing the first component *su* (water):

- (37) *su-ineş* (water + brook), *su-därya* (water + wide river) - 'rivers of all types';
- (38) *su-yangır* (water+ rain) - 'atmospheric precipitations';
- (39) *su-ut* (water+ fire) both - 'elements' and 'dangers";
- (40) *su-bolin* (water +meadow) - 'lands near the water, including water body';
- (41) *su-qamış* (water+ cane) -'characteristic elements of a water landscape';
- (42) *su-rızık* (water+ food), *su-azıq* (water + food) - 'food and drink as basic human needs';
- (43) *su-çay* (water+ tea) - 'beverages'.

So the component *su* joins words denoting streams (37), atmospheric precipitations (38), certain plants (41), food (42) and drinks (43), etc.

The interpretation of co-compounds named above is very rough, so the meaning of compound words is determined by the context. Code switching in any case is promoted by parts of co-compounds (features of its components) as well as by the specifics of the context and situation in which this composite functions.

In the formation of co-compounds denoting collection of some entities we see an attraction not to one central concept (monolexeme term), but to two central concepts, so the verbal sign, having collective meaning, remains bipolar, bifocal. In particular, each of its components may have its own affixes of cases or plurality:

(44) *yilgalarnı-küllärne*

<i>Yangır</i>	<i>suları</i>	<i>yilgalarnı-küllärne</i>	<i>basıp</i>	<i>kitär</i>
Rain	water-PL, POSS_3	river-PL, ACC, lake-PL, ACC	press-CONV	go away-FUT, 3

'Rain waters will flow into rivers and lakes'.

This property shows the grammatical independence of components of such constructs.

5. Conclusion

The paper represents only Tatar co-compounds formed by composing nouns denoting natural phenomena, although the analysis of other semantic classes of co-compounds leads to similar conclusions.

The meaning of a co-compound refers to the collection of its constituent members conceived as a

whole. The constituents of co-compounds express semantically closely associated concepts which may be on the same (usually) or different (on rare occasions) hierarchical levels. Co-compounds may combine two words denoting objects or phenomena, distinct or heterogeneous by their nature, so that a pair word has a more complicated meaning than the simple indication of collectivity. Formation of such constructs is based on encyclopaedic knowledge of the nature of things, on relations between various fragments of the world, and on comprehension of analogies and causal relations by human mind.

Nouns join to form compounds, and the meaning of the whole is defined by the area of overlapping component meanings. It is hard to describe the meaning of many co-compounds formed by way of joining two nouns, because this meaning is flexible and highly dependent on the context. In many cases composing components result in a new concept for designating a new class of objects. The openness of the list of co-compounds allows one to remove the boundaries of classes of objects and to form new concepts. Free variation of meaning of co-compounds within a wide area allows softening the discreteness of the world which is fixed in linguistic units.

In creating co-compounds with collective meaning a new concept may be formed around two central concepts so that the new unit remains bipolar, and each of its components may attach its own inflectional affixes.

6 References

- Bauer L. (2008). Dvandva. In *Word structure, I* (1), pp. 1-20.
- Ditrich T. (2006). Dvandva Compounds and the Chronology of the Rgveda. In *Crossroads*, pp. 26-35.
- Evans V. (2006). Lexical Concepts, Cognitive Models and Meaning-Construction. In *Cognitive Linguistics 17-4*, pp. 491–534.
- Geeraerts D. (2010). *Theories of lexical semantics*. Oxford University Press, Oxford.
- Kiparsky P. (2009). Verbal Co-compounds and Subcompounds in Greek. In *MIT working papers in linguistics: MIT workshop on Greek syntax and semantics*.
- Lévi-Strauss C. (1966). *The Savage Mind*. University of Chicago Press, Chicago.
- Mirgaleev R. (2011). On Issue of Semantics of Co-compounds in the Tatar Language [K voprosu o semantike parnykh slov v tatarskom yazyke]. In *Proceedings of the International Turkological Conference “Word Formation in Turkic Languages: Studies and Problems”* [“Slovoobrazovaniye v turkskikh yazykakh: issledovaniya i problemy”: materialy Mezhdunarodnoy t'urkologicheskoy konferentsii], Kazan, pp. 300-302.
- Paradis, C. (2003). Is the Notion of Linguistic Competence Relevant in Cognitive Linguistics. In *Annual Review of Cognitive Linguistics*, pp. 247-271.
- Paradis C. (2012). Lexical Semantics. In *The Encyclopedia of Applied Linguistics*. Oxford.
- Safiullina F. (2009). Pair Words in the Tatar Language [Tatar telendä parlı süzlär], *Problems of saving languages and cultures in the context of globalization* [Problemy sokhraneniya yazyka i kultury v usloviyakh globalizatsii]. Kazan, pp. 203-205.
- Suleymanov D., Nevezorova O., Gatiatullin A., Gilmullin R. & Khakimov B. (2013). National Corpus of the Tatar Language “Tugan Tel”: Grammatical Annotation and Implementation. In *Procedia - Social and Behavioral Sciences* (2013), pp. 68-74.
- Tatar Grammar [Tatarskaya grammatika] (1993). Tatar Book Publishers, Kazan, Vol. 1.
- “Tugan Tel” Tatar National Corpus. Accessed at: <http://corpus.antat.ru> [30/04/2014].
- Wälchli B. (2005). *Co-compounds and Natural Coordination*. Oxford: Oxford University Press.
- Zvegitsev V.A. (1973). *Language and Linguistic Theory* [Yazyk i lingvisticheskaya teoriya], Moscow University Press, Moscow.