

Daniela Majchráková

Ľ. Štúr Institute of Linguistics, Slovak Academy of Sciences, Bratislava

LEXICOGRAPHICAL TREATMENT OF SLOVAK ADJECTIVE-NOUN COLLOCATIONS

Abstract: In the article, we focus on the conceptual treatment of adjective-noun collocations in the *Dictionary of Slovak Collocations. Adjectives*. We deal with the analysis of the formal and semantic aspects of collocations, especially those collocations that lie on the border between fixed and free word combinations. We evaluate the formal side of collocations from the perspective of their formal preference and observe how it is reflected in their changed word-class features. We analyse the semantic side of the chosen collocations using their restrictions in collocability. We demonstrate that the formal and collocability preference of words reflects itself in their statistical characteristics, which is why we consider the preference an important lexicographical criterion for choosing collocation candidates.

Keywords: collocation, collocate, collocability, statistics, paradigm

1. Introduction

In general, collocation dictionaries provide evidence of the combinatorial potential of words, but primarily they serve the users as a tool to form correct word collocations. In order to create collocation dictionaries, lexicographers use an immense amount of corpus data and statistical tools, which help them to identify the collocation candidates. Although the frequency of words and the related statistics are the first step in identifying collocations, not all of them can be automatically included in a dictionary. Therefore, deciding which criteria must be applied is no easy task for lexicographers.

There are two types of the *Dictionary of Slovak Collocations* – a dictionary of noun collocations and a dictionary of adjective collocations. The main aim of these dictionaries is to create collocational profiles of words – both nouns and adjectives (see Ďurčo 2007). A collocational profile represents all the collocates which co-occur with the entry word. The distinction between these two versions lies in the definition of a collocation. In the *Dictionary of Slovak Collocations. Nouns*, we defined a collocation mainly through its statistical properties. In this approach, a collocation was defined as a statistically significant combination of words (see Majchráková and Ďurčo 2010; SSS-PM 2015). Here, collocates of nouns were primarily collected according to their frequency distribution and statistical features. This determined the extent (length) of the collocational profiles of the noun entries. On the other hand, the *Dictionary of Slovak Collocations. Adjectives* has more restricted criteria for distinguishing the so-called “true” collocations from common (non-significant) sets of words.

In this paper, we focus on adjective-noun collocations the way they are treated in the *Dictionary of Slovak Collocations. Adjectives*, and in particular on

those which are situated on the very narrow border between fixed and non-fixed combinations of words. For these purposes, we will analyse their morphosyntactic, semantic and pragmatic properties. We will also show that statistical significance can help to identify a collocation as a lexicographical unit.

2. The morpho-syntactic properties of collocations

One of the characteristic features of collocations which allow us to record them in the dictionary are the morphosyntactic preferences of their components. This is obvious for combinations that are evidently fixed – mainly for phrasemes and idioms, which are often petrified syntactically and/or morphologically.

On the other hand, there are combinations of words which traditional Slovak phraseology did not consider as phraseological units, mainly due to their literal meaning. These collocations typically occur in texts in a preferred grammatical form that can also be proved statistically. For example, the declension paradigm of the collocation *v plnej paráde* (*in full regalia*) is disrupted; the basic nominal form of the collocation simply does not exist and other cases cannot be formed either, e.g. **plná paráda* (**full regalia*), **plnou parádou* (**with full regalia*).

In addition to case preference, in adjective-noun collocations we can observe an absence of the grammatical category of comparison in the adjectival component, cf. *tmavé pečivo* – **tmavšie pečivo* (*dark bread* – **darker bread*). In fixed combinations, the disruption of the form paradigm is caused by the transposed meaning of the adjective. For example, in the collocation *tmavé pečivo* (*dark bread*), the adjective has the additional semantic component of “wholemeal”. Even in the case of non-phraseological collocations, we regard the inability of the adjective to change its grammatical form as a manifestation of its gradual lexicalisation, cf. *zlá predtucha* – **horšia predtucha* (*a gloomy foreboding* – **a gloomier foreboding*), *do najmenších detailov* – **do malých detailov* (*to the slightest detail* – **to a slight detail*).

During the process of lexicalisation, one form of the lexical combination sets itself apart and becomes an element which exists independently in the lexical system and has no potential for paradigmatic variations. On this level, a disruption of the paradigm occurs, which also brings about a change in the functional validity of the collocation. For example, in the syntagma *každý prípad* (*any case*), which was originally a free combination, the form *v každom prípade* (*in any case*) has left the morphological paradigm. This prepositional form became fixed when it started functioning as a multi-word particle. This way, various forms of the free combination *každý prípad* (*any case*) can coexist with the given prepositional fixed form (cf. *s čistým svedomím* – *čisté svedomie* (*with a clear conscience* – *a clear conscience*), *v plnom prúde* – *plný prúd* (*in full swing* – *full swing*), *na plný plyn* – *plný plyn* (*at full throttle* – *full throttle*)). These include the identical form with the literal meaning *v každom prípade* (*in any case*), which can also be identified using modifications such as *v každom jednom prípade* (*in each and every case*), *v každom tom prípade* (*in each of those cases*), etc.

Hence, when the paradigm of adjective-noun collocations gets disrupted, their functional (i.e. word-class) features are changed. Some lexicalised forms display functional and sometimes even semantic correspondence with one-word adverbial equivalents, cf. *v nedávnom čase – nedávno* (*in recent times – recently*), *do značnej miery – značne* (*to a significant degree – significantly*), *spoločnými silami – spoločne* (*with joint forces – jointly*), *v najhoršom prípade – prinajhoršom* (*in the worst case – (at) worst*), *v najlepšom prípade – prinajlepšom* (*in the best case – (at) best*). These are lexicalised analytical forms that function as adverbs and are composed of an adjective and a noun which is primarily abstract. From the standpoint of expressivity, they are neutral. The difference between a one-word adverb and its multi-word counterpart lies in their stylistic function, which also manifests itself in their preferential contextual occurrence – analytical forms are official and explicit, thus they are typical for technical, journalistic and administrative texts, cf. collocations with a temporal meaning used primarily in the journalistic style: *v ranných hodinách – ráno* (*in morning hours – (in the) morning*), *vo večerných hodinách – večer* (*in evening hours – (in the) evening*). This characteristic of theirs is also underlined by their lexical composition – analytical forms are formed exclusively using abstract nouns.

From the perspective of word-class features, we can categorize non-phraseological adjective-noun collocations into:

– collocations functioning as adverbs expressing various adverbial meanings:

time: *v dohľadnom čase* (*in the foreseeable future*), *v blízkom čase* (“*in the near time*” – *in the near future*), *v nedávnom čase* (*in recent time*), *v blízkej budúcnosti* (*in the near future*), *v ranných hodinách* (*in morning hours*), *vo večerných hodinách* (*in evening hours*), *hodnú chvíľu* (*for a long time*), *na poslednú chvíľu* (*at the last minute*);

manner: *na vlastné riziko* (*at one’s own risk*), *hrubou silou* (*by brute force*), *spoločnými silami* (*with joint forces*), *v nehybnom stave* (*in a stationary position*), *rovnakým spôsobom* (*the same way*), *právnou cestou* (*by legal means*), *po zrelej úvahe* (*after mature inquiry*), *z dlhodobého hľadiska* (*in the long run*);

reason: *cudzím pričinením* (*by the act of somebody else*), *vlastným pričinením* (*in one’s own right*), *cudzím zavinením* (*through the fault of somebody else*), *vlastným zavinením* (*through one’s own fault*);

place: *v tesnej blízkosti* (*in close proximity*);

degree: *v plnom rozsahu* (*in the full extent*), *do značnej miery* (*to a great extent*), *v nemalej miere* (*in no small way*);

purpose: *vo verejnom záujme* (*in the public interest*), *vo vlastnom záujme* (*in one’s own interest*).

– collocations functioning as particles:

with an explicatory meaning: *v každom prípade* (*in any case*), *v konečnom dôsledku* (*in the end / ultimately*), *v neposlednom rade* (*last but not least*), *na druhej strane* (*on the other hand*), *svojím spôsobom* (*in its own way*);

with an evaluating meaning: *v najlepšom prípade* (*in the best case*), *v najhoršom prípade* (*in the worst case*) (categorisation according to MSJ 1966: 750).

– a collocation functioning as a coordinating conjunction with an explicatory meaning:

v opačnom prípade (otherwise).

The formal preference of a collocation can have different forms. Collocations which lie on the border between being fixed and free have a partially limited form paradigm, which is related to their contextual occurrence. Thus, the extent of the paradigmatic anomaly is governed by textual usage, i.e. the way the collocation is included in the context. From a formal perspective, the so-called typical or usage collocations can appear as unmarked units. However, their significance lies in their tendency to frequently occur only in the form of a specific grammatical case. A typical example is the significantly more frequent plural form of certain collocations, e.g. *plané sľuby (empty promises)*, *snehové jazyky* (“*snow tongues*“ – *drifting snow*), *problémové partie (problem areas)*. The prepositional construction *v zrelom veku (in middle age)* occurs in texts preferentially (447 occurrences), as opposed to its nominative form *zrelý vek (middle age, 63 occurrences in Slovak National Corpus)*. In some cases, this tendency can signalize a shift from the area of typical combinations to the area of lexicalised combinations, cf. *plná rýchlosť (full speed, occurrences)* and *v plnej rýchlosti (at full speed, occurrences)*

It is often the case that formal preference is connected to a typical (sometimes exclusive) syntactic position in the sentence, which does not allow the lexical combination to realize any other paradigmatic forms. Hence, the regularity of its morphological paradigm is merely hypothetical. Thus, the inability to create certain forms can be possibly linked to the typical position of the object in a sentence or to participation in a predicate nominal, e.g. *bohapustá lož: to je (an outright lie: it is)*, *čistá hlava / čistú hlavu: mať / s čistou hlavou (a clear head / a clear head: to have / with a clear head)*, *nemý úžas / v nemom úžase: byť (blank bewilderment / in blank bewilderment: to be)*; *pozorovať niečo / sledovať niečo (observe something / watch something)*, *s nemým úžasom: stáť (with blank bewilderment: stand)*. In this connection, J. Mlacek (2001: 107) talks about the tendency to prefer a certain morphological form.

3. Anomalous collocability

To the greatest extent, the defective collocability found in the components of a combination manifests itself as a type of semantic incompatibility where no synthesis of meaning occurs, as is the case in ordinary combinations, cf. the phrasemes *salámová metóda* (“*salami method*“ – *salami tactics*), *živá mŕtvola* (“*living corpse*“ – *zombie*) or *medové týždne* (“*honey weeks*“ – *honeymoon*).

Collocational anomalousness is measured by the extent of the collocational paradigm – the smaller the number of components in the collocational paradigm, the greater the anomalousness (cf. Čermák 1985). Its borderline case is monoccollocability of one of the components found in a combination. A limited collocational paradigm is a consequence of the semantic uniqueness of a word. For

such a paradigm, the compatibility of its semantic components with those of other lexical units is limited to mostly one case or three cases at most, e.g. *pitná: voda, prameň, režim* (“drinking”: (potable) water, spring (with drinking water), (liquid) intake).

In adjective-noun combinations, it is the adjective that is monoclocable. From a semantic perspective, the adjective serves as the governing component of the lexical combination – it “chooses” the lexical partner of the collocational paradigm. The collocational paradigm of a monoclocable word does not necessarily only contain a single lexical unit. However, in a given limited set of units, there needs to be a semantic connection in the form of a common semantic element, cf. *v dohl'adnom: čase / budúcnosti, dobe* (in the foreseeable: time / future, period); *hurónsky: rev / krik / smiech* (“Huron” – uproarious: cries / shouts / laughter).

Monoclocability does not “guarantee” that a collocation will classify as a systemic unit, even though there are lexical combinations which satisfy this requirement, e.g. the lexicalized combinations *pitná voda* (drinking water), *pitný režim* (fluid intake). From the perspective of markedness and the relevance of the lexicographical recording of this type of collocations, we consider monoclocability a significantly anomalous feature.

In addition to limited collocability, we also distinguish preferential collocability. When identifying this type, we ask which component of the combination determines the other, or which one acts preferentially in relation to the other from a collocability standpoint. This type of anomaly usually does not “produce” fixed combinations. It is the so-called typical or usage collocations that are characterized by their feature of selective (i.e. preferential) collocability.

Selective collocability (cf. Jarošová: 2007) is seen as a type of anomalous combinatorial potential, which is:

– motivated by the meaning of that component in the combination which limits the collocational paradigm semantically, not quantitatively, e.g. the deverbative adjective *neskrývaný* (unconcealed) draws in nouns from the group of “expressing emotions”, such as *obdiv* (admiration) / *radosť* (joy) / *hrdosť* (pride) / *pohrdanie* (contempt) / *obava* (worry) / *nadšenie* (enthusiasm) / *záujem* (interest) / ...; nouns from the lexical-semantic class of “people” mainly with a negative semantic element mostly collocate with the deverbative *zarytý* (unrelenting): *odporca* (opponent) / *protivník* (rival) / *nepriateľ* (enemy) / *skeptik* (sceptic) / *pesimista* (pessimist) / *fanúšik* (fan) / ... (*zarytý optimista* (“unrelenting optimist”) is an example of marked usage of this adjective).

– defined by the conventional (i.e. usual) connection of lexical units. It is a statistically supported tendency of certain words to preferentially choose one or two collocational partners that occur considerably frequently (cf. Jarošová 2007: 96). In this case, we do not take into account the semantic restrictions that are in play when choosing lexical units. For example, we can observe significant preference in the collocation *slastný pocit* (blissful feeling), which is determined by the collocability of the adjective *slastný* (blissful): 1. *pocit* (feeling) 274 occurrences,

2. *chvíľa* (*moment*) 40 occurrences, 3. *vzdych* (*sigh*) 22 occurrences. Similarly, in the phrases *pokročilý vek* (*advanced age*) and *pokročilé štádium* (*advanced stage*), the nominal components of the collocation are the most frequent collocates of the adjective *pokročilý* (*advanced*).

Anomalous collocability can concern a single component of a combination as well as the combination itself. We distinguish between two types of anomalous collocability: either the combination occurs in a narrow, semantically defined group of words, or its collocability is preferential, characterised partly by a semantically defined group of lexical units from the collocational paradigm (typical lexical surroundings).

Limited collocability can mostly be found in combinations which have formal limits and are used in an adverbial or attributive function. Strict limitations in the external positions of these combinations often lead to discussions about their lexical limitation (cf. Ďurčo 1997: 51; Mlacek 2001: 29). These are phrases that have a fixed verbal component (usually they are part of a predicate nominal), e.g. *v plnom prúde: byť* (*in full swing: to be*); *v najlepších rokoch: byť* (*in the prime of life: to be*); *v druhom stave: byť / zostať* (*with child: to be / to end up*); *na plný plyn: ísť* (*at full throttle: go*); or phrases which have a limited range of lexical partners which is semantically specified, e.g. *s čistým svedomím: môcť povedať / vyhlásiť niečo / ...* (*with a clear conscience: to be able to say / proclaim something / ...*); *ukázať sa / predstaviť sa / ... v celej kráse* (*show oneself / present oneself / ... in all one's beauty*); *celým menom: volať sa / podpísať sa / osloviť niekoho* (*full name: be called / sign in / call somebody by*), *v plnej paráde: predstaviť sa / predviesť sa / ukázať sa / ...* (*"in full regalia" – in all one's beauty: present oneself / show oneself / appear / ...*), *celou váhou: oprieť sa do niečoho / zvaliť sa na niekoho / ...* (*all one's weight: push something with / throw ... on somebody / ...*).

The anomalous distributional features of collocations which manifest themselves in the form of preferential collocability are often indicators that the collocation is fixed. Even in this case, the typical lexical surroundings (which, however, are usually not limited and are greatly semantically "fragmented") are partly part of a common semantic paradigm. However, this criterion has a low weight, so it cannot be the only one used to determine whether a combination is fixed, cf. *vlastnými rukami: vyrobiť niečo / postaviť niečo / spraviť niečo / ...* (*with one's own hands: create something / build something / make something / ...*) In addition to the typical lexical surroundings, there is semantic coherence between the components, the proof of which is the one-word equivalent *vlastnoručne* (*oneself*).

In some cases, typical lexical surroundings can signalize a semantic shift of the collocational component. For instance, the verbal component of the collocation *na patričných miestach* (*"at the appropriate places"*, also used in the singular form *na patričnom mieste* – *"at the appropriate place"*): *sťažovať sa / odvolať sa / zmeniť sa o niečom / konzultovať niečo / ...* (*complain at / appeal to / mention something to / consult something with / ...*) is becoming fixed. This fact proves that the collocation is used in the sense of „*at official or governmental authorities*“,

but unlike the collocation *vyššie miesta* (“higher places” – higher authorities), or *na vyšších miestach* (“at higher places” – at higher authorities), the expressive component of its meaning is emphasised.

4. Conclusion

At present, the lexicographical treatment of collocations can no longer neglect taking into account the frequency characteristics of the collocational potential of words. The corpus approach to collocations primarily relies on various ways of using methodology and tools to examine the distributional properties of words. The application of these methods lies not only in the identification of collocations in huge corpus databases, but also in the examination of the mutual coherence between the individual components of the collocations. It is primarily thanks to text corpora and corpus tools that linguists and lexicographers have started focusing their attention on such collocations that are easily identified not only because of their statistical significance, but also because of their collocational markedness. As a result, they have started looking for methods that could be used to distinguish these collocations from free combinations of words.

In this article, we present the conceptual treatment of collocations in the *Dictionary of Slovak Collocations. Adjectives*, while emphasising those that lie on the border between texts and the system – the so-called typical collocations (cf. Jarošová 2007). We demonstrate that one of the reasons why these collocations are seen as usual lexical combinations is their anomalous combinatorial potential, on the one hand, on the formal level in the form of morphosyntactic preference (the combinatorial potential of grammatical categories), on the other hand, on the semantic level in the form of monocolocability or preferential collocability (lexical combinatorial potential). We try to emphasise that the anomalousness of these lexical units is not necessarily to be interpreted by applying linguistic rules, but rather that it can have pragmatic motivation. This fact is reflected in the frequency and statistical characteristics of words and to a large degree it influences the selective process when creating the collocational profiles of these words.

REFERENCES

- Čermák 1985:** Čermák, F. Frazeologie a idiomatika. – In: *Česká lexikologie*. Praha: Academia, 1985, s. 166 – 236.
- Ďurčo 1997:** Ďurčo, P. Gramatické vlastnosti frazém z porovnávacieho hľadiska. – In: *Frazeologické štúdie*. II. Ed. P. Ďurčo. Bratislava: Esprima, 1997, s. 45 – 62.
- Ďurčo 2007:** Ďurčo, P. *Zásady spracovania slovníka kolokácií slovenského jazyka*. 2007. <http://www.vronk.net/wicol/index.php/Main_Page>.
- Jarošová 2007:** Jarošová, A. Problém vymedzenia kolokácií. – *Jazykovedný časopis*, 2007, No 2, s. 81 – 102.

Majchráková and Ďurčo 2010: Majchráková, D. and Ďurčo, P. Compiling the first dictionary of Slovak collocations. – In: *Feste Wortverbindungen und Lexikographie. Kolloquium zur Lexikographie und Wörterbuchforschung. Lexicographica, Series Maior 138.* Berlin/New York: De Gruyter, 2010, p. 103 – 114.

Mlacek 2001: Mlacek, J. *Tvary a tváre frazém v slovenčine.* Bratislava: STIMUL, 2001.

MSJ: *Morfológia slovenského jazyka.* Ed. J. Ružička. Bratislava: Vydavateľstvo SAV, 1966.

SOURCES

Slovenský národný korpus – prim-6.1-public-all. Bratislava, Jazykovedný ústav Ľ. Štúra SAV 2013.<http://korpus.juls.savba.sk>.

SSS-PM: *Slovník slovných spojení. Podstatné mená.* Eds. P. Ďurčo, D. Majchráková. Bratislava: Veda, 2015.