

Iwona Kosek

Sebastian Przybyszewski

A Dictionary of Paradigms of Verbal Phraseological Units. An Introduction

1. Theoretical foundations

1.1. Terminology and definitions

This dictionary provides inflectional characteristics of verbal phraseological units and, specifically, explicit presentation of forms constituting inflectional paradigms of these units. The dictionary is a pioneering work: the analysed issues have not been investigated so far, with the exception of some introductory remarks and lexicographic notes for a number of units. The dictionary can be useful for a wide range of readers: not only for linguists but also for anyone interested in Polish phraseology, inflection of phraseological units and automated description of the Polish language. The dictionary can be used by people of different levels of knowledge about language and its algorithmization because information about inflectional paradigms of given units is presented in a number of ways and at different levels of description.

Due to its characteristics, the dictionary focuses both on syntax and phraseology and, consequently, is grounded in these two linguistic disciplines and levels of language description. Methodologically, the dictionary is based on the kind of analysis applied in structuralist descriptions of syntax. The most relevant study of words (lexemes) includes *The Grammatical Dictionary of Polish* by Zygmunt Saloni and collaborators (*Słownik gramatyczny języka polskiego*, SGJP), while the works by Andrzej Maria Lewicki, who laid the foundations for the grammatical description of Polish phraseological units, constitute the most important studies of multi-word lexical units (Lewicki 1976, 1984, 1986a, 1986b and many more). Our research has also been inspired by Lewicki, Pajdzińska & Rejakowa (1987) and Bąba (1997, 1998).

Phraseological units are defined as multi-word units of varying semantic irregularity, whose components belong to substitutionally closed-ended classes, i.e. the classes which, with the maintained meaning of the whole, can only be characterised as an inventory. This approach is based on the concept of a closed/open-ended class as defining a unit's boundary and makes reference to the idea of language units proposed by Andrzej Bogusławski (Bogusławski 1976, 1987, 1989 and other; see also Grochowski 1982). It means that phraseological units can also be

interpreted as language units. Consequently, following this definition, the expression *ktoś uderzył kogoś bykiem* (lit. *someone hit someone else as a bull*, meaning “someone head-butted someone”), where the verbal component *uderzyć* (“to hit”) tends to be replaced by other synonymic verbs, cf. *walnąć, palnąć, rąbnąć, trzasnąć, trafić, zaprawić... kogoś bykiem* etc. is not interpreted as a unit. Because the left-sided connectivity of *bykiem* can be characterized according to general principles, these phrase makes it possible to identify the adverb *bykiem* meaning “hitting an opponent with one’s head”. The situation is analogical in the case of the phrases *ktoś rąbnął (palnął, walnął, trzasnął...) kogoś z byka*, in which *z byka* has the same meaning as *bykiem* and is granted the unit status. On the other hand, the expressions *ktoś leży martwym bykiem* (lit. *someone is lying as a dead bull*, meaning “someone is lying doing nothing”) and *ktoś (chyba) z byka spadł* (lit. *someone has (apparently) fallen off a bull*, meaning “someone is behaving in a surprising and irrational way”), can be interpreted as whole units because the connectivity of their components is unique and none of the components, with the maintained meaning of the whole, belong to a substitutionally open-ended class.

The focus of description are phraseological units functioning as verbs, classified as phrases (Polish term: *zwroty*). We understand **phrases**, following Lewicki, as fragmentary units that require completion, functioning as the main part of a sentence. This understanding is narrowed in the dictionary by the morphological criterion: the focus is only on the phraseological units which have a verbal main component. Thus, what is excluded from the analysis are, on the one hand, such units as *klamka zapadła* (lit. *the door handle has been turned*, analogical to the idiom *the die is cast*) or *znam ten ból* (lit. *I know that pain* meaning “I understand you because I also suffered from that”), because they are not fragmentary and do not open slots for other units, on the other hand, such expressions as *ktoś w krzyk* (lit. *someone into shouting*, meaning “someone started shouting suddenly”) or *ktoś w nogi* (lit. *someone into feet*, meaning “someone started running suddenly”) because they do not display a verbal component.

The dictionary lists two types of verbal expressions: one with the open slot for the sentence subject in Nominative, e.g. *ktoś spija śmietankę* (lit. *someone drinks the cream*, meaning “someone takes undue credit”), *ktoś spuszcza kogoś na drzewo* (lit. *someone sends someone else up a tree*, meaning “someone gets rid of someone in an unpleasant way”), *ktoś/coś dolewa oliwy do ognia* (lit. *someone/something adds oil to the fire*, analogical to the English idiom *to add fuel to the fire*), *coś dochodzi do skutku* (lit. *something comes to a result*, meaning “something is

concluded or takes place”), and the other with the slot for the sentence subject lexically filled, e.g. *serce się komuś kraje* (lit. *someone’s heart is being cut*, meaning “someone feels regret or sadness because of some bad things that happen to someone else”), *język kogoś świerzbi* (lit. *someone’s tongue itches*, meaning “someone has a strong need to say something”), *kiszki komuś marsza grają* (lit. *someone’s guts are playing marches*, meaning “someone is very hungry”). Following Lewicki, such units can be classified respectively as phrases of the verbal group and phrases of the main agreement. Phraseological units, in which the sentence subject is one of the lexical components are far less frequent. In the dictionary, they constitute a group of 300 units. Because the definition of phraseological unit is based on open/close-endedness of substitutional classes, and not on their meaning, they differ in the “degree of semantic connectivity”, as Stanisław Skorupka puts it. Therefore, the dictionary lists both semantically non-compositional phrasemes, known as idioms, e.g. *ktoś puszcza pawia* (lit. *someone sends a peacock*, meaning “someone vomits”), *ktoś sadzi koperczaki* (lit. *someone lopes courtship*, meaning “someone courts”), *ktoś strzela sobie w stopę* (lit. *someone shoots oneself in a foot*, meaning “someone searching for personal benefits, does something that proves to be harmful to him/her”), and partially compositional phrasemes, in which at least one component maintains its lexical meaning, e.g. *coś bierze początek z czegoś* (lit. *something takes its origin in something*, meaning “something begins in something”), *ktoś robi sobie żarty z kogoś / z czegoś* (lit. *someone makes fun of someone/something*), *ktoś ucieka jak szczur z tonącego okrętu* (lit. *someone is running away like a rat from a sinking ship*, meaning “someone leaves a company, party etc. when it is somehow endangered”).

The dictionary is run by the application software *Werbosław* created by Piotr Sikora. *Werbosław* is designed specifically for the needs of a syntactic description of verbal phraseological units and is a modified version of *Toposław* created by the same author.

1.2. Inflectional paradigms of phrases

The dictionary is constructed on the basis of two theoretical assumptions:

- a) phraseological units, as lexical units, display the same characteristics as words (lexemes), and thus these two kinds of language units should be described in the same way;
- b) phrases, which are functional analogues of verbs and have structure motivated by verbal groups, have the inflectional paradigm determined by the main component of the whole, i.e. the

verbal component (Lewicki, Pajdzińska, Rejakowa 1987: 20). However, because their semantics is different, often idiomatic, they can display inflectional constraints and variants in comparison with their founding verbs. Consequently, the inventory of forms in inflectional paradigms is determined both by the verbal component and by individual semantic and functional properties of the whole, especially the semantic type of sentence subject determined by the phrase (Kosek, Przybyszewski, Czerepowicka 2017).

The units presented in the dictionary can be ascribed the inflectional categories of verb and the non-inflectional category of aspect (Perfective/Imperfective). For finite verbs, the inflectional categories are: person, mood, number, and gender (Saloni 2000). These categories, their functions and construction of specific forms, have been described in detail in the grammatical dictionary (sgjp.pl; Introduction to the 2nd edition by Saloni, sections 6.4. and 6.6.) so here we make only a few supplementary comments. We adopt the values for person and number standard for the Polish language. In the category of tense, the Past Perfect Tense is not included as not used in modern Polish. In the category of mood, complex conditional forms such as *ktoś byby poszedł pod nóż* (lit. *someone would have gone under the knife*, meaning “someone would have had an operation”) are not included as is the practice in SGJP. These forms are not included either for defective verb phrases and predicative verb phrases, which is where this dictionary differs from SGJP listing such forms. These forms are very regular and if a phraseological unit has in its paradigm a conditional form (which is the case with a vast majority of the units analysed in the dictionary), it also has complex forms, the form of which is easy to predict even though they may be very rarely used, if at all. Further, we did not include the analytic constructions such as *niech ktoś trzyma kciuki za kogoś* (lit. *let someone keep one’s thumbs for someone else*, meaning “someone wishes someone else good luck”), which constitute part of secondary function of the synthetic imperative mood. In the dictionary, they are treated as a combination of the particle *niech* with the 3rd person non-past tense (i.e. Present or Future Simple). The exception are isolated expressions, which function mainly in imperative mood, e.g. *spadaj na szczaw* (lit. *fall on the sorrel*, meaning “get lost!”). In this particular case, the particle *niech* was treated as a unit: *ktoś (niech) spada na szczaw* (lit. *(let) someone fall on the sorrel*), and the connection was highlighted in the section “Comments” available in the tab “Entry: General Description”).

In comparison to the way SGJP describes verbs, this dictionary offers a more detailed description of gender forms of phrases, namely, it follows the distinctions given in NKJP: m1 (virile / masculine personal gender), m2 (masculine animal gender), m3 (masculine objective gender), f (feminine), n (neuter). What stands behind such a decision (even though it yields multiple morphologically homonymic forms), is the possibility of showing semantics of the subject determined by the units, which is a feature relevant to the inflectional paradigm of a given phrase. This feature is not always obvious, especially, if an expression does not open a slot for personal nouns (i.e. slot of the type “someone”). However, it does not have to entail the lack of masculine personal gender forms. Individual units can differ in that respect, for example, the expression *coś cieszy (czyjeś) oczy* (lit. *something gladdens someone’s eyes*, meaning “something makes someone happy”) allows in its paradigm masculine personal gender forms e.g. *cieszyli oczy* (lit. *they_{virile} gladdened_{virile} their eyes*, meaning “they enjoyed something”) as in *Kossak i Fałat cieszyli oczy zwiedzających wyrazistością szczegółu i kolorystyką* (*Kossak and Fałat gladdened_{virile} the eyes of visitors with the sense for detail and colour*, meaning that the visitors enjoyed the sense for detail and colour in Kossak’s and Fałat’s paintings), which is not the case with *coś (np. awantura) wisi w powietrzu* (lit. *something, e.g. an argument, is hanging in the air*, meaning “something unpleasant is about to happen”) (Kosek, Przybyszewski, Czerepowicka 2017).

The inflectional paradigm also includes non-finite forms: infinitive, e.g. *uchylać rąbka tajemnicy*, *uchylić rąbka tajemnicy* (lit. *to unfold a hem of a secret*, meaning “to reveal partly something that is a secret”, Imperfective and Perfective forms), impersonal forms in indicative and conditional mood e.g. *uchylano rąbka tajemnicy* (*one was unfolding a hem of a secret*), *uchylano by rąbka tajemnicy* (*one would unfold a hem of a secret*); *uchylono rąbka tajemnicy* (*one unfolded a hem of a secret*), *uchylono by rąbka tajemnicy* (*one would unfold a hem of a secret*), contemporary adverbial participle, e.g. *uchylając rąbka tajemnicy* (*unfolding a hem of a secret*), and anterior adverbial participle e.g. *uchyliwszy rąbka tajemnicy* (*having unfolded a hem of a secret*). We, however, do not list regular derivatives: adjectival participle and gerund e.g. *uchylający rąbka tajemnicy*, *uchylanie rąbka tajemnicy* (*unfolding a hem of a secret*). We treat them as separate units, not inherent part of the analysed phrases, which are subject to declension not conjugation. As visible from the set of inflectional categories given above, the distinction into

passive and active voice is not included either. We treat it as a syntactic phenomenon, which is consistent with contemporary approaches to syntax.

A phrase can be said to display a full inflectional paradigm if, being founded on a finite verb, has both all the above-mentioned categories fully realized and the impersonal forms. It is known that in Polish, the scope of tenses and the kind of adverbial participle depend on aspect. This is why, in this dictionary, the full inflectional paradigm will include, for imperfective expressions: present tense, future periphrastic tense, past tense, synthetic conditional mood and imperative mood, with a maximum distinction for person, gender and number, plus impersonal forms with contemporary adverbial participle discussed above, and for perfective expressions: future simple (instead of present and periphrastic future), past tense, synthetic conditional mood and imperative mood, and anterior adverbial participle (yet not contemporary). An example of a unit with such a paradigm could be *ktoś/coś gra komuś na nerwach* (lit. *someone/something plays someone's nerves*, meaning “someone/something gets on someone's nerves”), *ktoś bierze coś na warsztat* (lit. *someone takes something to his workshop*, meaning “someone works on something”), *ktoś bije brawo* (lit. *someone beats bravo*, meaning “someone applauds”), *ktoś daje komuś nauczkę* (lit. *someone gives someone a lesson*, meaning “someone punishes someone who apparently deserves it”), *ktoś utrze komuś nosa* (lit. *someone will rub someone's nose*, meaning “someone will bring someone else down a peg”), *ktoś palnie sobie w łeb* (lit. *someone will shoot himself in his head*, meaning “someone will commit suicide”). Obviously, the forms differ in their frequency of occurrence (Przybyszewski 2015; Kosek, Przybyszewski, Czerepowicka 2018). Some forms, for example the gerund and adverbial participle, are very rare for most units.

Incompleteness of the inflectional paradigm of a phrase may involve missing only some forms, for example, impersonal forms as in *ktoś / coś idzie w zapomnienie* (lit. *someone/something goes into oblivion*, meaning “someone/something is getting to be forgotten”) or missing whole categories. Thus, expressions which open the slot for the sentence subject “something”, e.g. *coś dokonuje żywota* (lit. *something accomplishes its life*, meaning “something is over or set for destruction”), *coś dochodzi do skutku* (lit. *something comes to a result*, meaning “something is concluded or takes place”), *coś jest bez ceny* (lit. *something is without a price*, meaning “something is priceless”), do not have in their paradigm (next to other specific constraints) the category of person, imperative mood, and impersonal forms (Kosek 2013, 2017).

Consequently, their paradigm is simpler than that of units connected with a personal agent (such as presented in the previous paragraph).

A phrase can also have only non-finite forms. For example, the paradigm of the phraseological unit *komuś wilkiem z oczu patrzy* (lit. *someone's eyes look like wolf's eyes*, meaning “someone looks with hate or has hostile look”) generates only a few forms, which are: *komuś wilkiem z oczu patrzyło* (*someone looked with hate*), *będzie patrzeć / będzie patrzyło wilkiem* (*will look with hate*), *patrzyłoby wilkiem* (*would look with hate*), *patrzeć wilkiem* (*look with hate*). We treat such phrases analogically to defective verbs in word inflection. However, it does not entail that they are always motivated by verbs of such kind. *Brakować* (“to lack, to miss”), which is a component of the phraseme *komuś brakuje słów* (lit. *someone lacks words*, meaning ‘someone does not know what to say’), is described in SGJP as incorrect, unlike *patrzy* (*looks*) from the example above. Still, both phrases, just like defective verbs, function the same way as wholes, and display an analogical set of forms, which is the decisive factor when describing inflection of phraseological units. This is the reason why our dictionary approaches them in the same way and why they receive the comment: “a full paradigm of a defective verb”. The label “defective” means that the phrase does not permit a connection with a subject in Nominative and has the structure of inflectional forms analogical to finite verbs (i.e. synthetic inflection). Apart from the units mentioned above, such a description applies to phrases such as *komuś dzwoni w uszach* (lit. *it rings in someone's ears*, meaning “someone has their ears ringing”), *komuś rzuca się na mózg* (lit. *it throws itself on someone's brain*, meaning “someone acts irrationally”). A separate group consists of phrases motivated by the so-called secondary defective verbs, e.g. *komuś dużo nie potrzeba* (lit. *it is not much needed for someone*, meaning “someone does not have high expectations or requirements”), which are labelled as predicative expressions in the dictionary. Unlike other defective verbal phrases, they have only analytic inflection, but similarly to other defective verbal phrases, they have the same inflective categories, i.e. tense and mood.

A separate group consists of expressions with the subject as their component, e.g. *łza się komuś w oku kręci* (lit. *a tear is turning in one's eye*, meaning “someone feels touched”) or *mina komuś zrzędzie* (lit. *the face will get thin to someone*, meaning “someone's good mood or self-confidence will disappear”). In some studies on phraseology, such expressions are considered to be phrases. Lexical filling of this slot involves a considerable simplification of the paradigm,

which comprises only of tense forms, conditional mood and infinitive, e.g. *Mina mu może wtedy zrzędnąć* (lit. *his face may get then thin*), and number for some expressions, e.g. *Miny im zrzędyły* (lit. *their faces got thin*). The last two types are of special interest because they have been so far ignored in studies of inflection (with the exception of WSPJ, which comments on some of them).

Finally, phraseological units can display specific constraints, defying regularity, and characteristic only to individual examples, not whole groups. For instance, the expression *ktoś nie wygrał czegoś na loterii* (lit. *someone did not win something at a lottery*, meaning “someone will not strain or use something without an important cause”) has only the past tense forms; *ktoś zobaczy coś jak świnia niebo* (lit. *someone will be able see something as a pig will be able to see the sky*, meaning “something is highly unlikely”) has in its paradigm only the future tense, conditional mood and infinitive; and *ktoś nie jest z cukru* (lit. *someone is not of sugar*, meaning “someone should not be afraid of rain and may stay outside when it is raining”) whose paradigm is limited to the present tense.

All the described types and groups of constraints identified in the analysis of the collected several thousands of expressions were listed in our dictionary.

2. The dictionary parameters: sources scope and kind of provided information

The dictionary describes over 5000 expressions which come from different lexicographical studies, corpora (especially NKJP) and the authors’ own records. In collecting data, the most useful proved to be general dictionaries of the Polish language (ISJP, SJPD, SJPSz, SWJP, USJP, WSJP), phraseological dictionaries (PSF, SFJP, SFWP, WSF, WSFJP) and lexicographical compilations (Bogusławski, Garnysz-Kozłowska 1979; Bogusławski, Wawrzyńczyk 1993; Bogusławski, Danielewiczowa 2005). However, the form of the excerpted units has not always been kept unchanged, which results from the differences in approach (more on the discrepancies between different studies in relation to the present dictionary see Kosek, Przybyszewski, Czerepowicka 2018, forthcoming).

The main aim of the dictionary is to provide explicit information about the inflectional paradigm of a given expression. Such information is recorded in a number of ways and accessible from several levels: the paradigm is described by comments like, e.g. “full paradigm”, “no

impersonal form”, “present tense only”, by listing all generated forms with their morphological characteristics (tagged), and by presenting sub-sets (collections of flexemes representing the values depending on tense, indicative mood, the remaining moods, and non-finite forms). The forms creating specific sub-sets can be separately generated (we say more about levels of a paradigm description in the next section). Imperfective expressions with the full paradigm are characterized by a set of ten flexemes, perfective expressions by eight. The numbers are higher for the forms generated in the past tense, conditional mood and future periphrastic tense due to stem alternations as in *niosł-/niósł* (“carried”). In such cases, additional flexemes were included for masculine agglutinative forms (1st and 2nd person sing. m1, m2 and m3, e.g. *niosłem* [I carried], *niosłeś* [you carried]) and for masculine non-agglutinative (3rd person sing. m1, m2 and m3, e.g. *niósł* [he carried]). The forms in conditional mood and in future periphrastic tense, singular number and masculine gender exist only for non-agglutinative forms (the kind exemplified by *niósł*), which is why only one additional flexeme was generated.

The dictionary is not limited to listing inflectional forms of specific units. It also records the aspect value of a given phrase and provides an aspectual pair to the described unit (if such a pair exists). Each of the pair has a separate entry in the dictionary, which is motivated by its inflectional nature, on the one hand, and on the other, by the fact that the components of the pair do not have to display identical paradigms, for example, *ktoś obchodzi się smakiem* (lit. *someone makes do with the taste*, meaning “someone does not achieve what they hoped for”) has a full set of forms while its aspectual counterpart *ktoś obejdzie się smakiem* does not have an impersonal form.

Our study focuses on syntax but in the case of many of the units, meaning is also taken into account. As a rule, it is provided with homonymic units and with more recent units, excerpted from the corpora, such as *ktoś ugotuje kogoś na twardo* (lit. *someone will cook someone else hard*, meaning “someone will defeat someone”) or *ktoś spuszcza kogoś na drzewo* (lit. *someone sends someone else up a tree*, meaning “someone gets rid of someone in an unpleasant way”) would could be unfamiliar for the reader. If an expression has an aspectual pair, then its semantic characteristics is given with the imperfective realization.

An important section of the dictionary is the tab “Comments”, in which the reader can find additional information. Comments are grouped according to the type and labelled: F, P, G.

W, N and Np (meaning “e.g.”). Type **F** comments concern a description of a paradigm. Type **P** comments provide pragmatic information related to forms’ frequency of occurrence or scope of use, for example, “typically in perfective aspect”, “typically in past tense”, etc. However, most often the characteristics of imperative mood are included here as well. In the case of some expressions, such as *ktoś chowa się za czyimiś plecami* (lit. *someone is hiding behind someone’s back*, meaning “someone is trying to avoid something dangerous or risky by exposing someone else”), *ktoś buja w obłokach* (lit. *someone is swinging in the clouds*, analogical to the English idiom “someone has his/her head in the clouds”) or *ktoś czepia się czegoś jak pijany płotu* (lit. *someone is grabbing something as a drunk to a fence*, meaning “someone stubbornly insists on something and does not want to change his/her mind”), imperative mood is used mainly with negation, which is acknowledged in the comments as “imperative mood in the base function typically with negation”. The dictionary also acknowledges cases, in which the imperative is not in the base function, that is ordering or prohibiting, but is used as a warning, wish, threat, etc. (for a description of the imperative in relation to verbs see Laskowski 1998). Then in the comments there appears information “imperative mood in secondary function, cf. the description of such expressions as *ktoś stawia na złego konia* (lit. *someone bets on the wrong horse*, meaning “someone had made a wrong decision supporting someone who lost”) or *ktoś ma poślizg* (lit. *someone has a slip*, meaning “someone is late”).

Type **W** comments focus on non-verbal specific variants of a given unit, e.g. the expression *ktoś umiera własną śmiercią* (lit. *someone dies their own death*, meaning “someone dies of natural causes”) has the nominal element realized as *swoją śmiercią* (*their own death*) or *śmiercią naturalną* (*natural death*); *ktoś stapa po kruchym lodzie* (lit. *someone is treading on brittle ice*, meaning “someone is doing something risky”) can also appear with a component *po cienkim lodzie* (*on thin ice*). Variants are recorded either as separate lexemes or as combinations of several components. The decision was made on grounds of the number of components undergoing inflection and the sake of clarity of presentation. Classifying relationships between components as variants means that they can be interchanged irrespective of internal grammatical factors and the context of a sentence. For example, if for the expression *ktoś wycina komuś (jakiś) numer* (lit. *someone makes (some) stunt to someone*, “someone does something unexpected and harmful to someone else”) we have a variant *numery* (*stunts*), it means that both components are always interchangeable. The choice of a component depends only on the speaker’s preferences.

Variants of verbal components are not acknowledged in the comments but, for obvious reasons, treated as separate units, which is why *ktoś chowa coś pod korcem* (lit. *someone is hiding something under a bushel*, meaning “someone does not reveal something and keeps it in secret”) and *ktoś ukrywa coś pod korcem* (lit. *someone is concealing something under a bushel*) have in the dictionary separate entries and separate descriptions.

Type **G** comments concern syntax of specific components, that is the category of number (and sometimes gender) of nominal elements. It means that they refer to internal syntax of a phraseological unit (in the sense of Lewicki 1986a) and correlation options between the form of the expression and an external context. In the dictionary, there are two kinds of such comments:

- a) “the component X can be inflected for person”,
- b) “the value of number (and gender) of the component X correlates with the variable Y”.

Information of the type a) appears next to the phrases, in which the nominal component in singular can be inflectionally adjusted to the sentence context, that is, it can (but may not) occur in plural if the subject and (simultaneously) the verbal component are also in plural. Such situation is illustrated by the expressions: *ktoś chwyta się za głowę*, *ktoś łapie się za głowę* (lit. *someone is grabbing their head*, meaning “someone expresses his/her astonishment”), *ktoś łamie sobie język na czymś* (lit. *someone is breaking their tongue on something*, meaning “someone has difficulties in pronouncing some words or phrases”), which in discourse can be realised as *łapali / chwyтали się za głowę* (*they were grabbing their head*), *łamali sobie język na czymś* (*they were breaking their tongue*), but also as *łapali się za głowy* (*they were grabbing their heads*), *łamali sobie języki na czymś* (*they were breaking their tongues*). It is the speaker’s decision to choose between singular or plural of “head” and “tongue” with the plural form of the main component. In the case of such dependencies, only the base form of a component is listed among the forms constituting the paradigm.

Information of the b) type means that the number of an indicated subordinate component correlates with one of the slots opened by the expression. For example, in the forms of the unit *ktoś przechodzi samego siebie* (similar to English idiom: *someone outdoes himself/herself*), the number and gender of the component *samego* (*self*) is adjusted with the meaning and reference of the subject, e.g. *One przechodzą same siebie* (*They_{nonvirile} outdo themselves_{nonvirile}*), *Jarek i Marcin przechodzili już samych siebie w tych łgarstwach* (*Jarek and Marcin were outdoing_{virile}*

themselves_{virile} in their lies). In such cases, the paradigm given in the dictionary contains a form marked for number and gender matching the main component e.g. *przechodził sam siebie* (*he was outdoing himself*), *przechodziła sama siebie* (*she was outdoing herself*), *przechodziły same siebie* (*they_{nonvirile} were outdoing themselves_{nonvirile}*), *przechodzili samych siebie* (*they_{virile} were outdoing themselves_{virile}*) etc. We realise, however, that the situation is more complex: in elaborated nominal phrases, the value of number and gender of a nominal component can match not necessarily that of the subject, but that of one of the elements of the nominal phrase, as in, for example, *Większość z tych ludzi przechodziła samych siebie, gdy chodziło o uchylanie się od pracy* (Majority [feminine gender] of these people *outdid_{nonvirile}* themselves, when it came about avoiding work). The forms of expressions in which the verbal component is not inflected, such as infinitive, are presented with inflected subordinate components, e.g. *przechodzić samego siebie* (*to outdo himself*), *przechodzić samą siebie* (*to outdo herself*), etc. The value of number (and sometimes gender) of a subordinate component can also depend on a different slot opened by the expression, as in, for example, *ktoś pozostawia kogoś/coś samemu sobie* (lit. *someone leaves someone/something on one's/its own*, meaning “someone does not meddle”), in which the value of number and person of the component *samemu* (one's) directly depends on the variable “kogoś” (somebody), e.g. *pozostawili ją samej sobie* (*they left_{virile} her on her own*), *pozostawili ich samym sobie* (*they left_{virile} them on their own*). In the case of such dependencies, as in the type a), only the nominal component in singular is listed from among the forms constituting the paradigm.

As it becomes apparent, the degree of internal syntactic and inflectional dependencies among components of an expression can be very high. We have tried to acknowledge these complexities at one level of description in our dictionary. We have ignored, however, for obvious reasons, Genitive forms of units without a negative component, for example, *ktoś czyta grę przeciwnika* (lit. *someone is reading their opponent's games_{Accusative}*, meaning “someone understands and foresees their opponent's moves”) vs *ktoś nie czyta gry przeciwnika* (lit. *someone is not reading their opponent's games_{Genitive}*) and Accusative forms in units with a negative component, for example, *Nie mam teraz głowy do tej sprawy* (lit. *I don't have a head_{Genitive} for this case now*, meaning “I have no patience or interest for this now”) vs *Masz teraz głowę do tej sprawy?* (lit. *Do you have a head_{Accusative} for this case now?*). The change of the object in Accusative to the object in Genitive with negation for expressions of the structure V + N_{Accusative} (and vice versa for expressions of the structure *nie* V + N_{Genitive}) is regular and thus

constitutes an element of their paradigm (cf. Lewicki 1976, Ch. 4). It can be described, however, as a general rule because it derives from one of the most fundamental principles of Polish syntax.

The symbol **N** is used in comments to mark forms which are archaic or less acceptable in modern Polish, which is presented by the formula “in (an X category), possible archaic/less acceptable”. Such situation often obtains in imperative forms though not only (see, for example, the description of the units *ktoś weźmie kogoś na plewy* (lit. *someone will take someone else in with the chaff*, meaning “someone will beguile someone”); *ktoś/coś znajdzie się na czyichś ustach* (lit. *someone/something will appear on someone’s mouth*, meaning “someone/something will get popular and will be discussed”). If specific forms differ only in the frequency of occurrence (they are qualified as “rare” in dictionaries), such normative information is not provided.

The last type of comment, **Np.**, are usage examples. When added to descriptions of some units, they are intended to help the reader with understanding contexts. We aimed at providing typical examples. Examples come with data about their source (corpus type, from which they come, which is either NKJP or the Internet). Lack of such data means the authors’ own example.

3. The Structure of the dictionary

This a downloadable e-dictionary. It contains both data and metadata. The latter can prove interesting for the readers investigating mechanisms of algorithmic description of phraseological units. When the program is started a list of units appears. They are recorded in 3rd person form, which makes the identification of the slots opened by the expression possible. Such mode of recording is one way of presentation of semantic and syntactic requirements of units. Specific phraseological units can be located either by scrolling through the list or by using a browser.

When an expression is selected, the home page appears. It gives access to five sections of a unit’s description. The first one is “Entry: General Description”, which consists of: a form of an expression, aspect value, optionally the expression’s meaning, comments and information about an aspectual pair. In this part, the reader can also find “Comments on a counterpart”, that is, comments about the expression’s aspectual pair. It is used to present alternative aspectual counterparts. Thus, for example, in the description of the unit *ktoś pocałuje kogoś z dubeltówki* (lit. *someone will kiss someone [as if it were a shot] from a shotgun*, meaning “someone kisses loudly both cheeks of someone”), the “Comments on a counterpart” informs us about “two possible aspectual counterparts”, which means that the phrase which participates in aspectual

relation, i.e. *całować z dubeltówki*, can have an alternative counterpart (in this case, *ucalować z dubeltówki*).

The tab “Unit Description” contains an analysis of an internal structure of individual components, with the main component (the head) specified. Next, there is a set of flexemes realizing the unit, that is form types. At this level of description, one can get an idea of the degree of a paradigm’s completeness.

A set of lexemes can be generated from the tab “Lexeme’s Forms”. Selecting a given lexeme and activating the command “Generate” we obtain a set of forms together with their morphological characteristics (tagged). Thus, in this part of the dictionary, the paradigm is presented for finite forms from the level of specific categories of tense and mood, while for impersonal forms and in the case of expressions functionally corresponding to defective verbs and predicatives directly from the level of form. The whole paradigm can be accessed from the tab “All Forms of a Unit”. The forms are generated with the exceptions of the ones already displayed. The tab “Flexeme inflection: A Description” presents metadata.

References:

Bańko M., 2002, Wykłady z polskiej fleksji, Warszawa.

Bąba S., 1997, Z fleksji polskich frazeologizmów. Part II: Uwagi o odmianie fraz, [In:] Z polszczyzny historycznej i współczesnej, T. Ampel (ed.), Rzeszów, pp. 153–162.

Bąba, 1998a, Uwagi o odmianie polskich zwrotów przez czasy, Poznańskie Spotkania Językoznawcze, Vol. III, Z. Krążyńska, Z. Zagórski, eds., Poznań, pp. 7–14.

Bogusławski A., 1976, O zasadach rejestracji jednostek języka, Poradnik Językowy 8, pp. 356-364.

Bogusławski A., 1987, Obiekty leksykograficzne a jednostki języka, [In:] Studia z polskiej leksykografii współczesnej, Vol. II, Z. Saloni (ed.), Białystok, pp. 13-34.

Bogusławski A., 1989, Uwagi o pracy nad frazeologią [In:] Studia z polskiej leksykografii współczesnej, Vol. III, Z. Saloni (ed.), Białystok, pp. 13-30.

Bogusławski A., Garnysz-Kozłowska T., 1979, Addenda do frazeologii polskiej. Zeszyt wstępny, Edmonton.

Bogusławski A., Wawrzyńczyk J., 1993, Polszczyzna jaką znamy. Nowa sonda słownikowa, Warszawa.

Bogusławski A., Danielewiczowa M., 2005, Verba polona abscondita. Sonda słownikowa III, Warszawa.

Grochowski M., 1982, Zarys leksykologii i leksykografii. Zagadnienia Synchroniczne, Toruń.

Kosek I., 2013, Paradygmaty zwrotów frazeologicznych – problemy opisu leksykograficznego, [In:] Perspektywy współczesnej frazeologii polskiej. Między teorią a praktyką leksykograficzną, G. Dziamska-Lenart and J. Liberka (eds.), Poznań 2013, pp. 51-61.

Kosek I., 2017, *Pełny czy defektywny?* O ustalaniu paradygmatów związków frazeologicznych, Prace Filologiczne LXXI, pp. 201-213.

Kosek I., Przybyszewski S., Czerepowicka M., 2017, Realizacja pozycji podmiotu otwieranej przez zwrot frazeologiczny a jego paradygmat (na wybranych przykładach), Prace Językoznawcze XIX/3, 2017, pp. 113-122.

Kosek I., Przybyszewski S., Czerepowicka M., 2018, Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych, Biuletyn PTJ.

Laskowski R., 1998, Semantyka trybu rozkazującego, Polonica XIX, pp. 5-29.

Lewicki A. M., 1976, Wprowadzenie do frazeologii syntaktycznej. Teoria zwrotu frazeologicznego, Katowice.

Lewicki A. M., 1984, Wskaźniki utrwalenia jednostki słownikowej, [In:] Słownictwo w opisie języka, K. Polański (ed.), Katowice, pp. 42-555, reprint [In:] A. M. Lewicki, Studia z teorii frazeologii, Łask 2003, pp. 184-194 .

Lewicki A. M., 1986a, Składnia związków frazeologicznych, Biuletyn Polskiego Towarzystwa Językoznawczego XL, pp. 75-82.

Lewicki A. M., 1986b, Słownik a gramatyka w leksykograficznym opisie związków frazeologicznych, [In:] Typy opisów gramatycznych języka. Materiały polsko-czeskiej sesji

naukowej. Jabłonna 15-17 XI 1983, M. Basaj (ed.), Wrocław, pp. 157-170; reprint [In:] A. M. Lewicki, *Studia z teorii frazeologii*, Łask 2003, pp. 168-183.

Lewicki A. M., Pajdzińska A., Rejakowa B. 1987: *Z zagadnień frazeologii. Problemy leksykograficzne*, Warszawa.

Przybyszewski S., 2015, Some problems with the description of paradigms of Polish multi-word units, [In:] *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)*, b. 18, E. Gutierrez Rubio, M. Falkowska, E. Kislova, M. Stępień (eds.), Wiesbaden, pp. 213-223.

Saloni Z., 2000, *Wstęp do koniugacji polskiej*, Olsztyn.

Saloni Z., 2012, *Podstawy teoretyczne „Słownika gramatycznego języka polskiego”*, Warszawa, online:

<http://sgjp.pl/static/pdf/Wst%C4%99p%20do%20II%20wydania%20SGJP.pdf?1458313456>

Dictionaries and Corpora:

ISJP – Bańko M, ed., 2000, *Inny słownik języka polskiego, I–II*. – Warszawa: PWN.

NKJP – Przepiórkowski, A., Bańko M., Górski R. L., Lewandowska-Tomaszczyk B., (eds.), *Narodowy Korpus Języka Polskiego*, Warszawa: PWN; <http://www.nkjp.pl>

PSF – Bąba S., Liberek J., Dziamska-Lenart G. (eds.), 2001, *Podręczny słownik frazeologiczny języka polskiego*, Warszawa: PWN.

SFJP – Skorupka S., 1985, *Słownik frazeologiczny języka polskiego*, Warszawa: Wiedza Powszechna, wydanie IV.

SFWP – Bąba S., Liberek J. (eds.), 2001, *Słownik frazeologiczny współczesnej polszczyzny*, Warszawa: PWN.

SGJP – Saloni Z., Woliński M., Wołosz R., Gruszczyński W., Skowrońska D. (eds.), 2015, *Słownik gramatyczny języka polskiego, 3rd edition on-line*, <http://sgjp.pl>

SWJP – Dunaj B. (ed.), 1996, *Słownik współczesnego języka polskiego*, Warszawa: Wydawnictwo Wilga.

USJP – Dubisz S. (ed.), 2003, *Uniwersalny słownik języka polskiego, I–IV*, Warszawa: PWN.

WSF – Kłosińska A., Sobol E., Stankiewicz A. (eds.), 2005, Wielki słownik frazeologiczny z przysłowiami, Warszawa: PWN.

WSFJP – Müldner-Nieckowski P., 2003, Wielki słownik frazeologiczny języka polskiego, Warszawa: Świat Książki.

WSJP – Żmigrodzki P. (ed.), Wielki słownik języka polskiego, <http://www.wsjp.pl>