

MONIKA CZEREPOWICKA
IWONA KOSEK
SEBASTIAN PRZYBYSZEWSKI

O projekcie elektronicznego słownika odmiany frazeologizmów czasownikowych

1. Charakter i cel badań

W artykule zamierzamy przybliżyć projekt badań związany z charakterystyką gramatyczną związków frazeologicznych¹. Z konieczności niniejsza prezentacja będzie miała charakter szkicowy, skupimy się na przedstawieniu ogólnych założeń opisu, wskazaniu podstawowych problemów i własności jednostek, które opis będzie uwzględniał.

Głównym celem podjętych badań jest stworzenie elektronicznego słownika odmiany zwrotów frazeologicznych, czyli związków pełniących w zdaniu funkcję członu głównego (zob. Lewicki 1986)². Szczegółowych przyczyn, dla których postanowiliśmy zająć się taką tematyką badawczą, jest kilka. Problemy związane z opisem fleksyjnym zwrotów nie były do tej pory przedmiotem dokładniejszej analizy, a jej podjęcia domaga się i sam rozwój języka (w którym zaobserwować można wzrost liczby różnych wielowyrazowych całości), i rozwój leksykografii. Włączanie do słowników ogólnych, zarówno w wersji tradycyjnej, jak i elektronicznej, coraz większej liczby związków frazeologicznych wymaga ich precyzyjnej charakterystyki fleksyjno-składniowej, natomiast badania dotyczące cech gramatycznych frazeologizmów są w językoznawstwie niezbyt zaawansowane, zarówno w zestawieniu z opisem leksemów, jak i na tle pokazanego i zróżnicowanego dorobku naukowego dotyczącego jednostek wielowyrazowych. Bardzo wyraźnie pokazują to zarówno słowniki, jak i „Materiały do bibliografii frazeologii polskiej” Stanisława Bąby³.

¹ Projekt jest finansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/09/B/HS2/01222.

² Bardziej szczegółowe definicje i rozróżnienia terminologiczne wprowadzamy w p. 2 artykułu.

³ W indeksie rzeczowym poszczególnych części „Materiałów” analiz zaklasyfikowanych jako stricte fleksyjne jest łącznie zaledwie kilka. Obrazu całości nie zmieniają nieliczne prace o szerszej problematyce (z odniesieniami fleksyjnymi), uwzględnione w innych hasłach indeksu.

Fleksji związków frazeologicznych dotyczą fragmenty pionierskiej pracy Andrzeja Marii Lewickiego (1976), artykuły Stanisława Bąby (1996, 1997, 1998a), Jarosława Liberka (2006) oraz prace Iwony Kosek (2008, 2013). Fleksja frazeologizmów o funkcji rzeczownika jest zbadana bardziej szczegółowo (Kosek 2008), opracowana także leksykograficznie (zob. Czerepowicka, Savary 2012). Natomiast paradygmatyka zwrotów nie jest opisana, poza uwagami przyczynkowymi i adnotacją słownikową przy niektórych zwrotach. Jednak z dotychczasowych analiz cząstkowych Bąby (1998a) i Kosek (2013) wynikałoby, że paradygmaty licznych związków są defektywne. Stąd też potrzeba zweryfikowania tego poglądu, w oparciu o większą, co najmniej kilkutysięczną grupę jednostek. Analiza ma być w zamierzeniu dokonywana zarówno w oparciu o własną kompetencję językową, jak i na podstawie tekstowych wystąpień, zwłaszcza danych z Narodowego Korpusu Języka Polskiego.

Wykorzystanie przykładów internetowych (w tym dostępnych w korpusach tekstów) może stanowić istotną pomoc w określeniu form preferowanych przez użytkowników, a tym samym — wskazaniu typowych kontekstów gramatycznych, typowych konstrukcji składniowych, w których używane są poszczególne zwroty lub grupy zwrotów. Byłby to równocześnie komentarz do tez i sugestii wysuwanych w dotychczasowych opracowaniach. Ze wspomnianych wyżej analiz S. Bąby można mianowicie wyciągnąć wniosek, że między frekwencją poszczególnych form paradygmatu istnieją duże różnice: niektóre formy zwrotów są przez użytkowników w komunikacji pomijane lub używane bardzo rzadko, być może także z powodu wątpliwości co do normatywności ich użycia.

Charakteryzowane badania są ściśle związane ze zalgorytmizowanym opisem fleksji frazeologizmów, w pewnej mierze stanowią odpowiedź na rozwój i potrzeby językoznawstwa komputerowego. Narzędzia służące do sformalizowanego opisu gramatycznego polszczyzny były dotychczas rozwijane z myślą przede wszystkim o leksemach. Ich wykorzystanie w analizie frazeologizmów czasownikowych napotyka na trudności, przede wszystkim ze względu na specyfikę fleksyjno-składniową tych wielowyrazowych jednostek. Badania będą zatem stanowić również weryfikację istniejących narzędzi, konieczne okaże się zapewne ich zmodyfikowanie i rozwinięcie.

Jak wspomnieliśmy na wstępie, głównym celem opisywanych badań gramatycznych jest stworzenie elektronicznego słownika paradygmatów zwrotów frazeologicznych (ok. 5 000 jednostek). Wybór opracowania innego niż tradycyjne nie jest współcześnie niczym zaskakującym: „Przyszłość frazeografii, jak i w ogóle leksykografii, niewątpliwie należy do mediów i nośników elektronicznych” (Bańko 2013: 15). Takie rozwiązanie ma także swoje istotne zalety, pozwala na przykład łatwo uzupełniać, rozbudowywać opis o nowe jednostki. Elektroniczne opracowanie fleksji daje też możliwości, których nie oferuje wersja tradycyjna (np. różne poziomy „przeszukiwania” paradygmatu), jest wreszcie tańsze.

2. Założenia opisu

Spojrzenie na paradygmaty zwrotów, sposób ich analizy wynika z podstawowej cechy frazeologizmów — z faktu, że są one motywowane przez grupy składniowe. Przyjmujemy zatem jako punkt wyjścia założenie, że jeśli funkcja składniowa sfrazeologizowanej całości i fundującego ją komponentu są tożsame, a więc zachowana jest motywacja kategoriajno-gramatyczna (w rozumieniu Lewickiego (1982)), składnik czasownikowy wyznacza też schemat fleksyjny jednostek wielowyrazowych. Mając jednak odrębną semantykę i swoiste konteksty wystąpień, jednostki te mogą wykazywać ograniczenia i różnice fleksyjne w stosunku do odpowiednich czasowników. Problem zasadniczy sprowadza się zatem do odpowiedzi na pytanie, w jakim stopniu paradygmat poszczególnych zwrotów wyznaczany jest przez główny komponent, a w jakim jest on swoisty, ograniczony właściwościami indywidualnymi zwrotu.

Na poziomie rozróżnień głównych nawiązujemy do koncepcji jednostek języka Andrzeja Bogusławskiego (1976 i in.) oraz do prac A.M. Lewickiego (zwłaszcza 1986), za którym przyjmowane jest rozróżnienie składni wewnętrznej i zewnętrznej oraz klasyfikacja funkcjonalna związków frazeologicznych. Rozumienie zwrotu zostało natomiast nieco zmodyfikowane: tym mianem określamy jednostki fragmentaryczne (czyli niebędące frazami), pełniące funkcję członu głównego w zdaniu (jak wspomniano we wstępie) i opierające się w składni wewnętrznej na komponencie czasownikowym. Jest to zatem rozumienie funkcjonalne (tak jak w pracach Lewickiego), zawężone poprzez kryterium morfologiczne. Takie rozumienie pozwala wykluczyć z obszaru badań z jednej strony jednostki w rodzaju *niech Bóg broni*, z drugiej np. *ktoś w nogi*.

Przyjmujemy, że zwrotom można w punkcie wyjścia przypisać określone kategorie fleksyjne. Za kategorie takie uznajemy w odniesieniu do jednostek opartych na czasownikach właściwych (w rozumieniu Z. Saloniego (2000)) osobę, liczbę, rodzaj, czas i tryb. Jako posiadający pełny paradygmat jest zatem traktowany zwrot, który (jeśli opiera się na czasowniku właściwym) realizuje wszystkie wartości wskazanych kategorii. Do paradygmatu włączane są formy nieosobowe: bezokolicznik, bezosobnik i imiesłów przysłówkowy.

Zakładane jest także odnotowanie, czy badana jednostka tworzy parę aspektową. Niezależnie od tego, jak kategoriałnie potraktujemy aspekt, w opisie gramatycznym zwrotów wymaga on niewątpliwie uwzględnienia ze względu na brak całkowitej odpowiedniości między parami aspektowymi czasowników i parami frazeologizmów. Pokazują to choćby takie przykłady, jak *obrzucac* – *obrzucić kogoś błotem*, w którym mamy pełną analogię z czasownikową parą aspektową, czy *bić pianę*, *robić bokami*, w których takiej analogii już brak (podane zwroty nie tworzą par aspektowych, por. **ubić pianę*, **zrobić bokami*, w przeciwieństwie do homonimicznych z głównym komponentem czasowników).

Mówiąc o charakterze opisu, trzeba wspomnieć o jeszcze jednym założeniu: opisywaną klasę jednostek mają tworzyć zwroty zróżnicowane pragmatycznie, tzn. pochodzące z różnych rejestrów języka (wysokiego i potocznego), należące jednak do polszczyzny ogólnej, a nie specjalistycznych jej odmian.

W pierwszym etapie badań zamierzamy wykorzystać opracowanie paradygmatów czasownikowych dokonane w pracy Saloniego (2001) i w Słowniku gramatycznym (Saloni, Woliński, Wołosz, Gruszczyński, Skowrońska 2012), w kolejnych etapach — bazę danych, która zostanie stworzona na potrzeby opisu, i aplikację Toposław, o której szerzej piszemy w p. 4.

3. Paradygmaty zwrotów — problemy teoretyczne

W opisie fleksji zwrotów można wyróżnić dwie podstawowe grupy jednostek: o paradygmatach pełnych i defektywnych czy też, ujmując problem z innej perspektywy, jednostki dziedziczące paradygmat głównego komponentu i jednostki o paradygmacie swoistym⁴. Trudno bez badania większej liczby jednostek stwierdzić, jaki jest stopień owej regularności (dziedziczenia fleksji) bądź nieregularności. Paradygmat głównego komponentu powielają np. zwroty: *stać z boku, trzymać coś (nerwy, uczucia, język) na wodzy, chuchać i dmuchać na coś, brać się z czymś za bary, nadać czemuś bieg, lecieć sobie z kimś w kulki czy być na bani*⁵. Ostatni z podanych zwrotów nie ma formy bezosobnika, podobnie jak *być*, pozostałe mają pełny paradygmat, oczywiście z różną frekwencją poszczególnych form.

Swoistość paradygmatu zwrotu, łącząca się z cechami znaczeniowo-funkcjonalnymi jednostek, może dotyczyć zarówno pojedynczych form, jak i kategorii. Stosunkowo najprostsze w opisie są ograniczenia ogólne — łączące się z wypełnieniem pozycji mianownika w schemacie tworzonym przez zwrot. W przypadku nieosobowego podmiotu, np. *coś ma ręce i nogi, coś dostało nóg*, można w zwykłych użyciach założyć brak kategorii osoby. Zwroty takie nie mają też bezosobnika (co jest ograniczeniem ogólnym, wynikającym ze specyfiki form z wykładnikiem *-no, -to*, z odniesienia do osobowego wykonawcy czynności), a tryb rozkazujący, odpowiadający bardziej życzeniom niż rozkazom, tworzony jest tylko opisowo — z elementem posiłkowym *niech*.

Ograniczenia typowo indywidualne reprezentuje na przykład jednostka z komponentem *dać* (*ktoś da komuś coś / ktoś da komuś coś robić*) ze zdań typu *Ja ci dam chodzić na jabłka do cudzego ogrodu!* (wyrażająca brak aprobaty dla czyjogoś zachowania i groźbę kary). Jednostka ta nie ma m.in. kategorii fleksyjnej czasu, jej paradygmat ogranicza się do form czasu przeszłego, np. *Ja ci dam bawić się naftą i zapalkami!* (NKJP), *Poczekaj tylko, ojciec ci da te nocne eskapady!* (Bogusławski, Danielewiczowa 2005). Jako przykład ograniczeń można też wskazać jednostkę *ktoś z niejednego pieca chleb jadł* ('ktoś jest doświadczony, zna życie'), na której paradygmat składają się tylko

⁴ Paradygmaty zwrotów i ich ograniczenia były omawiane szerzej w artykule I. Kosek (2013), w tym miejscu więc ograniczamy się do zasygnalizowania podstawowych problemów i typów defektywności.

⁵ Przykłady zwrotów omawianych w tym punkcie pochodzą z trzech źródeł: wcześniejszej pracy Kosek (2013), sondy słownikowej A. Bogusławskiego i M. Danielewiczowej (2005), NKJP oraz tekstów ogólnie dostępnych w Internecie.

formy czasu przeszłego, czy mającą szczątkowy paradygmat jednostkę *coś piechotą nie chodzi* (np. **Pięć stów piechotą nie chodziło*)⁶.

Analizując zakres fleksyjności zwrotów, często stajemy wobec relacji system — tekst: rzadkość występowania niektórych form powoduje wątpliwości co do ich statusu. Problem dotyczy zwłaszcza bezosobnika i imiesłowów przysłówkowych, por. np. *dano nogę, chodzono z brzuchem czy zjadłszy ich w kaszy*⁷. Dotykamy tu istotnego problemu wyznaczenia granicy między tym, co możliwe w języku, ale bardzo rzadko wykorzystywane przez użytkowników lub wymagające specjalnego kontekstu, a tym, co dewiacyjne.

Odrębny problem w opisie fleksji zwrotów to tryb rozkazujący. Zależnie od znaczenia całości mamy tu różne jego odcienie i ograniczenia użycia. Zasadnicze wątpliwości tryb rozkazujący budzi tam, gdzie wykonawca czynności nie ma wpływu na zaistnienie danego stanu rzeczy, np. *złapać pana Boga za nogi, mieć długie ręce, mieć smykałkę / dryg do czegoś, czuć do kogoś miętę, robić bokami*⁸. W innych wypadkach użycia typowe to użycia z przeczeniem, np. *szlifować bruki, jechać na podwójnym gazie*. Poprawnie użyty tryb bez przeczenia ma charakter ostrzeżenia, przestrogi, życzenia, ale nie rozkazu czy zakazu. W takich sytuacjach dobrym uzupełnieniem opisu gramatycznego byłyby uwagi pragmatyczne.

Z tego krótkiego przeglądu problemów paradygmatyki zwrotów wynika, że badania będą oscylować między zjawiskami indywidualnymi i powtarzalnymi. Na przykład ograniczenia wskazane dla zwrotu *ktoś da komuś coś / ktoś da komuś coś robić* mają charakter typowo indywidualny, problemy z formami trybu rozkazującego dotyczą natomiast całej grupy zwrotów i nie są charakterystyczne jedynie dla jednostek frazeologicznych (co pokazują dotychczasowe analizy czasowników, zob. np. Laskowski 1998). Jednym z istotnych elementów badań będzie sprawdzenie na podstawie większej liczby jednostek, w jakiej relacji pozostają wobec siebie oba typy zjawisk: czy w paradygmatyce zwrotów dominują ograniczenia dające się przypisać klasom obiektów czy zjawiska całkowicie indywidualne.

4. Opis zalgorytmizowany

W projektowanym słowniku frazeologicznym jednostek werbalnych planujemy wykorzystać Toposław — aplikację (program komputerowy), która posłużyła do tworzenia „Słownika elektronicznego jednostek frazeologicznych” (SEJF). Słownik ten zawiera opis paradygmatów przede wszystkim jednostek rzeczownikowych oraz przysłówkowych i przymiotnikowych. W odniesieniu do tej grupy frazeologizmów Toposław

⁶ Inne przykłady jednostek defektywnych wraz z bardziej szczegółowym ich omówieniem zob. w: Kosek 2013.

⁷ Przykładów formy *zjadłszy ich w kaszy* nie ma nawet w internetowej masie tekstów (w przeciwieństwie do np. *nie dając sobie w kaszę dmuchać*).

⁸ Problem ograniczeń w użyciu trybu rozkazującego w odniesieniu do leksemów opisywał m.in. R. Laskowski (1998).

spełnia swoje zadanie, weryfikacji natomiast wymaga jego przydatność do opisu zwrotów, charakteryzujących się dużo większym stopniem nieciągłości.

Toposław korzysta z osiągnięć zarówno lingwistów, jak i informatyków zainteresowanych rygorystycznym opisem polszczyzny. Program wykorzystuje analizator morfologiczny Morfeusz, którego bazę stanowi Słownik gramatyczny języka polskiego, z pewnością najobszerniejsza baza informacji fleksyjnej o polskich leksemach. W obecnej postaci SGJP zawiera charakterystykę gramatyczną ponad 250 tys. leksemów (w tym blisko 30 tys. leksemów czasownikowych), zarówno odnotowanych w słownikach, tekstach, jak i potencjalnych. 250 tys. leksemów daje ponad 4,5 miliona form fleksyjnych. Rozmiary i zawartość SGJP nie gwarantują co prawda znalezienia i zidentyfikowania wszystkich form, komponentów zwrotów, ale trudno wyobrazić sobie obecnie większą bazę danych morfologicznych języka polskiego. Konieczne natomiast wydają się dodatkowe prace inżynierskie nad oprogramowaniem, tak by możliwa była ingerencja leksykografa w generowany przez program paradygmat. Obecny stan wiedzy na temat frazeologizmów czasownikowych pozwala przewidzieć, że wybrane formy trzeba będzie zablokować, wykluczyć, nie dopuścić do ich utworzenia — zgodnie z wynikami analizy lingwistycznej.

Toposław związany jest z Narodowym Korpusem Języka Polskiego. W aplikacji stosuje się ten sam system znaczników morfosyntaktycznych (technicznie zwany *tagsetem*), którymi zostały anotowane formy w NKJP. Podstawową jednostką w *tagsecie* są fleksemy. Termin ten wprowadził Janusz S. Bień (1991). Przez fleksem rozumie się zbiór form jednolicie lub niemal jednolicie zróżnicowanych ze względu na właściwe im kategorie gramatyczne. Traktowanie jako jednostek opisu leksemów, nie leksemów, niesie z sobą konsekwencje zarówno teoretyczne, jak i techniczne. Na leksem czasownikowy właściwy składa się kilkanaście fleksemów, na czasownikowy niewłaściwy (w NKJP stosuje się na ich oznaczenie termin *predykatyw*) oraz czasowniki typu WINIEN – po jednym fleksemie, por.:

Leksem	Fleksem	Symbol	Przykład
Czasownik	forma nieprzeszła	fin	<i>jadam</i>
	forma przyszła czasownika BYĆ	bedzie	<i>będę</i>
	aglutynant czasownika BYĆ	aglt	<i>-śmy</i>
	pseudoimiesłów	praet	<i>jadał</i>
	rozkaznik	impt	<i>jadaj</i>
	bezosobnik	imps	<i>jadano</i>
	bezokolicznik	inf	<i>jadać</i>
	imiesłów przysłówkowy współczesny	pcon	<i>jadając</i>
	imiesłów przysłówkowy uprzedni	pant	<i>zjadłszy</i>
	odsłownik	ger	<i>jadanie</i>
	imiesłów przymiotnikowy czynny	pact	<i>jadający</i>
	imiesłów przymiotnikowy bierny	ppas	<i>jadany</i>
	Czasownik typu WINIEN (forma teraźniejsza)	winien	<i>winna, powinna</i>
Predykatyw	pred	<i>trzeba, słyhać</i>	

Tabela 1. Klasy fleksyjne wyodrębnione w NKJP⁹

⁹ Przedruk z: NKJP: 63.

Aktualna wersja aplikacji uniemożliwia wprowadzenie kompletnego opisu jednostki wielosegmentowej, której główny składnik stanowi leksem werbalny. Dlatego by kodować w Toposławie jednostki czasownikowe, konieczna jest zmiana obecnej hierarchii opisu — z fleksemów na leksemy.

Zamierzamy również korzystać z innych funkcjonujących narzędzi komputerowych, zawierających opis gramatyczny leksemów lub jednostek wielowyrazowych — Kuźni¹⁰ oraz Słownika walencyjnego Walenty¹¹. Kuźnia jest narzędziem zbiorczej bazy Słownika gramatycznego, umożliwiającym pracę zespołową w środowisku webowym. Sprzężenie Słownika gramatycznego oraz Słownika frazeologizmów wydaje się naturalną konsekwencją badań językoznawczych prowadzonych przez niezależne zespoły, dzielące się wynikami swoich prac. Jakkolwiek zakres opisu gramatycznego zawartego w Słowniku walencyjnym różni się od celu, który sobie stawiamy (opisu paradygmatów, nie łączliwości składniowej jednostek wielosegmentowych), to z oczywistych powodów nie możemy pominąć zawartości słownika Walenty. Budowanie nowych narzędzi i baz w oparciu o już istniejące sprzyja także transferowi nauki i praktycznemu jej wykorzystaniu w przyszłości.

Elektroniczny słownik frazeologiczny, mający stanowić wynik omawianego projektu, będzie korzystał z prac teoretycznych poświęconych opisowi gramatycznemu współczesnej polszczyzny, jak również wpisywał się w badania nad jej automatycznym przetwarzaniem. Można przypuszczać, że przyczyni się do precyzyjnego rozpoznawania w tekstach wystąpień jednostek wielocłonowych (frazeologicznych), co ma wpływ na:

- zmniejszenie stopnia wieloznaczności morfologicznej analizowanych wypowiedzi (a więc podniesienie jakości przetwarzania poziomu wyrazowego);
- identyfikację wariantów i powiązanie ich z formą kanoniczną znajdującą się w zasobach leksykalnych (zwiększenie efektywności systemów odpowiedzi automatycznej);
- dostarczenie początkowego zbioru terminów i wyrażeń pewnych, który zwiększy jakość systemów uczących się, a także systemów wydobywania wzorców;
- tworzenie wspólnych ram opisowych dla jednostek o podobnych właściwościach.

Scharakteryzowany w artykule projekt będzie stanowił kolejny krok zarówno w opisie gramatycznym związków frazeologicznych (w mało zaawansowanym dotychczas opisie frazeologii werbalnej), jak i w zalgorytmizowanym opisie języka polskiego.

¹⁰ <http://zil.ipipan.waw.pl/Ku%C5%BAnia>

¹¹ <http://zil.ipipan.waw.pl/Walenty>

Bibliografia

- B a ń k o M., 2013, Za mało czy dość? Dylemat leksykografa w dobie rewolucji informacyjnej (na przykładzie słowników frazeologicznych), [w:] Perspektywy współczesnej frazeologii polskiej. Między teorią a praktyką frazeograficzną, red. G. Dziarska-Lenart, J. Liberek, Poznań, s. 9–16.
- B ą b a S., 1996, Z fleksji polskich frazeologizmów. Cz. I: Uwagi o odmianie wyrażen, [w:] Synchroniczne i diachroniczne aspekty badań polszczyzny, t. III, red. M. Białoskórska, Szczecin, s. 141–154.
- 1997, Z fleksji polskich frazeologizmów. Cz. II: Uwagi o odmianie fraz, [w:] Z polszczyzny historycznej i współczesnej, red. T. Ampel, Rzeszów, s. 153–162.
- 1998a, Uwagi o odmianie polskich zwrotów przez czasy, Poznańskie Spotkania Językoznawcze, t. III, red. Z. Krążyńska, Z. Zagórski, Poznań, s. 7–14.
- 1998b, Materiały do bibliografii frazeologii polskiej. Prace opublikowane w latach 1945–1995, Poznań.
- 2003, Materiały do bibliografii frazeologii polskiej. Prace opublikowane w latach 1996–2000, Poznań.
- 2008, Materiały do bibliografii frazeologii polskiej. Prace opublikowane w latach 2001–2005, Poznań.
- B i e ń J.S., 1991, Koncepcja słownikowej informacji morfologicznej i jej komputerowej weryfikacji, t. 383 serii Rozprawy Uniwersytetu Warszawskiego, Warszawa, <http://bc.klf.uw.edu.pl/12/>.
- B o g u s ł a w s k i A., 1976, O zasadach rejestracji jednostek języka, *Poradnik Językowy* 8, s. 356–364.
- B o g u s ł a w s k i A., D a n i e l e w i c z o w a M., 2005, *Verba polona abscondita*. Sonda słownikowa II, Warszawa.
- C z e r e p o w i c k a M., 2006, Opis powierzchniowoskładniowy wyrażen niestandardowych typu *na lewo, do dziś, po trochu, na zawsze* we współczesnym języku polskim, Warszawa.
- 2011, Toposław jako narzędzie znakowania jednostek wielocłonowych, [w:] Nowa zjawiska w języku, tekście, komunikacji III, red. I. Matusiak-Kempa, S. Przybyszewski, Olsztyn, s. 28–35.
- C z e r e p o w i c k a M., S a v a r y A., 2012, SEJF, <http://zil.ipipan.waw.pl/SEJF>.
- K o s e k I., 2008, Fleksja i składnia nieciągłych imiennych jednostek leksykalnych, Olsztyn.
- 2013, Paradygmaty zwrotów frazeologicznych — problemy opisu leksykograficznego, [w:] Perspektywy współczesnej frazeologii polskiej. Między teorią a praktyką frazeograficzną, red. G. Dziarska-Lenart, J. Liberek, Poznań, s. 51–61.
- K o s e k I., C z e r e p o w i c k a M., 2011, Problemy opisu związków frazeologicznych w formalizmie „Multiflex”, [w:] Różne formy. Różne treści. Tom ofiarowany Profesorowi Markowi Świdzińskiemu, red. M. Bańko, D. Kopcińska, Warszawa, s. 117–125.
- L a s k o w s k i R., 1998, Semantyka trybu rozkazującego, *Polonica* XIX, s. 5–29.
- L e w i c k i A.M., 1976, Wprowadzenie do frazeologii syntaktycznej, teoria zwrotu frazeologicznego, Katowice.
- 1982, O motywacji frazeologizmów, Z problemów frazeologii polskiej i słowiańskiej I, red. M. Basaj, D. Rytel, Wrocław, s. 33–47.
- 1986, Składnia związków frazeologicznych, *Biuletyn PTJ* XL, s. 75–82.
- L i b e r e k J., 2006, Zróżnicowanie aspektowe komponentów werbalnych w stałych związkach wyrazowych. Uwagi wstępne, [w:] Z zagadnień frazeologii, stylistyki i kultury języka, red. S. Bąba, P. Fliciński, Poznań, s. 44–66.
- NKJP: Narodowy Korpus Języka Polskiego, 2012, red. A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk, Warszawa, <http://www.nkjp.pl/index.php?page=3&lang=0>.
- S a l o n i Z., 2000, Wstęp do koniugacji polskiej, Olsztyn.
- 2001, *Czasownik polski*. Odmiana, słownik, Warszawa.
- S a l o n i Z., W o l i ń s k i M., W o ł o s z R., G r u s z c z y ń s k i W., S k o w r o ń s k a D., 2012, *Słownik gramatyczny języka polskiego*, wyd. II, b. w.

SUMMARY

On the project of digital inflectional dictionary of verbal phraseological units

Key words: verbal phraseology, paradigm, computational linguistics.

Słowa kluczowe: frazeologia werbalna, paradygmat, lingwistyka komputerowa.

In the paper authors present a research project defined as an inflection description of Polish verbal phraseological units and its implementation. In the first three points the main aims and methodological assumptions are discussed. The latest point applies to computer applications needed to the formal description of Polish verbal phraseology.

Analysis of the verbal Polish multi-word units as well as a dictionary comprising 5000 analysed units are the main aim of the project. The fundamental problem in the linguistic description is an answer to questions: How regular is a paradigm of verbal multi-word unit? How does it depend on the main segment of the unit? How is it determined in semantical and syntactical way? The project also assumes a development of available application, Toposław, used to formal description of Polish.