

## Opolska idiomatyka w wersji cyfrowej. Wokół słownika *Lidé – Média – Komunikace*

**Keywords:** lexicography, phraseography, transformability of dictionaries, online dictionary, reproduct, translational triads

**Słowa kluczowe:** leksykografia, frazeografia, przekształcalność słowników, słownik online, reprodukt, triady przekładowe

### Abstract

This article presents the process of creating a digital dictionary of translation triads in a database version (accessible through a search engine on the WWW). The author conducts discussion on practical and technical, theoretical, interlingual, as well as metalexicographical problems, related to the procedure described here, that is, the issue of – more generally – the transformability of dictionaries. As a theoretical basis and source of lexicographic practice in relation to the activities described here, the author points to a lexicographic tradition he calls *Opole idiomatics*, which is associated with the figure of Wojciech Chlebda, among others, the creator of phrasematics or the concept of the lexicographic idiomaton.

Niniejszy artykuł przedstawia proces powstawania cyfrowego słownika triad przekładowych w wersji bazodanowej (dostępnej poprzez wyszukiwarkę na stronie WWW). Autor przeprowadza dyskusję na temat problemów natury praktyczno-technicznej, teoretycznej, interlingwalnej, a także metaleksykograficznej, związanych z opisywaną tu procedurą, czyli kwestii – ogólniej mówiąc – przekształcalności słowników. Jako podstawę teoretyczną oraz źródło praktyki leksykograficznej względem opisywanych tu działań wskazuje autor tradycję leksykograficzną, którą nazywa *opolską idiomatyką*, a która związana jest z postacią Wojciecha Chlebdy, między innymi twórcy frazematyki czy koncepcji idiomatikonu leksykograficznego.

### Rys historyczny

Sformułowanie zawarte w pierwszych słowach tytułu niniejszego tekstu można rozumieć co najmniej dwojako:

1. W ujęciu teoretycznym *opolska idiomatyka* to nurt w badaniach frazeologicznych sięgający lat 70–80 XX wieku. Wojciech Chlebda, twórca opolskiej idiomatyki, w pierwszej kolejności rozumianej tu jako opolska tradycja frazeologiczna, odwołuje się w swoich rozważaniach między innymi do koncepcji jednostek języka Andrzeja Bogusławskiego (Bogusławski 1976) czy koncepcji jednostek leksykalnych Macieja Grochowskiego (Grochowski 1981). W latach 80 badacz czyni pierwsze spostrzeżenia dotyczące frazematyki (Chlebda 1985), które w roku 1991 formalizują się na kartach najbardziej znanej jego monografii (Chlebda 1991)<sup>1</sup>.
2. W ujęciu praktycznym *opolska idiomatyka* związana jest z działaniami w obszarze leksykografii, a dokładniej frazeografii<sup>2</sup>. Działania te nawiązują wprawdzie do pojęcia *frazem*, czyli przedmiotu rozważań teoretycznych wspomnianej w punkcie 1 frazematyki, aczkolwiek zorientowane są przede wszystkim na traktowanie frazemu jako obiektu opisu słownikowego. Warto dodać, iż prace opolskiego zespołu leksykograficznego były motywowane postulatem Andrzeja Bogusławskiego, dotyczącego potrzeby wyodrębnienia z tekstów (i kodyfikacji) jednostek języka (Bogusławski 1976, s. 357).

Z powodów, o których nie będziemy tu mówić, pojęcie *frazem* około roku 2010 zostało zastąpione w idiolekcie opolskich słownikarzy pojęciem *reprodukt*<sup>3</sup>. To właśnie reprodukt języka polskiego wraz z ich odpowiednikami przekładowymi w języku rosyjskim za-

---

<sup>1</sup> Por. rozszerzone wydanie tej publikacji, Chlebda W., *Elementy frazematyki. Wprowadzenie do frazeologii nadawcy*, Łask 2003. Spostrzeżenia te dotyczą głównie badania języka przez pryzmat osoby mówiącej.

<sup>2</sup> Rok 2006 to data wydania pierwszego zeszytu *Podręcznego idiomatikonu polsko-rosyjskiego* pod redakcją Wojciecha Chlebdy. Zeszyt ów rozpoczął serię 15 publikacji (12 słownikowych oraz 3 o charakterze teoretycznym) określanych tu przez nas właśnie mianem *opolskiej idiomatyki* w drugim – szerszym – ujęciu.

<sup>3</sup> Kwestie te szczegółowo omawia Wojciech Chlebda (Chlebda 2009, s. 11–26). Warto przy tym podkreślić, że zarówno *frazem*, jak i *reprodukt*, w założeniu twórcy

pełniły karty dziesięciu zeszytów *Podręcznego idiomatikonu polsko-rosyjskiego* (dalej *Idiomatikon* lub PIPR), a także dwóch tomów *Polsko-rosyjskiego słownika par przekładowych* (dalej PRSPP), wydawanych w latach 2006–2023. Przy okazji prac nad scalaniem pierwszych pięciu zeszytów *Idiomatikonu* i redagowaniem pierwszego tomu zbiorczego PRSPP pojawiła się idea cyfryzacji słownika, jednak nie doczekała się gruntownego opracowania<sup>4</sup>.

W roku 2013 opolski zespół nawiązał współpracę z pracownikami ówczesnej sekcji filologii polskiej Katedry Sławistyki na Wydziale Filozoficznym Uniwersytetu Palackiego w Ołomuńcu<sup>5</sup>. Owocem wspólnych prac był między innymi *Tematický česko-polsko-ruský, polsko-česko-ruský slovník pohraničí*, wydany dwa lata później. Publikacja ta bazowała na koncepcji znanej z PIPR oraz PRSPP. Po pięciu latach oba zespoły ponownie połączyły siły, publikując kolejny słownik, tym razem w dwóch wariantach: w formacie PDF oraz w wersji bazodanowej na stronie [idiomatikon.eu](http://idiomatikon.eu)<sup>6</sup>.

Powyższy rys historyczny stanowi niezbędne tło dla dalszych rozważań, które dotyczyć będą struktury słownika *Lidé – Média – Komunikace* (dalej LMK) w wersji online, procesu jego powstawania

---

obu tych pojęć, miały stanowić swoisty hiperonim dla wszystkich całości semantycznych (również jednowyrazowych) o charakterze frazeologicznym, idiomatycznym (Chlebda 2010, s. 8). Pojęciem wyjściowym wobec *reproduktu* jest *jednostka języka* przywołana w punkcie 2.

<sup>4</sup> Jest to istotne z dwóch względów. Podczas prac redakcyjnych obu tomów PRSPP wykrystalizowała się pewna metodologia digitalizacji, a także – co ważniejsze, choć głównie w odniesieniu do drugiego tomu zbiorczego – istniała surowa wersja materiału słownikowego właściwie gotowa do dalszej przeróbki i ewentualnej publikacji w wersji online (mamy tu na myśli wersję bazodanową).

<sup>5</sup> W ramach projektu *Uniwersytety na rzecz pogranicza*, finansowanego z unijnego Programu Operacyjnego Współpracy Transgranicznej Republika Czeska-Rzeczpospolita Polska, realizowanego w latach 2013–2015 (nr CZ.3.22/2.3.00/12.03352).

<sup>6</sup> W ramach projektu *Czesko-polsko-rosyjski idiomatikon*, współfinansowanego przez Agencję Technologiczną Republiki Czeskiej w ramach programu Éta (nr TL03000467).

i działania, a także problemów natury praktyczno-leksykograficznej i metaleksykograficznej, z tym wszystkim związanym.

## Struktura słownika LMK<sup>7</sup>

Struktura słownika obejmuje pięć poziomów – od najbardziej ogólnego (zakres) po poziom konkretnego materiału językowego (jednostki języka / reprodukt). Zakres słownika wyznacza jego tytuł; to ludzie, media, komunikacja. Zakres słownika realizowany jest w pięciu ogólnych obszarach tematycznych:

- I. Dzisiejsze media analogowe i cyfrowe,
- II. Komunikacja międzyludzka w dzisiejszych mediach,
- III. Media jako pośrednik w interesach między ludźmi,
- IV. Zagrożenia i bezpieczeństwo w Sieci,
- V. Polityka komunikacyjna Unii Europejskiej.

Każdy obszar tematyczny składa się z określonej liczby uszczegóławiających go pól. Na przykład obszar *I. Dzisiejsze media analogowe i cyfrowe* składa się z trzech pól:

1. Techniczna baza komunikacji analogowej i cyfrowej. Tradycyjny i współczesny sprzęt i osprzęt komunikacji.
2. Używanie środków komunikacji ogólnej, analogowej i cyfrowej. Działania ze sprzętem i osprzętem.
3. Napisy w przestrzeni komunikacyjnej.

---

<sup>7</sup> W całym tekście będziemy stosować skróconą wersję nazwy słownika. Jego pełna nazwa jest dwujęzyczna i brzmi *Lidé – Média – Komunikace. Česko-polsko-ruský a polsko-česko-ruský idiomatikon / Ludzie – Media – Komunikacja. Česko-polsko-rosyjski i polsko-czesko-rosyjski idiomatikon* (publikacja dostępna jest pod adresem <https://bg.uni.opole.pl/wp-content/uploads/LMK-ISBN-978-83-8332-051-9.pdf>). Nazwa ta jest w pewnym sensie problematyczna, bo w zasadzie dotyczy jedynie wariantu PDF, w którym językiem wiodącym jest albo język czeski, albo język polski. W wersji online po lewej stronie siatki słownikowej może stać również język rosyjski. O kwestiach tych będzie jeszcze mowa.

Każde pole tematyczne składa się z określonej liczby jeszcze bardziej szczegółowych działek tematycznych (na przykład pole 1 obszaru I składa się z takich szczegółowych działek, jak Prasa, Poczta, Telefony, Telewizja itp.).

Działki tematyczne wypełnione są konkretnym materiałem językowym – jednostkami języka (reproduktami), które stanowią hasła słownika LMK, oraz ich ekwiwalentami czeskimi i rosyjskimi. Nazwy działek w słowniku podane są tylko przy jednostkach, przy których jest to niezbędna informacja, pozwalająca na precyzyjną identyfikację tematyczną tych jednostek<sup>8</sup>.

### Proces powstawania słownika

Słownik LMK powstawał w dwóch podstawowych etapach: tworzenia siatki słownikowej i przetwarzania jej do wersji cyfrowej<sup>9</sup>. Siatka słownikowa była tworzona w kilku subetapach, na które składały się kolejno:

- 1) układanie hasłownika w języku polskim,
- 2) ustalanie ekwiwalentów przekładowych w języku rosyjskim,
- 3) ustalanie ekwiwalentów przekładowych w języku czeskim.

Na tym etapie prac uzyskaliśmy wersję siatki słownikowej, którą obrazuje poniższy jej fragment (trzy ostatnie kolumny dotyczą wspomnianych wcześniej obszarów, pól i działek, oznaczonych tutaj kolejno literami O, P i D):

<sup>8</sup> Podsumowując, obszary oznaczamy cyframi rzymskimi, pola tematyczne oraz działki – arabskimi. Szczegółowa struktura słownika dostępna jest zarówno w publikacji w formacie PDF (zob. przypis 7), jak i na stronie internetowej słownika pod adresem [idiomatykon.eu/dictionary](http://idiomatykon.eu/dictionary).

<sup>9</sup> W związku z tym, że układanie siatki słownikowej odbywało się w edytorze tekstu, na komputerze, i tę wersję można by uznać za cyfrową (tym bardziej, że w efekcie końcowym słownik ten został opublikowany również w formie PDF-u), niemniej jednak przymiotnik ten zarezerwujemy – dla uproszczenia – dla wersji bazodanowej, dostępnej pod adresem [idiomatykon.eu/search](http://idiomatykon.eu/search).

Tabela 1. Fragment siatki słownikowej w układzie trójkolumnowym

Język polski	Język rosyjski	Język czeski	O	P	D
asymilacja językowa	языковая ассимиляция	jazyková asimilace	V	2	–
asymilacja kogo w jakiej kulturze	ассимиляция кого в какую культуру	asimilace koho do jaké kultury	V	2	–
asymilować się; kto asymiluje się do jakich warunków	кто ассимилируется к каким условиям	asimilovat se; kdo se asimiluje do jakého prostředí	V	2	–
atlas językowy	языковой атлас	jazykový atlas	V	2	–
atlas języków jakich	атлас языков каких	atlas jazyků jakých	V	2	–

Jak widać, słownik LMK nawiązuje do zeszytów PIPR (czy też tomów PRSPP) również na poziomie formatowania jednostek słownikowych, co swoją drogą przysporzyło nam wielu problemów podczas przetwarzania materiału tekstowego w taki sposób, aby to formatowanie zachować w wersji cyfrowej. Proces ten wymagał obróbki półautomatycznej, z wykorzystaniem wyrażeń regularnych (w edytorze tekstu LibreOffice Writer), a także w pełni zautomatyzowanej (w ramach działania przygotowanego przez nas skryptu w języku programowania Python)<sup>10</sup>. Dla lepszego zobrazowania napotkanych trudności poniżej prezentujemy najbardziej złożone w tym aspekcie hasła (w różnych językach), dzięki czemu w sposób maksymalnie uproszczony przybliżymy etapy przetwarzania słownika bez wchodzenia w zbędne dla niniejszego wywodu szczegóły programistyczne.

[1] aktualizować: *kto aktualizuje informacje o kim / czym gdzie*

[2] общаться: *kto общается с кем по „Скайпу” / скайпу*

[3] vytvořit: *kdo si vytvořil účet kde*

<sup>10</sup> Piszący te słowa odpowiada w zasadzie za cały proces cyfryzacji materiału słownikowego (również w ujęciu programistycznym), a także projekt i wykonanie strony internetowej słownika, w tym interfejsu użytkownika, projekt i wykonanie bazy danych słownika oraz wykonanie skryptu przeszukującego bazę słownika. Ostateczne decyzje co do sposobu wyświetlania wyników kwerend słownikowych były omawiane zespołowo. Wykaz członków zespołu dostępny jest pod adresem [idiomatykon.eu/team](http://idiomatykon.eu/team).

[4] **FTL** [ang. *full truck load*, ładunek całopojazdowy]

[5] **Úřad pro ochranu osobních údajů** Cz. [ÚOOÚ]

Zarówno przykłady z tabeli 1, jak i powyższa lista pozwolą nam przypomnieć konwencję opisu leksykograficznego zaczerpniętą z PIPR i z PRSPP, którą z pewnymi modyfikacjami przyjęliśmy w słowniku LMK. W odniesieniu do formatowania tekstu konwencja ta obejmuje:

- 1) tekst jasny dla tzw. wyrazów naprowadzających (zakończone dwukropkiem; stanowią środek techniczny służący sortowaniu alfabetycznemu) oraz tzw. mikrodefinicji (ujęte w nawiasy kwadratowe; stanowią dodatkowe informacje niezbędne do prawidłowego odczytania sensu odpowiedników przekładowych lub wariant opisowy w wypadku ekwiwalentu zerowego<sup>11</sup>),
- 2) tekst pochylony dla sygnałów zaimkowych, wyrażen obcojęzycznych oraz kwalifikatorów stylistycznych<sup>12</sup>,
- 3) tekst pogrubiony dla oznaczenia komponentów obligatoryjnych,
- 4) nawiasy ostrokątne dla wyodrębnienia komponentów fakultatywnych,
- 5) ukośnik dla oznaczenia wariantów ortograficznych, akcentowych (w języku rosyjskim) oraz wariantów sygnałów zaimkowych.

W odniesieniu do znaków niealfabetycznych (symboli) kwestia przetwarzania komputerowego była nieco bardziej złożona i wiązała się potrzebą segmentacji takich całości<sup>13</sup>, jak [1]–[5] na dwóch poziomach. Po pierwsze, segmentacji na poziomie zewnętrznym, co oznacza oddzielenie wyrazów naprowadzających od części zasadniczej jednostki słownikowej, a także od jej mikrodefinicji<sup>14</sup>. Po drugie,

<sup>11</sup> Będzie o tym mowa w części *Przekształcalność słowników*.

<sup>12</sup> W LMK stosowane są jedynie kwalifikatory odpowiedników funkcjonalnych Cz., Pol. i Ros., por. (Chlebda, Dobrotová 2023, s. 373).

<sup>13</sup> Sformułowanie *całość* będziemy stosować zamiennie ze sformulowaniem *jednostka słownikowa* bądź – krócej – *jednostka*.

<sup>14</sup> Wyrazy naprowadzające oraz mikrodefinicje opolscy słownikarze stosują od pierwszego tomu zbiorczego PRSPP oraz od szóstego zeszytu PIPR, co redaktor serii opisał we wstępie do PRSPP (Chlebda 2014, s. 32–34, 45–49).

segmentacji na poziomie wewnętrznym, co oznacza podział segmentów wyższego rzędu na pojedyncze wyrazy graficzne. W tym miejscu proces odbywał się dwutorowo: jedna kopia jednostki była przetwarzana do postaci wyświetlanej, druga z kolei – sprowadzana do postaci odmienionej, czyli bez formatowania i z pominięciem wszelkich symboli, a ta – w następnym kroku – do postaci podstawowej (zlematyzowanej). Ostatnim etapem było nadanie jednostce słownikowej numeru identyfikacyjnego, przyporządkowanie jej do odpowiedniego obszaru, pola tematycznego oraz działki<sup>15</sup>, a także umieszczenie poszczególnych jej segmentów w odpowiednich polach w bazie danych. Całą procedurę, a tym samym strukturę bazy danych słownika, obrazuje tabela 2.

**Tabela 2.** Uproszczona procedura komputerowego przetwarzania słownika

Nazwa pola bazy danych	Efekt procesu zawartość pola bazy danych
Postać wyjściowa <sup>16</sup>	aktualizować: <i>kto</i> aktualizuje informacje <o</i> <i>kim</i> / <i>czym</i>> <i>gdzie</i>
Wyraz naprowadzający	aktualizować
Postać wyświetlana <sup>17</sup>	<i>kto</i> <b>aktualizuje</b> <b>informacje</b> <b>o</b> <i>kim</i> / <i>czym</i>> <i>gdzie</i>
Mikrodefinicja	–
Postać odmieniona	kto aktualizuje informacje o kim czym gdzie
Postać zlematyzowana	kto aktualizować informacja o kto co gdzie
Obszar	II
Pole	I
Działka	–

<sup>15</sup> Unikalny numer identyfikacyjny (liczba porządkowa) umożliwiał wyświetlenie pełnej triady przekładowej na podstawie zapytania dotyczącego tylko jednego języka.

<sup>16</sup> Otrzymana po etapie obróbki półautomatycznej. Jak widać postać ta zawiera znacznik języka HTML oznaczający kursywę (<i></i>).

<sup>17</sup> Postać ta w porównaniu do postaci wyjściowej (por. przypis 15) uwzględnia dodatkowy znacznik języka HTML odpowiadający za pogrubienie (<b></b>).

Komentarz <sup>18</sup>	–
Informacje dodatkowe	–
Numer identyfikacyjny	2534

## Działanie słownika LMK

Słownik online został wyposażony w wyszukiwarkę typu korpusowego, oczywiście znacznie mniej złożoną i oferującą ograniczone możliwości budowania zapytań. Podstawowe informacje o jej funkcjonowaniu przedstawiamy w formie listy:

- Hasła w bazie słownika występują w trzech postaciach:
  - wyświetlanej: blokovat: **kdo si blokuje čí číslo v telefonu**,
  - odmienionej: kdo si blokuje čí číslo v telefonu,
  - podstawowej: kdo blokovat si čí číslo v telefon.
- Niektóre hasła zawierają wyrazy naprowadzające, po których następuje dwukropek, np.: brát: **kdo si bere auto na leasing**.
- Niektóre hasła zawierają mikrodefinicje umieszczone w nawiasach kwadratowych, np.: **dlouhé vlny** [rádiové].
- Wyszukiwarka przeszukuje hasła w wariacie odmienionym oraz podstawowym (1.b–c), wyrazy naprowadzające (2), a także mikrodefinicje (3).
- Wyszukiwarka ignoruje wielkie litery.
- Zapytaniem mogą być pojedyncze wyrazy lub frazy.
- W zapytaniach można stosować symbol kropki (.) oraz gwiazdki (\*); symbol kropki zastępuje 1 lub więcej znaków alfabetu; symbol gwiazdki zastępuje 1 lub więcej dowolnych znaków, np.: „kdo . co” oznacza „kdo” + 1 lub więcej znaków alfabetu + „co”; „kdo \* co” oznacza „kdo” + 1 lub więcej dowolnych znaków (np. spacji) + „co”

<sup>18</sup> Pola *Komentarz* i *Informacje dodatkowe* zostały stworzone niejako na wszelki wypadek, jeśli w którymś momencie zaistniałaby potrzeba uzasadnienia takiego, a nie innego kształtu danej triady przekładowej czy jednego z jej komponentów.

- Wyrazy oraz połączenia wyrazowe z łącznikiem w bazie słownika zapisane są ze spacją i w takiej postaci powinny być wyszukiwane, np.: *e-mail* jako *e mail*, *czarno-biały* jako *czarno biały* itp.

Taki sposób zapisu pozwala na szybsze odnalezienie połączeń typu *e-mail*, *e-przelew*, *e-zwolnienie lekarskie* (dla podanych przykładów wystarczy zapytanie o postaci „e”), lecz również pozwala na odszukanie haseł z łącznikiem poprzez wpisanie tylko jednego ich członu, np. zapytanie „rozliczeniowy” wyświetli hasła zawierające połączenie *oszczędnościowo-rozliczeniowy*.

- Nie jest możliwe łączenie w zapytaniu form zlematyzowanych z formami odmienionymi, np. zapytanie o postaci „hledat rozhlasovou” jest błędne (powinno przyjąć postać „hledat rozhlasova” albo „hledá rozhlasovou”). Rzecz ta nie dotyczy nieodmiennych części mowy.

Przeszukiwanie słownika możliwe jest dzięki prostemu interfejsowi użytkownika, na który składają się przedstawione na poniższym obrazie pola: pole zapytania, pole wyboru języka zapytania, pole wyboru układu wyświetlanych wyników, pole wyboru obszaru (po wybraniu jednego z nich w interfejsie pojawia się pole wyboru pól tematycznych dotyczących danego obszaru), przycisk wyszukiwania.

Obraz 1. Interfejs użytkownika

Wyniki wyszukiwania wyświetlane są w dwu wariantach: jako lista (odpowiedniki przekładowe podawane są jeden pod drugim, a kompletne triady przekładowe wieńczy pozioma linia), a także jako tabela (odpowiedniki przekładowe podawane są w kolumnach, każda triada zajmuje osobny wiersz tabeli). Dodatkowo w obu układach jednostki słownikowe oznaczone są ikonką flagi narodowej. W układzie listy obszar, pole i działka, do którego należy dana triada, podawane są w postaci tekstowej pod ostatnią jednostką triady przekładowej, w układzie tabelarycznym natomiast – w postaci numerycznej w dodatkowych kolumnach po prawej stronie (por. tabela 1).

### Przekształcalność słowników

W tym miejscu przystąpimy do omówienia triad przekładowych w ujęciu interlingwalnym, praktyczno-leksykograficznym oraz meta-leksykograficznym. Sądzimy, że możliwe jest rozpatrywanie materiału słownikowego przez pryzmat ogólniejszej kategorii, którą Wojciech Chlebda nazywał „przekształcalnością słowników” (szczegółowy opis patrz: Chlebda 2013 i 2014). Kwestie te – międzyjęzykowe i techniczno-leksykograficzne – w zasadzie przeplatają się wzajemnie – w różnym stopniu i w różnych układach, dotyczą bowiem nie tylko faktu zestawienia ze sobą reproduktów trzech języków, lecz także wszystkich praktycznych i teoretycznych konsekwencji takiego zestawienia. Mogłoby się wydawać, że proces przekształcenia słownika tematycznego (PIPR) w słownik par przekładowych (PRSP) wyczerpuje te zagadnienia, jak się jednak okazało, proces digitalizacji słownika par przekładowych (por. tabela 1) przyniósł inny (bądź też poszerzony) zestaw problemów tej kategorii<sup>19</sup>.

<sup>19</sup> Przekształcalność słowników nie dotyczy zresztą jedynie PIPR i PRSP. Swego czasu Piotr Wierchoń dokonał przekształcenia *Wielkiego słownika rosyjsko-polskiego z kluczem polsko-rosyjskim* (Wawrzyńczyk, red., 2004) w *Indeks a tergo jednostek przekładowych Wielkiego słownika rosyjsko-polskiego z kluczem polsko-rosyjskim* (Wierchoń 2007). Znane są również inne działania, być może niebezpiecznie związane z przekształcalnością słowników, lecz w różnym stopniu

Słownik LMK w zasadzie „nie narzuca” użytkownikom konkretnego języka wyjściowego. Podobnie zresztą, jak w niektórych korpusach równoległych, język wyjściowy jest tutaj umowny, choć w istocie był nim język polski (dla którego ustalano ekwiwalenty kolejno rosyjsko- i czeskojęzyczne). Mamy zatem do czynienia ze swoistą rozbieżnością pomiędzy faktycznym kierunkiem ustalania odpowiedników przekładowych (właściwym zresztą ogólnie przyjętej praktyce przekładowej), a tym kierunkiem, jaki może uzyskać użytkownik słownika wybierając jeden z trzech języków zapytania. Fakt ten prowadzi do pewnych problemów technicznych. Słownik LMK dopuszcza istnienie tzw. ekwiwalencji zerowej (oznaczanej symbolem  $\emptyset$ )<sup>20</sup>, co za tym idzie dla niektórych haseł polskojęzycznych albo po stronie czeskiej, albo po stronie rosyjskiej, albo w obu tych wypadkach nie istnieje odpowiednik przekładowy danej jednostki<sup>21</sup>. Oznacza to, iż wypadki te powinny zostać warunkowo wykluczone z wyników wyszukiwania, jeśli użytkownik słownika za język wyjściowy przyjmuje język czeski lub język rosyjski<sup>22</sup>. Rozpatrzmy poniższy przykład z obszaru II (pole 3, *Etykieta w komunikacji realnej i wirtualnej*).

---

odnoszące się do ucyfrowienia materiału tekstowo-słownikowego, jego przekształcenia w taki sposób, aby powstał konstrukt innego typu powiązań niż w tradycyjnej leksykografii; por. choćby *Francusko-polski słownik (graficznych) odpowiedników przekładowych* (Bralewski 2013) lub *Hipersłownik – Гиперсловарь – Hyperdictionary* (Wawrzyńczyk 2018), broszurka zapowiadająca wydanie *Hipersłownika języka polskiego* (Wawrzyńczyk 2020; wymieniamy tylko jeden z wielu tomów).

<sup>20</sup> Więcej na ten temat, por. (Gasek 2019, s. 2–5). Ogólnie rzecz ujmując, ekwiwalencja zerowa oznacza tu brak relacji „reprodukt w języku wyjściowym” – „reprodukt w języku docelowym”.

<sup>21</sup> Swoją drogą również „strona” jest tu, by tak rzec, komponentem ruchomym, co więcej, w układzie listy w ogóle trudno mówić o „stronie” (w sensie wizualnym).

<sup>22</sup> Pozycje z tzw. zerem usunięto również z czesko-polsko-rosyjskiej części publikacji wspomnianej w przypisie 7.

**Tabela 3.** Translady zerowe a mikrodefinicje

Język polski	Język czeski	Język rosyjski
<b>Panie doktorze!</b> [na uczelni]	<b>Pane doktore!</b> [na vysoké škole]	☉ [обращённые к мужчине кандидату наук, имя + отчество Ros.]

W wypadkach z ekwiwalentem (translatem) zerowym słownik zawsze podaje jakiś rodzaj „odpowiednika” w nawiasach kwadratowych, czasem o charakterze informacyjnym, jak w powyższym przykładzie, czasem – przekładowym (np. kalkę translandu, czyli jednostki języka wyjściowego). W języku rosyjskim wzorec „imię + nazwisko” obejmuje wiele zwrotów adresatywnych realizowanych w języku polskim w ramach innego wzorca „Panie/Pani + zawód/stanowisko/funkcja”, a zatem nie istnieją takie reproduktury języka rosyjskiego, które można by uznać za leksykalne odpowiedniki systemowej językowe danych reproduktów języka polskiego („odpowiednikiem” będzie tu właśnie wzorec komponentów „do podstawienia”). Jak wspomnieliśmy wyżej, nasza wyszukiwarka nie dopuszcza zapytań o postaci ☉, jednak identyfikuje triady przekładowe poprzez wyrazy naprowadzające lub mikrodefinicje, co może doprowadzić do wyświetlenia t a k i c h wyników w układzie czesko-polskim, wśród których będą jednostki czeskie z zerem. Dla relacji polsko-czeskiej (przypomnijmy, że językiem wyjściowym przy ustalaniu odpowiedniości przekładowych był tutaj zawsze język polski) w bazie słownika odnotowujemy 34 takie przypadki. Rozpatrzmy kilka z nich.

**Tabela 4.** Wybrane pary polsko-czeskie (w układzie czesko-polskim) z translatem zerowym

Język czeski	Język polski
☉ [2. část zázilky]	<b>przesyłka uzupełniająca</b>
☉ [zázilka číslo 2]	<b>przesyłka uzupełniająca</b>
☉ [patologický streaming]	<b>patostreaming</b>
☉ [komerční platforma elektronických veřejných služeb]	<b>Platforma e-Uслуг Publicznych</b>

☉ [program pro rodičovskou kontrolu na internetu]	<b>Stražník Učnia</b> [program do kontroli rodičovské]
☉ [Centrální registr řidičů a Registr silničních vozidel]	<b>Centralny System Ewidencji Pojazdów i Kierowców</b> [CEPiK]

Przesyłka uzupełniająca w języku polskim funkcjonuje jako swego rodzaju termin z obszaru sprzedaży internetowej; zespół czeski nie odnalazł podobnego terminu w analogicznych tekstach języka docelowego i opatrzył go komentarzem objaśniającym (również po stronie rosyjskiej mamy tu translat zerowy)<sup>23</sup>. Co warto odnotowania, w tym konkretnym przypadku ekwiwalent zerowy posiada dwa warianty, które w nieco inny sposób precyzują, jakie znaczenie realizuje reprodukt języka polskiego. Swego rodzaju dookreślenie danej sytuacji przekładowej stanowi informacja o przynależności jednostki *przesyłka uzupełniająca* do obszaru III (*Média jako prostředník ve vztazích mezi lidmi*), pola 2 (*Obchodování v síti*), działki 1 (*Internetový obchod*).

*Patostreaming* z kolei, zjawisko popularne w Polsce, posiada odpowiednik funkcjonalny w języku rosyjskim, nie odnotowano go jednak w czeskojęzycznym dyskursie internetowym<sup>24</sup>. W wypadku nazw usług, programów czy też instytucji, charakterystycznych tylko

<sup>23</sup> W zasadzie od rozpoczęcia prac na zeszytach PIPR zespół autorski za podstawowe źródło ekwiwalentów przyjmował teksty (najczęściej internetowe, rzadziej drukowane), nierzadko także tzw. dwuteksty. Kwestie traktowania zasobów Internetu jako swego rodzaju korpus referencyjny podnoszone były między innymi w Kilgarrif, Grefenstette 2003. Problematykę wykorzystania dwutekstów (szerzej – tekstów zespolonych) do celów słownikarskich opisano w Borysowski 2020. W tym miejscu warto zwrócić uwagę jedynie na fragment rozważań piszącego te słowa na temat pojęcia *tekstowe pary przekładowe* (Borysowski 2020, s. 97–98), które w pewnym stopniu wyjaśnia właśnie tekstową naturę ekwiwalencji zerowej (czy też w jakimś stopniu ją uzasadniają) – czy przyjmujemy za jej źródło teksty zespolone, czy dwu- lub wielojęzyczne teksty porównywalne (w nawiązaniu do pojęcia *korpus porównywalny/porównywalny zasób Internetu*), czyli teksty w dwu- lub więcej językach o zbliżonej bądź identycznej tematyce.

<sup>24</sup> Pytanie „dlaczego?” nie mieści w ramach niniejszego tekstu. Możliwe, że fakt ten jest w jakiś sposób uwarunkowany kulturowo. Wyszukiwarki internetowe nie odnajdują także poświadczeń wariantu opisowego, ujętego tu w nawiasy kwadratowe.

dla realiów jednego kraju słownik LMK oferuje dwa rozwiązania. Pierwsze omówiliśmy pod tabelą 3; dotyczy ono mikrodefinicji. Obrazują to polskie reprodukt *Platforma e-usług Publicznych* oraz program *Strażnik Ucznia*, dla których nie odnaleziono nawet odpowiedników funkcjonalnych. Nieco inaczej sprawa ma się z ostatnim reproduktom z tabeli 4. W ramach czeskiego ministerstwa (Ministerstwo doprawy ČR) istnieją dwa oddzielne rejestry: Centrální registr řidičů oraz Registr silničních vozidel. W związku z tym nie ustalono relacji „reprodukt języka polskiego” – „reprodukt języka czeskiego”, ponieważ nazwa polska jest całością, podczas gdy dwie nazwy czeskie zostały zestawione w ramach mikrodefinicji niejako doraźnie, przez autorów części językowej części siatki słownika. Drugie rozwiązanie polega na podaniu przy odpowiednikach przekładowych kwalifikatorów *Pol.*, *Cz.*, *Ros.*, podkreślających ich funkcjonowanie tylko w realiach polskich, czeskich lub rosyjskich. Oznacza to, że dana usługa, program, instytucja itp., obejmuje ten sam obszar działań, ale w każdym z trzech (czasem dwóch) krajów istnieje pod nieco inną nazwą (por. **RODO** *Pol.* [rozporządzenie o ochronie danych osobowych] / **ONOOÚ** *Cz.* [obecné nařízení o ochraně osobních údajů] / **ЗПД** *Ros.* [приказ о защите персональных данных]). Na zakończenie rozważań na temat ekwiwalencji zerowej w perspektywie przekształcalności słowników przeanalizujemy jeszcze jeden przykład. Gdy użytkownik słownika zada do wyszukiwarki słownikowej zapytanie w języku czeskim o postaci „zdrav.”<sup>25</sup>, wśród wyników otrzyma poniższą triadę przekładową.

**Tabela 5.** Transland zerowy? Konsekwencje ucyfrowienia słownika triad przekładowych

Język czeski	Język polski	Język rosyjski
∅ [ZDRAVeL]	Internetowe Konto Pacjenta [IKP]	Личный медицинский кабинет пациента

<sup>25</sup> Przypomnijmy, iż znak kropki w zapytaniu oznacza jeden lub więcej znaków alfabetu czeskiego.

Fakt wystąpienia znaku ∅ po lewej stronie słownika może zaskakiwać, o czym wspominaliśmy wcześniej. Jest on przede wszystkim konsekwencją ucyfrowienia słownika przekładowego i umożliwienia wyboru dowolnego z trzech języków jako języka wiodącego (teoretycznie wyjściowego) i rzecz jasna nie świadczy o istnieniu w słowniku „translandu zerowego”<sup>26</sup>. Językiem wyjściowym w procesie tworzenia słownika LMK był, przypomnijmy, język polski. Podstawowym założeniem zespołu autorskiego było zestawianie *reproduktów* polskich, czeskich i rosyjskich, tworzących triady przekładowe niezależnie od kolejności ich wyświetlania. Kluczowa jest tu zasada *podstawialności*, sformułowana swego czasu przez Jolantę Mędelską i Jana Wawrzyńczyka (Mędelska, Wawrzyńczyk 1992, s. 16), która pozwala użytkownikowi słownika traktować poszczególne jego jednostki jako elementy wymienne (w tekstach). Dotyczy to również ekwiwalencji zerowej, która sygnalizuje brak odpowiedności na poziomie reproduktów właśnie, ale nie brak jakiegokolwiek możliwości przekładowej.

Jeśli chodzi o kwestie praktyczno-leksykograficzne, zestawienie z tabeli 5 należałoby odczytać w następujący sposób: w języku czeskim nie istnieje taki reprodukt, który realizowałby sens związany z polskim reproduktom *Internetowe Konto Pacjenta* (funkcjonuje natomiast – jak sugeruje podana tu relacja – grupowiec *ZDRAVeL*). W języku rosyjskim taki reprodukt istnieje, choć nie istnieje (czytaj: nie jest przez rodzimych użytkowników języka rosyjskiego stosowany) odpowiadający mu literowiec. Sądzymy jednak, że podana tu triada powinna zostać zmodyfikowana. *ZDRAVeL* nie jest grupowcem, jest pełną nazwą (teoretycznie przyjmuje postać grupowca, choć w praktyce nie funkcjonuje jego pełny zapis), a więc mogłaby być odpowiednikiem polskiego *Internetowego Konta Pacjenta*. Rozbieżności te dotyczą również innych triad, por. zakładat: *kdo zakládá účet v ZDRAVeL / kto zakłada konto w IKP / кто создаёт Личный медицинский кабинет пациента* oraz *ZDRAVeL / IKP* [Interneto-

<sup>26</sup> Celowo ujętego tu w cudzysłów, jako oczywisty oksymoron.



we Konto Pacjenta] / ∅ [Личный медицинский кабинет пациента]. Szczególnie ta ostatnia triada ponownie sugeruje, że ZDRAVeL nie jest w języku czeskim grupowcem odpowiadającym polskiemu literowcowi<sup>27</sup>. W tym i w wielu innych miejscach zapewne konieczna będzie korekta materiału słownikowego. Podobne przypadki można albo modyfikować, albo opatrywać odpowiednim komentarzem, co w bazodanowej wersji słownika jest bezproblemowe (por. odpowiednie pola bazy słownika wymienione w tabeli 2).

\* \* \*

Ostatnią kwestią, którą poruszymy są wyrazy naprowadzające. Przysporzyły one wprawdzie mniejszych problemów niż ekwiwalent zerowy, choć także dotyczyło to prac zarówno praktyczno-leksyko-graficznych, jak i przekształceniowych. Dla wielu jednowyrazowych czasowników języka polskiego nie istnieją jednowyrazowe ich odpowiedniki w języku czeskim i rosyjskim. W ramach opracowywania siatki słownikowej autorzy poszczególnych działów, co zresztą uzasadnione, formułowali wyrazy (całostki) naprowadzające odnosząc się do ich sensu wyjściowego. Por. choćby pary polsko-rosyjskie *udostępniać – открывать доступ, wnioskować – подавать заявление, autoryzować – проводить авторизацию, bankować – пользоваться электронным банком, doubezpieczyć – застраховать дополнительно* czy też polsko-czeskie, np. *subskrybować – přihlásit se k odběru, przeanalizować – provést analýzu, organizować się – dávat se dohromady* (o grupie ludzi), *skomunikować się – navázat komunikaci* i inne<sup>28</sup>.

Wyrazy naprowadzające pełnią w słowniku LMK funkcję środka technicznego. Umożliwiają prezentację haseł w porządku alfabetycz-

<sup>27</sup> Co więcej, nie udało nam się podjąć ostatecznej decyzji w tym zakresie, ponieważ strona internetowa zdravel.cz została wyłączona, a odpowiadająca jej aplikacja na telefon nie jest już dostępna.

<sup>28</sup> Oczywiście z wyłączeniem czasowników zwrotnych z partykułami *się /se*.

nym. Dotyczą przede wszystkim zwrotów czasownikowych z sygnałami zaimkowymi *kto/kdo, co*, które bez wyrazów naprowadzających zgrupowane byłby w słowniku pod literami „k” i „c” (rzadziej stosowane są też rzeczownikowe lub przymiotnikowe/imię-słowne wyrazy naprowadzające). Ponadto w początkowej wersji słownika język rosyjski nigdy nie miał pojawić się po jego lewej stronie<sup>29</sup>. Dopiero prace nad wersją cyfrową (online) uświadomiły autorom taką potrzebę. Zaistniała zatem konieczność dopisywania brakujących słów naprowadzających, aby możliwe było poprawne (tzn. alfabetyczne) wyświetlanie wyników również rosyjskojęzycznych kwerend słownikowych. Przy okazji tych działań złożone „całostki naprowadzające” zostały zredukowane tylko do czasownika, czyli dwuwyrazowiec *открывать доступ* został zastąpiony bezokolicznikiem *открывать*. W aktualnej wersji słownika podobnych zmian nie wprowadzono jeszcze dla jednostek języka czeskiego<sup>30</sup>.

Z wyrazami naprowadzającymi w języku czeskim wiązała się jeszcze jedna rozbieżność względem języka polskiego. Niektóre czasowniki w języku czeskim są dwuaspektowe – jednostki z tymi czasownikami miały tylko jedną postać, podczas gdy odpowiadające im jednostki w języku polskim – dwa. Por. poniższe wyrazy:

- *absolovovat* w znaczeniu ‘odbywać, odbyć praktyki’,
- *asimilovat se* w znaczeniu ‘asymilować się, zasymilować się’,
- *citovat* w znaczeniu ‘cytować, zacytować’,
- *vydavat se* w znaczeniu ‘podszywać się, podszyć się’ i in.

W obecnej wersji słownika wypadki te bywają rozproszone wownikach wyszukiwania. Być może powiązanie pewnych haseł ze sobą – na wzór *Hipersłownika* Jana Wawrzyńczyka (por. przypis 19) – w lepszy sposób porządkowałoby informację słownikową. Zresztą takie rozwiązanie mogłoby dotyczyć nie tylko par aspektowych, lecz

<sup>29</sup> Por. słownik w wersji PDF pod adresem <https://bg.uni.opole.pl/wp-content/uploads/LMK-ISBN-978-83-8332-051-9.pdf>.

<sup>30</sup> Jest to wersja oznaczona datą 24.01.2024 r. (por. stronę internetową słownika pod adresem [idiomatykon.eu/search](https://idiomatykon.eu/search)).

także większej liczby powiązań wewnątrzsłownikowych (np. na poziomie leksemów).

## Perspektywy

Zarysowane tutaj problemy będą miały zastosowanie w przyszłości. Planujemy cyfryzację obu tomów *Polsko-rosyjskiego słownika par przekładowych*, dla której niniejsze rozważania oraz prace związane ze stworzeniem słownika *Lidé – Média – Komunikace* mają charakter, by tak rzec, pilotażowy. Wydaje nam się, że dość wyraźnie pokazaliśmy, iż każde działania prowadzące do przekształcenia jakiegoś słownika w inny typ słownika (jak w wypadku PIPR i PRSPP) czy inną postać słownika (jak w wypadku wersji tekstowej i cyfrowej słownika LMK) dostarczają nowych doświadczeń nie tylko w obszarze praktyki słownikarskiej (układanie hasłownika, ekwiwalentyzacja i in.), lecz również na polu metaleksykografii przekładowej (problemy natury teoretyczno-językowej z tą praktyką związane). Opisane tutaj zagadnienia pomogą autorom słownika LMK przystąpić do prac korygujących poszczególne hasła, a także porządkujących bieżące rozwiązania.

## Bibliografia

- Bogusławski Andrzej, 1976, *O zasadach rejestracji jednostek języka*, „Poradnik Językowy”, nr 8, s. 356–364.
- Borysowski Daniel, 2020, *Teksty zespolone a problemy leksykografii przekładowej*, Opole.
- Bralewski Dariusz, 2013, *Francusko-polski słownik (graficznych) odpowiedników przekładowych*, Łask.
- Chlebda Wojciech, 1985, *Oksymoron. Z problemów językowego poznania rzeczywistości*, Opole.
- Chlebda Wojciech, 1991, *Elementy frazematyki. Wprowadzenie do frazeologii nadawcy*, Opole.
- Chlebda Wojciech, 2009, *Idiomatykon 4.: gdzie jesteśmy, dokąd zmierzamy (i parę zdań o tym, skąd przychodzimy)*, [w:] *Podręczny idiomatykon polsko-rosyjski*, z. 4, red. W. Chlebda, Opole, s. 9–38.

- Chlebda Wojciech, 2010, *Reprodukty na warsztacie (I)*, [w:] *Na tropach reproduktów. W poszukiwaniu wielowyrzowych jednostek języka*, red. W. Chlebda, Opole, s. 7–13.
- Chlebda Wojciech, 2013, *O przekształcalności słowników (frazologicznych): casus opolski*, [w:] *Perspektywy współczesnej frazeologii polskiej*, red. G. Dziamska, J. Liberek, Poznań, s. 29–37.
- Chlebda Wojciech, 2014, *Od tematycznego frazeologicznego słownika przekładowego do alfabetycznego słownika par przekładowych*, [w:] *Polsko-rosyjski słownik par przekładowych. Tom zbiorczy Podręcznego idiomatykonu polsko-rosyjskiego (z. 1–5)*, red. W. Chlebda, Opole, s. 9–67.
- Chlebda Wojciech, Dobrotová Ivana, 2023, *Założenia wyjściowe i zasady użytkowania idiomatykonu*, [w:] *Lidé – Média – Komunikace. Česko-polsko-ruský a polsko-česko-ruský idiomatikon / Ludzie – Media – Komunikacja. Czesko-polsko-rosyjski i polsko-czesko-rosyjski idiomatykon*, red. D. Borysowski, I. Danecka, I. Dobrotová, Opole, s. 365–380.
- Gasek Bogumił, 2019, *Ekwiwalent zerowy w polsko-rosyjskiej leksykografii par przekładowych (na materiale 8. zeszytu »Podręcznego idiomatykonu polsko-rosyjskiego«)*, „Applied Linguistics Papers”, 26/1, s. 1–14.
- Grochowski Maciej, 1981, *O wyróżnianiu jednostek opisu semantyki leksykalnej*, [w:] *Sbornik prací Filozofické fakulty brněnské univerzity. A, Řada jazykovědná*, vol. 30, iss. A29, s. 31–37.
- Kilgarrif Adam, Grefenstette Gregory, 2003, *Web as Corpus: Introduction*, „Computational Linguistics”, vol. 29/3, s. 333–347.
- Mędelska Jolanta, Wawrzyńczyk Jan, 1992, *Między oryginałem a przekładem. Rzecz o słownikach dwujęzycznych*, Kielce.
- Wawrzyńczyk Jan, 2018, *Hipersłownik – Гиперсловарь – Hyperdictionary*, Warszawa.
- Wawrzyńczyk Jan (red.), 2004, *Wielki słownik rosyjsko-polski z kluczem polsko-rosyjskim*, Warszawa.
- Wawrzyńczyk Jan, 2020a, *Hipersłownik języka polskiego*, Tom 1: A–Ć, Warszawa.
- Wierzchoń Piotr, 2007, *Indeks a tergo jednostek przekładowych Wielkiego słownika rosyjsko-polskiego z kluczem polsko-rosyjskim*, Warszawa.