

Natalia Kotsyba¹
(gnatko@gmail.com)
Instytut Sławistyki PAN
Magdalena Turska
(turska@domeczek.pl)
Uniwersytet Warszawski

Leksykografia polsko-ukraińska – stan obecny i perspektywy

1. Istniejące słowniki drukowane - przegląd²

Ukraińska leksykografia przez długi czas była zdominowana i ukierunkowana przez leksykografię rosyjsko-radziecką, co prawdopodobnie przyczyniło się do tego, że ukraińskie słowniki dwujęzyczne, tworzone w czasach radzieckich były wzorowane na odpowiednich słownikach rosyjskich i nie odzwierciedlają rzeczywistej leksyki ukraińskiej, nie tylko zawierając wiele rusycyzmów, ale też często nie podając wyrazów ukraińskich, będących dobrymi ekwiwalentami dla języka wyjściowego, tylko wyrazy, które są dokładnymi tłumaczeniami rosyjskich ekwiwalentów z odpowiednich słowników.

W ciągu ostatnich dwóch dekad, w związku z „okresem bardzo ożywionych kontaktów wzajemnych w dziedzinie polityki, gospodarki, handlu, nauki i kultury, a zwłaszcza turystyki między Ukrainą a Polską” [TPMNazJur98, s.3] pojawiło się wiele małych słowników dwujęzycznych, często układanych „na szybko” pod potrzeby rynku, bez udziału dużych zespołów zawodowych leksykografów, w których czasami ucziwie deklaruje się, że są to przedruki słowników wcześniejszych, obecnie już niedostępnych. Głównymi zaletami tych nowopowstałych słowników było zazwyczaj usunięcie leksyki o tematyce marksistowsko-leninowskiej i dodanie najpowszechniejszych neologizmów, z reguły z branży komputerowej lub ekonomicznej.

W celu sprawdzenia obecnego stanu słowników polsko-ukraińskich przeprowadziłyśmy kwerendę wśród zawodowych tłumaczy, pytając o ich zdanie na ten temat i aktualne zapotrzebowanie. Poniżej przedstawiamy niektóre typowe wypowiedzi dotyczące poszczególnych słowników polsko-ukraińskich:

„dobry, ale malutki, jak na moje potrzeby”

„mówią, że masa tam rusycyzmów, ale nic lepszego nie ma, więc czasem zaglądam”

„dobre, ale bardzo przestarzałe”

„straszne maleństwo”

„kompletna katastrofa, wzór, jak nie należy robić słowników”

„mityczny słownik... wszyscy mówią, że dobry, ja z niego nie korzystam, bo jest nie do zdobycia, a do kserowania za duży”

Co się tyczy słowników ukraińsko-polskich, wiele tłumaczy polskojęzycznych stwierdza, że korzysta głównie z ukraińskich słowników objaśniających, i samemu ustala odpowiedniki polskie. Te słowniki również nie są łatwe do odnalezienia („nieduży słownik... mam go w postaci ksero, bo jest nie do zdobycia”). Oprócz słowników jednojęzycznych – „podstawowe źródło

¹ Nazwiska autorek podane są w kolejności alfabetycznej.

² Pragniemy podziękować prof. dr hab. J. Linde-Usiekiewicz za podpowiedzi ze sfery teorii leksykograficznej oraz prof. dr hab. J. S. Bieniowi za uwagi praktyczne i dostarczanie lektur.

informacji językowej to znajomi i znajomi znajomych (...), bo bez tego rodzaju kontaktów trudno jest cokolwiek zrobić porządnie, a już w przypadku literatury pięknej...”

Jak widać, szczególne niedogodności w związku z sytuacją polsko\leftrightarrowukraińskich słowników mają tłumacze. Osoby uczące się języka ukraińskiego również powinny uważać, żeby nie korzystać ze słowników zawierających rusycyzmy.

Intuicyjne sądy tłumaczy i osób, uczących się polskiego lub ukraińskiego potwierdza nasza analiza, oparta na słownikach objaśniających obu języków oraz obiektywnych danych korpusowych³. Dla przykładu podajemy niektóre nasze obserwacje oraz analizę porównawczą dostępnych słowników. Słowniki podzieliłyśmy na grupy według języka wyjściowego, w obrębie grup są one przedstawione według nazwisk redaktorów w porządku alfabetycznym.

1.1. Słowniki polsko-ukraińskie

Praktyczny słownik polsko-ukraiński, red. O. Aleksijczuk, 2004. **(PUAleks04)**

Najnowszy słownik, obecny na rynku. Zawiera 45 tys. artykułów hasłowych. Według autora, obejmuje najczęściej używane słownictwo współczesnego języka polskiego oraz terminologię specjalistyczną i pewien zasób wyrazów potocznych. Dodatki zawierają spis imion świętych oraz postaci biblijnych, a także współczesnych i historycznych nazw geograficznych, w tym niewielkich miejscowości polskich i ukraińskich, których nazwy nie są zwykłą transliteracją (np.: Pełtwa, Zaleszczyki, Krzemieńczuk, Skole, Świtez, Jezioro Aralskie), co uważamy za bardzo pożyteczne.

Słownik polsko-ukraiński, red. S. Lewińska, T. Starak, 1998. **(PULewSta98)**

Ten słownik jest jednym z najnowszych i najlepiej wypadł w ocenach tłumaczy. Jego najważniejszą wadą jest mała objętość – tylko 16 tys. haseł. Niemniej jednak wygrywa dzięki przemyślanemu doborowi haseł – mieści nie tylko słownictwo współczesne, powszechnie używane, ale też dobrą jak na swój rozmiar reprezentację terminologii specjalistycznej oraz najczęściej spotykane połączenia wyrazowe i frazeologizmy.

Słownik polsko-ukraiński we dwóch tomach, red. Ł. Ł. Humecka, 1958-1960. **(PU2tHum58)**

Warto tutaj wspomnieć o jego istnieniu, biorąc pod uwagę fakt, że jest to największy i najbardziej rozbudowany pod względem mikrostruktury słownik polsko-ukraiński, mimo że nie udało się nam do niego dotrzeć bezpośrednio (tzw. „słownik mityczny”). Z relacji autorów TPMNazJur [s. 4] wiemy, że zawiera ok. 100 tys. haseł, z których wiele jest przestarzałych i jest oparty materiałowo na polskiej literaturze pięknej wieku XIX i pierwszej połowy wieku XX.

1.2. Słowniki „dwustronne”

Mały słownik ukraińsko-polski, polsko-ukraiński, red. M. Jurkowski i B. Nazaruk, 1998. **(TPMNazJur98)**

Jak świadczy nazwa, jest to słownik mały - część ukraińsko-polska zawiera ok. 16 tys. haseł, część polsko-ukraińska – ok. 12 tys. haseł.

Ukraińsko-polski słownik (cz. 1), red. J. Hrycak i K. Wasylewski, 1990. **(UPHryWas31)**
Słownik polsko-ukraiński, (cz. 2), red. J. Hrycak i K. Wasylewski, 1990. **(PUHryWas31)**

³ W przypadku języka polskiego korzystaliśmy głównie z Korpusu IPI PAN, w przypadku języka ukraińskiego z stron www przy użyciu wyszukiwarek internetowych typu: google.com.ua lub www.nigma.ru

Część ukraińsko-polska zawiera 50 tys. artykułów hasłowych, część polsko-ukraińska – ok. 30 tys. Jest to dokładny przedruk słownika z 1931 roku, co oczywiście świadczy o jego charakterze „archaicznym”. Zarówno hasła, dobrane w obu językach, jak i ich znaczenia są mocno przestarzałe.

Słownik biznesu ukraińsko-polski, polsko-ukraiński, red. S. Domagalski, 2002. (TPBiz02)

Jest to niewielki słownik specjalistyczny, zawierający terminologię finansowo-ekonomiczną i częściowo krajoznawczą. Występują w nim liczne rusycyzmy (por. 1.4.2).

1.3. Słowniki ukraińsko-polskie

Ukraińsko-polski słownik, red. A. Iwczenko, 2003. (UPIwcz03)

W porównaniu z poprzednikami, atutem tego słownika ma być współczesne słownictwo i objętość zwiększona o ok. 3 tys. haseł; łącznie 36 tys. haseł według danych podawanych przez autora w przedmowie. Jednak jeśli uważnie przyjrzeć się jego budowie, można mieć wątpliwości czy dodane słownictwo dużo wnosi jakościowo. Nazwy geograficzne są włączone do samego słownika (zazwyczaj są one w dodatku). Jako osobne hasła są podawane formy dokonane i niedokonane czasowników. Można tam też znaleźć dużo form imiesłowowych, rzeczowników odczasownikowych, rodzajów czynności, które są tworzone regularnie, i raczej należą do sfery gramatyki. To wszystko sztucznie zwiększa objętość. Słownik „grzeszy” też licznymi okazami rusycyzmów i regionalizmów. Przykłady podajemy niżej.

Ukraińsko-polski słownik terminów lingwistycznych, red. M. Aleksijenko, A. Horniatko-Szumiłowicz, 2005. (UPLing05)

Zawiera ok. 700 haseł, ściśle lingwistycznych. Jest pożyteczny dla względnie wąskiego kręgu odbiorców.

Ukraińsko-polski słownik syntaktyczny, red. O. Śpiwak i M. Jurkowski, 2003. (UPSynt03)

Bardzo pożyteczny dla nauki obu języków. Podano w nim dokładny opis walencji ok. 1700 wyrazów, głównie czasowników – ekwiwalentów tłumaczeniowych, różniących się charakterystykami walencyjnymi. Por. pol. *zakochać się w kim, czym* (narzędnik) vs ukr. *закохатися в кого, що* (biernik). Artykuły hasłowe są opatrzone obszernym materiałem ilustracyjnym.

Prace nad dużym słownikiem ukraińsko-polskim są obecnie prowadzone także w Instytucie Języka Polskiego Polskiej Akademii Nauk (kierownik projektu – Janusz Rieger).

1.4. Makrostruktura

Makrostruktura rozpatrywanych słowników była czasem zaskakująca. Mogły się znaleźć w nich wyrazy bardzo rzadko używane (przy ilości haseł od 10 do 45 tys.), dialektyzmy oraz regionalizmy, przedstawiane jako rdzenne ukraińskie wyrazy, ale nie można było znaleźć wyrazów często używanych, nie tylko potocznie. Słowniki te, z nielicznymi wyjątkami wciąż bardziej odzwierciedlają język literatury pięknej, niż najbardziej używany, zrównoważony zestaw leksykalny z różnych sfer życia. Zmiany wprowadzane w liście haseł można uważać za kosmetyczne, zaś te dotyczące makrostruktury nie są raczej oparte na żadnych obiektywnych miarach częstości użycia wyrazów, np. przy pomocy list frekwencyjnych wyrazów nawet niewielkich, lecz zrównoważonych korpusów któregośkolwiek z języków.

1.4.1. Pozostałości z słownictwa „socjalistycznego”

TPMNazJur w części ukraińsko-polskiej, mieszczącej według deklaracji autorów ok. 16 tys. haseł, wciąż zawiera wyrazy typu *марксист* i *марксистський*, chociaż we wstępie autorzy krytykują „przestarzałe” słowniki za nadmiar „słownictwa «socjalistycznego» (często już nieaktualnego)” [TPMNazJur98, 5].

1.4.2. Rusycyzmy i polonizmy

W niektórych słownikach bardzo licznie występują rusycyzmy lub polonizmy, np. UPIwcz zawiera: *переміна* w znaczeniu *przerwa w szkole*, gdy poprawnie jest *перерва*, por. rosyjskie *переміна*. Także: *чманіти* (ros. *чманеть*) jako *wariować* zamiast *дурити*, *онтом* (ros. *онтом*) jako *hurtem* zamiast *гуртом*, itd. W TPBiz znajdziemy takie wyrazy jak: *загрузка* jako *ładowanie* zamiast *завантаження* (ros. *загрузка*), *належність* w znaczeniu *kwoty do zapłacenia*, a nie *bycia w czyjejś posiadłości*, (por. pol. *należność*); *явка* (*stawiennictwo*) zamiast *присутність* (ros. *явка*); *умовитися* zamiast *домовитися* (pol. *umówić się*); *стислість* zamiast *точність* (pol. *ścisłość*) itd.

1.4.3. Wyrazy rzadko używane i archaizmy

W TPMNazJur pośród 16 tys. słów można znaleźć takie rzadko używane wyrazy jak *марс* ze znaczeniem *bocianie gniazdo (na statku)* jako jedynym albo *непевесло* (*powrósło*) oznaczające przedmiot, który już dawno wyszedł z życia codziennego. Niektóre wyrazy z kolei są podawane w swoich przestarzałych znaczeniach: *неперевертень* tłumaczy się tam jako *zdrajca*.

Зав'язка została przetłumaczona jako *wiązadło, tasiemka; węzeł*, por. opis znaczenia tego wyrazu w [SUM, wersja elektroniczna]⁴:

1. Початок, основа, вихідний пункт, момент чого-небудь.
2. Епізод, яким починається розвиток сюжету в творі; протилежне розв'язка.

Jednym z odpowiedników wyrazu *заиморг* jest *суwak, zamek błyskawiczny*, por. znaczenia z [SUM]:

1. Петля з мотузки, шнура і т. ін., зроблена таким способом, що затягується.
2. Те, що тримає кого-небудь у якійсь залежності (матеріальній, духовній).
3. Крива лінія у вигляді замкненої або напівзамкненої петлі.

Znajdują się tam także czasowniki *obuwać, obuć, rozzuwać się, prześpiwanie, zakipieć*, które, jak to wynika z niżej przedstawionych danych korpusowych, są archaizmami.

W Korpusie IPI PAN, zawierającym 70 mln. słów, nie ma ani jednego wystąpienia wyrazu *obuwać* w jakiejkolwiek z form. Użyć, sprowadzanych w Korpusie do formy podstawowej *obuć*, znaleziono 21, z czego 19 to imiesłowu typu *obuty*, a dwa użycia w formie czasu przeszłego *obuł*, jedno z których pochodzi z powieści „Faraon” Bolesława Prusa, napisanej w 1985 r., a drugie z tekstu na temat religijny, mocno nacechowanego stylistycznie.

Żadne z 10 korpusowych użyć wyrazu *zakupieć* (wszystkie w formie czasu przeszłego) nie dotyczy wody, ale wszystkie zostały zanotowane w latach 1895-1908 i znajdują się w utworach Prusa, Żeromskiego lub Reymonta.

⁴ Wyrażamy wdzięczność zespołowi ULIF za udostępnienie dla celów badawczych dużej części Słownika Języka Ukraińskiego w postaci elektronicznej, nad którym obecnie prowadzone są prace w Ukraińskiej Akademii Nauk.

1.4.4. Brak często używanych wyrazów

Nowe słowniki deklarują zazwyczaj dodawanie wyrazów języka potocznego, ale nie można znaleźć tam niezbędnego dla osoby podróżującej po ukraińskich miastach słowa *маршрутка* oznaczającego polski „busik”. Natomiast *електричка* jest tłumaczona jako *pociąg elektryczny*. Uważamy, że bliższym funkcjonalnym odpowiednikiem byłby *pociąg podmiejski*, ponieważ dla użytkownika rodzimego języka ukraińskiego ten wyraz kojarzy się właśnie z krótkim dystansem kursowania i związanymi z tym brakiem wygody i niską ceną biletów na taki pociąg. Z innej strony istnieje bardziej ogólne pojęcie *електропоїзд*, przeciwstawiane pociągom parowym czy spalinowym, dokładny odpowiednik podanego polskiego wariantu *pociąg elektryczny*.

1.5. Mikrostruktura

Ze względu na małe objętości struktura hasłowa większości wymienionych słowników nie jest rozbudowana – podawane są tylko najbardziej powszechne znaczenia, rzadko – najbardziej używane połączenia wyrazowe (np. [PULewSta])

1.5.1. Aspekty gramatyczne

Można zaobserwować zarówno brak informacji gramatycznych jak i ich nadmiar, a także niezgodność przyporządkowania sobie haseł według charakterystyk gramatycznych.

Do dobrej praktyki leksykograficznej należy umieszczenie w słownikach dwujęzycznych krótkiego przeglądu gramatycznego, z podaniem wszystkich możliwych koniugacji i odmian, itd. Dokładny opis gramatyki języka ukraińskiego znaleźliśmy tylko w HryWasUP. TPMNazJur w dodatku „O języku ukraińskim” podaje tylko 5 reguł z wybranymi formami gramatycznymi, np. „Pierwsza osoba liczby mnogiej czasownika ma końcówkę –*мо...*”.

Niektóre słowniki z kolei usiłują „wcisnąć” nadmiarowe informacje gramatyczne w same hasła, zamiast podawać tablice odmian na końcu słownika. Np. UPIwcz w hasle *домогтися* podaje prawie cały paradygmat jego odmiany:

домогтися – *доможуся, доможешся, przesz. домігся, домоглася, гоґ. доможися...*

плем'я – D. *племені*||*м'я*, N. *племенем*||*м'ям*, M. Im *племена*, D.Im. *племен*, C. Im. *племенам* – *n, plemię*

Niektóre nowe słowniki zawierają wiele tworzonych regularnie form gramatycznych, np. imiesłówów *лежачу* – *лежачий* [PUAleks], *запрошений* – *zaproszony, доварений, вульгаризований, втрачений, втомливий, втопаючий* [UPIwcz], derywatów odczasownikowych – *bicie* [PUAleks], tym samym sztucznie zwiększając ilość haseł. Zbędne wydaje się również podawanie czasowników o znaczeniu inchoatywnym, tworzonych w sposób regularny przy pomocy prefiksów, np. *загавкати, завитати, задрімати*.

Warto także podawać informacje o częściach mowy ekwiwalentów, jeżeli należą do różnych. Np. ekwiwalentem przymiotnika *заміжня* w TPMNazJur jest rzeczownik *мężatka*; ekwiwalentem przymiotnika *рównoległa* jest fraza nominalna *паралельна лінія*.

Znaczenia czasowników nie zawsze są oddawane dokładnie w języku polskim. Np. *задрімати* jest tłumaczone jako *zdrzemnąć się*, które z kolei ma bliższe odpowiedniki *здріснути* oraz *подпірмати*, należących do tego samego *Akzionsart* – czynności jednorazowych. Poza tym, wyraz ukraiński mieści element znaczenia „coś, co dzieje się mimo woli”, natomiast polski wyraz zawiera element „zrobić coś z własnej woli”, co dobrze widać z korpusowych kontekstów użyć

tego wyrazu: *postanowiliśmy...*, *musi Pani...*. Poza tym, polski wyraz dotyczy całego procesu drzemania, akcentując tylko jego krótkoterminowy charakter, a ukraiński jest formą inchoatywną i dotyczy tylko zdarzenia, które rozpoczyna stan drzemania.

Por. przykłady z Korpusu IPI PAN:

... postanowiliśmy późnym popołudniem **zdrzemnąć** się, by nabrać sił na nocne szaleństwo

*... ми вирішили задрімати

... ми вирішили подрімати/здрімнути.

Najchętniej głowę lwa złożyłby na piachu, by **zdrzemnąć** się co nieco, ale nie wypada.

*... щоб дещо/ трохи задрімати...

... щоб дещо здрімнути / трохи подрімати...

1.5.2. Pragmatyka

Pol. *dziękować* vs ukr. *дякувати*.

Polski wyraz *dziękować* charakteryzuje szerszy wachlarz znaczeń, niż jego ukraiński odpowiednik *дякувати*. Możemy mówić o rozbudowanym znaczeniu wyrazu w polskim przypadku, bowiem inne języki, np. angielskie *thank*, niem. *danken*, fr. *remercier (merci)* nie rozwinęły znaczenia, pragmatyczny sens którego sprowadza się do odmowy, jak w przypadku *Akwizytorom i roznosicielom ulotek dziękujemy*, *Kołodce*⁵ *dziękujemy*. Zazwyczaj słowniki objaśniające języka polskiego podają to znaczenie jako użycie ironiczne. Informacje takiego rodzaju, czy to traktowane jako semantyczne, czy jako pragmatyczne są istotnym elementem tła kulturowego, bez którego nie można osiągnąć zadowalającego poziomu wiedzy językowej, a w przypadku tłumaczy zawodowych pomaga to unikać gaf.

Pol. *ambitny* vs ukr. *амбітний*

Ważna jest też kwestia konotacji wyrazów. Polski wyraz *ambitny* ma konotacje pozytywne, jego ukr. odpowiednik – negatywne. Mimo że są ekwiwalentami formalnymi warto chyba w niektórych przypadkach używać innych ekwiwalentów tłumaczeniowych, o czym powinny się znaleźć informacje w słowniku objaśniającym.

1.6. Homonimia międzyjęzykowa

Ze względu na bliskość pochodzenia języków polskiego i ukraińskiego wielkim, choć w różnym stopniu, problemem wszystkich słowników jest przyporządkowanie sobie wyrazów, które mają podobną formę, ale różnią się treścią – tzw. homonimów międzyjęzykowych albo „fałszywych przyjaciół tłumacza”.

1.6.1. Test porównawczy

Poniższy test pokazuje jak sugestywna może być forma językowa podczas tworzenia słowników przekładowych dwóch spokrewnionych języków. Jednostki leksykalne do testu zostały wybrane ze względu na ich problematyczny charakter jako potencjalnych homonimów międzyjęzykowych. Informacje o nich podajemy z encyklopedii i słowników objaśniających:

⁵ Ostatni przykład jest wzięty z prasy, z czasów, kiedy minister finansów był zmuszony porzucić swoje stanowisko w związku z powszechnym niezadowolaniem ze sposobu, w jaki pełnił swoje obowiązki.

SJP PWN⁶:

gad 1. Reptilia, gromada lądowych **kręgowców** zmiennocieplnych, oddychających płucami, przeważnie jajorodnych, o skórze pokrytej łuskami, tarczками lub płytkami; należą tu: hatterie, żółwie, jaszczurki, węże i krokodyle

2. *pot.* zwierzę, stworzenie budzące obrzydzenie, grozę, najczęściej wąż

wąż: *zool.* Serpentes, podrząd **gadów** lądowych obejmujący ponad dwa tysiące gatunków, rozpowszechnionych na całym świecie, poza strefą polarną; należą do niego węże właściwe i żmije»

zaskroniec ... niejadowity **wąż**, z rodziny węży właściwych, o ubarwieniu szarozielonkawym, z dużymi żółtymi plamami na skroniach; dobrze pływa...

Encyklopedia PWN⁷

żmija 1. Vipera, gad z rodziny o tej samej nazwie, z rzędu **węży**, o płaskiej, sercowatej głowie, wyraźnie oddzielającej się od reszty ciała i występujących w szczęce zębach jadowych, których ukąszenie jest niebezpieczne; w *lm* rodzina tych gadów

SUM:

плазуни - клас **хребетних** тварин, до якого належать **змії**, ящірки, черепахи, крокодили та інші тварини, що пересуваються перев. плазом, на животі

гад - земноводна тварина або **плазун**.

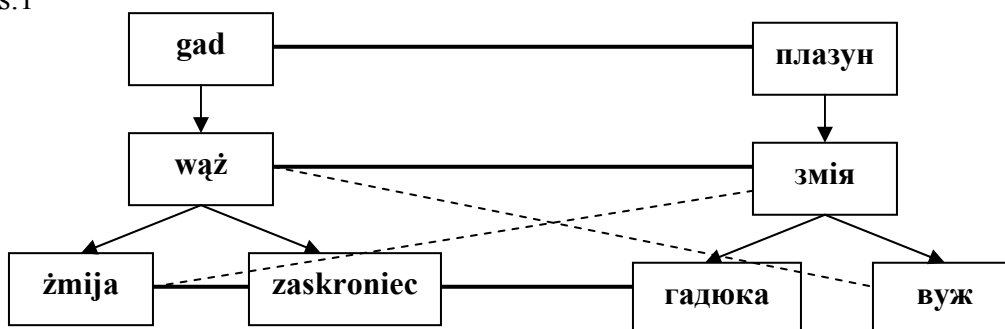
змія - **плазун** з видовженим тілом, укритим лускою.

гадюка - отруйна **змія** з плескатою головою у вигляді трикутника.

вуж - неотруйна **змія** середнього розміру, що має жовті плями з обох боків голови.

Jak widać, wybrane wyrazy należą do tego samego pola leksykalno-semantycznego i są między sobą w relacjach hipo-hiperonimicznych, co pokazałyśmy w rys. 1. Linie ciągłe łączą ekwiwalenty właściwe. Linie przerywane łączą wyrazy, które są często błędnie przedstawiane w słownikach jako ekwiwalenty, ponieważ mają podobną formę.

Rys.1



W tabelach podajemy tłumaczenia jednostek testowych w poszczególnych omawianych słownikach. Podkreśliłyśmy ekwiwalenty błędne.

Tab. 1

	PUNazJur	PUAleks	PU LewSta	PUHryWas
gad	1. плазун, 2.	1. плазун (зоол.)	зоол. плазун	гад, гady –

⁶ Tego skrótu używamy w odniesieniu do słownika internetowego znajdującego się pod adresem <http://sjp.pwn.pl>.

⁷ Korzystamy z wersji internetowej <http://encyklopedia.pwn.pl>

	гад, гадына	2. гад (перен.)		гадыя
wąż	1. вуж, змя. 2. ~ гумовый – шланг	1. змя (зоол.) 2. шланг, рукав	1. зоол. вуж, змя 2. шланг, рукав	вуж
zaskroniec	вуж звычайный	вуж звычайный (зоол.)	зоол. вуж	вуж
żmija	змя, гадюка	змя, гадюка (зоол., перен.)	змя, гадюка	змя, гадына

Tab. 2

	UPNazJur	UPHryWas	UPHrabiec	UPIwcz
вуж	<u>wąż</u>	<u>wąż</u>	<u>wąż</u>	<u>wąż</u>
змя	1. <u>zmija</u> , 2. wąż	<u>zmija</u>	<u>zmija</u> , wąż	<u>zmija</u> , wąż
гадюка	zmija	zmija, jędza	zool. zmija	1. zmija, 2. przen. zmija, gadzina
гад			gad, płaz, przen. gadzina	zw. lmn – gady, płazy, przen. gad, gadzina

Uważamy, że w artykułach hasłowych takich „fałszywych przyjaciół” tłumacza należy podawać odsyłacz także do znaczeń wyrazów podobnych pod względem formalnym, sygnalizując użytkownikowi możliwe niebezpieczeństwo błędnych skojarzeń.

1.7. Podsumowanie przeglądu

Podsumowując możemy powiedzieć, że mimo pojawienia się na rynku kilku nowych, rozszerzonych słowników polsko->ukraińskich, ich rozmiar i stopień rozbudowania struktur hasłowych wciąż nie są zadowalające dla użytkowników zawodowych. W dużym stopniu ich twórcy opierają się na własnej intuicji, a nie na obiektywnych źródłach językowych, co powoduje błędy, szczególnie w sferze homonimii międzyjęzykowej. Wspólny dla większości istniejących słowników jest dość przypadkowy wybór pozycji hasłowych, który w niedostatecznym stopniu odzwierciedla frekwencję użyć wyrazów. W nowych słownikach odczuwa się brak informacji gramatycznych, stylistycznych i pragmatycznych. Osiągnięcia teoretycznej i eksperymentalnej semantyki leksykalnej i gramatycznej drugiej połowy stulecia nie są wykorzystywane. Nadal istnieje zapotrzebowanie na większe i bardziej wiarygodne słowniki dwujęzyczne.

2. Zasoby elektroniczne

2.1. Słowniki

Polska leksykografia elektroniczna w porównaniu z ukraińską jest reprezentowana dość licznie, systematycznie i w sposób scentralizowany – większość jej twórców to duże wydawnictwa lub państwowe placówki naukowe. Wiele zasobów lingwistycznych jest powszechnie dostępnych przez Internet – od kilku lat w sieci istnieją przynajmniej dwa duże słowniki języka polskiego –

słownik wydawnictwa Wilga na portalu <http://slovniki.onet.pl>, oraz **Słownik Języka Polskiego PWN**, oparty w znacznej mierze na SJP pod red. W. Doroszewskiego – <http://sjp.pwn.pl>.

Internetowa i w ogóle elektroniczna leksykografia ukraińska ma charakter bardziej żywiołowy. Ze względu na brak wystarczająco solidnych zasobów tworzonych na poziomie państwowym, powstaje wiele inicjatyw oddolnych, które jednak nie mogą konkurować z projektami tworzonymi przez zespoły profesjonalne. Najbardziej poważnym ze względu na objętość i jakość materiałów językowych jest projekt zintegrowanych słowników elektronicznych pod kierownictwem Wołodymyra Szyrokowa realizowany przez **Ukraiński Fundusz Lingwistyczno-Informacyjny** przy Ukraińskiej Akademii Nauk w Kijowie. Słowniki: ortograficzny, synonimów (z objaśnieniami znaczeń spokrewnionych wyrazów), antonimów i gramatyczny są już dostępne na stronie internetowej ULIF – www.ulif.org.ua oraz w postaci płyty CD „**Słowniki Ukrainy**”. Prace nad nową wersją dwudziestotomowego objaśniającego słownika języka ukraińskiego w oparciu na obszerny jedenastotomowy słownik pod red. I. Biłodida są prawie ukończone. Autorzy zapowiadają, że w najbliższym czasie z tego słownika można będzie korzystać online albo z płyty CD.

Jeszcze jednym projektem, częściowo dotyczącym leksykografii ukraińskiej i zasługującym na szczególną uwagę, jest wielojęzyczny, komercyjny projekt **Lingvo ABBYY**. Przedsięwzięcie ma długie tradycje (ukazało się już ponad 10 wersji programu obsługującego słowniki) i cieszy się dużą popularnością. Jego główną zaletą jest to, że do programu można „doczepić” dowolny słownik jedno- albo dwujęzyczny, zawierający dowolne pary języków, pod warunkiem, że plik będzie skonfigurowany odpowiednio do wymogów programu. Ze strony internetowej Asocjacji Leksykografów (www.lingvoda.ru) można pobrać darmowy program pomocniczy, konwertujący odpowiednio sformatowany plik MS Word na plik słownikowy, który można podłączyć do programu ABBYY. Strona zawiera także bardzo bogaty zbiór gotowych darmowych plików słownikowych w takiej postaci. Większość z nich jest dostarczana przez użytkowników programu, dlatego one różnią się zarówno objętością jak i jakością. Znalazł się tam np. słownik ukraińsko-angielski na 30 tys. haseł (4770 Kb); angielsko-ukraiński ok. 17 tys. haseł (2170 Kb); kilka niedużych słowników terminologicznych i glosariów. Są to słowniki, oparte na źródłach „papierowych”. Program ABBYY zawiera z kolei wygodny interfejs graficzny, umożliwiający wyszukiwanie haseł, powrót do poprzednich wyszukiwań itd. Niestety sam program jest komercyjny.

Istnieją też liczne inicjatywy mniejszej rangi, jak np. **Polsko - ukraiński słownik internetowy**, liczący około 600 haseł, ułożony przez Romana Zacharija:

http://www.personal.ceu.hu/students/97/Roman_Zakharii/ukrainski.htm

Ten słownik ma postać tabeli, gdzie po lewej stronie znajdują się hasła polskie (bez dodatkowych kwalifikatorów części mowy itd.), a po prawej ich ukraiński wariant, zapisany czcionką łacińską. Każda para tłumaczeniowa jest podawana jako relacja jednojednoznaczna (jednemu polskiemu wyrazowi odpowiada jeden wyraz ukraiński). Można zatem stwierdzić, że słownik ma charakter amatorski i jest przydatny tylko dla tych użytkowników, którzy dopiero zaczynają poznawać język ukraiński, nie znając nawet jego alfabetu.

“**Tłumaczka**” (w oryginale – “Перекладачка”) <http://pere.slovnky.org.ua>

Jest to projekt, zadeklarowany przez autora – informatyka Walentyna Sołomki jako wielojęzyczny tłumacz-samouczek. Według informacji na stronie internetowej program dokonuje automatycznej ekstrakcji ekwiwalentów tłumaczeniowych na poziomie wyrazów i połączeń

wyrazowych, a także zdań z par tłumaczonych tekstów. Sądząc z możliwości programu, projekt jest w stadium początkowym. Np. przy próbie przetłumaczenia z angielskiego na ukraiński ciągu składającego się z dość popularnych wyrazów: *Good morning to you, I am glad to see you*, nie znalazł żadnego znajomego słowa. Natomiast z języka ukraińskiego na rosyjski niektóre wyrażenia tłumaczą się pomyślnie. Autor jednak twierdzi, że jest to tylko kwestia wystarczającej ilości tekstów równoległych.

Słownik ukraińsko-białorusko-angielsko-polsko-rosyjski Sołomki z możliwością edycji haseł przez dowolnego użytkownika Internetu, z wielojęzycznym interfejsem: <http://slovnky.org/dic>. Podajemy przykładowe tłumaczenie angielskiego wyrazu *water*:

WATER

Переклади >> [[Додати](#)]

Білоруська

ВАДА >> [[Виправити](#)] >> [[Вилучити](#)]

Польська

ОСІЕКАЇ >> [[Виправити](#)] >> [[Вилучити](#)]

PODLEWAĆ >> [[Виправити](#)] >> [[Вилучити](#)]

POIĆ >> [[Виправити](#)] >> [[Вилучити](#)]

POLEWAĆ >> [[Виправити](#)] >> [[Вилучити](#)]

WODA >> [[Виправити](#)] >> [[Вилучити](#)]

Російська

АКВАРЕЛЬ >> [[Виправити](#)] >> [[Вилучити](#)]

ВОДА >> [[Виправити](#)] >> [[Вилучити](#)]

ВОДНЫЙ >> [[Виправити](#)] >> [[Вилучити](#)]

ПОЛИВАТЬ >> [[Виправити](#)] >> [[Вилучити](#)]

ПОЛИТЬ >> [[Виправити](#)] >> [[Вилучити](#)]

РАЗВОДНЯТЬ >> [[Виправити](#)] >> [[Вилучити](#)]

Українська

АКВАРЕЛЬ >> [[Виправити](#)] >> [[Вилучити](#)]

ВОДА >> [[Виправити](#)] >> [[Вилучити](#)]

ВОДНИЙ >> [[Виправити](#)] >> [[Вилучити](#)]

ВОДЯНИЙ >> [[Виправити](#)] >> [[Вилучити](#)]

ПОЛИВАТИ >> [[Виправити](#)] >> [[Вилучити](#)]

ПОЛИТИ >> [[Виправити](#)] >> [[Вилучити](#)]

Słownik pozwala także na wyszukiwanie przy pomocy elementarnych wyrażeń regularnych. Niestety projekt jest niewystarczająco udokumentowany, zatem trudno cokolwiek powiedzieć o związanych z nim danych statystycznych.

Słowniki francusko-ukraiński i angielsko-ukraiński online firmy Trident Software, z wirtualną klawiaturą do wprowadzenia liter cyrylicy: <http://ukrainiandictionary.com>

Słownik ukraińsko-angielski online, 50 tys. haseł, z możliwością wyszukiwania wyrazów ukraińskich w dowolnej formie gramatycznej, wirtualną klawiaturą oraz podaniem akcentów: <http://lingresua.tripod.com/about.htm>

Kilka niewielkich słowników terminologicznych, jedno- i dwujęzycznych, na podstawie papierowych: <http://dict.linux.org.ua/dict/other/>

2.2. Korpusy jednojęzyczne

Język polski cieszy się dość różnorodną i bogatą objętościowo reprezentacją w postaci kilku dużych, anotowanych morfologicznie korpusów. Tutaj wymienimy tylko główne zasoby:

- 1) Korpus Języka Polskiego IPIAN [<http://korpus.pl>], zawiera ok. 70 mln. słów, morfologicznie anotowany, zaopatrzony w program wyszukiwawczy Poliqarp, dostępny publicznie. Projekt pod kierownictwem Adama Przepiórkowskiego.
- 2) Korpus referencyjny języka polskiego PELCRA: 100 mln. słów, zawiera dużo transkrypcji tekstów mówionych [<http://korpus.ia.uni.lodz.pl>]
- 3) Korpus PWN, 40 mln., z których bezpłatnie jest dostępna tylko wersja demonstracyjna wielkości 7,5 mln. słów. [korpus.pwn.pl]
- 4) Korpus „krakowski” – elektroniczny korpus ciągłych tekstów staropolskich (do 1500 roku). Autor projektu: Waław Twardzik. Na razie korpus nie jest dostępny publicznie, w Internecie można znaleźć tylko wzmiankę o nim [<http://www.jezyk-polski.pl/polski/00/01/granty.php>]

Korpus języka ukraińskiego, według naszego rozeznania, został stworzony przy wspomnianym wyżej ULIF, niestety tylko do użytku wewnętrznego (m.in. w celu stworzenia nowej rozszerzonej wersji dużego słownika objaśniającego). Planowane jest doprowadzenie go do postaci, która umożliwiłaby jego publiczne udostępnienie.

2.3. Korpusy równoległe

Dobrze rozbudowane korpusy równoległe istnieją jedynie dla par, w których co najmniej jeden z języków jest bardzo rozpowszechniony. Dla par języków ze względnie mniejszą liczbą użytkowników albo o mniejszym znaczeniu w skali światowej takie zasoby dopiero zaczynają się pojawiać. Do takich języków można zaliczyć polski i ukraiński. Ze względu na liczbę użytkowników są to języki raczej równorzędne, chociaż istniejące zasoby elektroniczne są bardziej obszerne w przypadku języka polskiego.

Z korpusów równoległych dla każdego z interesujących nas języków można wymienić polsko-angielski równoległy korpus, prace nad którym są dopiero planowane na Uniwersytecie Łódzkim. [<http://pelcra.ia.uni.lodz.pl/>]

Natomiast informacje o istnieniu jakichkolwiek równoległych korpusów przynajmniej średnich rozmiarów (do ok. 1 mln. słów) z udziałem języka ukraińskiego nie są nam znane. Poszczególne badacze układają niewielkie eksperymentalne korpusy dla konkretnych celów, bez dodatkowej anotacji (por. portal internetowy www.pere.slovnyk.org.ua, gdzie wspomina się o ukraińsko-rosyjskich oraz ukraińsko-angielskich zbiorach tekstów). Są to jednak jedynie wzmianki, bez dokładnych informacji statystycznych czy anotacyjnych.

2.3.1. PolUKR

W 2004 roku w Instytucie Sławiści Polskiej Akademii Nauk rozpoczęliśmy się prace nad PolUKR - polsko-ukraińskim korpusem równoległym, odzwierciedlającym w założeniach współczesny stan obu języków. Pomyślany on jest jako baza dla wielofunkcyjnego narzędzia dla tłumaczy, językoznawców i osób uczących się któregoś z wymienionych języków.

Korpus zawiera obecnie ok. 50 par tekstów, zestawionych (*aligned*) na poziomie akapitów i jest opatrzony w mechanizm wyszukiwania oparty na składni wyrażen regularnych (*regexp*). Ma on na razie charakter raczej eksperymentalny, umożliwiającą wstępną weryfikację projektu i testowanie pracy programistów. Przewidujemy jego rozbudowę po ukończeniu wstępnych prac projektowych oraz anotowanie na poziomie morfologicznym.

Szczegółową informację o PolUKR i jego możliwych zastosowaniach można znaleźć w [Kotsyba, Turska – PALC] albo na stronie <http://corpus.domeczek.pl>.

2.4. Podsumowanie

Podsumowując informacje o ukraińskojęzycznych zasobach elektronicznych możemy powiedzieć, że przeważnie są one oparte na wydaniach papierowych, dziedzicząc przy tym ich wady. „Tłumaczka” Sołomki jest wyjątkiem z tej reguły, ale zakres jej działania według naszych testów jest bardzo ograniczony, a dokumentacja projektu wyjątkowo skąpa. Nie można więc uważać tego programu za rzeczywistą pomoc w pracy tłumacza czy językoznawcy. Elektroniczna leksykografia polsko-<>ukraińska jest dopiero w powijakach, stanowi więc otwarte pole do działania.

3. Rozwiązania alternatywne

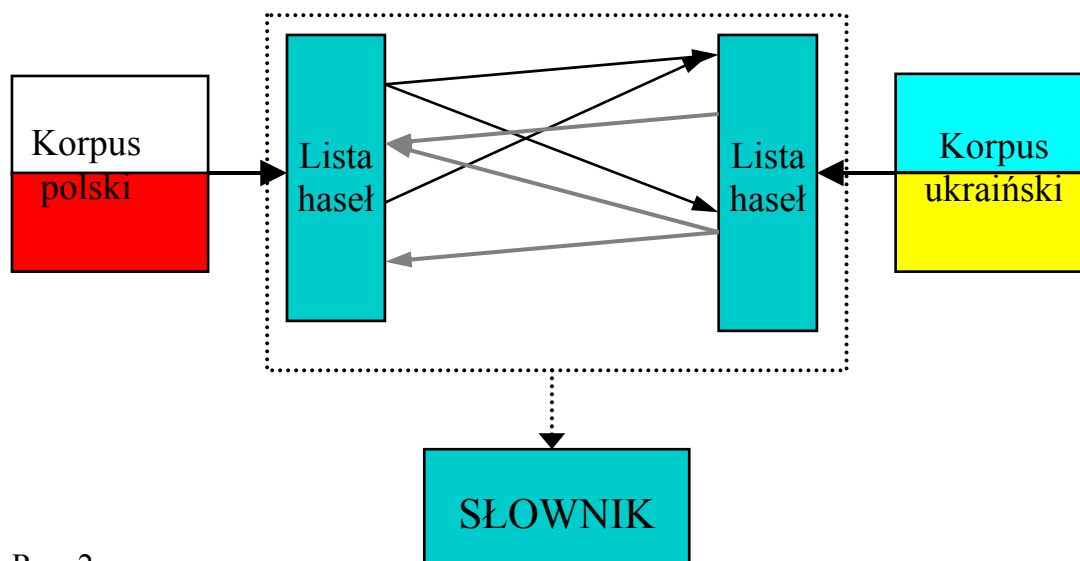
Nowoczesne metody lingwistyki komputerowej mogą nie tylko znacznie ułatwić pracę leksykografa, ale też odkrywają nowe możliwości badania i opisu leksyki i języka w całości. Między innymi, możliwe jest połączenie funkcji kilku słowników – synonimów, frazeologii, odmiany gramatycznej, dwujęzycznych itd. Częściowo takie systemy zostały już zrealizowane zarówno dla języka polskiego – portal PWN, jak i języka ukraińskiego [www.ulif.org.ua]. Natomiast zintegrowany system leksykograficzny, który by obejmował więcej języków istnieje według naszego rozeznania tylko dla angielskiego i niemieckiego [www.dict.leo.org]. My proponujemy jeszcze inne połączenie zasobów, które według naszego rozeznania jak dotąd nie zostało zrealizowane w ogóle. W przypadku minimalnym ma to być korpus równoległy, na podstawie którego tworzy się dwukierunkowy słownik dwujęzyczny i stopniowo integruje wszystkie zasoby w całość w postaci bazy danych. Taki system może też być rozbudowany o słowniki tematyczne dla obu języków, pliki dźwiękowe i graficzne – ilustrujące użycia wyrazów, odsyłacze od haseł mogą wreszcie prowadzić do odpowiednich artykułów encyklopedycznych.

3.1. Realizacja

Rys. 2 przedstawia schemat realizacji takiego systemu. Materiałem wejściowym (tzw. „surowcem”) jest:

- 1) korpus równoległy,
- 2) dostępne korpusy jednojęzyczne obu języków
- 3) duże elektroniczne słowniki objaśniające obu języków

Na podstawie tych materiałów tworzymy siatkę haseł, opierając się na najczęściej występujących w korpusie jednostkach i, ewentualnie, uzupełniając listę istniejących elementy makrostruktury istniejących słowników.



Rys. 2

Po wyznaczeniu listy haseł możliwie dobrze odzwierciedlającej współczesną leksykę, wyznaczamy obszary słownictwa, którymi zajmują się poszczególni redaktorzy. Wybieramy interesującą nas jednostkę językową co powoduje jednoczesne odwołanie się do wystąpień tejże jednostki w każdym z wymienionych źródeł. Informacje, związane z nią pojawiają się w odpowiednich częściach ekranu, umożliwiając redaktorom porównywanie tych danych i jednoczesną edycję hasła. Niżej omawiamy bardziej szczegółowo elektroniczne zasoby językowe, które zamierzamy stosować podczas edycji słownika dwujęzycznego.

3.1.1. Konkordancje równoległe

Anotowany morfologicznie w obu częściach korpus równoległy, którego segmenty (ekwiwalenty tłumaczeniowe) są przyporządkowane sobie na poziomie zdań oraz akapitów, uzupełniamy o wyszukiwarke, która znajduje wszystkie wyrazy lub połączenia wyrazów, odpowiadające zdefiniowanemu w zapytaniu kryteriom i porządkuje je w tabeli, podając przy tym lewy i prawy najbliższy kontekst każdego z wystąpień owych wyrazów (tzw. tabele konkordancyjne). Konkordanser dla korpusu równoległego umożliwia szukanie wyrazów w każdym z języków, przy czym wyświetla także wszystkie odpowiedniki tłumaczeniowe odnalezionych kontekstów. Już w takiej postaci konkordanser może być wielce przydatny zarówno dla zawodowych tłumaczy, jak i osób uczących się któregoś z wymienionych języków. Właściwie, przeglądając listę konkordancji, otrzymanej w wyniku zapytania, które zadajemy sami, dokonujemy analizy semantycznej całego hasła (jeśli szukamy odrębnego wyrazu), czyli wykonujemy pracę leksykografa. Różnica polega na tym, że dawną kartotekę hasła, która jest poprzednikiem konkordancji automatycznej, robi nam program komputerowy w nieporównywalnie krótszym czasie, z wysokim prawdopodobieństwem, że wszystkie możliwe znaczenia zostaną w niej ujęte, oczywiście, jeśli mamy wystarczająco duży korpus. Równoległe konkordancje umożliwiają nam tworzenie listy wyrażen będących tłumaczeniami hasła wejściowego w języku docelowym.

Konkordancje równoległe mają także „korzyści uboczne” w postaci twórczych wynalazków językowych tłumaczy – neologizmów, tworzonych i wprowadzanych do tekstów z konieczności (terminologia, frazeologizmy) lub trafnych opisów nowych pojęć (np. w przypadku nazw kulinarnych oraz innych realiów kulturowych). Z drugiej jednak strony nie wszystkie takie neologizmy mogą i muszą wejść do języka, czasami bowiem są to tylko okazjonalizmy lub odpowiedniki kontekstowe. Możliwość tłumaczenia wyrazów przy pomocy deskrypcji lub połączeń wyrazowych (kolokacji) jest dowodem na asymetryczność makrostruktur słowników dwujęzycznych tej samej pary językowej. Tłumaczenia mogą też być źródłem do ekstrakcji synonimów.

Konkordancje równoległe nie są absolutnie wiarygodnym źródłem odpowiedników tłumaczeniowych, nawet jeśli tłumaczenie jest zrobione dobrze. Jedną z głównych przyczyn jest to, że idealnie tłumaczy się nie osobne wyrazy, lecz sytuacje (*frames*) i tłumacz jest skazany w związku z tym na pewne konwencje. Np. w języku polskim zawiadamiamy o czyjejś śmierci, mówiąc *x nie żyje*. W języku ukraińskim – *nomep x* (dosł. *zmarł x*, szyk wyrazów jest tutaj istotny). W języku angielskim – *x is dead* (dosł. *x jest martwy*). Mimo że „technicznie” możliwe jest przetłumaczenie dosłowne z każdego i na każdy z wymienionych języków, i nie byłyby to struktury niegramatyczne, normy językowe wymagają użycia właśnie wymienionych wzorów. Potwierdzają ten fakt zarówno statystyczne dane korpusowe, jak i intuicyjne sądy rodzimych użytkowników tych języków.

Częste są też sytuacje, kiedy w tłumaczeniach jest stosowany nie dokładny odpowiednik tłumaczeniowy, lecz jego hiperonim lub hiponim, co też jest całkiem uzasadnione. W ten sposób unika się zbytecznego wgłębiania się w szczegóły lub, na odwrót, wyjaśnia się coś bardziej dokładnie. Często to jest sposób na tłumaczenie realiów językowych (por. anegdotę o czterdziestu nazwach śniegu u Eskimosów – przy tłumaczeniu na języki krajów, położonych bardziej na południu, użyje się hiperonimu, jeśli tylko nie zniekształca to przekazywanej treści). Tłumacze też mogą używać wyrazów bardziej oddalonych znaczeniowo (synonimów) w celu uniknięcia tautologii stylistycznych. Wymienione zamiany nie wpływają na jakość tłumaczenia, ale nie są satysfakcjonującym rozwiązaniem dla słowników dwujęzycznych.

Po odrzuceniu hiperonimów, hiponimów, synonimów oraz swobodnych tłumaczeń, oddalonych od oryginalnego znaczenia hasła, albo będących ekwiwalentami tylko kontekstowymi (możliwa jest częściowa automatyzacja tego procesu) otrzymujemy listę wyrazów – odpowiedników szukanego hasła. Warto zaznaczyć w tym miejscu, że podstawową jednostką językową, którą będziemy operować przy tworzeniu słownika, jest nie słowo, lecz jego **znaczenie**. Praca redaktora, jak widać, mimo wszelkich pomocy technicznych pozostaje procesem twórczym.

3.1.2. Słowniki objaśniające

Jak wspominałyśmy wcześniej, jedną z głównych wad istniejących słowników polsko$\langle\rangle$ukraińskich jest ich mała objętość i związana z tym niewystarczająco rozbudowana forma artykułów hasłowych. W związku z tym, strukturę haseł naszego słownika idealnie moglibyśmy wzorować na istniejących dużych słownikach objaśniających każdego z języków. Pożyteczność takich słowników polega jeszcze na tym, że pozwalają nawet bez zaglądan

konkordancji jednojęzycznych bądź równoległych⁸ uniknąć problemów homonimii międzyjęzkowej (patrz też 1.6., gdzie opisujemy przypadki sugestywności formy językowej w tworzeniu słowników dwujęzycznych).

Teoretyczne podstawy tworzenia słowników dwujęzycznych w oparciu o słowniki objaśniające każdego z języków opisywał jeszcze w latach 50-tych XX w. Lew Szczerba [Щерба, 1958]. Przyjęły się nawet odpowiednie terminy – słownik diagnostyczny (objaśniający) i słownik generacyjny (dwujęzyczny). W latach 80-tych XX w. ukazał się artykuł Hermana Grosbarta [1987] z opisem projektu elektronicznego słownika rosyjsko-polskiego właśnie na zasadach teorii Szczerby. Projekt ten jednak według naszego rozeznania nie został zrealizowany. Za korzystaniem ze słowników objaśniających jednak także kryje się kilka niebezpieczeństw dla leksykografa. Mianowicie: opisywane hasła albo ich znaczenia są często przestarzałe (wyodrębnione i egzemplifikowane na podstawie utworów klasycznych); znaczenia są niepotrzebnie powielane; znaczenia są niekonsekwentnie pogrupowane; kolejność opisywanych znaczeń nie zawsze odzwierciedla częstotliwość ich występowania w mowie. Problemy te są powielane, ponieważ próbując dopasować do siebie hasła różnych języków na poziomie znaczeń korzystamy z dwóch dużych słowników objaśniających. W związku z tym powstaje potrzeba uporządkowania językowego materiału wejściowego.

3.1.3. Próba uporządkowania danych przy pomocy klasyfikacji semantycznej

Praca redaktora, jak wspominałyśmy, polega w naszym przypadku głównie na wyznaczeniu, w których przypadkach, które odpowiedniki w konkordancjach i słownikach objaśniających są właściwe. Wyżej wymienione zasoby łącznie powinny dawać leksykografowi wyczerpującą informację potrzebną do ułożenia hasła słownika dwujęzycznego. Pojawia się jednak niebezpieczeństwo „utonięcia” w tym morzu informacji. Np. nie tylko w obrębie jednego słownika (objaśniającego) można dostrzec niekonsekwencje opisu haseł (co jest zrozumiałe przy takiej objętości), problem stanowi też różnica w zasadach opisu bliskoznacznych wyrazów w słownikach objaśniających zestawianych języków (szczegółowy opis przykładowego hasła podajemy w 3.1.5.). Uważamy, że konieczne jest uporządkowanie tych informacji, tym bardziej, że uporządkowaną informację językową można wykorzystać w innych celach, jak np. tworzenie słowników tematycznych, automatycznej ekstrakcji informacji, itd.

W tym celu wprowadzamy analizę semantyczną, opartą na pewnej uniwersalnej ze względu na języki klasyfikacji zarówno hasła wyjściowego, jak i jego ewentualnych odpowiedników. Wtedy ostatecznie przyporządkowujemy sobie ekwiwalenty językowe już na poziomie poszczególnych znaczeń szukanych haseł, należących do tych samych kategorii semantycznych. Znaczeniom przyporządkowujemy kategorie semantyczne, na których podstawie je wyróżniłyśmy, zapisując te relacje w bazie danych. To pozwala nam w miarę konsekwentnie rozróżniać odrębne znaczenia hasła. W ten sposób też dążymy do formalizacji zapisu znaczenia, co miałoby być pożytecznym materiałem do automatycznej analizy obu języków. Zakładamy, że mając konsekwentną klasyfikację, po opracowaniu większej ilości haseł przynajmniej jednego z języków możemy pogrupować znaczenia według przyporządkowanych kategorii semantycznych, co teoretycznie powinno nam dać pewnego rodzaju tezaursus oraz być podstawą do tworzenia ontologii i gramatyki semantycznej. Ze względu na wysoki stopień subiektywności materiałów

⁸ Tłumaczenia również zawierają niebezpieczeństwo homonimii międzyjęzkowej, tzw. „fałszywych przyjaciół” tłumacza, więc pod tym względem należy ostrożnie korzystać z konkordancji równoległych.

leksykograficznych i brak ogólnie przyjętych uniwersalnych klasyfikacji semantycznych⁹ ostateczna decyzja o dołączeniu bądź usunięciu znaczenia należy do redaktora.

3.1.4. Semantyczna klasyfikacja czasowników Galiny Zolotowej

Ostatnimi czasy prowadzone są intensywne badania w celu tworzenia ontologii stosowanych, które byłyby uniwersalnym opisem połączeń semantycznych słownictwa w całości [por. m.in. <http://www.loa-cnr.it>]. Zadowalających wyników na razie nie osiągnięto. Dlatego zaczynamy od opisu obszaru czasownikowego.

Główną zaletą klasyfikacji semantycznej czasowników [Zolotova et al.1999, 59-80], na podstawie której postanowiłyśmy zbudować swój system anotacji, jest to, że klasyfikuje znaczenia, a nie wyrazy. Dlatego zapewnia konsekwentny podział leksykalnych i gramatycznych znaczeń tych samych czasowników, przynajmniej na samym wierzchołku hierarchii. Nie pozostawia przy tym czasowników niesklasyfikowanych.

Podstawowe wyodrębnione grupy to **pełnoznaczące** (takie, które mogą samodzielnie pełnić funkcję predykatu/ orzeczenia w zdaniu; większość czasowników jest pełnoznacząca) oraz **niepełnoznaczące**. Te ostatnie są podzielone na **modyfikatory** i **kompensatory**.

Modyfikatory – fazowe (zaczynać, kontynuować); modalne (chcieć);

Do nich odnoszą się też złożone predykaty: zdążyć, spóźnić się, przyzwycząić się¹⁰

Kompensatory – zawsze czasowniki przechodnie, przy czym dopełnieniem jest nazwa nieprzedmiotowa, i właśnie na nim leży cały ciężar kategoriałno-semantyczny, czasownik zaś tylko „dodatkowo werbalizuje znaczenie procesu albo cechy i wyraża formalno-gramatyczne znaczenia predykatu” [Zolotova et al., 72] (*prowadzić badania/ zajęcia* = badać/ uczyć (wykładać); *ktoś wyróżnia się talentem* = jest utalentowany).

Pełnoznaczące – są dzielone na akcjonalne i nieakcjonalne

Akcjonalne: działalność fizyczna *myć, kosić*; czasowniki ruchu *biec*; mowy *krzyczeć, odpowiadać*; donatywne *kupować, odsyłać, dzielić się*; społecznych międzypodmiotowych działań, np. *klócić się, głosować, radzić się*; działalność mentalna; perceptywne; emocjonalne; fizjologiczne.

Nieakcjonalne (predykaty stanu): stanowe *rosnąć, cierpieć; kwitnąć; pada (deszcz)*; funkcyjne/ egzystencjonalne (pochodne od akcjonalnych, ale w połączeniu z podmiotami, będącymi nieanimowanymi nazwami przedmiotowymi mówią „o stanie funkcjonowania przedmiotów-artefaktów” [ibid. , s. 62]) *komputer działa, nóż kroi*; relacyjne: lokalizujące, partytywne, posesywne, porównawcze; komplikatory (kauzatory, autoryzatory).

Kauzatory są złożonymi predykatami i zawsze mają odpowiednik prosty (*jeść :: karmić, palić się :: zapalić, mieć :: dać, jechać :: wieźć, wisieć :: powiesić, klócić się :: skłócić kogoś z kimś*)

Autoryzatory różnią się od akcjonalnych predykatów mowy tym, że należą w wypowiedziach do sfery *modus*, a nie *dictum* (*oceniać jako, rozpatrywać, uważać za*)

Spróbujemy zilustrować atuty tej klasyfikacji za pomocą następującego przykładu:

⁹ Mamy nadzieję, że w trakcie prac nad słownikiem elektronicznym według opisanej tutaj metodologii uda nam się dostosować do potrzeb leksykograficznych i w miarę możliwości rozwinąć istniejącą klasyfikację semantyczną.

¹⁰ Zob. komentarze w 3.1.6.

prowadzić

1. wieść kogoś, coś do jakiegoś miejsca, do celu; przeprowadzać kogoś, kto nie zna drogi albo sam iść nie może, wskazywać komuś drogę - **akcjonalny ruchu + kauzator**
2. stawać się przyczyną czegoś; powodować, wywoływać coś, pociągać za sobą coś - **kauzator**
3. kierować pojazdem mechanicznym, statkiem, końmi =1
4. kierować partnerem w tańcu =1
5. sprawować nad czymś nadzór, zarządzać czymś, zajmować się, trudnić się czymś; kierować czymś - **niepełnoznaczny, kompensator**
6. w połączeniu z rzeczownikiem będącym dopełnieniem oznacza: realizować, kontynuować to, co jest wyrażone w dopełnieniu =5
7. ciągnąć się w jakąś stronę, stanowić dojście lub przejście do czegoś - **nieakcjonalny relacyjny, lokatywny**
8. o zawodnikach, drużynach sportowych: być pierwszym w klasyfikacji, mieć przewagę nad przeciwnikiem; przodować - **nieakcjonalny relacyjny, porównawczy**

Wyróżnione w ten sposób znaczenia, jak widać z przykładów poniżej, mają różne zachowanie składniowe oraz różne możliwości derywacyjne. Ciekawe, że formy dokonane od nich też są różne. Tego rodzaju informacji słowniki zazwyczaj nie podają, a ich brak stanowi źródło błędów językowych osób, uczących się języka. Por. przykłady modyfikacji fazowych, kauzatywnych, pasywizacji, możliwość tworzenia formy imiesłowu i form dokonanych do wybranych znaczeń *prowadzić*:

- 1) *Mama prowadzi dziecko do przedszkola – Dziecko jest prowadzone przez mamę – Prowadzenie dziecka przez mamę – *Mama zaczyna prowadzić dziecko do przedszkola – Mama powoduje, że dziecko idzie... (forma dk – odprowadzić; zaprowadzić)*
- 5) *Profesor prowadzi zajęcia – Zajęcia są prowadzone przez profesora – Prowadzenie zajęć/ profesora – Profesor zaczyna/chce prowadzić zajęcia – *Profesor powoduje, że... (forma dk – poprowadzić; przeprowadzić) Ścieżka prowadzi do lasu – *Prowadzenie przez ścieżkę do lasu – *Ścieżka zaczyna prowadzić do lasu – * Ścieżka powoduje... (brak formy dokonanej).*

W przypadku słowników dwujęzycznych takie rozróżnienie znaczeń często nakłada się na użycie różnych ekwiwalentów tłumaczeniowych.

3.1.5. Przykładowe hasło słownika dwujęzycznego, skonstruowane na podstawie klasyfikacji Zolotowej

Zobaczmy jak taki podział może być pomocny w konstruowaniu haseł słowników dwujęzycznych. Zaczniemy od analizy danych korpusowych.

3.1.5.1. Analiza konkordancji

Zapytanie do korpusu IPI PAN¹¹, które miało znaleźć wszystkie formy gramatyczne wyrazu *podejść* znajdujące się w tekstach stworzonych po 1960 roku (kryterium współczesności) zwróciło 244 próbki tekstów. Porównaliśmy użycia tego wyrazu z opisem jego znaczeń w [SJP PWN], który podajemy w tab. 3.

¹¹ Zapytanie zostało sformułowane następująco: [base="podejść"] meta created>1960; dokładny opis składni języka zapytań znajduje się pod adresem: <http://dach.ipipan.waw.pl/~adamp/Papers/2004-corpus/>

Zdecydowana większość użyć (ok. 80%) wystąpiła w znaczeniu 1; w najbliższej dystrybucji wskazywano docelowe miejsce ruchu „podejść do + nazwa przedmiotowa” (np. *do furtki, do koszyka, do nas; bliżej* = do mówiącego).

Podobnie z zapytaniem dotyczącym formy niedokonanej tego czasownika - 203 próbki, proporcje były te same. Użycia w pozostałych wymienionych w [SJP PWN] znaczeniach nie przekraczały 1% całości. Poza tym, wystąpiły użycia nie wymienione w tym słowniku w ogóle:

1. Mieliśmy grać to, co zawsze. I choć Pogoń bardzo poważnie **podeszła do meczu**, to my jesteśmy lepiej przygotowani. Mieliśmy wprowadzić drobny kryzys po meczu w Parmę, ale to już przeszłość. => NOWE ZNACZENIE „mieć nastawienie” (nieakcyjny, relacyjny) – ok. 20% użyć

2. W okolicy miasta Mechelen musieli awaryjnie lądować. **Podchodzący do lądowania** samolot nie uszedł uwagi belgijskich żołnierzy pełniących służbę na znajdującym się nieopodal posterunku granicznym. => NOWE ZNACZENIE „zaczynać” (niepełnoznaczny, modyfikator fazowy)

3. Na 5 pytań odpowiedział gładko, zapewniwszy sobie gwarantowany 1000 zł. Niestety, pytanie za 2000 zł, dotyczące El Greco, "nie **podeszło**". => NOWE ZNACZENIE, „pasować, odpowiadać” (nieakcyjny, relacyjny, porównawczy)

Tab. 3 Przykładowe hasło po analizie – wyraz polski

PODEJŚĆ - PODCHODZIĆ
(za słownikiem PWN)

1. «posunąć się (pójść, rzadziej: pojechać) w jakimś kierunku, zbliżyć się do kogoś lub czegoś»
2. «posunąć się pod górę; wspiąć się»
3. zwykle *dk* «postąpić wobec kogoś podstępnie, chytrze, zdradziecko; oszukać kogoś»
4. częściej *ndk* «zbliżać się do kogoś lub czegoś ostrożnie, ukradkiem, zwykle w celu dokonania napaści lub podpatrzenia; tropić»
5. «wypełnić się od spodu cieczą (zwykle w połączeniu z formą narzędnika)»

PODEJŚĆ - PODCHODZIĆ (po analizie konkordancji i klasyfikacji semantycznej)

1. posunąć się (pójść, rzadziej: pojechać) w jakimś kierunku, zbliżyć się do kogoś lub czegoś *akcyjny ruchu* = 1+2
2. mieć nastawienie do czegoś *nieakcyjny relacyjny* DODANE
3. oszukać kogoś *złożony akcyjny mentalny + autoryzator* = 3
4. zbliżać się do kogoś lub czegoś ostrożnie, ukradkiem, zwykle w celu dokonania napaści lub podpatrzenia; tropić *złożony akcyjny ruchu + mentalny + autoryzator* = 4
5. pasować, odpowiadać DODANE *złożony nieakcyjny, relacyjny + autoryzator*
6. zaczynać *niepełnoznaczny modyfikator fazowy* DODANE
7. wypełnić się od spodu cieczą (zwykle w połączeniu z formą narzędnika) *nieakcyjny stanowy* = 5

Tab.4 Przykładowe hasło po analizie – najbliższy odpowiednik ukraiński

ПІДЙТИ док. - **ПІДХОДИТИ** недок. (za słownikiem Білодиди)

ПІДЙТИ док. – **ПІДХОДИТИ** недок. (po zastosowaniu klasyfikacji semantycznej)

1. ідучи, наближатися до кого-, чого-небудь;

1. наближатися *akcyjny ruchu*

- наближатися підїжджаючи, підпливаючи, підлітаючи і т.ін. до кого-, чого-небудь; прибувати куди-небудь;
2. приступати до чого-небудь, братися за яку-небудь справу; виявляти своє ставлення до чого-небудь, оцінюючи;
 3. уміти привернути, прихилити кого-небудь до себе, завоювати довір'я; звернутися до кого-небудь з проханням, пропозицією, вимогою і т.ін
 4. наближатися, наставати (про час, події, явища і т.ін.)
 5. розміщуватися близько чого-небудь, бути в безпосередньому сусідстві з чимсь, межувати з ним
 6. бути придатним, прийнятним, відповідаючи яким-небудь вимогам; бути відповідним; личити; пристосовуватися, підроблятися
 7. переміщатися, підніматися догори
 8. ідучи, пройти яку-небудь відстань
 9. збільшуватися в об'ємі, підійматися (про тісто)
 10. насичуватися чим-небудь *Сніг підійшов водою.*
2. мати ставлення *nieakcjonalny relacyjny*
 3. пасувати *złożony nieakcjonalny, relacyjny + autoryzator*
 4. починати(ся) *niepełnoznaczny modyfikator fazowy*
 5. бути близько чого-небудь *nieakcjonalny relacyjny lokatywny*
 6. збільшуватися в об'ємі, підійматися (про тісто) *nieakcjonalny stanowy*
 7. насичуватися чим-небудь (рідиною) *nieakcjonalny stanowy*

Kojarzenie hasel pol. > ukr.

PODEJŚĆ - PODCHODZIĆ (analiza konkordancji i klasyfikacja semantyczna)

1. posunąć się w jakimś kierunku, zbliżyć się do kogoś lub czegoś *akcjonalny ruchu* **підійти 1**
2. mieć nastawienie do czegoś *nieakcjonalny, relacyjny* - **підійти 2**
3. szukać kogoś *złożony akcjonalny mentalny + autoryzator* **ошукати 1**
4. pasować, odpowiadać *nieakcjonalny, relacyjny* - **підійти 3**
5. zaczynać *niepełnoznaczny modyfikator fazowy* - **починати**
6. wypełnić się od spodu cieczą *nieakcjonalny stanowy* - **підійти 5**

Kojarzenie hasel ukr. > pol.

ПІДІЙТИ док. – **ПІДХОДИТИ** недок. (za słownikiem Biłodida)

1. наближатися *akcjonalny ruchu* - **podchodzić 1**
2. мати ставлення *nieakcjonalny relacyjny* - **podchodzić 2**
3. пасувати *nieakcjonalny relacyjny* - **podchodzić 5**
4. починати(ся) *niepełnoznaczny modyfikator fazowy* – **nadchodzić 5**
5. бути близько чого-небудь *nieakcjonalny relacyjny lokalizujący* - **znajdować się blisko**
6. збільшуватися в об'ємі, підійматися (про тісто) *nieakcjonalny stanowy* - **rosnąć 7**
7. насичуватися чим-небудь (рідиною) *nieakcjonalny stanowy* - **podchodzić 7**

3.1.6. Wady klasyfikacji Złotowej

Ponieważ dana klasyfikacja stosuje holistyczne podejście do leksyki, wychodzi nie od struktury predykatowo-argumentowej, lecz od wyrazu – hasła słownikowego, to nie udaje się przy jej pomocy dostrzec niektórych istotnych cech znaczeń czasowników, które mogłyby być przydatne w dalszym ich rozróżnieniu.

3.1.6.1. Problem polipredykcji

Czasowniki jednopredykatowe i wielopredykatowe są grupowane wspólnie. Czasowniki z elementem kauzatywnym są zaliczane do osobnej kategorii, mimo że pozostałe ich elementy semantyczne mogą należeć do kilku różnych kategorii. To powoduje niespójność klasyfikacji. W tej sytuacji możemy przyjąć podejście Małgorzaty Korytkowskiej [1992] i każdy podrodzaj predykatów podzielić na podkategorię, zawierającą element kauzatywny i taką, która go nie zawiera. Ponieważ nie tylko kauzatywność może być elementem dodanym do któregoś ze znaczeń czasownika, i tych elementów może być więcej niż dwa, wydaje się, że bardziej systematycznym podejściem byłoby rozdzielenie predykatów prostych i złożonych. W ten sposób traktujemy komponent kauzatywny jako czasownik niepełnoznacznym – modyfikator (razem z czasownikami fazowymi i modalnymi), a czasowniki, zawierające ten element jako złożone. Tę metodę stosowaliśmy w 3.1.5.

3.1.6.2. Reprezentacja struktury predykatowo-argumentowej

Nie są też rozróżniane czasowniki, zawierające któryś z argumentów w swojej strukturze semantycznej. Z reguły są to czasowniki ze znaczeniem „instrumentalnym”: *szczotkować, cesać, kosić*, często utworzone od nazw narzędzi, które są zawarte w ich strukturze semantycznej (por. *szczotkować* – *чумуму (шумкою)*). Polski wyraz zawiera argument instrumentalny w swoim znaczeniu i występowanie w zdaniu tego argumentu explicite byłoby niegramatyczne. Odpowiednik ukraiński natomiast dopuszcza jego użycie jawne, a nawet jest ono wskazane, ponieważ dokładniej oddaje treść polskiego wyrazu.

Jak zauważa Korytkowska, takie subtelności można zauważyć tylko porównując sposoby predykcji w różnych językach. Do wyróżnienia elementu instrumentalnego dochodzi w swoich rozważaniach i próbach formalnego zapisu semantyki czasowników Olgierd Wojtasiewicz [1975], m. in. także na podstawie porównywania z angielskimi czasownikami – typowymi reprezentantami konwersji słowotwórczej rzeczownik → czasownik.

3.1.6.3. Problem zakresu dekompozycji semantycznej

Krańcowym w tym przypadku byłoby przyjęcie całkowitej dekompozycji semantycznej – „Idealem, do którego w rekonstrukcji znaczeń języka naturalnego należy dążyć, jest posługiwanie się wyłącznie elementami absolutnie prostymi, przynależącymi do semantycznej bazy. To jednak na obecnym etapie zaawansowania teorii semantycznej nie jest możliwe...” [Bogusławski & Danielewicz 2005, 20]. Prace nad uniwersalnym zestawem takich elementów i opisywaniem znaczenia leksykalnego przy ich pomocy wciąż trwają – por. teoria prymitywów semantycznych Anny Wierzbickiej [1996], semantycznej struktury leksyki Raya Jackendoffa [2000].

Wydaje się jednak, że dla celów słownika dwujęzycznego rozkładanie znaczenia na najmniejsze jednostki byłoby zbędne, ponieważ struktury semantyczne jednostek – odpowiedników przekładowych byłyby w większości przypadków identyczne albo bardzo podobne. Niemniej jednak, podane przykłady pokazują, że dla obiektywnej reprezentacji struktury semantycznej czasowników nie możemy uniknąć przynajmniej częściowej dekompozycji znaczenia. Ważnym więc jest pokazanie właśnie różnic między nimi. Dlatego dekompozycja powinna być stosowana minimalnie i tylko z konieczności. Dokładna analiza większej ilości haseł powinna dać wskazówki co do optymalnego sposobu opisu znaczenia.

3.1.6.4. Problem uniwersalności

Klasyfikacja Złotowej obecnie obejmuje tylko czasowniki, więc pozostaje problem innych części mowy. Jest to kwestia teoretyczna, dla której idealne rozwiązanie na razie nie zostało znalezione. Pozostaje też problem traktowania różnych części mowy – czy należy semantycznie rozdzielać je na samym wierzchołku hierarchii, czy pozwalać na grupowanie różnych części mowy na podstawie wspólnego etymonu. Istnieją klasyfikacje, deklarowane jako ponadkategorialne, np. według teorii Leksykonu Generatywnego James'a Pustejovskiego, ale wymaga ona osobnego rozpatrywania i dodatkowych badań empirycznych. Pozostaje także problem systematyzacji informacji stylistycznych i pragmatycznych.

4. Wdrożenie metody w praktyce

Ze względu na ogrom pracy konieczny do wdrożenia prezentowanej metody w praktyce, bogaty zbiór informacji (w tym cennych danych będących efektami ubocznymi pracy edytorskiej) oraz najprawdopodobniej duży zespół redakcyjny konieczne jest udostępnienie redaktorom słownika narzędzia, które

- umożliwi pracę zespołową, zdalną, niezależną od położenia geograficznego użytkownika,
- zapewni równoczesny dostęp do jak najszerszej gamy zasobów lingwistycznych (słowniki objaśniające, dwujęzyczne, korpusy z wyszukiwaniem konkordancji),
- zautomatyzuje chociaż częściowo niektóre zadania (wyszukiwanie kandydatów na synonimy, wyszukiwanie informacji o walencji),
- uporządkuje pracę i narzuci pewne standardy formalne jej wykonywania i zapisu wyników (dobór i podział haseł, sformalizowana mikrostruktura, możliwość uporządkowanego zapisu notatek).

Najbardziej elastyczną formą dostępu dla użytkowników jest obecnie praca zdalna, przez internet, najlepiej przy użyciu oprogramowania pracującego w przeglądarce www (jak na przykład Internet Explorer), które nie wymaga żadnej instalacji na komputerze użytkownika. Dzięki temu możliwe jest korzystanie z rozmaitych platform sprzętowych, lub przenosiny z jednego stanowiska na drugie bez żadnych dodatkowych kłopotów, co jest istotne zwłaszcza w sytuacji, w której użytkownikami docelowymi są osoby o wykształceniu przede wszystkim humanistycznym. Przy rozwiązaniu internetowym kluczowe jest zapewnienie bezpieczeństwa danych, m.in. dzięki zróżnicowanemu systemowi uprawnień dla poszczególnych użytkowników.

Środowisko pracy redaktora powinno integrować jak najwięcej dostępnych zasobów lingwistycznych, takich jak istniejące już słowniki objaśniające i dwujęzyczne, a przede wszystkim korpusy tekstów jednojęzycznych i korpusy równoległe, stanowiące bazę dla stosowania prezentowanej przez nas metody. Korzystanie z tych zasobów powinno być jednocześnie możliwie jak najprostsze i najszybsze, z możliwością czytelnej prezentacji wyników a zarazem umożliwiające zadawanie stosunkowo szczegółowych zapytań. Żeby to umożliwić, biorąc pod uwagę nie tylko zbiory aktualnie dostępne, ale i takie, które mogą powstać w przyszłości, niezbędne jest stosowanie ogólnie przyjętych standardów umożliwiających eksport i import pomiędzy różnymi narzędziami i bazami danych. Proponujemy zapis danych w języku XML i standardzie CES, przy użyciu kodowania znaków w standardzie Unicode, co umożliwi pracę z wieloma językami i przyłączanie nowych zasobów lingwistycznych.

Dzięki sformalizowanemu zapisowi danych, obejmującemu przede wszystkim anotację morfologiczną możliwa jest częściowa automatyzacja niektórych zadań, przede wszystkim wyszukiwanie według kryteriów gramatycznych.

Sformalizowana, zhierarchizowana i skategoryzowana struktura danych gromadzonych w systemie pozwala na ich wyświetlanie w postaci widoków prezentujących najważniejsze w danej chwili informacje, przy zachowaniu możliwości jednoczesnego dostępu do wszystkich pozostałych powiązanych zasobów dzięki licznym odsyłaczom. Umożliwia to również szybkie tworzenie produktów zorientowanych na jedno zadanie, np. słowników tematycznych czy podręcznych dzięki ograniczeniu liczby haseł do tych należących do pewnej kategorii (słownik tematyczny), czy najczęściej występujących (podręczny).

4.1. Graficzny interfejs do pracy edytorskiej

Poniżej przedstawiamy szkic części interfejsu umożliwiającej edycję hasła słownikowego przy jednoczesnym przeszukiwaniu konkordancji równoległych.

forma **Szukaj** polskie
 hasło ukraińskie

Kontekst lewy

5 10 15

całe zdania

Kontekst prawy

5 10 15

całe zdania

oraz na lata następne konieczności nowego **podejścia** do problemu zaspokojenia potrzeb mieszkaniowych

przy Rynku Kleparskim dwie kobiety **podeszły** od tyłu do robiącej zakupy mieszkanki

також на наступні роки потреби нового **підходу** до проблеми задоволення житлових потреб

біля Клепарського ринку дві жінки **підійшли** ззаду до мешканки, котра робила покупки

1. «posunąć się (pójść, rzadziej: pojechać) w jakimś kierunku, zbliżyć się do kogoś lub czegoś»	>	<input type="text" value="підійти"/> Definicja
2. «mieć nastawienie do czegoś»	>	1. (до когось/чогось) наблизитися
3. «oszukać kogoś»	>	2. (до когось/чогось) пасувати пр.
4. «pasować, odpowiadać»	>	<i>ключ не підійшов</i>
5. «zaczynać»	>	3. (до когось/чогось) з певної точки зору
6. «wypełnić się od spodu cieczą (zwykle w połączeniu z formą narzędnika)»	>	4. піднятися (про дріжджове тісто)

4.2. Tematyczny podział pracy nad hasłami

Opracowany przez nas przykład homonimii międzyjęzykowej (zob. 1.6.) pokazuje, że warto pracować nad hasłami tematycznie. Redaktor ma możliwość lepiej wniknąć w wąski krąg tematyczny, zawierający np. terminologię z jakiejś dziedziny, postarać się ustalić związki logiczne między pojęciami, a zatem i wyrazami, które je opisują, i uniknąć w ten sposób typowych błędów, sugerowanych przez formę wyrazów. Ponadto, utworzone w ten sposób mini-słowniczki tematyczne same będą stanowić pewną wartość zarówno dla celów dydaktycznych jak

i w przypadkach tłumaczeń specjalistycznych. Oczywiście, pozostaje problem rozdziału słownictwa na grupy tematyczne, ale to jest temat na osobne opracowanie.

4.3. Zespół redaktorski

Możliwe są przynajmniej dwie drogi rozwijania zaproponowanego systemu elektronicznych zasobów lingwistycznych. Pierwszy sposób, być może zapewniający bardziej profesjonalne wykonanie prac leksykograficznych, to utworzenie zespołu zawodowych językoznawców, niekoniecznie leksykografów, a raczej specjalistów od semantyki leksykalnej i gramatycznej, z których każdy zajmowałby się pewnym obszarem tematycznym słownictwa. Ten sposób jednak, jest bardzo kosztowny i dlatego mało realny. Drugi sposób jest żywiołowy, którego rozwój jest trudno przewidzieć – to tworzenie słownika (głównie opis i klasyfikacja semantyczna znaczeń haseł) przez użytkowników korpusu równoległego, który byłby dostępny publicznie poprzez portal internetowy. Taki portal może zawierać dużo informacji dodatkowych z zakresu historii literatury, translatoryki, językoznawstwa oraz forum, gdzie zarówno użytkownicy, jak i twórcy słownika i korpusu mogliby wymieniać zdania. Stały dopływ tzw. informacji zwrotnej umożliwiłby natychmiastowe dostosowanie się do uwag i zapotrzebowań użytkowników i do dynamicznego rozwoju portalu. Znane są liczne udane projekty, gdzie praca społeczeństwa internetowego przyczyniła się do powstania bardzo pożytecznych produktów np. encyklopedia Wikipedia, istniejąca w wielu wersjach językowych, system operacyjny Unix / Linux, itd. Pojawiło się nawet określenie „wiktionary” (wiki-słownik) oraz przykładowe darmowe oprogramowanie, umożliwiające zdalną pracę grupową [Wiktionary]. Potrzebowałoby ono jednak poważnego rozbudowania, żeby można było go dostosować do potrzeb opisanej wyżej metodologii. Konieczna więc jest ścisła współpraca profesjonalnych informatyków i językoznawców, która według naszego głębokiego przekonania może przynieść bardzo owocne efekty.

Źródła leksykograficzne (z wykazem używanych skrótów)

- PUAleks04** – *Praktyczny słownik polsko-ukraiński. Практичний польсько-український словник*, red. Olega Aleksijczuka, Kraków, 2004.
- PULeWSta98** – *Польсько-український словник*, упорядники – С. Й. Левінська, Т. В. Старак, Львів, Світ, 1998.
- PU2tHum58** – *Słownik polsko-ukraiński we dwóch tomach* red. Ł. Ł. Humecka, Kijów, 1958-1960.
- PUHryWas31** – *Українсько-польський і польсько-український словник, ч. 2. Польсько-український словник* під ред. Євгена Грицака і Костя Василевського. Варшава: Міжмор'я, 1990.
- SJPDor** – *Słownik języka polskiego*, red. W. Doroszewskiego, Warszawa, 1963.
- SUM** – *Словник української мови в 11-у томах* під ред. І. Білодіда, Київ 1970-1980.
- TPBiz02** – *Słownik biznesu ukraińsko-polski, polsko-ukraiński* pod red. Stanisława Domagalskiego. W-wa: Wiedza Powszechna, 2002.
- TPMNazJur98** – *Mały słownik ukraińsko-polski, polsko-ukraiński* pod red. M. Jurkowskiego i B. Nazaruka, Warszawa, Wydawnictwa Szkolne i Pedagogiczne, 1998.
- UPHryWas31** – *Українсько-польський і польсько-український словник, ч. 1. Українсько-польський словник* під ред. Євгена Грицака і Костя Василевського, з передмовою Мар'яна Юрковського. Варшава: Міжмор'я, 1990.
- UPIwcz03** – *Ukraińsko-polski słownik. Українсько-польський словник* pod red. A. Iwczenki. Lublin: Wydaw. UMCS, 2003.
- UPLing05** – *Українсько-польський словник лінгвістичних термінів* під ред. М. Алексеєнко, А. Горнятко-Шумилович. Szczecin: Wydaw. Naukowe Uniwersytetu Szczecińskiego, 2005.
- UPSynt03** – *Ukraińsko-polski słownik syntaktyczny: (różnice w łączliwości syntaktycznej około 1700 słów i wyrażeń z ilustracjami z klasycznej i współczesnej literatury ukraińskiej oraz prasy)* pod red. Oresta Śpiwaka i Mariana Jurkowskiego. Warszawa: Tyrsa, 2003.

Bibliografia

- Bogusławski, A., Danielewiczowa M. 2005. *Verba polona abscondita. Sonda słownikowa III*. Warszawa.
- Grosbart H., 1987. *Szkic teoretycznych założeń projektu komputerowego słownika przekładowego i propozycja podjęcia prac nad prototypowym komputerowym słownikiem rosyjsko-polskim*, [w:] "Studia z leksykografii współczesnej" pod redakcją Zygmunta Saloniego.
- Jackendoff, R., 2000. *Semantic Structures*, MIT Press (Korean translation, Hanshin Publishing Co.)
- Korpus „Krakowski”(wzmianka): <http://www.jezyk-polski.pl/polski/00/01/granty.php>
- Korytkowska M. 1992. *Gramatyka konfrontatywna bułgarsko-polska, t. 5: Typy pozycji predykatowo-argumentowych*, Warszawa, SOW.
- Kotsyba N., Turska M. *Polish-Ukrainian Parallel Corpus and its Possible Applications*. Proceedings of the International Conference "Practical Applications in Language and Computers, 7-9 April, Łódź", Peter Lang GmbH. (w druku)
- Laboratory for Applied Ontology: <http://www.loa-cnir.it/>
- Lieber, Rochelle 2004. *Morphology and Lexical Semantics*, Cambridge Studies in Linguistics 104, CUP.

Lipczuk, R. „*Falszywi przyjaciele tłumacza*” w słownikach niemiecko-polskich, <http://germ.ger.univ.szczecin.pl/~lipczuk/ArtFA.htm>

Piotrowski, T. 1994. *Z zagadnień leksykografii*. Warszawa: Wydawnictwo Naukowe PWN.

Polsko-angielski korpus równoległy (wzmianka): <http://pelcra.ia.uni.lodz.pl/>

Przepiórkowski, A. *Korpus IPI PAN. Wersja wstępna / The IPI PAN Corpus: Preliminary version*: <http://dach.ipipan.waw.pl/~adamp/Papers/2004-corpus/>

Pustejovsky, J. 1995. *The generative Lexicon*. Cambridge, MA: MIT Press.

Wierzbicka, A. 1996. *Semantics, Primes and Universals*. (Oxford: OUP).

Wójtowicz, B. 2005. *Ogólnolingwistyczne i leksykograficzne podstawy organizacji słownika suahilijsko-polskiego*, nieopublikowana rozprawa doktorska.

Winogradnik, Jana. *Wstępna klasyfikacja homonimów polsko-ukraińskich*, „Adeptus”, SOW, Warszawa (w druku).

Wojtasiewicz O. A. 1975. *Sformalizowana semantyczna interpretacja czasowników*, „Studia Semiotyczne”, t. VI, Ossolineum, Wrocław.

Золотова Г., Онипенко Н., Сидорова М. 1998. *Коммуникативная грамматика русского языка*, Москва, ИРЯ – РАН.

Крылов С. А. *Некоторые особенности лингвистической концепции В. Г. Гака*: <http://www.lingvoda.ru/transforum/articles/Krylov.asp>

Селегей В. *Электронные словари и компьютерная лексикография*: http://www.lingvoda.ru/transforum/articles/selegey_a1.asp

Щерба, Л. В. *Избранные работы по языкознанию и фонетике*, Ленинград, 1958. (podajemy za H. Grosbartem)

Linkografia zasobów elektronicznych (stan na listopad 2005 roku)

ABBY Lingvo: www.lingvo.ru/; www.lingvoda.ru/

Encyklopedia PWN: <http://encyklopedia.pwn.pl>

Korpus IPI PAN: <http://korpus.pl>

Korpus PWN: korpus.pwn.pl/

Korpus referencyjny języka polskiego PELCRA: <http://korpus.ia.uni.lodz.pl/>

Polsko - ukraiński słownik internetowy: http://www.personal.ceu.hu/students/97/Roman_Zakharii/ukrainski.htm

PolUKR – Polsko-Ukraiński Korpus Równoległy: <http://corpus.domeczek.pl>

Słownik Języka Polskiego PWN: <http://sjp.pwn.pl>

Słownik ukraińsko-białorusko-angielsko-polsko-rosyjski: <http://slovnyk.org/dic>

Słowniki: francusko-ukraiński, angielsko-ukraiński, języka ukraińskiego: <http://ukrainiandictionary.com>

Słownik ukraińsko-angielski: <http://lingresua.tripod.com/cgi-bin/oluaen.pl>

Tłumacz-samouczek W. Sołomki: <http://pere.slovnyk.org.ua>

Ukraiński Lingwistyczny Portal: www.ulif.org.ua

Zintegrowany leksykograficzny system dla języka angielskiego i niemieckiego: www.dict.leo.org

Wikipedia: <http://pl.wikipedia.org>

Wiktionary: <http://en.wiktionary.org/>

Wybrane elektroniczne zasoby języka ukraińskiego: <http://dict.linux.org.ua/dict/other/>