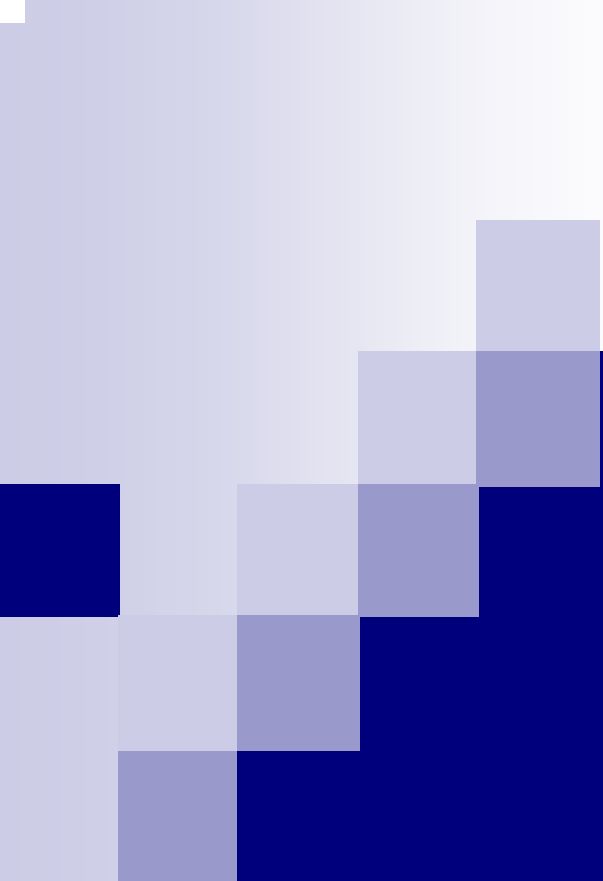




Semantyczne klasyfikacje czasowników dla potrzeb leksykograficznych

Natalia Kotsyba

Instytut Slawistyki PAN

Sesja MSH „Lingwistyka Korpusowa”, Warszawa, 27.V - 2.VI 2007

PoUKR

- Polsko-Ukraiński Korpus Równoległy
wersja prototypowa, 2004-2005

<http://corpus.domeczek.pl>

- Autorki:

Natalia Kotsyba, Instytut Slawistyki PAN

Magdalena Turska, Uniwersytet Warszawski

Polsko-ukraiński korpus równoległy

nie przecze mebel piękny na pewno *Не побрались, не знали,*
А он розпинають
 a jeszcze pies rasowy to już cywilizacja *Воду за подушине, а сина кують,*
Літа ти могоді
 Kolorowi, biali, czarni *... А оное під тиню*
 Chodźcie zwłaszcza wy, ludko *Опухла дитина - голоднес мре,*
 Przez na oścież bramy



Korpus

Autorki

Орест Друль

Orest Drul

Publikacje

Украина vs українськості

Ukraina vs. ukraińskość

Linki

Все дуже просто.

To bardzo proste.

Podziękowania

XXX

Простота вищого рівня – цивілізаційного масштабу: виклик глобалізації породжує посилення регіональної ідентифікації; посилення ідентифікації експлікується сепаристичними (у сенсі шлюбної сепарації) тенденціями, рухами [1], розмовами, статтями тощо; іншими словами – Європа додифундувала таки до Галичини у формі расла "Європа регіонів".

Prostota wyższa to prostota na skalę cywilizacyjną: wyzwania globalizacji rodzą nasilenie patriotyzmu regionalnego; nasilenie regionalnego patriotyzmu tłumaczy separatystyczne (w znaczeniu separacji małżeńskiej) tendencje, ruchy [1], rozmowy, artykuły itd.; innymi słowy, Europa dodyfundowała ostatecznie do Galicji pod postacią hasła „Europa regionów”.

Простота інша – на державну скалю: тенденція поглиблення не те що перманентного, але консеквентного невдоволення соціально-економічною ситуацією на тлі державної інституалізації України компрометує ідентифікацію з цією аутсайдерною країною; іншими словами – одержавлення тендітного поняття "Україна" як ідеалу всього, чого тільки можна уявити доброго: від мови (другої у світі за мітичним, але авторитетним французьким конкурсом) через кохання (ті ж Бітли співали про ukrainian girls – послухайте "Back in USSR" з т.зв. "White Album") до вершин демократії (тих, хто ще пам'ятає неспадкувану її конотацію, відсилаю до прецеденту конституції Орлика) – призвело, прогнозовано, зрештою, до настільки повної трансформації семантики слова, що утотожнюватися з новим його значенням стало вищою дзен-практикою. Зрозуміло, що вищі ступені цієї практики (як от, що рух – це застиглість) досяжні лише немоної щепті – тож питання вибору іншої ідентичності для неочудненої решти конгруентне kwestії часу.

Inny rodzaj prostoty to prostota w skali państwowej: tendencja pogłębiania się nie dość że permanentnego, ale też konsekwentnego niezadowolenia z sytuacji społecznej i ekonomicznej na tle państwowej instytucjonalizacji Ukrainy kompromituje poczucie związku z tym krajem-outsiderem; innymi słowy – upaństwowienie delikatnego pojęcia „Ukraina” jako ideału wszelkiego dobra, jakie tylko można sobie wyobrazić: od języka (drugiego na świecie według mitycznego, ale szanowanego francuskiego rankingu) przez miłość (nawet Beatlesi śpiewali o ukrainian girls – proszę posłuchać „Back in USSR” z tzw. „White Album”) do szczytów demokracji (pamiętających jeszcze pierwotne, czyste konotacje tego słowa odsyłam do precedensu konstytucji Filipa Orlika) – doprowadziło, zgodnie zresztą z przewidywaniami, do tak całkowitej przemiany semantyki słowa, że utożsamianie się z jego nowym znaczeniem to wyższa szkoła zen. Wiadomo, że wyższe arkana tej szkoły (jak np. to, że ruch jest bezruchem) dostępne są tylko nielicznej garstce – więc problem wyboru innej tożsamości dla niedouczonej reszty jest kongruentny z kwestią czasu.

Простота ще інакшіша – регіональна, прецінь галицька: Боже! а ми ж

Jeszcze inna prostota to prostota regionalna, wszakże galicyjska: Boże! a

Polsko-ukraiński korpus równoległy

nie przecze mebel piękny na pewno He pobralis',
a jeszcze ples rasowy to juz cywilizacja nie znalis',
Chodźcie wszystkie stany A on roztrinaють
Kolorowi, biali, czarni Witam ci mozozi
Chodźcie zwłaszcza wy, ludka Witam ci mozozi
Przez na oścież bramy ... A onoe nio minom
Отухла дитина - голоднес мре,

/kra[ʋw][ʌs]/ Szukaj

- Korpus
- Autorki
- Publikacje
- Linki
- Podziękowania
- XXX

Oryginał	Tłumaczenie
1 AIESEC	✗ AIESEC
7 czygier_op.xml	✗ czygier_tu.xml
9 downar_op.xml	✗ downar_tu.xml
5 Drul	✗ Drul
11 eurowybor_ou.xml	✗ eurowybor_tp.xml
13 finberg_ou.xml	✗ finberg_tp.xml
14 hrycak_ou.xml	✗ hrycak_tp.xml
16 jaworski_op.xml	✗ jaworski_tu.xml
29 kasjanow_ou.xml	✗ kasjanow_tp.xml
30 krawczyk_op.xml	✗ krawczyk_tu.xml
32 krym_ou.xml	✗ krym_tp.xml
27 kubisz_op.xml	✗ kubisz_tu.xml
3 List Burakowski	✗ List Burakowski
18 natoukr_ou.xml	✗ natoukr_tp.xml
20 nowystaw_op.xml	✗ nowystaw_tu.xml
34 obolonczyk_ou.xml	✗ obolonczyk_tp.xml
36 oniszczyk_ou.xml	✗ oniszczyk_tp.xml
37 pashawer_ou.xml	✗ pashawer_tp.xml
23 popowycz_ou.xml	✗ popowycz_tp.xml
39 riabczuk_ou.xml	✗ riabczuk_tp.xml
41 tozsamosc_op.xml	✗ tozsamosc_tu.xml
43 wolyn_ou.xml	✗ wolyn_tp.xml
45 wozniak_ou.xml	✗ wozniak_tp.xml
47 zacharow_ou.xml	✗ zacharow_tp.xml

© tuurma '05

Polsko-ukrai korpus r[eg]y - Microsoft Internet Explorer

Plik Edycja Widok Ulubione Narzędzia Pomoc

Wstecz Wyszukaj Ulubione

Adres http://localhost/corpus/index.php?option=query

Przejdź Łącz

oswici różnych kraj... mi j bo otrzymali дуже мало інформації про спроби організації шкільних занять з альтернативної історії за нашим прикладом. Наші мрії здійснилися б у тому разі, якби на подібних заняттях зустрічалися школярі з різних країн, наприклад, угорські та словацькі ліцеїсти. Думаємо, що варто було б "експортувати" ідею, яку ми реалізували під час семінарів-лабораторій: цікаво, чи виправдала б вона себе в інших культурних контекстах.

370 naszych marzeń byłaby sytuacja, w której uczniowie z różnych krajów, np. licealiści węgierscy i słowaccy, spotykali się

Годиться лише порушити питання про масштаби цього явища та можливості провадження аналогічних навчальних експериментів у системі шкільної освіти різних країн. Ми ж бо отримали дуже мало інформації про спроби організації шкільних занять з альтернативної історії за нашим прикладом. Наші мрії здійснилися б у тому разі, якби на подібних заняттях зустрічалися школярі з різних країн, наприклад, угорські та словацькі ліцеїсти. Думаємо, що варто було б "експортувати" ідею, яку ми реалізували під час семінарів-лабораторій: цікаво, чи виправдала б вона себе в інших культурних контекстах.

371 wpłynie zróżnicowanie programów i metod nauczania historii w krajach, których przedstawiciele zapraszaliśmy. Udało się to

Нас непокоїло питання про те, як на перебігу семінарів-лабораторій позначається розбіжності в програмах і методиках навчання історії у тих країнах, чіх представників ми запросили. Ці відмінності вдалося звести до мінімуму завдяки добору шкіл. Польщу репрезентували переможці останніх конкурсів "Шкільна історична лім" і в яких постійно перемагають школи з провінційних міст. А, наприклад, Україну представляли школи, що входять до елітарної Асоціації коледжів. Тож молоді поляки не мали переваги. Що більше, можна сказати, що на заняттях віддавали певну перевагу гостям.

372 prywatyzacja Przez ukraińskie społeczeństwo oraz krajowych i zagranicznych analityków prywatyzacja została przyjęta

Приватизація була сприйнята українським суспільством, вітчизняними й зарубіжними аналітиками вельми неоднозначне. Справедливі претензії висунуто і до технології приватизації, і до її економічних та соціальних наслідків.

373 bezpośredniego wykorzystania tradycyjnych metod prywatyzacji krajów rozwiniętych – sprzedaży na rynku papierów wartościowych

— неможливості безпосереднього використання традиційних для розвинутих країн методів приватизації — продажу на фондовому ринку та ринку нерухомості, оскільки ці ринки не були сформовані;

374 społeczne. 3. Znacznie gorsze niż w innych europejskich krajach postsocjalistycznych skutki prywatyzacji tłumaczy się przede

3. Значно гірші, ніж в інших європейських постсоціалістичних країнах, наслідки приватизації пояснюються насамперед не вадами обраної моделі, а специфічністю умов, в яких здійснювався процес приватизації. Україна об'єктивно мала чи не найгірші стартові умови для ринкових реформ у цілому і для приватизації зокрема. Основні підстави для такого висновку:

375 o ile we wszystkich krajach państwo odegrało kluczową rolę w procesie reform, to

— якщо в усіх країнах держава відігравала ключову роль в процесі реформ, то, Україна на початку 90-х років не мала не тільки сформованих державних інститутів, але й історично близького досвіду державного будівництва (на відміну від країн Прибалтики) або успадкованої від СРСР сформованої державної машини (як у Росії);

376 doświadczenia w tworzeniu państwowości (w przeciwieństwie do krajów bałtyckich) lub odziedziczonej po ZSRR, ukształtowanej

— якщо в усіх країнах держава відігравала ключову роль в процесі реформ, то, Україна на початку 90-х років не мала не тільки сформованих державних інститутів, але й історично близького досвіду державного будівництва (на відміну від країн Прибалтики) або успадкованої від СРСР сформованої державної машини (як у Росії);

377 s kraj nie niesprzyjające warunki społeczno-polityczne: dominacja

' — вкрай несприятливі суспільно-політичні умови: домінування соціалістичних та колективістських ідеологічних стереотипів у основній масі населення; довготривале домінування відвертих противників приватизації в представницьких органах влади, в тому числі у Верховній Раді.

378 Najwyższej. 4. W wyniku szeregu przyczyn obiektywnych, kraj e Europy Środkowej zdołały rozpocząć proces radykalnych reform

4. Внаслідок цілого ряду об'єктивних причин, країни Центральної Європи змогли розпочати процес радикальних ринкових реформ не з приватизації, а з так званої "маркетизації", яка включала лібералізацію та стабілізацію економіки. На відміну від цих країн, де приватизовані підприємства потрапили у вже підготовлене ринкове середовище і з перших же кроків проявляли свої переваги, в Україні створення потужного недержавного сектора стало необхідною політикоекономічною умовою для доведення ринкових реформ до тієї глибини, коли вони приносять плоди.

379 i stabilizację gospodarki. W przeciwieństwie do tych krajów, w których prywatyzowane przedsiębiorstwa trafiły w

4. Внаслідок цілого ряду об'єктивних причин, країни Центральної Європи змогли розпочати процес радикальних ринкових реформ не з приватизації, а з так званої "маркетизації", яка включала лібералізацію та стабілізацію економіки. На відміну від цих країн, де приватизовані підприємства потрапили у вже підготовлене ринкове середовище і з перших же кроків проявляли свої переваги, в Україні створення потужного недержавного сектора стало

Gotowe Lokalny intranet

Start C:\WINDOWS\... teksty 2 Internet Ex... Microsoft Power... PL 08:15

Czemu ma służyć korpus równoległy?

- Baza danych odpowiedników słów i wyrażeń oraz ich kontekstów dla tłumaczy
- Punkt wyjścia do konstrukcji wiarygodnych **słowników dwujęzycznych**
- Baza danych dla konfrontatywnej analizy lingwistycznej

Integrowany materiał językowy

- korpus przedstawia układ syntagmatyczny
- słownik/gramatyka – układ paradygmatyczny
- korpusy + słowniki = uniwersalny bank danych językowych
- im bogatsza anotacja, tym prostsze jest automatyczne przejście między układami
- integracja danych syntagmatycznych i paradygmatycznych ułatwia dalszą automatyczną anotację

Hasło z „papierowego” słownika polsko-ukraińskiego (red. Humeckiej)

- **popęłniać, popęłnić**, -nię, -nisz, -nił робити, -блю, -биш, зробити, чинити, -ню, -ниш, учинити; ~nić błąd зробити (учинити) помилку; ~nić nietakt **зробити** безтактність (нетактовність); ~nić zbrodnię **учинити** злочин; <*> *перен. жарт.* ~nić powieść розродитися романом; ~nić samobójstwo заподіяти собі смерть, накласти руки на себе, покінчити з собою, скінчити життя самогубством.
- **скоїти злочин, безтактно повестися**

Słownik papierowy

- **interferencja** sytuacyjna - „papierowy” leksykograf działa pod wpływem przeczytanego przed chwilą polskiego tekstu i używa wyrazów, których w jednojęzycznym środowisku nie użyłby;
- podstawą opisu są konkretne użycia językowe (**kontekst**), a nie wyodrębnione znaczenia
- często znaczenie jest wyrażane poprzez typową **ramę walencyjną** (komu, czemu; pod co; na czym; kogo, co, itd.)

Wiarygodność korpusów

- (Fillmore 1992) badanie mające na celu porównanie haseł słownikowych angielskiego wyrazu *risk* z danymi uzyskanymi korpusu
- hasła z 10 słowników vs rezultat, uzyskany poprzez analizę przykładów z 25-milionowego korpusu
- jedynie dwa słowniki odnotowywały wszystkie trzy znaczenia wyrazu
- żaden z nich nie odnotowywał wszystkich istotnych kolokacji (źródło - Wójtowicz, 170)

Wiarygodność korpusów

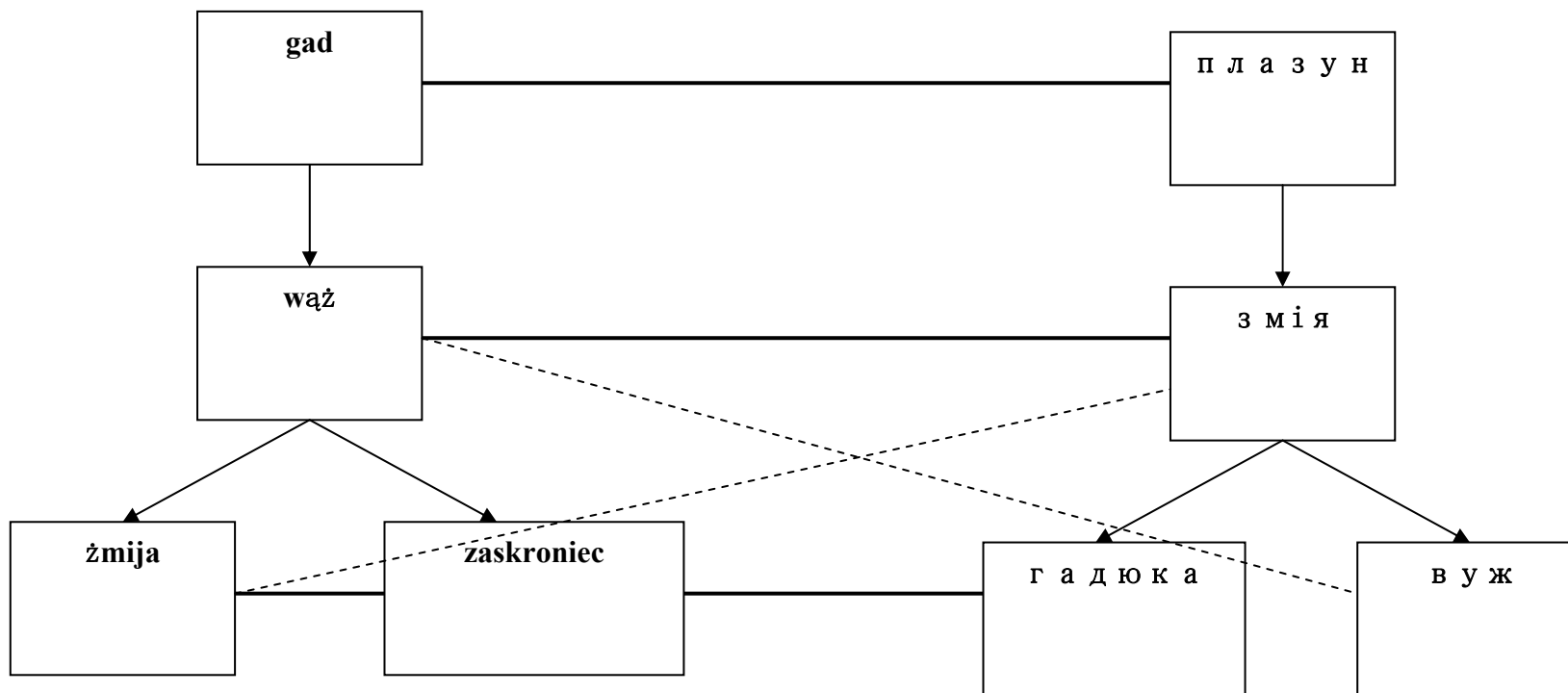
■ Podejść – podchodzić

„mieć stosunek do czegoś” np. *podejść do problemu* 20% użyć w Korpusie IPI PAN, nie zanotowano w SJP Doroszewskiego, SJP PWN

■ Kwiatek

„nieoczekiwany wynik”, iron.
10% użyć, też nie zanotowane

Homonimia międzyjęzykowa



	PUNazJur	PUAleks	PULewSta	PUHryWas
gad	1. плазун, 2. гад, гадин	1. плазун (зоол.) 2. гад (перен.)	зоол. плазун	гад, gady – гаддя
wąż	1. <u>вуж</u> , змія. 2. ~ гумовий – шланг	1. змія (зоол.) 2. шланг, рукав	1. зоол. <u>вуж</u> , змія 2. шланг, рукав	<u>вуж</u>
zaskroniec	вуж звичайний	вуж звичайний (зоол.)	зоол. вуж	вуж
żmija	<u>змія</u> , гадюка	<u>змія</u> , гадюка (зоол., перен.)	<u>змія</u> , гадюка	<u>змія</u> , гадин

	UPNazJur	UPHryWas	UPHrabiec	UPlwcz
вуж	<u>Wąż</u>	<u>Wąż</u>	<u>Wąż</u>	<u>Wąż</u>
змія	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. Imn – gady, płazy, przen. gad, gadzina

Czasowniki

- Nie zbyt łatwy materiał dla opisu leksykograficznego
- Problem przeładowanego opisu czasowników „lekkich” (light verbs)
- Problem różnicy kombinatoryki
- Przenoszenie znaczeń gramatycznych do dziedziny leksyki

Realizacje w świecie: kryteria i zasady klasyfikacji (czasowniki)

- klasyfikacja Beth Levin bazowana na zachowaniu składniowym czasowników (alternacje)
- Zeno Vendler (Tatiana Bułygina – Jurij Apresjan): kompatybilność z przysłówkami temporalnymi
- klasyfikacja Galiny Zołotowej: bazowana na znaczeniu oraz lingwistycznych charakterystykach podstawowych argumentów (podmiot, dopełnienie)

Beth Levin i Malka Rappoport-Hovav (1993, 1995)

■ Alternacje

Alternacja jest to regularna wariantywność w syntaktycznym uzewnętrznieniu argumentów czasownikowych. Zasada jest taka, że dwa zdania tworzą alternacje, jeśli ich znaczenie jest identyczne (bądź przynajmniej zbliżone) lub też znaczenie jednego jest uszczegółowieniem znaczenia drugiego (zawieranie). W pierwszym przypadku obie alternacje muszą mieć tę samą liczbę argumentów. (wg E. Hajnicz)

Przykłady alternacji (E. Hajnicz)

- alternacja **konatywna** (ang. *conative*):

Jan uderzył drzwi. - Jan uderzył w drzwi. - Jan uderzył się o drzwi.

- alternacja **datywna** (ang. *dative*):

Jan posłał książkę Piotrowi / do Piotra.

- alternacja **benefaktywna** (ang. *benefactive*):

Piotr zostawił/przekazał wiadomość Annie / dla Anny.

- alternacja **lokatywna** (ang. *locative*):

Piotr (za)ładował ziemniaki na wóz / do worka.

- alternacja **kauzatywna** (ang. *causative*):

Piotr zbił szybę (piłka).

Piłka zbiła szybę.

Semantyczna klasyfikacja czy składniowa klasteryzacja?

Czasowniki *ruchu* (***Motion***) Levin zawierają 9 klas:

- z wbudowanym znaczeniem kierunku *arrive, go, ...*
- czasowniki odejścia (leave verbs)
- sposób ruchu (manner of motion):
 - **toczenie się (roll verbs) *bounce, float, move, ...***
 - **bieganie (run verbs) *bounce, float, jump, ...***
- sposób ruchu z użyciem środka transportu:
 - czasowniki, nazywające środki transportu (vehicle name verbs) *bike, ...*
 - czasowniki, nie kojarzone z nazwami środków transportu *fly, ...*
- czasowniki walcowe (waltz verbs) *boogie, polka, ...*
- czasowniki ścigania (chase verbs) *follow, pursue, ...*
- czasowniki towarzyszenia (accompany verbs).

Klasyfikacja jest nieortogonalna i niepełna: brakuje innych sposobów ruchu, środków ruchu, itd.

Semantyczna klasyfikacja czy składniowa klasteryzacja?

Czasowniki *przekazywania własności* (*transfer of possession*) Levin:

- czasowniki dawania (give verbs) *feed, give, lease, ...*
- czasowniki wkładu (contribute verbs) *distribute, donate, submit, ...*
- dostarczenie (providing):
 - wykonanie (fulfilling verbs) *credit, provide, ...*
 - sprzętowe (equip verbs) *arm, invest, ...*
- otrzymanie (obtaining):
 - Get (**book, buy, call, cash, order, phone**, ...),
 - Obtain (accept, accumulate, seize, ...),
- czasowniki z komponentem przyszłościowym (future having verbs) *advance, assign, ...*
- czasowniki wymiany (exchange verbs)
- Berry verbs (nest, clam, ...).

Walencja: perspektywa wielojęzykowa

- szczotkować – чистити (щіткою), hiperonim

szczotkować

[arg1= x:agent]

[arg2 = y:object]

ЧИСТИТИ

[arg1= x:agent]

[arg2 = y:object]

[arg3 (opt) = z:instrument]

Walencja: perspektywa wielojęzykowa

- *Ukraińsko-polski słownik syntaktyczny*, red. O. Śpiwak i M. Jurkowski, 2003.
- zawiera 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.

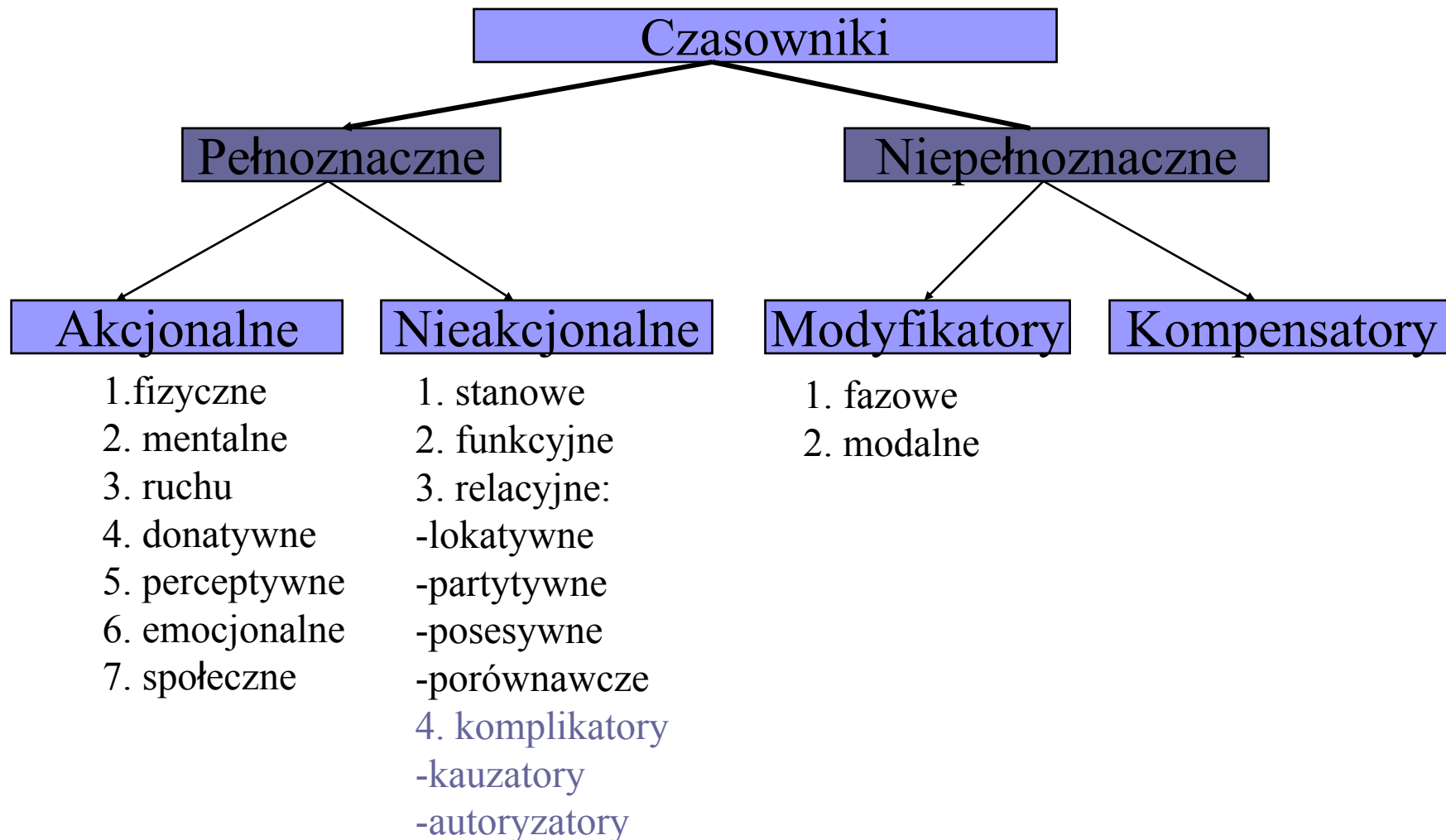
Klasyfikacja Galiny Zołotowej (1998)

„Gramatyka komunikatywna j.ros.”

Zalety:

- pozwala uwzględnić **wszystkie** czasowniki
- klasyfikuje **znaczenia**, a nie wyrazy
- **konsekwentny** podział znaczeń na gramatyczne (półpredykatywne) i leksykalne (*pełnopredykatywne)

Klasyfikacja Złotowej



Przykłady czasowników **pół**predykatywnych

- **x wyróżnia się** talentem = x (jest) utalentowany; x - талановитий (kompensator)
- **przewodzić** badania = badać (kompensator)
- **oceniać, rozpatrywać, uważać za** (autoryzatory)
- **zaciągać** kredyt, **zachodzić** w ciążę; **взяти** кредит, **завагітніти** (modyfikator fazowy)

Odpowiedniki formalne vs ekwiwalenty tłumaczeniowe

- **prowadzić**, -dzieć, -dzisz, -dził **вести** (веду, ведеш); (*проводити, робити* що — ще) **провадити**, -джу, -диш; *перен.* ~ do celu вести до мети; ~ dom вести gospodarstwo; ~ dysputę вести (провадити) диспут; *торг.* ~ interes вести (провадити) діло; *перен.* ~ kogo na rasku вести кого на поводі (на повідку); *бухг.* ~ księgę вести книги; ~ obradę вести (провадити) наради; ~ robotę вести (провадити) роботи; ~ rozmowę вести розмову; ~ wojnę вести (провадити) війну; ~ życie вести (провадити) життя; <*> ~ kogo wzrokiem **проводжати** кого очима (поглядом); ~ konie **правити** кіньми, **поганяти** коней; ~ łaską po piasku **водити** палицею по піску.
- Brak znaczenia „kierować pojazdem” – „їздити за кермом”

Przykłady czasowników **polipredykatywnych**

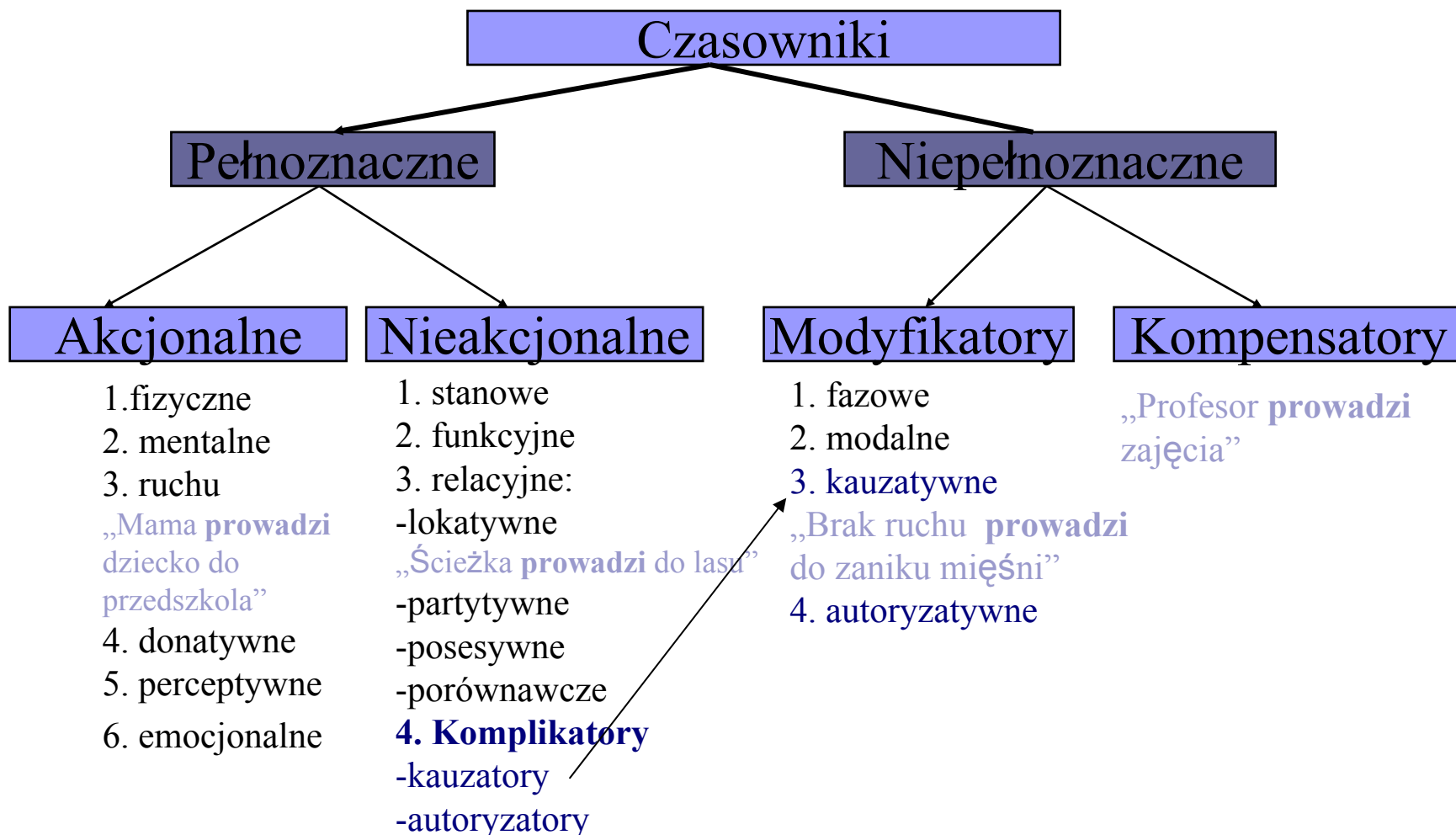
- **zaistnieć** = zacząć być (nieakc. relacyjny + modyfikator fazowy)
- **prowadzić** x = powodować, że x się porusza (akcjon. ruchu + kauzacja)

Przykład języka mandżurskiego

- om**i**umbi – poi, om**i**mbi – pije;
- gele**b**umbi – straszy, gele**m**bi – boi się;
- songǧo**b**umbi – doprowadza do łez, songǧo**m**bi – płacze;
- taci**b**umbi – uczy, taci**m**bi – uczy się;
- te**b**umbi – sadzi, tem**i**bi – siedzi.

Ekwiwalenty angielskie? Rosyjskie? Ukraińskie?
Białoruskie?

Klasyfikacja Zołotowej (w trakcie bootstrappingu)



Przykład analizy hasła ze zgramatyzalizowanymi znaczeniami

prować (słownik PWN)

- 1. «wieć kogoś, coś do jakiegoś miejsca, do celu; przeprowadzać kogoś, kto nie zna drogi albo sam iść nie może, wskazywać komuś drogę» - **pełnoznanzy, akcjonalny ruchu + kauzacja** (akcjonalny ruchu czy nieakcjonalny komplikator???) dk odprowadzić, zaprować - **вести**
- 2. «stawać się przyczyną czegoś; powodować, wywoływać coś, pociągać za sobą coś» - **pełnoznanzy, nieakcjonalny komplikator (kauzatywny)** dk doprowadzić/ zaprować - **вести**
- 3. «kierować pojazdem mechanicznym, statkiem, koźmi» =1 **brak dk (бути) за кермом, за штурвалом, верхи**
- 4. «kierować partnerem w tańcu» =1 **brak dk - вести**
- 5. «sprawować nad czymś nadzór, zarządzać czymś, zajmować się, trudnić się czymś; kierować czymś» - **niepełnoznanzy, kompensator** dk poprowadzić, przeprowadzić **проводити (заняття)**
- 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» - **pełnoznanzy, nieakcjonalny, relacyjny, lokatywny doprowadzić**
- 8. «o zawodnikach, drużynach sportowych: być pierwszym w klasyfikacji, mieć przewagę nad przeciwnikiem; przodować» - **pełnoznanzy, nieakcjonalny, relacyjny, porównawczy**

Bardziej subtelne rozbieżności w znaczeniach ekwiwalentów – sposoby na łatanie dziur

■ zabrać – забрати

zabrać kogoś do miasta/ klubu; dziecko do cyrku/kina

**забрати когось до міста/ клубу; дитину до цирку/ в кіно*

Próbka Korpusu IPI PAN - [base="zabrać"][orth="do"]:

Wkrótce na miejscu zdarzenia zjawiała się policja, a pijaka **zabrano do** izby wytrzeźwień. **akcjonalny, ruchu + kauzacja** **завезли**

... rozdzielono mnie z bratem, którego **zabrano do** Brzezinki. **забрали**

Mechanika **zabrano do** szpitala po zawinięciu do Los Angeles. Wymieniono też część załogi. **завезли**

Schematy struktury znaczeniowej

zabrać – забрати

- **wspólny** dla pol. i ukr.

arg1 = x:anim ◦ loc1

arg2 = y:anim ◦ loc2

y CAUSE x (loc1 → loc2)

- **tylko w jęz. pol.**

arg1 = x:anim ◦ loc1

arg2 = y:anim ◦ loc1

y CAUSE x^y (loc1 → loc2)

Czasowniki **ruchu** (kierunku), notacja Wojtasiewicza

x **PRZYBYŁ** tutaj: $\exists y, z \text{ Ag}(x, \text{Trans} (L(x, y), L(x, z)))$
 $\wedge \underline{L(r, z)} \wedge (x \text{ e Hum})$; r = punkt odniesienia

x **ODSZEDŁ** stąd: $\exists y, z \text{ Ag}(x, \text{Trans} (L(x, y), L(x, z)))$
 $\wedge \underline{L(r, y)} \wedge (x \text{ e Hum})$; r = punkt odniesienia

x **UDAŁ SIĘ** tam: $\exists y, z \text{ Ag}(x, \text{Trans} (L(x, y), L(x, z)))$
 $\wedge \underline{\sim L(r, z)} \wedge (x \text{ e Hum})$; r = punkt odniesienia

Zestaw podstawowych predykatów Wojtasiewicza

- **L(a,b)**, L – relacja lokalizacji w przestrzeni, a – n lub z, przedmiot lokalizowany lub zdarzenie lokalizowane, b – lokalizator (przedmiot lokalizujący).
- **T(a,b)** – relacja lokalizacji w czasie.
- **Ex(a)** – istnienie materialne (nie mylić z życiem, np. *zwłoki istnieją*)
- **Trans(a,b)** – przejście z jednego stanu do innego; a i b – zdania – stany. **ZMIANA**
- $\text{Trans}(a,b) \rightarrow \forall t, t' \text{ Trans}(at, bt) \wedge (t < t')$, < - relacja wcześniejszości; t, t' - czas.
- **Ag(a,b)** – relacja między agensem i wynikiem jego działania; a – agens, b – wynik;
- „a działa tak, **ŻE** b”, nie zaś „a działa tak, **ŻEBY** b” – nie mylić z **Caus**; $\text{Ag}(a,b) \rightarrow \forall t, t' \text{ Agt}(a, bt') \wedge (t \leq t')$ **ODDZIAŁYWANIE**
- **B(a,b)** – relacja zachodząca pomiędzy podmiotem żywiącym pewne przekonanie a tymże jego przekonaniem; ograniczenia selekcyjne – a=hum (istoty ludzkie lub grupy oraz przedmioty intencjonalne traktowane antropomorficzne);
- **S(a,b)** – relacja zachodząca pomiędzy nadawcą informacji a samą tą informacją; a – nadawca, b – treść.
- **Exp(a,b)** – podmiot doznający a doznanie
- **M(a,b,c)** – relacja przypisywania czemuś przez kogoś lub przez coś jakiejś miary; a – mierzący, b – obiekt pomiaru, c – miara w wyniku pomiaru.
- **V(a,b,c)** – przypisywanie czemuś przez kogoś jakiejś wartości; a=hum, b – zdarzenie wartościowania, c – nazwa wartości;
- **Prob(a,[i,j])** – relacja pomiędzy pewnym zdarzeniem a przypisywanym mu prawdopodobieństwem subiektywnym;
- [i,j] – przedział zamknięty będący miarą p ($0 \leq i \leq j \leq 1$) – warunki
- B(..., Prob(a,[i,j])...); podmiot przekonania, miara i fakultatywnie treść
- **Pi(a)** – a znajduje się w stanie i;
- **Po(a)** – a znajduje się w stanie dla siebie standardowym

Stany i zmiany stanów

- **x JEST ZDROWY:** $Po(x) \wedge (x \in Anim)$
- **x MA k y-ów:** $Po(x) \rightarrow ((y^* <^* x) \wedge M(s, y, k))$
- **x CHORUJE:** $\sim Po(x) \wedge (x \in Anim)$
- **x ZDROWIEJE:** $Trans(\sim Po(x), Po(x)) \wedge (x \in Anim)$
- **x ZACHOROWAŁ:** $Trans(Po(x), \sim Po(x)) \wedge (x \in Anim)$ NK: is falling ill
- **x CIERPI:** $\forall y Exp(x, \sim Po(x)) \wedge \forall (x, \sim Po(x), y) \wedge (y < 0)$ x doznaje tego, że jest w stanie niestandardowym, i przypisuje temu wartość ujemną
- **x PSUJE SIĘ:** $Trans(Po(x), \sim Po(x)) \wedge (x \in Org)$, gdzie Org jest substancją organiczną
- **x PSUJE y:** $Ag(x, Trans(Po(y), \sim Po(y)))$
- **x NAPRAWIA y:** $Ag(x, Trans(\sim Po(y), Po(y))) \wedge (x \in Hum \vee Aut)$ – automat
- **x LECZY y:** $Ag(x, Trans(\sim Po(y), Po(y))) \wedge (x \in Hum) \wedge (y \in Anim)$ – z domniemaniem, że skutecznie wylecza
- **x LECZY2 y:** $\exists i, j Py(z) \rightarrow Prob (Ag(x, Trans(Py(z), Po(z))), [i, j]) \wedge large(i) \wedge (x \in Med) \wedge (y \in Dis) \wedge (z \in Anim)$

Jeżeli ktoś ma zapalenie płuc, to jest wielce prawdopodobnie, że penicylina go z tego wyleczy

- **x SAMOREGULUJE SIĘ:** $Trans t(Po(x), \sim Po(x)) \rightarrow Ag t'(x, Trans(\sim Po(x), Po(x))) \wedge (t < t')$

Uproszczenie ról semantycznych, własności vs czynności

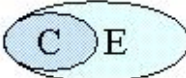
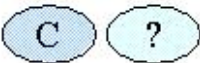
Atrybut (identyfikujący vs charakteryzujący)

- Różne podejścia ontologiczne: Jadacki – relacja przypisywania x-owi własności P; Wojtasiewicz – stan P, przypisywany x-owi
- $\sim Po(x)$ He is crazy
- $Trans (Po(x), \sim Po(x))$ He got crazy
- $Ag (y, Trans (Po(x), \sim Po(x)))$ He drives me crazy

(wariat, wariuje, doprowadza do szału)

Uniwersalny Język Sieciowy (UNL)

Concepts and English Words

	The English word has no ambiguity	'thing', 'scientist', 'beautiful'
	The English word needs to be disambiguated	'dance(ict>action)' 'dance(agt>volitional thing)'
	A non-English word: Extra UW	「畳」 <i>tatami</i> (a particular mat made of grass and used to floor a Japanese-style room)
	A collection of different concepts	「話す」 <i>hanasu</i> corresponds to: talk, chat, tell, speak, mention, state, narrate
	A phrase: E + additional words	「上下」 <i>jouge</i> corresponds to: the top and bottom, the upper and lower parts, high and low

Uniwersalny Język Sieciowy (UNL)

- UNL jest formalnym językiem dla reprezentacji informacji w sposób, pozwalający na generowanie tekstu, wyrażającego tę informację w dużej ilości języków naturalnych.
- wyrażenie (expression) UNL jest schematem (hyper-graph), który odpowiada zdaniu języka naturalnego pod względem ilości przekazywanej informacji.
- węzły (nodes) reprezentują jednostki specjalne, tzw. „słowa uniwersalne” Universal Words (UW), interpretowane jako pojęcia, albo grupy UW. Węzły mogą mieć atrybuty, dostarczające dodatkowych informacji o ich użyciu w danym zdaniu, np. *@imperative, @generic, @future, @obligation*.
- węzły są łączone łukami (arcs) interpretowanymi jako relacje semantyczne: *agent, object, time, place, manner*, etc.

Słownik UNL:

- NAME: divorce(agt>human)
- ZONE:EN
 - TRANS: divorce
- ZONE:RU
 - <none>

- NAME: marry(agt>human)
- ZONE:EN
 - TRANS: marry
- ZONE:RU
 - <none>

- NAME: marry(agt>male)
- ZONE:EN
 - <none>
- ZONE:RU
 - TRANS: zhenit'sja

Słownik angielski

NAME: divorce

ZONE: RU

TRANS: razvodit'sja

ZONE:UNL

TRANS: divorce(agt>human)

NAME: marry

ZONE: RU

TRANS: zhenit'sja / vyxodit' zamuzh

ZONE:UNL

TRANS: marry(agt>human)

Słownik rosyjski

- NAME: razvodit'sja
 - ZONE: EN
TRANS: divorce
 - ZONE:UNL
TRANS: divorce(agt>human)
-
- NAME: zhenit'sja
 - ZONE: EN
TRANS: marry
 - ZONE:UNL
TRANS: marry(agt>human)

Inwentarz deskryptorów semantycznych

- bazowany na wersji „fundamentalnej klasyfikacji predykatów” Jurija Apresjana
- klasy wspólne (i odpowiednie znaczniki) z podobnymi systemami (Jurij Maslov, Zeno Vendler, Tatiana Bulygina, Elena Paducheva, Charles Fillmore): ‘action - czynność’, ‘activity - działalność’, ‘process’, ‘state - stan’, ‘property - własność’, itd.
- nowe klasy: ‘occupation - zajęcie’, ‘behaviour – zachowanie się’, ‘impact - wpływ’, ‘spatial position – pozycja przestrzenna’, ‘interpretacja’, itd.

Apresian – klasyfikacja „fundamentalna”

- **Czynności** (действия): *atakować, iść, pisać*: redukują się do „robić” – sem prymityw, czas trwania nie przekracza jednego aktu obserwacji
- **Działalności**: *wojować, wychowywać, uczyć, handlować*: czynność wielokrotna, powtarzana; różnoczasowe; jeden wspólny cel
- **Zajęcia**: *spacerować, grać, odpoczywać, czytać* vi: czynność, której celem jest ona sama
- **Zachowania**: *marudzić, wygłupiać się*: czynność + ocena
- **Oddziaływania**: *słońce rozgrzewa powietrze do 20 stopni*; *feed::eat*, *rozgrzewa::rozgrzewa się*: redukuje się do „bycia przyczyną”, czyli kauzacji
- **Procesy**: *powracać do zdrowia, rosnąć, gotować się*: zmiana, która zachodzi bez woli podmiota; zawiera elementy fazowe
- **Przejawy**: *błyszcząć, śmierdzieć, kłóć się, brzmieć*: własność, stan albo proces, percepowane aktualnie przez obserwatora
- **Zdarzenia**: *napotkać (przypadkiem), znajdować, zdarzać się*
- **Położenia w przestrzeni**: *wisieć, leżeć, siedzieć, opierać się, stać* POSITIONER JAK, bardziej złożone; być położonym jakoś
- **Lokalizacje**: *okazać się, być², znajdować się* GDZIE: znajdować się. Między obiektem i miejscem lokalizacji nie ma związku organicznego (*paznokcie na stole vs na nogach*; *włosy na głowie vs na podłodze/ grzebieniu*)
- **Stany**: *wiedzieć, cieszyć się, chcieć, odczuwać*; statywy
- **Własności**: *jakać, tonać* (żelazo tonie w wodzie), *картавить, виться* (о волосах); potencjalność
- **Umiejętności**: *mówić po polsku, prowadzić, jeździć konno, chodzić* (dziecko już chodzi); inchoatywność
- **Parametry**: *ważyć, umieszczać, trwać, pojawiać się, liczyć*
- **Istnienia**: *bywać, pojawiać się, istnieć*
- **Stosunki**: *mieścić, zawierać, być równym* (равняться)
- **Interpretacje**: *grzeszyć, mylić się, zawodzić*: zawierają ocenę i porównywanie z normą

Funkcje leksykalne

- **OPER1** [robić X, mieć X albo być w stanie X (P1 jako podmiot gramatyczny czasownika wspomagającego i X jako dopełnienie główne)];

por. *dokonać (wyboru)*, *przeprowadzić (wywiad)*, *rzucić (spojrzenie / okiem)*,
mieć (większość) / być w (większości). kompensator czynny

- **OPER2** [poddawać się czynności X albo być w zasięgu X (P2 jako podmiot gramatyczny i X jako dopełnienie główne)];

por. *być pod (wpływem)*, *udzielać (wywiadu)*, *cieszyć się (powodzeniem u kogoś)*, *poddać się (badaniom)*. kompensator bierny

- **INCEOPER1** [zacząć robić X, mieć X albo być w stanie X (P1 jako podmiot gramatyczny i X jako dopełnienie główne)];

por. *złapać (grypę)*, *nabrać (wagi)*, *dojść do (władzy)*; ang. *to fall in (love)* (vs. pol *zakochać się*). Połączenie modyfikatora fazowego *inch.* i kompensatora

- **FINOPER1** [przestać robić X, mieć X albo być w stanie X (P1 jako podmiot gramatyczny i X jako dopełnienie główne)];

por. *zerwać ze (znajomością)*, *wyjść z (interesu)*, *stracić (kontrolę)*, *rozstać się z (życiem)*. Połączenie modyfikatora fazowego *term.* i kompensatora

Funkcje leksykalne c.d.

- **REAL1-M** [robić w stosunku do X to, **co jest zwykle oczekiwane** od P1 (P1 jako podmiot gramatyczny i X jako dopełnienie główne)];
por. *wykonywać (swoje obowiązki)*, *dotrzymać (obietnicy)*, *sprostać (oczekiwaniom)*.
- **INCEPREAL1-M** [zacząć robić w stosunku do X to, co jest zwykle oczekiwane od P1 (P1 jako podmiot gramatyczny i X jako dopełnienie główne)];
por. *dowiedzieć się (wiadomości)*, *przyjąć (zasadę)*, *świadczyć (przysługę)*.
- **FINREAL1-M** [przestać robić w stosunku do X to, co jest zwykle oczekiwane od P1 (P1 jako podmiot gramatyczny i X jako dopełnienie główne)];
por. *rzucić (próby)*, *zrzec się (tronu)*, *porzucić (nadzieję)*.
- **CAUSREAL1-M** [przyczynić P1 do robienia w stosunku do X tego, co jest zwykle oczekiwane od P1 (P0 jest podmiotem gramatycznym czasownika wspomagającego, P1 jest dopełnieniem głównym, X jest dopełnieniem wtórnym)];
por. *nałożyć na (kogoś) (obowiązek)*, *związać (kogoś) (obietnicą)*.
- **LIQUREAL1-M** [przyczynić P1 do nie robienia w stosunku do X tego, co jest zwykle oczekiwane od P1 (P0 jest podmiotem gramatycznym czasownika wspomagającego, P1 jest dopełnieniem głównym, X jest dopełnieniem wtórnym)];
por. *wybaczyć (komuś) (dług)*, *zwolnić (kogoś) z (obowiązku)*.

Literatura

- Jurij D. Apresjan, Igor M. Boguslavsky, Leonid L. Iomdin, Leonid L. Tsinman *Lexical Functions as a Tool of ETAP-3, SMTT 2003, Paris, 16-18 juin 2003*
- Hajnicz, E. Dobór czasowników do badań przy tworzeniu słownika semantycznego czasowników polskich, materiały seminarium NLE IPI PAN, 19 marca 2007.
- 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.
- Levin, B. i Rappaport Hovav, M. (2005) *Argument Realisation*, Cambridge University Press
- 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.
- 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.