

## Dwujęzyczna Słownosieć – możliwości wykorzystania w pracy tłumacza

---



Ewa Rudnicka,  
Tadeusz Piotrowski  
Politechnika Wroclawska  
Katedra Inteligencji  
Obliczeniowej  
Grupa Naukowa G4.19  
[ewa.rudnicka@pwr.wroc.pl](mailto:ewa.rudnicka@pwr.wroc.pl)

# Plan

---



- Słowność vs. WordNet Princeton:
  - specyfika struktury zasobów
  - różnice w opisie znaczeń jednostek leksykalnych i synsetów
- strategia powiązania zasobów i jej etapy
- system relacji międzyjęzykowych:
  - definicje, hierarchia oraz przykłady relacji
  - relacje specyficzne dla różnych części mowy
  - procedura rzutowania
  - liczności relacji

# Słownosieć vs Princeton WordNet

---



- Synsety - zbiory jednostek leksykalnych:

**pIWN:** wchodzących w te same relacje konstytutywne  
(np. **hiponimia, hiperonimia, meronimia, antonimia**)  
(cf. **Piasecki et al. 2009, Maziarz et al. 2012**)

**PWN:** będących w relacji synonimii, traktowanej jako relacja konceptualna ustalana na podstawie intuicji językoznawcy oraz definicji słownikowych (cf. **Fellbaum 1998**)

- Glosy:

**pIWN:** glosy dla jednostek;

**PWN:** definicje dla synsetów

- Relacje:

**pIWN:** bogaty zbiór synsetów i relacji jednostek leksykalnych

**PWN:** nieduży zbiór relacji synsetów i jednostek leksykalnych

# Liczności synsetów

---



- **Średnia liczba jednostek w synsecie:**
- Słownosieć: 1.32
- Princeton WordNet: 1.74
- Synsety rzeczownikowe Słownosieci: 1.35
- Synsety rzeczownikowe Princeton WordNet: 1.76
- Synsety przymiotnikowe Słownosieci : 1.16
- Synsety przymiotnikowe Princeton WordNet: 1.65

# Etapy budowy dwujęzycznej Słownosieci

---



Zaprojektowanie procedury rzutowania **czyli łączenia synsetów Słownosieci z synsetami WordNetu Princetonkiego na podstawie** podobieństwa znaczeń oraz miejsca w strukturze sieci

Zdefiniowanie zestawu relacji międzyjęzykowych oraz ustalenie hierarchii ich wprowadzania  
**(cf. Rudnicka et al. 2012)**

Zaprojektowanie systemów automatycznych podpowiedzi dla leksykografów  
**(Kędzia et al. 2013, Rudnicka et al. 2015)**

# Kierunek rzutowania

---



- Kierunek rzutowania: **Słowność > Princeton WordNet**  
od najniższych poziomów sieci stopniowo obejmując  
coraz wyższe
- Ustalenie kolejności rzutowania ze względu na  
kategorie gramatyczne i dziedziny wordnetowe
- Obecnie rzutowane: rzeczowniki i przymiotniki

# Relacje międzyjęzykowe

---

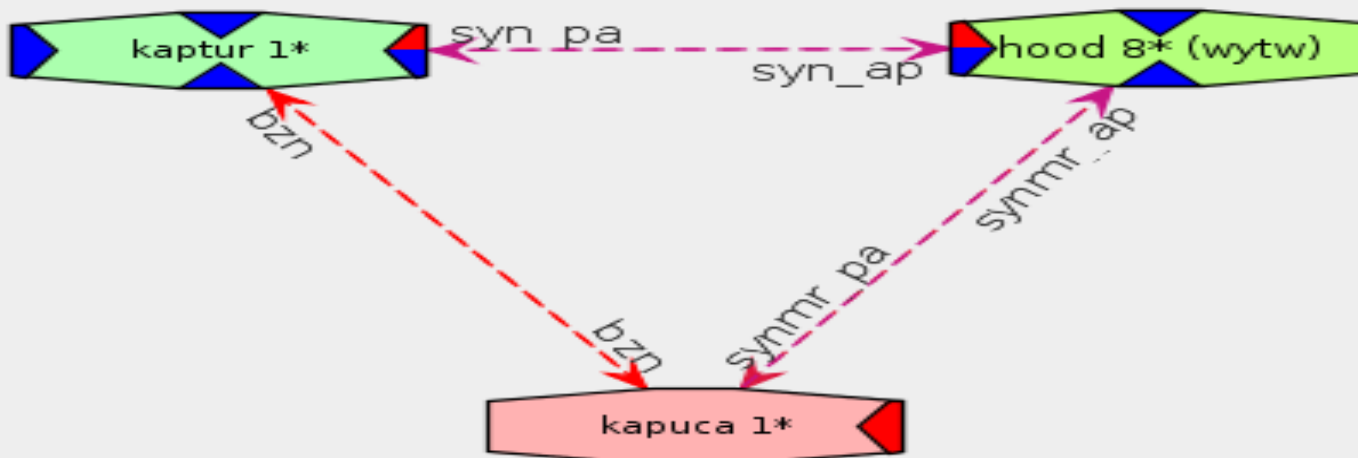


- **Zestaw relacji międzyjęzykowych inspirowany:**
  - relacjami międzyjęzykowymi z EuroWordNetu (Vossen 2002)
  - relacjami wewnątrzjęzykowymi ze Słownosieci (Maziarz et al. 2012)
- Synonimia (ekwiwalencja)
- Hiponimia i hiperonimia
- Meronimia i holonimia
- Synonimia częściowa
- Synonimia międzyrejestrów
- Synonimia międzyparadygmatyczna

# Synonimia



- Synonimia - **oznacza** dużą odpowiedniość znaczeń i struktur; **traktowana jako** podstawowa i priorytetowa **relacja**; **może być tylko jedna dla jednego synsetu**
- Synonimia częściowa - **stosowana w przypadku** częściowej odpowiedniości znaczeń i/lub struktur
- Synonimia międzyrejestrowa - **dla odpowiedników znaczeniowych różniących się** rejestrami stylistycznymi

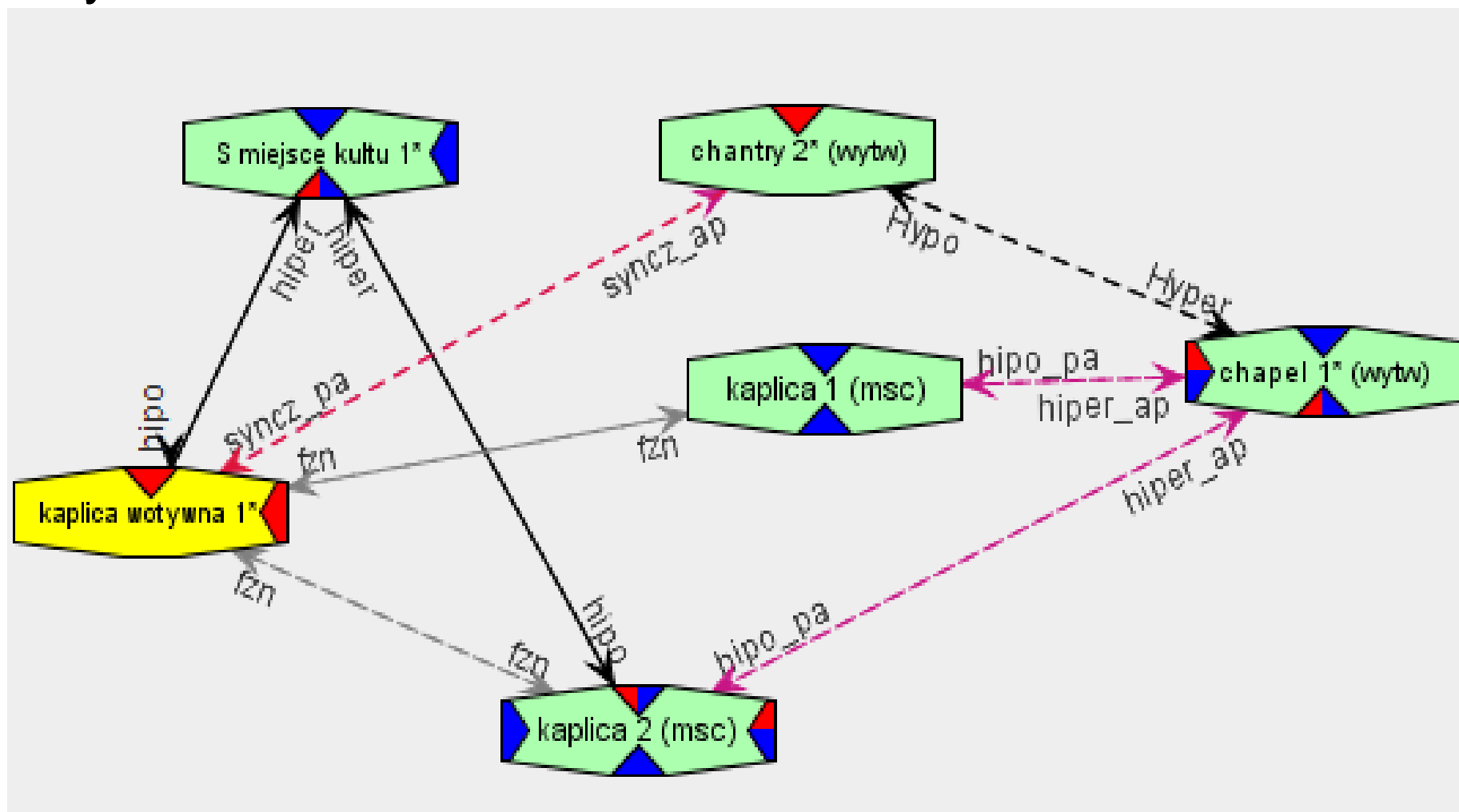




# Synonimia częściowa



## Przykład



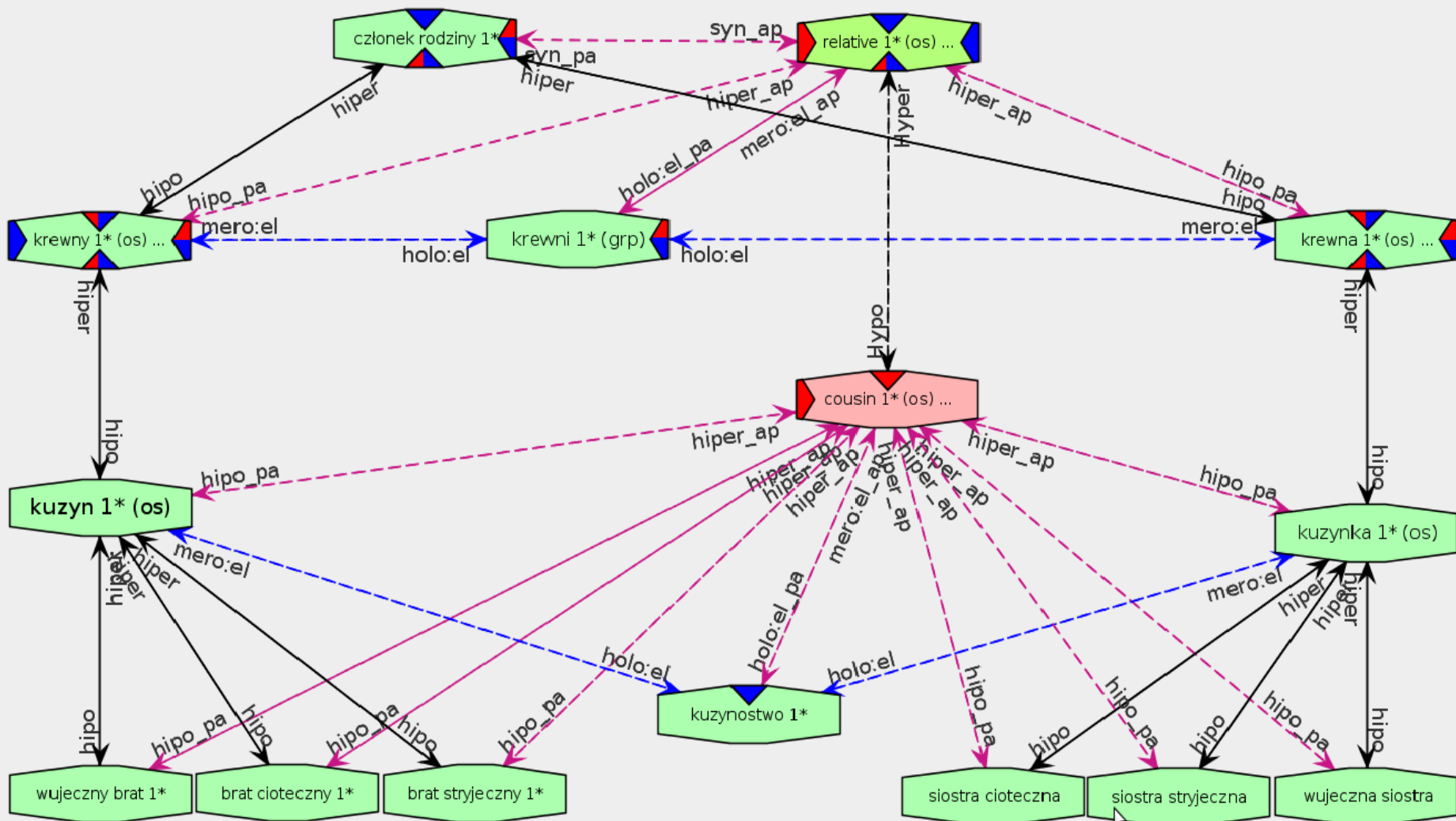
# Relacje międzyjęzykowe

---



- Hiponimia i hiperonimia zdefiniowane są w ramach zawierania się denotacji zbiorów: hiponim odnosi się do obiektu, który zawiera się w zbiorze denotatów hiperonimu
- Meronimia międzyjęzykowa – oznacza część większej całości
- Holonimia międzyjęzykowa – oznacza całość składającą się z mniejszych części

# Hyponimia i meronimia





# Synonimia międzyparadygmatica (3)

<x#> jest przymiotnikiem w języku polskim, <y#> jest rzeczownikiem w języku angielskim,

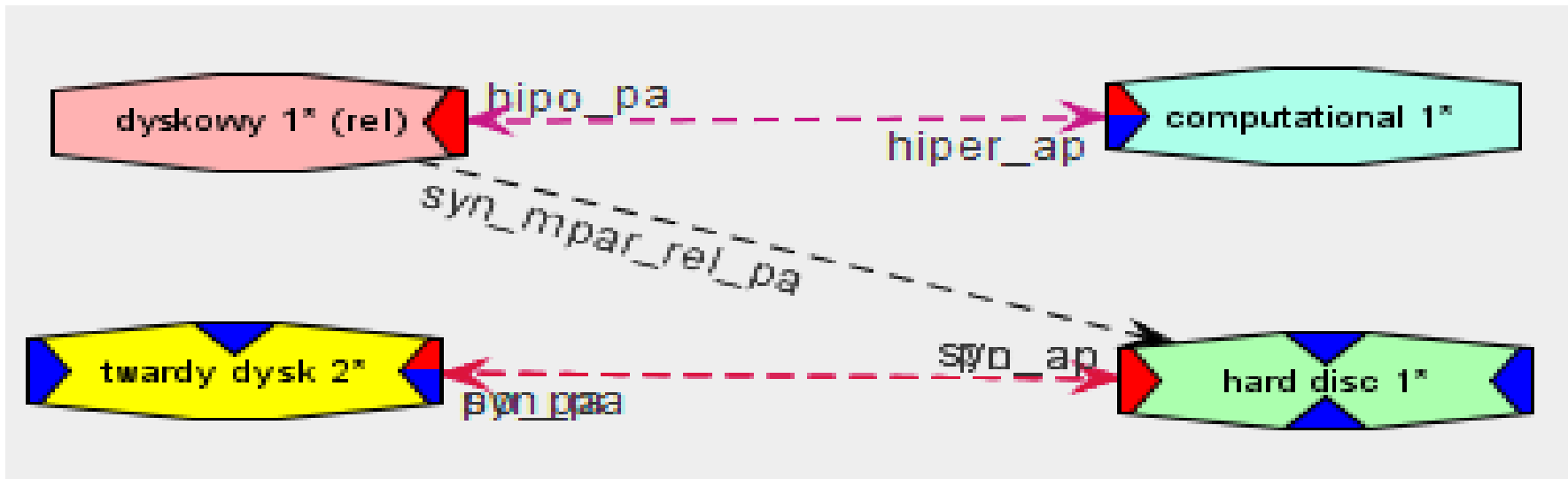
N jest rzeczownikiem w języku polskim

<x#> i <y#> łączy relacja międzyjęzykowej synonimii międzyparadygmaticznej typu "odnoszący się do" jeśli:

<x#> określa właściwości <y#>, które nie są jego cechami fizycznymi lub nie określają materiału, z którego jest wykonany

<x#> łączy relacja synonimii międzyparadygmaticznej lub derywacyjności z N

N i <y#> łączy relacja międzyjęzykowa



# Procedura rzutowania (1)

---



- Rozpoznanie znaczenia synsetu źródłowego:
  - miejsce w strukturze sieci - wszystkie istniejące relacje z naciskiem na hiperonim(y) i hiponimy;
  - definicje, komentarze;
  - porównanie pozostałych synsetów zawierających dany lemat)
- Przykład:  
{zagranica 1, obczyzna 1, obce terytorium 1}:
  - jest hiponimem {obszar 1, terytorium 1, obręb 1, strefa 1, zona 1, rejon 3} komentarz: 'ograniczona część przestrzeni, zwykle dużych rozmiarów, określona powierzchnia czegoś (np. obszar państwa, obszary niezalesione)
  - jest meronimem {świat 3, nieznanne 1} komentarz: 'iść w świat'
  - jest w relacji fazynimii z {granica państwa 1}

## Procedura rzutowania (2)

---



- Poszukiwanie synsetu docelowego:  
wytypowanie kandydatów na synset docelowy  
przy wykorzystaniu intuicji, podpowiedzi automatycznych  
oraz słowników  
  
{foreign country 1} 'any state of which one is not a citizen' –  
'każde państwo którego nie jest się obywatelem'  
  
jest hiponimem {state 1, nation 1, country 1, land 9,  
commonwealth 2, res publica 1, body politic 1} -  
'a politically organized body of people under a single  
government' – polityczna organizacja grupy ludzi skupiona  
pod jednym rządem'

## Procedura rzutowania (3)

---



- **weryfikacja kandydatów na synset docelowy:**  
porównanie struktur hipero i hiponimicznych (i innych jeśli istnieją) z synsetem źródłowym
- sprawdzenie istniejących i/lub potencjalnych relacji międzyjęzykowych;
- definicje, komentarze; słowniki  
np. {state 1, ..} jest międzyjęzykowym hiponimem {państwo 1, kraj 1} - 'zorganizowana politycznie społeczność, zamieszkująca określone terytorium, z niepodległą formą rządów'



# Procedura rzutowania (3)

---



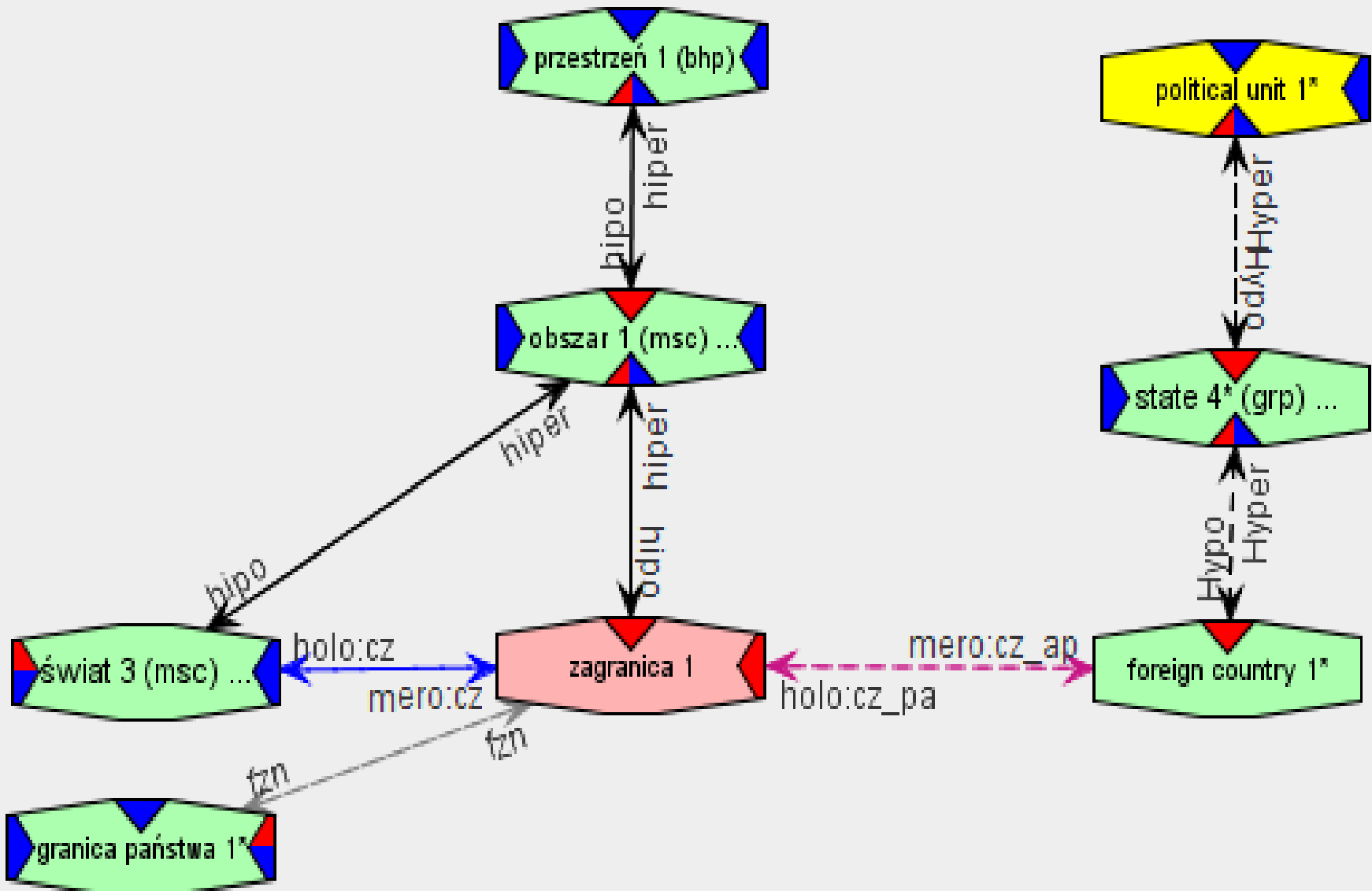
- Wytypowanie synsetu docelowego oraz relacji międzyjęzykowej:  
{foreign country 1}

Synonimia – nie (inne znaczenie, struktury i relacje)

Hiponimia – nie (znaczenie, struktury i relacje nie kwalifikują jako rodzaj)

Meronimia – tak (znaczenie, struktury i relacje kwalifikują jako część)

Połączenie synsetu źródłowego z synsetem docelowym



# Relacje międzyjęzykowe

---



Relacje międzyjęzykowe (pomiędzy synsetami): 148 387

	Rzeczowniki	Przymiotniki
● Synonimia	32 521	3 491
● Hiponimia	66 008	15 459
● Hiperonimia	4 068	51
● Meronimia	6 395	
● Holonimia	2 582	
● Synonimia częściowa	3 383	1 278
● Synonimia międzyrejestrowa	1 615	45
● Synonimia międzyparadygmatyczna		11 486
● Suma	116 577	31 810

# Jeszcze niezrzurowane

---



- Niezrzurowane synsety Słowosieci:
  - Rzeczowniki: 12 483
  - Przymiotniki: 16 886
  - Czasowniki: 21 596
  - Total: 50 965
  
- Niezrzurowane synsety Princeton WordNet:
  - Rzeczowniki: 44 380
  - Przymiotniki: 13 007
  - Przysłówki 3 619
  - Czasowniki: 13 789
  - Total: 74 795

# Praca tłumacza - sytuacje

---



- **Sytuacja 1 – tłumacz styka się z nowym dla siebie słowem w języku źródłowym (np. należącym do słownictwa specjalistycznego)**
- **Sytuacja 2 – tłumacz zna znaczenie słowa w języka źródłowym, ale nie zna jego odpowiednika w języku docelowym**
- **Sytuacja 3 – tłumacz ma kandydatów na odpowiednik w języku docelowym, ale nie potrafi wybrać najlepszego**
- **Sytuacja 4 - tłumacz zna słowo z języka źródłowego i jego znaczenie, ale nie spotkał się z**

# Przykład 1 (1)

---



- “We have services for all of them. Or maybe you don't believe in any God. In that case you go to the library during chapel. But we don't have any Buddhist.”

*Eagle Station*, Autor: Mark Berent

[https://books.google.pl/books?id=\\_Q-4BQEp3mMC&pg=PT3&lpg=PT3&dq=this+is+%22chapel+for+formation%22&source=bl&ots=KTJBbLI-Mc&sig=PdCWg77aFtIZ83Hx0YNK0U1tFx0&hl=pl&sa=X&ei=kwlZVb7IIYHT7Qa29oCwCQ&ved=0CDUQ6AEwAw#v=onepage&q=this%20is%20%22chapel%20formation%22&f=false](https://books.google.pl/books?id=_Q-4BQEp3mMC&pg=PT3&lpg=PT3&dq=this+is+%22chapel+for+formation%22&source=bl&ots=KTJBbLI-Mc&sig=PdCWg77aFtIZ83Hx0YNK0U1tFx0&hl=pl&sa=X&ei=kwlZVb7IIYHT7Qa29oCwCQ&ved=0CDUQ6AEwAw#v=onepage&q=this%20is%20%22chapel%20formation%22&f=false)

- Jednostki leksykalne Princeton WordNet:

chapel 1 (wytw)

chapel 2 (czy)

# Przykład 1 (2)

---



- Właściwy synset PWN:

{**chapel service 1, chapel 2**} (czy)

'a service conducted in a place of worship that has its own altar;

he was late for chapel'

- hiperonimia do  
{**service 3, religious service 1, divine service 1**}
- Synonimia międzyjęzykowa do {**nabożeństwo 1**}

## Przykład 2 (1)

---



- Stanowiły one bowiem, iż w takim wypadku zleceniodawca uprawniony będzie do obciążenia przewoźnika karą umowną w wysokości od 10 do 100 % **frachtu.**”

<http://www.transportoweprawo.pl/2013/kara-umowna-musi-byc-precyzyjnie-okreslona>

- **Jednostki SłowoSieci:**

fracht 1 (pos)

fracht 2 (pos)

fracht 3 (zvw)

- **{fracht 1} (pos): hiperonimia do {ładunek 3};**

Międzyjęzykowa synonimia częściowa do:

{cargo 1, lading 1, freight 1, load 3, loading 4, payload 2, shipment 1, consignment 1}



## Przykład 2 (2)

---



- Demand for container ships has pushed up **freight rates**.

<http://www.bbc.com/news/business-10999044>

'Popyt na kontenerowce zwiększył **stawki opłat frachtowych**.'

- Stanowiły one bowiem, iż w takim wypadku zleceniodawca uprawniony będzie do obciążenia przewoźnika karą umowną w wysokości od 10 do 100% **frachtu**."

'They held that in such case the customer would be entitled to charge the carrier with the contractual fine ranging from 10 to 100% of the **freight**

# Przykład 2 (3)

---



- {fracht 3} (zvw): hiperonimia do {żegluga 1};

Synonimia międzyjęzykowa do {freight 2, freightage 2}

hiperonimia do {transportation 5}

- “Jeśli polski przewoźnik załaduje fracht w Berlinie a miejscem przeznaczenia jest Poznań, to nie mamy do czynienia z kabotażem tylko znów z przewozem międzynarodowym.” <http://famos.net.pl/kabotaz-w-ue/>

'If a Polish carrier moves freight from Berlin and the destination is Poznań, we have to do not with cabotage, but again with international freight.'

- “Rzetelnie i kompleksowo obsługujemy wszelkie wyzwania jakie stawia przed nami fracht morski.”

[http://www.frachtfwo.pl/p=zakres\\_uslug&pm=fracht\\_morski](http://www.frachtfwo.pl/p=zakres_uslug&pm=fracht_morski)

# Przykład 3 (1)



## 'fafle'

(...) w zamyśleniu, z opuszczoną czarną głową przypominał psa myśliwskiego z obwisłymi **faflami**.” Borelowski Jan Maria Gisges, p. 70 [https://books.google.pl/books?id=j-pIAAAAIAAJ&q=fafle+psa+my%C5%9Bliwskiego&dq=fafle+psa+my%C5%9Bliwskiego&hl=en&sa=X&ei=B1AqVf-CCeXcywOIIIbY&redir\\_esc=y](https://books.google.pl/books?id=j-pIAAAAIAAJ&q=fafle+psa+my%C5%9Bliwskiego&dq=fafle+psa+my%C5%9Bliwskiego&hl=en&sa=X&ei=B1AqVf-CCeXcywOIIIbY&redir_esc=y)

Jednostki Słowosieci:

fafle 1(czc)

Synsety Słowosieci:

{fafle 1} (czc): hiperonimia do {warga 1}

(czc),

meronimia część do {pies myśliwski 1}

## Przykład 3 (2)

---



- 'His jowls hung in flaps like the **flews** of a hound, and he held the long gun behind him, run down back of his leg."  
*Cold Mountain* (przykład z wordnik.com)
- “Its eyes are one of the Bloodhound's most well-known characteristics. They are deeply sunk in the orbits, and the lids have a lozenge or diamond shape. This is because the lower lids are dragged down and turned outward by **the heavy flews**. This can cause eye problems if the flews are too heavy, so a good breeder goes for moderation rather than exaggeration.”

*Bloodhounds: Everything about Purchase. Care. Nutrition.*

# Przykład 4 (1)

---



- 'grzebień'

“Jednym z przyjaciół Frediego był kogut. Był to odważny, dumny ptak o imieniu Klemens. Miał on zawsze pierś dumnie wypiętą do przodu, dziób w górze i sterczący **grzebień.**”

*Magiczna bajaderka: Antologia bajek p. 184*

[https://books.google.pl/books?id=cDdiBgAAQBAJ&dq=magiczna+bajaderka&source=gbs\\_navlinks\\_s](https://books.google.pl/books?id=cDdiBgAAQBAJ&dq=magiczna+bajaderka&source=gbs_navlinks_s)

- Jednostki Słowosieci:

grzebień 1 (wytw)

grzebień 2 (**czc**)

grzebień 3 (**czc**)

## Przykład 4 (2)

---



- Synsety Słowosieci:

{**grzebień 2**} (czc)

'wrostek na głowie niektórych ptaków i gadów (u gadów może to być szereg wyrostków nie tylko na głowie, ale też wzdłuż grzbietu i ogona)'

hiperonimia do {wrostek 1} (czc),

{**grzebień 3**} (czc) hiperonimia do {sierść

1}

{**grzebień 6**} (czc)

meronimia części do {mostek 3}

## Przykład 4 (3)

---



- {**grzebień 2**} (czc)

Synonimia międzyjęzykowa do {**crest 5**}

'a showy growth of e.g. feathers or skin on the head of a bird or other animal'

hiperonimia do {process 5} (czc),

hiponimia do {**comb 2**, **cockscornb 4**, **coxcomb 3**} (zw),  
{topknot 2} (zw), {tuft 2} (zw)

“Now the rooster is an entirely different sort of bird from the hen. It is very proud and has a red **crest** on the top of his head.”

*The Benchley Roundup: A Selection by Nathaniel Benchley of His Favourite Robert C. Benchley, Nathaniel Benchley*

# Przykład 5 (1)

---



- “W stanach przewlekłych stosować można „wcierki” rozgrzewające, lub **blistry**.”

[http://starywww.up.wroc.pl/polish/struktura/wet/kkc/cwiczenia/ortopedia\\_kop.html](http://starywww.up.wroc.pl/polish/struktura/wet/kkc/cwiczenia/ortopedia_kop.html)

- Jednostki Słownosieci:

blistr 1 (sbst)

blistr 1 (wytw)

blistr 2 (il)

- Synsety Słownosieci:

{blistr 1} (wytw) hiperonimia do  
{opakowanie 2}

{blistr 2} (il) hiperonimia do



## Przykład 5 (2)

---



- “Harltons Liquid **Blister** is a counter irritant with penetrating action for use in the blistering of horses legs.”

<http://www.allvetsupply.com/halibl12.html>

“If a paint or **blister** has been used, the leg should not be bandaged as the chemicals under the bandages may severely irritate the horse’s skin even more.”

[https://books.google.pl/books?id=i\\_iN\\_eyz0NMC&pg=PA223&lpg=PA223&dq=apply+blister+horses+legs&source=bl&ots=vlf-LndrwT&sig=4B8w-f9P\\_IAAQQjZVP3gtJ6BMx4&hl=en&sa=X&ei=38MqVdPAKoPMygPXvIHoAg&ved=0CF4Q6AEwCA#v=onepage&q=apply%20blister%20horses%20legs&f=false](https://books.google.pl/books?id=i_iN_eyz0NMC&pg=PA223&lpg=PA223&dq=apply+blister+horses+legs&source=bl&ots=vlf-LndrwT&sig=4B8w-f9P_IAAQQjZVP3gtJ6BMx4&hl=en&sa=X&ei=38MqVdPAKoPMygPXvIHoAg&ved=0CF4Q6AEwCA#v=onepage&q=apply%20blister%20horses%20legs&f=false)

# Bibliografia

---



- **Fellbaum, Ch. (ed). (1998). *WordNet: An Electronic Lexical Database*. MIT Press: Cambridge, Massachusetts.**
- **Kędzia, P., Piasecki, M., Rudnicka, E., Przybycień, K. (2013). Automatic Prompt System in the Process of Mapping plWordNet on Princeton WordNet. *Cognitive Studies* 13: 123-141.**
- **Maziarz, M., Piasecki, M. and S. Szpakowicz. (2012). 'Approaching plWordNet 2.0.' *Proceedings of the 6th Global Wordnet Conference*, Matsue. pp. 189-196. accepted for publication.**
- **Piasecki, M., Szpakowicz, S. and B. Broda. (2009). *A WordNet from the Ground Up*. Oficyna Wydawnicza Politechniki Wrocławskiej: Wrocław.**
- **Princeton WordNet <http://wordnet.princeton.edu/wordnet/>**
- **Rudnicka, E., Maziarz, M., Piasecki, M., & Szpakowicz, S. (2012) 'A Strategy of Mapping Polish WordNet onto Princeton WordNet'. In *Proceedings of COLING 2012*. ACL.**
- **SłowoSieć <http://plwordnet.pwr.wroc.pl/wordnet/>**
- **Vossen, P. (ed). (2002). *EuroWordNet General Document***

Dziękuję bardzo za uwagę

---