

25.11.2015

**Recenzja rozprawy doktorskiej mgr. Agnieszki Płuwak**  
***Ramy semantyczne w tekstach prawnych – analiza konfrontatywna***  
***polsko-angielsko-niemiecka***

Praca stanowi teoretyczne studium poświęcone kontrastywnej analizie terminologii prawnej, którą stosuje się w polsko- angielsko- i niemieckojęzycznych umowach najmu. Rozprawa ma charakter interdyscyplinary w tym sensie, że w obszarze językoznawstwa należy do zakresu semantyki kontrastywnej, a jednocześnie stosuje metody informatyczne z dziedziny przetwarzania języka naturalnego (NLP). Praca wpisuje się w nurt semantycznych badań kognitywnych i stawia sobie za cel zastosowanie podejścia semantyki ramowej FrameNet do badania kontrastywnego, w którym ramy semantyczne są wykorzystane jako *tertium comparationis*. Autorka jest najprawdopodobniej pierwszą osobą, która z kontrastywnego punktu widzenia analizuje słownictwo prawnicze w ramach teorii FrameNetu, i która przewiduje implementację wyników badania kontrastywnego stosownie do zapotrzebowań branży zarządzania nieruchomości w celu praktycznego wykorzystania do automatyzacji analizy umów najmu. W rozdziale wstępnym Autorka przedstawia cztery podstawowe cele rozprawy:

- i) wkład w rozbudowę polskiego FrameNetu,
- ii) zastosowanie metody semantyki ramowej do badania stylu urzędowego (tj. w obrębie terminologii prawniczej),
- iii) przygotowanie materiału językoznawczego dla potencjalnego programu do automatycznego wyszukiwania informacji w tekstach specjalistycznych objętych tematem rozprawy,

iv) przyczynę do rozwoju metody kontrastywnej w ujęciu semantyki ramowej (oryginalnie FrameNet został stworzony dla języka angielskiego i na razie nie jest jeszcze prawdziwym językiem pośrednikiem).

Odpowiednio Autorka formułuje tezę główną o możliwości wpisania leksemów z różnych języków w ramy semantyczne zdefiniowane w (przede wszystkim angielskim) FrameNecie lub zbudowania nowych ram i zdefiniowania reguł ekstrakcji danych dla systemu automatycznej analizy w trzech językach będących przedmiotem rozprawy. Z tego wynika, że rozprawa ma zasadniczo orientację teoretyczną a nie empiryczną.

Składa się ona z ośmiu rozdziałów, podsumowania, wykazu literatury i pięciu załączników. W załącznikach zamieszczono anotację semantyczną, nieanotowane przykłady, symbole zastosowane w anotacji, prymitywy pojęciowe i elementy pozatematyczne.

W rozdziałach pierwszym i drugim przedstawiono historię rozwoju projektu FrameNet od początków sięgających amerykańskiego dystrybucjalizmu aż po aktualne etapy tworzenia zasobów leksykograficznych dla różnych języków. Rozdział zawiera bardzo obszerny przegląd literatury i wskazuje na dobrą znajomość zarówno najważniejszych klasyków teorii ram semantycznych, jak i jej prekursorów.

Rozdział trzeci poświęcono aktualnemu stanowi rozbudowy zasobów FrameNetu w celu przetwarzania języka naturalnego. Autorka zwraca uwagę na fakt, że projekty leksykograficzne typu FrameNet, w odróżnieniu od innych podejść semantycznych, dzięki podobieństwu ich struktur semantycznych do ontologii wypracowanych w sztucznej inteligencji, można stosować w NLP, dokładniej w funkcjach automatycznej ekstrakcji danych z tekstów, zadawania pytań do treści i streszczenia tekstów. Obie reprezentacje wiedzy opierają się na tych samych relacjach taksonomicznych między pojęciami (w przypadku ontologii informatycznych) i leksemami (w przypadku FrameNetu), takich jak hiponimia i meronimia. Autorka przekonująco pokazuje, że FrameNet który bazuje przeważnie na podejściu oddolnym, tj. wychodząc od leksemu (*bottom-up*), zajmuje centralne miejsce wśród ontologii informatycznych zbudowanych od ogółu do szczegółu (*top-down*). Zgadzam się z twierdzeniem, że projekt FrameNet, w rzeczy samej, oferuje możliwość zastawania do budowy ontologii dziedziczonej, co jest jednym z celów recenzowanej pracy.

W rozdziale czwartym dość obszernie przedstawiono procedurę opisu ram semantycznych i zasady anotacji używanej w projekcie FrameNet. Opisano także ogólne zasady budowy ram i system anotacji dla poszczególnych części mowy, głównie według systemu Ruppenhofera, który Autorka stosuje dla języka polskiego. W związku z tym, pragnę zwrócić uwagę na fakt, że wybrany model niekonsekwentnie rozróżnia między elementami ramowymi na poziomie konceptualnym i ich realizacjami syntaktycznymi na poziomie zdania. Należy podkreślić, iż spis przypadków, w których elementy ramowe nie uzyskują eksplicytniej realizacji formalnej w zdaniu w rozdziale 4.2.4 nie jest pełen, brakuje

typów CNI jak pro, PRO i inne ukryte subjekty. Rodzi to pytanie, w jakim stopniu analiza różnych typów elips i elementów zerowych potrzebna jest do analizy terminologii prawniczej. Chodzi przecież o czysto syntaktyczne struktury, niezależne od ram semantycznych. To samo dotyczy relacji gramatycznych subiektu i obiektu.

Następny rozdział zawiera przegląd osiągnięć FrameNetów powstających w różnych językach, z naciskiem na zagadnienie zastawiania ram semantycznych wypracowanych na podstawie języka angielskiego jako *tertium comparationis*. Przegląd ten uzupełniono opisem polskich prac w kontekście semantyki ramowej. Nie dziwi wniosek Autorki, że większość różnic wywnioskowanych na podstawie przedstawionych badań kontrastywnych dotyczy właśnie odmienniej struktury morfologicznej i składniowej. Nasuwa się pytanie, czy formalne struktury jak ww. konstrukcje (implicitne realizacje elementów ramowych, rola syntaktycznego podmiotu itp.) można w ogóle zaliczyć do *tertium comparationis*, tj. do języka pośrednika, czy nie należałoby ich wykluczyć z analizy.

Części zasadnicze rozprawy to rozdziały 7 i 8 poświęcone analizie empirycznej. W rozdziale 7 przedstawiono procedurę zastosowaną w badaniu, tj. założenia metodologiczne. Stosując zasady opisane w poprzednich rozdziałach, Autorka przewiduje trzy kroki: i) podejście oddolne, ii) podejście odgórne i iii) analiza konfrontatywna. O doborze tekstów Autorka pisze tylko, że zanalizowała „po 30 tekstów w języku polskim, angielskim i niemieckim, napisanych przez prawników według obowiązujących standardów prawa” i że „dla wersji angielskiej zebrano umowy zawarte głównie na terenie Stanów Zjednoczonych oraz Wielkiej Brytanii” (str. 203, 205). Uderza mnie brak potrzebnych informacji o danych empirycznych: i) liczba słów, ii) źródła, iii) informacja o tym, w jakim stopniu umowy są ustandaryzowane, iv) pochodzenie i autor tekstu (wiadomo, że w Stanach Zjednoczonych i Wielkiej Brytanii stosuje się różną terminologię). To samo dotyczy danych korpusowych, które Autorka wspomina mimochodem, natomiast nigdy ich nie opisuje. Załącznik nr. 1 zawiera anotowane przykłady z Narodowego Korpusu Języka Polskiego, ale nie podano w nim konkretnego zapytania, a w załączniku nr. 2 znajdują się nieanotowane przykłady dla zapytania „Umowa + zawarta + pomiędzy”. Dziwi, że te przykłady nie są bezpośrednio związane z obiektem badania (tj. z umowami NAJMU). Taki brak dokumentacji i niedokładność prezentacji danych empirycznych uważam za jeden z najważniejszych niedostatków pracy. Drugie zastrzeżenie dotyczące części empirycznej wiąże się z niedostatecznym opracowaniem aspektu języka dziedzinowego. Rozumiem, że intencją Autorki była przede wszystkim rozbudowa modelu FrameNet, lecz mimo to wydaje mi się, że skoro praca ma na celu opis nie leksyki ogólnej, lecz zasobu leksykalnego w dziedzinie komunikacji prawnej, Autorka powinna przedstawić najważniejsze cechy tego stylu funkcjonalnego i objaśnić jak ustaliła „powtarzające się i charakterystyczne dla języka umów prawnych jednostki leksykalne” (str. 205). Niestety Autorka poświęca aspektowi prawnicznemu jedynie krótki akapit na stronach 220-221, w którym przedstawia definicję umowy w dziedzinie prawa.

Sama analiza semantyczna leksemów w badanych trzech językach została przeprowadzana bardzo rzetelnie. Znając niemieckie umowy najmu, pytam się, dlaczego Autorka nie zajmuje się elementami dowolnymi, którymi umowy w poszczególnych krajach mogą się różnić (na przykład *koszty dodatkowe* – *Nebenkosten*). Wyniki części empirycznej wnoszą nowe spojrzenie na semantykę ramową z punktu widzenia przydatności do analiz kontrastywnych. Z wymienionych na str. 358-363 wniosków dla teorii semantyki ramowej chciałbym przytoczyć tylko następujące. Autorka zwraca uwagę, że po pierwsze ramy realizujące ten sam scenariusz w różnych językach mogą oddawać różne elementy lub etapy. Po drugie, styl zdaje się mieć znaczenie dla opisu ramy w tym sensie, że dopuszcza występowanie czasowników z inną strukturą walencyjną pomimo istnienia w tym języku dokładnych ekwiwalentów (kwestia uzusu). Ciekawym wydaje mi się także wniosek, że nie wykryto różnic między językami na poziomie pojęciowym i wyboru elementów ramy jako jądrowych. Można zatem stwierdzić, że postawione w pracy cele badawcze (i-iv, zob. wyżej) zostały w pełni osiągnięte.

Zauważono jednak kilka usterek terminologicznych; na przykład na str. 113 wspomina się „inne języki anglosaskie” (o jakie języki chodzi?), na str. 119 deixis i elementy ewaluatywne podsumowano pod „różnicami socjolingwistycznymi” (chodzi jednak o różnice czysto semantyczne); S. 130: w zdaniu *Cook for 20 minutes* mamy do czynienia nie tylko z brakiem obiektu wyrażanego dopełnieniem lecz i subiektu; podobnie na str. 130: w stronie biernej nie brakuje agenta, lecz raczej brakuje agenta wyrażonego powierzchniowo.

Pozytywnie oceniam natomiast warstwę formalną pracy doktorskiej, jak na przykład styl – o ile tylko mnie, nierodzimemu użytkownikowi polszczyzny, wolno to oceniać – czy staranność w zakresie redakcji: układ tekstu, cytaty i dane bibliograficzne. Tekst zawiera stosunkowo mało literówek.

Podsumowując, można powiedzieć, że recenzowana tu rozprawa doktorska daje wszechstronny obraz stanu badań w dziedzinie semantyki ramowej w ogóle, a w szczególności aktualnych prac w ujęciu FrameNetu. Moje zastrzeżenia dotyczą z jednej strony nierozróżnienia poziomów konceptualnych i czysto składniowych. Z drugiej strony, Autorka pominęła literaturę na temat stylu urzędowego (albo języków dziedzinowych); badając język umów najmu Autorka powinna była uwzględnić zarówno specyfikę terminologii, jak i przede wszystkim cechy języka prawnego. Niestety nie mogę nie zauważyć, że słabą stroną pracy jest jej wątkowa podbudowa empiryczna. Natomiast, na wysoką ocenę bez wątplenia zasługuje szeroki horyzont teoretyczny pracy. Choć zauważono kilka problemów w dokumentacji badanych tekstów i w opracowaniu części dotyczącej języka dziedzinowego, należy podkreślić, że rezultaty badania są relewantne dla przyszłych badań w ujęciu FrameNetu. Na koniec pragnę zauważyć, że wspomniane niedoskonałości nie wpływają na ostateczną ocenę pozytywną pracy.

Stwierdzam zatem, że przedstawiona rozprawa spełnia wymogi stawiane pracom doktorskim w zakresie językoznawstwa.

A handwritten signature in black ink, appearing to be 'B. W. Klen'.