

**Z5** Efektywny parser składniowo-semantyczny

**Z5.2** Rozbudowa kategorialnego parsera składniowo-semantycznego

**M6** Kwerenda biblioteczna, utworzenie bazy jednostek leksykalnych i wstępnej listy zjawisk semantycznych, wstępna wersja koncepcji metaopisu semantycznego

Okres sprawozdawczy: październik – grudzień 2016

Partner: Instytut Podstaw Informatyki PAN, Warszawa

Typ: raport + zasób

31.12.2016, Warszawa

Autor: Wojciech Jaworski, Jakub Kozakoszczak

# 1. Wprowadzenie

W ramach kamienia milowego opracowano listę jednostek leksykalnych wprowadzających niestandardowe zjawiska semantyczne. Zjawiska te wymagają pogłębionych badań i opracowania lub adaptowania specjalnych rozwiązań po stronie metajęzyka reprezentacji. Na podstawie listy oraz kwerendy literaturowej opracowano listę niestandardowych zjawisk semantycznych przedstawioną w rozdziale 2.2. Etapem wieńczącym pierwsze dwa było opracowanie rozbudowanej innowacyjnej koncepcji metaopisu semantycznego przystosowanej do celów implementacyjnych w parserze kategorialnym.

## 2. Opis zasobu/programu

### 2.1. Baza jednostek leksykalnych

Znajduje się w załączonym do niniejszego dokumentu pliku `base_LU.tab`. Zawiera listę jednostek leksykalnych przeznaczonych do szczegółowej analizy. Na listę trafiły następujące jednostki leksykalne:

- spójniki, przyimki, kubliki i wykrzykniki z SGJP oraz NKJP1M;
- przysłówki nieodprzymiotnikowe z SGJP;
- przyimki złożone z pliku z realizacjami z Walentego;
- wybrane spójniki złożone z NKJP i SJP;
- rzeczowniki, przymiotniki i przysłówki wybrane z listy 5000 najczęstszych leksemów NKJP1M.

W ostatnim punkcie głównym wyznacznikiem dla przymiotników było nie posiadanie roli tematycznej Attribute, a dla przysłówków Manner. Oprócz tego kryterium było posiadanie niestandardowej dla danej części mowy cechy semantycznej. W przypadku rzeczowników kryteria były następujące:

- SYTUACJE
  - rzeczowniki mające wiele wspólnego znaczeniowo z czasownikami, w szczególności odnoszące do sytuacji / stanów rzeczy, wprowadzające typowe dla czasownika role tematyczne, choćby modyfikatory czasu i miejsca
- NOŚNIKI TREŚCI PROPOZYCJONALNEJ
  - nazwy komunikatów / bytów językowych / informacji przyjmujące podrzędniki „o”, „że”, „żeby”
  - nazwy zjawisk podobnych do komunikacji, bo mających propozycjonalną treść, np. „wątpliwość(, że...)”
- MIARA, ILOŚĆ, LICZBA
  - pojemniki i rzeczowniki miary, jak „wysokość”, „powierzchnia”, „długość”, „częstka”, rzeczowniki implikujące kolejność lub porządek („etap”, „poprzednik”)
  - nazwy miar (niekonsekwentnie, bo przypomniały mi się w trakcie), np. „moment”
- UKRYTE ZMIENNE
  - rzeczowniki okazjonalne, jak „dzisiaj”
  - rzeczowniki z ważnymi ukrytymi zmiennymi, które często będą niewyrażone na powierzchni, jak „konkurencja”, „reszta”, „rodak”
- METAJĘZYK
  - metatekstowe, jak „konieczność”, „możliwość”, „nieprawda”

- nazwy ważnych klas ontologicznych (typów obiektów) w naszej ontologii, jak „cecha”, „miejsce”, „rodzaj”, ponieważ mogą koreferować z innymi częściami mowy odnoszącymi do obiektów tego typu, por. \*, „Był zielony i ta cecha rzucała się w oczy.”\*

Usunięte zostały rzeczowniki o formie synkretycznej z gerundium, np. „włamanie”.

Czasowniki nie zostały umieszczone na liście z uwagi na to, że są analizowane w Walentym i ich semantyka będzie badana w oparciu o ten zasób.

## 2.2. Lista zjawisk przeznaczonych do uwzględnienia w reprezentacji semantycznej parsera ENIAM

### 2.2.1. Wstęp

Lista zbiera zjawiska semantyczne, których spójne opisanie w ramach jednolitej implementowalnej reprezentacji znaczenia o szerokim pokryciu dla języka polskiego wymaga szczegółowych, głębszych badań. Są to przede wszystkim zjawiska polegające na wprowadzaniu przez wyrażenia elementów znaczenia do zdań w sposób inny podstawowy, czyli przez zawężającą modyfikację znaczenia bezpośredniego nadrzędnika składniowego. Ten podstawowy sposób nazywamy roboczo *lokalnym wkładem znaczeniowym*. Lista jest wstępnym rozpoznaniem, które posłuży jako mapa zadań w następnych etapach rozwoju kategoryjnego parsera składniowo-semantycznego ENIAM.

### 2.2.2. Lokalność i nielokalność składniowo-semantyczna

**Opis zjawiska** Podzbiorem logiki pierwszego rzędu, który wystarcza do wyrażenia znakomitej większości wypowiedzi w języku naturalnym, jest logika egzystencjalno-koniunkcyjna (logika  $\exists \wedge$ ). Język reprezentacji znaczenia parsera ENIAM jest bardziej złożony, bo wzbogacony między innymi o operator DSCR wiążący zmienną z formułą oraz pozalogiczne predykaty relacyjne, jak TYPE i COUNT i role tematyczne, jak AGENT i THEME. Istnieją jednak jednostki języka naturalnego, wśród nich kwantyfikatory, operatory logiczne i modyfikatory nieintersektywne, których znaczenie leksykalne (w sensie wkładu w znaczenie zdania) jest niewyraźne w logice  $\exists \wedge$  z operatorem DSCR, np. dla języka polskiego

- każdy, prawie każdy, każdy z powyższych, co dziesiąty, każdy ... ma inny ..., nawzajem, razem, codziennie, nigdy, zawsze, tylko, jakiś, pewien, jedyny, oba, inny nie tylko ... lecz także

Jednostki te nazywać będziemy *jednostkami o nielokalnym wkładzie składniowo-semantycznym* lub po prostu *operatorami nielokalnymi*. Pozostałe jednostki mają *lokalny wkład składniowo-semantyczny*.

Interesują nas następujące cechy operatorów nielokalnych:

- konstrukcje składniowe, w których występują,
- semantyka, czyli wprowadzane zmienne i liczba argumentów,
- odwzorowanie konstrukcji składniowych na semantykę.

Nie będziemy opatrywać operatorów nielokalnych pełnymi hasłami leksykalnymi, a jedynie zaznaczać ich wystąpienia w formach logicznych w sposób taki, by potencjalne dodanie znaczenia leksykalnego nie wymagało dalszego zmieniania formy logicznej zdania.

Z uwagi na to, że operatory nielocalne wchodzą ze sobą w interakcje w zdaniu (np. mogą stać w różnej kolejności), zadane przez nas odwzorowanie konstrukcji składniowych na formę logiczną nie będzie stanowić przepisu na jednoznaczny semantykę, w szczególności nie będziemy na tym etapie rozwiązywać problemu niejednoznaczności zakresu operatorów nielokalnych.

Osobnym problemem jest reprezentacja nieciągłych operatorów nielokalnych, np. *Każda strona ma inny kolor* i interakcji między operatorami nielokalnymi innych niż zagnieżdżanie się zakresów np. w przypadku kwantyfikatorów rozgałęzionych w zdaniu Hintikki.

**Negację** traktujemy jako operator nielokalny mający jeden argument i nie wprowadzający zmiennej.

Spójniki „**jeśli ..., to ...**” oraz „**jeśli ..., ...**” użyte w znaczeniu logicznym uznajemy za operatory nielokalne.

### **Zjawisko wyciągania kwantyfikatorów poza zakres operatora spowodowane koreferencją elementów argumentów operatora**

Jeśli rolnik ma krowę, to ją doi (1)

W powyższym przykładzie „jeśli” traktujemy jako logiczną implikację, „rolnik” i „krowa” są kwantyfikowane egzystencjalnie, a „ją” i „pro” (podmiot domyślny czasownika „doi”) są kwantyfikowane koreferencyjnie ( $\iota$ ). Aby rozwiązać koreferencję trzeba przenieść powyższe 4 kwantyfikatory ponad implikację co powoduje zamianę kwantyfikatorów egzystencjalnych przy „rolniku” i „krowie” na uniwersalne.

Analogiczne zjawisko będzie występować przy wszystkich operatorach nielokalnych mających więcej niż jeden argument.

#### **2.2.3. Niejawne argumenty semantyczne**

**Opis zjawiska** Niezbędne jest rozpoznanie i oznaczenie argumentów niejawnych w leksemach i wypracowanie skutecznej koncepcji ich opisu i przetwarzania.

Za niejawne argumenty semantyczne uznajemy te argumenty, które są niezbędne do ustalenia znaczenia wyrazu mającego ten argument. Przykłady występowania wyrażeń z różnymi argumentami niejawnymi:

— większy – większy niż słoń

Znaczenia wyrazu „większy” nie da się obliczyć bez argumentu „niż słoń”.

— Biegnie. – On biegnie. – Jaś biegnie.

W pierwszym zdaniu inicjatora (INITIATOR) reprezentujemy oznaczając go niemym zaimkiem (PRO), w drugim jest on wyrażony za pomocą zaimka „on”, w trzecim jest wskazany jawnie jako „Jaś”. Aby obliczyć znaczenie PRO i „on” potrzebna jest wartość koreferencyjnego lub deiktycznego argumentu.

— Karol chce. – Karol chce pracować.

Brakującego argumentu składniowego czasownika „chcieć” nie traktujemy jako niejawnego argumentu semantycznego.

**Lista niejawnych argumentów semantycznych** Po wstępnej analizie przyjmujemy istnienie sześciu typów niejawnych argumentów semantycznych:

1. Okazjonalne

Np. *tu*.

2. Deiktyczne

Np. *taki*.

3. Porównawcze

Leksemy z tym argumentem mogą przyjmować porównawczy argument przyimkowy.

4. Porządkowe  
Np. *kolejny*.
5. Koreferencyjne  
Np. *on*.
6. Relacyjne  
Aby określić znaczenie znaczenie przymiotników i liczebników z argumentem relacyjnym (*duży, sporo*) trzeba wiedzieć, jaki predykat jest modyfikowany przez dane słowo.

#### 2.2.4. Znaczenie relacji składniowych

**Uogólniona intersektywność** Ważną cechą semantyczną, która przysługuje nie wszystkim leksemom, a pozwala na reprezentowanie ich wkładu znaczeniowego jako zawężającego znaczenie szerszej frazy, jest *intersektywność*. Tradycyjnie rozumiana intersektywność przypisywana jest przymiotnikom i oznacza, że denotacja frazy rzeczownikowej z przymiotnikiem intersektywnym jest podzbiorem denotacji frazy bez niego.

Widzę czarnego kota.  $\rightarrow$  Widzę kota. (2)

Rozumienie to można uogólnić na przysłówki i mówić o przysłówkach intersektywnych. Zbadania wymaga, w jakich innych częściach mowy lub pozycjach składniowych da się mówić o podziale na wyrażenia nieintersektywne i intersektywne w sensie uogólnionym – wyrażenia takie powinny spełniać większość lub wszystkie testy:

*Test 1.* Zaczynając od korzenia drzewa rozbioru składniowego dodajemy kolejne węzły i sprawdzamy, czy ze znaczenia drzewa rozszerzonego wynika znaczenie drzewa mniejszego. Jeśli tak, dodawane węzły są intersektywne w sensie uogólnionym.

*Test 2.* Zaczynamy od liści i ucinamy je po jednym, sprawdzając, czy znaczenie zdania staje się bardziej ogólne. Jeśli tak, ucinane węzły są intersektywne w sensie uogólnionym.

*Test 3.* Porównujemy typy semantyczne wyrażenia złożonego i tego samego wyrażenia bez jednego z podrzędników. Jeżeli pierwszy z tych typów jest podtypem drugiego, podrzędnik jest intersektywny w sensie uogólnionym.

Jaś biegnie.  $\rightarrow$  Biegnie. (3)

cztery jabłka  $\subset$  jabłka (4)

Mieszkam nieopodal fontanny.  $\overset{?}{\rightarrow}$  Mieszkam nieopodal. (5)

Przykłady 3 i 5 ilustrują wynikanie zdań, przykład 4 mówi o zawieraniu się denotacji. Ostatni przykład pokazuje, że rozstrzygnięcia wymaga problem relacji znaczenia jednobrzmiących leksemów przysłówkowych i przyimkowych. Pożądaną opcją wymagającą zweryfikowania jest identyczność znaczeniowa wszystkich lub większości takich par *modulo* argument niejawny przysłówka.

**Rozbiór składniowy w podejściu semantycznym** Do wyliczenia reprezentacji semantycznej zdania potrzebne są dane wejściowe – rozbiór składniowy w podejściu semantycznym. Otrzymanie takiego rozbioru wymaga przekształcenia bardziej niskopoziomowego rozbioru w podejściu morfoskładniowym z użyciem informacji o znaczeniu wzajemnych rzędników. Przekształcenie jest potrzebne, gdy bezpośredni podrzędnik węzła  $n$  nie spełnia jego ograniczeń

selekcyjnych, ale podrzędnik następnego rzędu (wnuczek  $n$ ) je spełnia. Spełnianie ograniczeń selekcyjnych przez leksem polega na tym, że jego typ semantyczny jest podtypem jednego z typów wskazanych w ograniczeniach selekcyjnych. W takim przypadku podrzędnik pierwszego i drugiego rzędu są zamieniane miejscami w drzewie rozbioru.

Typowym przykładem zadania, w którym opisana wyżej procedura jest nietrywialna, jest rozpoznanie, czy wystąpienie przyminka przewidzianego w ramie składniowej nadrzędnego czasownika jest realizacją tej ramy, jak „w głowie”, czy jest to przymimek semantyczny, jak w wyrażeniu „w samolocie” w przykładzie:

**W samolocie** wielu osobom kręci się **w głowie**. (6)

W rozwijanej dla parsera ENIAM odmianie podejścia semantycznego o tym, czy bezpośrednim podrzędnikiem czasownika jest przymimek czy rzeczownik, decyduje ostatecznie to, czy preferencje selekcyjne argumentu ramy uzgadniają się z typem semantycznym rzeczownika czy frazy przyminkowej. Podobny problem dotyczy liczebników i rzeczowników w użyciu pojemnikowym (patrz sekcja 2.2.10)

**Zależność roli tematycznej od znaczenia rzędników** Podrzędniki dzielimy na argumenty i modyfikatory rozumiane semantycznie. Rola tematyczna podrzędnika zależy za każdym razem od jego typu zależności składniowej, a ponadto

- w przypadku argumentu – od znaczenia nadrzędnika. Przykładowo, odniesienie podmiotu czasownika *czuć* to EXPERIENCER zdarzenia, czasownika *myśleć* – to AGENT zdarzenia, a jego dopełnienia – to THEME zdarzenia.
- w przypadku modyfikatora – od jego znaczenia (podrzędnika). Przykładem jest modyfikator rzeczownikowy w narzędniku. Taki modyfikator może wprowadzać role INSTRUMENT, TIME lub MANNER w zależności od jego znaczenia:

Naprawiła szafę **młotkiem**. (7)

Naprawiła szafę **wieczorem**. (8)

Naprawiła szafę **własnym sposobem**. (9)

Oddzielnej uwagi wymagają też rzeczownikowe frazy wolne w celowniku, przymiotniki, które mają rolę tematyczną inną niż ATTRIBUTE i przysłówki, które mają rolę tematyczną inną niż MANNER.

### 2.2.5. Nieintersektywność

Modyfikatory nieintersektywne mają nielokalny wkład znaczeniowy. Przykładami modyfikatorów nieintersektywnych są

— „były”  
były prezydent (10)

— „fałszywy”  
fałszywy brylant (11)

— „co najmniej”  
co najmniej cztery dni (12)

*Były* i *fałszywy* są funktorami, które tworzą wraz ze znaczeniem nadrzędnika pojęcie na tyle zmodyfikowane, że jego denotacja nie musi lub nie może mieścić się w przecięciu oryginalnych denotacji. Funktor *co najmniej* bierze za argument nazwę liczby bądź miary i zwraca nowe określenie liczby bądź miary, które nie jest zawężeniem określenia oryginalnego.

### 2.2.6. Relacyjność

**Opis zjawiska** Przez relacyjność modyfikatora rozumiemy tradycyjnie tę jego cechę, że jego znaczenie zależy w pewnym ustalonym zakresie od znaczenia jego nadrzędnika. Dogodnym sposobem jej reprezentowania jest relacyjny argument niejawny (punkt 6 na liście niejawnych argumentów w części 2.2.3).

— „ledwo” w znaczeniu ‘prawie nie’

Ledwo zdał egzamin. (13)

Funktor *ledwo* ma domyślny argument – referencyjny porządek, zależny od nadrzędnika.

— niedużo

niedużo mleka (14)

— dużo

dużo złota (15)

Te typowe przykłady przymiotników relacyjnych mają znaczenie jednoznacznie dodeterminowane znaczeniem nadrzędnika. „Dużo złota” to inna ilość niż „dużo węgla”.

— już

już za daleko (16)

— jeszcze

jeszcze we wtorek (17)

— aż

aż cztery (18)

Te funkcory mają domyślny argument wskazujący skalę jako czasową, przestrzenną itp. ale niedostarczony przez sam nadrzędnik. Nadrzędnik ogranicza ich znaczenie, niekoniecznie je determinując.

### 2.2.7. Przyrematyczność

**Opis zjawiska** Przyrematyczne są te partykuły, które obejmują swoim zakresem zawsze remat zdania, choć może to być niewidoczne na poziomie morfoskładni i szyku, często swobodnego.

— chyba

**Oslo** jest chyba stolicą Norwegii. (19)

Oslo jest chyba stolicą **Norwegii**. (20)

— także

Pójdziemy także do kina. (21)

— również

Pojedziemy również do Stanów. (22)

— też

Polecimy też do Berlina. (23)

— nawet

Popłyniemy nawet do Oslo. (24)

## 2.2.8. Operatory nieprawdomówne (*nonveridical*)

**Opis zjawiska** Szereg wyrażeń języka można definiować jako funktory od argumentów zdaniowych, w tym przysłówki, partykuły i czasowniki przyjmujące dopełnienie zdaniowe. Kluczowym podziałem przy reprezentowaniu ich znaczenia jest to, czy prawdziwość frazy pociąga prawdziwość argumentu (*veridicality*) czy nie (*nonveridicality*). Przykładem jest większość partykuł i wyrażeń przyimkowych wyrażających modalność epistemiczną, jak

— „chyba”

Chyba idzie zima. (25)

— „pewnie”

Pewnie spadnie śnieg. (26)

— „może”

Może będzie przymrozek. (27)

— „być może”

Być może spadnie grad. (28)

— „prawdopodobnie”

Prawdopodobnie będzie ślisko. (29)

które stanowią ocenę stopnia uzasadnienia sądu. *Prawdopodobnie* może też być partykułą modalną aletyczną wyrażającą możliwość i w tym znaczeniu też jest operatorem nieprawdomównym.

Tryb rozkazujący także jest nieprawdomówny, choć zależy to dodatkowo od znaczenia czasownika: Ze zdania „Podaj mi sól” wynika pragmatycznie, że teraz odbiorca nie podaje soli, ale ze zdania „Bądź punktualny” nie wynika, że odbiorca punktualny nie jest.

**Niefaktywność** Najlepiej znanym podtypem nieprawdomówności jest niefaktywność, tradycyjnie przypisywana leksemom czasownikowym. W naszej analizie faktywność i niefaktywność przysługuje pozycji w ramie walencyjnej, co uwzględnia odmienne zachowanie podrzędników tego samego leksemu wprowadzanych przez różne spójniki podrzędne.

— „roić sobie, że\_”

Anna roi sobie, że pada. (30)

— „chcieć, żeby\_”

Anna chce, żeby nie padało. (31)

Oddzielnej analizy wymaga prawdomówność spójników leksykalnych i faktywność pozycji wypełnionych przez frazy czasownikowe niefinitywne – bezokolicznikowe, imiesłowowe, odśłownikowe, bezosobnikowe, którym nie przysługuje np. czas gramatyczny.

## 2.2.9. Cechy składniowo-semantyczne jednostek nieopisanych w zasobach

Słownik *Walenty*, który jest podstawowym źródłem sformalizowanej informacji o cechach składniowo-semantycznych jednostek współczesnego języka polskiego, nie obejmuje niektórych klas gramatycznych lub obejmuje je w niedostatecznym stopniu. Według wstępnego rozoznania wiele z brakujących leksemów rzeczownikowych, przymiotnikowych, przysłówkowych i przyimkowych wnosi wkład znaczeniowy lokalnie, a więc w sposób najbardziej podstawowy. Niezbędny jest jednak ich opis składniowo-semantyczny pozwalający przypisać im odpowiednią



rolę tematyczną. Dalsze potrzeby zasobowe obejmują typ semantyczny leksemów nieopisanych w Słownosieci.

**Lista cech składniowo-semantycznych** Docelowy opis cech składniowo-semantycznych obejmuje następujące cechy:

1. Rola tematyczna łącząca wyraz z nadrzędnikiem, jeśli leksem podrzędnika ma ją do siebie przypisaną oraz jest uogólnienie intersektywny
2. Struktura składniowa argumentów leksemu (rama składniowa), np. dla przyimka będzie to
  - a) Liczba argumentów jawnych przyimka (por. *między\_a*, *między\_i*)
  - b) Okalającość przyimka (por. *między\_a*)
  - c) Pozycyjność przyimka (por. postpozycyjne (*100 lat*) *temu*)
  - d) Przypadek wymagany przez przyimek
3. Preferencje selekcyjne i role tematyczne argumentów intersektywnych w sposób uogólniony
4. Typ semantyczny leksemu, jeśli leksem nie występuje w Słownosieci. Dla przyimków i spójników podrzędnych jest to typ semantyczny frazy, których są elementem głównym.

#### 2.2.10. Oznaczanie miary i liczebności referenta

**Opis zjawiska** Szczegółnej uwagi wymagają wyrażenia, których typem semantycznym jest ILOŚĆ dla wyrażen policzalnych lub MIARA dla wyrażen niepoliczalnych. Ich rolą tematyczną jest odpowiednio COUNT lub MEASURE.

**Użycie pojemnikowe** Użycie pojemnikowe rzeczownika to takie, w którym rzeczownik odnosi do miary referenta swojego podrzędnika rzeczownikowego w dopełniaczu w morfoskładniowym (niesemantycznym) podejściu do składni. Przykładem jest „łyżka” w wyrażeniu „łyżka cukru”. Żeby otrzymać rozkład zdania zawierającego rzeczowniki w użyciach pojemnikowych w składni semantycznej, jak Universal Dependencies, traktujemy pojemniki jak podrzędniki swoich tradycyjnych składniowych nadrzędników. Do tego celu niezbędne jest oznaczenie pojemników w języku polskim.

**Podział liczebników ze względu na dokładność** Liczebniki niewłaściwe (*trochę*, *sporo*) i większość rzeczowników w użyciu oznaczającym miarę nie podaje dokładnej wartości wielkości, którą opisuje. Żeby reprezentować ich znaczenia, trzeba je podzielić ze względu na dokładność, opisać zakres niedokładności i zaproponować spójną koncepcję reprezentacji znaczeń niedokładnych.

**Miara referentów niepoliczalnych** Do zdefiniowania, które leksemy wprowadzają w języku polskim relację MEASURE (np. *niedużo*), konieczne jest ich oznaczenie w zasobie typu słownikowego.

#### 2.2.11. Niedenotatywne funkcje języka

**Opis zjawiska** Szczegółnej uwagi wymagają wszystkie leksemy i konstrukcje pełniące funkcję fatyczną (*halo*), konatywną (np. tryb rozkazujący), emotywną (np. ekspresywizmy) lub metatekstową (np. *że tak powiem*). Do opisu ostatniej należą złożone zagadnienia znaczenia spójników, wyrażen przyrematycznych i przytematycznych.

Przykłady wyrażen pełniących pierwotnie funkcję niedenotatywną:

- „halo”
- „brawo”
- „brr”
- „miau”
- „że tak powiem”

**Metatekstowość** Wyrażenia metatekstowe, przede wszystkim partykuły i spójniki, odnoszą do fragmentu wypowiedzenia lub wyrażają komentarz na jego temat. Przykładem jest

- „więc”

Padało, więc wyszliśmy tańczyć. (32)

Spójnik *więc* informuje, że zachodzi związek quasi-przyczynowy między odniesieniami kordinowanych członów.

### 2.2.12. Podmiot epistemiczny

Niektóre wyrażenia i kategorie gramatyczne odnoszą do stanów mentalnych osoby niekoniecznie tożsamej z nadawcą. Należą tu

- partykuły modalne
- „chyba”

Piotr mówi, że chyba skończyli już układać. (33)

- „pewnie”

Piotr mówi, że pewnie skończyli już układać. (34)

W obu przykładach osoba referencyjna potwierdziłaby modyfikowany sąd, gdyby wiedziała, że pewna brakująca przesłanka jest prawdziwa, ale osoba tą nie jest nadawca zdania.

- partykuły miriatywne
- „nawet”

Anna narzekała, że nie przyszedł nawet przywitać się z rodzeństwem. (35)

Modyfikowany sąd dziwi podmiot epistemiczny, który ponownie nie musi być tożsamy z nadawcą.

- tryb rozkazujący, ponieważ wskazuje na intencje nadawcy
- spójniki przeciwstawne i wprowadzające okoliczniki przyzwolenia
- „choć”

Maria powiedziała, że wyszli tańczyć, chociaż padał deszcz. (36)

- „ale”

Maria powiedziała, że padał deszcz, ale oni tańczyli. (37)

- „a”

Maria powiedziała, że padał deszcz, a oni tańczyli. (38)

## 2.3. Koncepcja metaopisu semantycznego

### 2.3.1. Wprowadzenie

Znaczenie wypowiedzi reprezentowane jest w systemie dwupoziomowo. Najpierw za pomocą grafu semantycznego opisującego występujące w tekście pojęcia oraz relacje pomiędzy nimi a następnie w postaci formuły logiki pierwszego rzędu rozszerzonej o predykat metajęzykowy i dodatkowe kwantyfikatory.

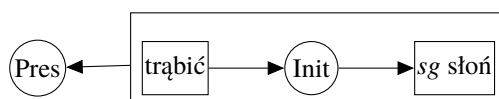
Różnica pomiędzy reprezentacjami polega na tym, że graf semantyczny jest bliższy składni, natomiast formuła logiczna posiada formalnie zdefiniowaną semantykę. Formuła logiczna jest generowana z grafu semantycznego za pomocą algorytmu, co pozwala przenieść formalną semantykę na graf. Niniejszy dokument skupia się na opisie poszczególnych zjawisk językowych za pomocą grafów semantycznych.

Grafy semantyczne inspirowane są pracami Johna Sowy (“Knowledge Representation: Logical, Philosophical, and Computational Foundations”)

Podstawy reprezentacji zostały opracowane w ramach projektu Clarin i są opisane w dokumencie: “Język reprezentacji znaczenia dla języka polskiego” oraz szkicowo opublikowane na konferencji COLING 2016 w pracy “ENIAM: Categorical Syntactic-Semantic Parser for Polish”. Poniżej zostanie opisane sposób w jaki poszczególne zjawiska składniowe są reprezentowane za pomocą grafów pojęć. Opis ma charakter techniczny i nie obejmuje rozważań dotyczących teoriomodelowej semantyki poszczególnych konstrukcji, analizy możliwych reprezentacji ani kontekstów literaturowych. Celem, oprócz opisanie zasad tworzenia form logicznych przez parser ENIAM jest określenie formatu zasobów leksykalnych potrzebnych do ich wygenerowania.

### 2.3.2. Pojęcia i relacje między pojęciami

Formuły naszego języka reprezentacji znaczenia wyrażamy graficznie w formie grafów semantycznych. Przykładowo dla zdania *Słoń trąbi* uzyskamy graf



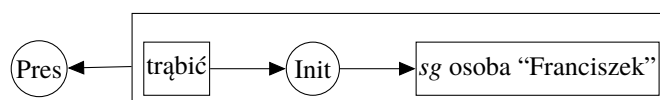
W powyższym grafie pudełka reprezentują obiekty, o których jest mowa. Występuje zatem obiekt *słoń* i zdarzenie *trąbić*. Symbol *sg* określa licznosc obiektów jako dokładnie 1.

Kółeczka reprezentują relacje między obiektami. *Init* wskazuje na to, że *słoń* jest inicjatorem *trąbienia*, a *Pres* na to, że zdarzenie jest równoczesne z czasem jego wypowiedzenia. Strzałka wchodząca to pierwszy argument, wychodząca drugi. Źródłem informacji o relacjach łączących czasowniki (a w przyszłości również rzeczowniki, przymiotniki i przysłówki) z ich argumentami jest słownik walencyjny Walenty.

Zewnętrzna ramka to kontekst. Reprezentuje on sytuację, czyli podzbiór czasoprzestrzeni, w którym istnieją byty wskazane przez pojęcia w pudełkach i zachodzą wymienione w kółeczkach relacje pomiędzy nimi.

Standardowo każdej jednostce leksykalnej (leksemowi lub wyrażeniu wielosłownemu) zawartej w zdaniu odpowiada pudełko, a relacji składniowej kółeczko. Pudełka zawierają sensy jednostek leksykalnych oraz ich liczbę, a w przypadku nazw własnych nazwę i typ nazwy wła-

snej np *Franciszek trąbi*:



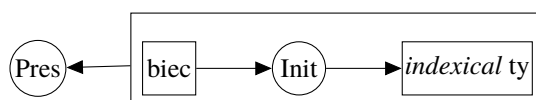
### 2.3.3. Wprowadzanie bytów do dyskursu

W języku logiki pierwszego rzędu wzmiankowane byty stanowią odniesienia zmiennych, a same zmienne wprowadzane są przez kwantyfikatory. Przy czym dostępne są dwa rodzaje kwantyfikacji: uniwersalna w której wartościowanie wprowadzanej zmiennej przebiega po wszystkich elementach uniwersum spełniających restrykcję oraz egzystencjalna, przy której zmienna wartościowana jest jednym z bytów spełniających restrykcję.

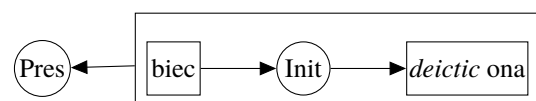
W przypadku grafów semantycznych zmienne nie są jawnie wskazane, tym niemniej zachodzi konieczność określenia w jaki sposób obiekty wskazywane przez pudełka są związane z uniwersum.

Domyślnie przyjmujemy kwantyfikację egzystencjalną, czyli np. zdanie *Słoń trąbi* stwierdza o istnieniu *słonia* i istnieniu zdarzenia *trąbienia*.

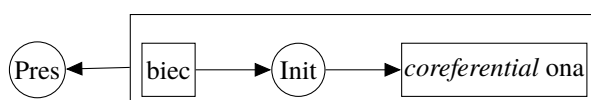
Z kolei dla zdania *Ty biegniesz* otrzymamy reprezentację z pojęciem okazjonalnym (indexical), który wiąże słowo *ty* z uczestnikami komunikacji, czyli odwołuje się do uniwersum na metapozymie.



a dla zdania *Ona biegnie* otrzymamy reprezentację z pojęciem deiktycznym



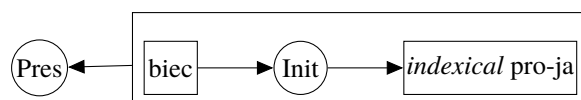
lub koreferencyjnym



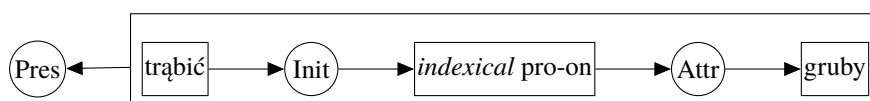
W powyższych trzech przypadkach nie następuje wprowadzenie nowego bytu do dyskursu.

Zaimki wnoszą informację o osobie, liczbie i rodzaju, która jest istotna przy ustalaniu ich referenta. Z tego względu lematyzujemy je do form uwzględniających te informacje (*ja*, *wy*, *one*, *która*), ewentualnie uzupełniając np. *ja-m*.

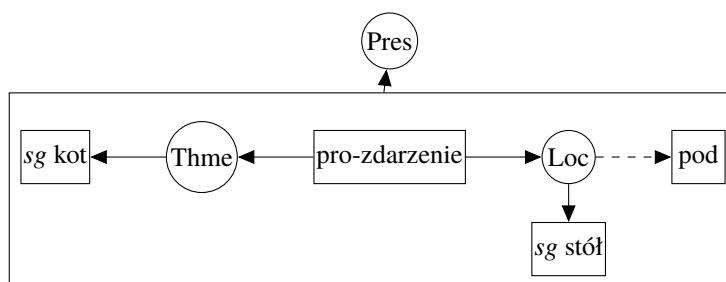
Kiedy podmiot nie jest dany w sposób jawny, jest reprezentowany za pomocą wariantu nie-mego zaimka *pro* np. *Biegnę*..



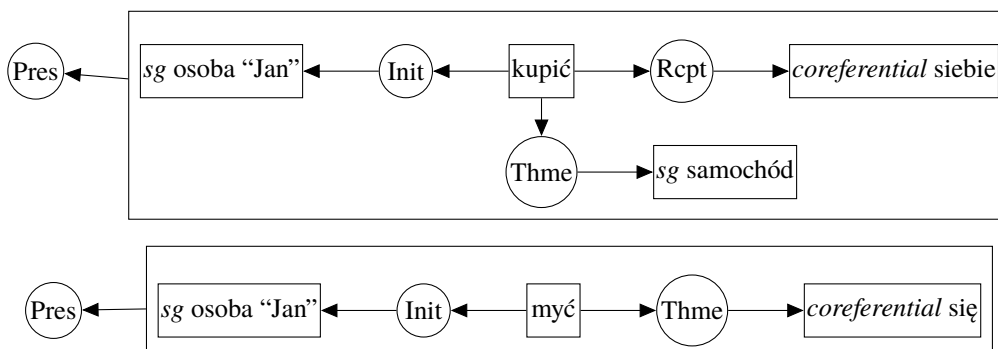
Zaimek *pro* używany jest również, kiedy przedmiot jest wprowadzany poprzez podanie swojej cechy, co ma miejsce we frazach rzeczownikowych, których głową jest przymiotnik np. *Gruby trąbi*.



Kolejnym zastosowaniem niemego zaimka są równoważniki zdań. Dodajemy do nich *pro*-zdarzenie, które wiąże poszczególne elementy zdania, np. *Kot pod stołem*



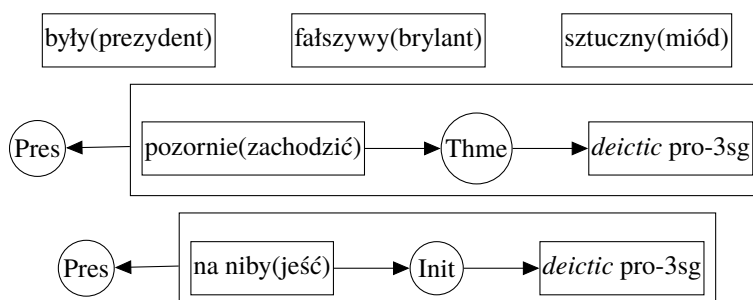
Leksemy *siebie* i *się* są koreferencyjne: *Jan kupił sobie samochód, Jan myje się*



W dalszym toku tekstu wprowadzimy jeszcze pojęcia interrogatywne (interrogative), występujące przy zadawaniu pytań, jak również kwantyfikację uniwersalną (przebiegającą po wielu bytach).

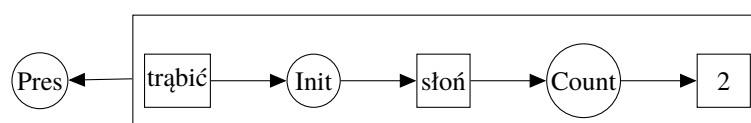
### 2.3.4. Modyfikatory nieintersektywne

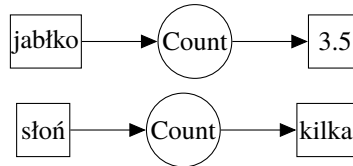
Funkcja modyfikacji nieintersektywnej zachodzi między wyrażeniem określającym a jego nieintersektywnym określnikiem, który może być przymiotnikiem (*były prezydent, fałszywy brylant, sztuczny miód*), przysłówkiem (*pozornie zachodzi*) lub wyrażeniem przyimkowym (*je na niby*).



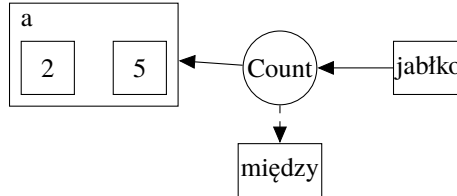
### 2.3.5. Liczebność i miara

Liczebność odniesienia rzeczowników policzalnych wskazuje relacja Count, np: *Dwa słonie trąbią, trzy i pół jabłka, kilka słoni*

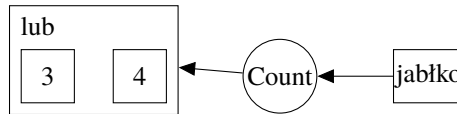




*między 2 a 5 jabłek*

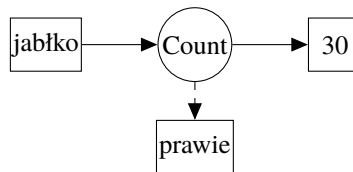


*3 lub 4 jabłka*



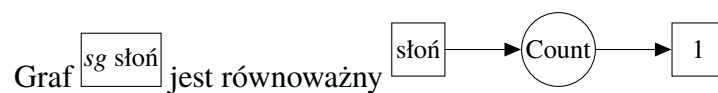
Liczebniki określone reprezentujemy za pomocą liczb, a nieokreślone (np *kilka*, *trochę*) za pomocą leksemów. Dwa ostatnie z powyższych przykładów wykorzystuj notację dla koordynacji opisaną szczegółowo w rozdziale 2.3.14.

Liczebniki mogą być modyfikowane przez operatory adnumeratywne np. *prawie 30 słoni*.

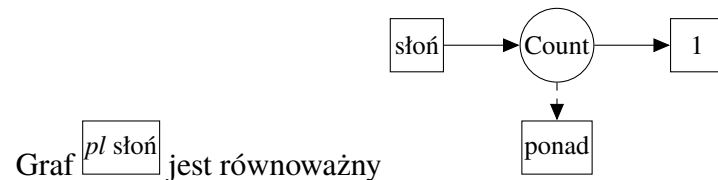


Podobnie zachowują się *około*, *koło*, *co najmniej*, *niemal*, *niespełna*, *co najwyżej*, *ponad*, *prawie*, *przeszło*, *przynajmniej*.

Liczość rzeczowników, które nie mają jej jawnie zadanej przez liczebnik wnioskujemy na podstawie liczby gramatycznej i zapisujemy za pomocą uproszczonej notacji. Symbol *sg* określa liczość obiektów jako dokładnie 1.

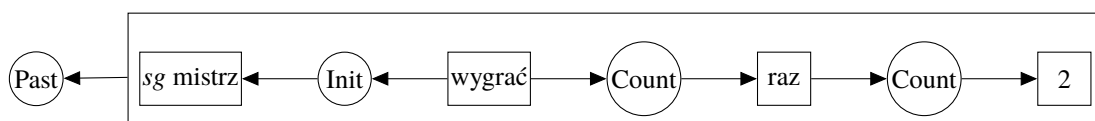


Oprócz niego stosujemy symbol *pl* na określenie liczości większej niż 1.

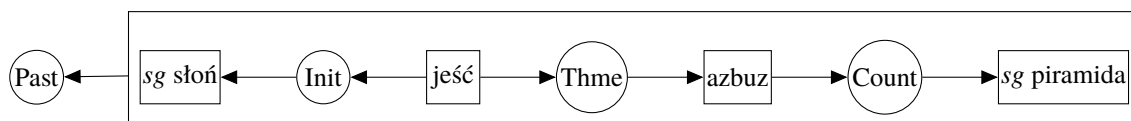


Rzeczowniki plurale tantum nie wnoszą informacji o liczości.

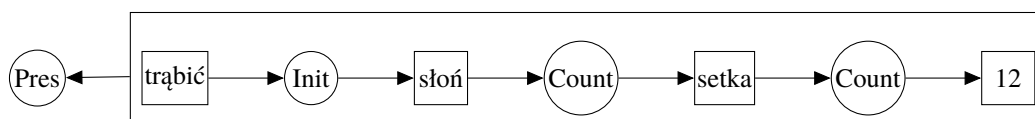
Dla innych części mowy liczość uznajemy za nieokreśloną, gdy nie jest jawnie wskazana. Liczebność czasowników wskazuje zazwyczaj leksem *raz*, np *Mistrz wygrał dwa razy*.



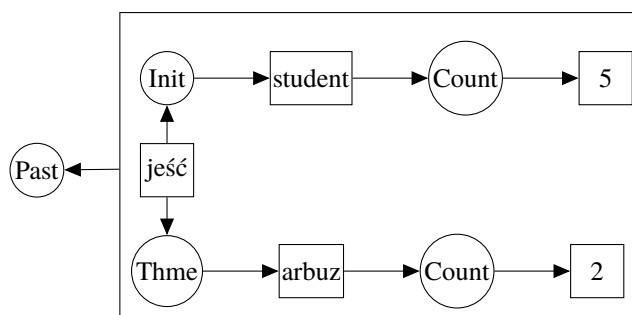
Oprócz liczebników liczność mogą wyrażać rzeczowniki użyte w znaczeniu pojemnikowym, np: *Słoń zjadł piramidę arbuźów*.



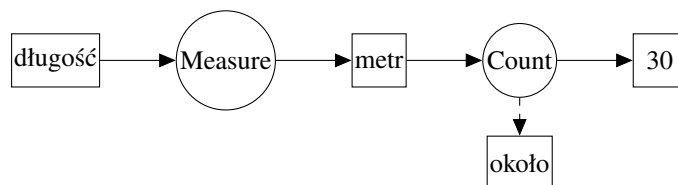
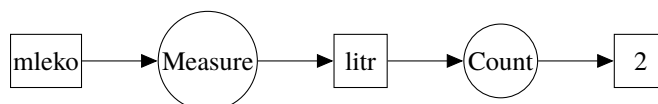
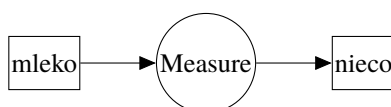
Słów takich jak *para*, czy *setka* nie zamieniamy na liczby by odróżnić je od liczebników głównych, np. *Dwanaście setek słoń trąbi*.



Nie rozróżniamy kolektywności i dystrybutywności, np. w dla zdania *Pięciu studentów zjadło 2 arbuzy* nie wskazujemy czy *studenci* zjedli w sumie 2 *arbuzy*, czy też każdy z nich zjadł 2 *arbuzy*.

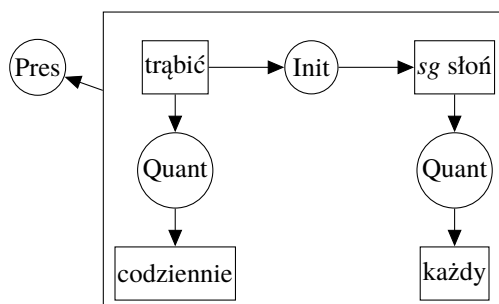


Pojęcia do których odnoszą się rzeczowniki niepoliczalne są oznaczane symbolem *mass*, a ich miara jest wyrażana relacją *Measure*. Miara może być wyrażona za pomocą liczebników np. *niecio mleka* lub pojemników np. *dwa litry mleka*, czy *długość około 30 metrów*



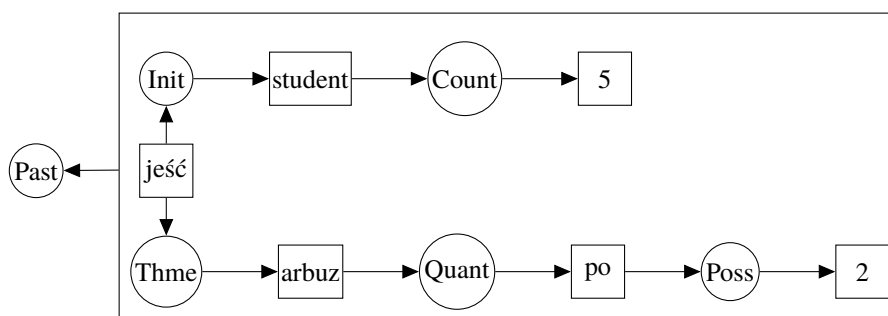
### 2.3.6. Kwantyfikatory

Kwantyfikatory i określenia częstości traktujemy jak pozostałe pojęcia np. *Każdy słoń trąbi codziennie*:

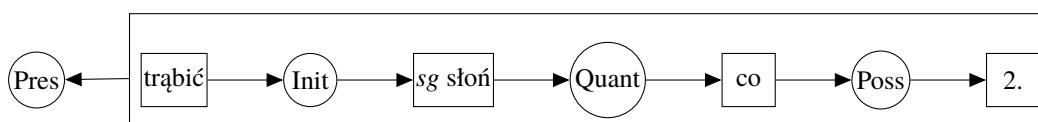
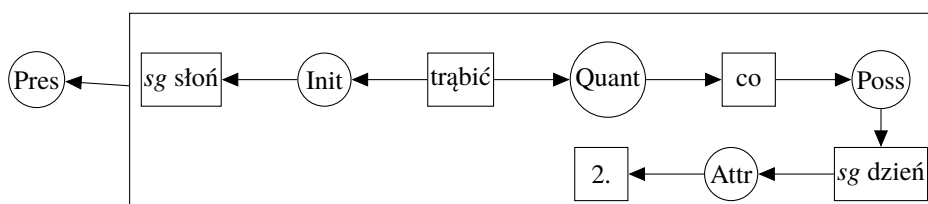


Podobnie zachowuje się *tylko*, *pewien*, *niektóre*, *wszystkie*, *zawsze*, *zwykle*, *czasami*

Dystrybutywne *po* traktujemy jako kwantyfikator, np. *Pięciu studentów zjadło po 2 arbuzy*



Kwantyfikujące częstość i porządek *co* reprezentujemy analogicznie jak *po*, np. *Słoń trąbi co drugi dzień*, *Co drugi słoń trąbi*.

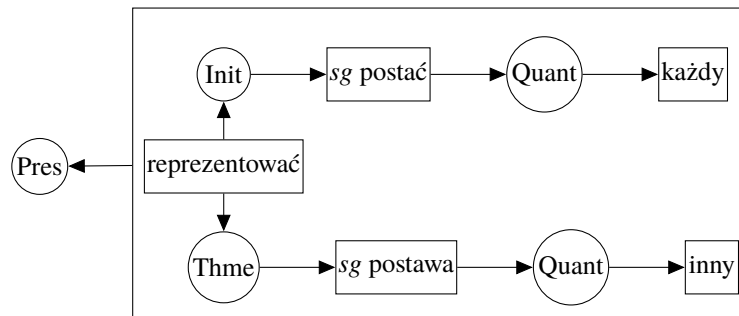


Podobnie we frazie *połowa słońi* występuje kwantyfikacja, zaś we frazie *połowa słońia* określona jest liczność.

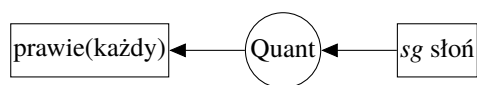
Nie wskazujemy zależności pomiędzy kwantifikatorami, czy też kwantyfikatorów wprowadzonych przez kilka leksemów rozsianych po zdaniu, tak jak w przypadku zdania *Każda postać*



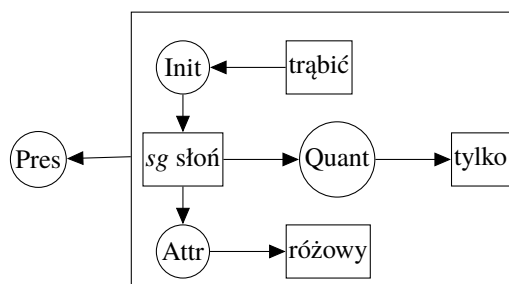
reprezentuje inną postawę



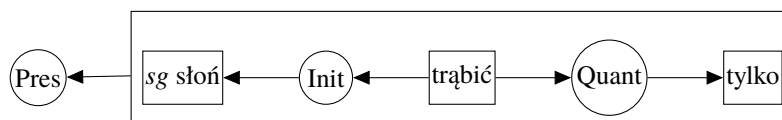
Kwantyfikatory można modyfikować nieintersektywnie np. *prawie każdy słoń*



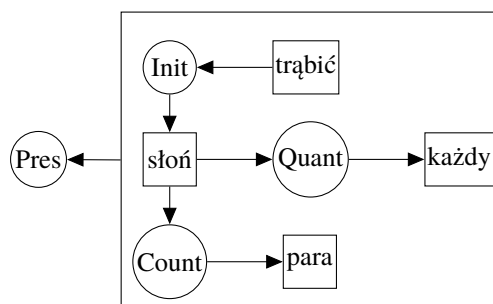
Kwantyfikatory można też modyfikować intersektywnie. Restrykcję kwantyfikatora wyznacza poddrzewo do którego przyłączona jest relacja Quant. Zakres kwantyfikatora nie jest jawnie wskazany. Np. zdaniu *Tylko różowy słoń trąbi* odpowiada graf



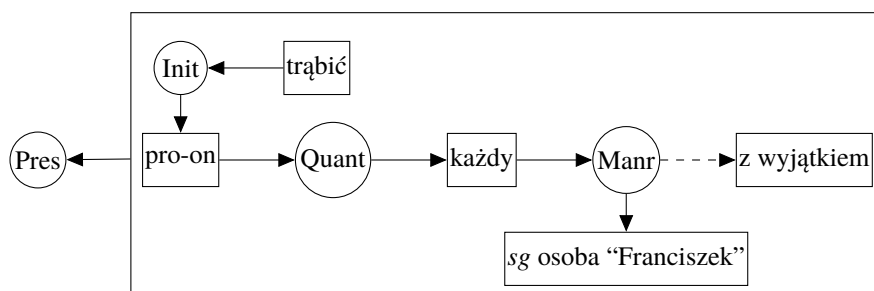
w którym restrykcja kwantyfikatora *tylko* obejmuje pojęcia *słoń*, *różowy* oraz relację między nimi. Zaś zdaniu *Słoń tylko trąbi* odpowiada graf



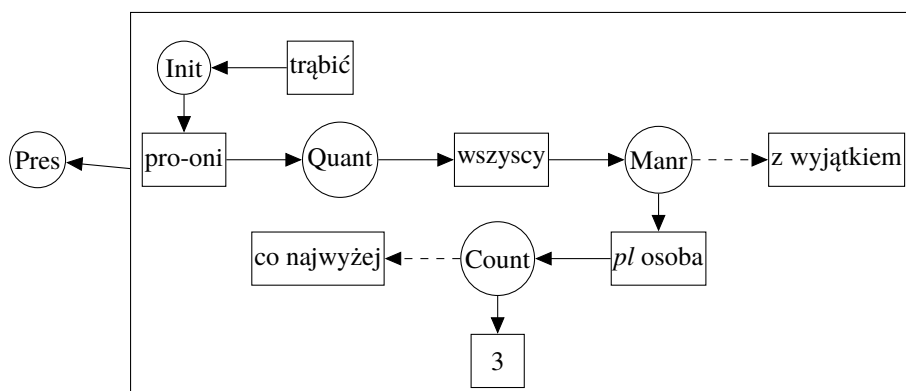
Kwantyfikacja może być połączona z licznością np. *Każda para słoń trąbi*



Kwantyfikatory mające dodatkowe argumenty, np: *Każdy z wyjątkiem Franciszka trąbi*,



*Wszyscy z wyjątkiem co najwyżej trzech osób trąbią.*

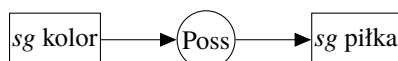


### 2.3.7. Relacja quasi-posiadania i identyczności

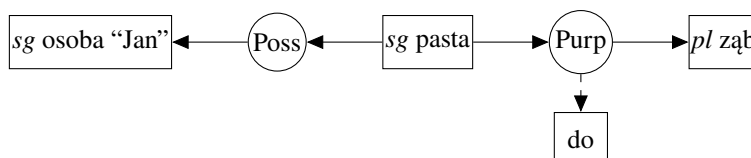
Predykat Poss symbolizuje relację posiadania (*piłka chłopca*) oraz pozostałe relacje wskazywane przez modyfikator rzeczownikowy w dopełniaczu (przydawkę dopełniaczową) (*sposób wykonania, prezeska organizacji, obrzeża Warszawy*). Poss jest domyślną rolą dla przydawki dopełniaczowej. W konkretnych przypadkach np. przy użyciu pojemnikowym jest zastępowana inną relacją.



Predykat Poss wykorzystujemy również w sytuacjach, gdy pojęcia mają charakter funkcji biorących odniesienie jednego ze swych podrzędników i określających swoje odniesienie na tej podstawie, np odniesieniem frazy *kolor piłki* mamy *piłkę*, z której wyłuskujemy cechę.

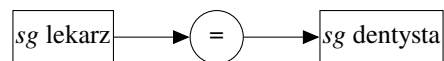


Przydawki przyimkowe domyślnie traktujemy tak jak argumenty czasownika. Na przykład *pasta do zębów Jana* jest reprezentowana jako



Relacja identyczności komunikowana jest przez apozycję oraz przydawkę rzeczoną i zachodzi między odniesieniami obu wyrażeń. Apozycja (dwa rzeczowniki uzgodnione pod względem

przypadka) wyraża dwa określenia tego samego obiektu. Może to być typ i nazwa, albo dwa różne typy, np. *lekarz dentysta* zapiszemy jako



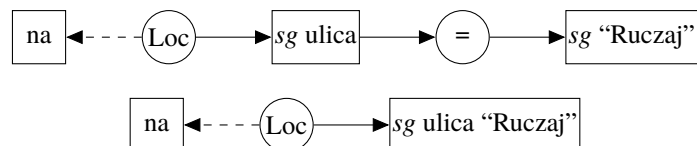
a w *mieście Warszawie*, zapiszemy jako



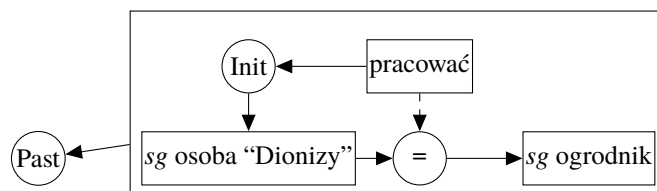
co możemy skrócić do



Przydawka rzeczowna (wyrażona w mianowniku) nadaje nazwę swojemu nadrzędnikowi, np. *na ulicy Ruczaj*, zapiszemy jako

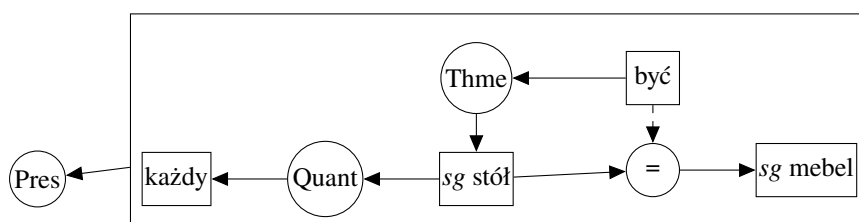
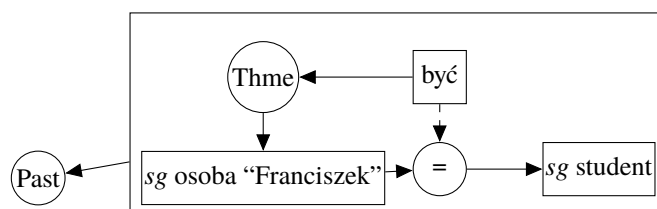


Relacja identyczności może zachodzić również pomiędzy podrzędnikami czasownika, np. w zdaniu *Dionizy pracował jako ogrodnik*. Walentym zjawisko to oznaczone jest rolą Attribute.



Przerywana strzałka pomiędzy *pracować* a relacją "=" oznacza, że relacja identyczności pomiędzy *Dionizym* a *ogrodnikiem* jest parametryzowana przez zdarzenie *pracowania*.

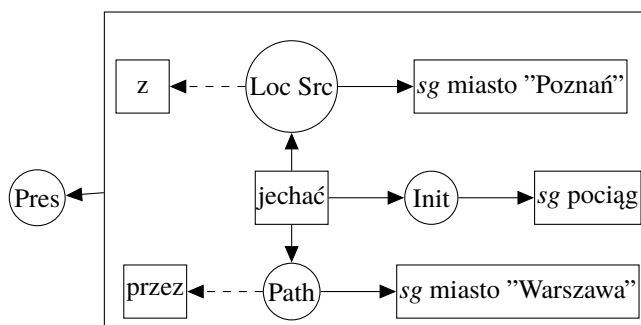
Analogicznie reprezentujemy konstrukcje predykatywne w zdaniach *Franciszek był studentem*, *Każdy stół jest meblem*.



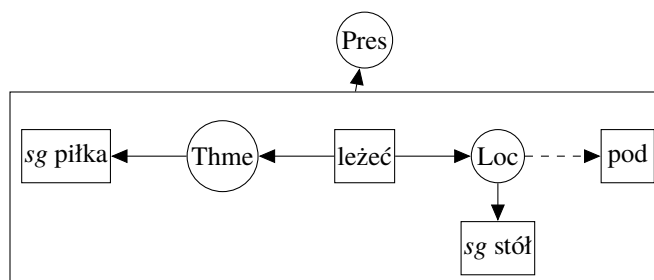
### 2.3.8. Relacje przestrzenne

Zazwyczaj wyrażane przez wyrażenia przyimkowe, przysłówki bądź zdania podrzędne. Przyimki lokatywne reprezentują relacje pomiędzy miejscami.

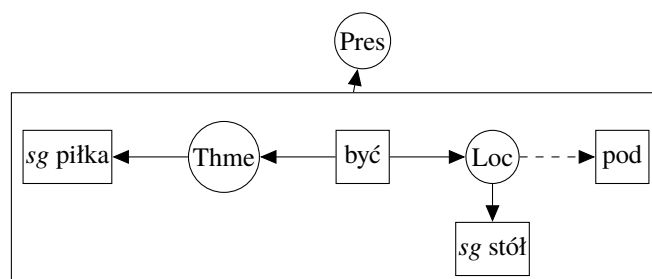
Relacja Location (Loc) wskazuje położenie sytuacji / zdarzenia. Relacje Location Source (Loc Src), Location Goal (Loc Goal), Path informują o obecności i kierunku ruchu. Z *Poznania jedzie pociąg przez Warszawę*:



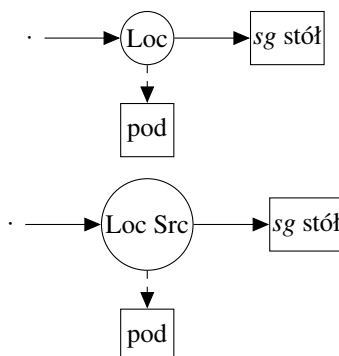
*Piłka leży pod stołem.*

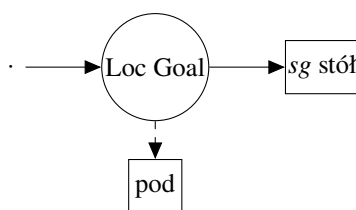


*Piłka jest pod stołem.*

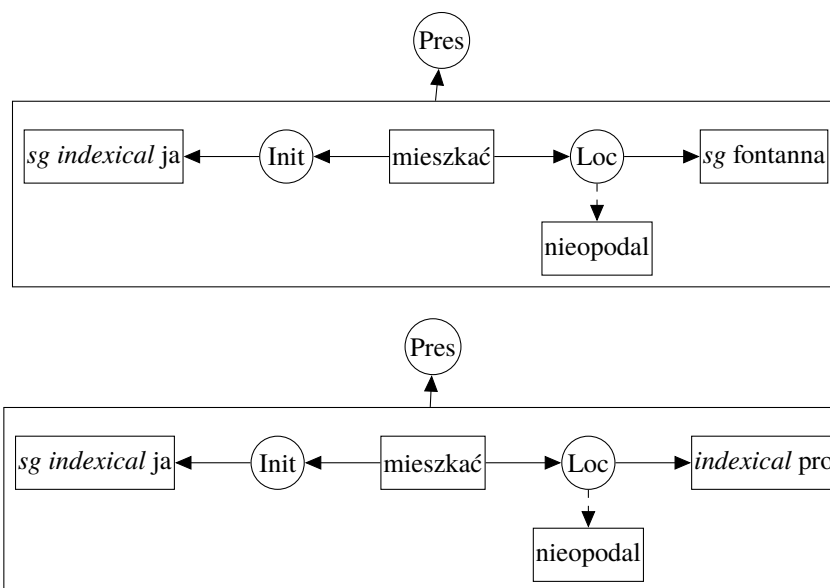


Przyjmujemy że odpowiadające sobie przyimki lokatywne, ablatywne i adlatywne mają ten sam sens, np *pod stołem*, *spod stołu*, *pod stół*:



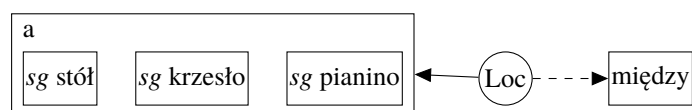


W niektórych sytuacjach argument przyimka opisującego relację przestrzenną nie jest jawny i staje się on przysłówkiem z argumentem okazjonalnym bądź koreferencyjnym, np: *Mieszkam nieopodal fontanny* vs *Mieszkam nieopodal*.

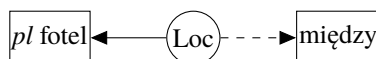


Przyjmujemy, że sens leksemu *nieopodal* jest identyczny w obu wypadkach. Podobnie zachowują się: *obok*, *blisko*, *dookoła*, *naokoło*, *naprzeciw*, *opodal*, *wewnątrz*, *wokół*, *wzdłuż*.

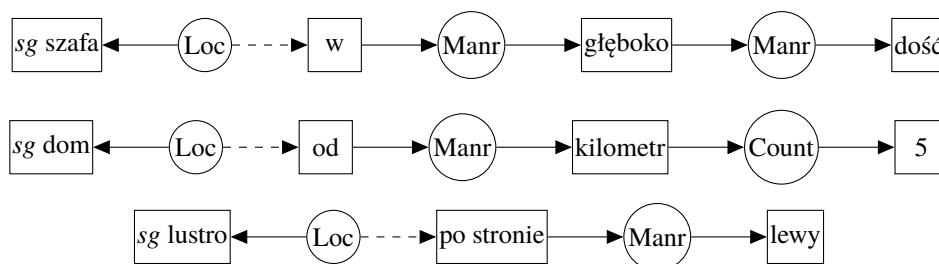
Wieloargumentowy przyimek *między* traktujemy jak przyimek jednoargumentowy, którego podrzędnikiem jest koordynacja, np. *między stołem, krzesłem a pianinem*



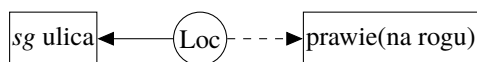
*między fotelami*



Przyimki lokatywne mogą być modyfikowane, np. *dość głęboko w szafie*, *5 km od domu*, *po lewej (prawej, drugiej) stronie lustra*.



Przymyki mogą być też modyfikowane nieintersektywnie np. *prawie na rogu ulicy*



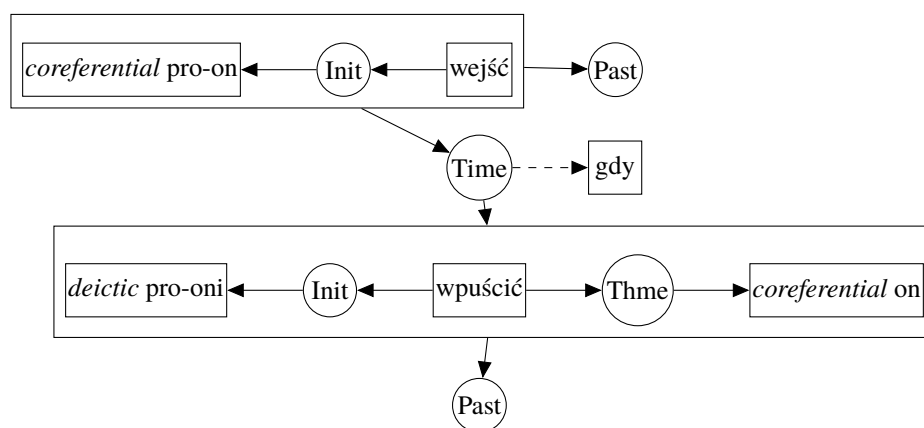
### 2.3.9. Sytuacje, procesy i relacje czasowe

Każdy proces wymieniowy w zdaniu umieszczamy w osobnym kontekście sytuacyjnym. W kontekstach sytuacyjnych uczestnicy istnieją a relacje między nimi zachodzą przez cały czas trwania sytuacji, chyba że uczestnik jest połączony relacją z procesem. W takiej sytuacji może on zostać w trakcie procesu stworzony, czyli zaistnieć dopiero na jego końcu, może powstawać stopniowo, może też przestać istnieć. Z kolei relacje wychodzące z procesów mogą się zmieniać w trakcie procesu. Natomiast relacje czasowe przypisane zdarzeniu dotyczą każdego uczestnika sytuacji a czas zdarzenia przysługuje całej sytuacji.

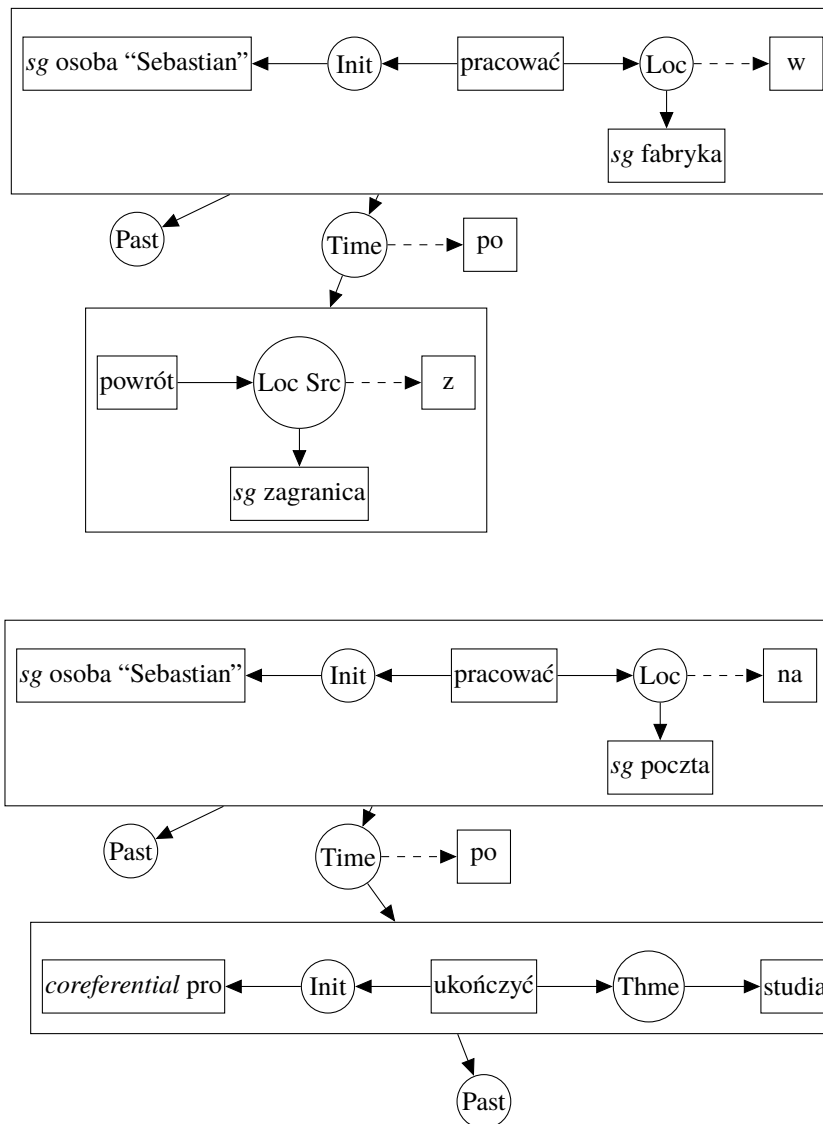
Na potrzeby reprezentacji za pomocą grafów semantycznych przyjmujemy, że relacje czasowe wiążą czas z sytuacjami, a znajdujące się w kontekstach pojęcia i relacje uznamy za fluenty niejawnie przez ten czas parametryzowane.

Relacje czasowe zazwyczaj wyrażane przez wyrażenia przymkowe, przysłówki bądź zdania podrzędne. Przymki temporalne reprezentują relacje pomiędzy punktami w czasie, interwałami oraz ich zbiorami. Relacje Time i Duration (Dur) informują o czasie, przypisanym do danego obiektu (zazwyczaj zdarzenia) oraz czasie jego trwania.

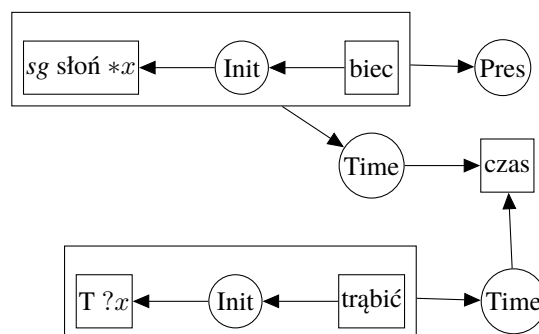
Spójnik *gdy* w jednym ze swoich znaczeń bierze sytuację (zdarzenie) i generuje jej czas, np. *Wszedł, gdy go wpuścili*:



Przymek *po* w znaczeniu czasowym odwołuje się do wcześniejszej sytuacji, bądź wcześniejszego zdarzenia. Może być ono wyrażone przez odśłownik, bądź rzeczownik, np: *Po powrocie z zagranicy Sebastian pracował w fabryce, Po ukończeniu studiów Sebastian pracował na poczcie.*

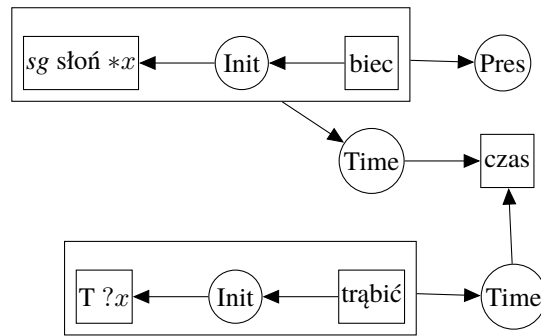


*Słoń biegnie trąbiąc.*

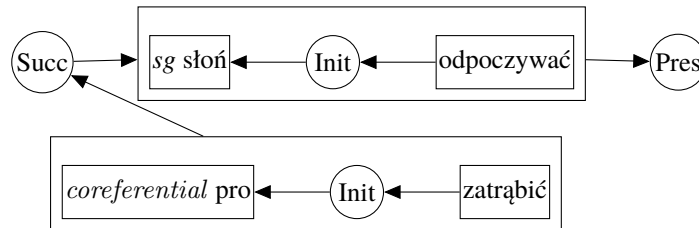


Symbol  $*x$  wprowadza zmienną  $x$ , która oznacza odniesienie pojęcia słoń, a symbol  $?x$  oznacza użycie zmiennej  $x$ . Razem wyrażają koreferencję. Pudełko *czas* służy do zareprezentowania równoczesności obu zdarzeń.

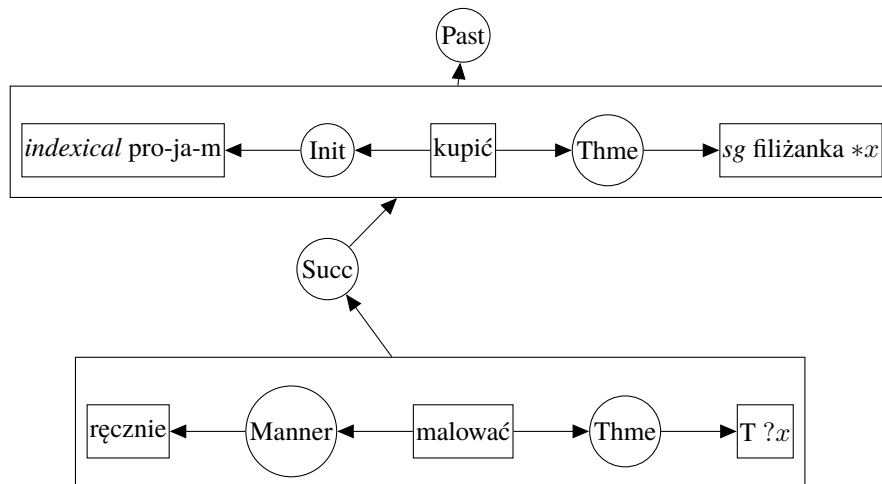
*Trąbiący słoń biegnie.*



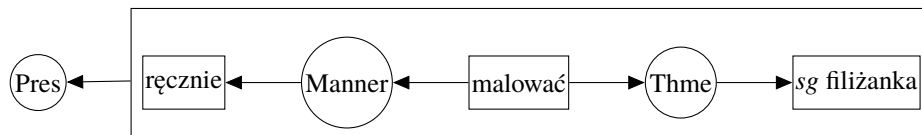
Następstwo zdarzeń reprezentujemy za pomocą relacji Succ, np *Słoń odpoczywa zatrąbiwszy*.



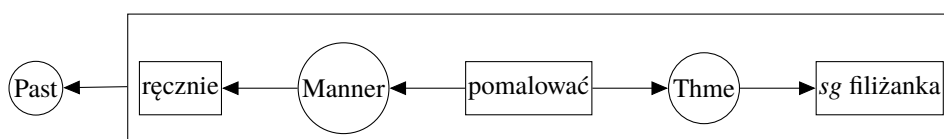
*Kupiłem ręcznie malowaną filiżankę.*



Czasowniki *być* i *zostać* w stronie biernej traktujemy jako czasowniki posiłkowe wnoszące do semantyki jedynie czas, np *Filiżanka jest malowana ręcznie*.

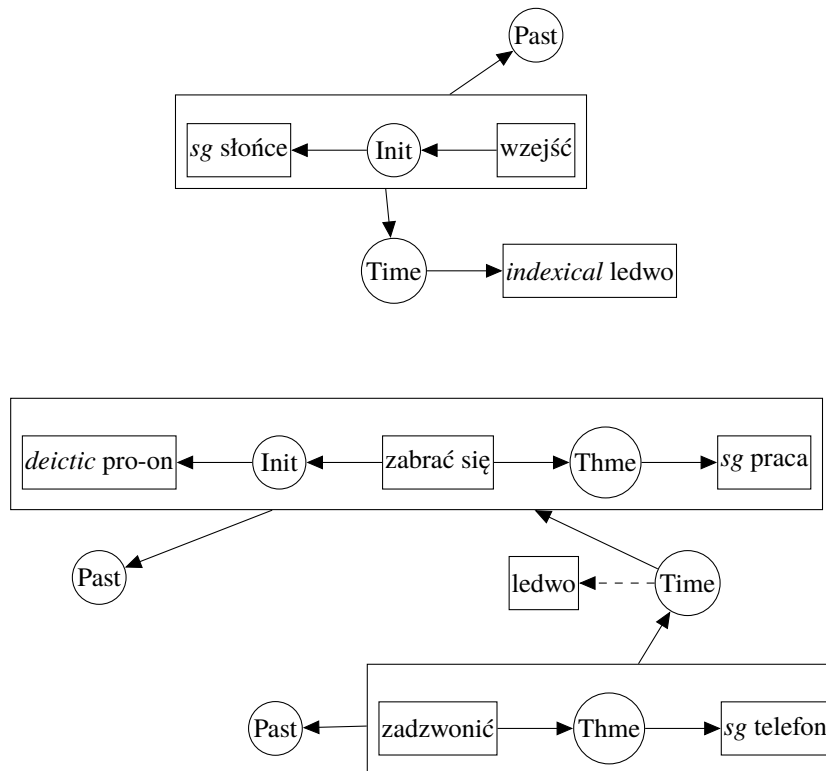


*Filiżanka została pomalowana ręcznie.*

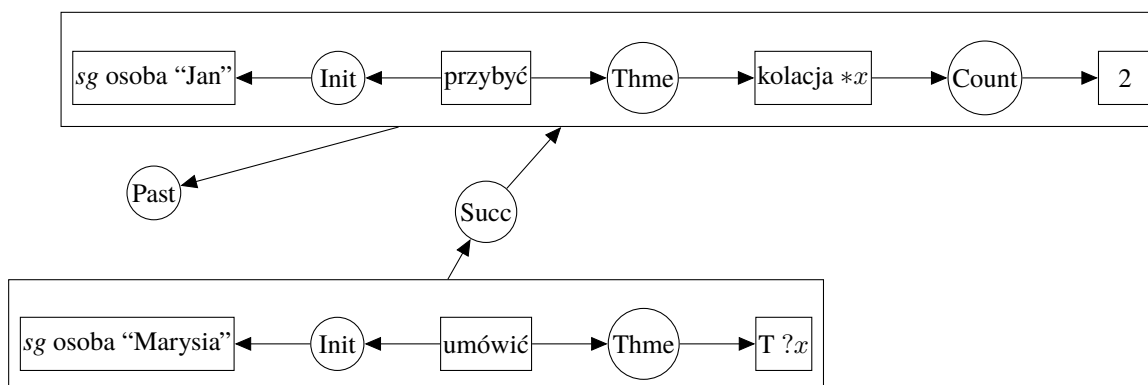




Leksem *ledwo* użyty w znaczeniu czasowym funktor biorący zdarzenie i generujący czas chwilę po nim. W funkcji spójnika podrzędnego, czas ten określa czas zdarzenia ze zdania nadrzędnego, w funkcji przysłówka czas ten określa czas zdarzenia wynikającego z kontekstu. Np.: *Słońce ledwo weszło*, *Ledwo zabrał się do pracy, zadzwonił telefon*:

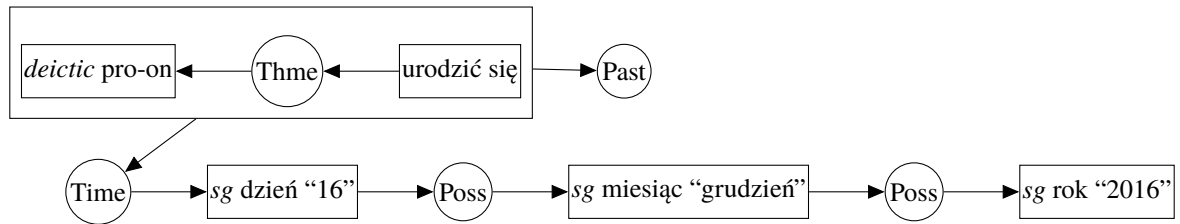


*Jan przybył na dwie umówione przez Marysię kolacje.*

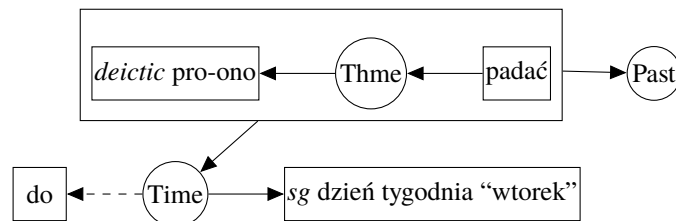


Poszczególne symbole czasowe (numery lat, dni tygodnia, numery godzin, minut, itp.) traktujemy jako nazwy interwałów. Wiele interwałów może nosić tą samą nazwę (np. co tydzień jest *poniedziałek*) podobnie jak wiele osób może mieć to samo imię, czy nazwisko. Złożone określenia czasu — 16 grudnia 2016 — traktujemy interwał odpowiadający dniowi o nazwie 16. Dzień ten jest dookreślony przez przynależność do miesiąca grudnia, który z kolei określo-

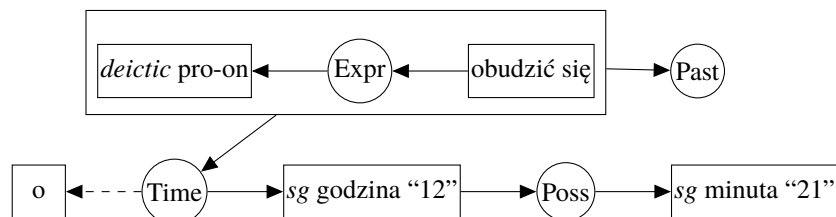
ny przynależnością do roku o numerze 2016. Zatem zdanie *Urodził się 16 grudnia 2016* będzie miało interpretację



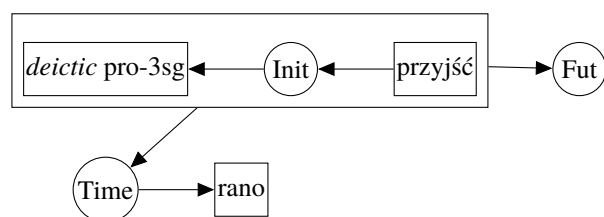
*Padało do wtorku.*



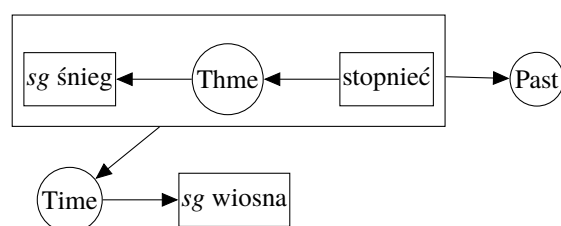
*Obudził się o 12:21.*



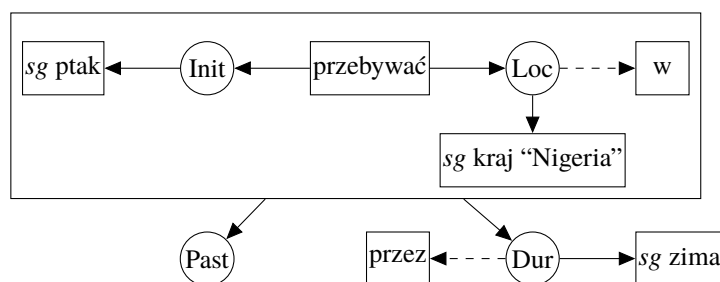
*Przyjdzie rano.*



*Śnieg stopniał wiosną.*

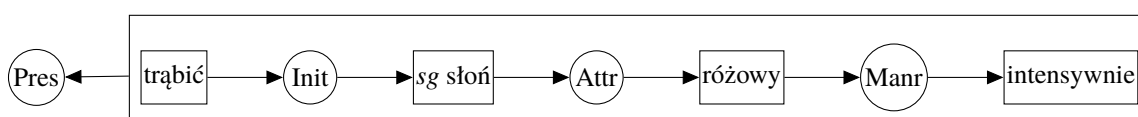


*Ptak przebywał w Nigerii przez zimę*



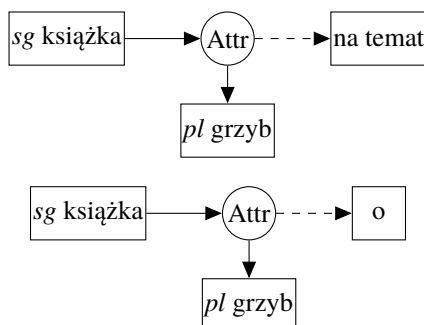
### 2.3.10. Cechy

Zazwyczaj cechy (atrybuty) wyrażane są przez przymiotniki, przysłówki lub wyrażenia przyimkowe, np.: *Intensywnie różowy słoń trąbił*:



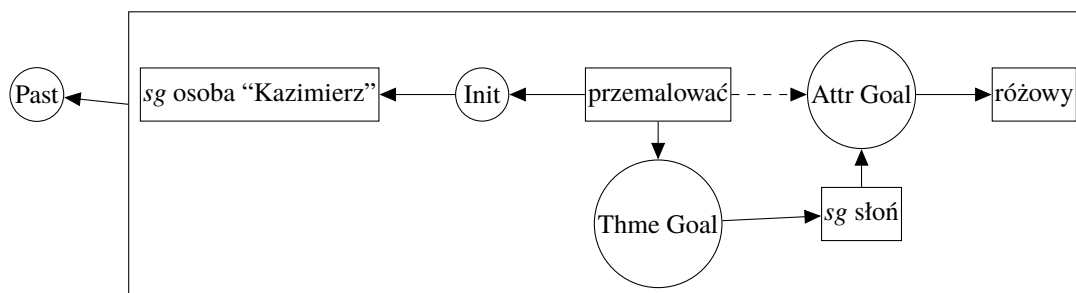
Cechę opisywaną przez przymiotnik domyślnie wyrażamy rolą Attribute (Attr) zaś cechę opisywaną przez przysłówki wyrażamy rolą Manner (Manr).

W przypadku wyrażen przyimkowych opisujących cechy rzeczowników również stosujemy rolę Attr, np.: *książka na temat grzybów, książka o grzybach*:

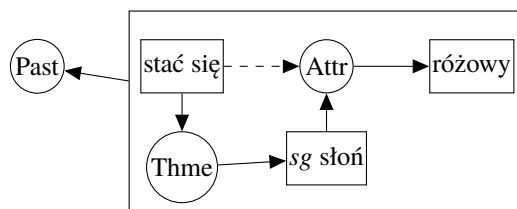


*na temat* jest przyimkiem złożonym traktowanym jako pojedyncza jednostka leksykalna, natomiast *o* w znaczeniu użytym w powyższym przykładzie jest jego synonimem. Przyimki złożone wskazują w swojej treści nazwę cechy, której dotyczą.

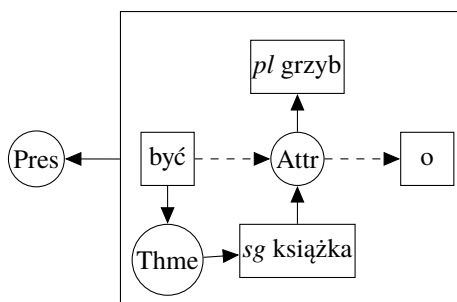
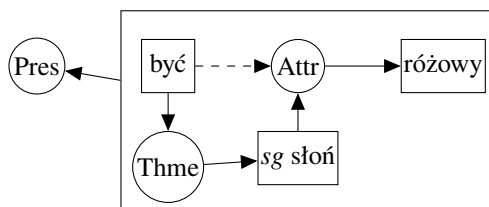
Cechy danego przedmiotu mogą być składniowo wyrażone przez podrzędnik czasownika. Słownik walencyjny Walenty sygnalizuje takie sytuacje nadając podrzędnikowi wyrażającemu cechę rolę Attribute oraz czyniąc go kontrolowanym przez podrzędnik będący nosicielem cechy, np. *Kazimierz przemalował słonia na różowo*:



Przerywana linia łącząca *przemaalować* i Attr Goal oznacza że relacja Attr Goal zmienia się w czasie trwania sytuacji w miarę postępów procesu malowania i nabierania przez słonia cechy bycia różowym. Analogicznie zinterpretujemy zdanie *Słoń stał się różowy*



Zdania z czasownikiem *być* stanowiące o cechach interpretujemy analogicznie do powyższego np. *Słoń jest różowy*, *Książka jest o grzybach*.

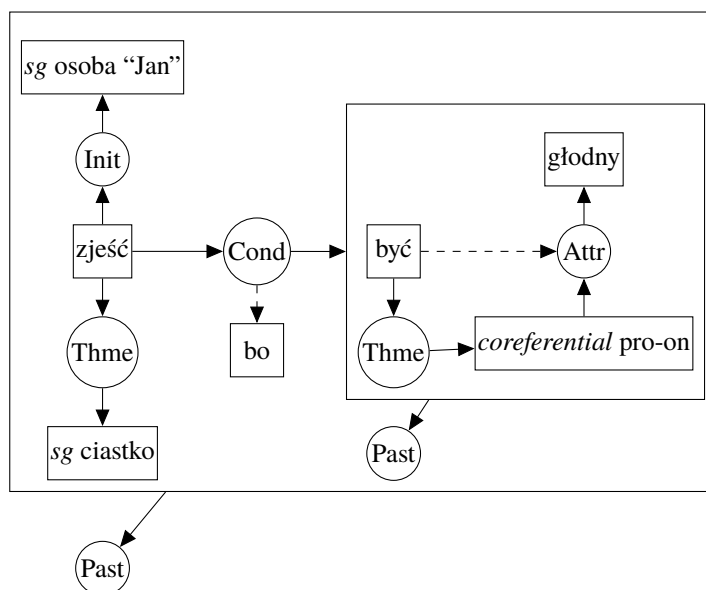
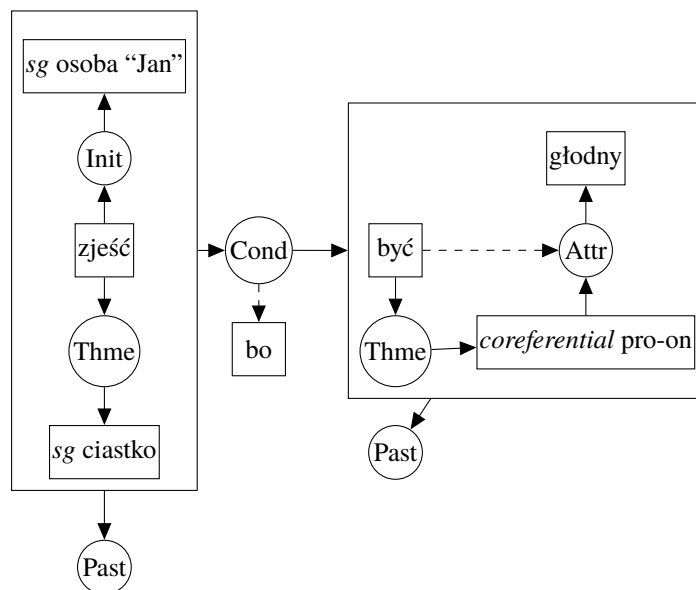


### 2.3.11. Przyczyna i cel

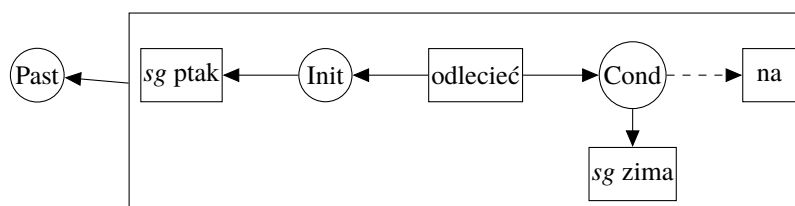
Rola Condition określa okoliczności, bez zaistnienia których nie doszłoby do zaistnienia akcji<sup>1</sup>. Informuje o przyczynie lub warunkach zaistnienia sytuacji. rola Purpose określa cel, który przyświeca sytuacji, na którą wskazuje predykat. Gdyby Initiatorowi nie przyświecał ten cel, nie doszłoby do danej sytuacji. Określa marzenia, życzenia i oczekiwania.

<sup>1</sup> Na podstawie instrukcji semantycznej Walentego

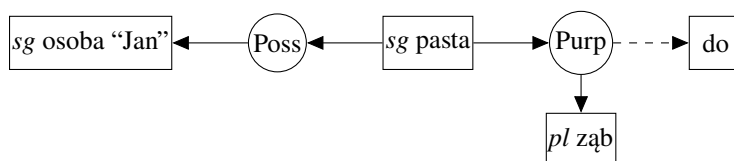
Role Condition i Purpose mogą wiązać zdarzenie bądź sytuację, np *Jan zjadł ciastko, bo był głodny*



Przyimek *na* w jednym ze swoich znaczeń bierze określenie czasu i wiąże ten czas z celem czynności. *Ptak odleciał na zimę*



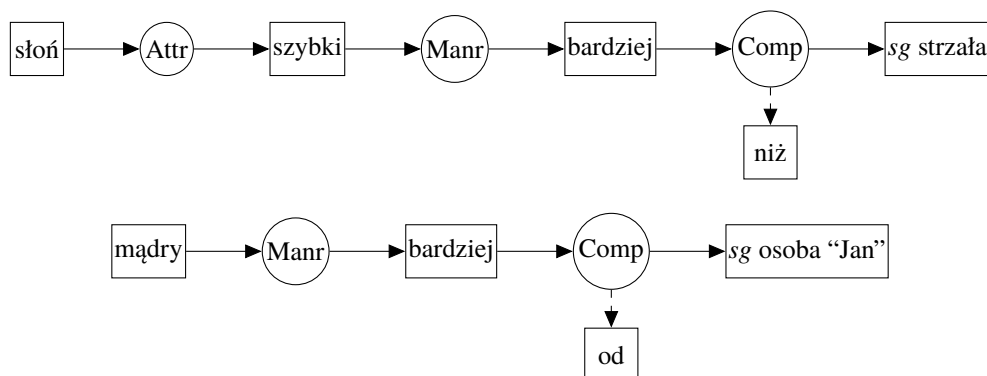
Rola Purpose użyta z rzeczownikami oznacza przeznaczenie obiektu, na przykład *pasta do zębów Jana* jest reprezentowana jako



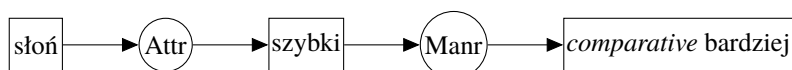
Pojęciem, które wiąże relacja Condition (typ semantyczny pojęć *bo*, *ponieważ*, *z powodu*) to pojęcie CZEMU, lub *przyczyna*. *przyczyna* to rola jaką pełni sytuacja.

### 2.3.12. Stopniowanie przymiotników i przysłówków oraz konstrukcje porównawcze, relacyjne i wprowadzające kolejność

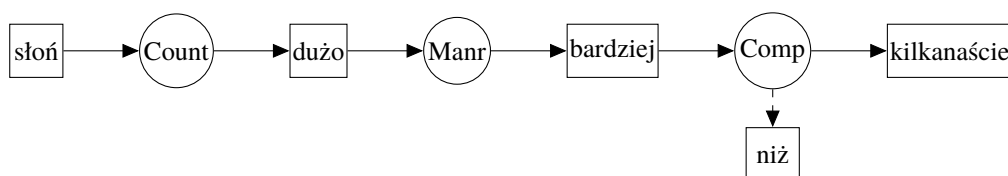
Zakładamy, że oba sposoby stopniowania przymiotników i przysłówków (przez dodanie *bardziej* lub *najbardziej* oraz przez dodanie afiksów) są równoważne semantycznie, np. *najweselejszy* i *najbardziej wesoły* znaczą to samo. Dlatego traktujemy wszystkie wystąpienia stopnia wyższego i najwyższego jakby były analityczne. Np. *słoń szybszy niż strzała*, *mądrzejsza od Jana*



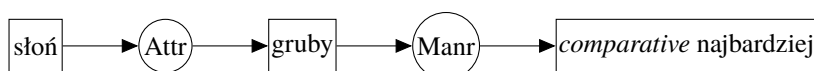
Comparative (Comp) oznacza argument porównawczy. Argument ten może pozostać niejawny, np. *szybszy słoń*



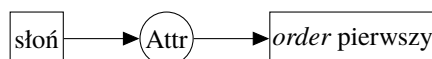
Liczność zadana przez konstrukcję porównawczą: *więcej niż kilkanaście słońi*



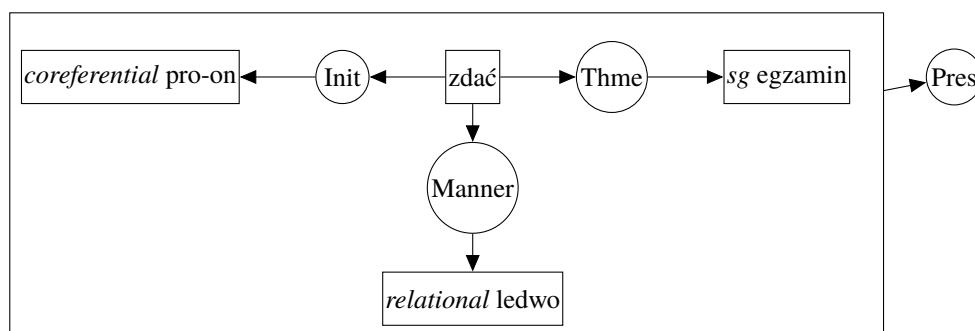
W przypadku stopnia najwyższego występuje niejawny argument porównawczy, np. *najgrubszy słoń*



Leksemy takie jak *pierwszy*, *drugi*, *ostatni*, *kolejny*, *jeszcze*, *już* mają niejawną, zależny od kontekstu argument porządkowy, np *pierwszy słoń*:



Znaczenie leksemów relacyjnych zależy w pewnym ustalonym zakresie od znaczenia nadrzędnika. Fakt ten jest reprezentowany przez argument relacyjny (*relational*), np. *ledwo* znaczące *prawie nie* w zdaniu *Ledwo zdał egzamin*



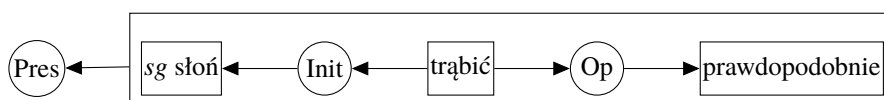
Inne leksemy relacyjne: *dużo*, *sporo*, *ledwo*, *niedużo*, *już*, *jeszcze*, *aż*.

### 2.3.13. Partykuły przyrematyczne i nieprawdomówne, podmiot epistemiczny

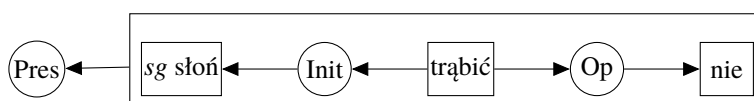
Możemy wyróżnić partykuły, których argument semantyczny jest zdeterminowany składniowo, *niejako*, *niemniej*, *oby*, *omalże*, *zwłaszcza*, *lada*, *ledwie*, *ledwo*, *nadto*, *aż*, *wprost*, *akurat*, *dosyć*, *dość*, *wszak*, *zaledwie*, *przecież*, *zaprawdę*, *zaiste*, *dopiero*, *jeszcze*, *już*, *doprawdy*, *ponadto* np *Jan lubi arbuzy zwłaszcza zimą*. Oraz partykuły, których argument semantyczny nie jest rozpoznawalny na poziomie opisu składniowego, w szczególności partykuły przyrematyczne modyfikujące akcentowaną, potencjalnie odległą część zdania (remat). Kublikami takimi są *otóż*, *notabene*, *owszem*, *prawda*, *skądinąd*, *również*, *także*, *też*, *znowu*, *znów*, *zarazem*, *tylko*, *niestety*, *jedynie*, *wręcz*, *naprawdę*, *oczywiście*, *naturalnie*, *wprawdzie*, *właśnie*, *nareszcie*, *wreszcie*, *bynajmniej*.

Niektóre spośród nich, gwarantują prawdziwość modyfikowanego modelu w modelu zewnętrznym, w którym same są interpretowane (partykuły faktywne): *Anna oczywiście zaśpiewała Marsyliankę*. Inne (partykuły nieprawdomówne) tworzą niefaktywny kontekst obejmujący fragment zdania, często jest nim remat.

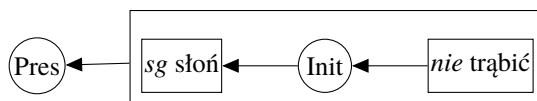
Partykuły, których zakres argumentu nie jest zadany przez morfoskładnię, będziemy reprezentować w sposób niedospecyfikowany podobnie jak kwantyfikatory. Np. *Słoń prawdopodobnie trąbi*



Negację traktujemy jako partykułę nieprawdomówną np. *Słoń nie trąbi*.

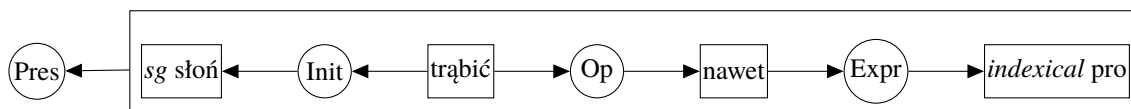


co możemy w skrócie zapisać jako



Podobnie zachowują się *chyba, także, również, też, nawet, pewnie, może, być może*.

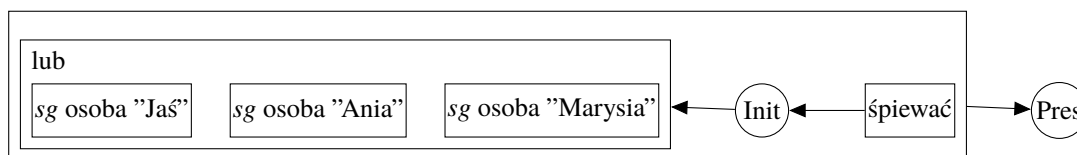
Dodatkowo, gdy wyrażenie odnosi do stanów mentalnych osoby niekoniecznie tożsamej z nadawcą (podmiot epistemiczny) dodajemy argument *pro* wskazujący tą osobę, zjawisko to występuje np. przy leksemach *chyba, pewnie, nawet, chociaż, ale, a*. *Słoń nawet trąbi*



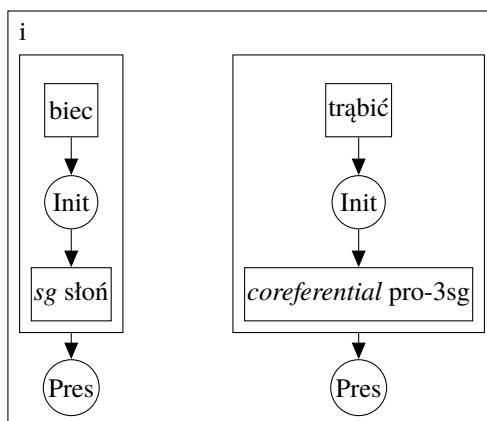
### 2.3.14. Koordynacja

Koordynację reprezentujemy jako kontekst etykietowany spójnikiem, zawierający listę koordynowanych obiektów. Graficznie kolejność elementów na liście jest wyrażona poprzez ich ułożenie od lewej do prawej.

Użycie kontekstu pozwala wyjść poza domyślną dla grafów pojęć zasadę łączenia poszczególnych węzłów za pomocą koniunkcji, np: w zdaniu *Jaś, Ania lub Marysia śpiewa*. kontekst *lub* zostanie zastąpiony w formule logicznej przez alternatywę.

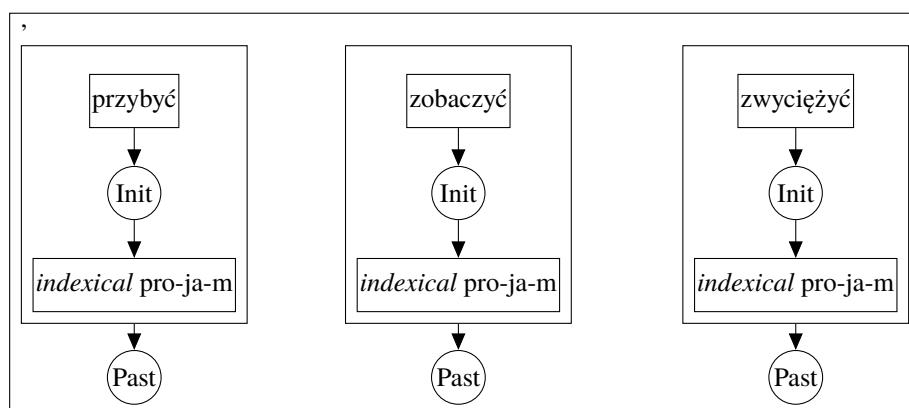


Współdzielone podrzędniki koordynacji reprezentujemy dodając koreferencyjny niemy zaimek *pro*. Dzięki temu, nie musimy rozstrzygać czy w zdaniu *Słoń biegnie i trąbi* trąbiącym jest *słoń*.

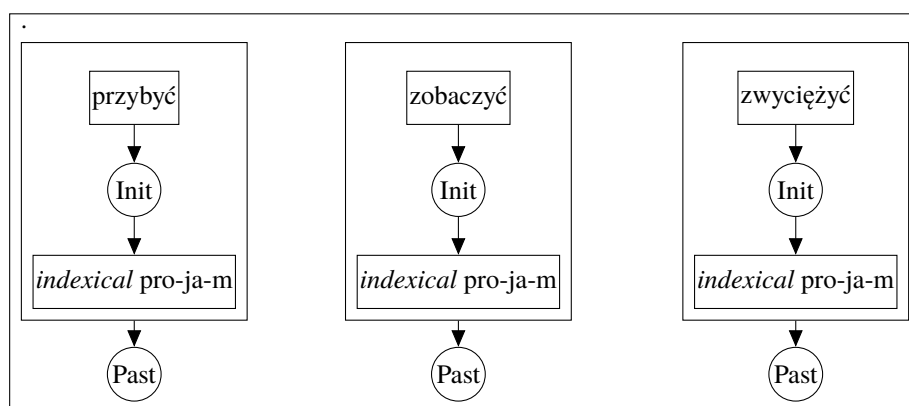




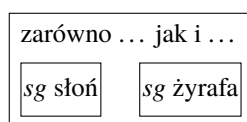
Sekwencje fraz rozdzielone przecinkami, bądź średnikami traktujemy tak jak frazy skoordynowane, np. *Przybyłem, zobaczyłem, zwyciężyłem*.



Sekwencje zdań również reprezentujemy jako koordynację, np. *Przybyłem. Zobaczyłem. Zwyciężyłem*.



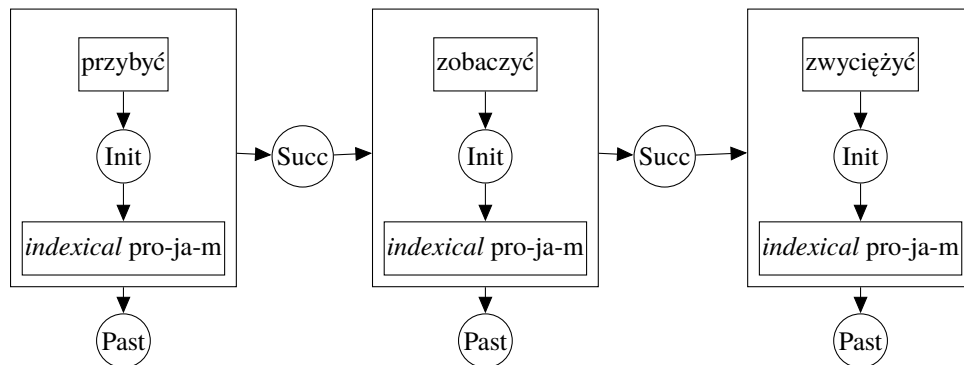
Leksemy spójników złożonych reprezentujemy umieszczając “...” w polach poszczególnych argumentów, np.: *zarówno słoń jak i żyrafa*. Argumenty występują na liście w kolejności takiej jak ich pola.



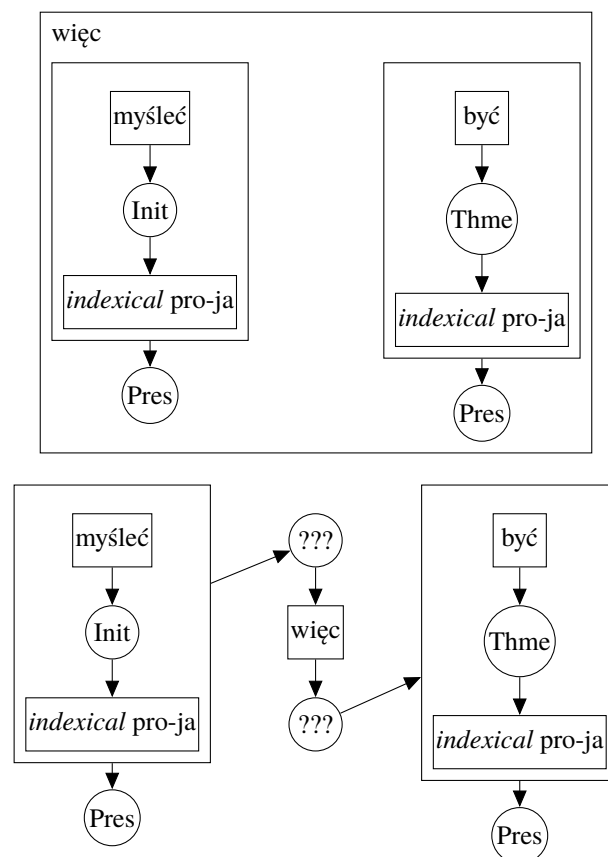
Spójniki współrzędne niosą znaczenie “i” oraz “lub” na poziomie zwykłej logiki uzupełnione o wkład metatekstowy. Te, które mają znaczenie “i” mogą zachowywać się addytywnie bądź multiplikatywnie. Powyższa reprezentacja nie wskazuje tych dwu znaczeń oraz pozostawia addytywność i multiplikatywność niedospecyfikowaną.

Poszczególne użycia spójnika “i” oraz innych spójników mających jego znaczenie można przetłumaczyć na reprezentację nie zawierających jawnego wystąpienia spójnika, np: kontekst

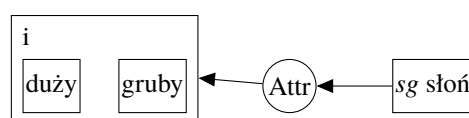
wprowadzany przez przecinek sygnalizujący następstwo zdarzeń w zdaniu *Przybyłem, zobaczyłem, zwyciężyłem*. możemy przetłumaczyć na

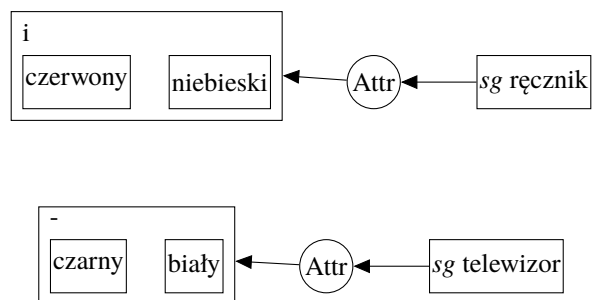


Zdanie *Myślę, więc jestem* z mającym wkład metatekstowy spójnikiem *więc* możemy zapisać na następujące sposoby:

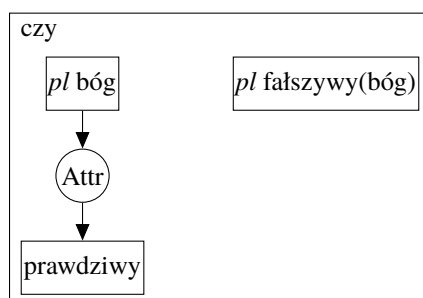


W ramach reprezentacji nie sygnalizujemy dystrybutywności i kolektywności koordynacji. Przykładowo dla zdania *Artur spał w Pile i Spale* nie zaznaczamy, że występują dwa miejsca i dwie czynności spania. Jeśli koordynowane podrzędniki wnoszą relacje wyciągamy je poza kontekst koordynacji np *duży i gruby stół*, *czerwony i niebieski ręcznik*, *czarno-biały telewizor*:

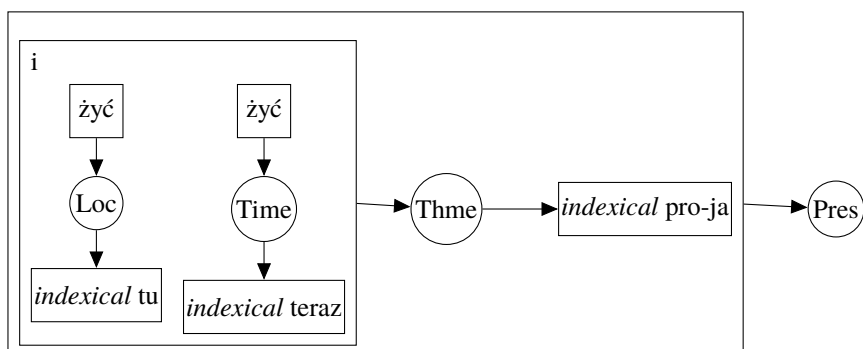




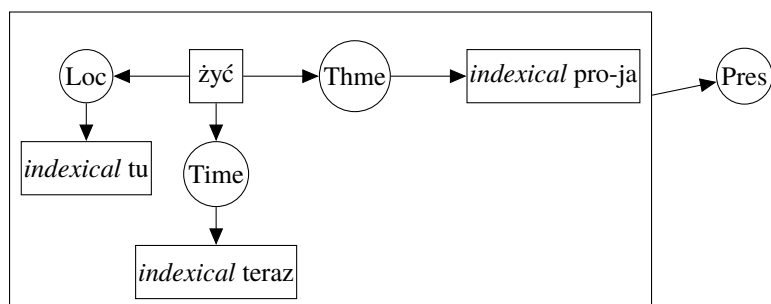
W przypadku, gdy człony koordynacji wnoszą różne relacje, nie są intersektywne lub jeden wnosi relację a drugi nie fraza zostaje przekształcona dystrybutywnie, np: *prawdziwi czy fałszywi bogowie*:



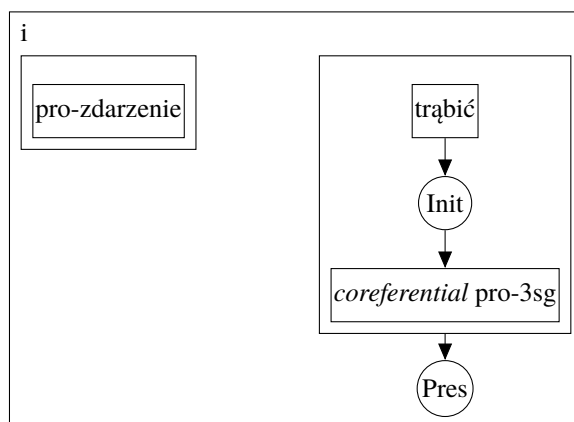
*Żyję tu i teraz*:



Jeśli dodatkowo uznamy, że mamy tu do czynienia z addytywnym *i* możemy usunąć koordynację:

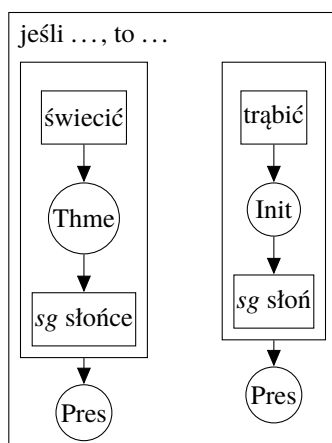


Zredukowane, mające tylko jeden argument spójniki, gramatycznie interpretowane jako ku-  
bliki np. *albo, ale, a, bo, chociaż, choć, i, czyli* reprezentujemy jako koordynację, której drugim  
argumentem jest *pro-zdarzenie* np. zdanie *I trąbi*:

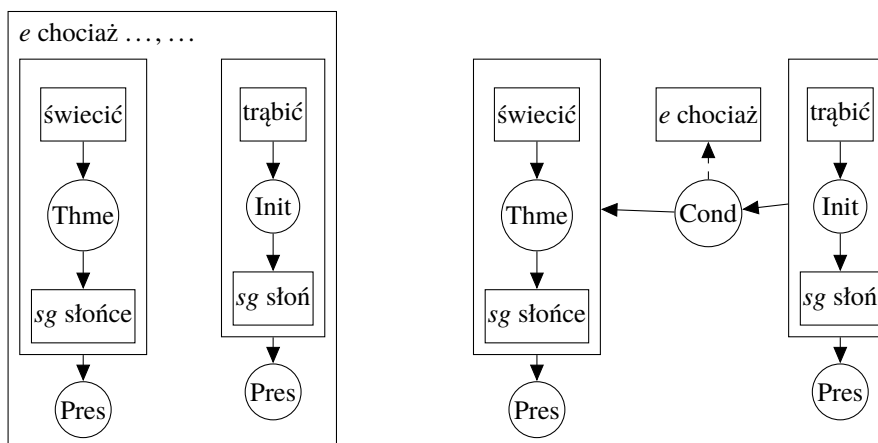


### 2.3.15. Spójniki podrzędne i zaimki względne

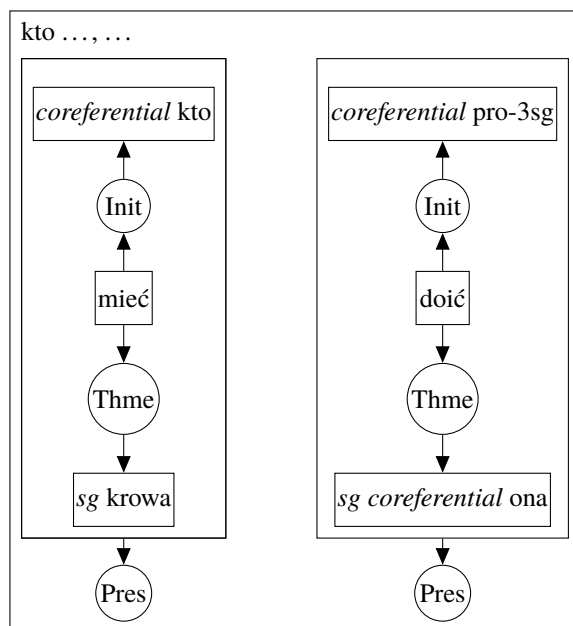
Spójnik *jeśli ..., to ...* użyty w znaczeniu logicznej implikacji reprezentujemy za pomocą kontekstu np. *Jeśli słońce świeci, to słoń trąbi*



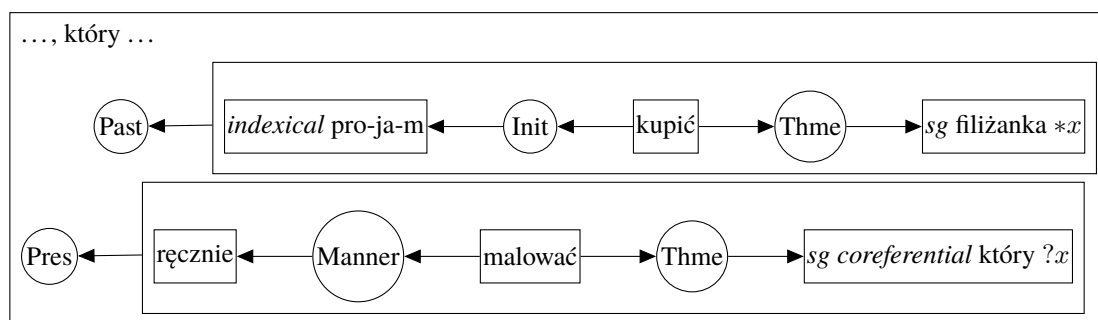
Zdania ze spójnikami podrzędnymi wnoszącymi znaczenie *i* możemy zapisać na dwa sposoby,  
np *Chociaż słońce świeci, słoń trąbi*:



Zaimki względne również mogą generować implikację. Dlatego będziemy reprezentować je jednocześnie za pomocą kontekstu i pojęcia np. *Kto ma krowę, doi ją*



*Kupiłem filiżankę, która jest ręcznie malowana.*

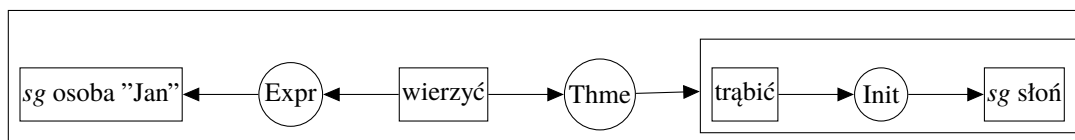


### 2.3.16. Dalsze zagadnienia dotyczące koordynacji

- zakres koordynacji a zakres kwantyfikacji *nie tylko ..., ale i ...*. Oceniano *nie tylko słowa, ale i głos*
- kompozycjonalność frazy koordynowanej, interpretacja “nie” w konstrukcjach “nie ... a ...”, “nie ... ani ...”.
- interakcja koordynacji i koreferencji przy spójniku *jeśli*, (zamiana kwantyfikatorów z egzystencjalnych na uniwersalne)
- koordynacja czasowników zanegowanych *Jan nie odpowiedział i wyszedł*
- Co zrobić przy koordynacji różnych kwantyfikatorów *Wszystkie słonie i niektóre samochody zatępiły*, czyli tak jak w zdaniu Hintikka (Za: Jakub Szymanik “PROBLEMY Z FORMĄ LOGICZNĄ”): *Pewien krewniak każdego wieśniaka i pewien krewniak każdego mieszczucha nienawidzą się nawzajem; Większość krewniaków każdego wieśniaka i większość krewniaków każdego mieszczucha nienawidzi się nawzajem.*
- Jak wyrazić ‘wkład meta językowy’, czyli odróżnić *Jan tańczy i śpiewa* od *Jan tańczy, ale i śpiewa*; jak wyrazić sekwencyjne ‘i’ *Jan wstał i poszedł*

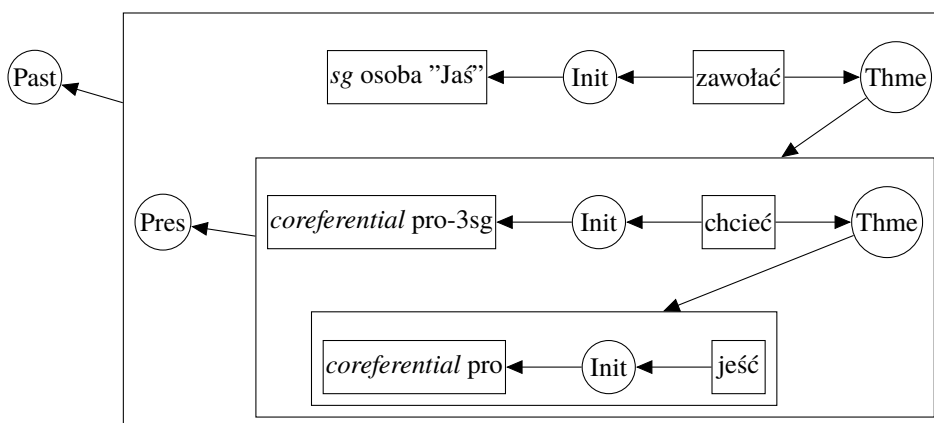
### 2.3.17. Wewnętrzne modele

Zdanie, które jest przedmiotem przekonań, pragnień, komunikacji nie musi być obiektywnie prawdziwe. Umieszczamy je w pudełku oznaczającym, że jego prawdziwość należy określać ze względu na subiektywny model świata, np *Jan wierzy, że słoń trąbi*.

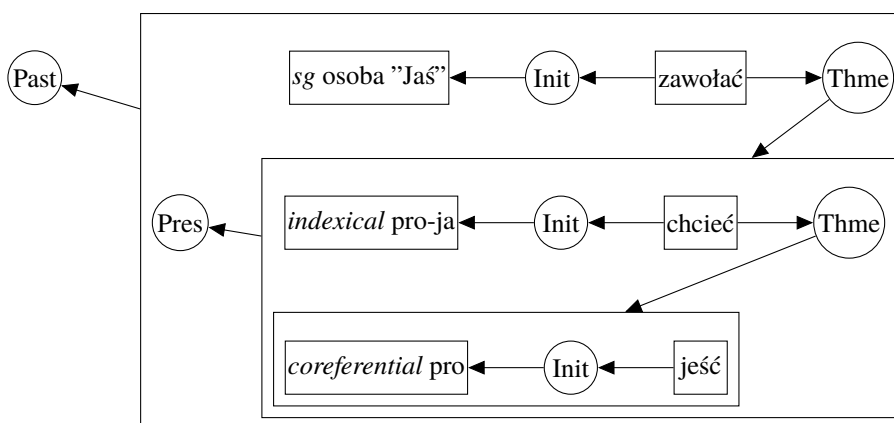


Podobnie jak przyimki, spójniki podrzędne dzielimy na semantyczne i niesemantyczne, zgodnie z tym, co stanowi o nich *Walenty*. Spójniki niesemantyczne nie są odzwierciedlane w grafach semantycznych.

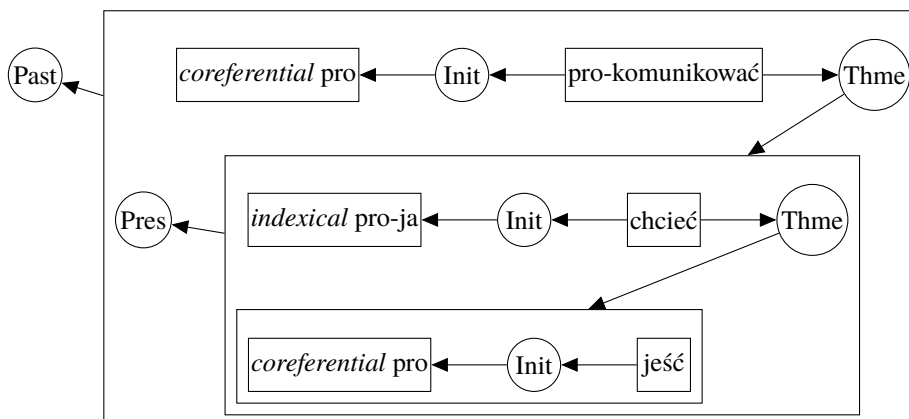
Mowa niezależna i zależna są interpretowane w sposób maksymalnie zbliżony, np. *Jaś zawałał, że chce jeść*



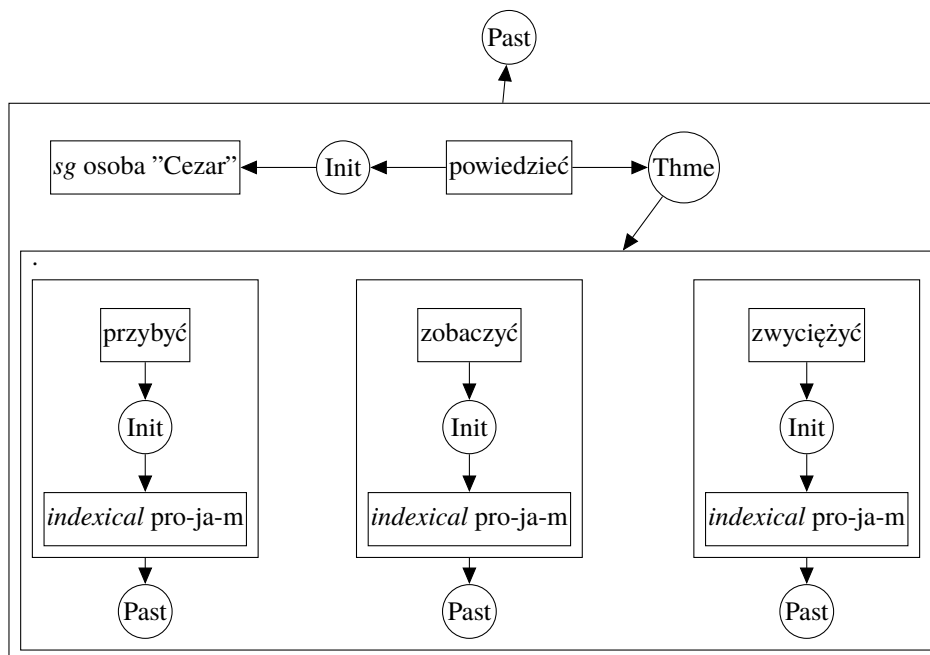
oraz - *Chcę jeść* - zawałał Jaś



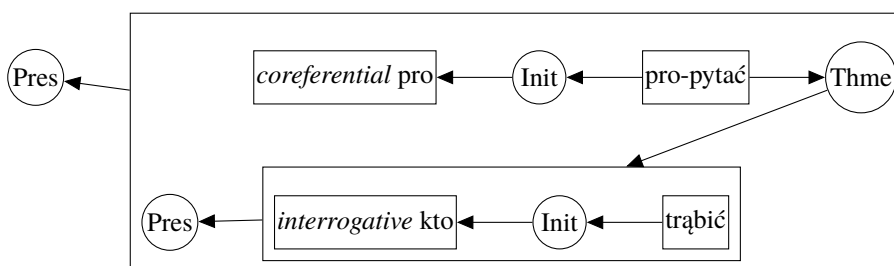
Kiedy w mowie niezależnej mówca nie jest wskazany dodajemy go oraz zdarzenie komunikowania do reprezentacji logicznej, np. - *Chcę jeść*

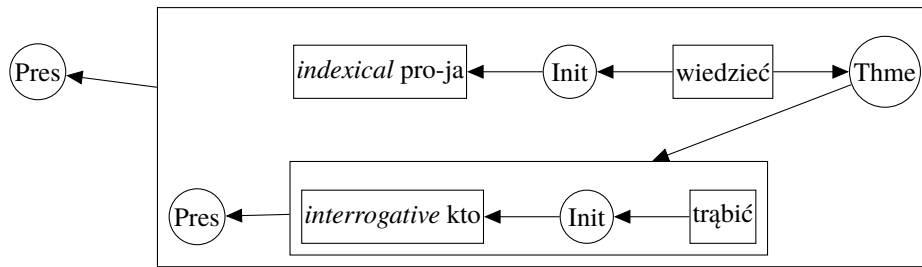


Mowę niezależną składającą się z wielu zdań reprezentujemy jako koordynację, np. - *Przybyłem - powiedział Cezar. - Zobaczyłem. Zwyciężyłem.*



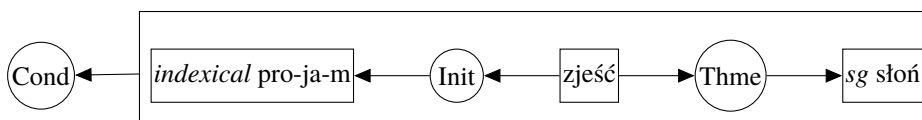
Pytania zadane wprost oraz wyrażone w mowie zależnej interpretujemy umieszczając je w wewnętrznym kontekście oraz nadając zaimkowi pytającemu symbol *interrogative* , np: - *Kto trąbi?*, *Wiem, kto trąbi.*





Faktywność możemy uwzględnić etykietując konteksty relacją wskazującą, czy treść kontekstu musi być prawdziwa w kontekście zewnętrznym.

Tryb przypuszczający zaznacza, że sytuacja nie ma miejsca w rzeczywistości, czyli czyni ją niefaktywną. Oznaczamy to jednoargumentową relacją Cond. *Zjadłbym słonia.*

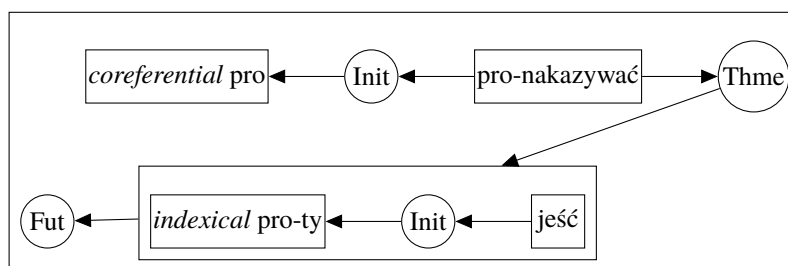


W przyszłości można zastąpić relację Cond poszczególnymi użyciami trybu przypuszczającego, np. wyrażeniem intencji.

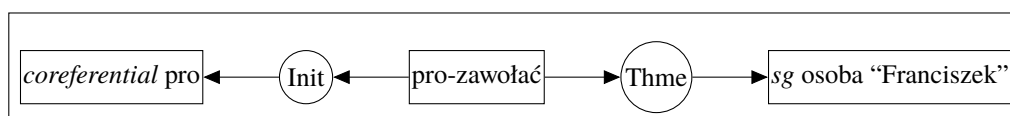
*Stanisław zawołał Franciszka i powiedział, żeby jadł*

### 2.3.18. Niedenotatywne funkcje języka

Semantykę trybu rozkazującego definiujemy przez sprowadzenie do mowy zależnej, np. - *Jedź!* przetłumaczmy na *Nakazał mu jeść*

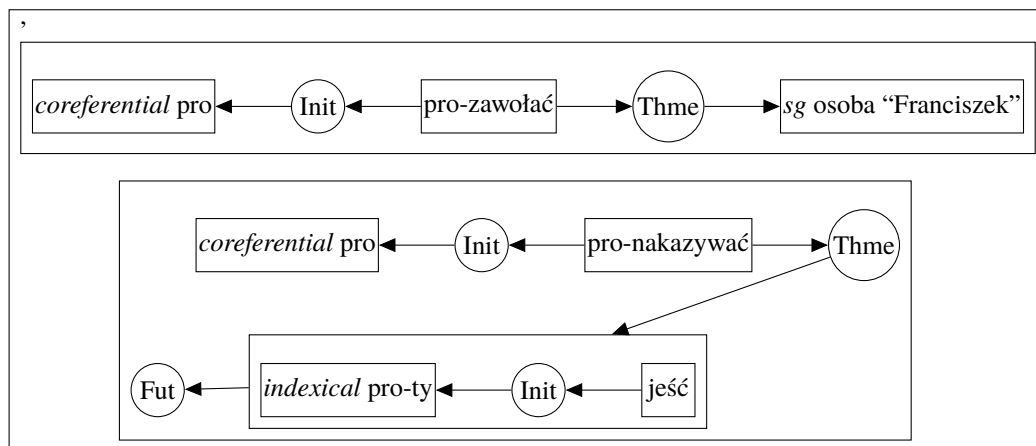


Podobnie semantyka wołacza jest określona przez sprowadzenie do mowy zależnej, np. - *Franciszku!* przetłumaczmy na *Zawołał Franciszka*

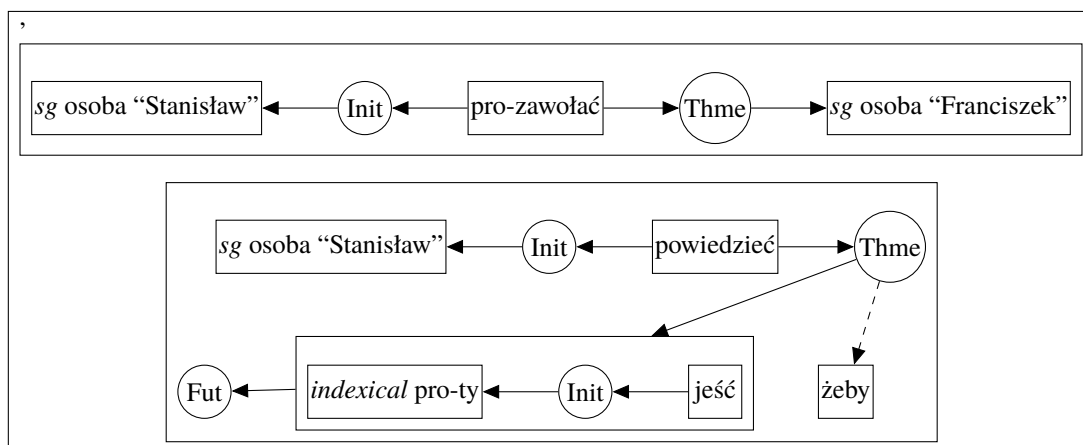




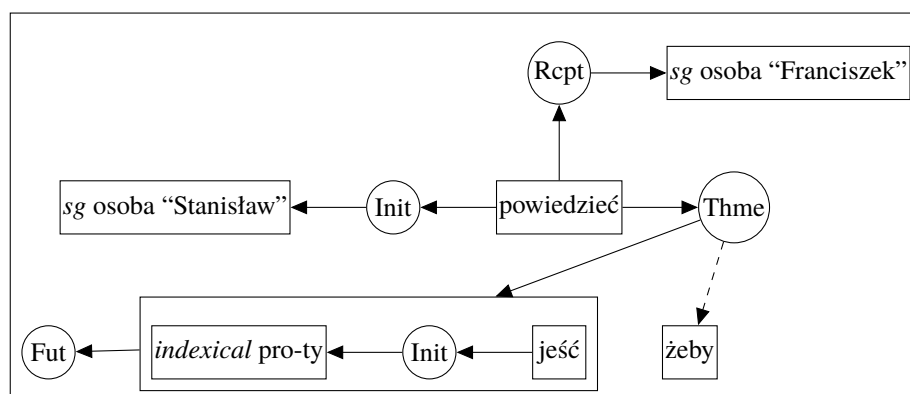
Połączenie wołacza i trybu rozkazującego daje następującą reprezentację - *Franciszku, jedz!*



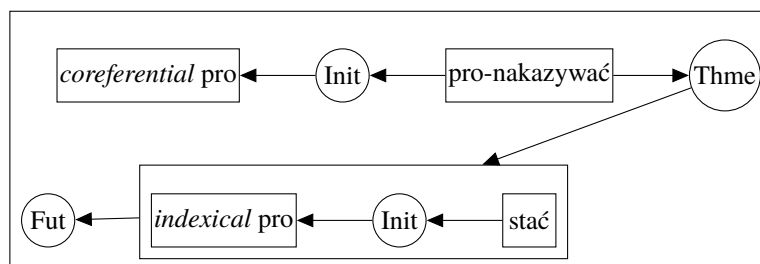
- *Franciszku, jedz!* - powiedział Stanisław. tłumaczymy na Stanisław zawołał Franciszka i powiedział, żeby jadł



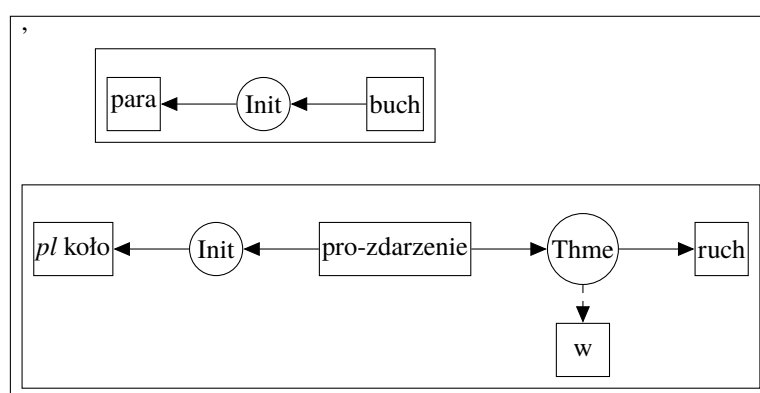
Alternatywnie można by powyższe zdanie zinterpretować jako *Stanisław powiedział Franciszkowi, żeby jadł*



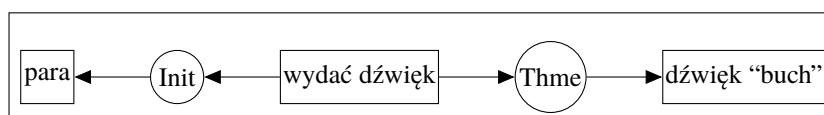
Zdania w bezokoliczniku oznaczające polecenia reprezentujemy analogicznie do trybu rozkazującego, np. - *Stać!*



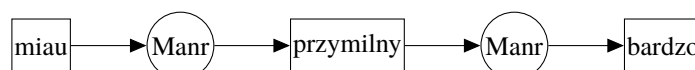
Wykrzykniki odnoszą do zdarzeń, dlatego reprezentujemy je analogicznie do czasowników i nadajemy kontekst sytuacyjny, np *Para buch, koła w ruch*



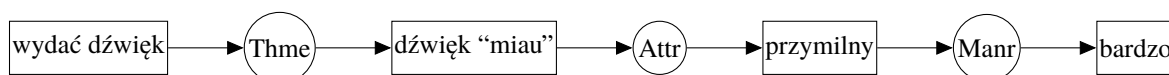
Znaczenie leksemu *buch* możemy zdefiniować jako wydać dźwięk “*buch*”. Pozwoli nam to zapisać frazę *Para buch* jako *Para wydała dźwięk “buch”* i reprezentować ją jako:



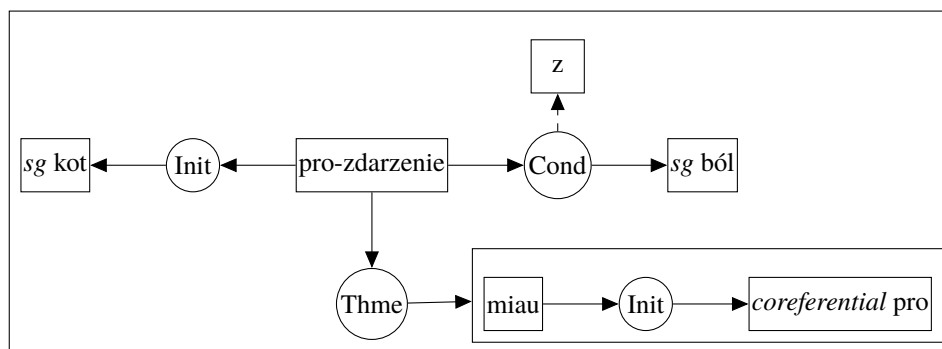
W powyższym grafie wydać dźwięk jest sensem ze Słownosieci, a dźwięk “*buch*” wygenerowanym dźwiękiem. Analogicznie możemy reprezentować pozostałe onomatopeje (np *miau*, *brzdęk*, *brr*, *be*, *baś*, *brrum*, *bum*, *cip-cip*). Zakładamy, że dla każdego zdarzenia sygnalizowanego przez onomatopeje istnieje jego Initiator. Frazę *bardzo przymilne miau* zapiszemy jako



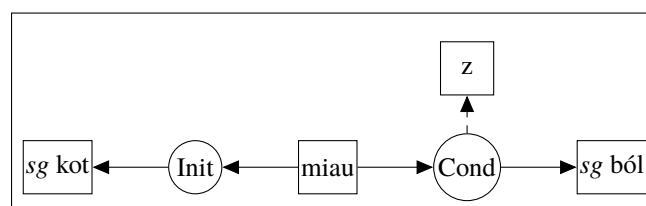
co jest równoważne:



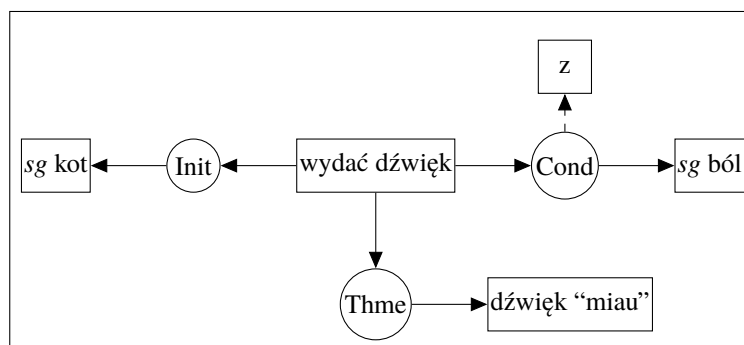
Z kolei zdanie *A kot z bólu: miau* zinterpretujemy tak jak mowę niezależną:



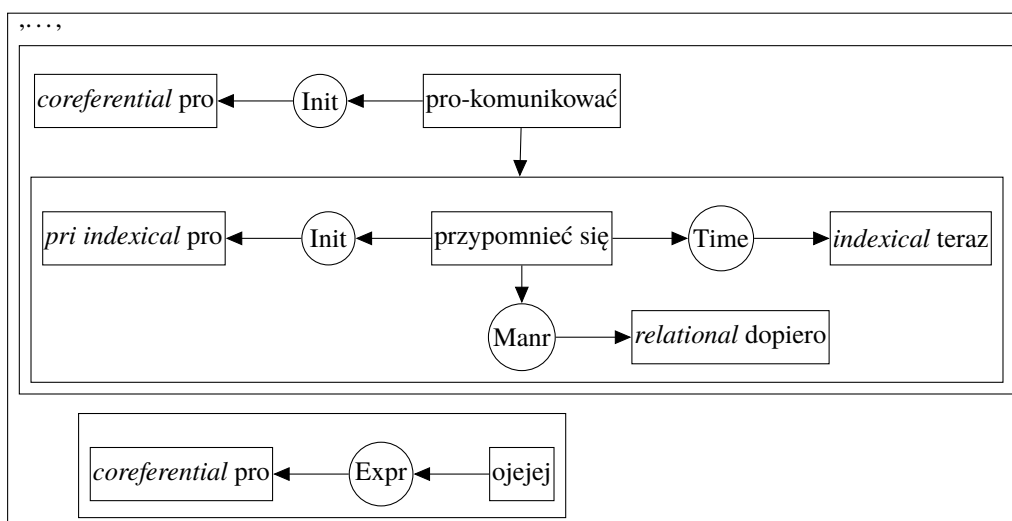
Następnie po rozwiązaniu koreferencji pomiędzy *pro-zdarzenie* i *miau* otrzymamy



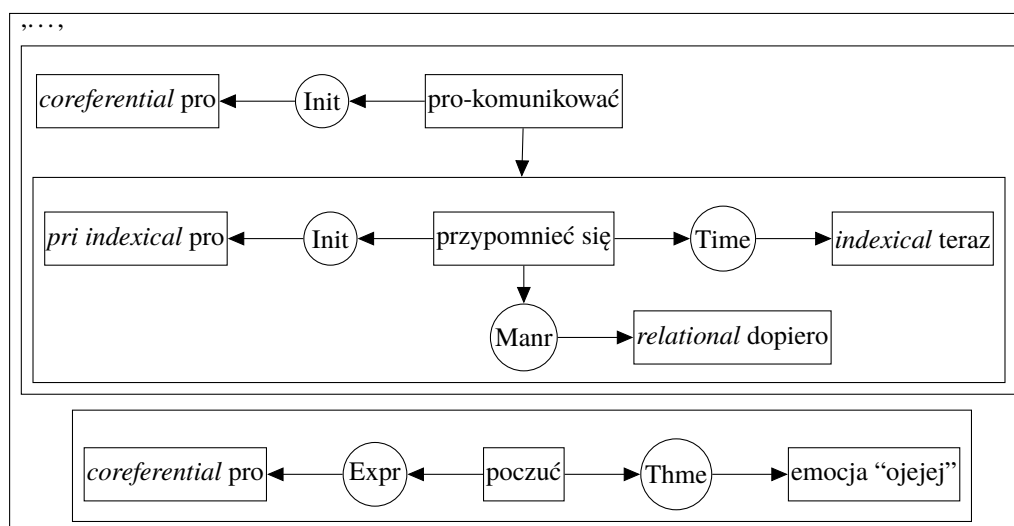
a po rozwinięciu *miau*:



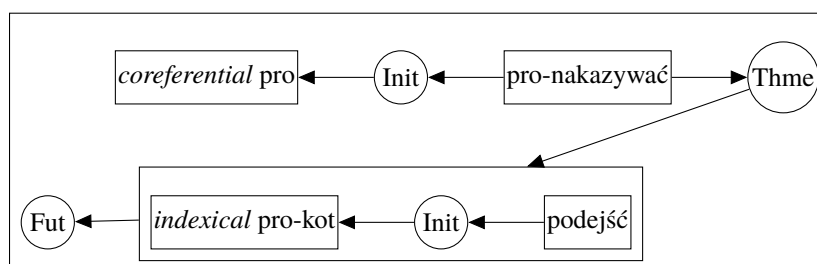
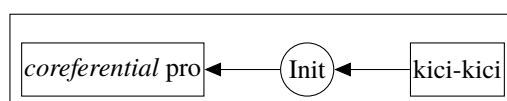
Z wykrzyknikami wyrażającymi uczucia i emocje np *hurra*, *fuj*, *ojej* związane są osoby, które ich doświadczają, np *Teraz mi się dopiero, ojej, przypomniało*



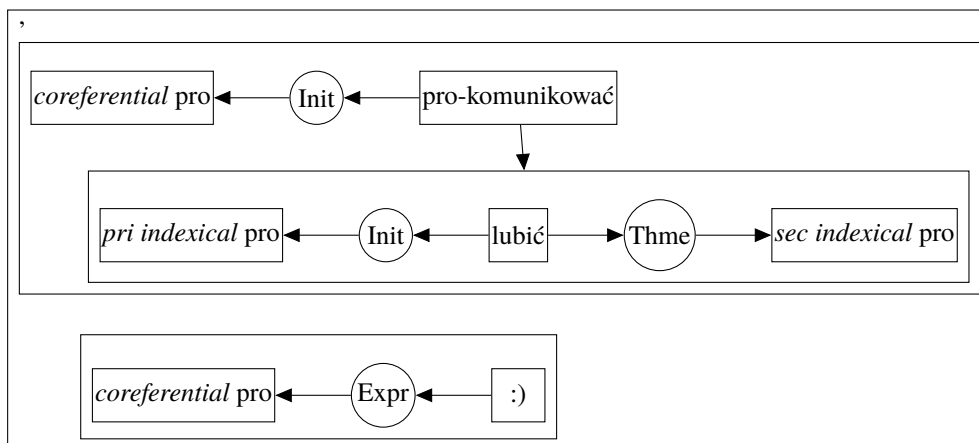
W powyższym zdaniu kontekst „...”, oznacza wtrącenie. Z uwagi na metatekstowy charakter *ojejej* trzeba jawnie wskazać, że *Teraz mi się dopiero przypomniało* jest treścią komunikatu. Możemy zinterpretować *ojejej* jako *poczuć emocję o nazwie “ojejej”*. Emocję “ojejej” można potem przetłumaczyć na np. *zaskoczenie*.



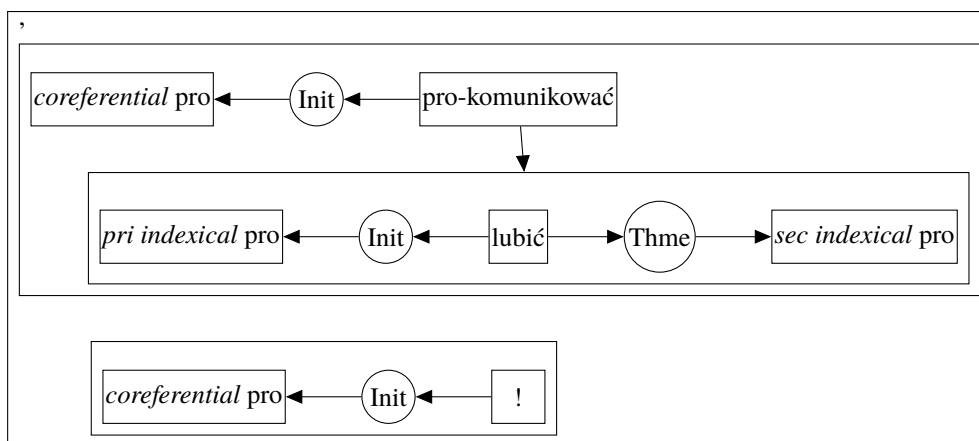
Wykrzykniki wyrażające wolę mówiącego (powitania *cześć*, kontaktu ze zwierzętami np. *kici-kici*), apelatywne *brawo*, *uwaga*, *precz*, *huzia*, *jazda*, *wara*, *won*, *a kysz*, *amen*, parentetyczne *halo*, *ałła*, *cholera*, potwierdzenia *spoko*, *tak* traktujemy jako zdarzenia, które mają swojego Initiatora. Wykrzykniki te mają charakter trybu rozkazującego, a ich Initiator to osoba, która wyraża wolę, np. - *Kici-kici!* można przetłumaczyć na - *Kocie podejdź!* i dalej na *Nakazał kotu podejść*.



Emotikony reprezentujemy tak jak wykrzykniki emotywne, *Lubię Cię :)*



Znak wykrzyknienia reprezentujemy podobnie jak emotikony, *Lubię Cię!*



### 3. Zawartość towarzyszącego pliku ZIP

Załączony plik zawiera opisaną w rozdziale 2.1 bazę jednostek leksykalnych.