

IZABELA GATKOWSKA

Uniwersytet Jagielloński, Kraków

Empiryczna sieć powiązań leksykalnych

1. Wprowadzenie

Potrzeba reprezentowania znaczeń jednostek języka za pomocą postaci sieci semantycznej ujawniła się w badaniach nad sztuczną inteligencją rozpoczętych na przełomie lat 60. i 70. XX wieku. Badacze poszukujący algorytmów rozumienia tekstu (Schank 1975) doszli do wniosku, że algorytm rozumiejący tekst lub rozpoznający obraz musi mieć dostęp do informacji, która nie występuje bezpośrednio w tekście czy obrazie. W przeciwnym przypadku algorytm nie potrafi poprawnie zinterpretować tak naturalnych w ludzkiej komunikacji sekwencji wypowiedzi: *Wczoraj dostałam spaniela — Cieszę się, że masz psa*. Chcąc umożliwić algorytmowi dostęp do informacji nieobecnej bezpośrednio w wypowiedzi, można posłużyć się siecią zbudowaną ze znaczeń i zależności (relacji) zachodzących pomiędzy znaczeniami. Algorytm, który zlokalizuje jednostkę analizowanej wypowiedzi w tak rozumianej sieci, zyskuje dostęp do wszystkich znaczeń, które wiążą się z analizowaną jednostką, a więc uzyska wiedzę, że np. *spaniel* łączy się bezpośrednio ze znaczeniem *pies* oraz pośrednio ze znaczeniem *zwierzę*.

Sieć semantyczną można zbudować ręcznie, analizując odpowiednio liczny zbiór zdań reprezentatywnych dla każdego z umieszczonych w sieci znaczeń — tak zbudowano i buduje się nadal wychodzące z różnych założeń sieci łączące znaczenia za pomocą relacji semantycznych i pragmatycznych np. WordNet (Miller i wsp. 1998), FrameNet (Fillmore 2000) czy Cyc (tzw. ontologia). Jednak różne założenia leżące u podstaw wymienionych sieci spowodowały, że sieci różnią się nie tylko zbiorem relacji semantycznych, np. WordNet ma ich kilkanaście, a Cyc ponad 20000 (Pohl 2009), ale też pojmowaniem obiektów, między którymi relacje zachodzą. WordNet i Cyc są słownikami pojęć i jednostek słownika, a FrameNet jest słownikiem struktur semantycznych, za pomocą których można opisać znaczenie jednostki leksykalnej w zdaniu.

Sieć reprezentującą znaczenia jednostek języka można także zbudować empirycznie, pozyskując naturalny zbiór powiązań pomiędzy znaczeniami jednostek słownika za pomocą eksperymentu psycholingwistycznego, np. poprzez test swobodnych skojarzeń słownych — w ten sposób zbudowano np. *Edinburgh Associative Thesaurus* (EAT, Kiss, Armstrong, Milroy, Piper 1973). Jednak sam eksperyment dostarcza tylko danych o powiązaniach, a sieć zyskuje przydatność wówczas, gdy dokonamy interpretacji powiązań pomiędzy znaczeniami. Jako pierwszy zwrócił na to uwagę H.H. Clark (1970), wskazując możliwość interpretacji skojarzeń za pomocą metod znanych językoznawstwu. Jednak praca Clarka była studium skoncentrowanym na objaśnianiu mechanizmów powstawania skojarzeń leksykalnych i nie analizowała konkretnej sieci. Chcąc opisać sposób reprezentowania znaczeń jednostek leksykalnych w sieci zbudowanej eksperymentalnie, musimy poklasyfikować ujawniające się w wyniku eksperymentu zależności pomiędzy znaczeniami, gdyż zakładamy, że nie tylko fakt powiązania z innymi znaczeniami, ale także jakość powiązań definiuje znaczenie jednostki leksykalnej w sieci. Zdefiniowane w ten sposób ‘naturalne’ znaczenie jednostki leksykalnej można porównać ze znaczeniem tej samej jednostki słownika zbudowanym przez leksykografów. Niniejszy tekst przedstawi budowę empirycznej sieci powiązań leksykalnych, próbę klasyfikacji występujących w sieci powiązań pomiędzy znaczeniami oraz wstępną charakterystykę definiowanego przez sieć znaczenia rzeczownika.

2. Empiryczna sieć powiązań leksykalnych

2.1. Podstawa eksperymentalna — test swobodnych skojarzeń słownych

Diagnostyczną przydatność analizy powiązań leksykalnych zauważyli amerykańscy psychiatrzy G. Kent i A.J. Rosanoff, którzy w roku 1910 opracowali i przeprowadzili test swobodnych skojarzeń słownych. W trakcie testu przebadano 1000 zdrowych osób o różnym wykształceniu i zawodach, prosząc badanych o podanie pierwszego przychodzącego na myśl skojarzenia wywołanego przez wyraz-bodziec. Zbadano w ten sposób skojarzenia dla 100 wyrazów-bodźców (głównie rzeczowników i przymiotników). Lista Kent-Rosanoffa została przetłumaczona na kilka języków, dla których powtórzono eksperyment, umożliwiając w ten sposób badania porównawcze. Badania skojarzeń słownych kontynuowano: Palermo, Jenkins (1964), Postman, Keppel (1970), Kiss, Armstrong, Milroy, Piper (1973), Moss, Older (1996), Nelson, McEvoy, Schreiber (1998), a powtarzalność wyników pozwalała zmniejszać liczbę badanych osób przy jednoczesnym zwiększaniu liczby wyrazów-bodźców, np. 500 badanych i 200 wyrazów (Palermo, Jenkins 1964) czy 100 badanych i 8400 wyrazów (Kiss, Armstrong, Milroy, Piper 1973). Zwiększanie użytej w testach liczby wyrazów sprawiło, że wyniki testów stały się materiałem ciekawym dla językoznawstwa, a prowadzony przez wiele lat eksperyment Kiss, Armstrong, Milroy, Piper (1973) pozwolił nawet na zbudowanie tezauryusa dla języka angielskiego —

Edinburgh Associative Thesaurus (EAT). Eksperymenty prowadzące do budowy sieci skojarzeń leksykalnych prowadzi się także współcześnie, np. De Deyne, Storms (2008). Po roku 2000 do badań nad swobodnymi skojarzeniami słownymi włączyła się też lingwistyka komputerowa, poszukując naturalnego modelu łączliwości znaczeń, np. Melinger, Weber (2006), S. Schulte im Walde, S. Borgwaldt, R. Jauch (2012).

W Polsce badania swobodnych skojarzeń słownych przeprowadziła w latach 1964/1965 I. Kurcz. Podstawą eksperymentu I. Kurcz była pełna (100 wyrazów) lista Kent-Rosanoffa, w tłumaczeniu na język polski. W teście uczestniczyło 1000 studentów Uniwersytetu Warszawskiego i Politechniki Warszawskiej. Kompletne wyniki oraz krótki artykuł wprowadzający opublikowano w VII tomie „Studiów Psychologicznych” z 1967r. Autorski eksperyment (Gatkowska 2014), który pozwolił zbudować sieć powiązań leksykalnych opisaną w niniejszym tekście, nawiązuje bezpośrednio do eksperymentu Kurcz, zob. 4.1.

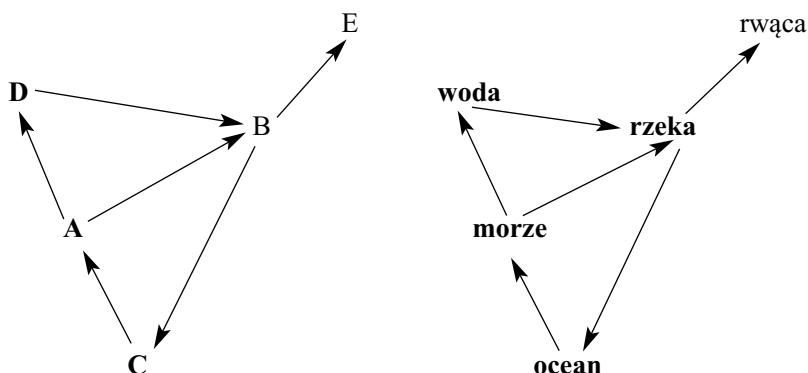
2.2. Charakterystyka empirycznej sieci leksykalnej

Wynikiem testu swobodnych skojarzeń słownych jest tzw. lista skojarzeniowa, tj. zbiór wyrazów stanowiących odpowiedzi związane z konkretnym bodźcem. Jednak w roku 1973 brytyjscy badacze (Kiss, Armstrong, Milroy, Piper 1973) zaobserwowali, że jeśli przebadamy odpowiednio liczny zbiór wyrazów bodźców, prowadząc eksperyment cyklicznie, tj. używając odpowiedzi uzyskanych w cyklu pierwszym jako bodźców w cyklu następnym, to bodźce i zbiory odpowiedzi tworzą sieć. Uzyskana eksperymentalnie sieć leksykalna to zbiór jednostek leksykalnych (wyrazów, np. *dom*, *pies* i jednostek wielosegmentowych, np. *panna młoda*, *czarna owca*) oraz zbiór powiązań między jednostkami leksykalnymi.

Sieć leksykalna. Jeśli przyjmiemy, że istnieją wyrazy A, B, C, D, E i skojarzenia, tj. pary bodziec — odpowiedź

A → D	np. morze → woda
A → B	np. morze → rzeka
B → C	np. rzeka → ocean
C → A	np. ocean → morze
D → B	np. woda → rzeka
B → E	np. rzeka → rwąca

To zbiór par bodziec — odpowiedź tworzy sieć o postaci



Wyrazy są **węzłami** sieci, a strzałki oznaczają **powiązanie** i jego kierunek — zawsze od bodźca do odpowiedzi. Powiązanie o tak zdefiniowanym kierunku zachodzące pomiędzy dwoma węzłami sieci np. $A \rightarrow B$, jest dla węzła A powiązaniem wychodzącym a dla węzła B powiązaniem przychodzącym. Tak więc w naszej przykładowej sieci węzeł *morze* ma dwa powiązania wychodzące: $\text{morze} \rightarrow \text{woda}$, $\text{morze} \rightarrow \text{rzeka}$ oraz jedno powiązanie przychodzące $\text{ocean} \rightarrow \text{morze}$. Sekwencja powiązań tworzy w sieci **ścieżkę**, np. $A \rightarrow B \rightarrow C$, gdzie A jest węzłem początkowym, a C węzłem końcowym ścieżki. Długość ścieżki określana poprzez liczbę powiązań w ścieżce nie ma formalnych ograniczeń. W sieci leksykalnej wyróżniamy dwa rodzaje ścieżek: zamknięte i otwarte.

Ścieżki zamknięte to takie, w których wyraz jest jednocześnie węzłem początkowym i końcowym węzłem ścieżki, np. $A \rightarrow D \rightarrow B \rightarrow C \rightarrow A$, czyli $\text{morze} \rightarrow \text{woda} \rightarrow \text{rzeka} \rightarrow \text{ocean} \rightarrow \text{morze}$. Konkretny węzeł może rozpoczynać wiele ścieżek zamkniętych; w naszym przykładzie będą to sekwencje powiązań: $A \rightarrow D \rightarrow B \rightarrow C \rightarrow A$, czyli $\text{morze} \rightarrow \text{woda} \rightarrow \text{rzeka} \rightarrow \text{ocean} \rightarrow \text{morze}$ oraz $A \rightarrow B \rightarrow C \rightarrow A$, czyli $\text{morze} \rightarrow \text{rzeka} \rightarrow \text{ocean} \rightarrow \text{morze}$.

Ścieżki otwarte to ścieżki, w których węzeł początkowy ścieżki jest różny od węzła końcowego; w naszym przykładzie będą to ścieżki np. $A \rightarrow D \rightarrow B \rightarrow E$, czyli $\text{morze} \rightarrow \text{woda} \rightarrow \text{rzeka} \rightarrow \text{rwąca}$ oraz np. $A \rightarrow B \rightarrow E$, czyli $\text{morze} \rightarrow \text{rzeka} \rightarrow \text{rwąca}$.

Każdy wyraz, który znajduje się w sieci i jest powiązany z innymi wyrazami, tworzy **węzeł leksykalny**. Istnieją węzły leksykalne, w których wyraz węzłowy ma powiązania wychodzące i przychodzące (wyraz mający ten typ węzła to wyraz, który był bodźcem), oraz węzły, w których wyraz węzłowy jest charakteryzowany tylko przez powiązania przychodzące, gdyż wyraz węzłowy nie był bodźcem tylko odpowiedzią do wyrazów bodźców.

Statystyczna analiza sieci leksykalnych (Steyvers, Tenenbaum 2005) pokazuje, że sieć leksykalna i słowniki semantyczne (Roget's Thesaurus, WordNet) mają wspólne cechy strukturalne i że cechy te odróżniają sieć powiązań leksykalnych od analizowa-

nych za pomocą tych samych metod statycznych sieci będących przedmiotem badań innych niż językoznawstwo dyscyplin naukowych. Tak więc analiza statystyczna potwierdza językoznawczą intuicję i można powiedzieć, że **wszystkie przychodzące i wychodzące powiązania wyrazu węzłowego oraz wszystkie ścieżki zamknięte i otwarte, w których wyraz węzłowy jest węzłem początkowym, definiują znaczenie wyrazu węzłowego. Jednak analiza językoznawcza węzła leksykalnego empirycznej sieci leksykalnej musi wykroczyć poza analizę statystyczną, powinna bowiem uwzględniać fakt, że uzyskane empirycznie powiązania pomiędzy znaczeniami wyrazów występujących w sieci mają właściwą sobie siłę i jakość.**

Siła i kierunek powiązania. Siła powiązania to stosunek liczby odpowiedzi łączących bodziec A z odpowiedzią B do liczby wszystkich odpowiedzi uzyskanych dla bodźca A. Przykładowo, gdy liczba wszystkich odpowiedzi dla bodźca A wynosi 95, z czego 49 przypada na odpowiedź B, to siła powiązania A–B wynosi 52% (49/95 po zaokrągleniu). Siła powiązania jest związana z jego kierunkiem, co znaczy, że powiązania wychodzące i przychodzące mają własną (odrębną) siłę powiązania. Przykładowo siła powiązań wychodzących i przychodzących przymiotnika *kiszona* i rzeczownika *kapusta* jest różna. Gdy bodźcem jest *kiszona*, tj. *kiszona* → *kapusta* siła powiązania wynosi 83,18%¹ (737 odpowiedzi *kapusta* z 886 odpowiedzi), zaś gdy bodźcem jest *kapusta*, tj. *kapusta* → *kiszona* siła powiązania wynosi już tylko 48,85% (426/886). Można więc powiedzieć, że zależność związana z kierunkiem siły powiązania określa stosunek pomiędzy znaczeniami — dla znaczenia *kiszona* połączenie ze znaczeniem *kapusta* jest najważniejsze, natomiast dla znaczenia *kapusta* powiązanie z *kiszona* jest mniej ważne, gdyż *kiszona* to tylko jedna z właściwości przysługujących *kapuście*, występująca obok: *warzywo* 4,47, *głowa* 3,9, *zielona* 2,87 itp. Jednak siła powiązania może być tylko wartością pomocniczą. Opis znaczenia węzła leksykalnego sieci możemy uzyskać, gdy uwzględnimy jakość powiązania.

Kierunek i jakość powiązania. Jakość powiązania to zależność semantyczna zachodząca pomiędzy znaczeniami, np. powiązanie *ręka* — *ciało*, *noga* — *krzesło* jest powiązaniem pomiędzy częścią a całością. Przyjmujemy, że zależność semantyczna ma swój kierunek, w naszym przykładzie od części do całości — w tym przypadku kierunek zależności semantycznej jest zgodny z kierunkiem skojarzenia bodziec — odpowiedź. Podobną zgodność kierunku skojarzenia i zależności semantycznej obserwujemy w parach *krzesło* — *noga*, *ciało* — *noga*, gdzie zależność prowadzi od całości do części.

Są jednak liczne powiązania semantyczne, których kierunek jest niezależny od kierunku skojarzenia, np. *ryba* — *morze* i *morze* — *ryba* wyrażają zależność pomiędzy zwierzęciem a miejscem, w którym zwierzę żyje; zależność ta prowadzi od zwierzęcia do miejsca. Nazwijmy tę zależność *miejsce_charakterystyczne* i przyjmijmy, że nazwy

¹ W dalszej części tekstu wartość liczbowa siły powiązania oznacza procent odpowiedzi.

jakości powiązań zachodzących pomiędzy węzłami sieci będziemy nazywać **wyznacznikami** jakości powiązania.

Istnieją wreszcie w sieci powiązania, np. *ręka — noga, noga — oparcie*, którym nie da się przyporządkować zależności semantycznej, a jedyną wspólną własnością wymienionych znaczeń jest to, że stanowią części określonej całości, tj. *ciała* i *krzesła*. Powiązania takie objaśnia ścieżka prowadząca od części-bodźca do części-odpowiedzi poprzez całość, a więc *ręka — ciało — noga, noga — krzesło — oparcie*. Podobne — choć mające innych mechanizm — zjawisko obserwujemy w powiązaniach *baranina — wełna, owca — sweter*; tu w każdym przypadku ścieżka ma inną strukturę, odpowiednio: *baranina — owca — runo — wełna — owca — runo — wełna — sweter*. Powiązania, w których zależność pomiędzy węzłami można objaśnić tylko poprzez ścieżkę w sieci, będziemy nazywać powiązaniem odległymi.

3. Jakościowa klasyfikacja powiązań

Klasyfikację jakościową powiązań prowadzoną z językoznawczego punktu widzenia jako pierwszy zaproponował Herbert H. Clark (1970). Zgodnie z psychologiczną tradycją badań nad skojarzeniami Clark przyjmuje, że skojarzenia dzielą się na semantyczne, tj. takie, w których pomiędzy bodźcem a odpowiedzią zachodzi zależność semantyczna, np. *dom — dach, łóżko — sen*, oraz dźwiękowe (clang and rhyme), czyli takie, w których pomiędzy bodźcem a odpowiedzią zachodzi podobieństwo brzmieniowe, np. *baranina — pani nina*. Dalej, zgodnie z tradycją, skojarzenia semantyczne Clark dzieli skojarzenia na paradygmatyczne i syntagmatyczne. Skojarzenia paradygmatyczne to takie, w których wyraz bodziec i wyraz odpowiedź należą do tej samej części mowy. Natomiast skojarzenia syntagmatyczne to takie, w których bodziec i odpowiedź należą do różnych części mowy. Jednak prowadzona przez Clarka analiza powiązania bodziec — odpowiedź jest już prowadzona z językoznawczego punktu widzenia. Znaczenie jest reprezentowane jako zbiór cech syntaktycznych (Chomsky 1965), np. *mężczyzna*: [+Noun, +Det, +Count, +Animate, +Human, +Adult, +Male], a zależność pomiędzy znaczeniami wyrażają operujące na cechach reguły, np. reguła minimalnego kontrastu, zmieniająca znak jednej cechy, powiedzmy +Male => -Male, przekształca jednostkę słownika *mężczyzna* w *kobieta*: [+Noun, +Det, +Count, +Animate, +Human, +Adult, -Male]. Clark postuluje szereg reguł opisujących zależności paradygmatyczne, stwierdzając zarazem, że znacznie trudniej zbudować reguły dla powiązań syntagmatycznych, gdyż te — zdaniem Clarka — powinny operować na syntaktycznych cechach selekcyjnych, które tylko częściowo reprezentują znaczenie wyrazów łączących się w zdaniu, a więc trudno za ich pomocą odróżnić na przykład szeroką łączliwość przymiotnika *biały* od silnie ograniczonej łączliwości przymiotnika *kiszony* — w zasadzie tylko: *kapusta, ogórek*. Drugim powodem utrudniającym formułowanie reguł generujących skojarzenia syntagmatyczne jest fakt, że wiele skojarzeń syntagmatycznych, np. *chleb — masło, igła — nitka, młotek — gwóźdź*, by pojawić się w zdaniu, wymaga wyrazu pośredniczącego, tj. *chleb z masłem, igła z nitką, młotek wbija gwóźdź*.

Dlatego w pracy Clarka pojawia się tylko ogólna reguła generująca skojarzenia syntagmatyczne. Szczególny przypadek skojarzeń syntagmatycznych tworzą skojarzenia składnikowe (*idiom completion*), w których bodziec i odpowiedź są składnikami wielowyrazowej jednostki leksykalnej, np. skojarzenie *owca — czarna* wobec jednostki leksykalnej *czarna owca* czy *owca — wilk* wobec *wilk w owczej skórce*. W konsekwencji Clark przyjmuje, że skojarzenia syntagmatyczne są wyabstrahowane ze struktur syntaktycznych². Dziś można powiedzieć, że skojarzenia paradygmatyczne i syntagmatyczne reprezentują zależności semantyczne przechowywane w słowniku, które można opisać za pomocą semantycznych relacji leksykalnych.

3.1. Relacje leksykalne

Od czasu fundamentalnego studium o słownictwie Platona, w którym Lyons usystematyzował strukturalistyczny model relacji paradygmatycznych, wprowadzając przy tym neogreckie nazwy relacji, relacje leksykalne stały się przedmiotem intensywnych badań prowadzonych w obrębie różnych dyscyplin naukowych. Bardzo obszernie omówienie rezultatów tych badań zawiera książka Murphy (2003). Z naszego punktu widzenia najistotniejsze są wyniki prac informatycznych zmierzających do budowy wielopoziomowych taksonomii służących reprezentacji wiedzy nazywanych ontologiami (Sowa 2000) i taksonomii kognitywistycznych, których celem była rekonstrukcja systemu pojęciowego — synsety, tj. zbiory quasi-synonimów w słowniku WordNet miały wyznaczać pojęcia (Miller i wsp. 1998). Budowa wielopoziomowych taksonomii często zmuszała do rezygnacji z klasycznych językoznawczych kryteriów różnicowania relacji paradygmatycznych i syntagmatycznych: “Sets of elements which can be substituted one for another in a given context are said to be in *paradigmatic* relationship; elements which combine to form a larger unit are said to be in *syntagmatic* relationship.” (Lyons 1970: 16). Rozluźnienie kryteriów ma swe konsekwencje, np. pragmatyczna założenia, zawierająca kilka tysięcy relacji, ontologia Cyc obok relacji paradygmatycznych ma też relacje operujące „w poprzek” taksonomii, pozwalające interpretować zjawiska syntagmatyczne (Pohl 2009). Stawiająca sobie cel poznawczy taksonomia WordNet interpretuje niektóre zjawiska syntagmatyczne za pomocą relacji paradygmatycznych, np. relacje *owca — baranina*, *owca — owczarnia* są interpretowane za pomocą relacji część — całość (holonimia)³, a przecież *baranina* to *mięso z owcy*, a *owczarnia* to *miejsce przeznaczone dla owiec*⁴. Z drugiej strony budowa taksonomii ujawniła szereg zjawisk (Murphy 2003), które były wcześniej zauważane (Lyons 1984), jednak w kontekście klasyfikacji powiązań w sieci leksykalnej mogłyby być przedmiotem

² “[...] syntagmatic associations are not merely continuous fragments of normal speech, [...] but rather responses that bear only an abstract relationship to normal speech” (Clark 1970: 283).

³ Tak Słowosieć: <http://plwordnet.pwr.wroc.pl/wordnet/3a00a6ac-c0f8-11e4-bdb2-7a5d273e87eb>. *Owczarnia* w użyciu metaforycznym: *boża owczarnia*, oznacza zbiór wiernych, a nie zwierząt.

⁴ Co potwierdza opis wyrazu *owczarnia*: <http://plwordnet.pwr.wroc.pl/wordnet/aad2e462-c0f8-11e4-aff-7a5d273e87eb>.

poważnych analiz językoznawczych. Wymieńmy najważniejsze, gdyż dotyczą zależności, wykazujących wysoką siłę powiązania. Pierwsza dotyczy zależności typ — okaz (hiponimia), gdzie zgodnie z definicją znaczenie podrzędne, np. *jamnik*, dziedziczy cechy nadrzędnika *pies*. Jednak są znaczenia nadrzędne, np. *mebel*, które nie mają cech dziedziczonych przez podrzędniki, np. *krzesło*, *kanapa*; podobnie *warzywo* — *kapusta*, *kolor* — *czerwony* itp. Można więc postawić pytanie, czy *mebel*, *warzywo*, *kolor* to rzeczywiste nadrzędniki, czy też tylko etykiety identyfikujące zbiory różnych obiektów lub stanów. Powiązania tego typu są obecne w sieci: (*krzesło* — *mebel* 869/47, tj. 5,41, *kapusta* — *warzywo* 872/39, tj. 4,47).

Kolejny problem stanowią zależności pomiędzy okazami, czyli podrzędnikami tego samego nadrzędnika. Jeśli przyjąć za Lyonsem, że zależność *krzesło* — *kanapa* określa relacja *incompatibility*, to możemy przyjąć, że jedynym wspólnym elementem łączącym oba znaczenia jest wspólny nadrzędnik *mebel*. Jednak powstaje wątpliwość, czy relacja *incompatibility* zachodzi także pomiędzy podrzędnikami dziedziczącymi cechy wspólnego nadrzędnika, np. *spaniel* — *jamnik*. Kolejne wątpliwości dotyczą relacji całość — część (meronimia). Zasadnicza wątpliwość odnosi się do sytuacji, w której całość nie ma wyodrębnionej struktury, np. *chleb* — *kromka*, *zboże* — *worek*, oraz sytuacji, gdy część maszyny lub urządzenia pełni określoną funkcję w stosunku do całości, np. *napęd* w powiązaniach *silnik* — *samochód*, *żagiel* — *statek*. Tu można bowiem upatrywać zależności syntagmatycznych. Jednak trudno rozstrzygnąć wspomniane wątpliwości bez osobnych badań, dlatego dla potrzeb jakościowej klasyfikacji powiązań przyjmujemy klasyczny Lyonsowski model relacji paradygmatycznych z jednym wyjątkiem: powiązania pomiędzy podrzędnikami tego samego nadrzędnika (ko-hiponimy), np. *wołowina* — *baranina*, oraz powiązania pomiędzy częściami tej samej całości, np. *siedzenie* — *oparcie* (ko-meronimy), będziemy traktować jako powiązania objaśniane przez ścieżkę otwartą, w której węzłem początkowym i końcowym są podrzędniki lub części, a węzłem pośredniczącym jest nadrzędnik lub całość, np. *baranina* — *mięso* — *wołowina* oraz *siedzenie* — *krzesło* — *oparcie*. Innymi słowy, powiązania te będziemy traktować jako odległe.

Poszukując modelu leksykalnych relacji syntagmatycznych napotykamy istotną trudność. Nie możemy się bowiem odwołać do skomplikowanego modelu tzw. słownika wyjaśniająco-kombinatorycznego (Meřčuk 2007), którego abstrakcyjne funkcje leksykalne, opisujące łączliwość znaczeń w zdaniu, nie nadają się do opisu powiązań w sieci leksykalnej. Drugie potencjalne źródło zachodzących pomiędzy znaczeniami relacji syntagmatycznych, czyli słownik FrameNet, także sprawia kłopot, gdyż FrameNet nie definiuje bezpośrednich zależności pomiędzy znaczeniami jednostek słownika (Lexical Units). Znaczenia te wchodzą w relacje pośrednie poprzez uczestnictwo w ramie (frame), która jest strukturą semantyczną podobną do scenariusza (Schunk, Abelson 1977), opisującą znaczenie abstrakcyjnego predykatu, np. *revenge* ‘zemsta’ w odniesieniu do konkretnej sytuacji za pomocą swoistych dla ramy ról semantycznych (FEs Frame Elements), np. *avenger* (mściciel), *offender* (krzywdziciel),

injury (krzywda), *injured party* (pokrzywdzony), *punishment* (kara), (Fillmore 2000). Jednostki leksykalne (jednostki słownika LU) wypełniają rolę w ramie dziedziczą zależności, w jakie wchodzi rola. Twórcy słownika FrameNet przyznają, że rzeczowniki mają także ramy o ograniczonej strukturze, gdyż często występują z wyrażeniami określającymi ich rodzaj (okaz), tworzywo, z którego zostały zrobione, sposób, w jaki zostały zrobione, ich przeznaczenie lub użycie. Jednak niestety właściwości te nie zostały zdefiniowane jako ramy, lecz jako role (Frame Elements) w konkretnych ramach⁵. Tak więc nie możemy sklasyfikować powiązań syntagmatycznych posługując się ramami, możemy jednak odwoływać się do idei ram i sposobu ich definiowania, definiując system wyznaczników opisujących powiązania syntagmatyczne występujące w empirycznej sieci powiązań leksykalnych.

3.2. Wyznaczniki jakości powiązania

Wracając do rozpatrywanego wcześniej przykładu powiązania *ryba — morze* możemy powiedzieć, że zależność *miejsce_charakterystyczne* zachodzi pomiędzy zwierzęciem a obiektem, będącym wytworem przyrody. Jednak dopiero rozpatrzenie kolejnych przykładów, np. *owca — łąka*, *sędzia — sąd*, *lekarz — szpital* prowadzi do wniosku, że zależność *miejsce_charakterystyczne* zachodzi pomiędzy obiektami żywymi: *ryba*, *owca* (zwierzę), *sędzia*, *lekarz* (osoba) a obiektami, które są wytworami natury *morze*, *łąka* lub obiektami wytworzonymi przez człowieka: *sąd*, *szpital*. Można więc powiedzieć, że powiązanie sklasyfikowane jako *miejsce_charakterystyczne* zachodzi pomiędzy znaczeniami należącymi do określonych klas. Innymi słowy, wyznacznik powiązania ma swój zakres wyznaczany przez klasy znaczeń. Dla określania zakresu wyznaczników jakości powiązania przyjmujemy istnienie następujących klas:

Obiekty, głównie rzeczowniki, są jednostkami leksykalnymi, które reprezentują elementy rzeczywistości percypowane przez ludzkie zmysły. Co oznacza, że za obiekt uznajemy każdą jednostkę języka, która wskazuje na osobę (*chłopiec*, *żołnierz*), zwierzę (*pies*, *owca*), roślinę (*drzewo*, *kwiat*), obiekt naturalny (*góra*, *rzeka*), zjawisko naturalne (*wiatr*, *światło*) lub artefakt, czyli rzecz (*krzesło*, *zamek*), dzieło (składające się z rzeczy i informacji abstrakcyjnej, np. *książka*, *muzyka*) lub artefakt abstrakcyjny (*religia*, *prawo*).

Własności, głównie przymiotniki, przysłówki i rzeczowniki odczasownikowe, są jednostkami leksykalnymi, które reprezentują percypowane za pomocą zmysłów lub umysłu stałe *cechy* obiektów (*duży*, *ciężki*, *biały*, *mądry*, *kapryśny*) lub percypowane za pomocą zmysłów lub umysłu czasowe stany obiektów (*choroba*, *głód*, *radość*, *złamanie*).

⁵ “We do recognize that artifact and natural kinds of nouns also have a minimal frame structure of their own. For example, artifacts often occur together with expressions indicating their sub-type, the material of which they are made, their manner of production, and their purpose/use; these are defined as FEs in the frames for various types of artifacts” (Ruppenthal in. 2010: 5).

Akcje, głównie czasowniki i rzeczowniki odczasownikowe, to jednostki leksykalne, reprezentujące działania wywołane przez obiekty i operujące na obiektach (*spać, sen, przybić, siedzieć*).

Znaczenia złożone, głównie rzeczowniki, to jednostki leksykalne reprezentujące sytuacje, czyli wywołane przez akcję lub zdarzenie statyczne stany więcej niż jednego obiektu (*kłopot, spadek*) oraz zdarzenia, czyli dynamiczne sytuacje zbudowane z obiektów, akcji i stanów powiązanych zależnościami przyczynowymi (*mecz, wykład, powódź*).

Powiązania paradygmatyczne

Jak powiedzieliśmy wcześniej, przyjmujemy klasyczny językoznawczy (Lyons 1963) model powiązań paradygmatycznych. Wyróżniamy więc następujące wyznaczniki powiązania:

Hiponimia, czyli powiązanie pomiędzy znaczeniem nadrzędnym i podrzędnym. Zależność ta zachodzi pomiędzy obiektami lub stanami i jest reprezentowana przez dwie relacje oraz powiązania odległe objaśniane przez ścieżkę w sieci:

hiponimia (typ — okaz) prowadząca od nadrzędnika do podrzędnika, np.: *pies — spaniel, zwierzę — owca, mebel — krzesło*;

hiperonimia (okaz — typ), prowadząca od podrzędnika do nadrzędnika, np.: *spaniel — pies, owca — zwierzę, krzesło — mebel*;

ko-hiponimia, powiązanie odległe objaśniane przez ścieżkę w sieci zachodzące pomiędzy podrzędnikami tego samego nadrzędnika, np.: *spaniel — jamnik, baranina — wołowina, zielony — czerwony*.

Meronimia, czyli powiązanie pomiędzy częścią a całością. Zależność ta właściwa obiektom jest reprezentowana przez dwie relacje i powiązanie odległe:

holonimia (część — całość) prowadząca od części do całości, np.: *noga — stół, dach — dom, pies — sfora, kromka — chleb*;

meronimia (całość — część) prowadząca od całości do części, np.: *góra — szczyt, chłopiec — głowa, ciało — ręka, owca — runo*;

ko-meronimia, objaśniana za pomocą ścieżki w sieci zależność pomiędzy częściami tej samej całości, np.: *ręka — noga, stopa — głowa, drzwi — okno*.

Synonimia, tj. zależność zachodząca pomiędzy obiektami, stanami lub akcjami, które w określonym kontekście mają znaczenia identyczne lub zbliżone, np.: *mnich — zakonnik, problem — kłopot, praca — robota*.

Antonimia, tj. zależność zachodząca pomiędzy obiektami, stanami lub akcjami, których znaczenia są przeciwstawne w określonym kontekście, np.: *kawaler — żonaty, dobry — zły, jasno — ciemno, wchodzić — schodzić*.

Komplemetarność, tj. zależność zachodząca pomiędzy obiektami lub stanami powiązanymi w sposób naturalny, których znaczenia wchodzą w opozycję, np.: *kobieta — mężczyzna, słońce — księżyc, dzień — noc, przyływ — odpływ*.

Powiązania syntagmatyczne

Cechy fizyczne

— **rozmiar** — określa zależność pomiędzy: osobami, zwierzętami, bytami nieożywionymi, w tym obiektami naturalnymi i artefaktami (rzeczami), a przysługującymi obiektowi cechami fizycznymi, określającymi percypowany za pomocą wzroku (względny) lub poddający się pomiarowi (bezwzględny) wymiar fizyczny, np.: *rzeka — szeroka* 6,02, *rzeka — głęboka* 2,89, *chata — mała* 0,6, *rzeka — wielka* 0,58, *rzeka — długa* 3,7, *rzeka — płytka* 0,23;

— **kształt** — określa zależność pomiędzy: osobami, zwierzętami, bytami nieożywionymi, w tym obiektami naturalnymi i artefaktami (rzeczami), a percypowaną za pomocą dotyku lub wzroku zewnętrzną formą obiektu, np.: *stół — okrągły* 10,43, *stół — kwadratowy* 0,23, *stół — płaski* 0,58, *rzeka — meandry* 0,23, *rzeka — kreta*⁶ 0,23, *wijąca się* 0,23;

— **kolor** — określa zależność pomiędzy: osobami, zwierzętami, obiektami naturalnymi i artefaktami (rzeczami) a właściwościami ich powierzchni percypowanymi za pomocą wzroku zależnie od tego, jak odbijają lub emitują światło, np.: *lodyga — zielona* 2,41, *ziemia — brązowa* 0,23, *księżyc — srebrny* 0,23, *żółty* 0,23;

— **tekstura** — określa zależność pomiędzy: osobami, zwierzętami, bytami nieożywionymi, w tym obiektami naturalnymi i artefaktami (rzeczami), a właściwościami powierzchni obiektu percypowanymi za pomocą wzroku lub dotyku, np.: *dywan — miękki* 6,34, *miętkość* 0,23, *dywan — kudłaty* 0,23, *dywan — puszysty* 1,38, *dywan — puch* 0,23, *dywan — puchaty* 0,23, *stół — twardy* 0,35;

— **waga** — określa zależność pomiędzy: osobami, zwierzętami, obiektami naturalnymi i artefaktami (rzecz) a względną masą obiektu, określaną poprzez siłę, jakiej trzeba użyć, by unieść obiekt, np.: *młotek — ciężki* 0,46, *motyl — lekkość* 2,27, *lekki* 1,7;

— **temperatura** — określa zależność zachodząca pomiędzy: osobami, zwierzętami, obiektami naturalnymi i artefaktami (rzeczami) a względnym stopniem intensywności ciepła obecnego w obiekcie lub substancji percypowanym przez dotyk lub wskazywanym przez termometr, np.: *rzeka — zimna* 0,23, *wódka — zimna* 0,92;

— **właściwość_swoista** — określa zależność pomiędzy: osobami, zwierzętami, obiektami naturalnymi lub artefaktami (rzeczami) a stanami lub akcjami, określającymi specyfikę obiektu, np.: *rzeka — rwąca* 5,67, *rzeka — wartka* 1,16, *rzeka — nurt* 0,93, *prąd* 0,81, *rzeka — mokra* 0,58, *rzeka — bystra* 0,23, *rzeka — dzika* 0,23, *rzeka*

⁶ Wyrazy przytaczamy w takiej formie, w jakiej zostały napisane przez uczestników eksperymentu, oznaczając gwiazdką (*) formy błędne.

— *sucha* 0,23; *dywan* — *brudny* 0,69, — *stary* 0,23; *ręka* — *złamana* 0,34, *odcięta* 0,23, *sprawna* 0,23; *butelka* — *rozbita* 0,46; *krzesło* — *stoi* 0,69; *lew* — *siła* 0,46, *silny* 0,46.

Cechy psychiczne

Cechy psychiczne przysługujące ludziom i zwierzętom:

— **charakter** — określa zależność pomiędzy: osobami, zwierzętami a sposobem bycia lub postawami wyróżniającymi obiekt, np.: *żołnierz* — *odwaga* 1,04, *waleczny* 0,23, *dzielny* 0,23, *złodziej* — *sprytny* 0,23, *lew* — *groźny* 0,57, *odważny* 0,34, *drapieżnik* 0,23, *walka* 0,34, *waleczny* 0,34;

— **umysł** — określa zależność pomiędzy: osobami a ich zdolnością rozumowania i rozumienia rzeczywistości, szczególnie w odniesieniu do zależności abstrakcyjnych (umysłowość), np.: *dziewczyna* — *mądra* 0,68, *głupia* 0,23.

Cechy przyporządkowane

— **nazwa** — określa zależność pomiędzy obiektami ożywionymi, osobami, zjawiskami naturalnymi, obiektami nieożywionymi a ich nazwami własnymi, np.: *rzeka* — *Wisła* 15,86, *miasto* — *Kraków* 30,62, *król* — *Maciuś* 11,25, *Kazimierz* 0,8, *lew* — *Simba* 0,57;

— **wiedza** — określa zależności pomiędzy obiektami naturalnymi, zjawiskami, osobami, zwierzętami, roślinami i artefaktami, stanowiące wynik poznania naukowego, doświadczenia zawodowego, znajomości języka obcego lub informacji przekazywanych przez media, np.: *pamięć* — *hipokamp* 0,34, *mózg* 10,52, *łodyga* — *biologia* 0,23, *ser* — *Szwajcaria* 0,46, *Francja* 0,23, *dywan* — *Persja* 0,69, *baranina* — *judaizm* 0,12, *rzeka* — *river** 0,12, *religia* — *Jezus* 0,56, *Chrystus* 0,23; *muzyka* — *Chopin* 0,35, *baranina* — *gangster* 0,46;

— **stereotyp** — określa zależność pomiędzy: osobami lub artefaktami a szeroko przyjętym uproszczonym wyobrażeniem, zbiorem sądów lub przekonań, np.: *ksiądz* — *pedofil* 1,28, *złodziej* 0,93, *złodziej* — *polityk* 0,57, *lekarz* — *konował* 0,46, *lekarz* — *łapówka* 0,23, *alkohol* — *śmierć* 0,34, *baranina* — *zoofilia* 0,12.

Stosunek obiektu do akcji lub zdarzenia

— **działanie** określa zależność zachodzącą pomiędzy osobami, zwierzętami, obiektami naturalnymi lub zjawiskami a typowymi, charakterystycznymi akcjami powodowanymi przez obiekt (obiekt sprawcą akcji), np.: *złodziej* — *kraść* 3,99, *krasc** 0,23, *ukraść* 0,34, *kradnie* 1,37, *doktor* — *badanie* 0,35, *leczy* 0,58, *leczyć* 0,46, *leczenie* 0,46, *żołnierz* — *walczy* 0,7, *walczyć* 0,46, *strzelać* 0,35, *strzela* 0,23, *zabijać* 0,23, *owca* —

beczy 0,43, *beczenie* 0,23, *bee* 0,23, *rzeka* — *płynię* 4,63, *płynię** 0,23, *płynąca* 0,23, *flow**⁷ 0,23, *rzeka* — *płynąc* 1,04, *kwiat* — *zapach* 4,08, *pachnie* 0,68;

— **udział** — określa zależność pomiędzy osobami lub zwierzętami a sytuacjami, stanami lub zdarzeniami, w których osoba uczestniczy, np.: *żołnierz* — *walka* 1,28, *złodziej* — *kradzież* 10,26, *kradzież** 0,91, *kradziez** 0,34, *rabunek* 0,34, a także pomiędzy osobami a zdarzeniami, w których uczestniczą, np. *żołnierz* — *wojna* 11,25, *złodziej* — *napad* 0,68;

— **przeznaczenie** — wyznacznik — określa zależność pomiędzy obiektami (rzeczy, dzieła, artefakty abstrakcyjne) a akcjami wyrażającymi powód, dla którego rzecz została zrobiona lub istnieje, oznacza: do czego obiekt służy, np.: *krzesło* — *siedzieć* 8,75, *igła* — *szyć* 0,93, *nożyczki* — *ciąć* 20,84, *dom* — *mieszkać* 0,57;

— **użyteczność** — zachodzi pomiędzy obiektami i zjawiskami naturalnymi, zwierzętami i roślinami a wykonywanymi przez osoby akcjami, określającymi użycie obiektu w konkretnym celu; ludzkie czynności nadają twórcom natury przeznaczenie, np.: *rzeka* — *pływanie* 0,12, *góra* — *wspinaczka* 0,5, *morze* — *pływać* 0,35, *słońce* — *opala* 0,34, *opalać się* 0,23.

Zależności funkcjonalne

— **miejsce charakterystyczne** — określa zależność pomiędzy: osobami lub zwierzętami a artefaktami lub obiektami naturalnymi, w których osoba lub zwierzę prowadzi działania charakterystyczne, np.: *król* — *zamek* 0,69, *ksiądz* — *parafia* 1,97, *plebania* 0,58, *owca* — *góry* 1,03, *łąka* 1,03, *hala* 0,91, *pole* 0,68, *pastwisko* 0,46, *owca* — *zagroda* 0,23%, *orzeł* — *gniazdo* 0,34;

— **domena** — określa zależność między osobą a artefaktem abstrakcyjnym, stanowiącym jej obszar działań, np.: *ksiądz* — *religia* 3,6, *sprawiedliwość* — *sędzia* 1,25, *doktor* — *medycyna* 2,88, *doktor* — *nauka* 0,69;

— **współdziałanie** — określa zależność pomiędzy osobami, zwierzętami lub osobami i zwierzętami, wykonującymi działania prowadzące do określonego wspólnego celu, np.: *ksiądz* — *katolik* 0,58, *ministrant* 0,35, *owca* — *bacha* 0,68, *pasterz* 0,57, *pies* 0,34;

— **cel** — zależność zachodząca pomiędzy: osobami a obiektami lub stanami, które są obiektem ludzkich działań lub stanowią wynik tych działań, np.: *złodziej* — *pieniądze* 3,88, *pieniądze** 0,23, *moneta* 0,23, *monety* 0,23, *portfel* 2,74, *torebka* 0,46, *czasu* 0,91, *dusz* 0,34, *łup* 0,8, *telefon* 0,23, *zegarek* 0,23, *złoto* 0,23, *samochodów* 0,23, *doktor* — *zdrowie* 1,15;

⁷ Forma użyta przez uczestnika eksperymentu.

— **posiadanie** — stan posiadania lub kontrolowania czegoś; określa zależność pomiędzy osobami lub zwierzętami a obiektami (osoby, zwierzęta, rośliny, rzeczy, obiekty naturalne), np.: *dziecko* — *rodzina* 1,7, *matka* 1,59, *rodzic* 1,47, *rodzice* 1,36, *mama* 1,25, *dom* — *własny* 0,8, *własny** 0,23;

— **użycie** — określa zależność pomiędzy osobami a obiektami naturalnymi i artefaktami, służącymi jako narzędzie, środek do realizacji celu, np.: *doktor* — *lek* 0,46, *lekarstwo* 0,35, *młotek* — *kowal* 1,04;

— wyznacznik **odżywianie** — określa zależność pomiędzy obiektamiżywionymi (osobami, zwierzętami i roślinami) a obiektami naturalnymi, artefaktami stanowiącymi pokarm, np.: *niemowlę* — *mleko* 0,35, *owca* — *trawa* 0,23, *pająk* — *owad* 7,31, *mucha* 1,14;

— wyznacznik **źródło** — zachodzi pomiędzy obiektami (osoby, zwierzęta, rośliny, obiekty naturalne, zjawiska, rzeczy lub miejsca charakterystyczne) a rzeczami, zjawiskami lub artefaktami abstrakcyjnymi; wskazuje na to, skąd coś pochodzi, gdzie ma swój początek lub z czego zostało uzyskane, np.: *baran* — *mięso* 3,6, *mleko* — *krowa* 0,5, *mleko* — *matka* 0,8, *owca* — *runo* 0,46, *chleb* — *piekarnia* 0,79, *muzyka* — *instrument* 0,81, *radio* 0,58, *religia* — *Bóg* 8,01, *słońce* — *światło* 5,15, *lampa* — *światło* 25,4;

— **materiał** — to zależność zachodząca pomiędzy obiektami naturalnymi, roślinami lub rzeczami stanowiącymi tworzywo (surowiec) a rzeczami, które są lub mogą być zrobione z tego tworzywa, np.: *kwiecie* — *bukiet* 3,51, *wianek* 1,7, *kapusta* — *bigos* 0,92, *kapuśniak* 0,8, *sałatka* 0,46, *gołąbki* 0,34, *surówka* 0,23, *z grochem* 0,23, *groch* 0,23, *zupa* 0,23, *pierogi* 0,23, *butelka* — *szklana* 1,3, *plastik* 0,46, *dom* — *drewno* 0,23, *cegła* 0,23, *mleko* — *masło* 0,46;

— **ubiór** — określa zależność pomiędzy osobami a rzeczami, w które osoba się ubiera, by chronić ciało, lub stanowiącymi składnik odzieży, np.: *chłopiec* — *spodnie* 1,35, *spodenki* 1,24, *szelki* 0,56, *ogrodniczki* 0,23, *piżama* 0,23;

— **atrybut** — określa zależność pomiędzy obiektami (osoby, zwierzęta, rośliny, rzeczy, miejsca charakterystyczne) a identyfikującymi częściami ciała, składnikami obiektu lub rzeczami charakterystycznymi dla obiektu określanego lub pomiędzy artefaktami abstrakcyjnymi a rzeczami symbolizującymi obiekt określanego, np.: *mężczyzna* — *broda* 0,91, *zarost* 0,57, *lew* — *grzywa* 6,09, *kobieta* — *szminka* 0,23, *złodziej* — *kominiarka* 0,46, *maska* 0,46, *opaska* 0,23, *worek* 0,23, *kościół* — *krzyż* 11,06, *sprawiedliwość* — *waga* 2,52, *temida* 0,8;

— **komplet** — określa zależność pomiędzy rzeczami, które łączy wspólne przeznaczenie; zazwyczaj jedna z rzeczy wnosi dodatkowe cechy podnoszące jakość całości, np.: *łóżko* (sen, spać) — *pościel* (sen, spać), *chleb* (jedzenie) — *masło* (jedzenie), *stół* — *krzesło*.

Wszystkie pary łączone przez wyznacznik komplet wyróżniają bardzo mocne skojarzenia zwrotne, np.: *łóżko — pościel* 8,5, *pościel — łóżko* 33,2, *chleb — masło* 6,09, *masło — chleb* 12,2, *stół — krzesło* 16,5, *krzesło — stół* 22,54, *igła — nitka* 32,71, *nitka — igła* 40,02, *młotek — gwóźdź* 21,2, *gwóźdź — młotek* 30. Często oba wchodzące w zależność obiekty mają w swoim węźle leksykalnym ten sam czasownik, oznaczający to samo przeznaczenie, np.: *łóżko — pościel*, *spać — spanie*. Są jednak przypadki, gdy oba obiekty wchodzące w zależność mają w swoich węzłach leksykalnych różne czasowniki, np. *stół — jedzenie*, określające jedną z możliwych czynności wykonywanych z udziałem stołu, *krzesło — siedzieć*;

— **zawartość** — określa zależność pomiędzy artefaktami lub obiektami naturalnymi a artefaktami lub obiektami naturalnymi, pełniącymi funkcję pojemnika lub ogranicznika, (obiekt przechowywany i obiekt przechowujący), np.: *chleb — chlebak* 0,23, *masło — maselniczka* 0,58, *wódka — kieliszek* 2,06, *butelka* 1,49, *bania* 0,34 [kieliszek — regionalizm krakowski], *kieliszki* 0,23, *100** 0,23, *rzeka — koryto* 0,81, *rzeka — brzeg* 0,23.

Zależności przestrzenne i czasowe

— **lokalizacja** — określa zależność przestrzenną pomiędzy obiektami (osoby, zwierzęta, rośliny, obiekty naturalne, artefakty); zależność jest względna, mamy obiekt lokalizujący i lokalizowany, np.: *dom — ogród* 2,41, *podwórko* 0,23, *osiedle* 0,23, *wieś* 0,46, *dywan — kurz* 0,58, *rzeka — dopływ* 0,23, *lampa — sufit* 0,79, *dywan — podłoga* 7,38;

— **orientacja** — określa położenie obiektu lub jego części względem stron ludzkiego ciała lub stron świata, np.: *ręka — lewa* 3,19, *baszta — północna*⁸;

— **lokalizacja_w_czasie** — określa zależność pomiędzy obiektami ożywionymi, obiektami nieożywionymi a stanami lub artefaktami abstrakcyjnymi (godzina, dzień, miesiąc), stanowiącymi określenia czasu, np.: *kwiecie — wiosna* 4,42, *maj* 1,36, *majowe* 0,57, *wiosenne* 0,45, *lato* 1,13, *księżyc — noc* 17,24, *północ* 0,23, *zmrzch* 0,34, *ciemno* 0,23, *zachód* 0,23.

Stosunek między stanami

— **współwystępowanie** — określa zależność pomiędzy stanami, sytuacjami, akcjami lub zdarzeniami a stanami towarzyszącymi lub współwystępującymi, np.: *praca — wysiłek* 1,58, *trud* 1,47, *obowiązek* 0,9, *obowiązki* 0,23, *pasja* 0,34, *pamięć — sesja* 0,34;

— **intensywność** — określa zależności pomiędzy stanem, sytuacją, akcją, zdarzeniem lub zjawiskiem a stanem określającym subiektywnie percypowany stopień nasilenia stanu, sytuacji, akcji lub zdarzenia, np.: *praca — ciężka* 11,63, *trudna* 1,69, *cięż-*

⁸ Przykład, który nie występuje w naszej sieci.

*ka** 1,13, *ciężka** 1,02, *męcząca* 0,34, *łatwa* 0, *radość* — *duża* 0,56, *ogromna* 0,45, *kłopot* — *duży* 2,78, *wielki* 2,67, *mały* 1,39, *spory* 0,23 *kłopot* — *poważny* 0,23, *światło* — *jasne* 11,06, *rażące* 0,34;

— **oznaka** — określa zależność pomiędzy ukrytymi (niepoddającymi się percepcji) stanami lub sytuacjami a percypowanymi sygnałami [symptomami] ich istnienia reprezentowanymi przez inny stan lub akcję, np.: *radość* — *uśmiech* 13,09, *śmiech* 2,6, *uśmiech* 1,2, *ły* 0,23, *hurra* 0,23, *śmiech** 0,23, *skok* 0,23, *złość* — *krzyk* 1,95, *grymas* 0,8, *czerwony* 0,8, *czerwień* 0,46, *placz* 0,34, *choroba* — *gorączka* 0,58, *ból* 11,45, *praca* — *pot* 0,11.

Przyczynowe

— **przyczynowość** — określa zależność zachodzącą pomiędzy stanami, sytuacjami, akcjami lub zdarzeniami, gdzie jedno ze znaczeń stanowi przyczynę, a drugie skutek, np.: *radość* — *życie* 2,93, *z życia* 1,81, *wiosna* 0,79, *zabawa* 0,79, *wesele* 0,56, *święta* 0,56, *dzieciństwo* 0,34, *lato* 0,34, *wolne* 0,23, *przyjaźń* 0,23, *wakacje* 0,23, *wygrana* 0,23, *praca* — *pieniądze* 5,64, *zarobki* 2,03, *ból* — *rana* 5,66, *złość* — *kłótnia* 0,8, *kłopot* — *troska* 2,2, *choroba* — *problem* 0,23.

Możliwa jest też przyczynowość niejawna (*implicit causation*) zachodząca wówczas, gdy obiekt sprawca wywołujący przyczynę (stan, sytuacja, akcja, zdarzenie) reprezentuje (zastępuje) przyczynę, lub gdy obiekt podlegający akcji lub zdarzeniu stanowiącemu przyczynę reprezentuje (zastępuje) skutek, tj. sytuację lub zdarzenie, np.: *radość* — *dziecko* 1,35, *dzieci* 1,13, *mama* 0,23, *radość* — *słońce* 2,03, *tęcza* 0,23, *praca* — *książka* 0,34, *złodziej* — *strata* 0,46, *słońce* — *opalenizna* 0,91, *niemowlę* — *opieka* 0,35, *muzyka* — *radość* 1,16, *muzyka* — *miłość* 1,04. Definicja przyczynowości implicitnej jest wzorowana na opisie FrameNet, ramy *causation_scenario*, *causation*.

Stosunek obserwatora do znaczenia wyrazu bodźca

— **identyfikacja** — określa zależność pomiędzy obiektami a ich klasyfikacją będącą wynikiem akcji lub procesu identyfikowania (rozpoznania) swoistości obiektu przez obserwatora, np.: *kobieta* — *ja* 0,23, *moja* 0,34, *męczyzna* — *on* 1,13;

— **ocena** — określa pozytywny lub negatywny stosunek do obiektu lub stanu, będący wynikiem subiektywnej postawy obserwatora, np.: *baranina* — *fuj* 0,35, *niedobre* 0,23, *kobieta* — *diabeł* 0,11, *alkohol* — *pyszności* 0,11.

Powiązania składnikowe

Składnik określa powiązanie pomiędzy wyrazami, które można zidentyfikować jak składniki wielosegmentowej jednostki leksykalnej, np.: *owca* — *czarna* => *czarna owca*, *owca* — *pęd* => *owczy pęd*, *owca* — *cała* => *wilk syty i owca cała*.

Przedstawiona lista wyznaczników uwzględnia wszystkie istotne właściwości powiązań pomiędzy znaczeniami, które ujawniły się podczas analizy węzłów zbudowanych wokół bodźców prymarnych. Zaproponowany zestaw wyznaczników traktujemy jako propozycję wstępną. Będzie ją można zweryfikować w trakcie analizy całej sieci empirycznych powiązań leksykalnych.

Zdajemy sobie sprawę z tego, że może dojść do sytuacji, w której konkretne powiązanie pomiędzy dwoma węzłami sieci można niesprzecznie opisać za pomocą dwu wyznaczników, np. powiązania *ksiądz* — *sutanna* 12,06, *żołnierz* — *mundur* 5,45, *lekarz* — *kitel* 0,58 można sklasyfikować za pomocą wyznacznika **ubiór** oraz za pomocą wyznacznika **atrybut**, bowiem ubiór jest ściśle związany z domeną aktywności konkretnej osoby. Jednak zjawisko to można prosto reprezentować w strukturze formalnej sieci. Zarówno siła powiązania, jak i wyznaczniki jakości powiązania są etykietami opisującymi powiązanie zachodzące pomiędzy dwoma węzłami sieci. Nie ma bowiem ograniczeń co do liczby etykiet i jakości informacji reprezentowanych przez etykiety.

4. Węzeł leksykalny rzeczownika a słownikowa definicja znaczenia

4.1. Podstawa empiryczna

Podstawą analiz jest sieć powiązań leksykalnych, którą zbudowano za pomocą eksperymentu autorskiego. W badaniu swobodnych skojarzeń słownych uczestniczyło 900 studentów UJ (Wydziału Zarządzania i Komunikacji Społecznej UJ, Wydziału Matematyki i Informatyki oraz Wydziału E AGH). Jako wyrazów bodźców użyliśmy 63 rzeczowniki z polskiego tłumaczenia listy Kent-Rosanoffa, którego dokonała I. Kurcz (1967), były to tzw. bodźce prymarne. Drugą grupę bodźców, tzw. bodźce sekundarne, stanowiły wyrazy będące najczęstszymi skojarzeniami w eksperymencie Kurcz, średnio 5 najczęstszych odpowiedzi do bodźca prymarnego. Jeśli dana odpowiedź pojawiała się dla różnych bodźców prymarnych, np. *biały* dla *doktor*, *ser*, *owca*, to wyraz jako bodziec sekundarny pojawił tylko raz. Bodźce sekundarne występowały w takiej formie gramatycznej, w jakiej pojawiły się jako odpowiedź w eksperymencie Kurcz, a więc np. *biała*, *wieku*, *do papieru*. W efekcie bodźce prymarne i sekundarne to 322 wyrazy, w tym 192 rzeczowniki, 103 przymiotniki, 15 czasowników (w tym imiesłowy czynne), 6 nazw własnych: *Maciuś*, *Lear*, *Tatry*, *Warszawa*, *Wieliczka*, *Wisła* oraz wyrażenia *do papieru*, *na świat*, *nie ma*, *z dziurami*, *wuja Toma*, *za wsią*. Uzyskana w wyniku w wyniku eksperymentu sieć liczy: 11 224 węzły leksykalne oraz 50 849 powiązań (relacji). Węzły leksykalne to wyrazy notowane w „Słowniku fleksyjnym języka polskiego” (Lubaszewski i in. 2001). W tym: 7 757 rzeczowników (69,1%), 2 459 przymiotników (21,9%), 744 czasowniki (6,6%) i 264 (2,4%) wyrazy należące do pozostałych części mowy. Sieć wykazuje duże podobieństwo strukturalne do współcześnie budowanej sieci języka holenderskiego (De Deyne, Storms 2008), gdzie: rzeczowniki to 72%, przy-

miotniki 18%, czasowniki 9%, a pozostałe 1%. Pozostałe 592 węzły sieci zbudowanej za pomocą eksperymentu stanowią nazwy własne, które wchodzą w 950 powiązań.

Graficzną reprezentację węzła leksykalnego przedstawimy na przykładzie wyrazu *dom*, gdzie wyraźnie widać, że sieć zawiera powiązania dla dwu znaczeń ‘budynek’ oraz ‘miejsce rodziny’. Chcąc zachować czytelność rysunku, pominieliśmy wszystkie powiązania, których siła była mniejsza niż 0,34.

4.2. Węzeł leksykalny a definicja słownikowa

Ze względu na szczupłość miejsca przeanalizujemy tylko dwa przykłady węzłów leksykalnych rzeczownika. Porównamy znaczenie słownikowe z zawartością węzła leksykalnego, identyfikując jakość powiązań w węźle za pomocą wyznaczników, podając siłę powiązania wychodzącego (wyrażoną w procentach). Opis będzie także sygnalizować powiązania przychodzące, jeśli wystąpią. Wyrazy przytaczamy w takiej formie, w jakiej zostały napisane przez uczestników eksperymentu, oznaczając gwiazdką (*) formy błędne.

4.2.1. Węzeł leksykalny *księżyc*

Węzeł leksykalny wyrazu *księżyc*, który był bodźcem prymarnym, zawiera powiązania ze 170 wyrazami, których siłę powiązania obliczono na podstawie 870 odpowiedzi — 30 badanych nie podało swojego skojarzenia do wyrazu *księżyc*. Porównywana z węzłem definicja znaczenia pochodzi ze „Słownika języka polskiego PWN” — dostęp *on-line*. Definicja ma postać: *Księżyc* — rzeczownik, r. m., l. p., l. «naturalny satelita Ziemi» 2. «ciało niebieskie krążące dookoła jakiejś planety» • *księżycowy*

Znaczenie pierwsze: ‘naturalny satelita Ziemi’ reprezentują powiązania sklasyfikowane jako:

wiedza *satelita* 2,53, *ziemia* 1,49; łącznie 4,02 oraz przychodzące *ziemia* 0,92.

Znaczenie *księżyc*: ‘ciało niebieskie krążące dookoła jakiejś planety’ ma powiązania sklasyfikowane jako:

wiedza *planeta* 1,03, *planety* 0,23, *ciało niebieskie* 0,23, *galaktyka* 0,23, *kosmos* 0,5; łącznie 2,29 oraz przychodzące *planeta* 0,11.

Łącznie oba znaczenia umieszczone w słowniku są identyfikowane w sieci przez powiązania wychodzące, których łączna siła wynosi 6,31 i powiązania przychodzące 1,03.

Najliczniejszy zbiór powiązań odnosi się do znaczenia wyrazu *księżyc*, które nie zostało bezpośrednio zdefiniowane w słowniku i można je zdefiniować jako obiekt będący naturalnym *źródłem światła*. Do tego znaczenia odnoszą się powiązania:

Paradygmatyczne:

komplementarność *słońce* 9,08, *słońce**, 0,69, *słońca* 0,23 łącznie: 10 oraz przychodzące *słońce* 5,38.

Syntagmatyczne:

kształt *pełnia* 7,24, *w pełni* 6,55, *okrągły* 0,57, *kula* 0,23, *nów* 4,94, *w nowiu* 3,1, *rogal* 0,69, *półksiężyc* 0,46, *sierp* 0,23, *rogalik* 0,46, *ser* 0,46 oraz przychodzące *pełnia* 34,11, *pełny* 8,35, *okrągły* 0,56;

kolor *srebro* 0,34, *jasny* 3,22, *żółty* 0,23;

właściwość swoista *zaciemnienie* 0,23;

działanie *świeci* 5,63, *świeci** 0,23 oraz przychodzące *świecić* 0,34;

źródło *światło* 0,92, *blask* 0,57, *jasność* 0,34 oraz przychodzące *światło* 0,11;

komplet *gwiazdy* 3,91, *gwiazda* 0,46;

lokalizacja *niebo* 2,87;

lokalizacja w czasie *noc* 17,24, *północ* 0,23, *zmiersch* 0,34, *ciemno* 0,23, *zachód* 0,23 oraz przychodzące *noc* 4,17;

wiedza *podróż* 0,34, *Armstrong* 0,57, *luna** 0,8, *moon** 0,23;

stereotyp *wilkolak* 0,23;

ocena *piękny* 0,57, *piękno* 0,23;

składnikowe *Czarodziejka* 0,23: „Czarodziejka z Księżycą” — japońska kreskówka.

Jak widzimy, encyklopedyczne definicje znaczenia wyrazu księżyc mają tylko niewielką liczbę powiązań wychodzących, których łączna siła powiązania wynosi 6,31 procent. Zdecydowanie liczniejsze są powiązania (łączna siła ponad 90 procent), odnoszące się do znaczenia: księżyc — naturalne *źródło światła*, którego nie podaje słownik. Powiązania paradygmatyczne stanowią niewielką część charakterystyki sieciowej tego znaczenia — tylko 10 procent. Ponieważ księżyc jest obiektem naturalnym percypowanym za pomocą wzroku, wśród powiązań syntagmatycznych dominują te, które można percypować za pomocą wzroku, a więc kolor, zmienny kształt i związana z tym właściwość swoista *zaciemnienie*. Księżyc działa: *świeci* i wytwarza *światło*, ma swoją lokalizację w przestrzeni wyrażaną przez stosunek do innych obiektów (gwiazdy) oraz lokalizację w czasie (*noc*), która stawia go w opozycji do *słońca* świecącego w dzień. Powiązania reprezentujące wiedzę, stereotyp, ocenę i składnikowe są nieliczne. Podsumowując, można powiedzieć, że księżyc jako węzeł sieci naturalnych powiązań leksykalnych to: obiekt widoczny na niebie w nocy, świeci jasno (światłem koloru srebrnego), jest widoczny na tle gwiazd. Kształt księżycy zmienia się od nowiu (rogal, rogalik, półksiężyc, sierp) do pełni (okrągły, kula, ser). Jest przeciwieństwem słońca.

4.2.2. Węzeł leksykalny *owca*

Węzeł leksykalny wyrazu *owca*, który był bodźcem prymarnym, zawiera powiązania ze 166 wyrazami, których siłę powiązania obliczono na podstawie 877 odpowiedzi — 27 uczestników nie podało swojej odpowiedzi do wyrazu *owca*.

Porównywane z węzłem znaczenie pochodzi ze „Słownika Języka Polskiego PWN” — dostęp *on-line*. Definicja ma postać *owca* — rzeczownik, r. ż., l. p., definicja SJP: «zwierzę roślinożerne o miękkiej, skręconej sierści, żyjące w stadzie».

Słownikowe znaczenie wyrazu *owca* reprezentują następujące powiązania:

Paradygmatische:

hiperonimia (okaz — typ): *zwierzę* 2,62 (hiperonim) oraz przychodzące *zwierzę* 3,72;

holonimia (część — całość): *stado* 0,8;

meronimia (całość — część): *szerść** 0,23.

Syntagmatyczne:

tekstura *mięka* 0,34, *kudłata* 0,23;

odżywianie *trawa* 0,23.

Powiązania wychodzące łącznie: 4,45% i powiązania przychodzące 3,72.

Pozostałe odpowiedzi rozszerzają definicję. W węźle leksykalnym wyrazu *owca* znajdujemy dodatkowe informacje, które etykietowane są przez następujące wyznaczniki:

Paradygmatische:

komplementarność *baran* 11,0 oraz przychodzące *baran* 28,19;

ko-hiponimia *koza* 1,94, *krowa* 0,57;

meronimia (całość — część): *skóra* 0,23.

Syntagmatyczne:

rozmiar *mała* 0,11;

kolor *biała* 5,25;

tekstura *puszysta* 0,23;

działanie *beczy* 0,43, *beczenie* 0,23, *bee* 0,23, *pasie się* 0,11;

użyteczność *strzyżenie* 0,23, *klonowanie* 0,23;

miejsce charakterystyczne *góry* 1,03, *łąka* 1,03, *hala* 0,91, *pole* 0,68, *pastwisko* 0,46, *zagroda* 0,23, *farma* 0,23;

współdziałanie *bacha* 0,68, *pasterz* 0,57, *pies* 0,34;

źródło *mleko* 0,11, *runo* 0,46 oraz przychodzące *baranina* 5,31, *mięso* 0,81;

wiedza *dolly*⁹ 1,37;

ocena *miła* 0,11;

składnikowe *wilk* 1,82 => *wilk w owczej skórze* WSF: s. 619; *czarna* 23,49 oraz przychodzące *czarna* 7,63 => *czarna owca* WSF s.332; *cała* 1,03 => *i wilk syty, i owca cała* WSF s. 619; *pęd* 0,23 => *owczy pęd* WSF s. 348; *zbląkana* 0,23 => *zbląkana owca* WSF s. 332, *błądna* 0,11 => *chodzić jak błędna owca* WSF s. 332;

odległe *welna* 10,83 oraz powiązania przychodzące: *welna* 43,2, *bawelna* 0,11, *ser* 0,34, *oscypek* 0,23.

⁹ *Owca Dolly* — owca domowa, pierwsze sklonowane zwierzę (1996–2003) przez naukowców Iana Wilmuta i Keitha Campbella (Edynburg, Szkocja).

Jak widzimy, powiązania wychodzące poza znaczenie zdefiniowane w słowniku stanowią zdecydowaną większość (ponad 90 procent) i powiązania te rozszerzają definicję słownikową wyrazu *owca*, wprowadzając kolejne elementy charakterystyki zwierzęcia, a więc: żyje w parze z *baranem*, dowiadujemy się o kolorze sierści owcy, poznajemy inne zwierzęta hodowane przez człowieka, dowiadujemy się, jak się zachowuje, jakie odgłosy wydaje, w jakich warunkach żyje, czym się żywi, w jakich miejscach przebywa, kto się nią opiekuje, jakie korzyści odnosi człowiek z hodowli owcy. Poza tym w węźle leksykalnym wyrazu *owca* występują powiązania wychodzące do wyrazów stanowiących fragment złożonych jednostek leksykalnych, które współtworzy wyraz *owca* (powiązania składnikowe).

Podsumowując, można powiedzieć, że nawet tak bardzo ograniczona analiza porównawcza pokazuje pewne właściwości opisu znaczenia w węźle leksykalnym empirycznej sieci powiązań leksykalnych. Widać bowiem wyraźnie, że sieciowy model znaczenia jest bogatszy od znaczenia definiowanego przez słownik.

5. Podsumowanie

Niniejszy tekst miał pokazać budowę i właściwości empirycznej sieci powiązań leksykalnych w taki sposób, by w podsumowaniu można było wskazać dalszy kierunek badań. Niewątpliwie dalsza analiza węzłów zbudowanych wokół bodźców prymarnych powinna pozwolić na stwierdzenie tego, czy sieć rozróżnia modele powiązań charakteryzujących np. osoby, zwierzęta, rośliny, rzeczy. Odrębna analiza powinna pokazać, jak jest reprezentowane znaczenie wieloznacznego wyrazu bodźca. Kolejnym pytaniem, na które może odpowiedzieć analiza węzła leksykalnego sieci, jest pytanie o to, czy sieć odwzorowuje ręcznie budowane taksonomie znaczeń obserwowane w słownikach, np. WordNet, czy ontologiach Cyc. Można wreszcie pokazać swoistość węzła leksykalnego definiowanego wyłącznie za pomocą powiązań przychodzących.

Empiryczna sieć powiązań leksykalnych dostarcza także materiału do innych analiz, które nie odwołują się do wszystkich właściwości sieci. Można bowiem badać właściwości tzw. powiązania zwrotnego, np. dla połączenia rzeczownik — przymiotnik, powiedzmy *kapusta* — *kiszona* i *kiszona* — *kapusta*. Można też badać, czy w sieci znajdują się sygnały tego, że grupa badanych miała kłopot ze zrozumieniem wyrazu bodźca, np. *plonna* — *nadzieja* 49,7 wobec *plonna* — *miłość* 1,95, *kobieta* 0,8, *wiara*, *łaka*, *świeca* wszystkie 0,7. Można wreszcie badać, jak cechy gramatyczne bodźca wpływają na cechy gramatyczne odpowiedzi, np. *biały* — *śnieg* 20,48 wobec *biały* — *mleko* 0,7, *chmura* 0,45 w zestawieniu z *biała* — *flaga* 9,57 wobec *biała* — *śnieg* 0,7, *mleko* 0,34.

Na koniec trzeba wspomnieć o coraz częstszych zastosowaniach empirycznej sieci leksykalnej w badaniach lingwistyczno-komputerowych. Najważniejsze wydają się prace poświęcone badaniu zależności tekst — sieć np. Rapp (2002), Wettler i inni (2005), Wandmacher i inni (2008), Schulte im Walde i inni (2012), także Gatkowska i inni (2013), Lubaszewski i inni (2015). Wydaje się, że analiza sieci, w której powią-

zania sklasyfikowano za pomocą wyznaczników, mogłaby wnieść nowy punkt widzenia pozwalający lepiej zrozumieć zależność sieć — tekst.

Bibliografia

- Bartmiński J., 1988a, Definicja kognitywna jako narzędzie opisu konotacji, [w:] Konotacja, red. J. Bartmiński, Lublin, s. 169–183.
- 1988b, Słownik ludowych stereotypów językowych. Założenia ogólne, [w:] Etnolingwistyka 1, red. J. Bartmiński, Lublin, s. 11–34.
- Chomsky N., 1965, *Aspects of the Theory of Syntax*, MIT Press, Cambridge Mass.
- Clark H.H., 1970, *Word Associations and Linguistic Theory*, [w:] *New Horizons in Linguistics*, red. J. Lyons, s. 271–286.
- De Deyne S., Storms G., 2008, Word associations: Network and semantic properties, *Behavior Research Methods* 40 (1), s. 213–231.
- Fillmore Ch.J., 2000, *FrameNet, The Work. The Product. The Applications*, <https://framenet.icsi.berkeley.edu>.
- Gatkowska I., Korzycki M., Lubaszewski W., 2013, Can Human Association Norm Evaluate Latent Semantic Analysis? [w:] *Proceedings of the 10th NLPCS Workshop, Marseille*, s. 92–104.
- Gatkowska I., 2014, Word Associations as a Linguistic Data, [w:] *Languages in Contact 2012*, t. 1, red. P. Chruszczewski, J. Rickford, K. Buczek, A. Knapik, J. Mianowski, Wrocław, s. 79–92.
- Geeraerts D., 2010, *Theories of Lexical Semantics*, Oxford Linguistics.
- Kent G.H., Rosanoff A.J., 1910, A study of association in insanity, *American Journal of Insanity* 67 (37–96), s. 317–390.
- Kiss G.R., Armstrong C., Milroy R., Piper J., 1973, An associative thesaurus of English and its computer analysis, [w:] *The Computer and Literary Studies*, red. A.J. Aitken, R.W. Bailey, Edinburgh, s. 153–165.
- Kurcz I., 1967, Polskie normy powszechności skojarzeń swobodnych na 100 słów z listy Kent-Rosanoffa, *Studia Psychologiczne* VIII, s. 122–255.
- Lubaszewski i in., 2001, *Słowniku fleksyjnym języka polskiego, dokument elektroniczny*, Kraków
- Lubaszewski W., Gatkowska I., Haręza M., 2015, Human Association Network and Text Collection, [w:] *Natural Language Processing and Cognitive Science*, red. B. Sharp, R. Delmonte, Berlin–New York, s. 101–114.
- Lyons J., 1963, *Structural Semantics. An Analysis of Part of the Vocabulary of Plato* (Publ. of the Philological Society 20), Oxford.
- (red.), 1970, *New Horizons in Linguistics*, Harmondsworth.
- 1984 (1977), *Semantyka*, t. 1 (Semantics 1), tłum. A. Weinsberg, Warszawa.
- Meřuk I.A., 2007, *Lexical Functions*, [w:] *Phraseology. An International Handbook of Contemporary Research*, red. H. Burger, D. Dobrovolskij, P. Kühn, N. Norrick, Berlin–New York, s. 119–131.
- Miller G.A., Beckwith R., Fellbaum Ch., Gross D., Miller K., 1998, *Introduction to WordNet: An On-line Lexical Database*, [w:] *WordNet An Electronic Lexical Database*, red. Ch. Fellbaum, Cambridge Mass.
- Melinger A., Weber A., 2006, *Database of Noun Associations for German*, <http://www.psycholing.es.uni-tuebingen.de/nag> [dostęp: 03.05.2015].
- Moss H., Older L., 1996, *Birkbeck word association norms*, Psychology Press.
- Murphy M.L., 2003 (2008), *Semantic Relations and the Lexicon: Antonymy, Synonymy, and other Paradigms*, Cambridge.
- Nelson D.L., McEvoy C.L., Schreiber T.A., 1998, *The University of South Florida word association, rhyme, and word fragment norms*, www.usf.edu/FreeAssociation/ [data dostępu: 6.06.2013].
- Palermo D.S., Jenkins J.J., 1964, *Word Associations Norms: Grade School through College*, Minneapolis.

- Postman L., Keppel G., 1970, Norms of Word Association, New York, Academic Press.
- Pohl A., 2009, Rozstrzygnięcie wieloznaczności, maszynowa reprezentacja znaczenia wyrazu i ekstrakcja znaczeń, [w:] Słowniki komputerowe i automatyczna ekstrakcja informacji z tekstu, red. W. Lubaszewski, Kraków.
- Rapp R., 2002, The Computation of Word Associations: Comparing Syntagmatic and Paradigmatic Approaches, [w:] Proceedings of the 19th International Conference on Computational Linguistics, Taipei.
- Ruppenhofer J., Ellsworth M., Petrucci M.R.L., Johnson C.R., Scheffczyk J., 2010, FrameNet II: Extended Theory and Practice, Berkeley.
- Schank R.C., 1975, Conceptual Information Processing, Amsterdam.
- Schank R.C., Abelson R., 1977, Scripts Plans Goals and Understanding, Hillsdale NJ.
- Schulte im Walde S., Borgwaldt S., Jauch R., 2012, Association Norms of German Noun Compounds, [w:] Proceedings of the 8th International Conference on Language Resources and Evaluation, Istanbul.
- Sowa J.F., 2000, Knowledge Representation. Logical, Philosophical, and Computational Foundations, Course Technology, Boston.
- Steyvers M., Tenenbaum J.B., 2005, The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth, *Cognitive Science* 29, s. 41–78.
- Uhr P., Klahold A., Fathi M., 2013, Imitation of the Human Ability of Word Association, *JSCSE* 3 (3), s. 2251–7545.
- Wandmacher T., Ovchinnikova E., Alexandrov T., 2008, Does Latent Semantic Analysis reflect human associations, [w:] Proceedings of the ESSLLI Workshop on Distributional Lexical Semantics.
- Wettler M., Rapp R., Sedlmeier P., 2005, Free word associations correspond to contiguities between words in text, *Journal of Quantitative Linguistics* 12 (2), s. 111–122.

Słownik języka polskiego, www.sjp.pwn.pl, on-line.

<http://www.opencyc.org/>

Wielki Słownik Frazeologiczny PWN z przysłowiami, oprac. A. Kłosińska, E. Sobol, A. Stankiewicz, Warszawa 2005.

SUMMARY

Empirical Lexical Network

Keywords: empirical network, lexical semantics, lexical relations, natural dictionary, computational linguistics.

Słowa kluczowe: sieć empiryczna, semantyka leksykalna, relacje leksykalne, słownik naturalny, lingwistyka komputerowa.

The properties of experimentally built association networks were studied from many different points of view. The results of those investigations show that if one aims to treat an empirical network as a ‘natural dictionary’ suitable for text processing algorithms, one should develop a consistent and linguistically valid set of semantic relationships which may explain information borne by word associations. The aim of this paper is describe a preliminary set of relationships which are applicable to word associations, and show how those associations, if classified, define the meaning of a noun which is a node of the network.