



You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice

Title: Kognitywna leksykografia komputacyjna : teoria Słownika Generatywnego Pustejowsky'ego i Boguraev'a

Author: Wiesław Banyś

Citation style: Banyś Wiesław. (1995). Kognitywna leksykografia komputacyjna : teoria Słownika Generatywnego Pustejowsky'ego i Boguraev'a. "Biuletyn Polskiego Towarzystwa Językoznawczego" (Z. 51 (1995), s. 49-66).



Uznanie autorstwa - Użycie niekomercyjne - Bez utworów zależnych Polska - Licencja ta zezwala na rozpowszechnianie, przedstawianie i wykonywanie utworu jedynie w celach niekomercyjnych oraz pod warunkiem zachowania go w oryginalnej postaci (nie tworzenia utworów zależnych).



UNIwersYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego

WIESŁAW BANYŚ
Sosnowiec

Kognitywna leksykografia komputacyjna: Teoria Słownika Generatywnego Pustejowsky'ego i Boguraev'a

Celem prezentowanego tekstu jest skrótowe przedstawienie i omówienie pewnego modelu reprezentacji wiedzy leksykalnej, dążącego do przezwyciężenia typowego dla słowników operacyjnych, realnie funkcjonujących w projektach zautomatyzowanego opisu słownictwa, którymi zajmuje się tzw. leksykografia komputacyjna, „listingu”, wyliczania, różnych możliwych znaczeń wyrażen traktowanych jako polisemiczne. Model ten opracowany został przez J. Pustejowsky'ego i B. Boguraev'a (cf. m.in. Pustejowsky, Boguraev (1993), Boguraev, Briscoe (1989), Boguraev (1992), Pustejowsky, Bergler (1992), Briscoe, Copestake, Boguraev (1990)), a oparty został m.in. na koncepcji tzw. *frame*, czyli zestawu typowych informacji, jakich powinno się standardowo udzielić wtedy, gdy mowa o danej rzeczy, pojęciu, etc. Model ten nazwany został *Teorią Słownika Generatywnego (Generative Lexicon Theory)*.

W pierwszej części tekstu przedstawimy motywację generalną, jaka przyświeca tego typu unifikacji znaczeniowej (sekcja 1), następnie przedstawimy proponowany sposób opisu znaczenia pozwalający na odejście od „listingu” znaczeniowego zaproponowanego przez J. Pustejowsky'ego i B. Boguraev'a (sekcja 2), a potem skonfrontujemy pokrótce prezentowany model opisowy z aktualnie opracowywanym modelem opisu reprezentacji wiedzy leksykalnej także opartym na koncepcji *frame* (sekcja 3).

1. W tradycyjnym, operacyjnym, ujęciu komputacyjnego opisu leksyko-
graficznego zakłada się, że znaczenie słowa może być w sposób wyczerpu-
jący zdefiniowane poprzez wyliczenie wszystkich podznaczeń szczegółowych
słowa opisywanego; w pewnym sensie, trzeba zatem zawczasu wyspecyfiko-
wać możliwe konteksty (bądź je znaleźć), w których słowo się pojawi; rezul-

tatem takiego podejścia do opisu znaczenia jest statyczna natura słownika, oddzielającego wyraźnie wydobywane podznaczenia jako odrębne podhasła. Jednakże rozdział na odrębne znaczenia ma swoje niedomagania, zwłaszcza dla badacza o orientacji kognitywistycznej. M.in. nie tylko nie pozwala na przechodzenie od jednego z nich do drugiego w sposób płynny, ale także nie pozwala na wyjaśnienie kreatywnego używania słów, w nowych kontekstach, co kłóci się z tym, co leksykograf komputacyjny z nastawieniem kognitywistycznym do analizy słownika mógłby zaakceptować. Nie jest to tyle wada słownika typu „listing”, który pozostaje operacyjny, i, z punktu widzenia swej skuteczności, spełnia pokładane w nim nadzieje, ile podejście konceptualne, które przewodzi jego powstawaniu może budzić wątpliwości u tych, którzy nie tylko szukają skuteczności analizy słownikowej, ale także oczekują by oddawała ona ogólne mechanizmy języka, jakie zachodzą w głowach ludzi posługujących się językiem.

Por. w tym kontekście np. dwuznaczne zdania (1) — (6)* (cf. Pustejowsky, Boguraev (1993: 198)), gdzie znaczenie predykatu zmienia się w zależności od rzeczownika, który jest modyfikowany:

1. *Wysłano po nas małą szybką łódź*
(łódź, która szybko płynie / która może szybko płynąć)
2. *Szybka maszynistka*
(„par défaut”: osoba, która szybko dokonuje czynności przepisywania na maszynie do pisania)
3. *Tenis jest szybką grą*
(ruchy wykonywane w trakcie gry są szybkie)
4. *Szybka książka*
(taka, którą można przeczytać w krótkim czasie)
5. *Paweł jest szybkim kierowcą*
(„par défaut”: takim, który szybko jeździ)
6. *Szybkie decyzje*
(takie, które zajmują mało czasu).

W powyższych zdaniach można wyróżnić przynajmniej cztery znaczenia słowa *szybki*:

- szybki* (1): szybko się poruszający
szybki (2): szybko wykonujący jakąś czynność
szybki (3): robienie czegoś wymaga krótkiego czasu
szybki (4): zawierający szybki ruch.

W *Małym Słowniku Języka Polskiego* pod redakcją S. Skorupki, H. Auerskiej i Z. Lempickiej, hasło *szybki* ma dwa podznaczenia:

szybki: 1. działający, poruszający się, wykonujący co, wykonywany w ciągu krótszego niż przeciętnie czasu: *szybki krok*, *szybki oddech*, *szybkie tempo*, 2. następujący, mający nastąpić zaraz, niedługo; natychmiastowy, niezwłoczny, raptowny, rychły: *szybka decyzja*, *szybki koniec*.

1.1. W bardziej szczegółowym słowniku komputacyjnym, znaczenia słów byłyby uzupełnione o restrykcje selekcyjne: *szybki* (1) może być predykowany, orzekany, jedynie o obiektach należących do klasy obiektów poruszających

się; *szybki* (3) natomiast łączy czynność zabierającą mało czasu (np. *czytanie* w (4)), z obiektem, który jest modyfikowany (por. np. z tego punktu widzenia niezwykle precyzyjne słowniki języka francuskiego DELAS, opracowywane w LADL w Uniwersytecie Paryż-VII i XIII pod kierunkiem Maurice Grossa (1975, 1989, 1990) i Gastona Grossa (1988, 1989, 1992).

Jeśli się przyjrzeć sposobowi w jaki *szybki* jest stosowany, to się okaże, że orzeka on nieco inaczej w każdym typie przypadku. Np. por. też *szybka droga* i *szybki garaż* (w sensie np. ang. czy fr. *garage*) (cf. Pustejowski, Boguraev (1993: 198)). W pierwszym przypadku *szybki* odnosi się do możliwości, jaką mają pojazdy do poruszania się szybko na danej jezdni, a w drugim — do okresu czasu wymaganego do naprawy w danym garażu. Traktując te użycia jako nowe dla *szybki*, trzeba szukać nowych znaczeń dla *szybki*, które nie są wymienione w *szybki* (1) — (4).

Można by zatem powiedzieć, że z prezentowanego tutaj punktu widzenia, granice między poszczególnymi znaczeniami dla *szybki* są zbyt sztywne, gdyż nie pozwalają na połączenie ich w jedną całość i na „spolegliwe” przechodzenie od jednego z nich do drugiego.

Ale nawet gdyby granice między nimi nie były sztywne, to nie zawsze wiadomo jak wybierać w danym kontekście sens, o który chodzi, por. np.:

7. *Jan upiekł kasztany*

7a. *Jan upiekł ciasto*

8. *Jan wykopał ziemniaki*

8a. *Jan wykopał dół*

(gdzie pojawia się różnica lektur *zmiana stanu / wytworzenie*).

Częściowe zachodzenie na siebie centrum i peryferii znaczeniowych — wedle przynajmniej koncepcji, które przyjmują, że istnieje centrum i peryferia znaczenia, a do takich zaliczane są koncepcje typu kognitywistycznego — powoduje, że tradycyjna analiza komputacyjna nie wydaje się być adekwatna.

1.2. Podobnie arbitralne wydaje się, z punktu widzenia kognitywisty, twórczenie oddzielnych znaczeń dla słowa tylko dlatego, że występuje ono w różnych kategoriach gramatycznych. Por. np. (9) — (13), gdzie syntaktyczna realizacja dopełnienia czasownika wskazuje jak ma być interpretowane zdanie. Dopełnienie wprowadzane np. przez *że* wskazuje na faktywne użycie predykatu a konstrukcja z bezokolicznikiem na implikatywne (tutaj tzw. implikatywne z negacją zwane także *neg-implikatywnym*) użycie predykatu (cf. w tym względzie np. Kiparsky (1970), Karttunen (1973a, 1973b), Givon (1973), Banyś (1986)).

9. *Jan zapomniał, że pies nic nie jadł przez cały dzień*
(faktywne)

10. *Jan zapomniał wyprowadzić psa na spacer*
(implikatywne z negacją).

Możliwość odróżniania przez system interpretacyjny faktywności i, ogólnie, nie-faktywności, wpływa w sposób oczywisty na interpretację pytań:

Jeśli pytanie będzie brzmieć:

11. *Czy Jan zamknął drzwi?*

to odpowiedź będzie *Tak* lub *Nie*, w zależności od tego czy, na wejściu systemu pojawiło się zdanie

12. *Jan zapomniał, że zamknął drzwi*

czy też zdanie

13. *Jan zapomniał zamknąć drzwi.*

Można oczywiście zdać sprawę z tej różnicy poprzez nadanie różnych sensów różnym konstrukcjom syntaktycznym, ale to nie oddaje jednocześnie więzi znaczeniowej między użyciami *zapomnieć* a zdanie z niej sprawy jest jednocześnie ambicją kognitywisty.

1.3. Nie jest zatem tradycyjne ujęcie leksykografii komputacyjnej zadawające dla badaczy z nastawieniem kognitywistycznym, ponieważ nie jest ono w stanie opisać (przewidzieć) wszystkich „sensów” danego słowa poprzez skończone ich wyliczanie a także nie jest w stanie wydobyć interesujących generalizacji między „sensami” danego słowa.

2. Proponowane zwykle rozwiązanie tego problemu polega na wprowadzeniu bogatszego zestawu wyróżników semantycznych dla wyboru uzupełnień, niż to, które jest dostarczane konwencjonalnie przez mechanizm restrykcji selekcyjnych.

Wprowadzany jest zestaw różnych poziomów opisu znaczenia:

— *struktura predykatowo-argumentowa*

— *struktura zdarzeniowa* (stan (S), proces (P), zdarzenie / tranzycja (T)); „fokus” zdarzenia (perspektywa percepcji zdarzenia (cf. np. Allen (1983), Desclés (1985, 1990a, 1990b, 1991, 1993))

— „*frame*” / *cadre* / *qualia structure* ((cf. p.ex. Minsky (1975, 1968), Schank (1972, 1977), Wilks (1973), Rumelhart (1980), Fillmore (1985), Pustejowsky, Boguraev (1993), czyli prototypowe predykaty, relacje, obiekty łączone z pojęciem wyrażanym przez dane słowo (cf. p.ex. Апресян (1974), Апресян, Мielczuk, Zolkowski (1969), Мielczuk (1974, 1986), Мielczuk, Zolkowski (1969, 1970), w tym m.in. tzw. „*verbes appropriés*” czasowniki reprezentujące pojęcie ogólne, którego forma językowa zależy od zwyczaju językowego i pojęcia, któremu towarzyszy, typu *opatulić szalem* także widziane od strony rzeczowników i tworzące tzw. *klasy obiektów* (cf. np. G. Gross (1992))

— *strukturę dziedziczenia leksykalnego* (określenie sposobu, w jaki dane słowo jest połączone z innymi w słowniku, hiponimia, hiperonimia, etc. (cf. p.ex. Woods (1975), Fahlman (1979), Touretzky (1984), Grégoire (1990)).

2.1. Przedstawimy poniżej pokrótce dwa ostatnie elementy w ujęciu Pustejowsky’ego, Boguraev’a (np. 1993, 1994) rozwijających tzw. *Teorię Słownika Generatywnego (Generative Theory Lexicon)*.

Qualia structure w wydaniu J. Pustejowsky i B. Boguraev (np. 1993) specyfikują cztery aspekty definicji słowa:

— relacje między obiektem a jego częściami konstytutywnymi (rola *Konstytutywna*) (CONST)

— relację obiektu do innych w ramach większego zbioru (rola *Formalna*) (FORMAL)

— cel i funkcję (rola *Teliczna*) (TELIC)

— czynniki biorące udział w jego powstawaniu lub tworzeniu (rola *Agentywna*) (AGENTIVE).

Por. np. Pustejowsky, Boguraev (1993: 205):

14a. *Jan przeszedł przez drzwi*

14b. *Jan pomalował drzwi na biało*

15a. *Czarny dym wypełnił pokój*

15b. *Pokój Jana znajdował się obok.*

W tym ujęciu powiedziano by się, że znaczenie obiektów może być funkcją relacji między różnymi rolami w strukturze, np. *Konstytutywną* i *Formalną*.

Por. np.:

drzwi (x,y)

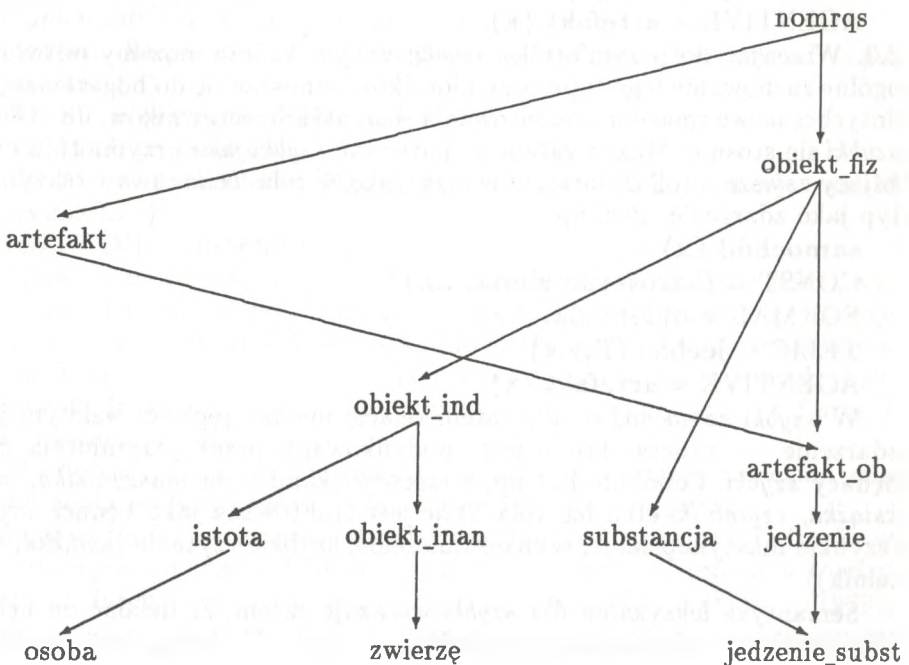
CONST = **otwór (x)**

FORMAL = **obiekt_fiz. (x)**

TELIC = **przechodzić_przez (P,a,y)**

AGENTIVE = **artefakt (x)**.

2.2. Strukturę typu *qualia* można też przedstawić jako zbiór restrykcji na typy reprezentujące bazę wiedzy leksykalnej (cf. np. Copestake (1992), Copestake, Briscoe (1992)):



Powyższa struktura przedstawia fragment hierarchii typów dla wiedzy o obiektach, kodujący „qualia structure”. Por. np. restrykcje na typy **artefakt** i **obiekt_fiz.**

artefakt (x)

TELIC = Pred (E,y,x)

obiekt_fiz.

FORMAL = **formafiz. (x)**

STAN-FIZYCZNY = **stały.**

Restrykcja na **artefakt_ob** dziedziczy informację z obu swych „przodków”: **artefakt** i **obiekt_fiz.**, por. np.:

artefakt_ob (x)

FORMAL = **formafiz. (x)**

STAN-FIZYCZNY = **stały**

TELIC = Pred (E,y,x).

Widać zatem wyraźnie dlaczego i jak się to dzieje, że np. zdania typu *Jan skończył kanapkę* otrzymują normalną interpretację *par défaut* typu: rzeczownik *kanapka* zawiera informację o „czynności jedzenia” jako restrykcję na wartość roli *Telic*, co jest związane z jego pozycją w strukturze typów, jako że **jeść (P,a,x)** określa proces, **P**, między jakimś indywiduum, **a**, a obiektem fizycznym **x**.

kanapka

CONST = (chleb,)

FORMAL = **obiekt_fiz.**

TELIC = **jeść (P,a,x)**

AGENTIVE = **artefakt (x)**.

2.3. Wracając do przymiotnika *szybki*: w tym kadrze możemy odtworzyć ogólne zachowanie tego typu przymiotników odnosząc się do bogatszego, niż dotychczasowe sposobu prezentowania semantyki rzeczowników, do których *szybki* się stosuje. Można zatem rozpatrywać *szybki* jako przymiotnik orzekający zawsze o roli *Telic* rzeczownika, jako że rola *Telic* zawsze otrzymuje typ jako zdarzenie. Por. np.:

samochód (x)

CONST = (karoseria, silnik,)

FORMAL = **obiekt_fiz.**

TELIC = **jechać (P,y,x)**

AGENTIVE = **artefakt (x)**.

W *szybki samochód* mamy zatem relację między **jechać**, wziętym jako zdarzenie — proces, który jest modyfikowany przez przymiotnik jako będący *szybki*. Podobnie jest np. z rzeczownikami typu *maszynistka*, *walc*, *książka*, *czytelnik*, etc.: ich rola *Telic* jest traktowana jako będąca *szybką* (szybkie maszynopisanie, szybkie tańczenie, szybkie czytanie (książka, czytelnik)).

Semantyka leksykalna dla *szybki* wskazuje zatem, że działać on będzie jako predykat zdarzeniowy, modyfikując czynność, która jest wartością

dla roli *Telic* rzeczownika (a więc, jeszcze inaczej, jako modyfikator czasownika „approprié” dla danego rzeczownika, wyrażany przymiotnikowo, wtedy, kiedy pojawia się na powierzchni — konceptualnie bowiem relacja jest *omniprésente* — rzeczownik, ale nie czasownik).

Z jednej zatem strony widzimy, że nie musimy dawać wielu różnych sensów dla *szybki*, a z drugiej strony, że ewentualne „nowe” znaczenia dla danego przymiotnika będą generowane, będą funkcją, współgry roli *Telic* dla rzeczownika, który ma być modyfikowany. Por. np. *szybka droga*: rola *Telic* dla *jezdni* specyfikuje jej cel i ten cel jest traktowany jako proces, który jest interpretowany jako *szybki*:

[*Telic*: *jeździć* (P, samochody) & po (P,x)].

2.4. Tego typu ujęcie pozwala nam oczywiście nie traktować jako odrębne sensory dla czasowników np. *zabić*, *zacząć*, *obudzić* realizacji tych czasowników wymagających różnych kategorii morfosyntaktycznych, co jest wymogiem normalnym dla tradycyjnej leksykografii, nie tylko komputacyjnej (por. np. G. Gross (1988, 1989), Gross, Dugas (1991), M. Gross (1975, 1989, 1990)), np.:

Jan zabił Marię (strzałem z pistoletu)

Kula zabiła Marię

Pistolet zabił Marię

Jan zaczął książkę

Jan zaczął czytać książki

Jan zaczął czytać książkę

Jan obudził Marię (dając jej do picia filiżankę kawy)

Filiżanka kawy obudziła Marię

Wypicie filiżanki kawy obudziło Marię.

I tak np. dla *pistolet* mielibyśmy *frame* czy też „qualia structure” typu:

pistolet (x)

CONST = (cyngiel, kula,)

FORMAL = obiekt_fiz.

TELIC = (strzelać (P,y,x,z), trafić (T,y,v,z,x): zabić (T,y,v,z,x)....)

AGENTIVE = artefakt (x),

co tłumaczy możliwość użycia jednego z elementów roli *Telic* jako „wypełniacza” dla pozycji przy *zabijać*: predykat *zabijać* wymaga, by na jego pierwszej pozycji argumentowej pojawiło się konceptualnie zdarzenie; ponieważ fraza nominalna *kula* nie spełnia tego kryterium, czasownik przekształca tę frazę w denotat zdarzeniowy, który jest dostępny z *frame* / „qualia structure” dla *pistolet*.

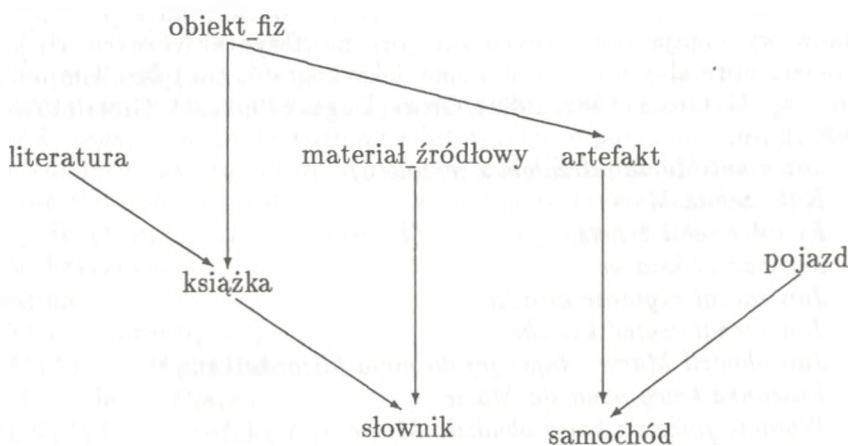
2.5. Na koniec tej zwięzłej prezentacji, dwa słowa jeszcze o mechanizmie dziedziczenia semantycznego w *Teorii Słownika Generatywnego*.

Wprowadzenie mechanizmu dziedziczenia semantycznego, już klasyka w komputacyjnych badaniach leksykalnych i systemach reprezentacji wiedzy w ramach Inteligencji Sztucznej (cf. np. Woods (1975), Fahlman (1979), Touretzky (1984), Grégorie (1990)), pozwala m.in. zwiększyć moc systemów

analizy tekstów, które dokonują syntaktycznej i semantycznej interpretacji poprzez odniesienie się do szczegółowej dekompozycji znaczenia.

Jedną z konsekwencji postulowania struktur typu *frames* czy *qualia structure* jest konieczność dysponowania, nałożonymi na leksykę, wieloma poziomami strukturalnymi.

Klasyczne podejście do hierarchizacji konceptualnych stosuje tzw. *dziedziczenie wielorakie* (*multiple inheritance*). I tak np. *książka jest literaturą*, równie dobrze jak i *obiektem_fiz.*; *słownik jest obiektem_fiz.* równie dobrze jak i *materiałem źródłowym* (dziełem do konsultowania); por. np. *konsultować słownik* ale nie bardzo: **konsultować książkę*; *samochód* jest tak *artefaktem* jak i *pojazdem*, etc., co można przedstawić w sposób następujący:



Pustejowsky, Boguraev (1993: 213–6) proponują zastosowanie oddzielnej struktury dla każdej roli ze struktury *qualia*. Dziedziczenie za pośrednictwem *qualia* sprowadzałoby się do relacji następujących w przypadku analizowanego przykładu (Pustejowsky, Boguraev (1993: 214):

Książka jest_formal obiektem_fiz.

książka jest_telic literatura

książka jest_agent literatura

słownik jest_formal książka

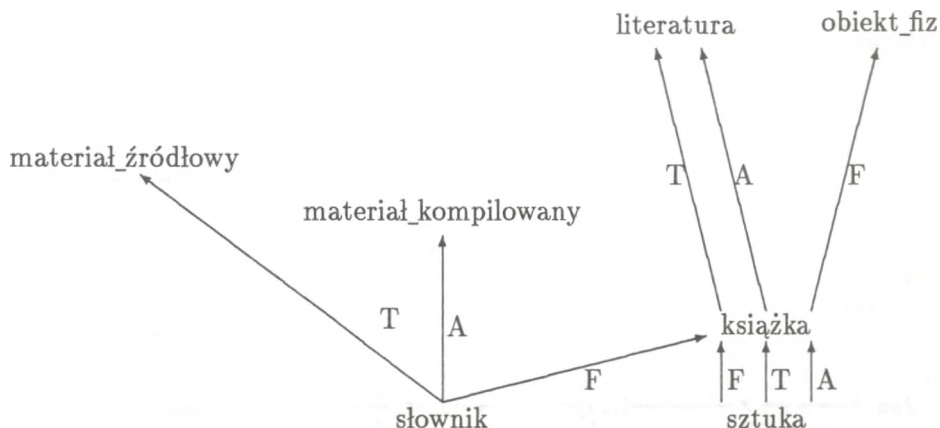
słownik jest_telic materiał źródłowy

słownik jest_agent materiał_kompilowany

sztuka jest_agent literatura

sztuka jest_telic książka,

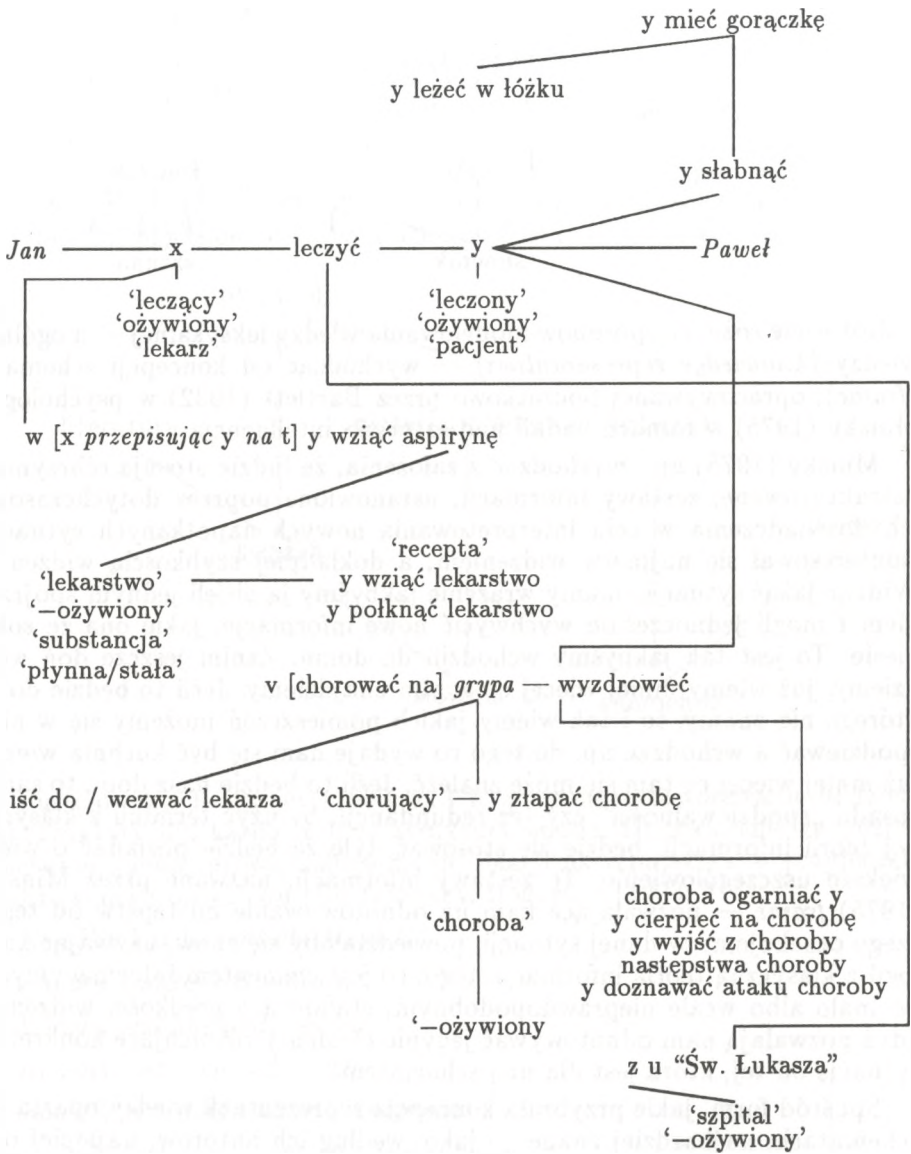
co można przedstawić w postaci grafu (Pustejowsky, Boguraev (1993: 215):



3. Jest wiele różnych sposobów modelowania wiedzy leksykalnej — a ogólnie wiedzy (*knowledge representation*) — wychodząc od koncepcji schematu (*frame*), opracowywanej początkowo przez Bartlett (1932) w psychologii, Minsky (1975) w ramach badań nad sztuczną inteligencją.

Minsky (1975) np., wychodząc z założenia, że ludzie stosują olbrzymie, ustrukturowane, zestawy informacji, ustanowione poprzez dotychczasowe ich doświadczenia w celu interpretowania nowych napotkanych sytuacji, zainteresował się najpierw widzeniem, a dokładniej szybkością widzenia. Widząc jakąś sytuację, mamy wrażenie jakbyśmy ją objęli jednym spojrzeniem i mogli jednocześnie wychwycić nowe informacje, jakie ona ze sobą niesie. To jest tak jakbyśmy wchodzili do domu. Zanim jeszcze doń wejdziemy, już wiemy mniej więcej co w nim znajdziemy. Jeśli to będzie dom, którego nie znamy, to i tak wiemy jakich pomieszczeń możemy się w nim spodziewać a wchodząc np. do tego co wydaje nam się być kuchnią wiemy już mniej więcej co tam się może znaleźć. Jeśli to będzie nasz dom, to sama zasada „spodziewalności” czy też redundancji, by użyć terminu z klasycznej teorii informacji, będzie się stosować, tyle że będzie posiadać o wiele większe uszczegółowienie. Te zestawy informacji, nazwane przez Minsky (1975) *frame* — pozwalające nam na odnotowywanie odstępstw od tego, czego oczekujemy w danej sytuacji, powiedziałyby się znów, używając analogii z klasyczną teorią informacji, tego, co jest elementem informacyjnym, bo mało albo wcale nieprawdopodobnym, stanowią o prędkości widzenia, gdyż pozwalają nam odnotowywać jedynie elementy różnicujące konkretną sytuację od tej, która jest dla niej schematem.

Spośród form, jakie przybrała koncepcja reprezentacji wiedzy oparta na schematach, najbardziej znane — jako, według ich autorów, najlepiej oddające idee analizy w kategorii schematów, bo przecież nie jedynie: żeby wspomnieć w tym kontekście np. KL-ONE R.J. Brachmana i J. Schmolza (1985) i in. — wydają się być te, które zostały opracowane przez Charniaka (1977, 1978), Abelsona (1975) i Schanka (1975, 1977). Fundamentalne



Rys. 1

założenia tego typu analizy leksykalnej — i nie tylko — można sprowadzić generalnie do charakterystyk następujących, muszą one posiadać:

- graf pojęciowy
- opis dołączony do każdego z pojęć, w postaci zbioru elementów istotnych
- atrybut (wartość)
- eksplikację aspektów proceduralnych w reprezentacji deklaratywnej.

3.1. Np. dla predykatów oznaczających sytuacje, których częścią jest sytuacja *leczenia* można np. zaproponować schemat następujący to znaczy także, że sama sytuacja *leczenia* jest częścią ogólniejszego scenariusza, *scriptu* (np. Schank (1977), Gioia, Poole (1984), Abbott, Black, Smith (1985), Fayol, Monteil (1988)), a generalnie znaczy tyle, że każda sytuacja jest potencjalnie częścią jakiegoś *scriptu* (patrz rys. 1), gdzie elementy między '.....' są traktowane jako atrybuty a elementy pisane kursywą są wypełnieniami / wartościami atrybutów (zauważmy, że z punktu widzenia relacji *leczyć* do innych czynności, *leczyć* byłoby też wypełnieniem).

3.2. Przedstawiona siatka pojęciowa może otrzymać postać bądź grafu bądź notacji nawiasowej, w stylu m.in. Mielczuk (1974), Mielczuk, Zolkowski (1970), z opisaniem przedstawionych relacji jako odpowiednich parametrów bądź zamian leksykalnych. Por. np. rys. 2, gdzie (RS) — rola semantyczna, a — agentywna, p — patientywna, i — instrumentalna, (rs) — restrykcje selekcyjne a [] wyrażają typowe relacje między elementami grafu, stąd też, brane poniżej jako relacje nie są umieszczone w '.....', co oznaczało przy zapisie poprzednim rodzaj restrykcji selekcyjnej i mogło się tam czasem zdarzyć, że status jednego i tego samego wyrażenia mógł być dwuznaczny (por. np. 'choroba' i [SO] *choroba*).

[Real2] *leczyć* (*x, y, v, w, z*)

(RSa) '*leczący*' (*x*)

(rs) '*ożywiony*' (*x*)

[Si1] *lekarz* (*x*)

[Real2] (*lekarz*) = *leczyć*

[Si1] *lekarstwo* (*x*)

[Sinst] *lekarstwo* (*w*)

[Real2] (*lekarstwo*) = *leczyć*

[Subst] (*lekarstwo*) = *substancja stała/płynna*

[Perf] [Real] (*lekarstwo*) = *wyleczyć*

[Rorm] (*lekarstwo, subst. stał*) = *tabletki*

[Real2] (RSi) '*przepisując wzięcie lekarstwa*'

(rs) '*ożywiony*' (*w*)

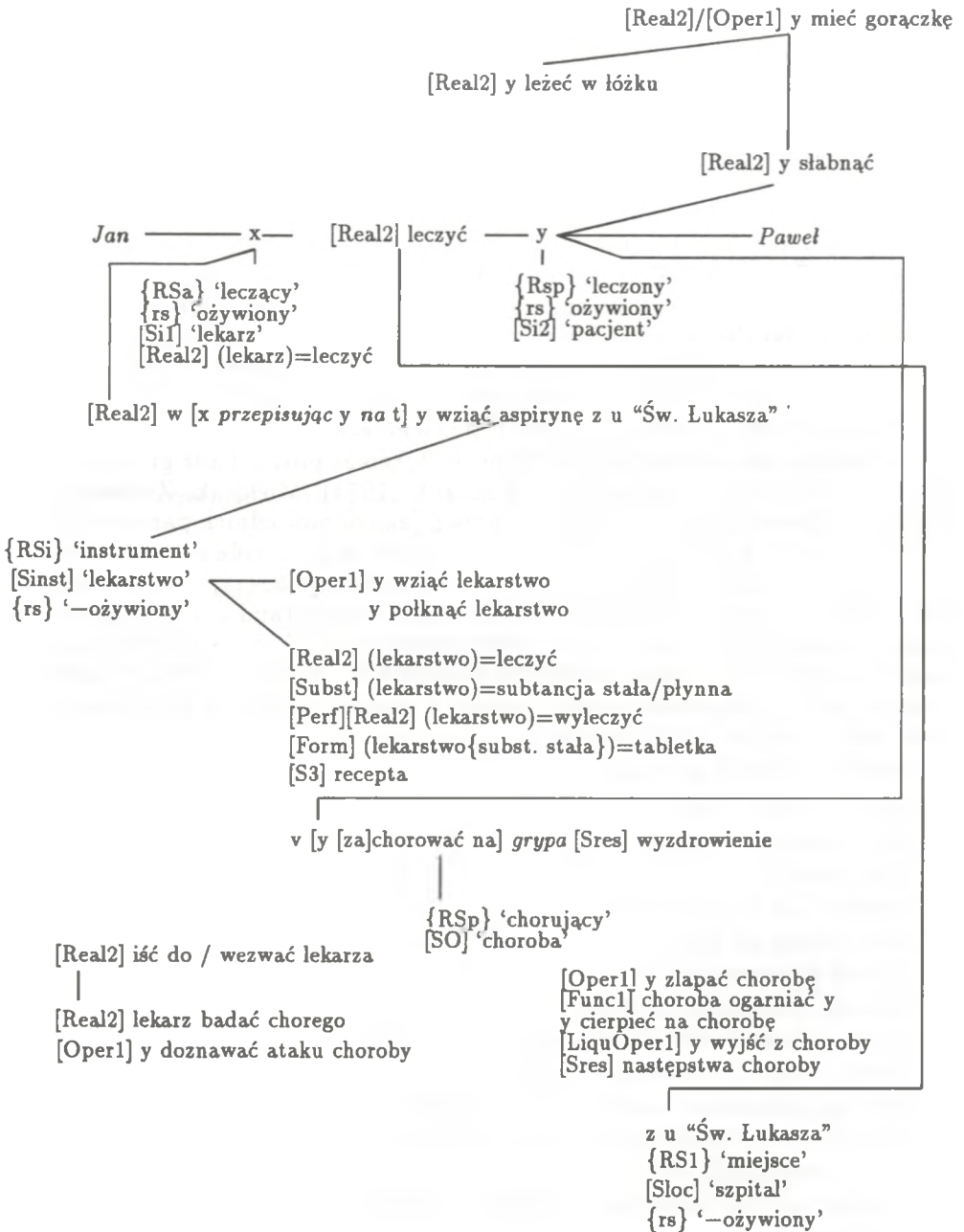
[Oper1] *y wziąć lekarstwo, y połączyć lekarstwo* (*w*)

[S3] *recepta* (*w*)

(RSp) '*leczoney*' (*y*)

(rs) '*ożywiony*' (*y*)

[Si2] *pacjent, chory*



Rys. 2

- (RSp) 'chorujący na' (v)
 (rs) '- ożywiony' (v)
 [SO] choroba (v)
 [Oper1] y złapać chorobę (v)
 [Func1] choroba ogarniać y (v)
 [Oper1] y doznawać ataku choroby (v), y cierpieć na chorobę (v)
 [LiquOper1] y wyjść z choroby (v)
 [Sres] następstwa choroby (v)
 [Real2]/[Oper1] y mieć gorączkę
 [real2] y leżeć w łóżku
 [real2] y słabnąć
 [Real2] iść do / wezwać lekarza
 [Real2] lekarz bada chorego
 (RS1) 'miejsce'
 (rs) '- ożywiony' (z)
 [Sloc] szpital, przychodnia, gabinet (z)
 [Sres] wyzdrowienie
 [Real1] wyzdrowieć
 Jan (x) [x przepisując y [y wziąć]] aspirynę (w)
 Paweł (y)
 [y chorujący na] grypa (v)
 u „Św. Łukasza” (z).

3.3. jak to już zauważyliśmy na początku omawiania proponowanego opisu, prezentowana koncepcja opisu znaczenia leksykalnego zawiera jednocześnie elementy *scriptalne* à la Schank (1977), wyrażające sekwencje czynności wykonywanych typowo w danych sytuacjach, które u Mielczuka (1974), Apresjana (1980) np. miały jeszcze nieśmiały wyraz jako funkcja *Real1/2*. Otrzymała ona charakterystykę typu „spełniać wymagania zawarte w znaczeniu nazwy sytuacji” (cf. np. Apresjan 1980: 69). Charakterystyka ta sama w sobie byłaby niepełna, gdyby nie to, że — jak podają autorzy w uzupełnieniu tego opisu (cf. Apresjan 1980: 69, n.32) — „...Real tworzy się od wyrazów dwu klas. Do znaczenia wyrazów pierwszej klasy wchodzi sensy pokrywające się z modalnym 'chcieć', 'móc', 'powinien' i pochodne od nich. Do pochodnych tych sensów należą np. 'cel', 'przeznaczenie', itp., z czego wynika, że *real-e* można tworzyć od dowolnych rzeczowników o znaczeniu narzędzia, urządzenia, środka i innych artefaktów (*Real* (nóż) = *krajać*, *Real* (pociąg) = *jechać*)”. W tym sensie, jeśli *powinien* traktować jako predykat wyrażający, w przypadku interesujących nas tutaj *scriptów*, że pewna sytuacja **powinna** się pojawić wraz czy po jakiejś innej sytuacji przy typowej realizacji pierwszej, a więc określona przewidywalność sytuacji w ciągu, ta część *Real-u* byłaby odpowiedzialna za *scriptalną* część zapisu hasła (por. np. *zachorować* — *mieć gorączkę* — *osłabnąć* — *leżeć w łóżku* — *iść/wezwać lekarza* — *być zbadanym przez lekarza* — *leczyć się* — *wyzdrowieć*). Stąd też i wiele różnych realizacji dla tej funkcji, które zależą

od tego, który z elementów *scriptu* / części grafu brany jest jako punkt wyjścia dla funkcji *Real* i rodzaj zamkniętego koła relacji tego typu: każda z nich może bowiem być realizacją, o czym wspominaliśmy wyżej, funkcji *Real* dla innej relacji, nawet jeśli nie jest ona wyspecyfikowana w danym grafie pojęciowym.

Opis znaczenia jest naturalnie uzupełniony o realizacje funkcji typu *IS-A*, *IS-NOT-A*, która tutaj ma postać rozwiniętą i uszczegółowioną, by móc stawić czoła problemom związanym z *dziedziczeniem leksykalnym*, rekurentnie stosowanej funkcji *Gener* (Mielczuk (1974), Apresjan (1980)), typu np. *dolegliwość — grypa — choroba zakaźna — choroba; niedomagający — chorujący — istota ożywiona — człowiek; lekarstwo — ciało stałe/płynne — obiekt fizyczny — artefakt — materia*, etc.

3.4. Jesliby porównywać moc opisową i eksplikacyjną modelu *Teorii Słownika Generatywnego* Pustejowsky'ego, Boguraev'a (np. 1993) i opracowanego przez nas aktualnie modelu opartego na koncepcjach Minsky (1975, 1986), Schank (1977), Mielczuk (1974), Desclés (1990a) to dwa typy uwag się tutaj narzucają.

Pierwszy dotyczy opisu i wyjaśnienia problemów nękających leksykoграфię komputacyjną typu „listing” z perspektywy Pustejowsky'ego i Boguraev'a a drugi dotyczy kwestii proponowanego opisu dziedziczenia semantycznego.

W uwagach poniższych omówimy tylko kwestie pierwszego typu; kwestie drugiego typu omówione zostaną później.

Z naszego punktu widzenia, łatwo jest przewidzieć wspólnotę znaczenia dla analizowanego przez Pustejowsky'ego, Boguraev'a (np. 1993) typu przymiotnikowego: to, co w ich systemie jest funkcją *Teliczną* nie jest niczym innym jak jednym z przejawów funkcji *Real*: tak jak np. *Real (nóż) = krajać* tak *Real (drzwi) = przechodzić*, *Real (kanapka) = jeść*, *Real (garaż [naprawczy]) = naprawiać samochody*, *Real (droga) = jeździć po*, etc.

Z drugiej strony, możliwość używania słów, np. *drzwi*, w różnych kontekstach typu np. cytowane *malować* czy *przechodzić*, jest funkcją wzięcia pod uwagę innego parametru leksykalnego, bądź *Real* bądź *Subst*, stanowiącego część opisu znaczenia słowa, np. *drzwi*, i w tym sensie, oczywiście, oba modele — nie będąc bynajmniej identycznymi — stosują paralelny mechanizm eksplikacyjny.

Model Pustejowsky'ego, Boguraev'a (np. 1993) musi także mieć gdzieś miejsce, którego w *qualia structure* przynajmniej na razie nie widać, dla składnika typu *scriptalnego*. Nie dlatego oczywiście, że każdy opis taki składnik mieć musi — to bowiem zależy przede wszystkim od celu, jaki sobie dany model stawia, choć możemy powiedzieć, że te, które mają większe ambicje opisowe i eksplikacyjne składnik taki posiadać powinny — ale dlatego, że sposób wyjaśniania, także ten, który stosuje Pustejowsky, Boguraev (1993), możliwości pojawienia się alternatywnego różnych kategorii gramatycznych na określonych pozycjach semantyczno-syntaktycznych predykatu, aby nie doprowadzić do „listingu” znaczeń, musi zawierać gdzieś nie

tylko formalną możliwość *type coercion*, o której mówią Pustejowsky, Boguraev (1993: 209–210), ale przede wszystkim opis relacji dopuszczających możliwość *type coercion*.

Z perspektywy prezentowanego tu modelu opisowego rozwiązanie raz jeszcze znajduje się m.in. w parametrze *Real*, który realizuje ciąg *scriptalny*, np. *Real (pistolet) = strzelać*, *Real (strzelać) = trafiać*, *Real (Real (strzelać)) = ranić/zabijać*, etc., oraz parametrach *S_o*, który realizuje derywaty syntaktyczne, np. *S_o (strzelać) = strzał*, *S_o Real (Real (strzelać)) = ranienie/zabicie*, etc., *S_{instr}*, realizujący typową nazwę narzędzia, sytuacji, np. *S_{instr} (strzelać) = pistolet/karabin/tuk*, etc., a także *Pars*, który realizuje typowe części składowe czegoś, np. *Pars (pistolet) = kula/cyngiel/lufa....*, etc.

Jakis tego typu mechanizm musi zapewne zakładać także model *Teorii Słownika Generatywnego*, bo inaczej nie byłby w stanie wyjaśnić — możliwość formalnego przeprowadzenia operacji w ramach *type coercion* jest jedynie wykazaniem poprawności formalnej samej operacji, ale jeszcze nie jej opisaniem i wyjaśnieniem — możliwości zamian typu cytowanych powyżej: *Jan strzałem z pistoletu zabił Pawła / Jan zabił Pawła / Strzał Jana z pistoletu zabił Pawła / Pistolet (Jana) zabił Pawła / Kula (wysstrzelona) z pistoletu Jana (przez Jana) zabiła Pawła*.

3.5. Tytułem konkluzji możemy stwierdzić, że proponowany przez Pustejowsky'ego i Boguraev'a opis słownictwa w ramach tzw. *Teorii Słownika Generatywnego* jest propozycją interesującą, pozwalającą przewyżczyć określone — przynajmniej z perspektywy kognitywnej leksykografii komputerowej — mankamenty tradycyjnych w komputerowych opracowaniach leksykograficznych ujęć znaczenia słów. Jednocześnie okazuje się, że ten sam cel osiągnąć można — po dopracowaniu i rozwinięciu składnika odpowiedzialnego za dziedziczenie leksykalne — w ramach opracowywanego aktualnie modelu opisu słownictwa integrującego elementy koncepcji Minsky (1975, 1986), Schank, Abelson (1977), Mielczuk (1974), Desclés (1990a), który zdaje się mieć większą moc opisową i eksplikacyjną.

Summary

The article briefly discusses the conception behind the *Generative Dictionary* developed by J. Pustejowsky and B. Boguraev. One of the main objectives is to overcome the listing of the meanings of an item, which has been the classic approach taken in computational lexical analyses.

The proposed approach attempts to put forward a cognitive solution of the polysemy of lexical items and takes account of the interrelations obtaining between various levels of semantic description, including the predicate-argument structure, event structure, qualia structure (roughly the equivalent of "frame"), and the structure of lexical inheritance.

The main focus in the approach is on qualia structure, whose components, e.g. telic and agentive functions, can account for the creative use of words representing artefacts (e.g. *szybki* in *szybki garaż*, *szybki samochód*, *szybka maszynistka*, etc.).

Bibliografia

- Abbott V., Black J.B., Smith E.E. 1985. The Representation of Scripts in memory, *Journal of Memory and Language*, 24.
- Abelson R. 1975. Representing Mundane Reality in Plans, in: Bobrow, Collins (eds).
- Apresjan J. 1974. Leksiczeskaja Semantika. Sinonimiczeskije sredstva jazyka, Nauka, Moskva. (Tlum. pol., Semantyka Leksykalna. Synonimiczne środki języka, 1980, Ossolineum, Wrocław).
- Apresjan J., Mielczuk I., Zolkovskij A.K. 1969. Semantics and Lexicography: Towards a New Type of Unilingual Dictionary, in: Kiefer (ed.).
- Allen J. 1983. Maintaining knowledge about temporal intervals, *Communications ACM* 26.
- Banyś W. 1986. Implications actualisationnelles des prédicats et structure thème-rhème. A propos de „si-verbos” positifs à la Karttunen, *Linguistica Silesiana*, 8.
- Bartlett F. 1932. Remembering: A Study in Experimental and Social Psychology, New York and London, Cambridge Univ. Press.
- Bates L., Weischedel R. (eds), 1992. Challenges in Natural Language Processing, Cambridge University Press, Cambridge.
- Bierwisch M., Heidolph K. (eds), 1970. Progress in Linguistics, The Hague, Mouton.
- Bobrow D., Collins A. (eds), 1975. Representation and Understanding: Studies in Cognitive Science, Academic Press, New York.
- Boguraev B., Briscoe E. 1989. Computational Lexicography for Natural Language Processing, Longman, London.
- Boguraev B. 1992. Building the lexicon: the contribution of computational lexicography, in: Bates, Weischedel (eds).
- Briscoe E., Copestake A., Boguraev B. 1990. Enjoy the Paper: lexical semantics via lexicology, Proceedings Thirteenth International Conference on Computational Linguistics, Helsinki.
- Brachman R.J., Schmolze J. 1985. An overview of the KL-ONE knowledge representation system, *Cognitive Sciences*, 9.
- Charniak E. 1977. A frame for PAINTING: the representation of a common sense knowledge fragment, *Cognitive Sciences*, 1 (4).
- Charniak E. 1978. On the use of frame knowledge for language comprehension, *Artificial Intelligence*, 11.
- Copestake A., 1992. The ACQUILEX LKB: representation issues in semi-automatic acquisition of large lexicons, Proceedings of Third Conference on Applied Natural Language Processing, Trento.

- Copestake A., Briscoe E., 1992. Lexical Operations in a unification-based framework, in Pustejovsky, Bergler (eds).
- Desclés J.-P. 1985. Representation des connaissances: archétypes cognitifs, schèmes conceptuels, schémas grammaticaux, Actes Sémiotiques, Documents VII, Paris, CNRS.
- Desclés J.-P. 1990a. Langages applicatifs, langues artificielles et cognition, Paris, Hermès.
- Desclés J.-P. 1990b. Les représentations intermédiaires, Revue de Synthèse, 1.2, Sciences cognitives: quelques aspects problématiques, Paris, Albin Michel.
- Desclés J.-P. 1991. La prédication opérée par les langues (ou à propos de l'interaction entre langage et perception), Langages, 103.
- Desclés J.-P. 1993. Interactions entre langage, perception et action, Faits de Langues, 1.
- Fahlman S.E. 1979. NETL: A System for Representing and Using Real-World Knowledge, MIT Press, Cambridge.
- Fayol M., Monteil J.-M. 1988. The Notion of Script: From general to developmental and Social Psychology, Cahiers de Psychologie Cognitive. European Bulletin of Cognitive Psychology, vol. 8 (4).
- Fillmore C. 1985. Frames and the Semantics of Understanding, Quaderni di Semantica, 6.
- Gioia D.A., Poole P.P. 1984. Scripts in behavior, Academy of Management Review, 9.
- Grégoire E. 1990. Logiques non-monotones et intelligence artificielle, Hermès, Paris.
- Gross G. 1988. Degré de figement des noms composés, Langages, 90.
- Gross G. 1989. Les constructions converses du français, Genève, Droz.
- Gross G. 1992. Forme d'un dictionnaire électronique, miméo.
- Gross G., Dugas A. 1991. Analyse des groupes N de N, Actes du Colloque Informatique et Langue Naturelle (23-24.01.1991), Presses Universitaires de Nantes, Nantes.
- Gross M. 1975. Méthodes en Syntaxe, Paris, Hermann.
- Gross M. 1989. La construction de dictionnaires électroniques, Annales des Télécommunications, 44 (1-2).
- Gross M. 1990. Grammaire Transformationnelle du français. 3-syntaxe de l'adverbe, Paris, ASSTRIL.
- Karttunen L. 1973a. La Logique des constructions anglaises ... complément prédicatif, Langages, 30.
- Karttunen L. 1973b. Presuppositions of Compound Sentences, Linguistic Inquiry, 4.2.
- Kiefer F., (ed.), 1969. Studies in Syntax and Semantics, Elsevier, Dordrecht.
- Kiparsky P., Kiparsky C. 1970. Fact, in: Bierwisch, Heidolph (eds).
- Mielczuk I. 1974. Opyt teorii lingwistycznych modelej „Smysł-Tekst”, Nauka, Moskwa.
- Mielczuk I., Zolkovskij A.K. 1970. Towards a functioning “meaning-tekst” model of language, Linguistics, 57.

- Mielczuk I. 1986. Dictionnaire explicatif et combinatoire du français contemporain, Presses de l'Université de Montréal, Montréal.
- Mielczuk I., Zolkovskij A.K. 1969. K postrojeniju dejstvujuszej modeli jazyka „Smysł-Tekst”, Maszynnyj Perevod i Prikładnaja Lingvistika. 11.
- Minsky M. 1975. A Framework for Representing Knowledge, in: Winston (ed.).
- Minsky M. 1986. The Society of Mind, New York, Simon & Schuster.
- Pustejowsky J., Boguraev B. 1993. Lexical Knowledge Representation and Natural Language Processing, Artificial Intelligence, 63.
- Pustejowsky J., Bergler S. (eds), 1992. Lexical Semantics and Knowledge Representation, Springer, New York.
- Rumelhart D.E. 1980. Schemata: The Building Blocks of Cognition, in: Spiro, Bruce, Brewer (eds).
- Schank R. 1972. Conceptual Dependency: a theory of natural language understanding, Cognitive Psychology, vol. 3, (4).
- Schank R. 1975. The Structure of Episodes in memory, in: Bobrow, Collins (eds).
- Schank R., Abelson R. 1977. Scripts, Plans, Goals and Understanding, Hillsdale, N.J., Lawrence, Erlbaum.
- Schank R., Colby F. (eds.), 1973. Computer Models of Thought and Language, Freeman, San Francisco.
- Spiro R.J., Bruce B.C., Brewer W.F. (eds), 1980. Theoretical Issues in Reading Comprehension, Hillsdale, New York, Erlbaum.
- Touretzky D.S. 1984. Implicit Ordering of Defaults in Inheritance Systems, Proceedings of the National Conference on Artificial Intelligence (AAAI-84).
- Wilks Y. 1973. An artificial Intelligence approach to machine translation, in: Schank, Colby (eds).
- Winston P.H. (ed.), 1975. The Psychology of Computer Vision, McGraw-Hill, New York.
- Woods W.A. 1975. What's in a link: foundations for semantic networks, in: Bobrow, Collins (eds).