

Bożenna Bojar

**JĘZYKOZNAWSTWO
DLA STUDENTÓW
INFORMACJI NAUKOWEJ**



NAUKA-DYDAKTYKA-PRAKTYKA

JĘZYKOZNAWSTWO

Polish Librarians Association
SCIENCE-DIDACTICS-PRACTICE

Bożenna Bojar

LINGUISTICS
FOR STUDENTS OF
SCIENTIFIC INFORMATION

WYDAWNICTWO

SBP



Warsaw 2005

Stowarzyszenie Bibliotekarzy Polskich
NAUKA-DYDAKTYKA-PRAKTYKA

Bożenna Bojar

JĘZYKOZNAWSTWO
DLA STUDENTÓW
INFORMACJI NAUKOWEJ



Warszawa 2005

Komitet Redakcyjny serii wydawniczej
<<NAUKA – DYDAKTYKA – PRAKTYKA>>

Marcin DRZEWIECKI (przewodniczący), Stanisław CZAJKA, Artur JAZDON,
Barbara SOSIŃSKA-KALATA, Danuta KONIECZNA, Krzysztof MIGOŃ, Mieczysław
MURASZKIEWICZ, Janusz NOWICKI (sekretarz), Joanna PAPUZIŃSKA-BEKSIĄK,
Wanda PINDŁOWA, Jadwiga SADOWSKA, Jan SÓJKA, Barbara STEFANIAK,
Elżbieta STEFAŃCZYK, Hanna TADEUSIEWICZ, Zbigniew ŻMIGRODZKI

Recenzenci:

Małgorzata KORYTKOWSKA
Wiesław BABIK

Redakcja techniczna i korekta:
Jarosław BOJAR

© Copyright by Stowarzyszenie Bibliotekarzy Polskich

ISBN 83-89316-35-8

CIP – Biblioteka Narodowa

Bojar, Bożenna

Językoznawstwo : dla studentów informacji naukowej / Bożenna Bojar ; Stowarzyszenie
Bibliotekarzy Polskich. – Warszawa : Wydaw. SBP, 2005. – (Nauka, Dydaktyka, Praktyka ; 74)

Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich
00-335 Warszawa, ul. Konopczyńskiego 5/7, tel. 827-52-96
Warszawa 2005. Wydanie I. Ark. wyd. 30. Ark. druk. 23,25
Skład i łamanie: Jarosław BOJAR
Druk i oprawa: Podlaska Spółdzielnia Producyjno-Handlowo-Usługowa
15-182 Białystok, ul. 27 Lipca 40/3, tel./fax (0-85) 675-48-02

Spis treści

Od autorki	7
I. Informacja	9
II. Znaki	22
III. Systemy znakowe. Kody i języki	33
IV. Cechy systemów znakowych	56
V. Funkcje systemów znakowych	65
VI. Metainformacja, metatekst, hipertekst	79
VII. System fonologiczny	101
VIII. System morfologiczny	117
IX. System syntaktyczny. Relacje syntagmatyczne. Typologia syntaktyczna języków	129
X. System semantyczny. Odwzorowanie rzeczywistości w języku	153
XI. System semantyczny. Relacje paradygmatyczne. Pole semantyczne	175
XII. Znaczenie	189
XIII. Prezentacja znaczenia – definicje	203
XIV. Prezentacja znaczenia – kody semantyczne	212
XV. System semantyczny języków informacyjno-wyszukiwawczych. Znaczenie. Prezentacja znaczenia	248
XVI. Tekst	289
XVII. Statystyczne prawidłowości w tekście	307
XVIII. Procesy informacyjne	335

Contents

From the Author	7
I. Information	9
II. Signs	22
III. Sign systems. Codes and languages	33
IV. Features of sign systems	56
V. Functions of sign systems	65
VI. Metainformation, metatext, hypertext	79
VII. Phonological system	101
VIII. Morphological system	117
IX. Syntactic system. Syntagmatic relations. Syntactic topology of languages ..	129
X. Semantic system. Image of reality in language	153
XI. Semantic system. Paradigmatic relations. Semantic field	175
XII. Meaning	189
XIII. Presentation of meaning – definitions	203
XIV. Presentation of meaning – semantic codes	212
XV. Semantic system of information retrieval languages. Meaning. Presentation of meaning	248
XVI. Text	289
XVII. Statistical regularities in texts	307
XVIII. Information processes	335

Od autorki

Językoznawstwo dla studentów informacji naukowej jest nową wersją wydane-
go przez Wydawnictwa Uniwersytetu Warszawskiego w 1991 roku podręcznika
Zarys językoznawstwa dla studentów bibliotekoznawstwa i informacji naukowej,
wersją o zmienionym układzie, uaktualnionych przykładach, poszerzoną o nowe
zagadnienia.

Te nowe zagadnienia dotyczą przede wszystkim użycia języka w konkretnych
aktach komunikacyjnych (aktach mowy) – bada je i opisuje prężnie rozwijająca
się dziedzina językoznawstwa zwana pragmatyką lingwistyczną. W podręczniku
zostały one omówione w nowym rozdziale *Procesy informacyjne*. W podręczniku
znalazł się jeszcze jeden nowy rozdział *System semantyczny języków informacyj-
no-wyszukiwawczych. Znaczenie. Prezentacja znaczenia*.

Podręcznik pisałam z myślą o studentach bibliotekoznawstwa i informacji na-
ukowej, których przez wiele lat swoimi wykładami starałam się wprowadzić w
zagadnienia językoznawstwa niezbędne w teorii i praktyce języków i systemów
informacyjno-wyszukiwawczych, stąd przykłady i odniesienia do tych właśnie
języków.

Ale podręcznik podaje przede wszystkim wiadomości o języku niezbędne także
studentom wszystkich filologii. Wiem, że z poprzedniego podręcznika korzystali
– mam nadzieję, że i ten będzie im służył jako wprowadzenie do dalszego zglę-
biania problematyki współczesnego językoznawstwa.

Informacja

Spośród dyscyplin naukowych, które rozwinęły się w minionym stuleciu, wiele zwróciło uwagę na to, że otaczający człowieka świat można traktować jako pewien zespół obiektów, między którymi dokonuje się nieustanny obieg informacji.

Znaczenie terminu **informacja** jest w tym wypadku nieco inne niż w języku potocznym. Potocznie przywykło się uważać, że informacja jest zawsze wiadomością o czymś, że informację uzyskać może tylko człowiek, otrzymując ją od innego człowieka bądź poprzez obserwację lub czynność umysłową, że wreszcie przekazać informację może tylko człowiek człowiekowi. W potocznym rozumieniu terminu informacja człowiek zajmuje więc miejsce wyjątkowe, uprzywilejowane, gdy w rzeczywistości cały świat jest systemem przekazów informacji, a człowiek tylko jednym z ogniw tego systemu. Będziemy więc używać terminu **informacja** w znaczeniu dużo szerszym, obejmując nazwą **procesu informacyjnego** „wszelkie uzyskiwanie danych nie tylko przez człowieka, lecz również przez dowolną istotę żywą, organ człowieka lub innej istoty żywej, a także przez maszynę lub jej organ (nie tylko każda wiadomość, lecz również każda decyzja, każdy rozkaz, nakaz, zalecenie, sugestia jest informacją) oraz wszelkie przekazywanie danych między ludźmi, istotami żywymi i maszynami (ewentualnie między organami istot żywych lub maszyn)” (Henryk Greniewski, *Sprawy wszystkie i jeszcze inne*).

W tym obiegu informacji możemy wyróżnić następujące **procesy**:

- nadawanie informacji,
- odbieranie informacji,
- przekazywanie („transportowanie”) informacji,
- przekształcanie (zwane przetwarzaniem lub transformowaniem) informacji,
- utrwalanie informacji,
- przechowywanie informacji.

Aby można było mówić o procesie informacyjnym, w otaczającej nas rzeczywistości musi zaistnieć sytuacja, w której dałoby się wyróżnić następujące elementy:

- źródło informacji (podklasą będzie tu klasa nadawców informacji),
- odbiorca informacji lub adresat informacji,

- informacja (przekaz, komunikat),
- język informacji (kod),
- kanał informacyjny (kanał łączności),
- materialny nośnik informacji (sygnał),
- instrumenty informacyjne.

Przyjrzyjmy się tym elementom i procesom dokładniej.

Źródło informacji (nadawca) i odbiorca informacji to dwa wyróżnione elementy świata pozostające ze sobą w kontakcie informacyjnym. Źródło informacji, a więc element rzeczywistości emitujący sygnał informacyjny, może w procesie informacyjnym pełnić rolę czynną lub bierną. Należałoby więc, aby unikać ewentualnych niejasności, używać terminu **nadawca informacji** tylko w takich sytuacjach, w których elementowi wysyłającemu informację można przypisać wyraźny zamiar przekazania informacji, intencję poinformowania (intencję znakową), w przypadku zaś braku takiej intencji używać terminu **źródło informacji**. I tak na przykład Jan, który mówi nam, że wyjeżdża na wczasy do Zakopanego, jest nadawcą informacji, natomiast słoje drzewa, których liczba i układ „informują nas” o wieku tego drzewa, są źródłem informacji. Możliwa jest również taka sytuacja informacyjna, w której nadawca informacji jest tożsamy z jej odbiorcą, na przykład człowiek, który myśli, jest równocześnie nadawcą i odbiorcą informacji. Człowiek, który coś mówi, zazwyczaj równocześnie słyszy to, co mówi, jest więc równocześnie nadawcą i odbiorcą komunikatu, choć nie jest jego adresatem.

Element rzeczywistości, do którego nadawca kieruje komunikat, to **adresat komunikatu**. Adresat komunikatu może komunikat odebrać i wtedy staje się odbiorcą komunikatu, ale może go nie odebrać, pozostając tym elementem, dla którego nadawca podjął pewne działania zmierzające do takiego ukształtowania elementów łączącego ich środowiska (zwanego kanałem informacyjnym), aby stały się one sygnałem informacyjnym. Odbiorcą takiego sygnału może stać się ktoś przypadkowy, niebędący jego adresatem.

Wszystko to, co nadawca komunikuje odbiorcy w procesie komunikacji, to informacja. Informacja przekazywana jest odbiorcy w postaci **komunikatu** (przekazu informacyjnego). Informacja zawarta w komunikacie dotyczy otaczającego nas świata (lub nas samych), wyróżnionych przez nadawcę lub odbiorcę elementów rzeczywistości i wyróżnionych relacji spośród wszystkich relacji zachodzących między tymi wyróżnionymi elementami (oczywiście mogą to być różne rodzaje rzeczywistości, nie tylko rzeczywistość rozumiana intuicyjnie jako to, co naprawdę jest).

Każdy komunikat ma postać **sygnału** i jest wyrażony w jakimś **kodzie** (języku), którego jednostki (układy tych jednostek) zastępują elementy (zjawiska) komunikowanej rzeczywistości. Kodem takim w szczególnym przypadku może być język etniczny (naturalny). Kod umożliwia więc sformułowanie komunikatu i przesłanie go na odległość. **Komunikat** jest więc tekstem przekazanym w jednym akcie komunikacyjnym. Jeśli jest to tekst języka z semantyką, czyli języka, dla którego wyrażen określono reguły przyporządkowania między tymi

wyrażeniami a elementami rzeczywistości pozajęzykowej, to elementy materialne komunikatu odwzorowują informację o tej rzeczywistości. Komunikatem może być zbiór znaków drogowych na słupie, wypowiedź ustna, artykuł w gazecie, wywieszona flaga, ukłon, palące się światło sygnalizacji ulicznej, odegrany utwór muzyczny (ten ostatni jest oczywiście komunikatem sformułowanym w języku bez semantyki).

Aby zakodować i przesłać informację, konieczne jest wybranie jednego z wielu możliwych kanałów informacyjnych, gdyż kod, którego używamy, zazwyczaj bywa związany z określonym rodzajem sygnału, choć teoretycznie może być realizowany w różnym materiale fizycznym. **Kanał informacyjny** to środowisko materialne łączące nadawcę z odbiorcą, w którym dokonuje się przemieszczenie sygnału informacyjnego. Z kanałem informacyjnym wiąże się ściśle **nośnik informacji**, który jest jej elementem materialnym. Nośnikiem informacji może być na przykład fala akustyczna, radiowa, świetlna. W niektórych wypadkach mamy również do czynienia z nośnikiem nośnika informacji – w wypadku informacji utrwalonej jest nim na przykład taśma magnetyczna, klisza fotograficzna, płyta kompaktowa czy papier.

Ten sam kod może być realizowany przy pomocy różnych kanałów. Takie różne postacie jednego i tego samego kodu, realizowane tylko za pomocą różnych kanałów informacji, nazywamy jego **subkodami**, zaś subkod najczęściej używany nazywamy **subkodem głównym**. Tak więc mamy różne subkody języka naturalnego, realizowane przez różne kanały informacyjne (podaję za Tadeuszem Milewskim, *Językoznawstwo*), na przykład:

- subkod główny wokalno-audycyjny – język mówiony,
- subkod wzrokowy (zazwyczaj – BB) trwały – pismo,
- subkod wzrokowy (zazwyczaj – BB) przemijający – mowa mimiczna głuchoniemych,
- subkod instrumentalno-audycyjny – mowa bębniowa Murzynów,
- subkod dotykowy, tj. posługujący się zmysłem dotyku jako kanałem informacyjnym (a właściwie jako odbiornikiem informacji – BB) – alfabet dla niewidomych Braille'a.

Nadawca posługuje się więc tym lub innym kanałem informacji zależnie od warunków technicznych, w jakich musi funkcjonować, tj. przekazywać informację, wskutek czego w miarę zmiany warunków wciąż powstają nowe subkody.

Spośród wymienionych wyżej elementów sytuacji informacyjnej pozostały do omówienia jedynie **instrumenty** używane w procesach informacyjnych. Zależnie od rodzaju procesu informacyjnego wyróżnić tu można:

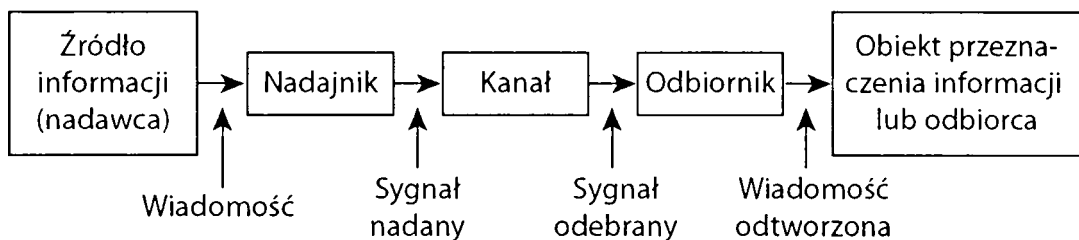
- instrumenty służące do nadawania komunikatu – **nadajniki**, których funkcją jest tworzenie sygnału, np. pióro, maszyna do pisania, antena radiowa, aparat mowny człowieka;
- instrumenty służące do odbierania komunikatu – **odbiorniki** postrzegające sygnał, np. wzrok, słuch, dotyk, węch, smak, będące instrumentami naturalnymi przeznaczonymi do odbierania informacji, oraz np. antena radiowa, będąca instrumentem sztucznym;

- instrumenty służące do przekazywania (transportowania) komunikatu – **przeźniaki**, np. układ nerwowy będący instrumentem naturalnym, przeźniaki sygnału telefonicznego, telegraficznego czy telewizyjnego będące instrumentami sztucznymi (przeźniaki łączą często funkcję odbiornika i nadajnika komunikatu, np. satelita telekomunikacyjny, głośnik);

- instrumenty służące do przechowywania informacji, urządzenia zwane **pamięcią**, albo będące częścią użytkownika komunikatu – zwane **pamięcią wewnętrzną**, np. mózg, pamięć wewnętrzna komputera, albo niebędące częścią użytkownika – zwane **pamięcią zewnętrzną**, którą są, lub stanowią jej główny element, dokumenty, przechowywać można bowiem tylko informację utrwaloną;

- instrumenty służące do transformowania komunikatu. Są to instrumenty różnego rodzaju, a podzielić je można zależnie od tego, czy dokonują przekształcenia tylko sygnału – materialnego nośnika informacji, czy też i sygnału, i samej informacji. Wśród tych pierwszych wyróżnić można instrumenty wzmacniające sygnał, np. okulary, mikroskop, tuba, epidiaskop, lornetka, oraz instrumenty dokonujące zamiany sygnału, transformujące informację odbieraną w jednym kanale i nadającą ją w innym kanale (jest to więc operacja złożona). Instrumentami takimi są np. termometr, barometr, mikrofon, magnetofon, gramofon, radioodbiornik, telewizor, telefon (jak widać z przytoczonych przykładów, taka transformacja towarzyszy często procesowi utrwalańia informacji). Istnieją także urządzenia dokonujące skomplikowanych procesów transformowania komunikatu zmierzających do zmiany jego zawartości informacyjnej, jak na przykład komputer.

Najprostszy schemat przekazywania informacji wygląda więc następująco:



Każdy organizm żywy, a więc także organizm człowieka, wyposażony jest w narządy zmysłów – **receptory**, przekazujące informacje płynące z kontaktu organizmu z otoczeniem, a także informujące organizm o jego własnym stanie wewnętrznym. Liczba tych informacji jest bardzo duża, jak wyliczono bowiem, tylko samo otoczenie organizmu dostarcza mu danych w liczbie rzędu 10^9 bit/sek. Wszystkie receptory organizmu podzielić można na trzy grupy:

- eksteroreceptory – informujące organizm o jego środowisku;
- interoreceptory – kontrolujące przemiany środowiska ustrojowego, a więc dostarczające informacji o stanie wewnętrznym organizmu, takim jak głód, zmęczenie itp.;
- proprioreceptory – kontrolujące postawę organizmu i przekazujące informacje do ośrodkowego układu nerwowego.

Każdy z tych receptorów odbiera informacje wyrażone we właściwym dla niego kodzie. Kody te Henryk Greniewski nazwał **kodami doznaniowymi**. Człowiek komunikuje się ponadto sam ze sobą za pomocą języka, na przykład robiąc notatki lub myśląc.

Należy przede wszystkim wyraźnie zdać sobie sprawę z tego, że język naturalny, stanowiący dla człowieka główne narzędzie przekazywania informacji, jest – w porównaniu z kodem doznaniowym – narzędziem mało sprawnym. Henryk Greniewski tak o tym pisze (*Sprawy wszystkie i jeszcze inne*) – „trudno jest zachwycić się dostatecznie bogactwem kodu doznaniowego; w porównaniu z kodem doznaniowym analogiczne partie najbardziej nawet rozwiniętych języków etnicznych wydają się ubogie. To relatywne ubóstwo języka w stosunku do kodu doznaniowego pociąga poważne skutki w życiu społecznym. Żadne doznanie nie jest bezpośrednio przekazywalne żadnemu innemu człowiekowi, wobec tego po to, aby się ludzie ze sobą porozumiewali (co jest, oczywiście, praktyczną koniecznością), odbywać się muszą nieustannie dwa przekłady:

- a) w głowie nadawcy – przekład z kodu doznaniowego na język oraz
- b) w głowie odbiorcy – przekład z języka na kod doznaniowy.

Warto się nad tym zastanowić, jak dalece niszczący jest przekład z bogatego (czterowymiarowego) kodu doznaniowego na (jednowymiarową – linearną – BB) mowę czy pismo, i ile dowolności zawiera w sobie przekład z (jednowymiarowej) mowy czy (również jednowymiarowego) tekstu pisanego na (znowu czterowymiarowy) kod doznaniowy”. O wszystkich tych kłopotach wiedzą najlepiej ci, którzy zetknęli się z informacją patentową, gdyż opis słowny niektórych przedmiotów i urządzeń omawianych w opisach patentowych mimo wysiłków autora bardzo często nie daje czytelnikowi jasnego obrazu tego, jak ten przedmiot czy to urządzenie ma w praktyce wyglądać.

Dokonując powyższej operacji przekładowej musimy więc dokonać pewnego wartościowania odebranej przez nasze receptory informacji i zdecydować się na przełożenie na język naturalny tylko tej części, która jest dla nas lub dla kogoś, kogo my z kolei informujemy, najbardziej istotna (**relewantna**). Selekcji tej dokonujemy przez odpowiednie „zagęszczanie” informacji (swoiste streszczanie), przeprowadzając redukcję elementów informacyjnych niekiedy tylko do jednego – w danych okolicznościach najważniejszego – elementu, który następnie otrzymuje swój językowy wyraz.

Rola odbiorcy w procesie informacyjnym nie zawsze polega na wykonywaniu takich samych czynności. Ilość i rodzaj pracy, jaką musi on wykonać, aby odebrać informację, jest bardzo różna i w dużym stopniu zależy od rodzaju odbieranej informacji oraz od wielkości szumu istniejącego w kanale informacyjnym (na szum składają się wszelkie sygnały niebędące nośnikami wiadomości użytecznej). O tym, że odbiorca odebrał informację, możemy mówić właściwie dopiero wtedy, gdy ogólna ilość informacji posiadanej przez odbiorcę po odebraniu komunikatu jest większa od ilości informacji, jaką posiadał przed jego otrzymaniem, a więc wtedy, gdy komunikat spowodował przyrost wiadomości i w pewien sposób zmienił odbiorcę pod względem informacyjnym. Nie wystarczy więc, aby

komunikat zawierający informację dotarł do receptorów odbiorcy i został przez niego zarejestrowany, musi być jeszcze zrozumiany, a zatem włączony do systemu poprzednio nagromadzonych informacji. Określenie kodu informacyjnego ze względu na odbiorcę informacji nie jest przypadkowe, on to bowiem jest najważniejszym uczestnikiem procesu informacyjnego i od jego możliwości zależy ilość i jakość odbieranych informacji. Kompetencje nadawcy informacji są w procesie informacyjnym w porównaniu z kompetencjami odbiorcy bardzo niewielkie, każdy organizm bowiem, a więc i człowiek, może przecież odbierać o wiele więcej i wiele bardziej różnorodnych informacji z otaczającego go środowiska niż może nadać, co więcej – każdy organizm zna o wiele więcej kodów odbiorczych niż kodów nadawczych. Przykładem mogą tu być kody doznaniowe człowieka, które są dla niego tylko kodami służącymi do odbierania informacji lub kody wyobrażeń i marzeń sennych, które są wyłącznie kodami wewnętrznymi organizmu, nie mogą więc być użyte do nawiązywania łączności z innymi organizmami. O funkcji odbiorcy w procesie informacyjnym świadczy również to, że warunkiem odebrania w pełni informacji, a więc zarejestrowania i odcodowania jej, jest posiadanie przez odbiorcę odpowiednich do tego kompetencji. Musi on w tym celu dysponować odpowiednimi receptorami umożliwiającymi zarejestrowanie informacji nadanej na tym właśnie kanale, znać kod, w którym informacja została nadana, czyli posiadać odpowiednią kompetencję językową, oraz mieć niezbędną ilość informacji umożliwiającą przyjęcie informacji nowej (oczywiście są to jedynie warunki konieczne, które musi spełniać odbiorca, aby mógł odebrać informację).

Już z powyższego widać, że odbiorca pełni w procesie informacyjnym rolę pierwszoplanową. Po dokładniejszym przyjrzeniu się temu procesowi możemy jednak zauważyć, że praca odbiorcy mająca na celu odebranie informacji zaczyna się właściwie nie w tym momencie, gdy tekst dociera do jednego z jego receptorów, ale znacznie wcześniej – można by właściwie powiedzieć, choć będzie to z pewnością określenie mało precyzyjne, że odbiorca musi się w jakiś sposób nastawić na odbieranie informacji. Musi to być przy tym nastawienie czynne, nie tylko bierne „przechodzenie na odbiór”. Aby sprawę wyjaśnić, przypomnijmy, że potencjalnym źródłem informacji jest dla danego odbiorcy cały postrzegalny dla niego świat, cała otaczająca go rzeczywistość, a także on sam, jego własny organizm. Potencjalna ilość informacji, jaka mogłaby być odebrana przez odbiorcę, jest więc ogromna, natomiast jego możliwości odbiorcze są dużo mniejsze, ograniczone już chociażby budową i sposobem działania receptorów. Odbieranie informacji to poza tym duży wysiłek dla organizmu. Musi się ten wysiłek przecież opłacać, chodzi więc o to, aby był współmierny do osiągniętych z tej pracy korzyści. Aby wysiłek ten był ekonomiczny, organizm musi nastawić się na odbieranie tylko tych informacji, które mają dla niego największą wartość, musi więc je wstępnie wyselekcjonować, skupiając uwagę tylko na tych obiektach, które mogą być źródłem użytecznej dla niego informacji, czyli informacji dla niego relewantnej. Taką selekcję przeprowadza się oczywiście stopniowo, przede wszystkim zaś wybiera się kanał, którym ma płynąć oczekiwana informacja.

Selekcja wszystkich bodźców, które mają potencjalną wartość informacyjną, jest także następstwem pewnej klasyfikacji otaczającej człowieka rzeczywistości, dokonywanej zgodnie z pojęciową siatką języka naturalnego, którym się dana jednostka posługuje. Dzieje się tak dlatego, że przede wszystkim wyodrębnia się z rzeczywistości te obiekty i cechy, którym przyporządkowane są wykładniki językowe, reszta bardzo często umyka uwagi obserwatora, a bywa i tak, że dostrzeżenie ich wymaga przewyciężenia utwalonych przez posługiwanie się językiem nawyków. Prowadzi to do niemal natychmiastowego przekładu odebranych w kodzie doznaniowym informacji na kod – mowę codzienną. Zwrócił na to uwagę Henryk Greniewski (*Cybernetyka niematematyczna*), pisząc: „W obrębie społeczeństwa nieustannie i masowo odbywa się przekład informacji – doznań na kod – mowę codzienną (czy kod – pismo) i przekład odwrotny (...). Nie należy zapominać, że przekład z kodu doznaniowego na kod – mowę czy kod – pismo odbywa się niekiedy nie dla celów łączności, lecz dla sformułowania przesłanek rozumowania (zwykle indukcyjnego). Nie należy też zapominać, że tak jesteśmy wprawieni w nieustanny przekład z kodu doznaniowego na kod – mowę, iż odbierając informację o środowisku czy własnym wnętrzu każdy z nas jednocześnie przekłada ją niezwykle szybko, lecz nieściśle na kod – mowę: doznawaniu towarzyszą niemal równocześnie informacje o wyobrażeniach dźwiękowych mowy”.

Poszukiwanie wciąż nowej informacji jest realizacją pasji odkrywczej, jaka cechuje każdego rozumnego człowieka. Bardzo często jednak nową informację człowiek otrzymuje nie bezpośrednio z otaczającego go świata za pomocą narządów zmysłów, ale dochodzi do niej na drodze bardzo skomplikowanych operacji myślowych przeprowadzanych na informacjach, które uprzednio uzyskał. W tym wypadku źródłem nowej informacji są inne informacje, które ulegają odpowiedniemu przetworzeniu. Te skomplikowane procesy myślowe, prowadzące do uzyskania nowych informacji, to nic innego jak po prostu odpowiednie **transformowanie** informacji wejściowych. Umysł ludzki jest więc systemem, który z pewnych informacji otrzymywanych na wejściu systemu daje na wyjściu nową informację.

Transformowanie informacji jest zazwyczaj rozumiane dość szeroko. Henryk Greniewski tak je określa w swej *Cybernetyce niematematycznej*: „Transformowanie informacji jest to operacja wykonywana na informacji (albo na informacjach) a dająca w wyniku informację. I tu można przedstawić liczne przykłady. Transformowanie informacji stanowi część składową każdego odruchu (obojętne, czy to odruch bezwarunkowy, czy warunkowy); wszelkie rozumowanie i wszelki rachunek jest transformowaniem informacji; wszelki przekład i wszelkie streszczenie jest transformowaniem informacji; wszelkie marzenie czy bieg wyobrażeń, obojętne, czy w stanie czuwania, czy we śnie – wszystko to jest transformowaniem informacji”.

W każdym językowym procesie informacyjnym funkcję informującą (z punktu widzenia odbiorcy) pełnią:

- Informacja świadomie (celowo) nadana przez nadawcę z intencją zmienienia pod względem informacyjnym odbiorcy.

- Sposób ukształtowania tej części materiału fizycznego będącego nośnikiem informacji, która nie jest bezpośrednio wykorzystywana do wyrażania językowej informacji semantycznej. Może to być zarówno ta część, która stanowi tworzywo językowe informacji, a więc na przykład wybór – świadomy lub nie – jednego z wielu możliwych równoznacznych sposobów skonstruowania przekazu informacyjnego, jak i część nienależąca do języka, na przykład barwa i modulacja głosu, tempo i natężenie mowy (również kształtowane świadomie lub nie). W przypadku tekstu pisanego taką rolę może pełnić charakter pisma, rodzaj użytej czcionki oraz rozmieszczenie tekstu.

- Stan fizyczny nadawcy w momencie przekazywania informacji, każdy człowiek bowiem w momencie, w którym coś mówi lub czegoś słucha, a więc w momencie, w którym czynnie uczestniczy w jakimś procesie informacyjnym, jakoś wygląda i jakoś się zachowuje. Wygląd i sposób zachowania nie są jednak przypadkowe, są wyraźnie skorelowane z sytuacją informacyjną, pełniąc w niej bardzo ważną funkcję informacyjną i przekazując wcale niemałą część całej informacji krążącej w danym układzie informacyjnym.

Po tych wszystkich rozważaniach na temat procesów informacyjnych możemy się pokusić o sformułowanie nieco dokładniejszych ich określeń.

Nadawanie komunikatu to proces informacyjny polegający na posłużeniu się przez nadawcę znakiem (zgodnie z jego funkcją znakową) lub utworzeniu sygnału (zgodnie z regułami danego języka). W wypadku języka z semantyką, a więc języka, którego wyrażenia zastępują elementy odwzorowywanej przez ten język rzeczywistości, nadawanie komunikatu połączone jest z kodowaniem informacji. Nadawaniem komunikatu jest na przykład wywieszenie białej flagi na znak poddania się, kiwanie głową na znak potwierdzenia, machanie ręką na pożegnanie, granie melodii, wypowiedzanie tekstu w języku naturalnym, tworzenie charakterystyki wyszukiwawczej dokumentu.

Czasami nadawanie informacji (komunikatu) rozumiane jest wężiej i ograniczone tylko do nadawania sygnału. Wprowadza się wtedy opozycję tworzenia komunikatu – równoznacznego z wyżej zdefiniowanym nadawaniem komunikatu, i nadawania komunikatu – rozumianego tylko jako emitowanie sygnału. W myśl tych rozróżnień autor tekstu wygłaszanego przez radio stworzył komunikat, a spiker tylko go nadał, kompozytor stworzył komunikat (utwór muzyczny), a wykonujący go muzyk tylko go nadał. Rozróżnienie takie pociąga za sobą modyfikację definicji nadawcy komunikatu i wprowadzenie opozycji **twórca komunikatu** (autor) i **nadawca komunikatu** (w węższym sensie).

Kodowanie informacji jest to proces informacyjny polegający na odwzorowywaniu przez nadawcę (na podstawie posiadanej kompetencji językowej) w komunikacji informacji zgodnie z regułami semantyki danego języka poprzez przyporządkowanie informacji odpowiedniego sygnału. Jest to więc kolejno odwzorowywanie informacji w planie treści języka, a następnie odwzorowywanie planu treści w planie wyrażania i realizacji go w postaci sygnału. Kodowanie informacji połączone jest więc z nadawaniem komunikatu. Kodowaniem informacji jest na przykład oznaczanie na planie znajdujących się na danym terenie obiektów

za pomocą specjalnych znaków, tworzenie tekstu w języku naturalnym, tworzenie charakterystyki lub instrukcji wyszukiwawczej dokumentu.

Przekazywanie komunikatu jest to proces informacyjny polegający na przemieszczeniu sygnału w wyróżnionej przestrzeni (kanale informacyjnym), na przykład w obrębie układu nerwowego, przez linię telefoniczną lub telegraficzną, drogą radiową. Procesowi temu towarzyszy często transformowanie sygnału – na przykład sygnał akustyczny przekształcany jest na sygnał radiowy. Przekazywanie informacji bywa również rozumiane szerzej, jako proces komunikacji. Potocznie przekazywaniem komunikatu lub informacji nazywa się też nadawanie informacji uzyskanej z innego komunikatu. Przekazywaniem informacji nazywa się niekiedy także proces określony wyżej jako nadawanie komunikatu w sensie węższym.

Komunikacja jest to proces informacyjny mający na celu przekazanie informacji. Proces komunikacji zachodzi między nadawcą a odbiorcą, któremu nadawca przekazuje informację w postaci komunikatu (informacji nie można przekazać inaczej niż pod postacią komunikatu, przekazać można bowiem tylko rzecz materialną, a więc sygnał, dlatego też używamy terminów nadawanie komunikatu i nadawca komunikatu, a nie nadawanie informacji i nadawca informacji). Na proces ten składają się nadawanie komunikatu, przekazywanie komunikatu oraz odbieranie komunikatu, połączone najczęściej z odkodowywaniem informacji.

Odbieranie komunikatu jest to proces informacyjny polegający na postrzeganiu przez odbiorcę sygnału i ewentualnie odkodowywaniu informacji w wypadku, gdy komunikat jest sygnałem semantycznym lub tekstem języka z semantyką.

Odkodowywanie (dekodowanie) informacji jest to proces informacyjny polegający na przyporządkowywaniu przez odbiorcę komunikatu odebranemu sygnałowi odwzorowanej w nim informacji (na podstawie znajomości reguł semantyki danego języka lub reguł posługiwania się danym znakiem). Proces ten musi być poprzedzony rozpoznaniem języka, w którym zakodowana jest informacja i odbywa się na podstawie kompetencji językowej odbiorcy. Odkodowywanie informacji bywa też nazywane odkodowywaniem komunikatu.

Transformowanie (transformacja) komunikatu. Nim określimy, na czym ten proces polega, musimy najpierw wyjaśnić, jak rozumiany jest termin transformacja. Najczęściej **transformacją** nazywa się przekształcenie jednej struktury w drugą, zazwyczaj ujęte w postaci zbioru operacji – reguł transformacyjnych. Transformowanie komunikatu jest więc przekształceniem struktury komunikatu albo jego sygnału, na przykład transformacja tekstu pisanego na mówiony, transkrypcja, transliteracja, tłumaczenie na inny język (przy uproszczeniu, że możliwe jest w procesie tłumaczenia odwzorowanie w obu tekstach, należących do dwóch różnych języków, dokładnie tej samej informacji przez proste przyporządkowanie sygnałów obu języków, a więc elementów ich planu wyrażania), albo zawartej w komunikacie informacji, albo sygnału i informacji równocześnie.

Utrwalanie komunikatu jest to proces informacyjny polegający na utrwaleniu sygnału, czyli **tworzeniu dokumentu**. Utrwalanie komunikatu może być połączone z jego jednoczesną transformacją materialną – transformacją sygnału, na przykład sygnału akustycznego na magnetyczny (utrwalanie na taśmie magnetycznej),

sygnału pneumatycznego na optyczny (np. w wypadku barografu) czy sygnału elektrycznego na optyczny (np. w wypadku elektroencefalografu), może to być też transformacja polegająca na zmianie liczby wymiarów sygnału (np. w wypadku fotografii).

Utrwalanie komunikatu jest więc procesem sporządzania **dokumentu**, gdyż dokument jest to utrwalona informacja wraz z materiałem, w którym została utrwalona. Dokumenty dzieli się zależnie od rodzaju utrwalonego sygnału, najczęściej wyróżnia się dokumenty: audialne (z utrwalonym sygnałem dźwiękowym, np. taśma magnetofonowa, płyta gramofonowa), wizualne (z utrwalonym sygnałem optycznym, np. fotografia, film, przezrocze), wśród których wyróżnia się tzw. dokumenty graficzne, gdzie informacja została utrwalona za pomocą znaków graficznych (np. książka, list), dokumenty dotykowe (np. książki drukowane pismem Braille'a). Oczywiście możliwe są typy mieszane, na przykład dokumenty audio-wizualne zawierające utrwalony dźwięk i obraz (np. film dźwiękowy). Dokumenty dzieli się również według cech materiału, w którym nośnik został utrwalony, czyli według cech nośnika informacji. Wyróżnić tu można np. tabliczki gliniane, papiirusy, książki, taśmy filmowe, taśmy magnetyczne, płyty gramofonowe, dyski magnetyczne, płyty kompaktowe itd. Inne typologie dokumentów dotyczą cech utrwalonej informacji.

Przechowywanie informacji – procesu tego nie będziemy definiować, ponieważ nie jest to proces swoisty dla informacji. Informację można przechowywać w specjalnych urządzeniach (instrumentach), tak jak można przechowywać inne przedmioty. Zwrócimy jedynie uwagę na to, że przechowywać można tylko informację utrwaloną. (Więcej o procesach informacyjnych powiemy w ostatnim rozdziale podręcznika, poświęconym użyciu języka naturalnego.)

Wszystkie poprzednie rozważania dotyczyły różnych rodzajów informacji, nie powiedzieliśmy jednak, co to właściwie jest **informacja**, nie podaliśmy definicji podstawowego dla nas terminu, zakładając niejako, że jest on intuicyjnie dla wszystkich zrozumiały. Zdefiniowanie tego terminu sprawia ogromne kłopoty wszystkim zajmującym się z jakichś względów informacją.

Jako pierwszy, w sposób naukowy, zaczął zajmować się informacją amerykański matematyk Claude Elwood Shannon, twórca matematycznej teorii informacji. Shannona interesowały jednak głównie zagadnienia optymalnych ekonomicznie warunków przenoszenia informacji, a więc raczej przekazywania przenoszących informację sygnałów, nie zaś zagadnienia samej informacji. Określił on, opierając się na prawdopodobieństwie zdarzenia, co to jest ilość informacji, samej informacji jednak nie zdefiniował, nie było to bowiem dla niego konieczne. Inni autorzy piszący o informacji albo posługują się tym terminem bez uprzedniego definiowania, albo wyjaśniają go za pomocą wyrażen synonimicznych, takich jak „treść”, „wiadomość”, co nie jest przecież definiowaniem. I tak na przykład Norbert Wiener pisze: „Informacją nazywamy treść zaczerpniętą ze świata zewnętrznego w procesie naszego dostosowywania się do niego i przystosowywania do niego naszych zmysłów” (N. Wiener, *Cybernetyka a społeczeństwo*).

W *Wielkiej encyklopedii powszechnej* informacja jest zdefiniowana w sposób następujący: „Informacja (łac. *informatio* ‘wyobrażenie, wyjaśnienie, zawiadomienie’), w cybernetyce i teorii informacji każdy czynnik, który ludzie, organizmy żywe lub urządzenia automatyczne mogą wykorzystać do bardziej sprawnego, celowego działania (w szczególności sterowania); zob. też «wiadomość»”, zaś pod hasłem „wiadomość” czytamy: „w cybernetyce czynnik, na podstawie którego ludzie, organizmy żywe lub urządzenia automatyczne mogą w sposób sprawny przeprowadzać celowe działania; wiadomością mogą być np.: obraz, dźwięk, zapach, wykres, wskazanie przyrządu pomiarowego, słowo, liczba (tu niekonsekwencja, bo przecież wykres i wskazanie przyrządu pomiarowego jest obrazem, słowo zaś dźwiękiem lub obrazem – BB); źródłem wiadomości jest podający je obiekt, np. mówiący człowiek, miernik dołączony do układu pomiarowego; przekazywaniem wiadomości zajmuje się teoria informacji; wiadomość traktuje się w niej jako zmienną przypadkową; własności statystyczne wiadomości zależą od własności podającego je źródła, a więc mogą stanowić podstawę klasyfikacji źródeł na binarne, ziarniste i ciągłe; wiadomość przed przekazaniem podlega często przekształceniom postaci w celu nadania jej dogodnej formy; na ogół w celu przekazania wiadomości przyporządkowuje się jej pewien nośnik zwany sygnałem; operację tę wykonuje się w nadajniku”.

Niewiele więcej możemy się dowiedzieć ze *Słownika terminologicznego informacji naukowej* (red. Maria Dembowska), gdzie wprowadzono aż cztery hasła: „informacja – pot. powiadomienie o czym, zakomunikowanie czego, wiadomość”, „informacja – cybern. (1) każdy czynnik zmniejszający stopień niewiedzy (nieokreśloność) o badanym zjawisku, umożliwiający człowiekowi i innemu organizmowi żywemu lub urządzeniu polepszenie znajomości otoczenia i w sprawniejszy sposób przeprowadzenie celowego działania; (2) miara zorganizowania układu”; „informacja – informat. znaczenie przypisywane danym, z uwzględnieniem konwencji stosowanych do ich wyrażania”; informacja – inte (1) każda wiadomość będąca odbiciem rzeczywistości lub ją zastępująca”.

Więcej wyjaśniać zdają się definicje informacji Tadeusza Wójcika i Henryka Greniewskiego – przytoczymy je poniżej.

Wójcik w *Prakseosemiotyce* pisze: „Rzecz poznawana P jest informatorem zmysłowym (informatorem bezpośrednim). Jego powierzchnia wysyłająca fale optyczne lub akustyczne (albo inne) – to emitor informujący, emitor informatora zmysłowego. Owa fala wysyłana przez emitor – to emisja informacyjna. Poznawanie zmysłowe (lub informowanie zmysłowe) można zdefiniować jako iloczyn względny obu stosunków: emitowania i postrzegania. Emitor informacyjny lub emisję informacyjną nazywam informacją”. I dalej: „Rzecz Z jest informatorem komunikatowym (informacją komunikatową) rzeczy P dla osoby U zawsze i tylko, gdy: dla każdego X – jeżeli X jest Z, to X jest rzeczą przyporządkowaną konwencjonalnie rzeczy P (lub rzeczą częściowo podobną do rzeczy P) i nadaje się do wyrażania myśli osoby U o rzeczy P”. Ta druga definicja jest definicją informacji mającej nadawcę.

Henryk Greniewski określa informację jako odbicie zasilenia (przy czym przez zasilenie rozumie dawkę energii lub masy), posługując się porównaniem: świeca (zasilenie) – odbicie świecy w zwierciadle (informacja o zasileniu, czyli o świecy), (H. Greniewski, *Cybernetyka niematematyczna*).

Z tych ostatnich rozważań wyciągnąć możemy dwa ważne wnioski:

- informacją może być tylko taka rzecz (układ rzeczy), która jest przez kogoś postrzegana – ma swego odbiorcę (lub jest skonstruowana, przeznaczona dla jakiegoś odbiorcy przez jakiegoś nadawcę);
- informacją jest rzecz (układ rzeczy), będąca odbiciem (modelem) rzeczywistości, tworzona albo przez nadawcę informacji z przeznaczeniem dla odbiorcy – a wtedy jej funkcją jest zastępowanie na zasadzie czysto konwencjonalnej (na zasadzie umowy, przy czym stronami tej umowy są nadawca i odbiorca) tej rzeczywistości, która przecież nie może być przesyłana, albo tworzona przez odbiorcę w miarę postrzegania poszczególnych elementów rzeczywistości.

Na zakończenie naszych rozważań o informacji należałoby jeszcze zwrócić uwagę na to, że termin informacja używany bywa w dwóch znaczeniach, co innego oznacza on bowiem w zdaniach:

- (1) *Na ekranie pojawiła się informacja o jutrzejszym programie telewizyjnym. Na Radzie Wydziału podano informacje o nowych zasadach rekrutacji. Wszystkie agencje przekazały dziś informację o straszliwym trzęsieniu ziemi w jugosłowiańskiej Czarnogórze.*

co innego zaś na przykład w zdaniach:

- (2) *Tekst ten zawiera wiele interesujących informacji. Zdania w stronie czynnej oraz zdania w stronie biernej zawierają taką samą informację. Ilość informacji zawarta w danym komunikacie jest odwrotnie proporcjonalna do prawdopodobieństwa wystąpienia tego komunikatu.*

Wszystkie podane wyżej przykłady to wypowiedzi o informacji rozumianej w sposób dwojaki: bądź jako rzecz lub układ rzeczy będące odbiciem (modelem) jakiejś rzeczywistości, tworzone przez nadawcę z przeznaczeniem dla odbiorcy (a wtedy ich funkcją jest zastępowanie w sposób konwencjonalny tej rzeczywistości), bądź też jako abstrakcyjny system opozycji odpowiadający takiemu materialnemu układowi. W tym pierwszym znaczeniu informacja traktowana jest jako połączenie pewnej abstrakcyjnej struktury (informacji w znaczeniu drugim) z jej materialnym nośnikiem (sygnałem), jest więc synonimem komunikatu, w tym drugim jedynie jako sama taka abstrakcja. Słowa informacja potocznie używa się właśnie w tych dwóch znaczeniach: w pierwszym, gdy mówimy, że ktoś nadał jakąś informację (nadać można tylko sygnał), w drugim, gdy mówimy, że sygnał jest nośnikiem informacji lub gdy stwierdzamy, że dwa teksty zawierają tę samą informację.

Nie zamierzamy tu oczywiście definiować terminu informacja – nikomu jeszcze zadowolającej definicji tego terminu stworzyć się nie udało – chcemy jedynie

zdać sprawę z różnego rozumienia tego podstawowego terminu. Na koniec należałoby odnotować, że termin informacja przyjmowany bywa czasem jako termin pierwotny (obok terminów materia i energia), a więc niewymagający definicji.

Literatura

Eco U.: *Pejzaż semiotyczny*. Warszawa 1972.

Greniewski H.: *Cybernetyka niematematyczna*. Warszawa 1969.

Greniewski H.: *Sprawy wszystkie i jeszcze inne*. Warszawa 1970.

Milewski T.: *Językoznawstwo*. Warszawa 1976.

Prieto L.J.: *Przekazy i sygnały*. Warszawa 1970.

Wójcik T.: *Prakseosemiotyka. Zarys teorii optymalnego znaku*. Warszawa 1969.

II

Znaki

Potencjalnym źródłem informacji dla odbiorcy jest cały otaczający go i postrzegalny dla niego świat. Każdy element rzeczywistości (przedmiot lub sytuacja) jest więc potencjalnym **źródłem informacji** o samym sobie dla jakiegoś odbiorcy. Na przykład, stół i krzesło stojące w pomieszczeniu nie są jeszcze źródłem informacji, staną się nim jednak w chwili, gdy ktoś je spostrzeże. Są one źródłem informacji o samych sobie. Patrząc na nie (nastawiam swoje receptory – oczy na odbiór informacji płynącej kanałem wizualnym) i widzę (odbieram tę informację), jaki mają kształt, wielkość, kolor, z jakiego materiału są wykonane, w jaki sposób są umiejscowione w przestrzeni względem innych przedmiotów, dotykam (nastawiam na odbiór informacji swój zmysł dotyku) i czuję (odbieram informację), że powierzchnia ich jest gładka, chłodna etc. I tak, w miarę napływu informacji, tworzę sobie wyobrażenie o jakiejś rzeczywistości (model). Model ten jest jednak modelem tego i tylko tego wycinka rzeczywistości składającego się ze stołu i krzesła.

Czasami jednak bywa inaczej. Gdy widzę, na przykład, że w jakiejś sali ustawione są na stołach krzesła do góry nogami, to wtedy wiem, że w sali tej na pewno będą sprzątać lub właśnie sprząтали. Gdy wchodzę do restauracji i widzę, że krzesło jest oparte o stolik tak, że oparcie jego dotyka blatu stolika, a tylne nogi nie dotykają podłogi, to wiem, że ktoś już ten stolik zajął, choć tego kogoś przy stoliku nie ma. W tych dwóch ostatnich przypadkach stół i krzesło są źródłem informacji nie tylko o samych sobie, ale również o jakiejś innej sytuacji (tu wcześniejszej lub planowanej). Są więc one jej modelem, komunikują ją, choć za każdym razem w nieco inny sposób. Sprzątania domyślimy się, gdy zobaczymy ustawione do góry nogami na stołach krzesła, ponieważ wiemy, że najczęściej tak bywa, że gdy się sprząta podłogę, to krzesła, żeby nie przeszkadzały, ustawia się na stołach. Informacja o sprzątaniu jest więc wnioskiem, który wyciągamy z innej informacji – przesłanki (jest nią informacja o wzajemnym położeniu krzesel i stołów), przy czym było to wnioskowanie uprawdopodobniające, nie zaś logiczne, gdyż wniosek ten z przesłanki nie wynikał w sposób konieczny (nie można bowiem wykluczyć innej przyczyny takiego rozmieszczenia krzesel i stołów – mógł się po prostu ktoś w ten sposób bawić, tworząc z nich budowle, choć taki wniosek

jest, w oparciu o nasze doświadczenie, mało prawdopodobny). W drugim wypadku, żeby wiedzieć, że stolik jest zajęty, nie musieliśmy przeprowadzać żadnego rozumowania, informację tę uzyskaliśmy, ponieważ wiemy, że taki jest zwyczaj zajmowania miejsc w restauracjach, taka jest po prostu konwencja. Oczywiście dla kogoś, kto by tej konwencji nie znał, stolik i krzesło komunikowałyby tylko same siebie i żadne wnioskowanie nic by tu nie pomogło. W pierwszym wypadku nikt nie zadawał sobie trudu ustawienia krzesła na stolikach, żeby nas poinformować o sprzątnięciu. Rzeczywistość nie była więc przez nikogo kształtowana specjalnie dla celów informacyjnych, nie było więc nadawcy informacji. W drugim wypadku ktoś specjalnie tak ustawił krzesło (uksztaltował rzeczywistość, a więc posłużył się znakiem), żeby ktoś inny wiedział, że miejsce jest zajęte – był więc nadawca informacji (komunikatu), informacja (komunikat) i potencjalny odbiorca (adresat) – był nim każdy, kto chciałby zająć to miejsce. Ustawienie krzesła i stołu było więc tylko o tyle ważne, o ile coś komuś komunikowało. W pierwszym wypadku wzajemne położenie stołów i krzesła było symptomem, oznaką tego, że się sprząta, w drugim było znakiem, sygnałem, że miejsce jest zajęte. Z takimi sytuacjami komunikacyjnymi, sytuacjami, w których ktoś posługuje się znakiem, spotykamy się bardzo często.

Znakami i systemami znakowymi zajmuje się dziedzina wiedzy nazywana **semiotyką**. Stworzyła ona ogólną teorię znaków, zajmującą się istotą znaków, ich funkcjonowaniem oraz ich użytkownikami. W zależności od aspektu, pod jakim rozpatruje się znaki i systemy znakowe, wyróżnia się trzy podstawowe działy semiotyki:

- **semantykę**, badającą relacje między znakami a odwzorowywaną przez nie rzeczywistością, a więc naukę o znaczeniu,
- **syntaktykę**, zajmującą się relacjami między elementami tworzącymi system znakowy,
- **pragmatykę**, która zajmuje się funkcjonowaniem systemów znakowych, relacjami między znakiem a jego użytkownikiem, opisuje więc użycie znaków w sytuacjach komunikacyjnych.

Po omówionych wyżej przykładach sytuacji znakowych należy się czytelnikowi kilka słów objaśniających użyte terminy, które w literaturze filozoficznej, semiotycznej, logicznej, językoznawczej i teorii informacyjnej często występują w różnych znaczeniach.

Najwięcej nieporozumień może powodować sam termin **znak**. Poniższe rozważania nie pretendują do miana rozważań rozstrzygających dotychczasowe spory, mają jedynie zdać sprawę z różnic terminologicznych i umożliwić czytelnikowi szersze spojrzenie na interesującą nas tu problematykę. W *Malej encyklopedii logiki* znak został zdefiniowany jak następuje: „*x* jest znakiem dla członków grupy *G*, gdy *x* jest przedmiotem spostrzegalnym zmysłowo, przy czym między członkami grupy *G* istnieje umowa, wyraźna lub domyślna, ustanawiająca między *x* i jakimś *y* stosunek szeroko pojętego oznaczania. (...) Np. ruch milicjanta oznacza zakaz przejazdu, napis sklep oznacza sklep, itp. Warunek przyporządkowania *x* do *y* w drodze umowy jest warunkiem koniecznym bycia znakiem w omawianym

sensie”. Definicja ta odpowiada więc tylko drugiej z omawianych wyżej sytuacji. Często jednak, co również uwzględniono w *Małej encyklopedii logiki*, termin znak rozumiany jest szerzej, a wówczas za znak moglibyśmy uznać każdy przedmiot postrzegalny zmysłowo, który komunikuje nie tylko sam siebie (a właściwie nie przede wszystkim sam siebie), ale zastępuje jakiś inny element rzeczywistości (często w tym momencie niepostrzegalny zmysłowo), przy czym zastępowanie to nie musi wynikać tylko z umowy, lecz może się również opierać na relacji naturalnej, relacji współwystępowania (lub następstwa) w czasie lub w przestrzeni. Taka definicja obejmowałaby również pierwszą z omawianych wyżej sytuacji. Czasami jednak termin znak bywa rozumiany jeszcze szerzej, wtedy, gdy mówi się, że każdy przedmiot, gdy jest postrzegany, jest swoim własnym znakiem (źródłem informacji o samym sobie), i takie rozumienie, bardzo rzadkie, wiąże się z najogólniejszym rozumieniem terminu informacja.

Niżej przedstawiamy najczęściej stosowaną **klasyfikację znaków**, starając się wszędzie tam, gdzie mogą się nasuwać jakieś wątpliwości, wyjaśniać używaną w literaturze (często niekonsekwentnie) terminologię.

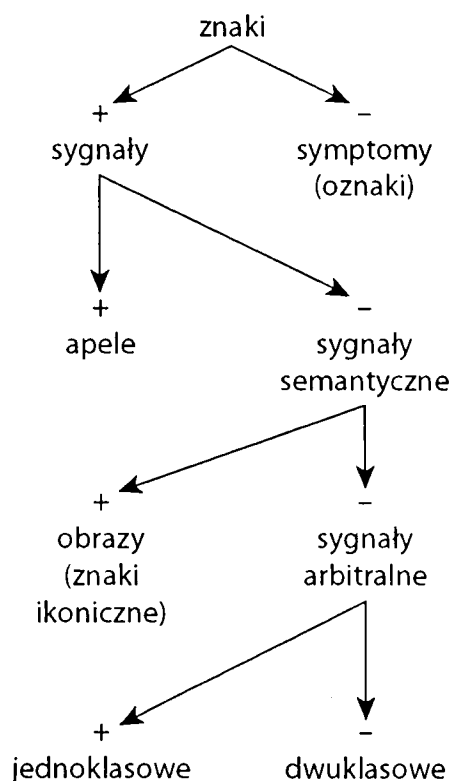
Kryterium podziału

1. Czy mają nadawcę?

2. Czy komunikują tylko same siebie?

3. Czy są podobne do komunikowanej rzeczywistości?

4. Czy funkcję semantyczną pełnią jednostki elementarne?



Znaki w znaczeniu szerszym dzieli się najczęściej w zależności od tego, czy zostały przez kogoś nadane, czy też nie, na **symptomy**, zwane też często **oznakami** (w znaczeniu węższym w ogóle nie należą one do znaków) i **sygnały** (termin sygnał jest tu użyty w innym znaczeniu niż w teorii informacji, por. rozdz. I).

Różnice między tymi dwoma typami podaję za Tadeuszem Milewskim (*Językoznawstwo*):

„1. Symptomy są nie wyodrębnioną częścią kompleksu zjawisk, którego zasadnicza funkcja nie jest znakowa, lecz najczęściej biologiczna, funkcja znakowa zaś występuje tylko ubocznie, natomiast sygnały nie spełniają żadnej innej funkcji prócz znakowej, są one w tej funkcji wyspecjalizowane.

2. Symptomy przez nikogo nie zostały nadane. Są to po prostu zjawiska wywołane przez pewne przyczyny, które znakami są jedynie dla odbiorcy wiążącego je z pewną treścią. Symptomy to znaki jednostronne. Natomiast sygnały zostały celowo nadane przez nadawcę dla oddziaływania na odbiorcę. Są one znakami dla obu, którzy je wiążą z tą samą treścią, a zatem są dwustronne.

3. Symptomy nie są celowo nadawane przez nadawcę, którego w normalnym znaczeniu nie posiadają, wobec czego nie mogą być przez niego przyjęte jako znaki, są one zatem niezwrotne. Człowiek może częściowo rozpoznać symptomy, których jest podłożem, ale będzie to zawsze tylko częśćka tego, co spostrzegają inni. Natomiast sygnały są zwrotne. Nadawca, wysyłając je celowo, sam musi je odbierać, by je skontrolować. Sygnałów, których sami nie możemy odbierać, nie możemy również swobodnie nadawać. Symptomy mogą tworzyć rozległe kody, ale także tylko w jego (odbiorcy – BB) psychice, np. system objawów chorobowych w pamięci lekarza lub system śladów nóg różnych zwierząt w umyśle myśliwego”.

Przykładami **symptomów** mogą być: huk w górach będący oznaką lawiny, dym w lesie wskazujący miejsce, w którym pali się ogień, rozpalona buzia dziecka będąca oznaką wysokiej gorączki, która z kolei jest oznaką jakiejś choroby.

Sygnały, a więc znaki mające swego nadawcę, bywają z kolei dzielone w zależności od tego, czy komunikują tylko same siebie, czy też są nośnikami informacji o jakimś innym wycinku rzeczywistości (stanowią jego odwzorowanie), na **apele** i **sygnały semantyczne**. Termin apel jest tu użyty w węższym rozumieniu, często bywa rozumiany szerzej, a wtedy jest równoznaczny ze zdefiniowanym wyżej terminem sygnał. Przykładem apeli, a więc znaków mających wprawdzie nadawcę, ale komunikujących jedynie same siebie, jest utwór muzyczny, który przecież nie nie znaczy (utwór muzyczny jest tekstem języka bez semantyki). W związku z powyższym należałoby jeszcze zmodyfikować definicję terminu znak, która apeli w żadnym z podanych wyżej znaczeń nie obejmuje, tu bowiem znakiem nazywamy każdy przedmiot (w najszerszym rozumieniu tego wyrażenia), który został przez kogoś wykonany (ma swego nadawcę) tylko po to, żeby komunikował albo sam siebie, albo jakiś inny element rzeczywistości. Tym co przesądza o zaliczeniu danego elementu do klasy znaków, jest tu więc tzw. **intencja komunikacyjna** jego twórcy. Definicja taka nie objęłaby oczywiście symptomów, a stworzenie definicji obejmującej także tę klasę nie wydaje się w ogóle możliwe, co przemawiałoby za wyłączeniem symptomów z klasy znaków.

Sygnały semantyczne komunikujące jakiś wycinek rzeczywistości mogą być w jakiś sposób do niej podobne i wtedy nazywane bywają **obrazami** albo **znakami ikonicznymi (ikonami)**. Takim znakiem ikonicznym jest na przykład rzeźba,

Google PR3F3R3N3Z

534R(h H4K5 | 411 450u7 Google

SAV3 j00r PR3F3R3N3Z WfjBN pH1N15H9D 4ND r3TURN 2 534RCH

54V3 PR3F3R3N3Z

G10b41 Pr3f3r3n3z ((-141N93z 4pp3y3 2 4LL g00g13 53M(Gz)

1N73RF4C3 L4NGU4G3 D15PL4Y Google 71Pz N' M3554G3z 1N H4x3r

534RCH L4NGU4G3

534R(h W3B P4G3Z VWR1773N 1N 4NY L4NGU4G3 (R3CUMM3ND3R3D)

534RCH DN1Y 7H3Z3 53L3C73D L4NGU4G3(Z)

- (4744N
- (04714n
- (z30k
- 1(3141N41(
- 3570N14N
- 3Ng15h (English)
- 4r4b1(
- 53rB14n
- 5V3gd15h
- 5I0V3N14N
- 5I0V4K
- 5p4N15h
- 7ur1<15K
- >>>p0L1sH<<<
- bUjG4r14N
- CH1N953 (51MPL1PH13D)
- 4r4b1(
- CH1N953 (7R4D1710N4L)
- d4N15h
- Du7(h
- f1N1N15h
- fz3Nch
- g3r1V14N
- Gz3K
- 174114N
- Ind0n35i4n
- CH1N953 (7R4D1710N4L)
- K0R34N
- L17j04N14N
- L47Y14N
- P0r7u6353
- RO1V4N14N
- Ru5514N
- N0VV3614N
- j3br3w
- j01Ng4r14N

NumB3R 0F R3ZuL7Z

Google's D3F4UL7 (10 R3ZUL7Z) PR0V1D3Z 04 F4573S7 R3ZuL7Z, j0. D1ZPL4Y | 10 R35uL7z p3R P4G3

R35UL7Z W1ND0W

0P3N 534RCH R35UL7Z 1N 4 N3W BR0W53R W1ND0W

SAV3 j00r PR3F3R3N3Z WfjBN pH1N15H9D 4ND r3TURN 2 534RCH

54V3 PR3F3R3N3Z

(NOTS 53TTING PR3F3R3N3Z VV1LL ND7 W0RK 1F j00 H4V3 d154E134 C00K13z 1N j00r b0jNz3r)

fotografia, mapa jakiegoś terenu, sylwetki pojazdów i ludzi na znakach drogowych, niektóre oznaczenia na mapach topograficznych. Znakami ikonicznymi są też używane coraz częściej w Internecie i wiadomościach SMS, tzw. mordki (emotikony), np.:

*</:-)	św. Mikołaj	>:)	diabełek	0:-)	anioł
:3-<	pies	8)	żaba	====:)	wąż
=:0]	Clinton	B)	Batman	+0<:-)	papież
%-)	zażenowanie	:-0	zszokowanie	:-X	pocałunek
:-p	sekret	:'-)	uśmiech przez łzy	:-(zmartwienie
:ll	zły uśmiech	8-0	zdziwienie	;-)	oczko

Z Internetu pochodzi też zamieszczony na poprzedniej stronie przykład znaków ikonicznych, wykorzystujący podobieństwo użytych znaków do liter alfabetu łacińskiego (odczytanie tekstu pozostawiamy czytelnikom):

Sygnaly, które nie są podobne do komunikowanej rzeczywistości i funkcjonują jedynie na zasadzie umowy, nazywane są **znakami arbitralnymi** albo **konwencjonalnymi** (synonimem jest tu termin znak według pierwszej z przytoczonych definicji). Klasa znaków arbitralnych jest o wiele liczniejsza. Najważniejsze są tu znaki tworzące system języków naturalnych (etnicznych). Znakami arbitralnymi są też na przykład światła sygnalizacji drogowej, kształt i kolor tablic znaków drogowych, znaki alfabetu Morse'a i Braile'a, systemy dystynkcji wojskowych komunikujące stopnie, kolory wypustek w togach sądowych (np. czerwone prokuratorów, zielone adwokatów, fioletowe sędziów), kolory wypustek, pasów, su-tann, guzików i piusek, które oznaczają godność w hierarchii kościelnej (np. fiolet – biskupa, czerwień – arcybiskupa, biel – papieża). Granica między znakami ikonicznymi a arbitralnymi nie jest oczywiście granicą ostrą i zależy od tego, w jaki sposób określona jest relacja podobieństwa. I tak na przykład mapa, choć jakoś podobna do terenu, który przedstawia, jest w pewnym stopniu również konwencjonalna, zaś odczytywania map trzeba się po prostu nauczyć. W tym miejscu dokonywany przez nas podział znaków nie spełnia już warunków poprawnej klasyfikacji, która musi być rozłączna, i zbliża się raczej do typologii znaków.

Należy jeszcze zwrócić uwagę na to, że terminy znak arbitralny i znak konwencjonalny, choć najczęściej desygnują tę samą klasę, nie są równoznaczne w tym sensie, że używając terminu znak arbitralny zwracamy uwagę przede wszystkim na brak relacji „naturalnej” między znakiem a elementem komunikowanej przezeń rzeczywistości, natomiast używając terminu znak konwencjonalny mówimy o istnieniu pewnego rodzaju umowy między użytkownikami znaku ustanawiającej relację między znakiem a jego denotatem (tym, co on oznacza), a więc ustanawiającej konwencję użycia znaku.

Sygnaly arbitralne podlegają dalszemu podziałowi na **sygnaly jednoklasowe** i **sygnaly dwuklasowe**, zależnie od tego, czy funkcję informacyjną (lub znakową) pełnią już jednostki elementarne, czy też dopiero jednostki z nich złożone. Zgodnie z tym kryterium sygnałami dwuklasowymi, zwanymi w wypadku języków naturalnych, których głównym subkodem jest subkod akustyczny, **znakami**

fonemowymi, są wszystkie języki naturalne w ich podstawowym subkodzie, każdą bowiem wypowiedź języka naturalnego można podzielić na głoski, które same nie znaczą – tworzą tzw. klasę **znaków pustych semantycznie** i służą jedynie do budowania jednostek bardziej złożonych, zwanych **znakami pełnymi semantycznie**, którym dopiero przyporządkowana jest funkcja znacząca (w odniesieniu do systemu języka naturalnego należałoby mówić nie o głoskach, ale o fonemach, bo to fonemy są jednostkami systemowymi, a głoski tylko ich tekstową realizacją – dokładniej problem ten wyjaśnimy w rozdziale omawiającym system fonologiczny języka). I tak na przykład, polskie głoski *a*, *k*, *t*, (podobnie litery) nie znaczą nic, można z nich jednak zbudować jednostki znaczące, takie jak: *akt*, *kat*, *tak*, *tka*. Te nie znaczące jednostki elementarne nazywane bywają **znakami pustymi** lub **diakrytami**, zaś pełniące funkcję semantyczną jednostki złożone nazywa się **znakami pełnymi**.



Języki naturalne w subkodzie wizualnym, a więc języki pisane, mogą być jednoklasowe lub dwuklasowe, zależnie od typu używanego w nich pisma. I tak te języki, w których znaki graficzne odwzorowują, mówiąc najogólniej, ich kształt fonetyczny (tzw. pismo alfabetyczne), należą do klasy języków dwuklasowych, natomiast te, w których znaki graficzne odwzorowują bezpośrednio znaczenie ich jednostek leksykalnych (tak jest w tych językach, w których używa się pisma ideograficznego), należą, przynajmniej w pewnym stopniu, do klasy języków jednoklasowych.

Systemem znaków jednoklasowych jest najprostszy system ulicznej sygnalizacji świetlnej, gdzie każdemu światłu przyporządkowana jest jedna informacja, mówiąc najogólniej: zielonemu – wolno się poruszać, czerwonemu – nie wolno się poruszać, ewentualnemu żółtemu – przygotować się do przejścia w stan następny.

Jednoklasowość i dwuklasowość systemów znakowych bywa czasem rozumiana inaczej: za **systemy jednoklasowe** uważa się te systemy, w których występuje tylko jedna klasa znaków, a mianowicie tylko klasa znaków prostych, z których nie można tworzyć wyrażen złożonych, jedyną regułą gramatyki w takich systemach jest więc tzw. reguła o niedołączaniu, **systemami dwuklasowymi** są natomiast te, w których ze znaków prostych (elementarnych) tworzy się drugą klasę znaków, klasę znaków złożonych. I w tym rozumieniu języki naturalne są systemami dwuklasowymi. Znaki należące do jednego systemu mogą mieć charakter mieszany – na przykład w systemie znaków drogowych znajdują się zarówno znaki jednoklasowe, na przykład znak podporządkowania ‘uwaga na drogę z pierwszeństwem przejazdu’, jak i znaki dwuklasowe, do których należy większość pozostałych znaków (w rozpatrywanym tu znaczeniu tych terminów).

Każdy **znak** jest jakimś przedmiotem A, któremu może być przyporządkowany jakiś inny przedmiot B (oczywiście nie musi to być przedmiot materialny, równie dobrze może nim być pojedyncza cecha, relacja lub sytuacja, lub też wyróżniony zbiór obiektów, kompleks cech lub sekwencja sytuacji), przy czym przedmiot A, jego kształt, materiał, z którego został zrobiony, cała jego „istota” zazwyczaj nie są ważne (wyjątkiem są tu znaki ikoniczne), ważne jest tylko to, że przed-

miot A zastępuje przedmiot B, jest jego symbolem, wskazuje nań, pozwala na jego identyfikację. Dlatego często mówi się, że znak jest **przezroczysty semantycznie**, że używając go, nie myśli się o nim samym, a jedynie o tym, co on oznacza. Działa więc trochę tak, jak szkło powiększające, które pozwala dojrzeć to, co bez jego użycia nie byłoby widziane, ale samo nie zatrzymuje naszego spojrzenia.

Dlatego ten sam przedmiot B, należący do komunikowanej rzeczywistości, może być komunikowany za pomocą różnych przedmiotów A: A_1, A_2, \dots, A_n , należących do tego samego systemu znaków lub też należących do różnych systemów. Na przykład przedmiot, którego uproszczonym rysunkiem jest  może być w języku polskim komunikowany przez przedmiot *dom* oraz przez przedmiot *budynek*, a rysunek  (będący w pewnym stopniu znakiem ikonicznym) na mapie topograficznej, rysunek o kształcie *kościół* (składający się z liter alfabetu łacińskiego) będący wyrazem języka polskiego, rysunek o kształcie *church* będący wyrazem języka angielskiego oraz rysunek o kształcie *Kirche* będący wyrazem języka niemieckiego oznaczają jeden przedmiot występujący w rzeczywistości, a mianowicie budynek przeznaczony do tego, aby zbierali się w nim wyznawcy jakiejś religii po to, aby modlić się do boga.

Zazwyczaj nie jest tak, że tylko jeden przedmiot A oznacza tylko jeden przedmiot B, ale tak, że cała klasa przedmiotów A – wszystkie egzemplarze przedmiotów A odpowiednio do siebie podobnych – oznacza całą klasę przedmiotów B – wszystkie egzemplarze przedmiotów B pod jakimś – określonym – względem do siebie podobne (przy czym ta relacja podobieństwa może być różnie zadana, co umożliwia między innymi tworzenie przenośni i co w ogóle umożliwia posługiwanie się językiem), gdyby bowiem każdy wyodrębniony w rzeczywistości przedmiot miał mieć w języku swoją nazwę własną – mają ją tylko niektóre przedmioty, takie jak rzeki, góry, miasta, no i oczywiście ludzie – językiem takim nie można by się było w ogóle posługiwać ze względu na ogromne rozmiary jego słownika. Zwrócił na to uwagę już John Locke (1690), pisząc (cyt. za Wikto-rem Nalimowem, *Probabilistyczny model języka*): „To, że dźwięki można uczynić znakami idei, nie wystarcza, iżby język był doskonały, o ile znaków tych nie można używać w ten sposób, aby obejmowały łącznie wiele rzeczy poszczególnych. Mnożenie bowiem słów utrudniałoby operowanie nimi, gdyby każda rzecz poszczególna potrzebowała oddzielnej nazwy. Nicwygodę tę usuwa dalsze udoskonalenie języka przez wprowadzenie nazw ogólnych, dzięki czemu każde pojedyncze słowo oznacza mnogość bytów poszczególnych”.

W tym miejscu należy zwrócić uwagę na jeszcze jedno rozumienie znaków: znak jako typ i znak jako okaz. **Znaki okazy** to realne (materialne) przedmioty tworzące konkretny komunikat – **znaki typy** mają charakter abstrakcyjny w tym sensie, że stanowią jak gdyby matrycę powielaną w każdym konkretnym akcie tworzenia komunikatu. Znaki typy należą do systemu znakowego, znaki okazy do konkretnych komunikatów będących materialną realizacją tego systemu. Znaki typy składają się tylko z cech odróżniających je od wszystkich innych przedmiotów (w tym od innych znaków systemu), znaki okazy są realizacją tych cech w konkretnym materiale. W tym sensie znaki drogowe ustawione w konkretnych

miejscach na drodze są znakami okazami, natomiast znaki drogowe ustanowione przez kodeks drogowy są znakami typami. Wyrazy języka polskiego: JEZYK, język, JEZYK, język, JEZYK, język, JEZYK, język są znakami okazami, natomiast odpowiadający im znak typ jest strukturą konstytuowaną przez te ich cechy, które stanowią ich część wspólną i pozwalają na odróżnienie od wszystkich innych elementów rzeczywistości, w tym innych wyrażenń języka polskiego. (Do problemu tego wrócimy w rozdziale omawiającym system fonologiczny języka.)

Zwróćmy uwagę na to, że wyraz **znak** jeszcze z innego powodu ma różne znaczenia: inne kiedy mówimy, że coś jest znakiem czegoś, że dany znak ma dane znaczenie, a inne, kiedy mówimy o systemie znaków. W tym drugim wypadku znak traktowany jest jako twór złożony z dwóch na stałe ze sobą związanych elementów, przy czym jeden z tych elementów może być przypadkowo obrany, a jedynym warunkiem, jaki musi spełniać, jest to, żeby łatwo można go było odróżnić od innych przedmiotów. Element ten nazywany bywa **częścią znaczącą** (oznaczającą, formą), natomiast przyporządkowany mu element rzeczywistości nazywa się **częścią znaczoną** (oznaczaną, treścią) znaku. Taka koncepcja znaku wiąże w jedność znak w sensie węższym (takim jak w poprzednich naszych rozważaniach, gdzie oznaczał tylko element oznaczający) z jego znaczeniem. We współczesnym językoznawstwie koncepcję tę nazywa się **bilateralną teorią znaku** i wiąże się z Ferdinandem de Saussure'em, choć koncepcję taką sformułował już św. Augustyn.

Języki informacyjno-wyszukiwawcze należą do klasy sygnałów semantycznych, wyrażenia ich są bowiem wykładnikami informacji o innej niż one same rzeczywistości (tzw. rzeczywistości pozajęzykowej lub pozaznakowej). Określenie charakteru tej rzeczywistości niezbędne jest do dokonania dalszego podziału znaków, kryterium zakwalifikowania znaków albo do klasy znaków ikonicznych, albo do klasy znaków arbitralnych opiera się bowiem na ich podobieństwie lub braku podobieństwa do komunikowanych elementów rzeczywistości pozaznakowej. Języki informacyjno-wyszukiwawcze pełnią funkcję metainformacyjną (zob. rozdz. VI) – rzeczywistość, o której informują, jest rzeczywistością informacji utrwalonej w postaci dokumentów. Języki informacyjno-wyszukiwawcze ikoniczne będą to więc takie języki, których znaki podobne są do dokumentów, zaś języki informacyjno-wyszukiwawcze arbitralne takie, w których znaki nie są podobne do dokumentów. Aby wszystkie elementy naszego rozumowania uczynić jasnymi, należałoby jeszcze określić, na czym ma polegać owo podobieństwo między znakami języków informacyjno-wyszukiwawczych a dokumentami, podobieństwo jest bowiem zawsze podobieństwem pod pewnym względem. W tym wypadku chodzi oczywiście o podobieństwo czysto formalne – podobieństwo formy (kształtu) znaków języka informacyjno-wyszukiwawczego do formy znaków zawartych w dokumentach. Przy tak określonej relacji podobieństwa ikonicznymi będą więc znaki tzw. paranaturalnych języków informacyjno-wyszukiwawczych, a więc język hasel przedmiotowych, język deskryptorowy, język słów kluczowych, zaś arbitralnymi języki o notacji sztucznej, takie jak Uniwersalna Klasyfikacja Dziesiąta (z wyjątkiem poddziałów wspólnych czasu) czy Klasyfikacja

Bibliograficzna Blissa. Charakter mieszany będą miały niektóre języki spośród kodów semantycznych, np. kod semantyczny Perry'ego i Kenta (zwany też językiem ASM-WRU), który składa się zarówno ze znaków ikonicznych, jak i znaków arbitralnych. I tak wchodzące w skład jego leksyki tzw. mnożniki (czynniki) semantyczne, takie jak np. F-SH – ryby, M-LT – wojsko, S-CN – nauka, B-TN – rośliny mają, z założenia, charakter ikoniczny (w stosunku do dokumentów w języku angielskim), natomiast inne używane w nim znaki mają charakter arbitralny.

Należy jednak pamiętać o tym, że przyjętym tu kryterium rozstrzygającym o ikonicznym lub arbitralnym charakterze języków informacyjno-wyszukiwawczych jest podobieństwo do dokumentów, a właściwie ich tekstów, podobieństwo to występuje jednak tylko w przypadku, gdy są to teksty języka naturalnego tego samego, który posłużył jako materiał wyjściowy przy tworzeniu danego paraturalnego języka informacyjno-wyszukiwawczego. Język taki używamy do opisu dokumentów pisanych w innych językach naturalnych lub do dokumentów, których teksty należą do innych kodów, nie ma już charakteru ikonicznego – jest w pełni arbitralny.

Zupełnie inaczej wyglądałyby nasze rozstrzygnięcia dotyczące charakteru ikonicznego i arbitralnego znaków języków informacyjno-wyszukiwawczych przy innym określeniu komunikowanej przez nie rzeczywistości. Jeżeli przyjmiemy, że teksty języków informacyjno-wyszukiwawczych, choć bezpośrednio odnoszą się do rzeczywistości dokumentów, to pośrednio komunikują coś o tej rzeczywistości, której dotyczą zawarte w dokumentach informacje (por. rozważania o znaczeniu w językach informacyjno-wyszukiwawczych, rozdz XV), to wówczas wszystkie znaki wszystkich języków informacyjno-wyszukiwawczych, podobnie jak znaki języków naturalnych, w których zakodowane są informacje zawarte w dokumentach, będą miały charakter arbitralny.

W dalszym podziale znaków wyróżniliśmy znaki (systemy znakowe) jednoklasowe i znaki (systemy znaków) dwuklasowe, zależnie od tego, czy funkcję informacyjną pełnią już jednostki elementarne (najniższego poziomu), czy też dopiero jednostki złożone. Wśród języków informacyjno-wyszukiwawczych mamy przedstawicieli obu tych klas. Językami dwuklasowymi będą wszystkie te języki informacyjno-wyszukiwawcze, w których znaki są równokształtne ze znakami języków naturalnych o piśmie nieideograficznym. Ich elementarne znaki graficzne, podobnie jak litery alfabetu odpowiednich języków naturalnych, tworzą klasę znaków pustych semantycznie, znakami pełnymi semantycznie są złożone z tych liter elementarne jednostki leksykalne. W wypadku języków o piśmie ideograficznym, na przykład używających pisma chińskiego (np. chiński i japoński) sprawa jest bardziej skomplikowana, gdyż znaczna część ich znaków elementarnych pełni funkcję semantyczną.

Języki typu kodów semantycznych mogą mieć różny charakter, zależnie od przyjętych w nich reguł semantycznych – choć zgodnie z założeniem powinny być językami jednoklasowymi. Wymieniony wyżej kod semantyczny Perry'ego i Kenta ma charakter mieszany, gdyż znaki tworzące klasę tzw. mnożników semantycznych są znakami dwuklasowymi, natomiast znaki tworzące klasę wykładników

tw. relacji analitycznych (litery: A, E, I, O, Q, W, X, Y, Z) są znakami jednoklasowymi (np. I jest wykładnikiem relacji część-całość).

Podobny charakter ma Uniwersalna Klasyfikacja Dziesiętna. Elementarne jednostki graficzne – cyfry 0, 1, ... , 9 są znakami pustymi, znakami pełnymi semantycznie są dopiero złożone z nich wyrażenia, zwane w UKD symbolami. Ta część systemu ma więc charakter dwuklasowy. Znakami jednoklasowym są w UKD wykładniki niektórych relacji syntagmatycznych (tzw. wskaźniki).

Literatura

Guiraud P.: *Semantyka*. Warszawa 1976.

Milewski T.: *Językoznawstwo*. Warszawa 1976.

Pelc J.: *Wstęp do semiotyki*. Warszawa 1982.

Pelc J.: *O klasyfikacji i typologii znaków*. Studia semiotyczne IX, Wrocław 1979.

Saussure F.: *Kurs językoznawstwa ogólnego*. Warszawa 2002.

III

Systemy znakowe. Kody i języki

Zbiór znaków wyróżnionych ze względu na pełnioną przez nie funkcję tworzy **system znakowy** (o funkcji systemów znakowych powiemy w następnym rozdziale). Przykładem takiego systemu jest dwuelementowy system ulicznej sygnalizacji świetlnej, na który składają się światło czerwone i światło zielone. Elementy systemu tworzą zbiór spójny, każdy z każdym powiązany jest bowiem relacją w tym sensie, że określoną funkcję może pełnić tylko w opozycji do pozostałych elementów zbioru (nie ma w nim elementów izolowanych): światło zielone może oznaczać „wolno się poruszać” tylko dlatego, że światło czerwone znaczy „nie wolno się poruszać”.

Systemy znakowe nazywane są **kodami** lub **językami**, w najszerszym rozumieniu tych terminów. Często między kodami a językami czyni się dodatkowe rozróżnienia, zawiązując ich użycie do systemów znakowych o określonych cechach. Dzieje się tak dlatego, że systemami znakowymi zajmują się przedstawiciele różnych nauk, w których termin ten, w zależności od potrzeb, jest nieco inaczej rozumiany. O językach piszą filologowie, językoznawcy, teoretycy literatury, filozofowie, logicy, a także informatycy i matematycy.

Językoznawcy tradycyjnie zajmowali się tylko jednym, choć mającym najwięcej przedstawicieli, typem języków, a mianowicie **językami naturalnymi**, a więc takimi, które wykształciły się w procesie ewolucji społeczeństw ludzkich i służą ludziom do komunikowania się między sobą. Językoznawcy zakładali więc, że wszyscy, będąc użytkownikami języka, znamy go z autopsji, nie dbali więc o podawanie ścisłych definicji języka, nierzadko zresztą rezerwując ten termin jedynie na określanie naturalnych systemów porozumiewania się społeczeństw ludzkich, a więc ograniczając go tylko do języków etnicznych – wszystkie inne systemy porozumiewania się nazywali kodami.

Terminy kod i język bywają jeszcze inaczej używane, przy czym językami nazywa się wtedy najczęściej tylko systemy dwuklasowe, którymi oczywiście są języki naturalne. **Kodami** czasami nazywa się systemy mniej skomplikowane, o czym świadczy następująca definicja: „Kod. W teorii informacji oraz w dyscyplinach wykorzystujących jej aparat pojęciowy system środków służący do przekazywania określonego typu informacji, z reguły stworzony sztucznie i mający

budowę prostszą, zwłaszcza zaś mniejszą liczbę elementów prostych (liter, znaków, sygnałów elementarnych) niż system wyjściowy. (...) Wtórnie kodami nazywa się wszelki system przekazywania informacji. Kodem w tym sensie jest m.in. język naturalny.” (*Encyklopedia językoznawstwa ogólnego pod red. Kazimierza Polańskiego*).

Językoznawcy są często autorami określeń opartych na funkcji języka, takich jak np.:

„Język jest to zespół środków, za pomocą których ludzie porozumiewają się ze sobą posługując się głosem lub pismem (...)” (*Wielka encyklopedia powszechna* – autor hasła Witold Doroszewski).

„Język (...) zasób wyrazów, zwrotów i form używanych w celu porozumiewania się.” (*Słownik języka polskiego pod red. W. Doroszewskiego*).

„Język jest to właściwa gatunkowi homo sapiens i wyróżniająca ten gatunek wśród innych istot żywych – zdolność reagowania na bodźce zewnętrzne artykułowymi dźwiękami i operowania tymi dźwiękami jako znakami treści psychicznych kojarzącymi się i wywołującymi skojarzenia z przedmiotami percepcji i przedmiotami myśli na podstawie mechanizmów odruchów warunkowych.” (W. Doroszewski: *Elementy leksykologii i semiotyki*).

„Język – zasób wyrazów, zwrotów i form określanych przez reguły gramatyczne, funkcjonujący jako narzędzie porozumiewania się przez członków jednego narodu, społeczeństwa; mowa.” (*Słownik języka polskiego pod red. Mieczysława Szymczaka*).

„Język (...) to zespół społecznie wytworzonych i obowiązujących znaków dźwiękowych (względnie wtórnie pisanych) oraz reguł określających ich użycie, a funkcjonujących jako narzędzie komunikacji społecznej (porozumiewania się). Jest to więc pewien system norm będący tworem abstrakcyjnym i społecznym (...). Podstawową funkcją języka jest przedstawianie, czyli zastępowanie, za pomocą umownych znaków zjawisk świata obiektywnego.” (Z. Gołąb, A. Hein, K. Polański: *Słownik terminologii językoznawczej*)

„Język – 1. System znaków (prymarnie dźwiękowych, wtórnie pisanych i innych) służących do porozumiewania się. Obok języków naturalnych należą tu także j. programowania, j. formalne, j. zwierząt.” (*Encyklopedia języka polskiego*)

„Język – w najogólniejszym znaczeniu (...) pojmujemy jako abstrakcyjny, dynamiczny (tj. podlegający rozwojowi) system porozumiewania, na który składają się słownictwo i normy gramatyczne, realizujący się w konkretnych aktach mówienia (...).” (S. Sierotwiński: *Słownik terminów literackich*).

„Język. 1. System znaków (prymarnie dźwiękowych, wtórnie pisanych i in.) służący do porozumiewania się w obrębie danej społeczności. Jęz. jest tworem społecznym, tj. wspólnym wszystkim członkom danej społeczności (nazywanej społecznością jęz.), w przeciwnym bowiem wypadku niemożliwe byłoby porozumienie.

2. W semiotyce i w teorii informacji – każdy system znaków, który służy do porozumiewania się. Obok języków naturalnych należą tu także jęz. programowania, jęz. formalne, jęz. zwierząt.” (*Encyklopedia językoznawstwa ogólnego*).

Bogaty wybór wypowiedzi o języku podaje W. Nalimow w pracy pt. *Probabilistyczny model języka*. Oto najbardziej interesujące z nich:

„Jako że słowa mogą być porównane z literami używanymi w algebrze, to sam język można traktować jako jeden z rodzajów algebry i na odwrót, algebra to nic innego, jak język przystosowany specjalnie do objaśniania wszelkiego rodzaju wielkości.” (D. Hartley, 1749).

„Język – cały zbiór słów i wyrażeń, z którego korzysta zazwyczaj lud, naród albo rasa celem wyrażenia lub przekazania myśli, szerzej – zdolność werbalnego wyrażania myśli.” (*Encyclopaedia Britannica*, 1910-1911).

„Język jest to system znaków, którego istota polega wyłącznie na związku znaczenia i obrazu akustycznego i w którym obie części znaku są w równej mierze psychiczne (...).” (F. de Saussure, 1915).

„Pełny zbiór wypowiedzi, z których korzysta dana społeczność językowa, stanowi język tej społeczności językowej.” (L. Bloomfield, 1926).

„Język jest czysto ludzkim, nieinstynktownym sposobem przekazywania myśli, emocji i pragnień, za pomocą arbitralnie wytwarzanych symboli.” (E. Sapir, 1929).

„(...) język jawi się jako złożony system powiązanych trwale, wzajemnie od siebie uzależnionych czynników, których najbardziej nawet dokładna lingwistyka nie jest w stanie podzielić na niezależne kategorie.” (W. Matezius, 1942).

„W najszerszym tego słowa znaczeniu język jest środkiem wyrażania myśli. Krzyki niższych zwierząt stanowią język o tyle, o ile wyrażają ich stan; istnieje język kwiatów, itd.” (*Encyclopaedia Americana*, 1944).

„Termin język właściwie w pewnym sensie jest inną nazwą komunikacji, jak również słowem używanym do określania kodów, za których pośrednictwem porozumiewanie dochodzi do skutku (...) język nie jest wyłącznym atrybutem istot żywych, lecz czymś, co, w pewnym sensie, mogą one dzielić z maszynami zbudowanymi przez człowieka.” (N. Wiener, 1954).

„Język można (...) przedstawić określając pewną strukturę (kolejność klas elementów) i szereg reguł (równoważności), za pomocą których strukturę tę można przekształcić w nieskończoną ilość połączeń elementów.” (B. Uspieński, 1965).

„Język – system znaków, będący środkiem ludzkiego porozumiewania się, myślenia i wyrazu. Za pomocą języka realizuje się poznanie świata, w języku realizuje się samoświadomość jednostki. Język jest specyficznym, społecznym środkiem przechowywania i przekazywania informacji, a także narzędziem sterowania ludzkim zachowaniem.” (radziecka *Encyklopedia filozoficzna*, 1970).

„Najogólniejszym określeniem języka, obejmującym tzw. potoczne albo naturalne języki, w których operuje się słowami i zdaniami, a także sztuczne języki nauk ze specyficzną symboliką, może być następująca definicja: język to forma istnienia wiedzy pod postacią systemów znaków. Stąd również i sama wiedza istnieje zawsze w formie jakiegoś języka.” (P. Kopnin, 1971).

Na podstawie powyższych określeń niełatwo byłoby stworzyć taką definicję języka, która pasowałaby nie tylko do języków naturalnych, ale również do języków sztucznych, takich jak język matematyki, chemii czy muzyki (mówiąc

o języku matematyki i chemii mamy tu oczywiście na myśli nie terminologię tych nauk, lecz specjalnie stworzoną w tych dziedzinach sztuczną notację, czyli język wzorów matematycznych i chemicznych). Warunki, jakie powinna spełniać taka definicja, próbował sformułować Charles Morris (1946, cyt. za W. Nalimowem):

„Proponuje się pięć następujących kryteriów, które powinny być zawarte w definicji języka:

Pierwsze – język składa się ze zbioru znaków.

Drugie – w języku każdy znak posiada znaczenie wspólne dla wielu interpretatorów. Oprócz wspólnych dla wielu interpretatorów znaczeń językowych znaków mogą również istnieć znaczenia odmienne dla poszczególnych interpretatorów. Lingwistyka nie interesuje się jednak takimi znaczeniami. Ten fakt, że znak jest w jakimś stopniu indywidualny nie kłóci się z wymogiem, zgodnie z którym znak winien być zrozumiały dla grupy interpretatorów, jako że wymaga się jedynie tyle, by znak był zrozumiały dla tej grupy w pewnym tylko stopniu.

Trzecie – znaki, z których składa się język, powinny być comsigns, tzn. winny odznaczać się odtwarzalnością przez interpretatorów i posiadać to samo znaczenie dla interpretatorów i dla tych, którzy je odtwarzają. Comsigns są działaniami organizmu (np. gesty) albo rezultatami tej działalności (takimi jak dźwięki, ślady pozostawione w środowisku materialnym, skonstruowane obiekty). Na przykład zapach może być jednakowo interpretowany przez kilka organizmów w określonej sytuacji, a więc będzie zrozumiały dla grupy jednostek i mimo to nie jest on comsigns. Zapachy stałyby się znakami językowymi, gdyby poza tym, że są one zrozumiałe dla grupy jednostek, charakteryzowały się odtwarzalnością przez interpretatorów.

Czwarte – znaki, z których zbudowany jest język, są znakami wielosytuacyjnymi tzn. znakami charakteryzującymi się względną stałością znaczenia w każdej sytuacji, w której pojawia się znak należący do danej rodziny znaków. Jeśli na przykład słowo przy każdym wystąpieniu tego znaku posiadałoby nowe znaczenia, wówczas nie byłoby ono znakiem językowym, nawet gdyby było zrozumiałe dla grupy jednostek. Tak więc znak językowy to tyle, co rodzina znaków, a nie po prostu znak jednosytuacyjny, tzn. jest to środek wyrazu, jednostka mowy.

Piąte – znaki języka winny stanowić system wzajemnie powiązanych znaków, łączonych w pewien określony, lecz nie inny, sposób, tak że powstają z nich zbiory znaków złożonych, czyli procesów.

Łącznie wymagania te umożliwiają skonstruowanie definicji języka. Język jest zbiorem wielosytuacyjnych znaków, charakteryzujących się zrozumiałymi dla grupy jednostek znaczeniami, wspólnymi dla interpretatorów, przy czym znaki te mogą być odtwarzane przez interpretatorów i mogą być łączone za pomocą określonych, lecz nie innych sposobów, tworząc znaki złożone. Albo prościej, język to zbiór wielosytuacyjnych comsigns o ograniczonych sposobach połączeń. Jeśli ograniczenia nakładane na sposoby połączeń oznaczamy za pomocą słowa system, wówczas możemy powiedzieć, że język jest systemem wielosytuacyjnych comsigns. Skoro zaś rodzina znaków jest wielosytuacyjna, to najprostsza definicja będzie następująca: język to system comsigns.”

Tak sformułowana definicja języka wydaje się jednak za wąska, ogranicza bowiem stosowalność tego terminu tylko do takich systemów, które cechuje:

- mnogość użytkowników,
- posiadanie znaczenia,

nie obejmuje więc ani takich tworów nazywanych powszechnie językami, jak na przykład muzyka czy matematyka (nie posiadają znaczenia – muzyka wcale, a matematyka w powyższym znaczeniu), ani też na przykład systemów znaków używanych przez jednego tylko użytkownika – np. system stenografii używany przez żonę Dostojewskiego w jej pamiętnikach. Zgodnie z taką definicją, nie byłyby też językami języki sztuczne dopóty, dopóki nie byłyby używane przez więcej niż jednego użytkownika. Spróbujmy więc podać definicję języka wolną od tych ograniczeń.

Najogólniej mówiąc, **język** jest to system znaków, złożony z wyróżnionego zbioru obiektów oraz zbioru reguł określających dozwolone (przyjęte dla tego języka, uznane za poprawne w tym języku) połączenia tych obiektów w kompleksy. Taki wyróżniony zbiór obiektów (będący jakimś podzbiorem uniwersum) nazywany bywa, zależnie od przyjętej konwencji terminologicznej, **słownikiem**, **leksykiem** lub **alfabetem** danego języka, same zaś obiekty należące do tego zbioru (będące jego elementami) nazywane są odpowiednio: **słowami**, **wyrazami**, **wyrażeniami elementarnymi** lub **literami**. Zbiór ten najczęściej tworzy się w taki sposób, że po prostu zadaje się go listą (wybór obiektów na zasadzie często arbitralnej), można jednak określić przynależność obiektów do słownika określając jedynie cechę (zbiór cech), którą muszą one posiadać. Wtedy słownik nie stanowi klasy zamkniętej, gdyż może być zawsze powiększony o każdy nowy element posiadający tę cechę (przykładem tak skonstruowanego słownika może być podstawowy materiał dźwiękowy języka muzyki europejskiej w systemie dur-moll).

Zbiór reguł określających, jakie połączenia elementów należących do słownika uznano za należące do języka, nazywa się **gramatyką** (**składnią**, **syntaktyką**, **syntaksą**) tego języka, natomiast same te połączenia (struktury) nazywa się **wyrażeniami złożonymi** (**frazami**, **zdaniami**, **tekstami**) tego języka. Czasami jedyłą regułą gramatyki jest tzw. **reguła o niedołączaniu**, a wtedy zdania tego języka składają się tylko z jednego wyrazu (tak jest na przykład w jednoelementowym systemie sygnałów ulicznych składającym się tylko ze światła żółtego). Jeżeli wśród reguł jakiegokolwiek gramatyki istnieje reguła wykluczająca powtarzanie się wyrazów w zdaniu, to wtedy zbiór zdań tego języka jest zawsze skończony. Ograniczona jest też automatycznie długość takiego zdania i jest równa co najwyżej (jeśli nie ma dodatkowych reguł ograniczających długość) liczbie wyrazów w słowniku.

Elementy słownika danego języka teoretycznie można łączyć w dwojaki sposób: w uzyskanych tworach złożonych może być określona struktura, np. porządek liniowy (następstwo w czasie) lub mogą to być zbiory utworzone tylko na zasadzie określenia tworzących je elementów. Reguły łączenia elementów mogą dotyczyć obu tych możliwości. Przykładem języka wykorzystującego obie te możliwości może być wspomniany już język muzyki europejskiej opartej na systemie

tonalnym dur-moll, gdzie alfabet stanowią używane w nim dźwięki, zaś reguły gramatyki określają zarówno dozwolone połączenia jego elementów w kompleksy słów o dowolnej strukturze – połączenia współwystępujących w czasie dźwięków w akordy, zbudowane zgodnie z zasadami harmonii, jak i połączenia elementów o zadanej strukturze – połączenia dźwięków następujących w czasie, które to następstwo określają reguły melodii (organizacja wysokościowa następujących po sobie dźwięków) oraz reguły rytmu (organizacja czasowego następstwa dźwięków). Językiem posiadającym oba te typy reguł jest język naturalny, co pokażemy w rozdziałach omawiających budowę języka naturalnego.

Język naturalny jest systemem, w którym z jednostek niższego rzędu tworzy się, zgodnie z obowiązującymi na tym poziomie prawami składniowymi, jednostki wyższego rzędu, stanowiącymi z kolei jednostki elementarne na następnym poziomie, składane następnie w jednostki złożone, które znów mogą stanowić jednostki elementarne kolejnego poziomu. Spróbujmy dokładniej prześledzić ten proces składania „budowli” językowych z elementarnych „cegielek”, co pozwoli nam zrozumieć **strukturę języka naturalnego**.

Język naturalny, jak już wiemy, jest językiem dwuklasowym, a więc językiem, w którym wyrażenia znaczące złożone są z mniejszych jednostek, tzw. wyrażeń pustych, które same nic nie znaczą, a jedyną ich funkcją jest funkcja diakrytyczna (odróżniająca) wyrażenia znaczące (w tym i tylko w tym sensie pełnią one funkcję semantyczną). Taka budowa systemu językowego jest najbardziej ekonomiczna, pozwala bowiem z niewielu tylko „klocków”, jednostek elementarnych najniższego poziomu, uzyskać ogromne bogactwo złożonych struktur językowych. Język naturalny ma aż dwa poziomy takich wyrażeń pustych. Wyrażeniami elementarnymi języka naturalnego, a więc wyrażeniami prostymi, są tzw. **cechy dystynktywne fonemów**, realizowane w procesie mówienia w postaci elementarnych artykulacji. Te cechy dystynktywne łączą się zgodnie z prawami gramatyki tego poziomu języka w większe całości – **fonemy**, którym w tekście odpowiada ją głoski, będące najmniejszymi jednostkami tekstu mającymi przebieg czasowy. Fonemy również nic nie znaczą, służą jedynie do odróżniania bardziej skomplikowanych, złożonych z nich wyrażeń, tzw. morfemów. Te dwa rodzaje tworów językowych: cechy dystynktywne fonemów i fonemy tworzą tzw. **system fonologiczny języka**, system, jak to widzieliśmy, składający się z dwóch poziomów, a ściślej mówiąc, z dwóch generacji struktur.

Struktury wyższych poziomów języka są coraz bardziej skomplikowane, zawierają coraz więcej jednostek wyjściowych, a zasady ich łączenia są coraz trudniejsze do sformalizowania.

W opisie struktury większości języków naturalnych następnym poziomem, na którym występują już wyrażenia pełne semantycznie, a więc takie, którym przypisane są reguły interpretacji, jest **poziom morfologiczny**. Z **morfemów**, wyrażeń trzeciej generacji, najmniejszych jednostek znaczących, tworzy się **wyrazy** – wyrażenia czwartej generacji. W języku polskim operacjami tymi zawiadują dwa rodzaje reguł gramatycznych: reguły słowotwórstwa i reguły fleksji, język polski jest bowiem językiem fleksyjnym. Reguły fleksji są związane z następnym

poziomem języka, **poziomem składniowym** (jest to tradycyjna nazwa tego poziomu języka) i można je uznać za reguły transformacji jednostek leksykalnych (wyrazów) przekształcających je w postać wymaganą przez reguły nazywane w tradycyjnych gramatykach regułami składni. W językach niefleksyjnych występują innego rodzaju wykładniki relacji między wyrazami tworzącymi zdanie – omówimy je dokładniej w rozdziale poświęconym relacjom syntagmatycznym. Na tym poziomie języka uzyskuje się już jednostki pełniące funkcję informacyjną, nazywane tradycyjnie **zdaniami** – są to wyrażenia piątej generacji, najmniejsze jednostki, które zawierają element predykcji (orzekania czegoś o rzeczywistości). Zdania te stanowią z kolei jednostki wyjściowe następnego poziomu języka i łączą się w tzw. **teksty spójne**, będące wyrażeniami szóstej generacji. Niektórzy językoznawcy wyróżniają tu trzy podpoziomy: poziom zdań złożonych, poziom akapitów i poziom tekstów spójnych.

W języku naturalnym wyróżnić więc można następujące generacje wyrażen:

- I. cechy dystynktywne
- II. fonemy
- III. morfemy
- IV. wyrazy
- V. zdania
- VI. teksty spójne.

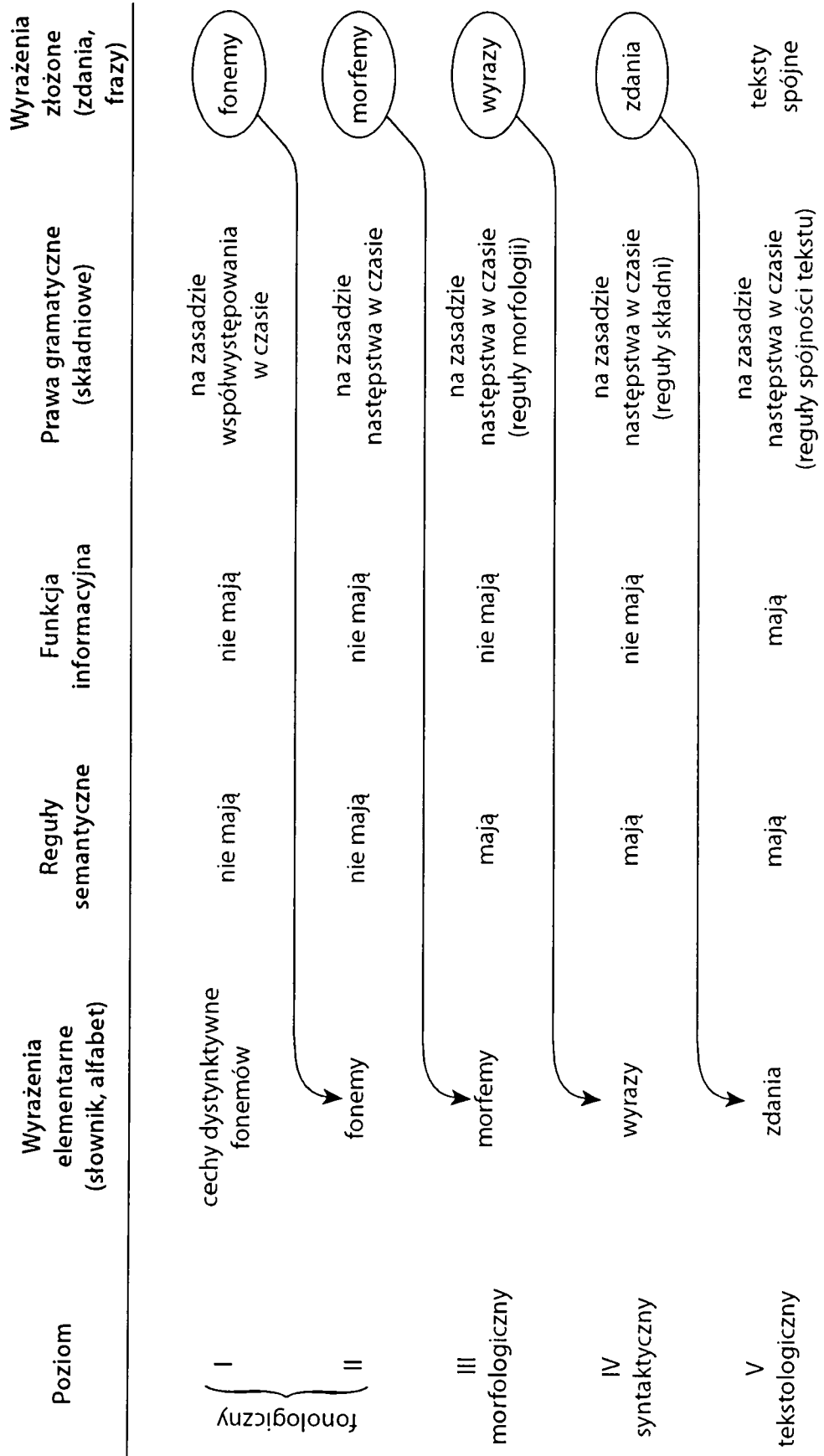
Omówioną tu strukturę języka naturalnego ilustruje załączony schemat. W następnych rozdziałach omówimy ją dokładniej.

Podobnie jak języki naturalne, również **języki informacyjno-wyszukiwawcze** mają strukturę wielopoziomową. Zazwyczaj wyróżnia się w nich następujące generacje wyrażen:

- I. elementarne znaki graficzne
- II. elementarne jednostki leksykalne
- III. zdania
- IV. teksty (charakterystyki i instrukcje wyszukiwawcze).

Przyjmuje się tu istnienie tylko jednego zbioru wyrażen semantycznie pustych – **elementarnych znaków graficznych** stanowiących alfabet języka informacyjno-wyszukiwawczego. Alfabet ten, zwany też **notacją**, jest inny dla każdego języka. Uznanie za wyrażenia proste tego typu jednostek wydaje się w wypadku tych języków teoretycznie uzasadnione i nie miałoby sensu wyodrębnianie minimalnych różnicujących je cech graficznych (pamiętajmy, że subkodem głównym tych języków jest subkod graficzny) na wzór cech dystynktywnych fonemów języka naturalnego (choć praktycznie jest to możliwe i realizowane na przykład w automatycznych czytnikach tekstów), gdyż cechy te nie byłyby elementami konstytutywnymi tych tylko języków, ich elementarne znaki graficzne są bowiem zapożyczone z innych systemów – najczęściej są to: cyfry arabskie, duże i małe litery dowolnego alfabetu języka naturalnego, np. łacińskiego, arabskiego, greckiego, cyrylicznego, znaki interpunkcyjne : , ; - / . () , symbole matematyczne > < = + , logiczne $\wedge \vee \sim$, chemiczne lub dowolne znaki zaprojektowane specjalnie dla danego języka, np. @, * itp.

Schemat struktury języka naturalnego



Z elementarnych znaków graficznych tworzy się wyrażenia pełne języków informacyjno-wyszukiwawczych. Zazwyczaj przyjmuje się, że najmniejszymi wyrażeniami pełnymi tych języków są **elementarne jednostki leksykalne** (wyrazy) i nie wyróżnia się poziomu najmniejszych jednostek znaczących, jakimi w językach naturalnych są morfemy, choć wyróżnianie morfemów wydaje się niezbędne przy opisie języków informacyjno-wyszukiwawczych zwanych kodami semantycznymi, gdyż obecnych w nim wyrażen, takich jak: np. w kodzie Perry'ego-Kenta: P-ST o znaczeniu 'klasa dokumentów, których treść ma jakikolwiek związek z owadami', B-TN o znaczeniu 'klasa dokumentów, których treść ma jakikolwiek związek z roślinami', A – oznaczającego relację przynależności do klasy, I – oznaczającego relację bycia częścią, nie można uznać za wyrazy (samodzielne elementarne jednostki leksykalne), są nimi bowiem dopiero wyrażenia z nich złożone, a więc:

BITN					klasa dokumentów, których treść ma związek z roślinami,
BITN.013	– " –	– " –	– " –	– " –	kwiatami,
BITN.026	– " –	– " –	– " –	– " –	korzeniami,
PAST	– " –	– " –	– " –	– " –	owadami,
PAST.004	– " –	– " –	– " –	– " –	muchami

Na tej samej zasadzie za morfemy w językach informacyjno-wyszukiwawczych należy uznać wskaźniki więzi, wskaźniki roli i wskaźniki wagi języków deskryptorowych oraz symbole podziałów Uniwersalnej Klasyfikacji Dziesiątej, nigdy bowiem nie stanowią one samodzielnych jednostek leksykalnych tych języków. Do problemu tego wrócimy przy omawianiu systemu morfologicznego.

Trzecim wyróżnianym poziomem wyrażen języków informacyjno-wyszukiwawczych jest poziom zdań, czwartym zaś poziom tekstów spójnych, którymi są **charakterystyki wyszukiwawcze** i **instrukcje wyszukiwawcze**. Te dwa poziomy wyrażen wprowadzono ze względu na możliwość tworzenia w niektórych językach wielozdaniowych charakterystyk wyszukiwawczych, np. w językach deskryptorowych o postaci:

TRANSPORT₁. MLEKO₁. PRZECHOWYWANIE₂. MASŁO₂.

lub

1. TRANSPORT. MLEKO.
2. PRZECHOWYWANIE. MASŁO.

(charakterystyki takie tworzy się dla dokumentów wielotematowych).

Jeżeli wyrażenia jakiegoś języka nie komunikują wyłącznie same siebie, lecz odsyłają odbiorcę do jakiejś rzeczywistości pozajęzykowej, to wówczas taki język jest **językiem z semantyką**, czyli posiada **reguły interpretacji** wyrażen, które przypisują każdemu obiektowi (zespółowi obiektów) należącemu do języka jakiś obiekt (zespół obiektów, jakąś sytuację) spoza tego języka, który jest przez to wyrażenie symbolizowany, a mówiąc ściślej, jakąś informację o tym obiekcie. Reguły te, zwane **regułami semantyki** (skrótowo **semantyką**), ustanawiają więc

znaczenie wyrażen danego języka. Taki język jest już systemem znaków w węższym sensie. Bardzo często terminu język używa się właśnie w tym węższym znaczeniu, nawet jeśli język taki jest systemem jednoklasowym. Przykładem języka niesemantycznego jest już tu wymieniany język muzyki, którego frazy nic nie znaczą.

Każdy język jest systemem znaków, a więc każdy język posiadający reguły interpretacji można uważać za twór dualny, co jest prostą konsekwencją przyjęcia bilateralnej koncepcji znaku, która traktuje znak jako twór złożony z części znaczącej – **formy**, i części oznaczanej – **treści**. Język semantyczny byłby więc systemem złożonym ze wszystkim form jego znaków tworzących jego **plan wyrażania (plan ekspresji)** i wszystkich treści jego znaków, tworzących **plan treści** języka (**plan semantyczny**). Taka definicja znaku pochodzi od szwajcarskiego językoznawcy Ferdynanda de Saussure’a, a taka definicja języka od językoznawcy duńskiego Louisa Hjelmsleva.

Poszczególne języki naturalne różnią się między sobą zarówno planem wyrażania, jak i planem treści, bo pomimo że każdy język naturalny służy do odwzorowania tej samej rzeczywistości (jest nią całe uniwersum), to każdy język, jak zobaczymy później (rozd. X), rzeczywistość tę odwzorowuje w sobie tylko właściwy sposób, wyznaczając w niej własny zbiór cech dystynktywnych. Wszystkie te cechy tworzą zbiór elementarnych informacji o rzeczywistości, które można w danym języku przekazać – przyporządkować jej jednostki planu wyrażania języka i wraz ze sposobem realizacji w systemie języka owego przyporządkowania (o czym powiemy dokładniej, omawiając sposoby odwzorowania rzeczywistości w języku, jego relacje paradygmatyczne i syntagmatyczne) nazywane bywają **siłą semantyczną języka**. Języki sztuczne również różnią się od siebie zarówno planem wyrażania, jak i planem treści, różna jest też ich siła semantyczna.

Tak zdefiniowany język (jako system) jest tworem abstrakcyjnym, gdyż ważne są w nim jedynie relacje zachodzące między należącymi doń obiektami, nie zaś same te obiekty, które mogą mieć dowolną naturę fizyczną, byleby tylko pozwalała ona na wydzielenie jednostek dyskretnych. Języki naturalne realizowane są albo w materiale akustycznym (subkod akustyczny – mowa), albo w materiale świetlnym (subkod wizualny – pismo), przy czym subkod akustyczny jest subkodem pierwotnym, głównym. Pamiętajmy jednak, że, ściśle mówiąc, te dwa subkody są różnymi systemami znakowymi, również w ich warstwie abstrakcyjnej, między tekstami języka mówionego a tekstami języka pisanego nie zachodzi bowiem relacja izomorfizmu i nawet w takich językach jak język polski, gdzie pismo odwzorowuje w pewien sposób kształt foniczny komunikatu, nie tylko brak jedno-jednoznacznego przyporządkowania między elementami tekstów różnych subkodów, ale różny jest zbiór ich jednostek elementarnych i inne reguły ich łączenia. Do problemu tego wrócimy w rozdziale poświęconym procesom komunikacyjnym.

Język jako system realizuje się w konkretnej sytuacji procesu komunikacyjnego w postaci tekstu, który z kolei może być zrozumiany jedynie przez odwołanie się do systemu językowego. W przypadku języków naturalnych mamy więc do czy-

nienia jedynie z tekstami tych języków, a z nich dopiero staramy się wyabstrahować strukturę samego języka, z której nie zawsze zdajemy sobie sprawę, choć jest ona w jakiś sposób w mózgu każdego mówiącego danym językiem zapisana (stanowi kompetencję do porozumiewania się tym językiem). Zapisanie tej struktury w sposób formalny jest sprawą bardzo trudną i daleką jeszcze od rozwiązania (na marginesie należy zauważyć, że pełna formalizacja języka naturalnego nie jest możliwa, przede wszystkim z powodu jego semantyki).

Język jako system znaków jest więc tworem abstrakcyjnym, realizowanym w postaci materialnej przez konkretnych nadawców w postaci **tekstów**, w konkretnych aktach komunikacyjnych skierowanych do konkretnych adresatów w konkretnej sytuacji komunikacyjnej. Taki tekst to komunikat (przekaz językowy). Język jest więc jak gdyby kodeksem pewnego typu zachowań mówiącym o tym, w jaki sposób wolno się danym ludziom w danej sytuacji pod względem językowym zachowywać, a w jaki nie wolno. Do wyróżnionych wyżej reguł językowych dochodzą więc jeszcze **reguły użycia języka**, zwane **regułami pragmatyki językowej** lub **regułami komunikacji**.

Językoznawcy stosunkowo późno zaczęli konsekwentnie odróżniać jednostki języka od jednostek tekstu, a odróżnianie języka jako systemu i będącego jego konkretną realizacją tekstu (w językoznawstwie znane jako opozycja pod francuską nazwą *langue – parole*) jest dziełem szwajcarskiego językoznawcy Ferdynanda de Saussure’a (1857-1913).

Tekst jest czymś istniejącym w czasie, tworem materialnym o charakterze przestrzennym – w języku natomiast wszystkie jednostki współwystępują ze sobą. Jednostki języka to jednostki, spośród których możemy wybierać te, które są najbardziej odpowiednie przy danej realizacji – jednostki konkretnego tekstu stanowią już rezultat tego wyboru. Jednostki języka są niepowtarzalne (w języku jest „przechowywany” tylko jeden wzorcowy egzemplarz danej jednostki stanowiący jak gdyby szablon, który jest powielany w tekstach) – w tekście natomiast poszczególne jednostki mogą się powtarzać. W języku jest tylko to, co wspólne wszystkim realizacjom, co w nich najbardziej istotne – w tekście również to, co niejako nieistotne, przypadkowe, uwarunkowane na przykład sąsiedztwem innych jednostek tekstu. Tak więc poszczególnym idealnym jednostkom języka odpowiadają ich konkretne realizacje w tekście. A oto odpowiedniości dla języka naturalnego (użyte tu terminy wyjaśnimy dokładniej w dalszej części podręcznika):

Język:	Tekst:
cechy dystynktywne fonemów	elementarne artykulacje
fonemy	głoski
morfemy	rdzenie, afiksy (np. formanty, końcówki)
schematy syntaktyczne (zdaniowe)	zdania (wypowiedzenia)

W podanej wyżej definicji języka przyjęliśmy za synonimiczne terminy:

- słowa, wyrazy, wyrażenia proste, wyrażenia elementarne, litery;
- słownik, leksyk, alfabet;

- fraza, zdanie, wyrażenie złożone, tekst,
- gramatyka, składnia, syntaksa,

należy jednak pamiętać, że nie zawsze terminy te używane bywają w taki właśnie sposób. Aby uniknąć nieporozumień, warto może wyjaśnić inne (zazwyczaj węższe) znaczenie tych terminów, szczególnie częste w literaturze omawiającej wielopoziomowe systemy językowe, a przede wszystkim języki naturalne.

Bardzo często terminu **litera** używa się na określenie elementarnego znaku graficznego tzw. pisma alfabetycznego lub sylabicznego języka naturalnego, teoretycznie odpowiadającego segmentowi ciągu fonicznego, a więc głosce lub sylabie, choć w praktyce litera może oznaczać kilka głosek lub też jedna głoska może być oznaczana przez kilka liter (możliwe są jeszcze inne kombinacje, o czym dokładniej w rozdziale XVIII). Litera jest więc jednostką subkodu wizualnego języka naturalnego lub innego kodu, którego wyrażenia elementarne są tożsame lub podobne do liter języka naturalnego.

Literami nazywa się również elementarne jednostki języka łączące się na zasadzie następstwa w czasie. Połączenia takich liter (a więc już jednostki złożone) nazywa się najczęściej **słowami** lub **wyrazami**. Także ten termin w odniesieniu do języka naturalnego używany jest niejednoznacznie, z czego zdaje sprawę *Encyklopedia językoznawstwa ogólnego*, gdzie czytamy:

„Wyraz. Rozumiany potocznie jako najmniejsza znacząca jednostka języka wykazująca względną samodzielność składniową jest w rzeczywistości terminem wieloznacznym, którego definicja zależna jest od płaszczyzny języka, dla której pojęcie wyrazu ma być definiowane. Niezbędne jest wyróżnienie co najmniej czterech odrębnych pojęć obejmowanych tradycyjnym terminem «wyraz»: wyraz fonologiczny, wyraz ortograficzny, wyraz gramatyczny, wyraz słownikowy, czyli leksem, wyraz tekstowy. (...) Jako synonimu terminu «wyraz» używa się niekiedy określenia «słowo».” (do problemu tego wrócimy przy omawianiu poziomu morfologicznego języka naturalnego). Przy takiej konwencji terminologicznej alfabetem nazywa się zbiór liter (w tym znaczeniu), zaś zbiór słów leksykiem lub słownikiem (słownictwem).

Tradycyjnie w językoznawstwie inaczej również używa się terminu **zdanie**, które, jak podaje *Encyklopedia językoznawstwa ogólnego* w odpowiednim artykule, „Doczekało się w literaturze językoznawczej ponad dwustu definicji, z których każda ujmowała jakiś istotny jego aspekt, ale z trudem mogła objąć wielką różnorodność struktur intuicyjnie uznawanych za zdanie.” Zdaniem, mówiąc najogólniej, nazywa się ciąg wyrazów (słów), któremu możemy przypisać wartość informacyjną, a więc zawierający leksykalny wykładnik predykatywności (czyli orzeczenie).

Jeszcze inaczej używa się terminu zdanie w logice, gdzie terminem tym obejmuje się tylko zdanie nazywane przez językoznawców zdaniem oznajmującym, a więc takie, któremu można przypisać wartość logiczną prawdy lub fałszu. O językach informacyjno-wyszukiwawczych czasami mówi się, że nie ma w nich zdań (są tylko nazwy), co oczywiście w świetle powyższego nie jest sformułowaniem poprawnym.

Reguły składania wyrazów w tak rozumiane zdania nazywa się **składnią** języka. W językoznawstwie termin fraza nie jest zazwyczaj równoznaczny z terminem zdanie, **frazą** nazywa się bowiem zgodne z prawami składniowymi połączenie wyrazów, które jednak może nie stanowić zdania, lecz jedynie jego część. Czasami frazą nazywa się ustalony konwencjonalnie zwrot językowy, np. *głowa mi pęka*. W kilku różnych znaczeniach używane bywają również terminy gramatyka i składnia (syntaktyka). Poza znaczeniem podanym wyżej, gdzie **gramatyka** oznacza właściwie całościowy kształt reguł gramatycznych, a więc reguł służących do odróżnienia znaku złożonego od przypadkowego zestawienia znaków, terminu tego używa się w językoznawstwie przynajmniej jeszcze w dwóch znaczeniach (por. Z. Gołąb, K. Polański, A. Heinz, *Słownik terminologii językoznawczej*):

„1. Gramatyką nazywa się uporządkowany i wewnętrznie powiązany ściśle określonymi stosunkami hierarchii zespół morfemów i reguł gramatycznych danego języka, przy czym przez morfemy gramatyczne należy rozumieć środki formalno-funkcyjne (np. afiksy) służące do tworzenia z morfemów leksykalnych określonych kategorii gramatycznych, przez reguły zaś zasady, według których operujemy morfemami leksykalnymi i gramatycznymi dla tworzenia wyrazów, grup wyrazowych (fraz) i zdań.

2. Gramatyką nazywa się naukę o morfemach i regułach gramatycznych (np. gramatyka historyczna, gramatyka opisowa, gramatyka normatywna, gramatyka języka polskiego). Tak rozumiana gramatyka dzieli się tradycyjnie na morfologię, czyli naukę o budowie (słowotwórstwo) i ewentualnie odmianie (fleksja) wyrazów branych w izolacji oraz na składnię, czyli naukę o wyrazach branych w relacji, czyli o budowie grup wyrazowych i zdań.”

Jeszcze inaczej termin gramatyka definiowany jest w *Encyklopedii językoznawstwa ogólnego*:

„Gramatyka. 1. Tradycyjnie nauka o środkach morfologicznych i syntaktycznych języka. Tak rozumiana gramatyka nie obejmuje ani nauki o dźwiękach (fonetyki i fonologii), ani nauki o stronie znaczeniowej wyrazów czy zdań (semantyki), ani leksykologii.

2. Kompletny opis języka. W tym znaczeniu w zakres gramatyki wchodzi: fonetyka (fonologia), morfologia, składnia, leksykologia i semantyka. Wewnętrznie powiązany ściśle określonymi stosunkami środków morfologicznych i reguł (zasad) ich łączenia (środki syntaktyczne) określa się jako system gramatyczny. (...)”

Z różnego użycia terminu **składnia** zdaje sprawę definicja podana w *Encyklopedii językoznawstwa ogólnego*:

„Składnia (syntaktyka). W tradycji gramatycznej obok morfologii dział gramatyki, którego przedmiot stanowiło łączenie się wyrazów w zdaniu. (...) Współcześnie przez składnię rozumie się teorię, która formułuje reguły tworzenia wyrażen złożonych z wyrażen prostych o funkcji ekstratekstualnej (semantycznej). Jest więc składnia zbiorem reguł (formuł) odzwierciedlających prawidłowości łączenia się takich wyrażen prostych w wyrażenia złożone. Jest to rozumienie szersze od tradycyjnego. (...)”

Zwróćmy uwagę na to, że i taka definicja składni jest węższa od przyjętej w tym podręczniku, nie obejmuje bowiem reguł łączenia wyrażen pustych kodów z semantyką i w ogóle nie odnosi się do języków niepełniących funkcji semantycznej.

W dawniejszej literaturze omawiającej języki informacyjno-wyszukiwawcze termin gramatyka używany był często niepoprawnie przy omawianiu języków deskryptorowych, a oznaczał najczęściej tylko wskaźniki więzi i roli, a więc wykładniki pewnych relacji między elementami tekstu lub między desygnatami tych elementów, natomiast języki deskryptorowe pozbawione tych wykładników nazywane bywały (niepoprawnie) językami bez gramatyki. Oczywiście, w myśl powyższych rozważań, języki bez gramatyki nie istnieją i wszystkie języki informacyjno-wyszukiwawcze mają swoją gramatykę, a więc mają ją także i języki deskryptorowe, tyle że w ich wypadku reguły składniowe rzadko kiedy bywają explicite sformułowane. Jedną z takich reguł jest reguła, w myśl której każdy deskryptor może w zdaniu wystąpić tylko jeden raz – taka reguła nie dotyczy języka naturalnego, gdzie wyrazy w zdaniu mogą się powtarzać.

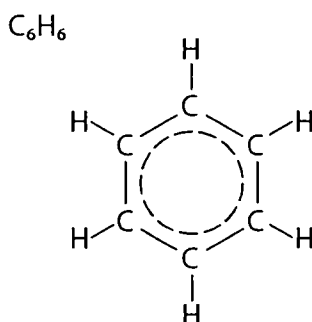
Języki, zaleźnie od tego, w jaki sposób powstały, dzieli się na naturalne (najczęściej, choć niezupełnie słusznie, utożsamiane z językami etnicznymi) i sztuczne. **Języki naturalne** są tworamii powstałymi w wyniku utrwalenia się pewnych norm dotyczących tego wszystkiego, co wspólne w procesie kontaktowania się członków określonych grup społecznych, natomiast w przypadku **języków sztucznych** konstrukcja określających je reguł poprzedza zawsze użycie tych języków.

Języki naturalne są **językami uniwersalnymi** w tym sensie, że pole semantyczne każdego z nich obejmuje całą rzeczywistość (całe uniwersum). Nie wszystkie jednak języki naturalne funkcję odwzorowywania elementów rzeczywistości pozajęzykowej pełnią wystarczająco sprawnie i nie zawsze formułowane w nich komunikaty są wystarczająco ekonomiczne w przekazywaniu informacji relewantnej dla użytkowników, nie wszystkie bowiem języki naturalne są kodami rozwiniętymi. **Kodami rozwiniętymi**, w przeciwieństwie do tzw. **kodów ograniczonych**, są tylko te języki naturalne, które cechuje bogate słownictwo abstrakcyjne, słownictwo specjalistyczne (tzw. terminologia), słownictwo konkretne o rozwiniętych, wielostopniowych relacjach paradygmatacznych, bogata frazeologia oraz odpowiednio rozbudowane struktury składniowe, pozwalające na odwzorowywania różnorodnych relacji zachodzących w rzeczywistości pozajęzykowej. Kodami rozwiniętymi są języki naturalne ludów o rozwiniętej kulturze i nauce, a naukowa wiedza pozajęzykowa odwzorowana jest w strukturze ich pola semantycznego. Ale nawet kody rozwinięte nie w każdej sytuacji komunikacyjnej okazują się wystarczająco precyzyjne i ekonomiczne i dlatego właśnie tworzy się języki sztuczne, lepiej przystosowane do pełnienia danej funkcji.

Większość **języków sztucznych** wyspecjalizowanych w funkcji semantycznej ma w porównaniu z językiem naturalnym znacznie ograniczoną rzeczywistość komunikacyjną, zawężoną do ściśle wyznaczonego fragmentu rzeczywistości odwzorowywanej przez dany język – na przykład język formuł chemicznych do świata pierwiastków, związków i reakcji chemicznych, język znaków dro-

gowych do określonej liczby sytuacji powstających na drogach i relewantnych (istotnych) ze względu na bezpieczeństwo poruszania się po nich. Dzięki takiemu ograniczeniu języki sztuczne informacje relewantne dla użytkowników mogą odwzorowywać znacznie precyzyjniej i bardziej ekonomicznie, niż czynią to języki naturalne, co wyraźnie widać, gdy porównamy chociażby następujące wyrażenia języka sztucznego i języka naturalnego:

NaCl – chlorek sodu,
H₂SO₄ – kwas siarkowy,
H₂O – woda,
C₆H₆ – benzen.



Na marginesie zwracamy uwagę na różnicę między językami specjalistycznymi a językami sztucznymi. **Języki specjalistyczne** stanowią pewien podzbiór języka naturalnego, a od języka potocznego różnią się głównie słownictwem (**terminologia naukowa**), natomiast języki sztuczne są tworem zupełnie odrębnymi, celowo konstruowanymi przez człowieka dla formułowania takich informacji, których nie da się sformułować dostatecznie precyzyjnie w języku naturalnym lub gdy sformułowania takie byłyby nieekonomiczne. Językami sztucznymi są języki informacyjno-wyszukiwawcze. Są to języki, których wyspecjalizowaną funkcją jest funkcja metainformacyjna, polegająca na odwzorowaniu treści i/lub cech formalnych dokumentów przez tworzenie ich zwężonej reprezentacji w postaci charakterystyki wyszukiwawczej, oraz funkcja wyszukiwawcza, polegająca na wyrażaniu zapytań informacyjnych i umożliwianiu odnalezienia w zbiorze informacyjnym dokumentów odpowiadających tym zapytaniom.

Języki sztuczne bywają również nazywane językami formalnymi. **Język formalny** ma więc wyraźnie określony słownik (wraz z ewentualnymi regułami interpretacji) oraz *explicite* sformułowane reguły składniowe. Współczesne językoznawstwo dąży do **sformalizowania języków** naturalnych, czyli do opisania w taki sposób ich słownika, semantyki i reguł składniowych, w jaki opisuje się języki formalne. Tak opisane języki naturalne nazywa się **językami sformalizowanymi**.

Aby sformalizować dany język (zdefiniować go) należy podać *explicite* jego słownik, zbiór reguł gramatyki, a w wypadku języków pełniących funkcję semantyczną także zbiór reguł semantyki. Należy również podać zbiór zasad użycia

komunikatów językowych w konkretnych sytuacjach komunikacyjnych, czyli reguły pragmatyki danego języka.

Język można również zdefiniować podając zbiór wszystkich jego zdań. Taki opis konkretnego języka nie jest sprzeczny z tym, co powiedzieliśmy wyżej, słownik każdego języka składa się bowiem z występujących w tych zdaniach wyrazów, reguły gramatyki zaś stanowią rezultat uogólnienia występujących we wszystkich zdaniach danego języka prawidłowości.

Wśród języków informacyjno-wyszukiwawczych możemy znaleźć przykłady obu sposobów prezentacji języka. Dotyczy to języka haseł przedmiotowych. Zwyczaj jest on definiowany poprzez podanie słownika tematów i określników oraz reguł tworzenia hasła przedmiotowego (zdań tego języka), ale język haseł przedmiotowych używanych do prezentacji zbiorów literatury medycznej występuje również w postaci zbioru zdań tego języka. W takiej formie język haseł przedmiotowych prezentowany jest także w kartotece haseł wzorcowych.

Języki naturalne tym się różnią od języków sztucznych, że reguły ich gramatyki nie są *explicite* zadane. Należą one do kompetencji językowej użytkowników tych języków, ale nie stanowią wiedzy wyraźnej, dlatego też przy opisywaniu języków naturalnych, a więc ich formalizacji, reguły składniowe muszą zostać wyabstrahowane z tekstów tych języków, bo tylko teksty mogą być poddane bezpośredniemu oglądowi. Czyniąc to, warto zastanowić się na następującymi prawidłowościami, poszukując odpowiedzi na pytania:

- Jaka jest najmniejsza liczba wyrażeń elementarnych tworzących zdanie?
- Jaka jest największa liczba wyrażeń elementarnych tworzących zdanie?
- Czy wśród wyrażeń elementarnych są takie, które muszą wystąpić w zdaniu (są obligatoryjne)?

Jeśli odpowiedź na to pytanie będzie twierdząca, to kolejne zadanie będzie polegało na podziale wyrażeń elementarnych na kategorię syntaktyczną wyrażeń obligatoryjnych i kategorię syntaktyczną wyrażeń fakultatywnych.

- Czy, i ewentualnie które, wyrażenia elementarne mogą się w zdaniu powtarzać?
- Czy wyrażenia elementarne tworzące zdania występują w określonym porządku lub tworzą określoną strukturę?

Odpowiedź na to pytanie wymaga określenia relacji, w jakich pozostają w stosunku do siebie wyrażenia elementarne tworzące zdanie (relacje takie nazywa się relacjami syntagmatycznymi).

- Czy, a jeśli tak, to jakie obowiązują prawa korelacji, czyli czy wystąpienie w zdaniu jakiegoś jednego wyrażenia elementarnego pociąga za sobą wystąpienie jakiegoś innego wyrażenia elementarnego?

Chodzi tu oczywiście tylko o zależność jednostronną, typu jeżeli A to B, która nie pociąga za sobą zależności jeżeli B to A. W wypadku zależności dwustronnej takie skorelowane na stałe elementy traktuje się jak jeden element.

Przy opisie systemu językowego dokonuje się zwykle **kategoryzacji syntaktycznej** elementów, dzieląc je na klasy abstrakcji grupujące wszystkie te elementy, które mogą występować w tym samym miejscu (może to być miejsce okre-

ślone względnie), czyli w tym samym typie kontekstu (takim samym otoczeniu) w danej formule zdaniowej. Takimi klasami syntaktycznymi są na przykład w języku polskim klasy rzeczowników, przymiotników, czasowników, przysłówków, przyimków. W wypadku tych dwóch ostatnich to względne miejsce w kontekście znalazło wyraz w nazwie kategorii syntaktycznej – przyimki to te wyrazy, które stoją przy imieniu (kategoria ta obejmuje rzeczowniki, przymiotniki i niektóre zaimki), natomiast przysłówki to te wyrazy, które stoją przede wszystkim przy słowie, który to termin oznaczał kiedyś czasownik. Takimi klasami syntaktycznymi poziomu morfologicznego języka polskiego są morfemy leksykalne (rdzenie wyrazów) oraz afiksy, dzielone na klasy zależnie od miejsca, jakie zajmują w strukturze wyrazu (np. przedrostki, przyrostki, końcówki fleksyjne – omówimy je dokładniej w rozdz. VIII).

Jeżeli w danym wyrażeniu, na przykład w składzie morfemowym wyrazu (w strukturze wyrazu), zawarty jest element (np. morfem) wskazujący na jego przynależność do danej klasy, semantycznej lub syntaktycznej, to klasę taką nazywa się **kategorią gramatyczną** (odpowiednio gramatyczną kategorią semantyczną lub gramatyczną kategorią syntaktyczną, czyli składniową), natomiast pełniący tę funkcję morfem – wykładnikiem kategorii gramatycznej lub syntaktycznej. W języku polskim takie wykładniki przynależności do kategorii syntaktycznej ma na przykład większość przysłówków – są nimi morfemy *-e* lub *-o*, (np.: *ładnie, dobrze, źle, brzydko, wesoło, smutno*), choć oczywiście nie wyznaczają one tej kategorii jednoznacznie, gdyż nie wszystkie wyrazy tradycyjnie zaliczane do kategorii przysłówków mają taki wykładnik (np.: *wyżej, wczoraj, dwakroć*), a morfemy *-e, -o* są poza tym wykładnikami innych kategorii gramatycznych, na przykład kategorii przypadku (w wyrazach *matce, matko*), w języku naturalnym bowiem nie ma zazwyczaj jedno-jednoznacznego przyporządkowania pomiędzy funkcją a jej wykładnikiem. Nie mają natomiast takiego wykładnika przyimki (np. *w, z, pod, do, od*) – przy formalizacji języka elementy tej klasy syntaktycznej muszą być zadane listą.

Zgodnie z powyższym, za błędne należy uznać spotykane w literaturze poświęconej językom informacyjno-wyszukiwawczym sformułowanie, że na przykład w językach tych występują tylko rzeczowniki lub że w językach tych nie ma czasowników, gdyż języki informacyjno-wyszukiwawcze są językami sztucznymi o zupełnie innych niż języki naturalne regułach składniowych, nie można więc przenosić na nie kategoryzacji syntaktycznej tych języków, mimo że na przykład jednostki leksykalne języków deskryptorowych są w wielu wypadkach równokształtne z rzeczownikami języka naturalnego.

Kategoria syntaktyczna (zwana też **kategorią składniową** lub **kategorią syntagmatyczną**) jest to więc klasa wyrażenń danego języka o takiej samej **dystrybucji zdaniowej**, czyli mogących występować w zdaniu w takiej samej funkcji (pamiętajmy, że termin zdanie używany tu jest w znaczeniu szerszym niż w tradycyjnym opisie języka naturalnego i oznacza każde wyrażenie złożone danego poziomu struktury języka – pokażemy to dokładniej, omawiając strukturę języka naturalnego).

W zależności od tego, czy dane wyrażenia mogą samodzielnie tworzyć zdanie, czy też nie, wyróżnia się kategorie **wyrażeń autosyntaktycznych** i **wyrażeń synsyntaktycznych**. W języku polskim do pierwszej kategorii należą na przykład czasowniki w formie słowa osobowego, gdyż frazy *Pada. Świta. Grzmi. Czytam. Odpoczywamy.* uznawane są za zdania, natomiast do kategorii drugiej należą na przykład przyimki, spójniki, przysłówki, przymiotniki, które samodzielnie zdania nie tworzą. Na poziomie morfologicznym kategorię autosyntaktyczną tworzą morfemy leksykalne (rdzenie wyrazów) – istnieją bowiem tzw. wyrazy rdzenne, np.: *las, wóz, śnieg, kwiat, dom*, natomiast morfemy gramatyczne, a więc przyrostki, przedrostki i końcówki fleksyjne tworzą kategorię wyrażen synsyntaktycznych, samodzielnie bowiem nie stanowią wyrazu.

Zależnie od tego, czy dane wyrażenia muszą występować w zdaniu, czy też tylko mogą, wyróżnia się kategorię **wyrażeń obligatoryjnych** i kategorię **wyrażeń fakultatywnych**. W języku polskim do kategorii wyrażen obligatoryjnych należą na przykład czasowniki w formie osobowej, pozostałe kategorie syntaktyczne tworzą klasę wyrażen fakultatywnych. Na poziomie morfologicznym kategorię obligatoryjną tworzą rdzenie wyrazów, zaś przyrostki i przedrostki są kategorią fakultatywną (przy pewnym opisie języka kategorią obligatoryjną są również końcówki fleksyjne, jeżeli włącza się do nich końcówkę zerową – tzw. zero morfologiczne). Na poziomie fonologicznym języka polskiego (zob. rozdz. VII) klasę wyrażen obligatoryjnych tworzą dystynktywne cechy fonemów: miejsce artykulacji i stopień zbliżenia narządów mowy, pozostałe cechy dystynktywne: udział wiązadeł głosowych (opozycja bezdźwięczność : dźwięczność), udział podniebienia miękkiego (opozycja twardość : miękkość) oraz udział jamy nosowej (opozycja ustność : nosowość) należą do kategorii fakultatywnej. Kategorii wyrażen autosyntaktycznych na poziomie fonologicznym w języku polskim brak.

W zależności od tego, czy dane wyrażenia pełnią funkcję konstytuującą zdanie, czy też nie, wyróżnia się kategorię **wyrażeń zdaniotwórczych** i kategorię **wyrażeń niezdaniotwórczych**. Przykładem mogą tu być funktory zdaniotwórcze i nazwotwórcze wyróżniane w języku logiki, w języku polskim kategorię wyrażen zdaniotwórczych tworzą czasowniki w formie osobowej.

A oto przykłady kategorii syntaktycznych wyróżnianych w językach informacyjno-wyszukiwawczych.

W języku haseł przedmiotowych wyróżnia się kategorie syntaktyczne tematów i określników. Tematy tworzą kategorię zdaniotwórczych obligatoryjnych wyrażen autosyntaktycznych, określniki kategorię niezdaniotwórczych synsyntaktycznych wyrażen fakultatywnych.

W Uniwersalnej Klasyfikacji Dziesiątej wyróżnić można na przykład kategorię syntaktyczną symboli głównych, kategorię syntaktyczną podziałów wspólnych i kategorię syntaktyczną podziałów analitycznych. Symbole proste są kategorią obligatoryjną autosyntaktycznych wyrażen zdaniotwórczych, pozostałe tworzą kategorie fakultatywnych wyrażen synsyntaktycznych, choć w niektórych przypadkach niektóre z tych wyrażen mogą być traktowane jako wyrażenia autosyntaktyczne, na przykład podziały wspólne miejsca czy podziały wspólne formy.

W niektórych językach deskryptorowych wyróżnia się na przykład kategorię deskryptorów – autosyntaktycznych obligatoryjnych wyrażen zdaniotwórczych i kategorię modyfikatorów – synsyntaktycznych wyrażen fakultatywnych. W poszczególnych językach informacyjno-wyszukiwawczych mogą występować zarówno obie te kategorie wyrażen, jak i same tylko jednostki leksykalne auto-syntaktyczne – na przykład w językach słów kluczowych i w wielu językach deskryptorowych, lub same tylko wyrażenia synsyntaktyczne – tak jest na przykład w kodach semantycznych.

Opisując język, często podaje się również **reguły transformacji** (przekształcania), które pozwalają na otrzymanie z danych wyrażen języka innych wyrażen należących do tego samego języka i zachowujących jakąś określoną cechę wyrażenia transformowanego (wejściowego). Chodzi więc o takie przekształcenia wyrażen, które nie prowadziłyby poza język, a więc uniemożliwiałyby otrzymywanie z wyrażen poprawnych wyrażen niepoprawnych. Transformacje są swego rodzaju procedurami oszczędnościowymi, pozwalającymi na tworzenie zdań ze zdań bez konieczności uprzedniego rozkładania ich na wyrażenia elementarne. Reguły transformacji są wywodliwe z reguł gramatyki. Przykładem transformacji mogą być proste przekształcenia, z jakimi mamy do czynienia w arytmetyce, na przykład zamiana ułamków zwykłych na ułamki dziesiętne. W wyniku takiej transformacji otrzymujemy np.:

$$\frac{1}{5} = 0,2$$

$$\frac{37}{8} = 4,625$$

lub sprowadzanie ułamków zwykłych do wspólnego mianownika, np.:

$$\frac{1}{3} + \frac{7}{8} + \frac{2}{5} = \frac{40}{120} + \frac{105}{120} + \frac{48}{120}$$

Przykładem transformacji wyrażen języka naturalnego jest chociażby zamienianie zdań sformułowanych w stronie czynnej na zdania w stronie biernej, dające w języku polskim na przykład następujące przekształcenia:

Jan czyta książkę. → Książka jest czytana przez Jana.
Pilka wybiła szybę. → Szyba została wybita przez piłkę.

Tę ostatnią regułę transformacyjną (zamiana strony czynnej na bierną) spróbujemy zapisać w sposób formalny, dla uproszczenia tylko dla form nieprzeczących trybu oznajmującego:

$$X_{M'L'R'O'} (V_{O'L'C'A' (CP \rightarrow R') (AND \rightarrow CT \vee CP)} \vee (BYĆ_{O'L'CF} (V_{Bo AND} \vee IPV_{L'R'AND}))) Y_{B'L'R''O''} \rightarrow$$

$$Y_{M'L'R''O''} (BYĆ_{O''L''C' (CP \rightarrow R')} \vee (V_{AD} \Rightarrow ZOSTAĆ_{O''L''C' (CP \rightarrow R'')})) IPBV_{M'L''R''A''} \langle PRZEZ X_{B'L'R'O'} \rangle$$

gdzie przyjęliśmy następujące oznaczenia:

X, Y – rzeczowniki (np. *dziecko, książka*) lub zaimki (np. *ja, to*),
V – czasownik (np. *czyta*),

IP^v – imiesłów przeszły od czasownika V (np. *czytało, przeczytali*),

IPB^v – imiesłów przymiotnikowy bierny od czasownika V (np. *czytana, przeczytana, wybita*).

Indeksy dolne przy podanych wyżej oznaczeniach oznaczają następujące kategorie:

M – mianownik,

B – biernik,

L – liczba,

R – rodzaj gramatyczny,

A – aspekt ($A = AD \vee ANd$),

AD – aspekt dokonany,

ANd – aspekt niedokonany,

C – czas ($C = CP \vee CT \vee CF$),

CP – czas przeszły,

CF – czas przyszły,

CT – forma gramatyczna czasu teraźniejszego (np. *piszę*) lub czasu przyszłego czasowników dokonanych (*napiszę*),

O – osoba

Bo – bezokolicznik.

Nazwy kategorii gramatycznych pisane u dołu (bez nawiasów) oznaczają, że dany element musi występować w formie danej kategorii. Zgodność indeksów górnych oznacza, że dane elementy muszą występować w tej samej formie danej kategorii, np.:

$$(X_l = \text{dziecko}) \Rightarrow (V_l = \text{czyta})$$

gdzie mamy zgodność co do liczby pojedynczej;

→ – oznaczenie transformacji (przekształcić na);

∨ – albo, przy czym możliwa (i konieczna) tylko jedna z dwóch możliwości;

⟨ ⟩ – ujmuje parę uporządkowaną (porządek członów taki, jak podano w nawiasie);

⇒ – znak implikacji (jeżeli to).

Transformacja ta pozwala na przykład na następujące przekształcenia:

Dziecko czyta książkę. → *Książka jest czytana przez dziecko.*

Dzieci czytały książki. → *Książki były czytane przez dzieci.*

Dziecko przeczytało książkę. → *Książka została przeczytana przez dziecko.*

Dzieci będą czytały książki. → *Książki będą czytane przez dzieci.*

Dzieci przeczytają książkę. → *Książka zostanie przeczytana przez dzieci.*
lub: *Książka będzie przeczytana przez dzieci.*

Powyższe formuły zdaniowe podają kolejność elementów w szyku zwykłym, nienacechowanym, poprawne jest jednak w języku polskim również połączenie tych elementów w dowolnej kolejności, z wyjątkiem elementów ujętych w nawias ⟨ ⟩, np.:

Książkę dziecko czyta.
Dziecko książkę czyta.
Będzie książkę czytało dziecko.
Książkę będzie dziecko czytać.
Przez dziecko zostanie książka przeczytana.
Była czytana przez dziecko książka.

W stronie czynnej możemy więc uzyskać dla czasu teraźniejszego i przeszłego 3^2 , czyli 9 poprawnych zdań, zaś dla czasu przyszłego strony czynnej oraz dla strony biernej 4^2 , czyli 16 poprawnych połączeń.

Jak widać, zapis formalny reguł transformacyjnych języka naturalnego jest zazwyczaj dość skomplikowany, dla równowagi podajemy więc jeszcze kilka prostszych zapisów reguł transformacyjnych dla języka algebry:

$$\begin{aligned}
 a + b &= b + a \\
 a \cdot b &= b \cdot a \\
 (a + b) + c &= a + (b + c) \\
 a + (-b) &= a - b \\
 a - (-b) &= a + b \\
 (a + b)^2 &= a^2 + 2ab + b^2 \\
 (a - b)^2 &= a^2 - 2ab + b^2 \\
 (a + b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\
 (a - b)^3 &= a^3 - 3a^2b + 3ab^2 - b^3
 \end{aligned}$$

Znakiem transformacji jest w tym wypadku znak =.

Po tych wszystkich rozważaniach spróbujmy zbudować sztuczny język.

Słownik tego języka zadamy listą, na którą wpisujemy następujące elementy (wyrazy):

- § . , : ; / ! " ? %

Reguły gramatyki:

- zdaniem należącym do języka będzie tylko taki ciąg elementów należących do słownika, w którym na pierwszym miejscu będzie stał element /, a na ostatnim element %.

Wtedy zdaniami tego języka będą np. ciągi:

/ %
 / § ! ? " % / / %
 / / / %
 / § ! - § : ? , % %
 / - § . : ; ! - %
 / § ! %

Nie będą natomiast zdaniami tego języka ciągi:

/ § ! % /
 % § " ! %

% _ - . : : /
§ !
/
%

Nie będzie zdaniem tego języka oczywiście również ciąg / - ! ? . . , & %, występuje w nim bowiem element &, który nie należy do słownika tego języka.

Regułą gramatyki oczywiście może być więcej. Do poprzedniej dodajmy na przykład regułę ograniczającą długość zdania, mówiącą, że:

- zdanie nie może składać się z więcej niż pięciu elementów.

Wówczas ze zdań uprzednio poprawnych za poprawne uznałibyśmy obecnie jedynie zdania:

/ %
/ / / %
/ § ! %

Możemy jeszcze dodać regułę wykluczającą powtarzanie się elementów słownika w jednym zdaniu, a wtedy pozostaną już tylko dwa poprawne zdania:

/ %
/ § ! %

Na koniec spróbujemy odtworzyć system językowy (sformalizować język) reprezentowany przez następujące zdania:

1. П 9 ▽ М
2. ↑ 5 88 7 2 9 П 2 3 3 6
3. ↑ ▽ 5 П 9 4 П
4. П 2 М 5 П 88 ▽ 6
5. П 9 М П 88 9
6. П ▽ 5 6
7. ↑ П 2 2 4 М 3 3
8. ↑ 9 6 5 7 2 4 9 3 3
9. ↑ 5 6 5 7 2 9
10. П 5 88 2 9
11. ↑ П ▽ 7 6
12. П 2 М 5 6 ▽ 88 9
13. ↑ П 6
14. ↑ 2

Słownik:

П ↑ 2 2 3 3 4 4 5 6 6 ▽ 7 88 9 9

Gramatyka:

- Zdanie rozpoczyna się albo od wyrazu П albo od wyrazu ↑,
- Jeżeli zdanie rozpoczyna się od П, to mogą w nim występować tylko: П 2 М 5 6 ▽ 88 9,

- Jeżeli w zdaniu występuje \mathfrak{R} , to występuje w nim także \mathfrak{S} .
- Wyrazy \mathfrak{E} i \mathfrak{X} występują tylko obok siebie (w dowolnej kolejności).
- Zdanie ma więcej niż jeden element.

Czytelnik zechce sam sprawdzić, na podstawie których zdań sformułowano poszczególne reguły gramatyczne oraz sformułować ewentualne pozostałe reguły.

Literatura

Bielicka L., Ścibor E.: *Wprowadzenie do języków informacyjnych*. Warszawa 1981.

Chmielewska-Gorczyca E.: *Problemy pre- i postkoordynacji w teorii języków informacyjnych*. „Zagadnienia Informatyki Naukowej” 1978, nr 2.

Nalimow W.: *Probabilistyczny model języka*. Warszawa 1976.

Saussure F.: *Kurs językoznawstwa ogólnego*. Warszawa 2002.

Sigurd B.: *Struktura języka*. Warszawa 1975.

IV

Cechy systemów znakowych

Język z semantyką jest narzędziem służącym do przekazywania informacji w tym sensie, że ustanawia wzajemne przyporządkowanie w komunikacji między sygnałem a odwzorowywaną przezeń rzeczywistością. Język, jak każde narzędzie, powinien być tworem optymalnym, powinien więc spełniać określone warunki operatywności. Tadeusz Wójcik, twórca prakseosemiotyki, czyli teorii **optymalnego znaku**, pisze o tym tak:

- „1. Komunikat międzyludzki jest narzędziem stosowanym w procesie poznawania i informowania – narzędziem semiotycznym.
2. Podstawowe postulaty operatywności (optymalności) komunikatu są uszczegółowieniem postulatów operatywności (optymalności) narzędzia. Są to:
 1. postulat trójstronnego optymalnego dostosowania komunikatu do: użytkownika komunikatu, jego twórcy i do komunikowanej rzeczywistości, jest postulatem funkcjonalnym optymalnego komunikatu,
 2. postulat wystarczania i niezbędności składników komunikatu jest postulatem jego optymalnej struktury,
 3. postulaty jednoznaczności, przejrzystości, ekonomii wysiłku twórcy i odbiorcy, izomorfizmu – to bardziej szczegółowe postulaty optymalnego komunikatu.
3. kod komunikatowy jest obrazem lub wzorem tworzywa komunikatu. Jest on klasyfikacją zbioru składników komunikatu, przedstawioną w symbolice równokształtnej z tymi składnikami.
4. optymalność kodu jest warunkiem niezbędnym (ale nie wystarczającym) optymalności komunikatu.
5. kodem optymalnym jest klasyfikacja zbioru składników optymalnego komunikatu – zbioru będącego zbiorem składników wystarczających i niezbędnych do budowy optymalnego komunikatu.” (*Prakseosemiotyka*).

Przyjrzymy się teraz cechom języka naturalnego i postaramy się ocenić, czy spełniają warunki stawiane kodom optymalnym.

Najważniejszą cechą języka naturalnego jest **dwoistość (dwuklasowość, podwójne rozczłonkowanie, podwójna artykulacja)**, która, jak wiemy, polega na tym, że funkcję znaczącą pełnią dopiero wyrażenia złożone (stanowią one tzw.

wyrażenia pełne), natomiast wyrażenia proste są **wyrażeniami pustymi**, które same nic nie znaczą i służą jedynie do różnicowania form pełnych (znaczących), należących już do klasy poziomu wyższego. I tak, wszystkie teksty języka polskiego zbudowane są z różnego rodzaju połączeń trzydziestu kilku zaledwie elementów – w tekście są to głoski (które, jak dalej zobaczymy, same są także wyrażeniami złożonymi z jeszcze mniejszych nic nieznaczących wyrażen), które same pojedynczo nic nie znaczą – znaczące są dopiero ich połączenia. W języku dwuklasowym na każdym poziomie obowiązują właściwe dla niego reguły składni, zaś wyrażeniami elementarnymi wyższego poziomu są wyrażenia złożone niższego poziomu.

Dwuklasowość jest podstawą **ekonomii systemu językowego**, ona to bowiem sprawia, że możemy przekazać dużą liczbę informacji, używając do tego jedynie niewielu elementów, stanowiących jak gdyby cegielki służące do tworzenia skomplikowanych budowli. Wszystkie języki sztuczne, które mają przekazywać dużą liczbę informacji, powinny mieć strukturę dwuklasową, jeśli mają spełniać postulat dostosowania kodu do potrzeb użytkowników komunikatu (nadawcy i odbiorcy) i komunikowanej rzeczywistości. Języki jednoklasowe byłyby w wypadku konieczności przekazywania dużej liczby informacji nieekonomiczne, musiałyby bowiem zawierać zbyt dużą, jak na możliwości operatywne użytkowników, liczbę znaczących jednostek elementarnych. O wyższości systemów dwuklasowych nad jednoklasowymi możemy się przekonać, porównując liczebność jednostek elementarnych subkodu pisanego języka polskiego i języka chińskiego. Problemem przy budowie języków dwuklasowych jest liczebność zbioru jednostek pustych. Obowiązuje tu następująca zależność: im mniejszy ich zbiór, tym dłuższe są teksty języka. Do problemu tego wrócimy przy omawianiu systemu fonologicznego języka.

Drugą ważną cechą języka jest jego **produktywność**. Cecha ta przysługuje wszystkim językom naturalnym – powinna ona cechować zresztą wszystkie języki służące do komunikowania nieograniczonej z góry rzeczywistości – a polega ona na tym, że mówiący danym językiem może powiedzieć coś, czego nigdy dotąd nie mówił ani nie usłyszał i może być przez innych znających język zrozumiany. Produktywność umożliwia też tworzenie nowych wyrażen bez naruszania systemu języka. Wyrażenia takie, dopóki nie zostały użyte w tekście, stanowią tzw. **wyrażenia potencjalne**.

Produktywność w języku naturalnym zapewniona jest na wszystkich poziomach jego struktury, choć aktualizacja wyrażen potencjalnych odbywa się przede wszystkim począwszy od poziomu morfologicznego – z morfemów tworzy się nowe wyrazy, tzw. **neologizmy**, a z nich nowe zdania i teksty. Język jest produktywny, jeżeli nie zostały w nim zrealizowane wszystkie wyrażenia złożone będące zdaniami na wszystkich poziomach struktury języka. Produktywny jest oczywiście każdy język, którego gramatyka umożliwia powtarzanie tych samych wyrażen w zdaniu (**reguła rekurencji**) – tak jest w języku naturalnym – zbiór zdań takiego języka jest potencjalnie (teoretycznie) nieskończony, teoretycznie, co więcej, nieskończona jest wtedy także długość zdania, choć oczywiście prak-

tycznie jest zawsze skończona. Przykładem może tu być notacja dziesiętna, język mający tylko dziesięć wyrażen elementarnych, służący do zapisywania liczb, których zbiór jest nieskończony.

Każdy język naturalny jest narzędziem umożliwiającym przekazywanie za pomocą ustalonej, stosunkowo niewielkiej liczby jednostek elementarnych, nieskończonej liczby nowych informacji. Takiej możliwości nie mają wszystkie języki, w niektórych bowiem liczba i treść informacji, która ma być przekazana, musi być z góry przewidziana i każdej informacji musi być z góry przyporządkowany jej wykładnik. Przykładem takiego języka może być chociażby najprostszy system ulicznej sygnalizacji świetlnej dla pieszych. Taki system jest nieproduktywny, bo nie można za jego pomocą przekazać żadnej innej informacji poza z góry przewidzianymi. Gdyby zaistniała taka potrzeba, trzeba by po prostu system rozbudować (a więc zmienić), dodając mu nowy wykładnik tej nowej informacji.

Przekazywanie nieskończonej liczby informacji w tym sensie, że każda informacja o jakichś obiektach i sytuacjach jest inną informacją, umożliwia również cecha znaku, na którą już zwracaliśmy uwagę, a mianowicie ta, że forma znaku przyporządkowana jest nie jednemu obiektowi, lecz całej klasie elementów, konstytuowanej przez relację podobieństwa między nimi ze względu na jakieś cechy relewantne dla użytkowników tego znaku. Dzięki temu na przykład wyraz *góra* oznaczać może nie jakąś jedną konkretną górę, ale wszystkie. Gdyby tak nie było, język komunikujący nieskończoną rzeczywistość musiałby się składać z nieskończonej liczby wyrażen, co uniemożliwiłoby jego funkcjonowanie (a nawet skonstruowanie).

Trzecią cechą przysługującą językom naturalnym jest **arbitralność**. Polega ona na tym, że w języku nie ma żadnego związku naturalnego ani koniecznego między obiektami należącymi do języka a obiektami rzeczywistości pozajęzykowej przez nie oznaczanymi. Przyporządkowanie to polega jedynie na konwencji – obiekty należące do języka symbolizują jedynie obiekty należące do rzeczywistości pozajęzykowej. Twierdzeniu temu nie przeczy występowanie we wszystkich chyba językach naturalnych znaków częściowo ikonicznych, takich jak np. w języku polskim dźwiękonaśladowcze wyrazy *kukulka*, *gwizdać*, *świst*, *szeptać*, *zgrzytać*, na co najlepszym dowodem jest to, że znaczenia tych wyrazów musimy się uczyć tak samo, jak znaczeń innych jednostek leksykalnych.

Wydawać by się mogło, że systemy złożone ze znaków ikonicznych lepiej spełniają postulat dostosowania do użytkowników komunikatu i do komunikowanej rzeczywistości – są podobne do komunikowanych elementów, łatwiej je zatem zapamiętać i domyślić się ich znaczenia. Ikoniczność możliwa jest jednak tylko w wypadku małych systemów znakowych, o ograniczonej liczbie odwzorowywanych informacji – znaki ikoniczne funkcjonują w różnego typu systemach informacji wizualnej, na przykład w systemie znaków drogowych. Nie jest to możliwe w wypadku systemów dużych, gdyż duża liczba znaków prowadzi do zatarcia się cechy podobieństwa. Przykładem może tu być ewolucja pierwotnie opartych na podobieństwie subkodów pisanych języków naturalnych, jak na przykład pisma chińskiego.

Czwartą cechą języków naturalnych jest **zdolność do wzajemnej wymiany**, która polega na tym, że każda jednostka uczestnicząca w procesie informacyjnym (nadawca i odbiorca komunikatu) może, korzystając z tego samego kodu, występować w obu rolach, a więc nadawać i odbierać komunikat w tym samym kodzie. Język naturalny jest więc **kodem nadawczo-odbiorczym**. Cechy tej nie posiadają wszystkie języki, na przykład języki programowania służą programiście tylko do nadawania informacji.

Język naturalny jest też systemem **wyspecjalizowanym pod względem informacyjnym**, a specjalizacja ta polega na tym, że zachowanie językowe człowieka jest głównie zachowaniem płynącym z chęci uczestniczenia w procesie komunikacyjnym i nie jest niczym w sposób konieczny uwarunkowane (mogę daną informację nadać, jeżeli zechcę, ale mogę jej nie nadawać). Również odebranie komunikatu nie powoduje w sposób konieczny żadnej innej niż informacyjna zmiany stanu odbiorcy nawet wtedy, gdy komunikat pełni funkcję ekspresywną, jak to na przykład dzieje się w przypadku wszelkiego typu rozkazów i zaleceń – odbiorca rozkazu może go przecież nie wykonać, jeśli nie zechce, sam rozkaz bowiem, jako twór językowy nie zmusza go do tego.

Wartość informacyjną mogą mieć również niejęzykowe zachowania człowieka. I tak, na przykład, informację o tym, że Jan ma zamiar położyć się spać, mogę uzyskać obserwując, jak Jan ściele łóżko, myje się, ubiera w piżamę. Jednak w tym wypadku Jan wykonywał wszystkie te czynności nie po to, żeby mi zakomunikować swój zamiar pójścia spać, tylko po to, żeby faktycznie przygotować się do snu. Zachowanie jego nie było więc wyspecjalizowane pod względem informacyjnym, a wartość informacyjną miało jedynie dla odbiorcy (na zasadzie symptomu). Tę samą informację mogę uzyskać jednak również wtedy, gdy Jan po prostu powie mi o swoim zamiarze pójścia spać, przy czym takiemu zachowaniu językowemu nie muszą towarzyszyć żadne inne czynności. Oczywiście nie wszystkie języki są wyspecjalizowane pod względem informacyjnym w pierwszym z omawianych tu znaczeń tego terminu – nie mają jej przede wszystkim te języki, których główną funkcją jest sterowanie jakimiś procesami, a więc niewyspecjalizowany informacyjnie jest np. kod genetyczny oraz języki programowania, gdyż w ich wypadku odebranie informacji równoznaczne jest z podjęciem przez odbiorcę określonego działania (zmiany stanu).

Językowi naturalnemu przysługuje również **zdolność przemieszczania komunikatu** w stosunku do komunikowanej rzeczywistości, która sprawia, że przekaz językowy może odnosić się do zjawisk w rzeczywistości pozajęzykowej oddalonych od zachowania informacyjnego zarówno w przestrzeni, jak i w czasie. Możemy więc przekazywać informację o takich zdarzeniach, których nie byliśmy świadkami, możemy mówić o wydarzeniach przeszłych i przewidywanych przyszłych, o wydarzeniach dziejących się gdzieś indziej. Ta ostatnia cecha być może wydaje się na tyle oczywista, że niewarta jest wymieniania, gdyż trudno sobie nawet wyobrazić, aby mogło być inaczej. Jednak bywają wypadki, że ludzie, którzy wskutek afazji stracili część swoich kompetencji językowych, mogą wypowiadać tylko zdania mówiące o tym, co w danej chwili się dzieje. Na przykład

cechy zdolności przemieszczania komunikatu nie ma system znaków, jakim pies komunikuje się ze swoim panem, pies może go bowiem poinformować swym zachowaniem tylko o tym, że na przykład w chwili nadawania komunikatu chce iść na spacer, nie może go jednak poinformować o tym, że chciał tego przedwczoraj lub że chce pójść na spacer jutro.

Ostatnią z ważnych cech wyróżniających języki naturalne spośród innych systemów komunikacyjnych funkcjonujących w przyrodzie jest **przekazywanie języka w ramach kultury (transmisja kulturowa)**. Polega ona na tym, że znajomość języka naturalnego nie jest człowiekowi przekazywana genetycznie ani też nie jest genetycznie uwarunkowana zdolność do nauczenia się określonego języka, choć jest w ten sposób przekazywana zdolność do nauczenia się języka w ogóle. Konwencje języka naturalnego są przekazywane z pokolenia na pokolenie w ramach kultury – każdy człowiek może się nauczyć dowolnego języka, czyli uzyskać **kompetencję językową**, i musi się nauczyć, nie może bowiem w żaden inny sposób posiadać zdolności do posługiwania się nim.

Językowi naturalnemu przypisuje się zwykle cechę **linearności**, rozumiejąc zazwyczaj przez to, że jednostki języka łączą się ze sobą w następstwie czasowym. Twierdzenie to jest prawdziwe, z dwoma wszakże uzupełnieniami:

- elementarne jednostki języka, którymi są cechy dystynktywne fonemów, łączą się na zasadzie współwystępowania w czasie (o czym dokładniej w rozdziale omawiającym system fonologiczny języka).
- elementy suprasegmentalne, takie jak intonacja zdaniowa, choć same mogą mieć charakter linearny, łączą się na zasadzie współwystępowania w czasie z elementami leksykalnymi. I tak, różne intonacje można złożyć z tym samym ciągiem leksykalnym, np.

czytales tę książkę

uzyskując w ten sposób komunikaty o różnym znaczeniu, czemu odpowiadają różne postaci graficzne:

Czytales tę książkę.

Czytales tę książkę?

Czytales tę książkę!

Czytales? Tę książkę?

Często zdarza się, że podczas przesyłania przez kanał sygnał nadany ulega zniekształceniu, co może spowodować stratę informacji. Przykładem takiego zniekształcenia może być zatarcie liter tekstu pisanego, zniekształcenie tekstu przesyłanego telegraficznie, zniekształcenie tekstu słyszanego na tle dużego hałasu. Usuwaniem takich zniekształceń zajmuje się odbiornik (lub odbiorca), który w oparciu o znajomość systemu językowego i znajomość komunikowanej rzeczywistości stara się odtworzyć zniekształcony tekst w formie najbardziej wiernej systemowym cechom tekstu nadanego. Tak postępuje człowiek odczytujący niewyraźnie napisany list lub usiłujący zrozumieć słowa słyszane na tle dużego hałasu. Wynikiem takiej operacji przeprowadzanej przez odbiorcę na odebranych

komunikacie może być tekst zawierający więcej cech systemowych niż tekst faktycznie nadany, nadawca mógł bowiem wyprodukować tekst obciążony różnymi wadami, jak to się dzieje na przykład przy niestarannej artykulacji tekstu mówionego lub niewyraźnym charakterze pisma.

Każdy dobry język musi być więc wyposażony w specjalny mechanizm obronny zapobiegający stracie informacji w przypadku niezrealizowania przez nadawcę lub nieodebrania przez odbiorcę wszystkich cech przekazu informacyjnego. Mechanizmem takim jest tzw. **redundancja systemowa** – nadmiarowość informacji, polegająca na tym, że komunikat zawiera więcej wykładników informacji niż jest to konieczne do jej przekazania. Dzieje się tak wtedy, gdy:

- poszczególne jednostki złożone języka różnią się od siebie więcej niż jedną cechą;

albo:

- ten sam element rzeczywistości komunikowany jest w tekście przez kilka wykładników, na przykład w polskim zdaniu:

Umorusani, zziębnięci, ubrani tylko w lachmany, straszliwie zabiedzeni, wychudzeni mali chłopcy nie odstępowali nas ani na chwilę i przez całe popołudnie laźli za nami ulicami miasta.

Informacja o tym, że chodzi tu o więcej niż jedną osobę, przy czym przynajmniej jedna z tych osób musi być płci męskiej, podawana jest aż dziewięć razy – wykładnikami tej informacji są końcówki liczby mnogiej rodzaju męskoosobowego przymiotników, rzeczownika i czasowników, przy czym ani razu nie można tej informacji pominąć, gdyż naruszyłoby to poprawność zdania. W tym wypadku redundancja założona jest więc przez system języka. Gdybyśmy jednak zdanie to przekształcili na czas teraźniejszy

Umorusani, zziębnięci, ubrani tylko w lachmany, straszliwie zabiedzeni, wychudzeni mali chłopcy nie odступują nas ani na chwilę i przez całe popołudnie lażą za nami ulicami miasta.

to informację o tym, że chodzi przynajmniej o jedną osobę płci męskiej moglibyśmy już podać tylko siedem razy, bo końcówki czasownika czasu teraźniejszego nie wyróżniają męskoosobowości, choć w dalszym ciągu wyrażają mnogość.

Redundancję mają również takie wyrazy języka polskiego, jak np.: *system, język, most*, nie ma bowiem w języku polskim takich wyrazów, które różniłyby się od nich tylko jednym elementem (jedną cechą). Nie mają natomiast redundancji wyrazy: *kość, kosa, trasa*, bo tylko jedną cechą różnią się od nich wyrazy *gość, koza, krasa*.

Redundancja w językach naturalnych występuje na wszystkich poziomach, na których istnieją jednostki złożone i wynosi około 50%. Łatwo się o tym przekonać, usiłując odczytać tekst języka pisanego, w którym usunięto co drugą literę lub tekst, w którym zakryto górną lub dolną połowę wiersza – w ten sposób sprawdzimy oczywiście redundancję występującą w znakach graficznych służących do utrwalania tekstów języka naturalnego. Będzie to tym łatwiejsze, im

dłuższy jest tekst, mechanizm redundancji przejawia się bowiem także w określonym prawdopodobieństwie współwystępowania i następstwa w tekstach określonych jednostek języka. Potwierdza to niżej przytoczony przykład (pochodzący z Internetu), którego tekst, właśnie dzięki temu mechanizmowi, bez trudu można odczytać.

*Zdognie z nanjwoymszi baniadmai perzporawdzomyni na bytyrijskch uwe-
niretasytch nie ma zenacznia kojnolesć ltier przy zpiasie dengao solwa.
Nwajżanszyeim jest, aby prieszwa i otatsnia lteria byla na siwom mijsecu,
ptzosaloe mgoq być w niaedzile i w dszaly m cągu nie pwinono to sawrztać
polbemórw ze zozumierniem tksetu. Dziejee się tak datgelo, że nie czamyty
wyszistkch lteir w solwie, ale clae solwa od razu.*

Ten sam mechanizm działa w wersji angielskiej:

*Aoccdrnig to a rscheearch at an elingsh uinervtisy, it deosn't mtttaer in waht
oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht frist and lsat
ltteer is at the rghit pclae. The rset can be a toatl mses and you can sitll raed
it wouthit porbelm. Tihis is bcuseae we do not raed ervey lteter by istlef but
the wrod as a wlohe.*

Nie mają natomiast redundancji niektóre języki sztuczne – do takich języków należy Uniwersalna Klasyfikacja Dziesiętna, w której każda cyfra w symbolu pełni funkcję znaczącą (redundancja w formie pisanej tego języka wynika tylko z zapisu cyfr). Redundancja w języku deskryptorowym wynika z redundancji języka naturalnego, na którym dany język deskryptorowy jest oparty.

Termin redundancja bywa używany, szczególnie w tekstach technicznych, także w innym znaczeniu, gdzie redundantnym nazywa się tekst pozwalający na więcej niż jedną interpretację, a więc po prostu tekst wieloznaczny.

Dobry język to taki język, który oszczędza wysiłek jego użytkowników. W procesie odbierania informacji bardzo ważną operacją jest podział porcji sygnału kodującego informację na poszczególne części, którym przyporządkowane są elementy odwzorowywanej w komunikacie rzeczywistości. Sprawa podziału tekstu (**delimitacji, segmentacji**) na mniejsze całości jest niezwykle ważna, prawidłowe rozczłonkowanie tekstu jest bowiem warunkiem koniecznym jego rozumienia. Najlepiej o tym świadczą następujące przykłady różnych możliwości podziału, a co za tym idzie i różnych znaczeń, następujących ciągów:

n a w ó z

nawóz

na wóz

p a l i s a d y

palisady

pali sady

b a r k a r o l a

barkarola

bar Karola

barka rola

<i>k o c h a n a d a r e m n i e</i>	<i>kocha nadaremnie</i> <i>kochana daremnie</i>
<i>j a n i e k o c h a m c i ę</i>	<i>ja niekocham cię</i> <i>Janie kocham cię</i>
<i>a d a m n i e m a</i>	<i>a dam nie ma</i> <i>Adam nie ma</i> <i>Ada mniema</i> <i>Ada mnie ma</i> <i>a da mnie ma</i>

Możliwość różnego podziału tekstu wykorzystana została przez Jerzego Ficowskiego w wierszu *Dziwna rymowanka*:

*Pewien żarłok nienażarty
raz wygłodniał nie na żarty
i wywiesił szyld na płocie,
że ochotę ma na płocie.
Tutaj na brak ryb narzeka,
bo daleko rybna rzeka.
Więc się zgłosił pewien żebrak
i rzekł żarłokowi, że brak
płoci, karpia oraz śledzi,
ale rzeki pilnie śledzi
i gdy tylko będzie w stanie,
to o świcie z łóżka wstanie,
poczym ruszy na Pomorze
i w zdobyciu ryb pomoże ...
Odtąd żarłok nasz jedynie
zamiast smacznych ryb je dynie.*

W każdym dobrym języku powinny istnieć środki umożliwiające podział tekstu (w języku naturalnym, jak widzieliśmy, nie zawsze jednoznaczny). Przede wszystkim istnieją środki umożliwiające wyodrębnienie w tekście jednostek niosących informację, czyli zdań – służy temu akcent zdaniowy. Zdania trzeba jednak podzielić na jeszcze mniejsze jednostki – wyrazy – temu celowi służy, na przykład, akcent wyrazowy. W tych językach, gdzie ma on ustalone miejsce, jak na przykład w polskim, gdzie z reguły pada na przedostatnią sylabę wyrazu, lub w czeskim, gdzie pada na sylabę pierwszą, akcent dość dokładnie wyznacza granicę między wyrazami. W językach, w których miejsce akcentu jest zmienne, jak na przykład w rosyjskim, pełni on funkcję liczącą jednostki leksykalne, a więc wyznacza, w przybliżeniu, liczbę segmentów, na jaki tekst należy podzielić. Funkcję delimitatora tekstów języka naturalnego w niektórych językach, na przykład w języku węgierskim i tureckim, pełni tzw. harmonia wokaliczna, która polega na tym, że morfemy tworzące jeden wyraz mają samogłoski należące do tego samego szeregu artykulacyjnego (wszystkie przednie albo wszystkie tylne), przy czym

wybór wariantu morfemu gramatycznego zależy od samogłoski morfemu leksykalnego, np. w tureckim wykładnik liczby mnogiej ma dwie postaci *-ler* i *-lar*, a wybór zależy od postaci rdzenia – mamy więc *evler* ‘domy’ i *oklar* ‘strzały’.

Definicja wyrazu oparta na kryterium akcentu nie jest jednak wystarczająca, za jeden wyraz należałoby bowiem uznać znajdujące się w jednym zestroju akcentowym przyimki z imionami, z którymi się łączą, na przykład z rzeczownikami: w *mieście*, *na wsi*. Dlatego też często wprowadza się kryterium dodatkowe – za jeden wyraz uważa się więc tylko taki ciąg znajdujących się w jednym zestroju akcentowym morfemów, w którym nie da się już wstawić żadnego dodatkowego samodzielnego morfemu. Na mocy tego kryterium połączenia morfemów w *mieście*, *na wsi* nie będą pojedynczymi wyrazami, można bowiem powiedzieć w *naszym mieście*, *na polskiej wsi*. Dwoma wyrazami będą natomiast pisane w języku polskim łącznie morfemy typu *nieładny*, można bowiem powiedzieć *nie bardzo ładny*.

Rolę wskaźnika granic wyrazu pełni również jego schemat strukturalny – wskaźnikiem końca wyrazu będzie w tym wypadku końcówka fleksyjna lub postfiks, natomiast wskaźnikiem początku pierwszy przedrostek.

Wszystkie te kryteria – ale tylko zastosowane łącznie – pozwalają na w miarę dokładną delimitację (segmentację) tekstu oraz na wydzielenie jednostek poszczególnych poziomów systemu językowego.

Jednym z najważniejszych postulatów optymalnego kodu, sformułowanych przez Tadeusza Wójcika, jest jego izomorfizm względem odwzorowywanej rzeczywistości. **Izomorfizm** to relacja między dwoma zbiorami polegająca na tożsamości ich struktury, a więc taka, że składniki jednego z nich można tak przyporządkować składnikom drugiego, że formalne własności relacji zachodzących między składnikami wewnątrz jednego z tych obiektów są takie same jak formalne własności relacji, które zachodzą między odpowiednikami tych składników w drugim obiekcie. Warunkiem izomorfizmu jest przyporządkowanie wzajemnie jednoznacznie (jedno-jednoznaczne) relewantnych elementów tych obiektów. W wypadku języków z semantyką izomorfizm powinien więc zachodzić między językiem a odwzorowywaną przez niego rzeczywistością. Do problemu izomorfizmu wrócimy przy omawianiu struktury semantycznej języka naturalnego. Tu ograniczymy się tylko do stwierdzenia, że języki naturalne warunkiem tego nie spełniają.

Literatura

Hockett Ch.F.: *Kurs językoznawstwa współczesnego*. Warszawa 1968.

Weinsberg A.: *Językoznawstwo ogólne*. Warszawa 1983.

Wójcik T.: *Prakseosemiotyka. Zarys teorii optymalnego znaku*. Warszawa 1969.

V

Funkcje systemów znakowych

Język, rozumiany jako abstrakcyjny system, realizowany jest w materialnej postaci tekstu przez konkretnego użytkownika w konkretnej sytuacji komunikacyjnej. Taka realizacja nazywana jest **komunikatem (przekazem)**. Dotyczy to wszystkich języków (kodów, systemów znakowych) i w tym sensie wszystkie one pełnią **funkcję komunikatywną** (w najszerszym rozumieniu tego terminu), komunikatem jest bowiem zarówno tekst języka naturalnego, na przykład prognoza pogody nadana przez radio lub wydrukowana w gazecie, tekst języka informacyjno-wyszukiwawczego (np. charakterystyka wyszukiwawcza konkretnego dokumentu), jak i utwór muzyczny w zapisie nutowym lub nagrany przez konkretnego wykonawcę, a także znak drogowy ustawiony przy drodze.

Oprócz tak zdefiniowanej funkcji komunikatywnej (bywa ona rozumiana wężziej, jako komunikowanie czegoś, a wtedy jest bliska omówionej dalej funkcji semantycznej), systemy znakowe mogą pełnić inne funkcje, bardziej wyspecjalizowane, określane ze względu na relacje łączące język z poszczególnymi elementami sytuacji komunikacyjnej. Funkcje te omówimy przede wszystkim na przykładzie funkcji pełnionych przez język naturalny, dla niego bowiem zostały one określone.

Najważniejszą funkcją języka naturalnego jest **funkcja semantyczna** polegająca na tym, że elementy języka naturalnego służą do odwzorowania w komunikacie jakiejś rzeczywistości. Funkcję taką pełni nie tylko język naturalny, ale wszystkie języki z semantyką. Język naturalny jest kodem najbardziej uniwersalnym w tym sensie, że odwzorowane jest w nim całe uniwersum, w każdym języku naturalnym można bowiem wyrazić informacje o relewantnych dla użytkownika tego języka elementach rzeczywistości. Stąd funkcja semantyczna nazywana bywa również **funkcją informacyjną, funkcją reprezentatywną** (bo wyrażenia języka reprezentują elementy innej rzeczywistości), **funkcją przedstawieniową, referencjalną, denotatywną, kognitywną** lub **symboliczną**. Wszystkie te terminy uznać należy za bliskoznaczne, bo choć odzwierciedlają tę samą relację, relację między elementami języka a elementami rzeczywistości, są jednak nieco inaczej definiowane ze względu na różne określenie samej reprezentowanej przez język

rzeczywistości (może to być świat rzeczywisty lub uniwersum informacyjne), jak i ze względu na różne pojmowanie samej relacji reprezentacji.

Inne języki z semantyką są językami wyspecjalizowanymi w tym sensie, że służą do odwzorowywania jakiegoś określonego (wybranego) wycinka rzeczywistości (komunikowania określonego zbioru informacji). I tak, na przykład, język znaków drogowych komunikuje wybrane informacje relewantne dla poruszających się po drogach, język symboli chemicznych relewantne informacje o strukturze związków chemicznych, a notacja muzyczna relewantne cechy dźwięków danego systemu muzyki. Funkcję semantyczną języka naturalnego omówimy dokładniej przy omawianiu systemu semantycznego.

Funkcja semantyczna, polegająca na komunikowaniu rzeczywistości, choć jest najważniejszą funkcją każdego języka naturalnego, nie jest jedyną funkcją pełnioną przez język. Jedną z nich jest **funkcja ekspresywna**, polegająca na tym, że komunikat językowy może być źródłem informacji o swoim nadawcy bądź przez fakt celowego wybrania do realizacji tych, a nie innych elementów systemu językowego, bądź poprzez pozasystemowe cechy sygnału tworzącego komunikat. Każdy komunikat języka naturalnego (przekaz językowy) nie tylko bowiem stanowi odwzorowanie jakiegoś wycinka rzeczywistości, ale jest dla odbiorcy również źródłem dodatkowych informacji o samym nadawcy, który często sam, przez specjalne ukształtowanie komunikatu językowego, chce zakomunikować odbiorcy na przykład swój stosunek do komunikowanych elementów rzeczywistości. Funkcję ekspresywną pełni wiele elementów systemu językowego, począwszy od elementów fonetycznych, poprzez leksykalne, aż do składniowych. W ogólnej ekspresji komunikatu ogromną rolę odgrywa przede wszystkim intonacja oraz tempo i natężenie mowy, przy czym środki te mogą być wykorzystywane przez nadawcę zarówno w sposób świadomy, jak i nieświadomy, a dla odbiorcy komunikatu stanowią informację o stanie nadawcy na zasadzie symptomatycznej. W sposób świadomy natomiast może kształtować nadawca komunikatu jego warstwę głoskową, co wykorzystywane jest na przykład wtedy, gdy się mówi do dzieci lub o dzieciach (tzw. spieszczenia). Ekspresywnymi wykładnikami w komunikacie mogą więc być zarówno systemowe środki języka, jak i środki pozasystemowe, a więc w komunikacie mówionym te cechy akustyczne, które nie są wykorzystywane w danym systemie języka do kodowania informacji – o ekspresywnej funkcji języka należałoby mówić tylko w przypadku tych pierwszych. O wadze funkcji ekspresywnej języka świadczy najlepiej fakt, że wykształciły się w nim wyspecjalizowane środki gramatyczne, na przykład w języku polskim formanty, za pomocą których można tworzyć tzw. zdrobnienia lub zgrubienia, takie jak np.: *-ek, -eczek, -uś, -unio, -qtko, -ulo, -ulek, -usio* (np.: *kot, kotek, koteczek, kotuś, kotunio, kociątko, kotulo, kotulek, kotusio*) wyrażające pozytywny stosunek mówiącego do komunikowanego elementu rzeczywistości (w tym wypadku do kota); *-isko* (np. *kocisko*) wyrażający negatywny stosunek mówiącego do elementów rzeczywistości (jest to funkcja prymarna tego przyrostka, choć możliwe jest również użycie go w funkcji pozytywnej – tego typu zamiana funkcji jest w języku polskim zjawiskiem częstym) lub *-ina* (np. *kocina*) wyrażający współczucie mó-

wiącego dla elementów rzeczywistości, o których mowa w komunikacie (może ono oczywiście być wyrażane pośrednio, np. *sukienczyna*, *paltocina*, w przypadku których chodzi oczywiście o współczucie dla osoby w nie ubraniej, a nie dla sukienki czy palta).

W języku polskim na zasadzie słowotwórczej można tworzyć także formy ekspresywne przymiotników i niektórych przysłówków, np.: *ladniutka*, *ladniusia*, *ladniunia*, *milutko*, *milusio*. Do wyrażania ekspresji służy także bardzo wiele elementów słownika, takich jak np.: *bachor* (dziecko, którego nie lubię, które mnie denerwuje), *chabeta*, *szkapa*, *kiecka*. W tej funkcji używane są też wyrazy o innym podstawowym znaczeniu, np.: *kundel* (pies), *dno* (coś złego), *bomba* (coś dobrego), *wapniak* (ktoś starszy), *chata* (mieszkanie), *ochlap* (mięso), *pala*, *lufa*, *bomba* (ocena niedostateczna w szkole). W przykładach tych wykładniki funkcji ekspresywnej często „nakładają się” na wyrażenia pełniące funkcję semantyczną, istnieją jednak w języku naturalnym i takie wyrażenia, które służą wyłącznie ekspresji. Tu należą wykrzykniki, takie jak *ojej!*, *och!*, oraz przekleństwa, jak na przykład *cholera!*, *psia kość!*, *kurczę!*. Te leksykalne środki ekspresji bardzo prędko wymieniają się w każdym języku naturalnym, gdyż skutkiem częstego używania szybko tracą swój ładunek ekspresywny (powszednieją) i muszą być zastąpione wyrazami nowymi. Funkcję ekspresywną, zwaną też **funkcją emotywną**, pełnią często języki bez semantyki, takie jak muzyka, której komunikaty (utwory muzyczne) wyrażają w jakiś sposób nadawcę (kompozytora lub wykonawcę), a odbiorca takich komunikatów może z nich uzyskać informację na przykład o stanie uczuciowym twórcy lub wykonawcy.

Z ekspresywną funkcją języka naturalnego wiąże się poprzez wspólną część wykładników inna jego funkcja, **funkcja impresywna**, czyli **funkcja apelatywna**, służąca do wpływania przez nadawcę komunikatu na odbiorcę, u którego nadawca chce wywołać pewne reakcje (na przykład postawy). Funkcji tej również służą różne elementy języka, mające zarówno charakter niekonwencjonalny, jak i konwencjonalny, począwszy od intonacji, poprzez środki leksykalne, aż do specjalnych konstrukcji składniowych. Do tych ostatnich należą dwie konstrukcje: zdania pytające, które służą do sterowania zachowaniem językowym odbiorcy (żądanie odpowiedzi) oraz zdania rozkazujące (tryb rozkazujący) służące wywołaniu odpowiedniego zachowania odbiorcy (żądanie działania).

Bardzo wiele środków leksykalnych służy do wywołania odpowiednich postaw odbiorcy komunikatu względem elementów komunikowanej rzeczywistości, których dotyczy komunikat, głównie poprzez sugerowanie pewnych ocen, które to oceny wbudowane są w bardzo wiele wyrazów. I tak, na przykład, tylko tym różnią się takie pary wyrazów, jak: *wywiadowca* i *szpieg*, *kolektyw* i *klika*, *żołnierz* i *żoldak*, *wojownik*, *bojownik* i *terrorysta*, *lewicowy* i *lewacki*, *Związek Radziecki* i *Związek Sowiecki*.

Skala ocen może być oczywiście szersza – dobro i zło mogą być stopniowane – tym na przykład różnią się wyrazy *napastnik* i *agresor*.

Wpływowi na odbiorcę służą również liczne środki słowotwórcze, najczęściej te same, które pełnią funkcję ekspresywną, bo sugerowana ocena musi być

najpierw wyrażona. Liczne środki impresywne obecne są w języku reklamy i propagandy, np.:

Kreacja za grosik;

Tylko u nas świeżutki chlebuś;

W dniu wczorajszym dokonano na granicy wymiany złapanych na gorącym uczynku szpiegów na pracowników naszego wywiadu.

Funkcja impresywna polega więc na tym, że poprzez odpowiednie komunikaty nadawca może sterować zachowaniem odbiorcy komunikatu. Języki naturalne nie są jednak wyspecjalizowane w pełnieniu tej funkcji w tym sensie, że odebranie komunikatu nie powoduje automatycznie pożądanej przez jego nadawcę zmiany stanu odbiorcy. Na przykład odbiorca prośby lub rozkazu może, ale nie musi się do nich zastosować. Są jednak języki, gdzie funkcja ta jest ich główną funkcją, języki te służą bowiem do sterowania odbiorcą, a przykładem mogą być języki programowania służące do sterowania komputerem czy języki informacyjno-wyszukiwawcze w swej funkcji wyszukiwawczej, która jest jedną z dwóch podstawowych funkcji tych języków. Funkcję impresywną pełnią także inne kody, na przykład muzyka, której komunikaty mają wywoływać u odbiorców określone stany uczuciowe.

Od języków informacyjno-wyszukiwawczych wymaga się, aby pełniły funkcję semantyczną i nie sugerowały żadnych postaw, dlatego każdemu wyrazowi języka naturalnego wybieranemu na słowo kluczowe lub deskryptor trzeba się dokładnie przyjrzeć, czy nie zawiera ukrytych elementów oceny.

Choć język naturalny wyspecjalizowany jest w funkcji semantycznej i jedynym skutkiem w rzeczywistości pozajęzykowej użycia języka, a więc aktu komunikacyjnego, jest zmiana stanu odbiorcy komunikatu pod względem informacyjnym, to bywa tak, że niektóre komunikaty językowe powodują w rzeczywistości innego rodzaju zmiany, a mianowicie stwarzają pewne fakty. Komunikaty takie, nadane w pewnych ściśle określonych sytuacjach komunikacyjnych, pełnią więc **funkcję sprawczą**. Wiara w funkcję sprawczą języka leżała u podstaw tzw. **magii językowej**, przejawiającej się w zaklęciach, klątwach, a także w **tabu** nakładanym na użycie nazw bytów wrogich lub niepożądanych (por. przestrożę *Nie wywołuj wilka z lasu.*), co miało wpływ na wykształcenie się w języku eufemizmów lub wyrazów etymologicznie peryfrastycznych, takich jak polski wyraz *niedźwiedź*. Pozostałością takiej wiary jest zwyczaj składania sobie życzeń z różnych okazji.

Wykładnikiem **siły sprawczej** komunikatów bywają specjalne wyrażenia obecne w tych komunikatach, np.:

Chrzczę cię imieniem...

Ogłaszam unieważnienie małżeństwa między...

Mianuję cię moim pełnomocnikiem.

Niniejszym ogłaszam wybory za nieważne.

Sąd uznaje oskarżonego winnym zarzucanych mu czynów.

Skazuję oskarżonego na...

Aby takie działania dokonywane przez wypowiedzenie były skuteczne, muszą być spełnione pewne ściśle określone warunki. Sformułował je angielski językoznawca i filozof języka John Langshaw Austin:

1. Musi istnieć ściśle określona konwencjonalna procedura, jak np. ceremonia udzielania ślubu, udzielania rozgrzeszenia przy spowiedzi, ogłaszania wyroku przez sąd, chrztu w kościele, nadawania imienia statkom, udzielania błogosławieństwa. Procedura taka wyraźnie określa, jakie warunki musi spełniać osoba, żeby być nadawcą takiego komunikatu, jakie warunki musi spełniać osoba, żeby być adresatem takiego komunikatu, w jakiej sytuacji taki komunikat może zostać nadany oraz określa niezmiennie elementy wszystkich komunikatów pełniących tę samą funkcję oraz kategorię i formę wyrażen wypełniających „puste miejsca” takiego komunikatu.

2. Procedura musi być przeprowadzona dokładnie tak, jak to zostało określone, a treść i forma komunikatu nie mogą być zmienione.

3. Komunikatowi musi towarzyszyć odpowiednia intencja, nadawca musi posiadać odpowiednie kompetencje (pozakomunikacyjne), a adresat musi świadomie się jej poddać i być poinformowany o skutkach pozainformacyjnych takiego komunikatu.

Tego typu wypowiedzi nazywane są **wypowiedziami performatywnymi**, a wyrażenia będące wykładnikami ich funkcji sprawczej **performatywami**, w odróżnieniu od wypowiedzi, w których orzekamy coś o rzeczywistości, konstatujemy coś o czymś, zwanych **konstatywami**. Wśród performatywów wyróżnić należy specjalną klasę wyrażen odnoszących się do komunikatów, a więc o **funkcji metatekstowej (metainformacyjnej)**, takich jak: *przysięgam, przyrzekam, obiecuję, wybaczam ci, ostrzegam*, ustanawiających (explicite) specjalny status wypowiedzenia. Zauważmy, że w wypadku komunikatów zawierających takie performatywne wykładniki ich funkcja jest inna niż funkcja zwykłych komunikatów, co widać przy porównaniu następujących komunikatów:

Niniejszym mianuję cię moim pełnomocnikiem.

(niniejszym, czyli właśnie tym wypowiedzeniem)

Kowalski mianował Nowaka swoim pełnomocnikiem.

Ten ostatni komunikat pełni tylko zwykłą funkcję informacyjną.

Takimi funkcjami aktów mowy, czyli komunikatów (wypowiedzeń) nadawanych w konkretnej sytuacji komunikacyjnej, przez konkretnego nadawcę i skierowanych do konkretnego odbiorcy (adresata) zajmował się Austin. Wyróżnił on trzy funkcje aktów mowy (trzy kategorie aktów mowy):

1. **lokucja** (od łac. *locutio* – mówienie) – czyli powiedzenie czegoś („wygłoszenie pewnych dźwięków, pewnych wyrazów w pewnej konstrukcji z pewnym „znaczeniem” w ulubionym filozoficznym znaczeniu tego słowa, czyli z pewnym sensem i pewnym odniesieniem” – Austin 1962, cyt. za J. Lyons: *Semantyka*);

2. **illokucja** (od łac. *in + locutio = illocutio*) – akt dokonywany przez powiedzenie czegoś, a więc na przykład poprzez nadanie komunikatu, poinformowanie,

obiecanie, przyrzeczenie, pytanie, ostrzeżenie, a także ochrzczenie kogoś, skazanie, niewinnienie, przebaczenie, przeproszenie;

3. **perlokucja** (od łac. *per* + *locutio*) – wywołanie przez komunikat zmiany stanu odbiorcy, na przykład pocieszenie kogoś, rozgniewanie kogoś, rozbawienie kogoś, ale także spowodowanie zmiany jego przekonań lub wywołanie zachowania niekoniecznie zamierzonego przez nadawcę komunikatu.

Wśród aktów illokucyjnych wyróżnia się akty illokucyjne *implicite* i akty illokucyjne *explicite*, te drugie mają wbudowany wykładnik illokucyjności (tak zwanej siły illokucyjnej), np.: *przyrzekam*, *obietuję*, *ostrzegam*, te pierwsze takiego wykładnika nie mają, dlatego wypowiedzenie *Wrócisz do domu* może być w różnych sytuacjach odczytane jako obietnica, przyrzeczenie, ostrzeżenie, groźba, a nawet prośba. Zależy to od sytuacji, ale także od innych niż leksykalne wykładników siły illokucyjnej, takich jak intonacja oraz na przykład postawa, gesty i mimika nadawcy komunikatu. Wykładniki siły illokucyjnej zawarte *explicite*, np.:

<i>Ostrzegam cię: wrócisz do domu.</i>	<i>Ostrzegam cię, że wrócisz do domu.</i>
<i>Obietuję ci: wrócisz do domu.</i>	<i>Obietuję ci, że wrócisz do domu.</i>
<i>Przyrzekam ci: wrócisz do domu.</i>	<i>Przyrzekam ci, że wrócisz do domu.</i>

mogą być interpretowane tylko w jeden sposób.

Wykładnik językowy rodzaju aktu mowy może przybierać różną formę, zawsze daje się jednak sprowadzić do postaci kanonicznej, zawierającej **czasownik performatywny** opisujący dany typ aktu w odpowiedniej formie. Na przykład wypowiedzenia takie jak:

Czy możesz otworzyć okno?
Otwórz okno, jeżeli mogę cię prosić.
Okropnie tu duszno.

mogą w określonych sytuacjach pełnić taką funkcję jak:

Proszę cię, żebyś otworzył okno.

Jeśli akt mowy nie jest wyrażony *explicite*, istnieje możliwość, że komunikat zostanie różnie zinterpretowany przez różnych odbiorców lub że odbiorca nie będzie pewien swej interpretacji. Na przykład pracownik, który usłyszy od swego szefa zdanie:

Jeżeli twój projekt nie zostanie przyjęty, pojedziesz na urlop.

może nie być pewien, czy spotyka się z obietnicą, czy z groźbą, zaś wypowiedzenie takie jak

Czy mogę Państwu przypomnieć, że posiedzenie zaczyna się o piątej?

to zarazem pytanie, przypomnienie i informacja.

Przyjmuje się, że rozróżnienie podstawowe można tu przeprowadzić między **aktem mowy bezpośrednim**, wynikającym z dosłownej interpretacji struktury

i wyrazów użytych w zdaniu, a **aktem mowy pośrednim**, czyli każdą interpretacją niebezpośrednią (w przykładzie powyżej pytanie to akt bezpośredni, przypomnienie i informacja to akty pośrednie). Pośredniość aktu wiązana jest często z wymogami grzeczności.

Możliwe jest także, że interpretacja przyjęta przez odbiorcę komunikatu nie będzie zgodna z intencją nadawcy komunikatu, czyli **celem illokucyjnym**. Zdarza się także, iż właściwa intencja nadawcy nie jest tożsama z celem, któremu zwykle służy dany typ aktu mowy.

Teoria aktów mowy rozwijana była przez innych językoznawców, a przede wszystkim przez Johna Rogersa Searle'a. Searle (na podstawie rozróżnień Austina) zaproponował następujący podział aktów mowy:

1. **opowiedzenia** (ang. *representatives*) – komunikaty, w których nadawca opowiada się za prawdziwością wyrażonego w komunikacie sądu (np. stwierdzenie, wnioskowanie);

2. **polecenia** (ang. *directives*) – komunikaty mające na celu skłonienie adresata do określonego działania (np. prośba, rozkaz);

3. **zobowiązania** (ang. *commissives*) – komunikaty, w których nadawca zobowiązuje się do wykonania pewnych działań w przyszłości (np. obietnica);

4. **wyrażenia** (ang. *expressives*) – komunikaty, które wyrażają stan psychiczny nadawcy komunikatu związany z jego relacją w stosunku do adresata (np. podziękowania, przeprosiny, gratulacje);

5. **deklaracje** (ang. *declarations*) – komunikaty, które powodują natychmiastową zmianę w rzeczywistości pozajęzykowej i zwykle odnoszą się do instytucjonalizowanych form życia społecznego (np. ekskomunika, ogłoszenie stanu wojny, chrzest, wypowiedzenie z pracy).

Typologia ta jest powszechnie stosowana w pragmatyce lingwistycznej, a zarazem krytykowana jako niewyczerpująca, niepoparta szeroką analizą danych z różnych języków, nieodwołująca się do jasno sformułowanych kryteriów. Najwięcej wątpliwości budzi grupa deklaracji, jako że ujmowane pod tym terminem zjawiska nie należą do opisu językoznawczego sensu stricto.

Ilość, a także typy aktów mowy są różne dla różnych języków, także poszczególne typy aktów są różne w różnych kulturach.

Każdy typ aktu określony jest przez zbiór warunków, które muszą być spełnione, aby akt mógł zostać określony jako udany (warunki illokucyjne). Proponuje się również, aby poszczególne akty mowy opisywane były poprzez opis czasowników performatywnych, czyli, tak jak inne jednostki leksykalne, poprzez eksplikacje semantyczne.

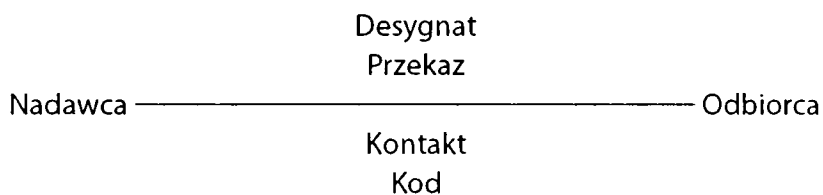
W każdym języku naturalnym zawarte są również elementy służące do nawiązania lub podtrzymywania kontaktu między nadawcą i odbiorcą komunikatu albo sprawdzenia, czy kontakt taki nie został przerwany. Do tego między innymi służą różnego typu sposoby zwracania się do odbiorcy, np.: *wiesz, słuchaj, wie pan, aha*, będące wykładnikami tzw. **funkcji fatycznej** języka.

Komunikat może być odpowiednio ukształtowany zgodnie z pewnymi kryteriami estetycznymi. Jest to **funkcja estetyczna** języka, zwana też w wypadku języka

naturalnego **funkcją poetycką** lub **funkcją artystyczną**. Funkcja ta realizowana jest w utworach literackich, a pełnią ją różnego rodzaju środki językowe, począwszy od fonetycznych (np. onomatopeja), poprzez gramatyczne, leksykalne, frazeologiczne i składniowe. Funkcją tą zajmuje się specjalny dział lingwistyki i teorii literatury – poetyka.

W każdym języku naturalnym formułowane są również wypowiedzi na temat tego języka jako systemu (kodu) – każdy język naturalny pełni więc również **funkcję metajęzykową** – oraz wypowiedzi na temat tekstów języka (informacji) – każdy język naturalny pełni więc również **funkcję metainformacyjną**. Funkcje te omówimy dokładniej w następnym rozdziale.

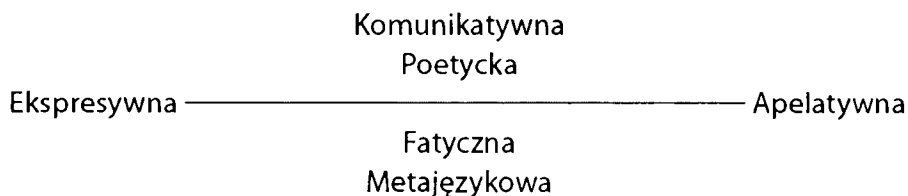
Funkcje języka naturalnego stały się przedmiotem badań wielu językoznawców. Karl Bühler wyróżnił trzy funkcje: przedstawieniową (semantyczną), ekspresywną i impresywną. Schemat ten rozbudował Roman Jakobson, odnosząc funkcje języka naturalnego do poszczególnych elementów sytuacji informacyjnej:



wyróżniając:

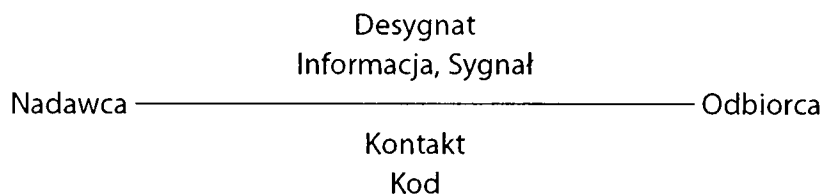
- nastawienie na desygnat (komunikowanie rzeczywistości) – funkcja komunikatywna (semantyczna, przedstawiająca, symboliczna, informacyjna);
- nastawienie na wyrażenie nadawcy – funkcja ekspresywna;
- nastawienie na odbiorcę – funkcja apelatywna (impresywna);
- nastawienie na kod – funkcja metajęzykowa;
- nastawienie na kontakt – funkcja fatyczna,
- nastawienie na przekaz – funkcja poetycka (estetyczna).

Schematowi układu informacyjnego odpowiada więc według Jakobsona następujący schemat funkcji językowych:



Termin funkcja komunikatywna jest terminem niefortunnym, może być bowiem interpretowany nie jako funkcja komunikowania rzeczywistości – jak tu została zdefiniowana, lecz jako funkcja komunikowania się nadawcy z odbiorcą, a więc nadawania i odbierania przekazu (komunikatu), dlatego też lepszymi terminami są terminy funkcja informacyjna, semantyczna, przedstawiająca czy nawet symboliczna, niedające możliwości takiej podwójnej interpretacji.

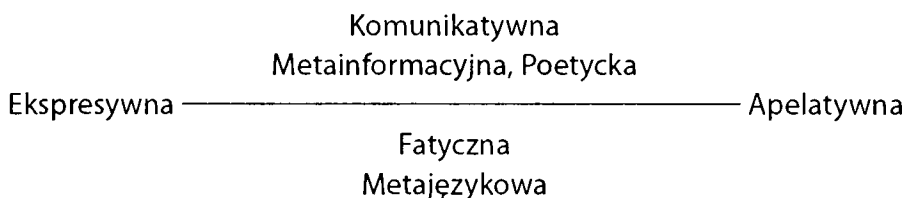
Wydaje się, że schematy Jacobsona należałoby nieco zmodyfikować, wyraźnie wyróżniając w przekazie jego element materialny (sygnał), związany z wyborem określonego kanału informacyjnego. Uzyskalibyśmy w ten sposób następujący schemat sytuacji informacyjnej:



i odpowiednio:

- nastawienie na sygnał (formę przekazu) – funkcja zwana poetycką lub estetyczną (nie jest to termin fortunny, gdyż nie zawsze jest to funkcja ściśle estetyczna);
- nastawienie na informację (zawartość treściową przekazu) – funkcja metainformacyjna.

W ten sposób uzyskalibyśmy następujący zmodyfikowany schemat funkcji języka informacyjnego:



Każdy komunikat języka z semantyką traktować można jako twór złożony z tekstu i z wyrażanej przez ten tekst informacji. Do tych dwóch składowych mogą też odnosić się komunikaty języka naturalnego. Odpowiednio wyróżnia się więc funkcję metatekstową i funkcję metainformacyjną. Funkcji metainformacyjnej, jako szczególnie ważnej w informacji naukowej, poświęcimy następny rozdział podręcznika, funkcję metatekstową omówimy w rozdziale zajmującym się tekstem spójnym. Tu chcielibyśmy omówić jeszcze funkcję metajęzykową języka naturalnego.

Każdy komunikat jest odbiciem (modelem) jakiejś rzeczywistości, muszą więc w nim znaleźć wyraz poszczególne relewantne elementy tej rzeczywistości oraz istniejące między tymi elementami relewantne relacje. Tym, co pozwala na przyporządkowanie elementom rzeczywistości zastępujących je elementów tekstu, jest język. Czasami jednak bywa tak, że w danym tekście niektórym jego elementom przyporządkowane są nie elementy jakiejś rzeczywistości pozajęzykowej, ale tożsame z elementami tekstu elementy samego języka. Wtedy mamy już do czynienia nie z wypowiedzią o rzeczywistości (zewnętrznej w stosunku do języka – tzw. rzeczywistości pozajęzykowej), ale z wypowiedzią o języku, a więc z wypowiedzią metajęzykową.

Dla lepszego uzmysłowienia sobie problemu porównajmy następująca zdania języka polskiego:

Z roztargnieniem przeglądał leżące na stole czasopismo.

Czasopismo, periodyk, wydawnictwo periodyczne (...), którego zakończenia nie przewiduje się, ukazujące się w postaci osobnych zeszytów (nrów), objętych wspólnym tyt., w regularnych odstępach czasu (...).

Czasopismo jest wyrazem złożonym, kalką z niemieckiego Zeitschrift.

Pierwsze zdanie mówi coś o rzeczywistości, dwa następne są zdaniami o języku (dokładniej: o pewnych jego elementach). Zdanie drugie mówi o tym, co wyraz *czasopismo* znaczy, zdanie trzecie o tym, jak wyraz ten jest zbudowany i skąd się wziął w języku polskim.

Zdanie pierwsze należy do języka (rozumianego jako zbiór wszystkich zdań), zwanego też czasami **językiem przedmiotowym**, ponieważ jego wyrażenia odnoszą się do przedmiotów spoza języka, natomiast zdania drugie i trzecie należą do **metajęzyka**, zaś wyraz *czasopismo* występuje w nich, jak mówią logicy, w tzw. **supozycji materialnej** (ważna jest tylko jego forma, materialna strona znaku), podczas gdy w pierwszym zdaniu użyty był w **supozycji zwykłej** (ważna jest tylko jego treść).

Na tej opozycji, supozycji zwykłej i supozycji materialnej, oparta jest zagadka we *Władcy pierścieni* Tolkiena (*Wyprawa*), której rozwiązanie miało otworzyć drzwi do Morii. Napis na nich wykuty (w zapomnianym już języku Śródziemia) „Brzmi po prostu tak: «Drzwi Durina, Władcy Morii. Powiedz przyjacielu i wejdź».(...)»

– Co to ma znaczyć: «Powiedz przyjacielu i wejdź»? – spytał Merry.

– Sens dość jasny – rzekł Gimli. – Jeżeli jesteś przyjacielem, powiedz hasło, a drzwi się otworzą i będziesz mógł wejść.

(...) Hasło było wzięte z mowy elfów. (...) próbował różnych zaklęć, jednego po drugim, wymawiając je to szybko i głośno, to znów powoli i cicho. Wreszcie zaczął rzucać pojedyncze słowa z języka elfów. Nic się nie działo. Urwisko piętrzyło się przed nimi w ciemnościach, gwiazdy bez liku rozbłyły na niebie, wiatr dmuchnął chłodem, ale drzwi wciąż trwały niewzruszone.

(...) Nagle Czarodziej zerwał się na równe nogi, a było to tak niespodziane, że wszyscy drgnęli. Gandalf się śmiał!

Już wiem! – krzyknął. – Oczywiście, oczywiście! Śmiesznie proste, jak większość zagadek, jeżeli się zna rozwiązanie.

(...) Hasło było wypisane na łuku, widzieliśmy je od pierwszej chwili. Napis trzeba tłumaczyć tak: «Powiedz <przyjacielu> i wejdź». Kiedy wymówiłem w języku elfów słowo «przyjacielu», drzwi się otwały. Po prostu!»

Język naturalny jest oczywiście i językiem, i metajęzykiem, można w nim bowiem formułować zarówno zdania o rzeczywistości pozajęzykowej, jak i zdania o nim samym, natomiast do opisu innych języków buduje się specjalne, różne od nich same, metajęzyki. Metajęzyk względem określonego języka L jest to więc taki zinterpretowany semantycznie język L₁, w którym występują nazwy

wszystkich wyrażen języka L (można je tworzyć np. przez ujęcie w cudzysłów wyrażen języka L), służący do opisu języka L.

Ścisłego odróżnienia metajęzyka od języka dokonano stosunkowo niedawno. Sam termin pochodzi od Alfreda Tarskiego (*Pojęcie prawdy w językach nauk dedukcyjnych*), i utworzony został na wzór wprowadzonego przez Davida Hilberta terminu metamatematyka. Rozróżnienie to wprowadził również Rudolf Carnap, nie używał jednak tego terminu (w pracy *Filozofia jako analiza języka nauki*), zaś sugestią zbudowania hierarchii języków wysunął po raz pierwszy Bertrand Russell. Do rozróżnienia języka i metajęzyka doprowadziły badania nad językami pozwalającymi konstruować pewne antynomie, nad którymi rozważania prowadzono już w starożytności.

Metajęzyki służą do opisywania innych języków, tak więc wszystkie reguły składniowe i semantyczne danego języka sformułowane są w metajęzyku względem tego języka. Zdania w metajęzyku zawierają także podręczniki gramatyki, słowniki, dzieła językoznawcze oraz pewne partie różnego typu encyklopedii. Metajęzyki mogą tworzyć hierarchię, ponieważ każdy z nich może być opisywany przez metajęzyk wyższego stopnia. Do problemów metajęzyka wrócimy przy omawianiu systemu semantycznego języka naturalnego.

Omówione tu funkcje języka naturalnego nie są jedynymi funkcjami wymienianymi w literaturze przedmiotu, są jednak jego funkcjami najważniejszymi. I tak, na przykład, niektórzy semiotycy wyróżniają jeszcze **funkcję ludyczną** (zabawową) języka, realizowaną na przykład w dziecięcych wyliczankach, nie wyróżniamy jej jednak tutaj, gdyż nie jest to, naszym zdaniem, funkcja dla języka specyficzna – bawić się można tekstami języka tak samo, jak innymi przedmiotami. To, co w tej zabawie specyficznie językowe, daje się sprowadzić do funkcji nastawienia na formę przekazu, który, jako całkowicie skonwencjonalizowany, pozbawiony jest w tym wypadku wartości informacyjnej.

Definiowanie poszczególnych funkcji języka naturalnego w odniesieniu do elementów układu informacyjnego nasuwać może uzasadnione wątpliwości, czy mówić należy o funkcjach języka (systemu znakowego), czy też raczej o funkcjach wynikających z użycia języka czy znaku, a więc o funkcjach komunikatów (tekstów języka). Jedynie słusznym wydaje się tu rozstrzygnięcie, że są to zarówno funkcje komunikatów, a więc konkretnych tekstów językowych nadanych w konkretnych warunkach, jak i funkcje leżących u podstaw tych realizacji systemów znakowych. W wypadku języka naturalnego, jak widzieliśmy, w systemie języka, zarówno leksykalnym, morfologicznym i syntaktycznym, a nawet fonologicznym, znajdują się elementy wyspecjalizowane w pełnieniu różnych funkcji.

Każdy język (system znakowy) pełniący funkcję semantyczną dokonuje pewnej klasyfikacji reprezentowanej przez jego wyrażenia rzeczywistości, wyznaczając jej elementy odpowiednio w danym języku nazywane. Każdy taki język, a więc i język naturalny, pełni wobec tego **funkcję klasyfikacyjną**. W odróżnieniu od funkcji poprzednio omawianych, jest to funkcja systemu językowego, a nie jego tekstów, a omówimy ją dokładniej w rozdziałach omawiających odwzorowanie rzeczywistości w języku.

Języki naturalne są kodami uniwersalnymi w tym sensie, że pełnią wszystkie z wymienionych wyżej funkcji i służą do opisywania całego uniwersum.

Języki informacyjno-wyszukiwawcze, jak sama nazwa wskazuje, pełnią dwie podstawowe funkcje: informacyjną i wyszukiwawczą.

Pierwsza funkcja, **funkcja informacyjna**, polega na reprezentowaniu przez teksty tych języków, czyli opisy dokumentacyjne (formalne i treściowe) elementów zbioru dokumentacyjnego opisywanego przez dany język (bezpośrednio), a pośrednio opisywanej w tych tekstach rzeczywistości. W ten sposób dla odbiorców pełnią funkcję metainformacyjną, informują ich bowiem o informacji zawartej w tekstach dokumentów. Wynika to jasno z poniższych definicji języka informacyjno-wyszukiwawczego, choć autorzy tych definicji przeważnie nie zdawali sobie sprawy z tego, że jest to funkcja metainformacyjna.

A oto przykładowe definicje (cyt. za E. Ściborem: *Typologia strukturalna języków informacyjnych*):

„Wyspecjalizowany język sztuczny, przeznaczony do wyrażania podstawowej znaczeniowej zawartości dokumentów i/lub tylko zapytań informacyjnych w celu odszukania w pewnym zbiorze dokumentów tych, które odpowiadają na postawione zapytanie informacyjne” (A.I. Michajłow, A.I. Černyj, R.S. Giljarevskij).

„Wyspecjalizowany język sztuczny, przeznaczony do opisywania (wyrażania) głównych tematów lub przedmiotów i charakterystyk formalnych dokumentów w celu późniejszego odszukania potrzebnych dokumentów wśród mnogości innych – i/lub do wyrażania treści zapytań informacyjnych i wyszukania potrzebnych dokumentów” (Černyj).

„Język sztuczny, który pozwala tworzyć sformalizowaną reprezentację dokumentów i zapytań interesujących pewną grupę użytkowników w celu wyszukania dokumentów, które odpowiadają na zapytania” (Y. Courier).

„System znaków i reguł operowania nimi, rozumianych tak samo przez pewną grupę ludzi, specjalnie zbudowany w celu przedstawienia w skondensowanej postaci istotnych (z punktu widzenia późniejszej identyfikacji) elementów zawartości lub formy informacji przechowywanej w pamięci systemu (SWI), zapewnienia przepływu informacji wewnątrz tego systemu oraz wyszukiwania w pamięci systemu informacji relewantnej do zapytania użytkownika” (O. Ungurian).

„Język informacyjny (JIW) jest instrumentem każdego systemu informacyjnego służącym do przedstawiania opisów dokumentów i formułowania pytań informacyjnych w celu wyszukania dokumentów” (R. Steiger).

„Język indeksowania zdefiniowany jest jako wszystkie słowa przeznaczone albo do opisywania poszczególnych dokumentów, albo do konstruowania pytań informacyjnych w celu wyszukania w kartotece dokumentacyjnej oraz zbiorów reguł określających jak te słowa powinny być używane i w jakich pozostają względem siebie relacjach” (D.B. Cleveland, A.D. Cleveland).

„JIW to język sztuczny, którego wyspecjalizowaną funkcją jest opis dokumentów w celu późniejszego odszukania ich na żądanie użytkownika. Język informacyjny wyraża treść i/lub charakterystyki (...) formalne (...) dokumentów” (E. Ścibor).

„Specjalistyczny język sztuczny, przeznaczony do odtwarzania podstawowej treści dokumentu i (lub) kwerendy w celu wyszukania ze zbioru informacyjnego tylko tych dokumentów, które odpowiadają na kwerendę, tzn. których charakterystyka wyszukiwawcza pokrywa się z instrukcją wyszukiwawczą” (*Słownik terminologiczny informacji naukowej*).

„Język informacyjno-wyszukiwawczy, język informacyjny, JIW – język sztuczny, którego wyspecjalizowaną funkcją jest funkcja metainformacyjna polegająca na odwzorowywaniu cech informacji i/lub cech nośników informacji, na których są one utrwalone, przez tworzenie ich reprezentacji zwanej charakterystyką wyszukiwawczą dokumentu oraz funkcja wyszukiwawcza polegająca na wyrażaniu zapytań informacyjnych i umożliwianiu odnalezienia w zbiorze wyszukiwanym tych informacji, które spełniają warunki wyrażone w zapytaniu. (...)” (*Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*).

Wszystkie przytoczone tu definicje zwracają też uwagę na drugą funkcję języka informacyjno-wyszukiwawczego, **funkcję wyszukiwawczą** realizowaną w pytaniach informacyjnych. Jest to oczywiście funkcja impresywna, pytania informacyjne skierowane do systemu informacyjnego mają bowiem za zadanie sterowanie zachowaniem tego systemu zmierzającym do udzielenia żądanej odpowiedzi. Impresywną funkcję pełnią oczywiście również instrukcje wyszukiwawcze, których język pokrywa się przynajmniej częściowo z funkcjonującym w systemie językiem informacyjno-wyszukiwawczym – w systemach zautomatyzowanych bywa on rozszerzany o odpowiednie komendy wyszukiwawcze, wśród których mogą znajdować się tzw. operatory logiczne (boolowskie) AND (i), OR (lub), NOT (nie, oprócz), które są wykładnikami rozkazów wydawanych systemowi dotyczących wykonywania pewnych operacji na zbiorach dokumentów.

Języki informacyjno-wyszukiwawcze pełnią jeszcze jedną funkcję, analogiczną do funkcji klasyfikacyjnej języka naturalnego, tzw. **funkcję organizacyjną**, która jest oczywiście wtórna w stosunku do jego funkcji prymarnej, funkcji metainformacyjnej.

Bywa i tak, że funkcję tę spełnia język naturalny, ale zawsze należy pamiętać o tym, że używany jest wtedy nie tak, jak używa się normalnego języka przedmiotowego, albowiem jego elementom leksykalnym przyporządkowuje się wtedy (w najprostszym przypadku) klasy dokumentów, w których charakterystykach wyszukiwawczych słowa te występują. Mamy tu więc do czynienia z zupełnie innymi regułami przyporządkowania między elementami rzeczywistości a elementami opisującego ją języka – jest to inna rzeczywistość, rzeczywistość informacji – czyli po prostu z inną semantyką, a więc nawet w wypadku identyczności zasobów leksykalnych oraz reguł składniowych mamy do czynienia właściwie z dwoma różnymi językami.

Języki informacyjno-wyszukiwawcze mogą też być używane do opisu zakresu i struktury systemów informacyjno-wyszukiwawczych, a więc do opisu ich pola informacyjnego, również w takich sytuacjach pełnią więc funkcję metainformacyjną.

Literatura

- Austin J.: *Performatywy i konstatacje*, w: Hempoliński M. (red.): *Brytyjska filozofia analityczna*. Warszawa 1974.
- Grzegorzczkova R.: *Problem funkcji języka i tekstu w świetle teorii aktów mowy*, w: Bartmiński J., Grzegorzczkova R. (red.): *Język a kultura*, t. IV. Wrocław 1991.
- Lalewicz J.: *Krytyka teorii funkcji mowy Bühlera-Jacobsona*. „Teksty” 1976, nr 6.
- Searle J.: *Czynności mowy. Rozważania z filozofii języka*. Warszawa 1987.
- Wierzbicka A.: *Akty mowy*. W: Mayenowa M. R. (red.): *Semiotyka i struktura tekstu*. Wrocław 1973.
- Zdunkiewicz D.: *Akty mowy*, w: Bartmiński J. (red.): *Encyklopedia kultury polskiej XX wieku*, t. 2. *Współczesny język polski*. Wrocław 1993.

VI

Metainformacja, metatekst, hipertekst

Henryk Greniewski, twórca polskiej szkoły cybernetycznej, napisał kiedyś: „Zdolność do tworzenia i transformowania metainformacji wydaje się pierwszą istotną różnicą między człowiekiem a zwierzęciem” (*Cybernetyka niematematyczna*).

Czym jest metainformacja? Nim odpowiemy na to pytanie, spróbujmy zastanowić się, czym różnią się od siebie następujące zdania:

- 1 (a) *22 bm Prezydium Rządu na swoim posiedzeniu rozpatrzyło sprawy związane z sytuacją paliwowo-energetyczną kraju.*
- 1 (b) *Jak informuje rzecznik prasowy rządu, 22 bm Prezydium Rządu na swoim posiedzeniu rozpatrzyło sprawy związane z sytuacją paliwowo-energetyczną kraju.*
- 2 (a) *Jaworski zgromadził liczne dowody winy byłego prezydenta.*
- 2 (b) *Jaworski, jak wynika z informacji prasowych, zgromadził liczne dowody winy byłego prezydenta.*
- 3 (a) *Politechnika Warszawska kształci lepszych niż inne politechniki elektroników, mechaników, architektów.*
- 3 (b) *Wiadomo, że Politechnika Warszawska kształci lepszych niż inne politechniki elektroników, mechaników, architektów.*
- 4 (a) *W Uniwersytecie Warszawskim, który w porównaniu z innymi uniwersytetami zatrudnia najwięcej pracowników naukowych, przeciętna zgłoszeń na jedno miejsce jest dwukrotnie wyższa niż przeciętna krajowa.*
- 4 (b) *W ostatnim numerze Gazety studenckiej czytamy, że na Uniwersytecie Warszawskim, który w porównaniu z innymi uniwersytetami zatrudnia najwięcej pracowników naukowych, przeciętna zgłoszeń na jedno miejsce jest dwukrotnie wyższa niż przeciętna krajowa.*

Zdania (a) mówią o jakichś sytuacjach lub zdarzeniach, które mają miejsce w jakiejś rzeczywistości, podają więc jakąś informację o tej rzeczywistości, ale o tym, skąd nadawca tych zdań zaczerpnął tę informację, czy sam był obserwatorem zdarzeń, o których komunikuje, czy też dowiedział się o nich z jakichś innych źródeł, w zdaniach tych nic się nie mówi. Inaczej wszystko wygląda w zdaniach (b).

Zdania te nie tylko, podobnie jak zdania (a), zawierają informację o jakichś zdarzeniach, ale ponadto mówią coś jeszcze o tej informacji. I tak:

- zdanie 1 (b) mówi, że nadawcą informacji o posiedzeniu Prezydium Rządu był rzecznik prasowy rządu;

- zdanie 2 (b) mówi o tym, że informacja o tym, że Jaworski zgromadził liczne dowody winy byłego prezydenta, jest wnioskiem z informacji prasowych – informuje o tym czasownik *wynikać*, który jest wykładnikiem takiej relacji między co najmniej dwiema informacjami, że jedna jest przesłanką rozumowania, a druga wyciągniętym z tej przesłanki wnioskiem;

- zdanie 3 (b) mówi, że pewna liczba osób ma informację o tym, że Politechnika Warszawska kształci lepszych niż inne politechniki elektroników, mechaników, architektów (czyli, że informacja ta jest utrwalona w pewnych pamięciach) – mówi o tym słowo *wiadomo*;

- zdanie 4 (b) mówi o tym, że informacja o Uniwersytecie Warszawskim została przez kogoś odebrana, przy czym była to informacja nadana w subkodzie pisanyim języka naturalnego – o tym wszystkim mówi czasownik *czytać* – a ponadto, że źródłem tej informacji jest ostatni numer *Gazety Studenckiej*.

Wszystkie zdania (b) podają więc pewną informację o innej informacji. Taka informacja o informacji jest informacją „o piętro wyższą”, nadbudowaną nad informacją, o której informuje, dla odróżnienia została więc nazwana (po raz pierwszy przez Henryka Greniewskiego, przez analogię do wcześniej zdefiniowanych terminów metamatematyka i metajęzyk) **metainformacją**.

Każda metainformacja komunikuje coś o innej informacji, może więc komunikować coś o:

- rodzaju procesu informacyjnego, czyli o:
 - pobieraniu informacji z jakiegoś źródła, np.:

Z komunikatu prasowego dowiedzieliśmy się, że rok akademicki rozpocznie się 4 października.

- nadawaniu informacji, np.:

Sławomir Gilonka z SGH mówił o nowych formach pracy Studenckiego Ośrodka Dyskusyjnego.

- utrwalaniu informacji, np.:

Nasz reporter zanotował, że w kwadrans po rozpoczęciu pracy na budowie brakowało jeszcze przeszło 10% załogi.

- przetwarzaniu (transformowaniu) informacji, np.:

Z zachowania się zwierząt leśnicy wnioskuje, że zima w tym roku będzie długa i ostra.

- elementach sytuacji informacyjnej:
 - źródle informacji, np.:

Z komunikatu PAP dowiadujemy się, że ambasador Chińskiej Republiki Ludowej złożył wizytę pożegnalną.

- nadawcy informacji, np.:

Carnap sformułował swą myśl w postaci zasady tolerancji, która głosi całkowitą dowolność wyboru języka.

- odbiorcy informacji, np.:

Nasz wysłannik dowiedział się, że Roman Polański wkrótce wybiera się do kraju.

- potencjalnym odbiorcy informacji (obiekcie przeznaczenia informacji, jej adresacie), np.:

Związek zawodowy powiadomił wszystkich pracowników, że zawarł ugodę z nowym właścicielem fabryki.

- języku informacji, np.:

Trener Gmoch powiedział, że na sukces greckich piłkarzy pracowali również polscy trenerzy.

(powiedzieć informuje, że informacja sformułowana została w języku naturalnym),

- kanale informacyjnym, np.:

Kowalski zobaczył, że z bramy wyszedł młody mężczyzna.

(zobaczyć implikuje kanał wizualny),

Trener Jacek Gmoch powiedział, że był to prawdziwy egzamin dla naszej jedenastki.

(powiedzieć implikuje kanał akustyczny).

- cechach sygnału stanowiącego materialny nośnik informacji, np.:

Halo, mostek! – krzyczy Jan do mikrofonu.

(krzyczeć informuje o dużym natężeniu sygnału akustycznego).

- instrumentach informacyjnych, np.:

*Jan dopiero przez lupę zobaczył zatarty odcisk palca włamywacza.
Kowalski telefonował, że wrześniowa konferencja rozpocznie się o tydzień wcześniej.*

- wartości informacji, np.:

Jan skłamał, że w czwartek nie wychodził z domu.

(kłamać informuje, że nadawca celowo nadał nieprawdziwą informację, chcąc wprowadzić odbiorcę w błąd).

a także o innych jeszcze aspektach informacji.

Tak więc człowiek przekazuje innemu człowiekowi nie tylko treść odebranych przez siebie informacji, ale bardzo często udziela o tej informacji dodatkowych wiadomości, powiadamia na przykład o sposobie, w jaki sam stał się posiadaczem tych informacji. Takich informacji o informacji, czyli metainformacji, udziela my odbiorcy zawsze wtedy, gdy sądzimy, że na przykład interesuje go źródło informacji lub gdy sami wolimy je podać, jak gdyby uwalniając się w ten sposób od odpowiedzialności za jej prawdziwość – jeżeli bowiem powiemy:

Kowalski obronił wczoraj doktorat.

a okaże się to nieprawdą, to możemy spotkać się z zarzutem celowego wprowadzenia odbiorcy w błąd, czego unikniemy, mówiąc:

Słyszałam, że Kowalski obronił wczoraj doktorat.

Podobno Kowalski obronił wczoraj doktorat.

gdyż słowa *słyszałam, podobno* zwalniają nas od odpowiedzialności za prawdziwość tej informacji.

Udzielając metainformacji często wskazujemy więc poprzednie ogniwo (lub ogniwa) procesu informacyjnego, w którym sami uczestniczyliśmy jako odbiorcy informacji, którą teraz z kolei dalej przekazujemy. Na przykład zdanie:

Nasz korespondent donosi, że w obozie uchodźców panuje spokój.

wskazuje poprzednie ogniwo procesu informacyjnego, identyfikując jego nadawcę oraz rodzaj procesu, natomiast zdanie:

Nasz korespondent donosi, że z wiadomości opublikowanych przez źródła dobrze poinformowane wynika, iż w obozie uchodźców panuje spokój.

wskazuje kilka poprzednich ogniw łańcucha informacyjnego, mamy tu bowiem wymienione:

- I ogniwo procesu informacyjnego, dla którego podano:
 - odbiorcę informacji: źródła dobrze poinformowane,
 - rodzaj procesu: odbieranie informacji;
- II ogniwo procesu informacyjnego, dla którego podano:
 - nadawcę informacji: źródła dobrze poinformowane,
 - proces informacyjny: nadawanie informacji,
 - odbiorcę informacji: każdy, kto chciał ją odebrać (potencjalnie wielu odbiorców – mówi o tym czasownik *opublikować*);
- III ogniwo procesu informacyjnego, dla którego podano:
 - proces informacji: transformowanie informacji,
 - informację wejściową (przesłankę): wiadomość opublikowana przez źródła dobrze poinformowane,
 - informację wyjściową (wniosek): w obozie uchodźców panuje spokój,
 - urządzenie transformujące informację: prawdopodobnie nasz korespondent (jest to sugerowane, nie zaś stwierdzone, możliwe bowiem, że korespondentowi ktoś przekazał informację, że z wiadomości opu-

blikowanych przez źródła dobrze poinformowane wynika, że w obozie uchodźców panuje spokój);

IV ogniwo procesu informacyjnego, dla którego podano:

- nadawcę informacji: nasz korespondent,
- proces informacyjny: nadawanie informacji,
- informację: z wiadomości opublikowanej przez źródła dobrze poinformowane wynika, że w obozie uchodźców panuje spokój;

V ogniwo procesu informacyjnego, dla którego podano:

- nadawcę informacji: ten, kto napisał zdanie:
Nasz korespondent donosi...
- proces informacyjny: nadawanie informacji,
- informację: Nasz korespondent donosi, że z informacji...

Jak z podanego przykładu widać, pojemność metainformacyjną jednego tylko zdania – i to wcale nie tak bardzo skomplikowanego – może być bardzo duża. Oprócz wyżej wymienionych metainformacji (czytelnikowi pozostawiamy identyfikację, które z nich mają status metainformacji i które ogniwa procesu informacyjnego są procesami metainformacyjnymi) podana jest w tym zdaniu także metainformacja dotycząca oceny wartości tej informacji poprzez określenie wiarygodności jej źródła – mówi o tym fraza: *źródła dobrze poinformowane*. Wymienione w tym zdaniu I ogniwo łańcucha informacyjnego musiało poprzedzać jeszcze co najmniej jedno ogniwo, w którym informację o stanie obozu uchodźców uzyskał informator owych dobrze poinformowanych źródeł. Mógł to uczynić albo przez bezpośrednią obserwację rzeczywistości – wtedy byłoby to tylko jedno ogniwo, ogniwo pobierania informacji w kodzie doznaniowym, albo też uzyskać odpowiednią informację językową, a wtedy łańcuch informacyjny uległby dalszemu przedłużeniu (teoretycznie o dowolną liczbę ogniw).

Metainformacyjną funkcję pełni też w języku naturalnym tzw. **kategoria określoności i kategoria nieokreśloności**, która w niektórych językach, np. w angielskim, niemieckim i bułgarskim, otrzymuje gramatyczny wykładnik w postaci tzw. rodzajników, w innych, np. w polskim, wykładniki leksykalne lub pozycyjne. Informują one o tym, czy nadawca komunikatu, ewentualnie również jego odbiorca, posiadają o obiekcie (szeroko rozumianym, a więc nie tylko materialnie), nazywanym (denotowanym) przez towarzyszący takiemu wykładnikowi rzeczownik, informację pozwalającą na jego identyfikację oraz ewentualnie o tym, czy nadawca komunikatu zamierza, chce, ją udostępnić odbiorcy komunikatu, czy też nie. Za ilustracje, mogą tu posłużyć następujące zdania:

- (1) *Uczeń ten przyniósł mi list od znajomych.*
- (2) *Ten uczeń przyniósł mi list od znajomych.*
- (3) *Jakiś uczeń przyniósł mi list od znajomych.*
- (4) *Pewien uczeń przyniósł mi list od znajomych.*

które informują o następujących sytuacjach informacyjnych:

- (1) informację pozwalającą na zidentyfikowanie ucznia nadawca przekazał odbiorcy we wcześniejszym komunikacie;

(2) informację pozwalającą na identyfikację ucznia nadawca posiada i przekazuje ją odbiorcy (często w akcie komunikacyjnym towarzyszącym zachowaniu językowemu wyrażonym w innym kodzie komunikacyjnym, na przykład za pomocą gestu);

(3) informacji pozwalającej na identyfikację ucznia nadawca komunikatu nie posiada i informuje o tym odbiorcę komunikatu;

(4) informację pozwalającą na identyfikację ucznia nadawca posiada (uzyskał ją we wcześniejszym komunikacie, np. przez obserwację rzeczywistości), ale nie zamierza przekazać odbiorcy komunikatu, o czym go jednocześnie informuje.

Podobnie można wyeksplikować metainformację zawartą w innych leksykalnych wykładnikach określoności/nieokreśloności, np.:

któryś – Któryś z uczniów przyniósł mi list od znajomych.

– informacji pozwalającej na identyfikację ucznia nadawca komunikatu nie posiada i informuje o tym odbiorcę komunikatu;

jeden z – Jeden z uczniów przyniósł mi list od znajomych.

– informację pozwalającą na identyfikację ucznia nadawca posiada, ale nie zamierza przekazać jej odbiorcy komunikatu, informując go jednocześnie o tym;

ktoś – Ktoś spośród uczniów przyniósł mi list od znajomych.

– co wydaje się bardziej skomplikowane, może bowiem odpowiadać bądź sytuacji 3, bądź sytuacji 4.

Oczywiście we wszystkich tych wypadkach określenie tego, co należy uznać za informację pozwalającą na identyfikację denotatu występującego w wypowiedzeniu rzeczownika, nie stanowi problemu lingwistycznego i nie wchodzi do eksplikacji towarzyszących rzeczownikowi wykładników metainformacyjnych, nie wchodzi więc także do eksplikacji znaczenia **rodzajnika określonego**, który zawiera następującą metainformację: informację pozwalającą na zidentyfikowanie elementu, którego dotyczy informacja nadawana w komunikacie, wśród wszystkich elementów zbioru denotowanego przez dany rzeczownik nadawca komunikatu przekazał jego odbiorcy we wcześniejszej części aktu komunikacyjnego, a obecnie informuje go o tym, że rzeczownik użyty aktualnie (z rodzajnikiem określonym) ma takie samo odniesienie (określa ten sam element denotowanego zbioru).

Jak widzimy, metainformowanie jest na tyle ważne w procesie komunikowania się członków ludzkich społeczności, że w wielu językach naturalnych wykształciły się specjalne struktury gramatyczne służące właśnie przekazywaniu metainformacji. W języku polskim temu celowi służą specjalne struktury składniowe, tzw. **mowa niezależna**, służąca do przekazywania informacji z dokładnością co do kształtu jej warstwy leksykalnej, np.:

Janek powiedział: Wyjeżdżam do Krakowa w przyszłym tygodniu we wtorek rano o godzinie 7:20.

oraz **mowa zależna**, służąca do przekazywania informacji z dokładnością co do treści relewantnej dla odbiorcy, np.:

Janek powiedział, że wyjeżdża do Krakowa we wtorek rano o godzinie 7:20.

Janek powiedział, że wyjeżdża do Krakowa we wtorek rano.

Janek powiedział, że wyjeżdża do Krakowa we wtorek.

Janek powiedział, że wyjeżdża we wtorek.

Janek powiedział, że wyjeżdża do Krakowa.

Janek powiedział, że wyjeżdża.

W niektórych językach, np. tureckim, bułgarskim i macedońskim, w funkcji metainformacyjnej wyspecjalizowały się osobne formy gramatyczne czasownika. Jest to tzw. **auditivus** (lub **imperceptivus**), który wyraża czynności i stany niezauważane przez nadawcę informacji, ale zasłyszane (od łacińskiego *audire* – słuchać), a więc takie, o których informacja powtarzana jest na podstawie cudzej relacji (**tryb nieświadka**) oraz **perceptivus** służący do przekazywania informacji o stanach i czynnościach bezpośrednio obserwowanych przez nadawcę (**tryb świadka**). Tak więc w tych językach każde zdanie zawiera metainformację, formułując je bowiem trzeba obowiązkowo wybrać jedną z dwóch możliwych do użycia form czasownikowych: perceptywną, która implikuje, że informację tę nadawca odebrał bezpośrednio, po prostu postrzegając rzeczywistość (a więc bierze odpowiedzialność za jej prawdziwość) lub imperceptywną, która mówi o tym, że informacja jest przez nadawcę jedynie za kimś powtarzana (a więc, że za jej prawdziwość nadawca informacji odpowiedzialności przyjąć nie może). W językach tych nie można więc sformułować zdania niezawierającego metainformacji, typu polskiego zdania:

Slawa bawi się lalką.

bo, na przykład w bułgarskim, trzeba wybrać między formą:

Slawa igraje s kuklata.

co znaczy: Slawa bawi się lalką, a informację o tym uzyskano w kodzie doznaniowym, przez obserwację rzeczywistości – mówi o tym, na przykład, polskie zdanie:

Widziałam, że Slawa bawi się lalką.

a formą:

Slawa igrajela s kuklata.

co odpowiada polskiemu zdaniu:

Mówiono mi, że Slawa bawi się lalką.

W języku polskim taką metainformację możemy oczywiście w razie potrzeby wprowadzić – właśnie za pomocą podanych wyżej zdań, ale nie mamy obowiązku za każdym razem tego czynić.

W funkcji metainformacyjnej wyspecjalizowało się szczególnie wiele czasowników, niezwykle „pojemnych” pod względem metainformacyjnym. Wszystkie te czasowniki, poza komunikowaniem rodzaju procesu informacyjnego, zawierają ponadto informację o innych elementach układu informacyjnego. Mamy tu na przykład:

- czasowniki relacjonujące proces odbierania informacji niesformułowanej w języku naturalnym dotyczącej otaczającego odbiorcę środowiska; rozkład tych czasowników w zależności od rodzaju receptora wygląda w języku polskim następująco:

Wzrok	Słuch	Smak	Węch	Dotyk	Bez określenia receptora
<i>widzieć</i>	<i>słyszeć</i>	<i>czuć</i>	<i>czuć</i>	<i>czuć</i>	<i>spostrzec</i>
<i>zobaczyć</i>	<i>usłyszeć</i>	<i>wyczuć</i>	<i>wyczuć</i>	<i>wyczuć</i>	<i>zauważyć</i>
<i>ujrzeć</i>	<i>posłyszeć</i>	<i>poczuć</i>	<i>poczuć</i>	<i>poczuć</i>	<i>zorientować się</i>
<i>dojrzeć</i>	<i>wysłuchać</i>				
<i>ogłądać</i>	<i>podслуuchać</i>				

- czasowniki mówiące o odbieraniu informacji dotyczącej stanu organizmu odbiorcy, np.: *czuć, poczuć*:

Jan czuł, że plecak coraz większym ciężarem pochyła jego plecy.

- czasowniki relacjonujące proces odbierania informacji, której źródła nie można bliżej sprecyzować, np. *czuć, wyczuć*:

Czuję, że stało się coś niedobrego.

- czasowniki mówiące o nastawieniu na odbiór informacji:
 - wizualnej, np.: *patrzeć, przyglądać się, wypatrywać, rozglądać się, podglądać,*
 - akustycznej, np.: *sluchać, przysłuchiwać się, nadsluchiwać, wsluchiwać się, podsluchiwać, osłuchiwać,*
 - za pomocą węchu, np.: *wąchać,* a w stosunku do zwierząt także *węszyć i niuchać,*
 - za pomocą dotyku, np.: *macać, obmacywać,*
 - bez określenia kanału informacyjnego, np.: *poszukiwać, wyszukiwać,*
- czasowniki relacjonujące proces przetwarzania informacji:
 - określające rodzaj transformacji, np.: *domyślić się, obliczyć, odgadnąć, wydedukować, wywnioskować,*
 - wskazujące relację „bycia informacją wejściową”, np.: *widzieć, wynikać,*
- czasowniki relacjonujące językowe procesy informacyjne, takie jak:
 - proces odbierania informacji sformułowanej w języku naturalnym, np.: *sluchać, słyszeć, wysłuchać* – w wypadku informacji akustycznej, *czytać, przeczytać, odczytać* – w wypadku informacji wizualnej,
 - proces nadawania informacji językowej, a wśród nich czasowniki:

- charakteryzujące stronę fizyczną tekstu, np.: *bazgrać, gryzmolić, kaligrafować* – w wypadku tekstu pisanego; *belkotać, burczeć, chrypieć, jękać się, krzyczeć, paplać, syczeć, szeptać, szwargotać* – w wypadku tekstu mówionego,
- czasowniki mówiące o istnieniu szumu w kanale informacyjnym, np.: *belkotać, mamrotać, seplenić, dukać, bąkać,*
- czasowniki charakteryzujące informację, a wśród nich:
 - określające formę i zawartość treściową informacji, np.: *adresować, kolegdować, dedykować, przeklinać, zaklinać, protokółować.*
 - określające informację dotyczącą zdarzeń przyszłych, np.: *przepowiadać, prorokować, wróżyć,*
 - określające stosunek treściowy informacji do innych informacji, np.: *argumentować, komentować, objaśniać, potwierdzać, negować, przeczyć, tłumaczyć,*
- czasowniki określające stosunek nadawcy informacji do obiektów, o których informuje, np.: *drwić, żartować, szydzić, ironizować,*
- czasowniki charakteryzujące funkcję informacji, np.: *polecać, zalecać, kazać, prosić, zachęcać, namawiać, żądać* – w wypadku funkcji apelatywnej; *obiecywać, przyrzekać, ślubować, zobowiązywać się* – w wypadku funkcji deklaratywnej,
- czasowniki charakteryzujące stan umysłowy nadawcy informacji, np.: *majaczyć, pleść,*
- czasowniki charakteryzujące sytuację informacyjną, np.: *zagaić, zagadnąć, zwrócić się* – mówiące o rozpoczęciu procesu komunikacyjnego; *odezwać się, wtrącić* – mówiące o przystąpieniu do procesu nadawania informacji; *odpowiedzieć, odburknąć, odrzec* – mówiące o tym, że dana informacja jest odpowiedzią na pytanie,
- czasowniki wskazujące narzędzie informacji, np.: *telefonować, telegrafować, teleksować,*
- czasowniki charakteryzujące stan pamięci urządzenia informacyjnego; np.: *pamiętać, zapomnieć, przypominać sobie, wiedzieć, nie wiedzieć, dowiedzieć się, znać.*

Każda przekazana informacja przedstawia dla jej odbiorcy określoną wartość.

Wartość informacji zależy przede wszystkim od następujących czynników:

- od prawdziwości informacji,
- od zgodności przekazu informacyjnego ze stanem faktycznym tego wycinka rzeczywistości pozajęzykowej, który dana wypowiedź komunikuje,
- od dokładności odwzorowania rzeczywistości w komunikacie,
- od stanu poinformowania odbiorcy, posiadanych już przez niego informacji,
- od prawdopodobieństwa uzyskania przez odbiorcę danej informacji,
- od stopnia redundancji informacji, rozumianej jako nadmiarowość sygnału, a więc stosunek elementów jego formy do niesionej przez nie treści,
- od stopnia zniekształcenia sygnału informacyjnego, tzw. szumu informacyjnego.

Oceny wartości informacji można dokonywać z dwóch punktów widzenia:

- punktu widzenia **odbiorcy informacji**, przy czym oczywiście nie musi nim być, choć może, odbiorca, dla którego informacja została przeznaczona, czyli adresat informacji,
- z punktu widzenia **nadawcy informacji**, który równocześnie z nadaniem informacji komunikuje lub sugeruje adresatowi jej wartość.

Najważniejszym kryterium w ocenie wartości informacji jest niewątpliwie jej **prawdziwość**. Znajduje to potwierdzenie w języku: stosunkowo dużo wyrazów języka polskiego zawiera informację dotyczącą prawdziwości odebranej informacji. Zależnie od innych ich komponentów semantycznych, ale tylko tych, które związane są w pewien sposób z oceną prawdziwości (pozostałe komponenty znaczenia nie będą nas tu interesować), wyrazy zawierające tego typu metainformację podzielić można na różne klasy.

Opinia odbiorcy o prawdziwości przekazanej przez danego nadawcę informacji zawarta jest w wielu wyrazach języka polskiego. Jak zobaczymy komunikowana jest tu przede wszystkim fałszywość informacji – na przykład stopień tej fałszywości – co jest chyba skutkiem tego, że komunikowanie informacji prawdziwych leży u podstaw językowej sytuacji komunikacyjnej, stąd relewantne dla odbiorcy są przede wszystkim odstępstwa od tej normy. Komunikowanie, że nadana informacja jest prawdziwa, poza szczególnymi wypadkami, którymi zajmujemy się dalej, byłoby po prostu nieekonomiczne, jako że dotyczyłoby większości komunikatów.

Wyrazy zawierające komponent semantyczny mówiący o tym, że nadawca nadał informację nieprawdziwą niecelowo, nie wiedząc o tym, że jest ona nieprawdziwa, to: *mylić się, pomylić się, pomyłka, błędny, błąd*.

O wiele więcej wyrazów zawiera komponent semantyczny mówiący o tym, że nieprawdziwą informację nadawca nadał celowo, wiedząc o tym, że jest nieprawdziwa, zakładając, że będzie przez adresata przyjęta jako prawdziwa. Wyrazy te podzielić można na kilka klas. Są to wyrazy komunikujące ocenę informacji nadanej w języku naturalnym, na przykład nazywające proces, w którym została nadana informacja nieprawdziwa, np. *klamać, okłamać, okłamywać, skłamać, łgać, zelgać, blagować, bujać, zbujać, bajcować, cyganić, ocyganić, ludzić*; wyrazy nazywające taki nieprawdziwy komunikat, np. *klamstwo, łgarstwo, blaga, bujda*; wyrazy nazywające cechę fałszywości takiego komunikatu, np. *klamliwy*; wyrazy nazywające nadawcę takiego nieprawdziwego komunikatu, np. *klamca, kłamczuch, łgarz, blagier, cygan* (w znaczeniu przenośnym). Wyrazy te różnią się między sobą komponentami innymi niż wyżej wymienione, np. oceną moralną postępowania nadawcy – takiego komponentu można się chyba dopatrzeć w opozycji *klamać* : *blagować*, oceną intencji towarzyszących postępowaniu nadawcy – np. opozycja *klamać* : *ludzić*, albo nacechowaniem środowiskowym lub stylistycznym, np. opozycja *klamać* : *bajerować*, dokładniejsza analiza tych komponentów wykracza jednak poza ramy nakreślone tematem, nie są one bowiem wykładnikami wartości informacji, choć mogą być oczywiście wykładnikami wartościowania innych zjawisk.

Odrębną podklasę stanowią wyrazy oceniające informację niekoniecznie sformułowaną w języku naturalnym, np. wyrazy nazywające procesy informacyjne: *oszustwo, oszukiwać, oszukać, kiwać, kiwnąć, wykiwać, kantować, okantować, nabierać, nabrać, oszwabiać, oszwabić, mamić, omamić, wabić, zwabić* oraz wyrazy nazywające wykonawcę takiego procesu, np. *oszust, kanciarz*.

Z podanych tu przykładów widać, że informację rozumiemy w najszerszym znaczeniu tego słowa – zależnie od jej charakteru możemy tu wyróżnić: wyrazy oceniające informację w zasadzie niejęzykową, a wśród nich nazywające proces informacyjny: *udawać, symulować, symulacja, pozorować, upozorować, fingo- wać, sfingować, fałszować, fałszerstwo*, wyrazy oceniające nadawcę takiego komunikatu, np. *udawacz, symulant, pozorant, fałszerz*, a także wyrazy nazywające taki komunikat, np. *falsyfikat*.

Bywa i tak, że nadawca przekazuje informację nieprawdziwą nie tylko po to, żeby wprowadzić adresata w błąd, ale również po to, aby w wyniku tego wywołać odpowiednią jego reakcję. Komunikat ma tu więc spełniać odpowiednią rolę sprawczą. O takich sytuacjach informacyjnych mówią na przykład następujące wyrazy języka polskiego: wyrazy nazywające sytuację informacyjną, np. *blefować, blef, podstęp, prowokować, prowokacja, fortel*, wyrazy nazywające nadawcę takiego komunikatu, np.: *prowokator*. Innymi komponentami znaczenia tych wyrazów, na przykład zawierającymi ocenę moralną takiego zachowania, tak jak w opozycji *fortel* : *prowokacja*, nie będziemy się tu zajmować.

Celem nadania informacji nieprawdziwej bywa ukrycie przed adresatem informacji prawdziwej. O takich procesach informacyjnych mówią na przykład następujące wyrazy: *zakłamywać, zakłamać* (np. *historię*), *zwieść, zwodzić*, a także: *maskować, zamaskować, tuszować, zatuszować, kamuflować, kamuflaż*, które nazywają procesy informacyjne nastawione na potencjalnego odbiorcę (informacja nie musi być odebrana).

Również ukrycie informacji prawdziwej ma na celu nadawanie informacji z celowo wprowadzonym szumem informacyjnym w plamie treści (do zjawiska szumu informacyjnego jeszcze wrócimy). O takich zachowaniach informacyjnych mówią na przykład następujące wyrazy: *dezorientować, dezinformować, kręactwo, kręcić, matactwa*, a także wyrazy nazywające nadawcę takiej informacji, np. *kręacz, matacz*. Używane również w takiej funkcji wyrazy *gmatwać, zagmatwać* nie przesądzają chyba o celowym działaniu, choć informują oczywiście o istnieniu szumu informacyjnego.

Tu należałoby wspomnieć o takich sytuacjach, gdy informacja prawdziwa celowo nie zostaje przez jej dysponenta nadana, choć przedstawia sobą dla odbiorcy określoną wartość – o takich sytuacjach mówią na przykład wyrazy: *ukryć, ukrywać, zatajać, zataić*.

W omawianej tu klasie wyrazów komunikujących celowe nadanie informacji nieprawdziwej po to, żeby wprowadzić adresata w błąd, wyróżnić można podklasę wyrazów zawierających ponadto komponent znaczeniowy komunikujący treść nadanej informacji, który najogólniej można by określić jako: informacja podaje nieprawdziwy, bo nieistotny w danej sytuacji, powód zachowania nadawcy

w sytuacji, gdy nadawca nie chce zachować się w sposób od niego oczekiwany, O tym informują na przykład wyrazy: *wykręcać się, wykręcić się, wykręt, wymawiać się, wymówić się, wymówka*, a także *pretekst*.

Potrójna ocena wartości – na poziomie metainformacji oraz na poziomie informacji – komunikowana jest przez niektóre wyrazy języka polskiego. Owa zbitka występuje w wyrazach, których zasadniczą funkcją jest funkcja metainformacyjna – podobnie jak poprzednio omawiane komunikują one pewne celowe działania informacyjne podające informację nieprawdziwą w celu wprowadzenia adresata w błąd, zawierają jednak ponadto wartościowanie takiego zachowania informacyjnego pod względem moralnym – oceniają je jako złe – oraz informują, że oceniana informacja jako nieprawdziwa sama zawierała ocenę komunikowanej sytuacji, wartościując czyjeś zachowanie jako złe. Te wykładniki językowe zmieniają więc wartość owej komunikowanej oceny. Chodzi tu o takie wyrazy, jak: *obmawiać, obmówić, obmowa, pomawiać, pomówić, pomówienie, oczerniać, oczernić, obsmarowywać, obsmarować, oszczerstwo, oszczerczy, szkalować, oszkalować, spotwarzać, spotwarzyć, potwarz, kalumnia, inwektywa, osławiać, osławić, osławienie, znieślawiać, znieślawić, znieślawienie*, a także chyba *deprecjonować* (wahanie dotyczy obecności oceny moralnej zachowania informacyjnego) oraz o wyrazy nazywające nadawcę takiej nieprawdziwej informacji: *oszczerca, potwarca*.

Omówione wyżej wyrazy języka polskiego podawały ocenę prawdziwości informacji dokonywaną z punktu widzenia odbiorcy informacji niekoniecznie będącego jej adresatem – to odbiorca dysponował odpowiednią wiedzą o rzeczywistości pozwalającą na weryfikowanie informacji – nie komunikowały natomiast możliwości weryfikacji informacji przez jej adresata – była ona w tym wypadku irrelevantna. Są jednak wśród wyrazów języka polskiego oceniających prawdziwość informacji również takie, gdzie informacja taka jest relevantna, komunikują one, że adresat zna informację prawdziwą (stan faktyczny), co presuponuje nadawca metainformacji dokonujący jednocześnie jej oceny jako nieprawdziwej i komunikujący, że została ona nadana w celu, aby wmówić adresatowi jej prawdziwość. Taką informację zawierają na przykład następujące wyrazy: *wmawiać, wmówić, imputować, insynuować, insynuacja* oraz nazywający nadawcę takiej informacji rzeczownik *insynuator*.

Rozpatrywane dotąd wyrazy języka polskiego komunikujące ocenę prawdziwości informacji jednoznacznie wartościowały nadaną informację jako nieprawdziwą, istnieją jednak w systemie leksykalnym języka polskiego i takie wyrazy, które komunikują niepewność nadawcy metainformacji co do prawdziwości informacji. Tego typu metainformacja komunikowana jest na przykład przez wyrazy: *plotkować, plotka, pogłoska, kmeraże*. Wydaje się, że wyrażenia te presuponują jednocześnie niski stopień wiarygodności informacji – nadawca metainformacji sądzi, że raczej jest ona nieprawdziwa, choć nie wyklucza jej prawdziwości. Podobna niepewność co do prawdziwości informacji komunikowana bywa również w tych sytuacjach, gdy nadawca metainformacji jest równocześnie nadawcą informacji i ustosunkowuje się do informacji nadawanej w tym samym akcie komunikacyjnym (poprzednio rozpatrywaliśmy sytuacje, gdy nadawca metainformacji oceniał

prawdziwość informacji nadawanej w innym akcie komunikacyjnym, zazwyczaj przez innego nadawcę). Taka niepewność komunikowana jest w języku polskim na przykład przez wyraz *jakoby*, komunikujący jednocześnie, że nadawca metainformacji uzyskał ocenianą informację od innego nadawcy w postaci komunikatu językowego.

Inne wyrażenia zawierające komponent semantyczny informujący o niepewności nadawcy informacji co do jej prawdziwości komunikują jednocześnie, która z dwóch wartości (prawda czy nieprawda) jest zdaniem nadawcy metainformacji bardziej prawdopodobna. I tak informację o tym, że nadawca sądzi, że informacja jest nieprawdziwa, zawierają na przykład następujące wyrazy: *rzekomo, pozornie, na pozór, wątpliwe, wątplić, wątpliwość, powątpiewać, powątpiewanie*, natomiast komponent mówiący o tym, że nadawca metainformacji sądzi, że informacja jest prawdziwa, zawierają na przykład następujące wyrazy: *chyba, prawdopodobnie, wydawać się, poniekąd, widocznie, zapewne, być przeświadczonym, mieć przeświadczenie, być przekonanym, mieć przekonanie, niewątpliwie*, a chyba także *podejrzewać*. Wyrazy te różnią się między sobą stopniem przekonania nadawcy metainformacji (zakładanym prawdopodobieństwem) prawdziwości informacji.

Ocena prawdziwości informacji może być też dokonana pośrednio, poprzez podanie oceny jej nadawcy, kwalifikującej go jako nadawcę komunikującego zazwyczaj informacje prawdziwe – taka ocena zawarta jest w wyrazach *prawdomówny* i *wiarygodny* lub jako komunikującego informacje zazwyczaj nieprawdziwe – obok wymienionych już wyżej wyrazów *klamca, klamczuch, łgarz, blagier* podać tu jeszcze można *niewiarygodny*, przy czym *wiarygodny* i *niewiarygodny* mają także inne znaczenia metainformacyjne, mogą bowiem określać informację jako pochodzącą od scharakteryzowanych wyżej nadawców. Przymiotnik *niewiarygodny* może też być użyty w znaczeniu nieprzesądającym o prawdziwości informacji, a wówczas jego znaczenie metainformacyjne można wyeksplikować jako: informacja, w której prawdziwość trudno uwierzyć.

Jak już widzieliśmy, komunikat przedstawia sobą dla odbiorcy różną wartość zależnie od tego, czy zawarte w nim informacje są zgodne z rzeczywistością, czy też nie. Zgodność z rzeczywistością komunikuje przede wszystkim słowo *prawda*, niezgodność słowa *nieprawda, fałsz*.

Prawda jest zawsze tylko jedna, natomiast odstępstwa od prawdy mogą mieć różny wymiar i różny charakter, dlatego też odbiorca informacji często zmuszony jest do dokonywania odpowiednich korektur przekazywanych mu informacji, szczególnie zaś do korygowania komunikowanych mu ocen rzeczywistości. O konieczności dokonania takich poprawek, z równoczesnym wskazaniem na to, że podana wartość 'dobre' jest zawyżona, zaś podana wartość 'złe' jest zaniżona, komunikują na przykład następujące wyrażenia: *przejaskrawiać, wyjaskrawiać, przesadzać, przesada, przesadny, przerysowywać, wyolbrzymiać, koloryzować, wypaczać* oraz przymiotnik *tendencyjny*. Tu także należałoby chyba wymienić takie wyrażenia, jak *parodiować, parodia, parodystyczny, karykaturować, karykatura*, a także *reklamować, reklama* i *propaganda*, które komunikują, że wyrażone w komunikacie cechy rzeczywistości mają wartość zawyżoną, a także

komunikujące chyba zaniżenie wartości cech wyrażenia: *przybliżenie, uproszczenie, symplifikacja* oraz *eufemizm, eufemistyczny*.

Metainformację mówiącą o tym, że podana w komunikacie ocena rzeczywistości jest zaniżona ze względu na wartość dobre : złe, zawierają wyrażenia: *pesymistyczny, pesymizm, pesymista, katastroficzny, katastrofizm, czarnowidztwo, czarnowidz, malkontenctwo, malkontent*, różniące się, jak się wydaje, między innymi miarą dystansu od oceny faktycznie przysługującej, natomiast informacje o tym, że podana w komunikacie ocena rzeczywistości jest zawyżona ze względu na wartość dobre : złe komunikują na przykład wyrażenia: *optymistyczny, optymizm, optymista, idealistyczny, idealizować, wyidealizować, idealizacja*. Do tych dwóch klas należałoby chyba również włączyć takie nazwy jak *pamflet, paszkwil* – wartość zaniżona oraz *panegiryk, dytyramb* – wartość zawyżona.

Wartość, jaką przedstawiają dla odbiorcy informacji wyrażone w komunikacie oceny rzeczywistości, komunikują także wyrażenia *obiektywny, obiektywizm* mówiące o braku konieczności wprowadzania do oceny jakichkolwiek korektur, oraz *subiektywny, subiektywizm* mówiące o możliwości potrzeby ich uwzględnienia, przy czym użycie tych wyrażen niekoniecznie musi przesądzać o faktycznej prawdziwości czy fałszywości przekazanej w komunikacie informacji.

Wartość komunikatu oceniana na podstawie relacji zachodzącej między informacją a rzeczywistością nie zawsze może być opisywana w kategoriach prawdy i fałszu. Wydaje się, że jest to możliwe tylko wtedy, gdy zakłada się istnienie jakiegoś przyporządkowania między elementami komunikatu i elementami faktycznie istniejącej rzeczywistości, czyli konkretnej sytuacji. Czasami bywa jednak tak, że dziedzina będąca podstawą weryfikacji prawdziwości przekazywanej w komunikacie informacji jest zbiorem pustym, czyli, mówiąc inaczej, informacji, której dotyczy dany komunikat metainformacyjny, nie odpowiada żadna realnie istniejąca sytuacja. Prymarnym źródłem informacji jest w tym wypadku jej nadawca, nie uzyskał jej bowiem w wyniku żadnego innego wcześniejszego procesu informacyjnego. O takim właśnie braku odniesienia informacji do rzeczywistości komunikują następujące metainformacyjne wykładniki języka polskiego: *roić, uroić, urojenie, ubrdać sobie, wyimaginować, imaginacja, wyobrażać sobie, wyobrazić sobie, wyobrażenie, mrzonka, fantazjować, złudzenie, złuda, majaczyć, majak, konfabulacja, konfabulować, halucynacja, mara, iluzja, miraż, omam, fatamorgana, zmyślać*.

Większość z omawianych dotąd wyrażen metainformacyjnych oceniało wartość informacji z punktu widzenia jej odbiorcy – punkt widzenia nadawcy reprezentowany był tylko przez klasę wyrażen komunikujących niepewność co do prawdziwości przekazywanej informacji. Nie są to jednak jedyne wypadki, gdy nadawca informacji sugeruje adresatowi informacji ocenę jej prawdziwości.

Wśród wyrażen metainformacyjnych oceniających prawdziwość informacji przez jej nadawcę wyróżnić można klasę wyrażen deklaratywnych – istnieją takie sytuacje, których nadawcy informacji szczególnie zależy na tym, aby adresat komunikatu uwierzył, że przekazywana mu informacja jest prawdziwa. W takich wypadkach nadawanej informacji towarzyszy specjalny komponent semantyczny

będący apelem do adresata o treści: „Mówię prawdę. Uwierz mi!” Taka metainformacja zawarta jest, naszym zdaniem, na przykład w następujących wyrażeniach języka polskiego: *przysięgać, przysiąc, przysięga, gwarantować, zagwarantować, gwarancja, ręczyć, zaręczać, deklarować, zadeklarować, deklaracja, zaklinać się, zakląć się*, a także *zaświadczać, zaświadczyć*, który choć ocenia prawdziwość informacji komunikowanej przez innego nadawcę, zawiera równocześnie wspomniany wyżej apel do adresata. Wyrażenia te w interesującym nas tu aspekcie różnią się między sobą przede wszystkim mocą owego apelu: „Uwierz mi!”

Wśród wyrażeń zawierających w swej strukturze semantycznej komponent „Mówię prawdę. Uwierz mi!” wyraźnie wyodrębnia się podklasa wyrażeń, które w strukturze powierzchniowej zdania konotują zawsze proces nadawania informacji, które dotyczą zdarzeń przyszłych w stosunku do sytuacji informacyjnej i to tylko takich zdarzeń, które mogą zaistnieć tylko wtedy, jeśli przyczyni się do tego nadawca informacji. Wykształcenie się w języku takiej klasy wyrażeń metainformacyjnych wydaje się uzasadnione tym, że rzeczywiście zwykle najbardziej zależy nam na przekonaniu adresata o prawdziwości naszych słów, gdy dotyczą one naszej gotowości do działania, o niej bowiem adresat informacji w żaden inny obiektywny sposób przekonać się nie może, gdyż informacja taka nie jest weryfikowalna przed upływem czasu, w którym mają zajść zdarzenia, o których mowa w komunikacie. W takich wyrażeniach metainformacyjnych ów apel nadawcy do adresata zostaje rozszerzony do postaci: „Mówię prawdę, że chcę to zrobić, o czym mówię. Uwierz mi!” Apel taki, towarzyszący nadawanej informacji, może być wyrażany bardzo różnymi środkami, zarówno językowymi jak i pozajęzykowymi, na przykład gestem, odpowiednią intonacją lub nawet sytuacyjnie, często nawet bywa wyrażany leksykalnie, na przykład przez użycie następujących wyrazów: *obietnicować, obiecać, obietnica, przyobiecycić, przyobiecować, naobiecycić, przyrzekać, przyrzec, przyrzeczenie, zobowizywać się, zobowizać się, zobowizanie, ślubować, ślubowanie, śluby*. Apel: „Uwierz mi, że mówię prawdę!” może mieć różną moc, która jest chyba cechą różniącą wszystkie omawiane tu wyrażenia.

Szczerość tego apelu może być podważona przez odbiorcę informacji i zakomunikowania np. przez wyrażenia: *obietcaniki* lub *obietcanki cacanki*, z komponentem semantycznym: „To nie jest prawda”.

Poszukiwanie wciąż nowej informacji jest realizacją pasji odkrywczej, jaka cechuje każdego rozumnego człowieka. Bardzo często nową informację człowiek otrzymuje nie bezpośrednio z otaczającego go świata za pomocą narządów zmysłów, ale dochodzi do niej na drodze, często bardzo skomplikowanych, operacji myślowych przeprowadzanych na informacjach, które uprzednio uzyskał. W tym wypadku źródłem nowej informacji są inne informacje, które ulegają odpowiedniemu przetworzeniu. Te skomplikowane procesy myślowe, prowadzące do uzyskania nowych informacji, to po prostu odpowiednie transformowanie informacji wejściowych. Umysł ludzki jest więc systemem, który z pewnych informacji otrzymywanych na wejściu systemu, daje na wyjściu nową informację. Informacja wejściowa to nic innego jak przesłanka, informacja wyjściowa jest

konkluzją, do jakiej na mocy tej przesłanki dochodzi się w procesie wnioskowania. Ta nowa informacja przedstawia sobą różną wartość, gdyż tylko wyjątkowo mamy tu do czynienia z wnioskowaniem logicznym, najczęściej zaś jest to wnioskowanie uprawdopodobniające, a więc wnioskowanie redukcyjne, wnioskowanie przez analogię lub wnioskowanie statystyczne. Do takiego samego wniosku można dojść na podstawie różnego rodzaju przesłanek, mających niejako różną „moc”, a od tego zależy wartość uzyskanej w procesie transformacji informacji. Określenie rodzaju przesłanki w pewnym sensie określa więc, w sposób pośredni, stopień prawdopodobieństwa konkluzji, a więc i wartość informacji. Taki komponent znaczeniowy zawierają na przykład następujące wyrażenia języka polskiego: *dowodzić, dowieść, dowód, implikować, implikacja, oznaczać, oznaka, świadczyć, świadectwo, uzasadniać, uzasadnić, uzasadnienie, wróżyć, wróżba, wskazywać, wskazówka, sygnalizować, sygnał, zapowiadać, zapowiedź, zwiastować, zwiastun.*

Wartość uzyskanej w wyniku transformacji informacji może też być określona bezpośrednio, przez podanie stopnia prawdopodobieństwa konkluzji. Taki komponent znaczeniowy zawierają na przykład następujące wyrażenia języka polskiego: *wniosek, hipoteza, teza, przypuszczenie, przypuszczać, twierdzenie, domniemywać, domniemanie, domysł, konkluzja.*

Pośrednio na wartość uzyskanej w wyniku przetworzenia informacji wskazuje również określenie procesu transformacyjnego, a więc rodzaju transformacji. Funkcje tę pełnią na przykład wyrażenia: *domyślić się, domyślać się, obliczać, obliczyć, obliczenie, wyliczać, wyliczyć, wyliczenie, oceniać, ocenić, odgadywać, odgadnąć, taksować, otaksować, szacować, oszacować, dedukować, wydedukować, kalkulować, wykalkulować, kombinować, wykombinować, koncytować, wykoncytować, spekulować, wyspekulować, zgadywać, zgadnąć.*

Wartość informacji jest wielkością względną i zależną od stanu poinformowania odbiorcy. Największą wartość ma oczywiście dla niego informacja nowa, jeszcze nieznaną. Nowość informacji komunikują słowa *nowina, odkrycie*, a także *wykryć, wyjaśnić, wysledzić, wytropić, wywąchać, odgadnąć, dociec*, które informują ponadto, że informacja była oczekiwana przez odbiorcę. Wykładniki wartości informacji języka naturalnego pokazują, że potwierdzenie prawdziwości informacji będącej hipotezą odbiorcy, a więc zniesienie jego stanu niepewności, przedstawia dla niego bardzo dużą wartość. Inną oczywiście wartość przedstawia sobą informacja dobrze już odbiorcy znana. Tu rzeczywiście ma ona dla niego wartość minimalną. Taką wartość informacji komunikują wyrażenia: *banal, banalny, trywialny, frazes, slogan, komunal, ogólnik*, a także *oczywisty, oczywiście, oczywistość, rzeczywiście, istotnie.*

W języku polskim istnieją również wyrażenia komunikujące niezgodność informacji z informacjami posiadanymi przez odbiorcę. Są to na przykład wyrazy: *niedorzeczny, niedorzeczność, paradoksalny, paradoks, absurdalny, absurd, nonsensowny, nonsens, bezsensowny* a także *herezja*, oraz wyrazy komunikujące niespójność informacji w komunikacie, np. *brednia, bredzić, banialuki, androny i ambaje.*

W kategoriach nowości informacji można chyba także rozpatrywać cechę oryginalności informacji – informacja nieoryginalna jest zawsze informacją znaną odbiorcy. Językowo opozycja ta rozbudowana jest nierównomiernie – wyrażeniom *oryginalny*, *oryginal* przeciwstawia się wiele wyrażen, np. *naśladownictwo*, *trawestować*, *trawestacja*, *parafrazować*, *parafraza*, *apokryficzny*, *apokryf*, *pastisz*, *plagiat*. Tu także chyba trzeba zaliczyć wyrazy *cytować*, *cytat*, *przyczytać*, *przyczytać*, *przyczytanie* jako wykładniki informacji o informacji powtórzonej za innym nadawcą z dokładnością co do jej warstwy leksykalnej lub tylko co do uznanych przez nadawcę za ważne dla adresata treści.

Wartość informacji dla odbiorcy zależy także od obecności lub braku szumu informacyjnego, przede wszystkim w planie treści. O istnieniu takiego szumu, wprowadzonego specjalnie przez nadawcę w celu ukrycia informacji prawdziwej, komunikują wymienione już wyrażenia *kręactwo*, *krętać*, *gmatwać*, *zagmentać*, *mactwa*, *mactać*. Istnienie szumu w planie treści, ale już bez przesądzania o jego celowym wprowadzeniu, komunikują także wyrazy: *mętniactwo*, *mętniak*, oraz komunikujące w swym podstawowym znaczeniu szum w planie wyrażenia wyrazy *belkotać* i *belkot* (na marginesie należy dodać, że znacznie więcej wyrażen języka polskiego informuje o istnieniu szumu informacyjnego w planie wyrażenia, czyli o występowaniu pewnych zakłóceń w sygnale).

Jedną z cech informacji decydujących o ocenie jej wartości jest redundancja. Redundancja, rozumiana jako stosunek między liczbą elementów planu wyrażania komunikatu a liczbą odwzorowywanych przez nie elementarnych informacji, jest jedną z cech każdego dobrego kodu, decydującą o jego optymalnym funkcjonowaniu, wartość jej nie jest więc dla odbiorcy relewantna. Mówiąc o redundancji mamy na myśli stosunek informacji odwzorowanej w komunikacie do ilości informacji ważnych dla odbiorcy, a taka redundancja jest oczywiście dla odbiorcy relewantna. Interesujące, że znacznie więcej wyrażen komunikuje brak takiej redundancji niż jej obecność. Do pierwszej z wymienionych tu klas należą wyrażenia *zwięzły*, *zwięzłość*, *treściwy*, *rzeczowy*, *rzeczowość*, *lapidarny*, *lapidarność*, *skrótowy*, *skrót*, do drugiej chyba tylko *rozwlekły*, *rozwlekłość*, *wodolejstwo* oraz użyty metaforycznie wyraz *woda*.

Ostatnią już chyba cechą, według której określa odbiorca wartość informacji, jest stosunek między liczbą elementów i łączących je relacji występujących w odwzorowywanej w komunikacie rzeczywistości a liczbą odwzorowanych elementów i relacji w komunikacie. Wartość tego współczynnika komunikują takie wyrażenia jak *szczegółowy*, *drobiazgowy*, *dokładny* w opozycji do *ogólny*, *szkicowy*, *szkic*, *zarys*.

Wymienione wyżej wykładniki wartości informacji nie uwzględniają współczynnika relewancji ich dla odbiorcy. Tę wartość komunikują takie wyrażenia jak *zajmujący*, *ciekawyy*, *interesujący*, *pasjonujący*, w opozycji do *nudny*, choć ich funkcja metainformacyjna nie jest być może ich funkcją prymarną.

Jeśli informacja relewantna dla odbiorcy jest warunkiem koniecznym do zmiany jego stanu, to informację taką określa się mianem *istotna*, *ważna*, *znacząca*. Wyrażenia *blahy*, *blahostka* znajdują się w opozycji znaczeniowej nie tylko do wyrażen

istotny, ważny, znaczący, ale również do wyrażen określających w ogóle informację relewantną dla odbiorcy.

Jak widzimy, bardzo wiele różnorodnych elementów języka pełni funkcję wykładników wartości informacji. Znajomość tych wykładników jest istotnym wskaźnikiem w ocenie wartości informacji przygotowywanych dla systemów informacji faktograficznej.

Metainformacja odgrywa tak ważną rolę w procesie informacyjnym, że właściwie w każdym dłuższym tekście możemy wydzielić te jego elementy, które pełnią funkcję metainformacyjną, będąc bardzo często wypowiedzią o samej wypowiedzi. Elementy te Anna Wierzbicka nazwała **metatekstem** (*Metatekst w tekście*), podając takie jego przykłady:

W tym rozdziale będę mówić o...

Potem przejdę do...

Na zakończenie przedstawię...

Mówilem dotychczas o...

Z kolei zajmę się...

Omówię teraz...

Podsumujmy dotychczasowe rozważania...

Pora sformułować wnioski...

Aby mówić o..., było konieczne zająć się najpierw...

Zanim przejdę do, muszę zatrzymać się na chwilę na...

Dam przykład...

Można ująć to inaczej...

Proszę mi pozwolić użyć tu pewnej metafory...

Chciałbym się teraz skupić na...

Nie mogę nie wspomnieć o...

Zacznę od...

a także:

przypominam, że...; podkreślam...; odpowiedź...; jeśli chodzi o...; tzn...; inaczej mówiąc...; prawdę powiedziawszy...; a mianowicie...; a propos...; notabene...; ściśle mówiąc...; innymi słowy...; na przykład...; przede wszystkim...; rzekomo...; jakoby...; wątpię, czy...; podobno...; co się tyczy...; co do...

i wiele jeszcze innych. Takie wyrażenia metatekstowe są ważnymi wykładnikami spójności tekstu.

Ważną funkcję metainformacyjną pełni również kształt intonacyjno-dynamiczny tekstu mówionego (ukształtowanie sygnału) oraz układ graficzny tekstu pisanego, stosowane w nim wykładniki przebiegu intonacyjnego wypowiedzenia, takie jak np. nawias, cudzysłów (cudzysłów bywa też oczywiście znakiem przytaczania cudzego tekstu i to jest jego podstawowa funkcja metainformacyjną), a także użycie różnego rodzaju druku (np. druk wyłuszczonej, druk rozstrzelony), zastosowanie podkreśleń, światła, akapitu i wszelkiego innego typu wyróżników

graficznych. Tego typu metainformacje zawarte w tekście mają przede wszystkim za zadanie sterowanie procesem odbierania i przyswajania informacji przez odbiorcę, zwracają bowiem jego uwagę na założony przez nadawcę stopień ważności dla odbiorcy (przydatności) poszczególnych informacji zawartych w danym komunikacie, informując go często także o stopniu nowości tych informacji.

Wydobyte z tekstu elementy metatekstowe stanowią obraz jego struktury, pokazują, jakie relacje łączą zawarte w tekście informacje (**relacje intratekstualne**). Tekst języka naturalnego ma postać linearną, dla jego pełnego odbioru niezbędne bywa korzystanie z zawartych w nim informacji w porządku innym niż ten, który jest narzucony porządkiem wynikającym z czasowego (linearnego) przebiegu komunikatu. To właśnie elementy metatekstu ustanawiają inne możliwe uporządkowanie informacji zawartych w danym tekście. O wiele więcej możliwości ma tu oczywiście tekst utrwalony na jakimś nośniku, na przykład tekst pisany, ale i w tekście nieutrwalonym, bywa, iż nadawca komunikuje odbiorcy, że informacja potrzebna do pełnej interpretacji tekstu aktualnego podana zostanie dopiero w dalszej części komunikatu, np.:

Problem ten wyjaśnimy w dalszej części wykładu.

O tym za chwilę.

Do sprawy tej jeszcze wrócimy.

lub w innym komunikacie, np.:

O tym na następnym wykładzie.

Problem ten omówiony będzie na zajęciach z logiki.

Bywa, że do interpretacji informacji zawartej w danym komunikacie mogą być potrzebne informacje podane we wcześniejszej partii aktualnego komunikatu lub w innych komunikatach. Wówczas także nadawca może na taką konieczność zwrócić uwagę odbiorcy w wypowiedzi metainformacyjnej, np.:

Mówił o tym już...

Jak już mówiłem...

Takie odesłania do innych tekstów, ustanawiające zależność między tekstami, to **relacje intertekstualne**. W tekście pisanim, szczególnie o charakterze naukowym, są one często wyprowadzane poza zasadniczy tekst i komunikowane w przypisach. Charakter tych relacji określany bywa przez system odsyłaczy takich jak *zob.*, *zob. też*, *por.* (*porównaj*). Odsyłacze takie mogą łączyć także elementy wewnątrz tego samego tekstu – są wtedy wykładnikami **relacji intratekstualnych** i często ustanawiają inny niż linearny porządek w tekście.

Podobną funkcję w obrębie tekstu pełnią indeksy (skorowidze), których rolą jest lokalizacja w tekście informacji mających wspólny temat nazwany przez wyrażenie indeksowe. Najczęściej jest to lokalizacja przybliżona, w tekście drukowanym z dokładnością co do strony.

Nieco inną rolę metainformacyjną pełnią w tekście pisanim takie strukturalizujące go elementy jak tytuły rozdziałów, podrozdziałów, paragrafów, a także

żywa pagina i żywy margines. Ich funkcją jest sygnalizowanie tematu, którego dotyczą informacje znajdujące się we fragmencie tekstu objętego ich zasięgiem. Zasięg ten wyznaczany jest najczęściej przez formalny podział tekstu na akapity lub większe całości. Funkcję delimitacyjną pełnią tu same teksty metainformacyjne – w wypadku tytułów rozdziałów lub przerwy w tekście wyznaczające akapity – w wypadku żywego marginesu. Taka strukturalizacja tekstu i jej wykładniki umożliwiają odbiorcy nieliniarny odbiór informacji, pełnią również funkcję wyszukiwawczą w zbiorze informacji, jakim jest każdy tekst, pozwalając użytkownikowi na odnalezienie relewantnej dla niego informacji bez konieczności jej liniarnego przeglądania.

Wszystkie te towarzyszące tekstowi wykładniki metainformacyjne stanowią ważną informację dla osób przygotowujących opracowania dokumentów dla potrzeb systemów informacyjnych, są też niezwykle pomocne przy komputerowym streszczaniu i indeksowaniu dokumentów.

Obie funkcje metainformacyjne: strukturalizacji w obrębie tekstu poprzez wskazanie relacji intratekstualnych oraz strukturalizacji zbioru tekstów poprzez wskazanie łączących je relacji intertekstualnych realizowane są w coraz bardziej popularnych **hipertekstowych systemach informacyjnych**, przy czym przy strukturalizacji tworzących system wyszukiwawczy zbiorów tekstów korzysta się nie tylko z wykładników metainformacyjnych zawartych bezpośrednio w tekstach (z ich metatekstu), ale znaczną ich część ustanawia się na podstawie relacji skojarzeniowych wynikających ze znajomości tekstów wchodzących w zakres systemu. Ani zakres systemu, ani jego struktura metatekstowa często nie są z góry zadane – system taki ciągle jest uzupełniany nowymi tekstami i nowymi strukturalizującymi go relacjami przez administratorów lub użytkowników systemu. Realizowanie i korzystanie z takich systemów hipertekstowych umożliwiają sieci komputerowe, gdzie zbiór wyszukiwawczy systemu mogą tworzyć teksty realizowane często równocześnie w różnym materiale, a więc nie tylko teksty pisane języka naturalnego, ale również ich realizacja dźwiękowa, obrazy, różnego typu animacje, filmy, zapisy dźwiękowe utworów muzycznych itp. Takie systemy multimedialne nazywane bywają **hipermediami**. Zna je każdy użytkownik Internetu.

Definicję **hipertekstu** podajemy za *Słownikiem encyklopedycznym informacji, języków i systemów informacyjno-wyszukiwawczych*:

„HIPERTEKST – wielopoziomowa struktura, będąca odwzorowaniem cech relewantnych tekstu (także tekstu mieszanego) lub zbioru tekstów, umożliwiająca przekształcenie tekstu liniarnego w postać nieliniarną (najczęściej za pomocą drzew hierarchicznych oraz sieci) w celu ułatwienia odbiorcy wyszukiwania zawartych w nim informacji oraz uzupełniania zbioru informacji wyselekcjonowanych odpowiednio do potrzeb informacyjnych użytkownika. Struktura taka może różnego typu relacjami intertekstualnymi łączyć w jedną całość części jednego dokumentu lub dokumenty zawierające informację dokumentacyjną z odpowiednimi dokumentami źródłowymi lub ich fragmentami i odpowiednią informacją faktograficzną, przy czym w jednym hipertekście mogą występować dokumenty różnych typów.

Podstawowymi **elementami hipertekstu** są węzły i powiązania. **Węzeł** to wydzielony, ze względu na zawartość informacyjną fragment tekstu (słowo, fraza, zdanie, akapit, rysunek, schemat, wideoklip itp.) lub dokument w bazie danych. Węzły są połączone **powiązaniem** (ang. links, cross-references) wyrażającymi: relacje semantyczne, relacje syntagmatyczne i relacje pragmatyczne (np. po zapoznaniu się z informacją w węźle A warto zapoznać się z informacją w węźle B). System hipertekstowy może zawierać powiązania jednego lub wielu typów, istnieje też możliwość określania przez użytkownika typów powiązań, które chce uaktywnić oraz rozbudowywania hipertekstu (interakcja systemu i jego użytkownika). Wyróżnia się następujące podtypy hipertekstu: **mikrotekst** – tekst z wyróżnionymi explicite powiązaniem pomiędzy jego fragmentami (relacje intratekstualne), **makrotekst** – zbiór dokumentów połączonych relacjami intertekstualnymi, **eksperttekst** – połączenie hipertekstu z systemem ekspertowym, którego działanie zostaje ulepszone w wyniku wykorzystywania sieci semantycznej hipertekstu. System hipertekstowy może obejmować jedną lub wiele tekstowych albo multimedialnych baz danych (**hipermedia**).”

Metainformacja odgrywa od dawna tak ważną rolę w funkcjonowaniu społeczeństw ludzkich, że poświęcono jej specjalne wydawnictwa zawierające wyłącznie informację o różnego typu informacjach. Wydawnictwa te – dokumenty zawierające metainformację – to różnego typu bibliografie, indeksy spisów treści czasopism, indeksy słów kluczowych oraz cytowania bibliograficzne i indeksy cytowań. Metainformacja o dokumentach zgromadzonych w dużych zbiorach zawarta jest w katalogu alfabetycznym, katalogu rzeczowym i różnego typu kartotekach tematycznych. Na sporządzaniu metainformacji polega też oczywiście praca dokumentalisty.

Wśród systemów wyszukiwania informacji istnieją systemy wyszukujące właściwie metainformację (tzw. systemy informacji dokumentacyjnej), wydające w odpowiedzi na zapytanie użytkownika informację, w jakich dokumentach można znaleźć informację na określony przez niego temat.

Wypowiedzi metainformacyjne, jak już wiemy, mogą tworzyć swojego rodzaju poziomy – informacja o metainformacji to **metametainformacja**. Wydawnictwa metainformacyjne, czyli zawierające informację o wydawnictwach zawierających metainformację, to różnego typu spisy wydawnictw informacyjnych, takie jak np. bibliografie bibliografii.

Tak jak metajęzyk jest językiem służącym do formułowania informacji o języku, tak języki informacyjno-wyszukiwawcze (zwane też językami informacyjnymi) są językami specjalnie skonstruowanymi do wyrażania metainformacji w tym jej zakresie, który dotyczy informacyjnej zawartości dokumentów. Językami sztucznymi wyspecjalizowanymi w funkcji metainformacyjnej są, jak wiemy, języki informacyjno-wyszukiwawcze.

Literatura

- Wierzbicka A.: *Metatekst w tekście*, w: Mayenowa M.R. (red.): *O spójności tekstu*. Wrocław 1971.
- Bojar B.: *Polskie czasowniki dotyczące procesów informacyjnych. Elementy metainformacji w tekstach języka naturalnego*, w: *Studia językoznawcze. Streszczenia prac doktorskich*, t. 3. Warszawa 1978.
- Bojar B.: *Ocena wartości informacji w leksyce języka polskiego*, w: Bogusławski A. i inni (red.): *Co badania filologiczne mówią o wartości*. Warszawa 1987.

VII

System fonologiczny

Nim przejdziemy do omówienia systemu fonologicznego języka naturalnego (na przykładzie języka polskiego), musimy wyjaśnić pewne sprawy. Wszystkie wiadomości dotyczące fonologii mogą się wydawać czytelnikom, którzy nie są językoznawcami, zbędne. Nie chcemy uczyć fonologii, lecz jedynie wyjaśnić budowę ekonomicznego systemu językowego. Chcemy również pokazać, w jaki sposób można opisywać język naturalny, czyli tworzyć model języka na podstawie jego tekstów (zasady tworzenia modelu każdej rzeczywistości są takie same). Przedstawimy tu model najmniej skomplikowanej warstwy języka (w tym wypadku języka polskiego) – model jego substancji dźwiękowej, starając się znaleźć najbardziej ekonomiczny sposób opisu całego inwentarza głosek.

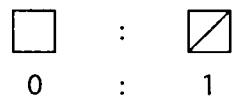
Głoski, czyli najmniejsze elementy tekstu mające przebieg czasowy, są albo tonami, albo połączeniami tonów i szmerów. Wydawałoby się więc, że powinny być opisywane w terminach dotyczących ich budowy akustycznej, opisujących ukształtowanie fali głosowej. My jednak będziemy je opisywać nie w terminach odnoszących się do ich realizacji w konkretnym materiale fizycznym, lecz w terminach dotyczących sposobu ich realizowania – układu narządów mowy przy artykulacji poszczególnych głosek – opisywać więc będziemy w istocie stan urządzenia w momencie realizowania konkretnej głoski. Postępowanie takie jest usprawiedliwione, jeżeli zachodzi izomorfizm między stanem urządzenia produkującego a produktem. Ponadto chodzi przecież jedynie o stworzenie modelu, a ten może być realizowany w dowolnych kategoriach (realizowany w dowolnym materiale), byleby tylko zachowane zostały wszystkie relacje zachodzące między elementami opisywanej rzeczywistości (model powinien być izomorficzny względem modelowanej rzeczywistości).

Każdy **model** jest pewną konstrukcją myślową będącą uproszczonym obrazem badanego fragmentu rzeczywistości, z którego eliminuje się wszystko to, co jest nieistotne z punktu widzenia określonego celu. W wypadku materiału dźwiękowego języka owym kryterium uznawania danej cechy badanego elementu za istotną (relewantną) lub nieistotną (nierelewantną) jest pełnienie przez nią funkcji semantycznej, rozumianej jako zdolność do odróżniania znaczenia jednostek pełnych semantycznie (języki naturalne są, jak pamiętamy, językami dwuklasowymi).

Aby ułatwić czytelnikowi zrozumienie zasad tworzenia modelu, dalsze rozważania poprzedzimy przykładem.

Załóżmy, że tworzymy model odzwierciedlający stany relewantne (czyli uznane za ważne dla danej rzeczywistości, mające zdolność różnicowania jej elementów) nadwozi samochodowych w różnych stadiach ich produkcji. Przyjrzyjmy się zbiorowi takich nadwozi. Znajdują się w nim: nadwozia zabezpieczone podkładem antykorozyjnym i nadwozia niezabezpieczone; nadwozia zabezpieczone podkładem antykorozyjnym i polakierowane oraz nadwozia niepolakierowane (zabezpieczone podkładem antykorozyjnym lub nie), przy czym te polakierowane mogą mieć różne kolory, np.: żółty, czerwony, zielony, granatowy, szary, beżowy, pomarańczowy, biały, wiśniowy; nadwozia już wypalone lub w trakcie wypalania piecowego, wypalone mniej lub bardziej; nadwozia wypolerowane lub będące w trakcie polerowania; nadwozia uzbrojone lub będące w trakcie uzbrajania (uzbrojone mniej lub bardziej). Wszystkie te różnice dadzą się sprowadzić do następujących **opozycji**:

1. Porównywane elementy różnią się tym, że jeden z nich posiada daną cechę, a drugi jej nie posiada. Opozycję taką przedstawimy schematycznie:



Jest to opozycja dwuczłonowa, zero-jedynkowa, a nazywa się ją **opozycją prywatną**, przy czym człon opozycji, w którym nie ma cechy, nazywa się członem nienacechowanym opozycji, zaś człon posiadający cechę członem nacechowanym. Tak więc:

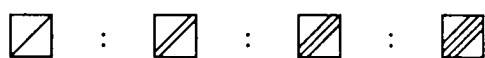
$$\text{człon nacechowany opozycji} = \text{człon nienacechowany} + \text{cecha}$$

Opozycja prywatna jest opozycją wyraźną, łatwo bowiem stwierdzić, które z elementów rzeczywistości stanowią jej człony.

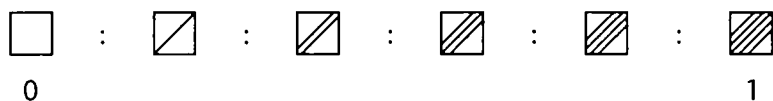
W powyższym przykładzie tworzą ją następujące elementy:

nadwozia zabezpieczone antykorozyjnym podkładem	:	nadwozia niezabezpieczone antykorozyjnym podkładem
nadwozia polakierowane	:	nadwozia niepolakierowane
nadwozia wypalone	:	nadwozia niewypalone
nadwozia wypolerowane	:	nadwozia niewypolerowane
nadwozia uzbrojone	:	nadwozia nieuzbrojone

2. Porównywane elementy różnią się tym, że każdy z nich posiada daną (tę samą) cechę w różnym stopniu jej natężenia. Opozycję taką, zwaną **opozycją stopniową (gradualną)**, ilustruje poniższy schemat:



Można za nią uznać również sytuację przedstawioną na schemacie:

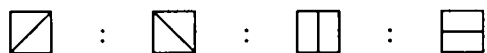


gdzie skrajne człony układają się w opozycję prywatywną.

Opozycja stopniowa daje teoretycznie możliwość skali ciągłej i wymaga uznania pewnych, wystarczająco od siebie różniących się punktów, za dystynktywne (wartości progowe), przy czym w wypadkach granicznych trudność może sprawiać przydział danego elementu do jednej z dwóch sąsiadujących ze sobą klas. W naszym przykładzie opozycję stopniową tworzą następujące elementy:

- nadwozia w różnym stopniu wypalone,
- nadwozia w różnym stopniu wypolerowane,
- nadwozia w różnym stopniu uzbrojone.

3. Porównywane elementy różnią się tym, że w każdym z nich obecna jest pewna cecha, przy czym są to cechy zupełnie różne, ale traktowane równorzędnie pod względem funkcjonalnym. Tego typu opozycja nosi nazwę **opozycji równorzędnej (ekwivalentnej)**, a schematycznie przedstawia ją poniższy rysunek:



Opozycja ta, podobnie jak opozycja stopniowa, jest opozycją wielocłonową. W naszym przykładzie w opozycji takiej występuje różny kolor nadwozi, np.:

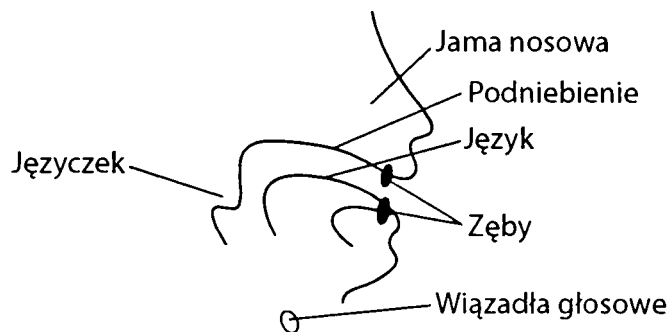
- nadwozia żółte
- nadwozia czerwone
- nadwozia zielone
- nadwozia granatowe
- nadwozia białe

Tę ostatnią cechę można by oczywiście, biorąc pod uwagę jej fizyczne właściwości, rozpatrywać w kategoriach opozycji stopniowej, wiadomo bowiem, że różnica barwy polega na różnicy długości fali świetlnej, w naszym modelu możemy ją jednak uznać za opozycję równorzędną, gdyż uznanie jej za opozycję stopniową wymagałoby zastosowania skomplikowanych urządzeń pomiarowych przy kwalifikowaniu konkretnych obiektów do poszczególnych jej członów.

Wszystkie cechy różniące elementy rzeczywistości sprowadzić się daje do omówionych wyżej opozycji, różnic takich będziemy szukać również w warstwie fonetycznej języka naturalnego (polskiego).

Aby ułatwić zrozumienie zasad tworzenia modelu substancji dźwiękowej języka polskiego, czyli jego systemu fonologicznego, dalsze nasze rozważania poprzedzimy przypomnieniem **budowy narządów mowy**, których schemat przedstawia rysunek.

Narządy mowy człowieka są tak zbudowane, że mogą pozostawać w stosunku do siebie w różnym położeniu, co pozwala na zmianę kształtu rezonatora, a co za tym idzie na zmianę barwy dźwięku. Wydostające się z płuc powietrze, które



jest nośnikiem fali akustycznej, napotyka na swojej drodze różnego rodzaju przeszkody. Pierwszą z nich stanowią wiązadła głosowe, które mogą zajmować dwa zasadnicze położenia:

- rozsunięte nie stanowią zapory dla wydychanego powietrza, a powstała fala głosowa nie stanowi wtedy tonu, lecz jest jedynie szmerem (powstają wtedy głoski bezdźwięczne);
- zsunięte, a wtedy są przez wydychane powietrze wprawiane w drgania, których częstotliwość może się zmieniać zależnie od napięcia wiązadeł głosowych, powodując przy tym zmianę wysokości tonu lub zmiany melodyczne w obrębie jednej sylaby (co w języku polskim nie jest cechą dystynktywną, ale może nią być w innych językach, np. w chińskim), a więc dając w efekcie głoski dźwięczne.

Opozycja dźwięczności : bezdźwięczności jest szeroko wykorzystywana w języku polskim do odróżniania głosek, z których większość różni się tą jedną cechą, tworząc pary bezdźwięczna : dźwięczna, np.:

- p : b (*pas* : *bas*),
- f : w (*fory* : *wory*),
- t : d (*tom* : *dom*),
- s : z (*kosa* : *koza*),
- k : g (*koi* : *goi*).

Opozycja dźwięczność : bezdźwięczność jest dobrą (z punktu widzenia ekonomii języka) opozycją, jak zresztą wszystkie opozycje prywatywne.

Opozycją prywatywną jest w języku polskim również **opozycja twardości : miękkości** głosek, która to cecha akustyczna uzyskiwana jest przez wzniesienie środkowej części języka do podniebienia (miękkość) lub nie (twardość). I w tym przypadku również większość głosek języka polskiego daje się ułożyć w pary twarda : miękka, np.:

- p : p' (*pył* : *pił*),
- c : ć (*cało* : *ciało*),
- b : b' (*był* : *bił*),
- s : ś (*kos* : *koś*).

Za tego samego typu opozycję może być uznana **opozycja głosek ustnych : nosowych**, która to cecha akustyczna uzyskiwana jest przez wyłączenie (ustne)

lub włączenie (nosowe), dzięki odpowiedniemu ustawieniu języczka, dodatkowego rezonatora, jakim jest jama nosowa. A oto pary głosek ustnych : nosowych w języku polskim:

b : m (*był : mył*),
b' : m' (*biły : miły*),
d : n (*dam : nam*),
o : ɔ (*żono : żoną*),
e : ɛ (*te : tę*).

Nie wszystkie opozycje fonologiczne są jednak tak wyraźne jak opozycje prywatywne. Bywa i tak, że człony opozycji przeciwstawiają się sobie różnym stopniem natężenia tej samej cechy, tworzą więc **opozycję stopniową**. Tak jest w wypadku artykulacji głosek różniących się **stopniem zbliżenia narządów mowy**, gdzie opozycję tworzą głoski:

otwarte : półotwarte : szczelinowe : zwarte (zwarto-szczelinowe i zwarto-wybuchowe).

Stopniem otwarcia różnią się też w języku polskim samogłoski – choć wszystkie, w porównaniu ze spółgłoskami, są mniej lub bardziej otwarte, np.:

u : o : a (*buty : boty : baty*)
i : e : a (*malusi : malusie : malusia*)

(uszeregowane są w miarę zwiększania się stopnia otwarcia).

Wśród opozycji fonologicznych języka polskiego są również **opozycje równorzędne**. Przyjęto, że w takie opozycje wchodzi głoski różniące się od siebie **miejscem artykulacji**, np.

p : t : k (*pory : tory : kory*)
(wargowa : przedniojęzykowo-zębowa : tylnojęzykowa).

Ten ostatni przykład jest dobrą ilustracją konieczności pewnych rozstrzygnięć arbitralnych przy budowie każdego modelu, gdyż opozycja między głoskami p : t : k równie dobrze mogłaby być uznana za opozycję stopniową, z następującymi członami: p – przednie : t – środkowe : k – tylne (a więc: przednie : mniej przednie : najmniej przednie).

Każdy **fonem** można więc traktować jako **wiązkę cech dystynktywnych**, składającą się z tylu cech, w ile różnych opozycji wchodzi w danym systemie fonologicznym – oczywiście, niektóre z tych cech będą przyjmowały wartości zerowe. Tak więc fonem, mimo że jest najmniejszym wyrażeniem języka naturalnego stanowiącym odcinek tekstu linearnego, jest najmniejszym wyrażeniem złożonym. Ponieważ zasadniczą funkcją fonemu jest odróżnianie morfemów, fonemy natomiast różnią się od siebie cechami dystynktywnych, więc inwentarz cech dystynktywnych danego języka ustala się na podstawie porównywania morfemów najmniej się od siebie różniących (tylko jedną cechą). Spróbujmy ustalić takie opozycje dla polskiej głoski b, a tym samym ustalimy, z jakich cech dystynktywnych składa się w języku polskim fonem b.

- I. Opozycje prywatywne:
 1. bezdźwięczny : dźwięczny – p : b (*pył* : *był*),
 2. twardy : miękki – b : b' (*był* : *bił*),
 3. ustny : nosowy – b : m (*był* : *mył*);

- II. Opozycje stopniowe:

1. zwarty : szczelinowy – b : w (*był* : *wył*);

- III. Opozycje równorzędne:

1. wargowy : przedniojęzykowo-zębowy – b : d (*bał* : *dał*).

Gdybyśmy postępowali podobnie z innymi głoskami języka polskiego, to okazałoby się, że wszystkie dałoby się opisać za pomocą tych pięciu opozycji. A oto wszystkie wartości, jakie mogą one przyjmować:

- I. Opozycje prywatywne:

- udział wiązań głosowych:
 - + dźwięczne
 - bezdźwięczne
- udział podniebienia miękkiego:
 - + miękkie
 - twarde
- udział jamy nosowej:
 - + nosowe
 - ustne

- II. Opozycje stopniowe:

- stopień zbliżenia narządów mowy:
 - otwarte
 - półotwarte
 - szczelinowe
 - zwarto-wybuchowe
 - zwarto-szczelinowe

Wyróżniliśmy tu dwa rodzaje zwarcia, które w zależności od sposobu rozwiązania może być mniej lub bardziej gwałtowne i przechodzić w szczelinę lub w rozwarcie. W pierwszym wypadku mamy głoski zwarto-szczelinowe, w drugim zwarto-wybuchowe. Sposób rozwiązania zwarcia jest w języku polskim dystyngtywny, różni bowiem takie głoski, jak t : c (*matka* : *macka*).

- III. Opozycja równorzędna:

- miejsce artykulacji:
 - przednie (wargowe)
 - średnio-przednie (zębowe)
 - środkowe (dźwiękowe)
 - średnio-tylne (środkowo-językowe)
 - tylne (tylnojęzykowe)

W taki schemat wpisują się zarówno polskie fonemy spółgłoskowe:

- przednie = spółgłoski wargowe,
- średnio-przednie = spółgłoski zębowe,
- środkowe = spółgłoski dźwiękowe,

- średnio-tylne = spółgłoski środkowo-językowe,
- tylne = spółgłoski tylnojęzykowe,

jak i fonemy samogłoskowe:

- przednie i/ y
- średnio-przednie e, ɛ,
- środkowe a,
- średnio-tylne o, ɔ,
- tylne u.

Fonemy samogłoskowe, jak widzieliśmy wyżej, można także opisywać w opozycji stopniowej, wymagałoby to jednak znacznego zwiększenia liczby członów opozycji, co dawałoby model zbyt (niepotrzebnie) skomplikowany, byłoby więc nieekonomiczne. Opis, jaki zaproponowaliśmy tu, nie wymaga powiększenia liczby cech dystynktywnych, pozwalając równocześnie na wystarczające z punktu widzenia modelu różnicowanie polskich fonemów samogłoskowych.

W ten sposób uzyskaliśmy dla języka polskiego pięciowymiarową matrycę dystynktywnych cech fonemów, która teoretycznie daje możliwość utworzenia aż 200 fonemów ($2 \times 2 \times 2 \times 5 \times 5$). Jest to doskonała ilustracja ekonomii systemu dwuklasowego, gdzie przy niewielu jednostkach podstawowych można utworzyć bardzo wiele jednostek złożonych (w języku polskim tych jednostek elementarnych jest trzynaście), nawet przy ograniczeniu liczby elementów w każdej jednostce złożonej. W wypadku tworzenia fonemów ograniczenie oczywiście istnieje, gdyż cechy dystynktywne nie mogą się w jednym fonemie powtarzać ze względu na to, że łączą się ze sobą na zasadzie synchronicznej. Oczywiście 200 fonemów można by utworzyć jedynie wtedy, gdyby nie istniały tu specjalne prawa gramatyki (czyli składania cech dystynktywnych w fonemy), a takie prawa obowiązują w każdym języku, co od razu znacznie ogranicza liczbę faktycznie występujących w nim fonemów, nie wszystkie bowiem połączenia należą do danego języka. (Na marginesie należy zauważyć, że ponieważ cechy dystynktywne łączą się ze sobą na zasadzie współwystępowania w czasie, nie można do odróżniania jednostek złożonych używać permutacji elementów – gdyby nie to, możliwość wszystkich połączeń byłaby jeszcze znacznie większa). Pewne (znaczące) ograniczenia w liczbie możliwych połączeń narzuca również sama budowa narządów mowy człowieka, a więc urządzenia realizującego model w konkretnym materiale fizycznym. I tak, na przykład, znacznie więcej jest fonemów przednich niż tylnych, bo najbardziej ruchome części narządów mowy położone są w przedniej części jamy ustnej (wargi i język), w tyle jamy ustnej jest o wiele mniej rozróżnień, bo i mniej możliwości ich realizacji, tył języka jest bowiem mało ruchliwy. Wszystko to sprawia, że liczba fonemów faktycznie występujących w danym języku jest znacznie mniejsza niż by to wynikało z reguł gramatyki poziomu fonologicznego.

Prawa gramatyczne składania cech dystynktywnych są różne dla różnych języków naturalnych, tak jak różny jest zestaw samych cech. Niżej podamy kilka podstawowych reguł gramatyki tego poziomu dla języka polskiego. Reguły te spróbujemy odtworzyć, odpowiadając na następujące pytania:

- Jaka jest najmniejsza liczba cech dystynktywnych w jednym fonemie?

- Jaka jest największa liczba cech dystynktywnych w jednym fonemie?
- Czy wśród cech dystynktywnych są takie, które muszą wystąpić w fonemie (są obligatoryjne)?

Jeśli odpowiedź na to pytanie będzie twierdząca, to kolejne zadanie będzie polegało na podziale cech dystynktywnych na kategorię (syntaktyczną) obligatoryjnych i kategorię (syntaktyczną) fakultatywnych fonologicznych cech dystynktywnych.

- Czy określona jest liczba cech dystynktywnych jednej kategorii w jednym fonemie – ile, a jeżeli więcej niż jedna, to które?
- Czy, a jeśli tak, to jakie obowiązują prawa korelacji, czyli czy wystąpienie w fonemie jakiejś jednej (której lub z jakiej kategorii?) cechy dystynktywnej pociąga za sobą wystąpienie jakiejś innej (której lub z jakiej kategorii?) cechy dystynktywnej?

Przypominamy, że chodzi tu tylko o zależność jednostronną, typu jeżeli A to B, która nie pociąga za sobą zależności jeżeli B to A. W modelu w wypadku zależności dwustronnej takie skorelowane na stałe elementy traktuje się jak jeden element.

Reguły gramatyki odtwarzające budowę fonemów języka polskiego przedstawimy w sposób schematyczny, przedtem jednak wprowadzimy umowne oznaczenia wyrażeń elementarnych tego poziomu języka, czyli cech dystynktywnych języka polskiego, tworząc w ten sposób:

Słownik:

△ dźwięczność	}	kategorie fakultatywne	}	kategorie obligatoryjne
○ miękkość				
~ nosowość				
□ otwartość	}	kategoria syntaktyczna □ stopnia otwarcia		
□ półotwartość				
□ szczelinowość				
□ zwartoszczelinowość				
□ zwartowybuchowość	}	kategoria syntaktyczna • miejsca artykulacji		
• przedniość				
: średniopredniość				
•• środkowość				
::: średniotylność				
:::• tylność				

Dla uproszczenia od razu dokonaliśmy tu kategoryzacji syntaktycznej słownika, a więc podziału cech dystynktywnych fonemów języka polskiego na kategorie syntaktyczne.

Aby odtworzyć reguły **gramatyki poziomu fonologicznego** języka polskiego należy ustalić wszystkie opozycje fonologiczne, jakie tworzą wszystkie fonemy, tak ja to zrobiliśmy dla fonemu b. Wynika z tego, że gramatyka składania fonologicznych cech dystynktywnych w fonemy języka polskiego zawiera następujące reguły:

- Najprostszy fonem składa się z dwóch cech dystynktywnych, po jednej z kategorii stopień otwarcia i kategorii miejsce artykulacji.

- Najbardziej złożony fonem składa się z pięciu cech dystynktywnych, przy czym każda cecha reprezentuje inną opozycję.

- W fonemie zawsze występuje jedna cecha dystynktywna z opozycji stopień otwarcia i jedna cecha dystynktywna z opozycji miejsce artykulacji.

Z powyższego możemy wyciągnąć wnioski, że:

- Każda opozycja fonologiczna może być reprezentowana w fonemie tylko przez jedną cechę,

- Obligatoryjnymi kategoriami syntaktycznymi polskiego systemu fonologicznego są opozycje miejsce artykulacji i stopień otwarcia. Pozostałe opozycje są kategoriami fakultatywnymi. Tworzą je nacechowane człony opozycji prywatywnych dźwięczność : bezdźwięczność, miękkość : twardość, nosowość : ustność. Są to więc kategorie jednoelementowe.

- W fonemie dana cecha dystynktywna może wystąpić tylko jeden raz.

- Jediną relacją syntagmatyczną zachodzącą pomiędzy wyrażeniami elementarnymi tworzącymi zdanie systemu fonologicznego, czyli pomiędzy fonologicznymi cechami dystynktywnymi w fonemie jest relacja współwystępowania – elementy tworzące takie zdanie (tu: fonem) nie występują w żadnym ustalonym porządku ani nie tworzą żadnej innej struktury z ustalonym miejscem w tej strukturze. Dlatego fonemy są określane jako wiązka współwystępujących w czasie cech dystynktywnych, a jednocześnie jako elementarne wyrażenie linearnej struktury języka pełniące funkcję dystynktywną, czyli odróżniającą (lub konstytuującą) znaczenie.

- W zbiorze fonemów rozpatrywanych jako struktury złożone występują też następujące korelacje:

- fonemy otwarte są zawsze dźwięczne,
- fonemy półotwarte są zawsze dźwięczne,
- fonemy nosowe są zawsze dźwięczne,
- fonemy nosowe są zawsze półotwarte lub otwarte,
- fonemy środkowo-językowe są zawsze miękkie,
- fonemy szczelinowe są zawsze ustne,
- fonemy zwarto-szczelinowe są zawsze ustne.

Zapiszemy je w postaci sformalizowanej, używając oznaczeń:

\Rightarrow – implikacja (jeżeli to),

\vee – alternatywa rozłączna (jedno z dwojga),

\wedge – koniunkcja (i),

() – nawias oznacza kolejność działania funktorów.

Reguły:

- $\sim \Rightarrow \Delta$ (fonemy nosowe są wszystkie dźwięczne),
- $\sim \Rightarrow \square \vee \square$ (nosowymi są tylko fonemy otwarte lub półotwarte),
- $\square \Rightarrow \Delta$ (fonemy otwarte są zawsze dźwięczne),

$\square \Rightarrow \Delta$ (fonemy otwarte są zawsze dźwięczne).

Te dwie reguły można zapisać łącznie z pierwszą regułą: $\sim \vee \square \vee \square \Rightarrow \Delta$


































$(\square \vee \square) \wedge :: \Rightarrow \circ$ (fonemy szczelinowe lub zwartoszczelinowe średnio-tylne (średniojęzykowe) są zawsze miękkie),

$:: \Rightarrow \square \vee \square \vee \square$ (fonemy średniotylnie są albo otwarte, albo szczelinowe, albo zwartoszczelinowe).

Te dwie ostatnie reguły można zapisać w odpowiednio większej liczbie reguł prostych.

Nie są to jeszcze wszystkie reguły gramatyki obowiązujące na tym poziomie języka. Zapisywaliśmy je w kilku wypadkach celowo w różnych postaciach, żeby pokazać, że te same reguły gramatyki mogą się różnić sposobem sformułowania.

Zgodnie z powyższymi regułami składniowymi utworzyć można na przykład następujące zdania-fonemy:

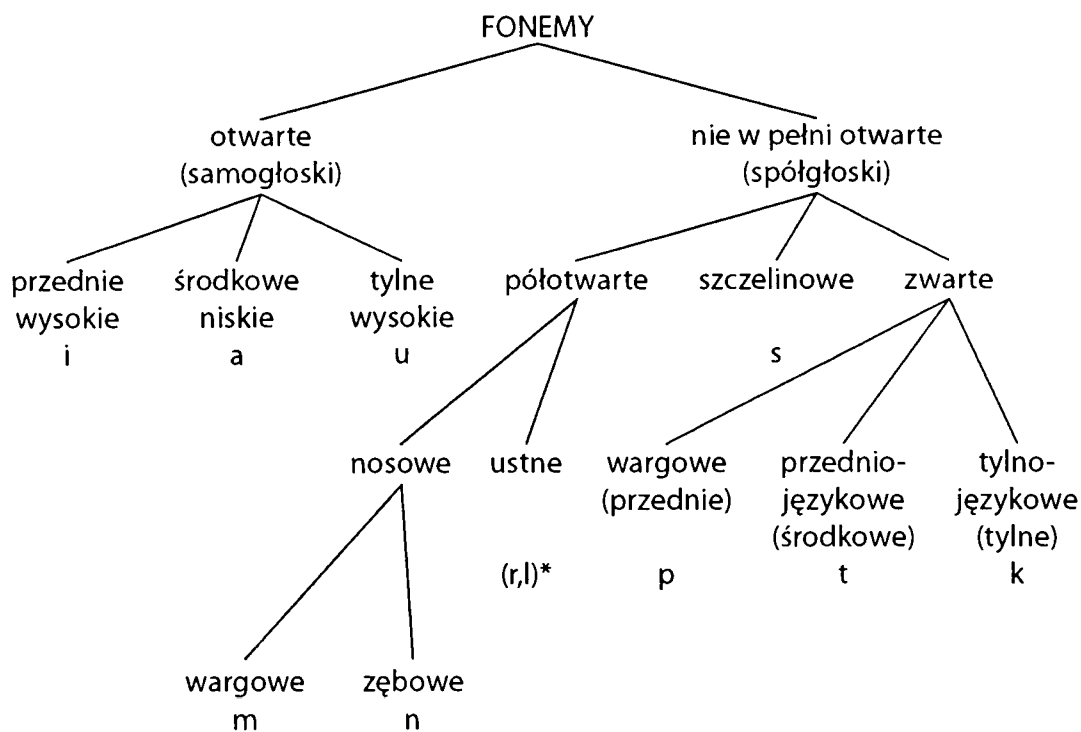
								
f	p	s	c	t	ś	ć	x	k
							(ch)	
								
w	b	z	3	d	ż	ź		g
			(dz)			(dź)		
								
	p´	ś	ć					
								
w´	b´	z´						
								
m	n		i/y	e	a	o	u	
								
m´	n´			ę		ø		

Prawa gramatyczne składania cech dystynktywnych znacznie redukują liczbę fonemów, mimo to i tak bardzo wiele możliwych, zgodnych z prawami składni fonemów nie jest faktycznie w języku realizowanych, stanowią one formy potencjalne, które są pewnego rodzaju „żelaznym zapasem” języka i mogą się aktualizować w miarę potrzeby. Takim **fonemami potencjalnymi** są na przykład w języku polskim r´ (miękki odpowiednik r) i t´ (miękki odpowiednik t), które aktualizują się w jednej tylko parze wyrazów *trik* i *tryk*, *plastik* i *plastyk* (i oraz y

stanowią warianty kombinatoryczne jednego fonemu, uzależnione od sąsiedztwa innych fonemów: y może występować tylko po spółgłoskach twardych, natomiast i po miękkich oraz w nagłosie (na początku wyrazu). Schemat systemu fonologicznego języka polskiego przedstawiamy na oddzielnym wykresie.

Każdy język naturalny ma własny **system fonologiczny**, czyli zbiór opozycji fonologicznych, tworzących je cech dystynktywnych i reguł ich składania w fonemy, a więc i własny zbiór fonemów. Są języki wykorzystujące opozycje fonemów słabych akustycznie i mocnych (np. język niemiecki), fonemów różniących się długością trwania, fonemów różniących się wysokością tonu samogłosek, np. niski, średni, wysoki (np. język ibo, hausa w Afryce) lub fonemów różniących się zmianami tonu, np. ton wznoszący się (rosnący), równy, opadający, wznosząco-opadający, opadająco-wznoszący (np. chiński).

Chociaż każdy język ma swój własny system fonologiczny, to jednak pewne opozycje fonologiczne powtarzają się prawie we wszystkich językach świata. Są to oczywiście opozycje najlepsze z punktu widzenia dobrego języka, najwyraźniejsze, a więc najlepiej przystosowane zarówno do nadawcy (oparte na najbardziej się od siebie różniących cechach artykulacyjnych), jak i do odbiorcy komunikatu (dające w wyniku artykulacji głoski najbardziej się różniące pod względem akustycznym). Warto się zapoznać z takim **prymarnym systemem fonologicznym**, uwzględniającym opozycje ogólnoludzkie. System ten podaję, z pewnymi modyfikacjami, za Tadeuszem Milewskim (*Językoznawstwo ogólne*).



* najczęściej jeden fonem różnie realizowany.

Jak widzimy, system ten opiera się na przeciwieństwie najwyrazistszych cech akustycznych i artykulacyjnych. Obejmuje on 10 fonemów, w tym 3 samogłoskowe i 7 spółgłoskowych. Język zbudowany tylko z fonemów systemu prymarnego nie byłby jednak językiem najbardziej ekonomicznym, gdyż teksty takiego języka musiałyby być stosunkowo długie. W języku obowiązuje bowiem zależność (odwrotnie proporcjonalna) pomiędzy liczbą jednostek elementarnych w systemie a liczbą jednostek elementarnych w jednostkach złożonych, czyli mówiąc inaczej: im mniej fonemów w systemie, tym więcej musi być fonemów w morfemie, aby uzyskać w rezultacie wystarczająco różniące się od siebie i wystarczająco liczne jednostki złożone. Systemy fonologiczne zbyt ubogie (13-20 fonemów) nie są więc ekonomiczne. Nie są również ekonomiczne systemy zbyt bogate (45-75 fonemów), bo przyjęcie dużej liczby jednostek elementarnych powoduje, iż różnią się one od siebie cechami zbyt mało wyrazistymi, są więc trudne zarówno dla nadawcy, jak i dla odbiorcy. Najlepsze są więc systemy pośrednie, zawierające 20-45 fonemów, gdyż różnice pomiędzy poszczególnymi fonemami mogą być jeszcze wystarczająco wyraziste i dają ciągi średniej długości o wystarczającej redundancji.

Z tego wynika, że w języku obowiązuje prawidłowość polegająca na tym, że zbyt duża oszczędność w jednym miejscu (na jednym poziomie) pociąga za sobą zwiększenie wysiłku w jakimś innym miejscu, dlatego, konstruując języki sztuczne, szczególnie rozważnie należy określać optymalną liczbę elementarnych jednostek systemu.

Funkcjonowanie tej zasady najlepiej widać przy porównaniu długości zapisów tych samych liczb w różnych systemach, np. dwójkowym (którego słownik liczy tylko dwie jednostki elementarne: 0 i 1), dziesiętnym, którego słownik zawiera dziesięć jednostek elementarnych) oraz szesnastkowym (którego słownik liczy szesnaście jednostek elementarnych); zgodnie z tą zasadą system dwójkowy daje napisy najdłuższe, natomiast system szesnastkowy najkrótsze. A oto przykład:

Liczba	System dwójkowy	System dziesiętny	System szesnastkowy
cztery	100	4	4
dwanaście	1100	12	C
trzydzieści siedem	100101	37	25
sto dziewięćdziesiąt osiem	11000110	198	C6
trzyście tysięcy osiemset dwa	11010111101010	13802	35EA

Omówioną tu zależność można więc w przybliżeniu określić wzorem:

$$\text{liczba jednostek w systemie} \times \text{liczba jednostek w zdaniu} = \text{constans}$$

Dla ilustracji powyższej zasady podamy jeszcze przykład zdania ze stosunkowo bogatego języka naturalnego (około 45 fonemów). Jest to język czynuk z Północnej Ameryki, w którym na jeden morfem przypada przeciętnie 1,25 fonemu, podczas gdy na przykład w języku ubogim fonologicznie, jakim jest aranta (Hawaje), który ma jedynie 13 fonemów, na jeden morfem przypada 4 fonemy.

A oto przykład z języka czynuk. Zdanie równoważne pod względem ilości informacji polskiemu zdaniu: 'Tyś miał w zwyczajcu brać to od niej':

a – m – L – a – x – cg – ām' – x

(tekst zapisany w rozbiciu na morfemy), gdzie:

a – wykładnik czasu przeszłego dokonanego,

m – ty (odbiorca informacji),

L – (wymawiane tl, ale stanowiące jeden fonem) – to,

a – jej,

x – (wymawiane ch) – wykładnik przynależności; to należy do niej,

cg – brać (rdzeń formy),

am' – wykładnik trybu oznajmującego,

x – (ch) – wykładnik tego, że czynność jest zwyczajem.

W zdaniu tym 8 morfemów oddane jest przez 10 fonemów, podczas gdy w odpowiadającym mu pod względem zawartości informacji zdaniu polskim jest aż 26 fonemów:

tyś miał w zwyczajcu brać to od niej

(liczba morfemów w zdaniu polskim jest nieco większa, gdyż niektóre z nich są redundantne, np. *ty* i *-ś*).

Fonemy są najmniejszymi jednostkami realizowanymi w ciągu linearnym w tekście. Nie mają znaczenia, lecz służą do odróżniania najmniejszych jednostek znaczących – morfemów. **Łączenie fonemów w morfemy** odbywa się zgodnie z prawami składniowymi danego systemu fonologicznego. I tak, na przykład, w języku polskim obowiązują, między innymi, następujące reguły ich składania:

- Fonemy spółgłoskowe szczelinowe, zwarte i zwarto-szczelinowe można łączyć ze sobą tylko wtedy, jeśli są jednorodne pod względem dźwięczności (tzn. tylko dźwięczne z dźwięcznymi lub tylko bezdźwięczne z bezdźwięcznymi), np.:

stać, wzdłużny (wzdłużny), uafka (ławka), kštuśić (krztusić)

(jak widzimy pisownia polska nie jest pod względem fonologicznym konsekwentna).

- Jeżeli wyraz kończy się fonemem spółgłoskowym szczelinowym, zwartym lub zwarto-szczelinowym, to może to być tylko fonem bezdźwięczny, np.:

śnek (śnieg), vus (wóz), śfat (świat)

(i tu pisownia polska nie jest konsekwentna).

- W jednym morfemie nie występują obok siebie dwa fonemy samogłoskowe.

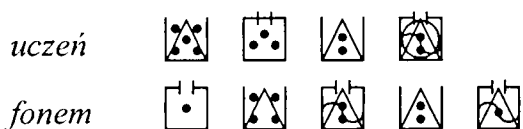
- Morfem może być utworzony przez jeden tylko fonem. Na przykład przyimki *z, w*, spójniki *i, a*, przedrostki *z-, w-, o-*, końcówki fleksyjne *-i, -a, -u, -ą, -ę*.

Reguły składniowe poziomu fonologicznego muszą także uwzględniać kontekst, czyli sąsiedztwo danego fonemu w ciągu fonemów tworzących morfem lub wyraz – są więc regułami gramatyki zależnej od kontekstu. Dotyczą one tzw. **allofonów**, czyli wariantów tego samego fonemu, których wybór bywa uwarunkowany kontekstem. W języku polskim takimi wariantami są na przykład *n* (przedniojęzykowo-zębowe), występujące w takich wyrazach, jak np. *nauka, noga, fonetyka, n[ɔ]dziąsłowe*), występujące w takich wyrazach, jak *Andrzej, mądry*, a więc tylko przed fonemem dźwiękowym *i* *ɲ* (tylnojęzykowe) występujące na przykład w wyrazie *mąka*, a więc przed fonemem tylnojęzykowym. Są to **warianty kombinatoryczne**, ich użycie uzależnione jest bowiem sąsiedztwem fonemów następujących. Jednostki te nie stanowią oddzielnych fonemów, gdyż nie pełnią funkcji odróżniającej (dystynktywnej), nie ma bowiem w języku polskim takich morfemów, których znaczenie różniłoby się tylko w zależności od tego, czy użyte zostanie w nich *n, n[ɔ]* czy *ɲ*. Wariantami na poziomie fonologicznym są również w języku polskim *l* (przedniojęzykowe) i *ɫ* (tylnojęzykowe, wymawiane przez większość Polaków), są to jednak **warianty fakultatywne** i mogą stanowić przedmiot arbitralnego wyboru (choć zazwyczaj wybór uwarunkowany jest nawykami fonetycznymi indywidualnego nadawcy, a te związane są z wymową uwarunkowaną terytorialnie).

W każdym systemie ważne są tylko opozycje, a nie materiał, w którym są one realizowane, dlatego wszystkie teksty języka polskiego można by przekodować, używając do tego zupełnie innego materiału niż fala głosowa.

Zaprezentowany tu model jest modelem wyłącznie polskiego języka mówionego – dla języka pisanego należałoby zbudować model uwzględniający dystynktywne cechy liter w oparciu o ich funkcję odróżniającą w obrębie morfemu. Model fonologiczny polskiego języka mówionego i model „grafologiczny” polskiego języka pisanego różniłyby się znacznie, ponieważ w pisowni polskiej nie ma jedno-jednoznacznego przyporządkowania pomiędzy fonemem i literą.

A oto te dwa wyrazy w zaprojektowanej poprzednio reprezentacji polskiego systemu fonologicznego:



System fonologiczny jest modelem fonetycznej (dźwiękowej) materii języka naturalnego, ujmującym tylko te cechy głosek, które w danym języku naturalnym służą do odróżniania znaczenia (w tym sensie pełnią funkcję semantyczną), czyli do kodowania informacji. Taki model jest więc systemem abstrakcyjnym, niezależnym od realizacji w konkretnym materiale. Jego realizację w materiale akustycznym przy użyciu narządów mowy człowieka stanowią głoski tworzące teksty mówione języka (teksty subkodu akustycznego). Głoski różnią się od fonemów tym, że zawierają również takie cechy, które nie są wykorzystywane w danym

systemie językowym do odróżniania jednostek znaczących, nie biorą więc udziału w kodowaniu informacji. W konkretnej realizacji tekstowej głoski nie zawsze realizują wszystkie cechy fonologiczne (cechy dystynktywne odpowiednich fonemów), konkretna realizacja zawsze jest bowiem uwarunkowana sąsiedztwem innych głosek (kontekstem), tempem mowy i warunkami artykulacyjnymi mówiącego. Przed stratą informacji w takim wypadku częściowo chroni redundancja wbudowana w system fonologiczny języka.

O systemie fonologicznym mówimy tylko w wypadku języka naturalnego, sam termin fonem jest bowiem związany z foniczną warstwą języków naturalnych. Dział językoznawstwa opisujący systemy fonologiczne języków naturalnych nazywa się **fonologią**, rzadziej **fonematyką** lub **fonemiką**. Modele tego podsystemu języka naturalnego w ujęciu różnych szkół językoznawców różnią się założeniami – tu zaprezentowany oparty jest na założeniach wypracowanych przez tzw. praską szkołę strukturalistyczną, a przede wszystkim przez Nikołaja Trubieckiego i Romana Jakobsona. Model ten nie uwzględnia całej substancji akustycznej języków naturalnych, nie ujmuje bowiem ani akcentu wyrazowego w jego funkcji delimitacyjnej, ani tzw. **suprasegmentalnych wykładników akustycznych** (nakładających się na całe segmenty, czyli określone ciągi linearne wypowiedzenia), stanowiących o tzw. **prozodii** języka naturalnego, określających przebieg melodyczny fraz i zdań (**intonacja**), również pełniących funkcje semantyczne (np. w języku polskim tylko intonacja zdaniowa może odróżniać zdania pytające od oznajmujących, lub rozkazujących, por. mówioną wersję zdań *Idziemy do kina. Idziemy do kina?* *Idziemy do kina!*).

Dwuklasowe języki sztuczne mogą mieć analogiczną strukturę, jeśli posiadają dwie klasy jednostek pustych semantycznie, nie muszą być jednak realizowane w materiale dźwiękowym, dlatego w ich wypadku nie mówi się o systemie fonologicznym. Języki informacyjno-wyszukiwawcze nie mają własnego systemu fonologicznego, mimo że ich wyrażenia (teksty) bywają realizowane w subkodzie akustycznym, zawsze jednak jest to subkod akustyczny odpowiedniego języka naturalnego. W wypadku języków informacyjno-wyszukiwawczych o notacji sztucznej są to zazwyczaj nazwy cyfr i innych znaków graficznych, a w wypadku tak zwanych paranaturalnych języków informacyjno-wyszukiwawczych są to odpowiadające (równokształtne w subkodzie graficznym) wyrażeniom tych języków wyrażenia sub kodu mówionego odpowiedniego języka naturalnego (na przykład słowa kluczowe lub deskryptory równokształtne z odpowiednimi wyrażeniami języka polskiego są czytane tak, jakby były wyrazami języka polskiego).

Literatura

Trubiecki N.S.: *Podstawy fonologii*. Warszawa 1970.

Klebanowska B.: *Interpretacja fonologiczna zjawisk fonetycznych w języku polskim*. Warszawa 1987.

Ostaszewska D., Tambor J.: *Fonetyka i fonologia współczesnego języka polskiego*. Warszawa 2000.

Wierzchowska B.: *Fonetyka i fonologia języka polskiego*. Warszawa 1980.

VII

System morfologiczny

Najmniejszą jednostką znaczącą języka naturalnego, jednostką, której przyporządkowany jest jakiś wycinek rzeczywistości, jest morfem. Morfemy łączą się ze sobą w wypowiedzeniu w większe jednostki tekstu, dając w rezultacie zdania mające już wartość informacyjną. W niektórych językach jednak, a do takich należy również język polski, morfemy bardzo często łączą się ściśle w jednostki mniejsze niż zdania, różniące się od nich funkcją – w wyrazy. Na czym więc polega różnica między morfemem, wyrazem a zdaniem?

Określenie różnicy między tymi trzema jednostkami zakłada, że potrafimy dokładnie określić, czym są morfemy, wyrazy i zdania, a tymczasem jak dotąd nie ma wśród językoznawców zgody co do definicji wyrazu i zdania. Jednostką dobrze zdefiniowaną jest jedynie morfem: **morfem** jest to najmniejsza jednostka planu wyrażania języka, której przyporządkowana jest reguła interpretacji (najmniejsza jednostka znacząca).

Morfemy dzieli się na kilka różnych kategorii.

I tak, ze względu na sposób **przyporządkowania znaczenia**, wyróżnia się:

- **morfemy leksykalne**, czyli **rdzenie** wyrazów, reprezentujące znaczenie indywidualne, np.: *dom-*, *las-*, *woz-*, *bieg-*, *książ-*, *raz-*, *leksyk-*, *język-*,
- **morfemy gramatyczne**, czyli stałe wykładniki kategorii językowych. W języku naturalnym do tej kategorii należą morfemy będące **wykładnikami relacji syntagmatycznych**, np. morfemy **aglutynacyjne** i **fleksyjne** (końcówki), np.: *-i*, *-a*, *-owie-* *-ą*, *-my*, *-cie*, *-m*, *-asz* (omówimy je dokładnie w rozdziale poświęconym relacjom syntagmatycznym) oraz **morfemy słowotwórcze**, tzw. **afiksy**, czyli morfemy, które w połączeniu z morfemami leksykalnymi tworzą nowe wyrazy.

Afiksy dzieli się na kategorie syntaktyczne zależnie od ich usytuowania w wyrazie względem morfemu leksykalnego i innych morfemów gramatycznych. I tak wyróżnia się:

- **prefiksy** czyli **przedrostki** – morfemy, które znajdują się w **prepozycji** (przed rdzeniem), np.: *za-* w *zapisywać*, *przy-* w *przyjechać*, *wy-* w *wybiec*, *naj-* w *najładniejszy*;
- **sufiksy**, czyli **przyrostki** – morfemy, które znajdują się w **postpozycji** (po rdzeniu), np.: *-arz*, w *pisarz*, *-acz* w *pakowacz*, *-ać* w *pisać*;

- **infiksy**, czyli **wrostki** – morfemy znajdujące się w **impozycji**, wstawiane pomiędzy fonemy tworzące rdzeń, np.: w łacinie *-n-* w wyrazie *vinco* (zwyciężam) w opozycji do *vic* w *victus* (zwyciężony) oraz morfemy gramatyczne w językach alternacyjnych (omówimy je w rozdziale o relacjach syntagmatycznych);

- **postfiksy** – które stoją w wyrazie po końcówce fleksyjnej, np. *-ś* w wyrazach *ktoś*, *kogoś*;

- **interfiksy** – morfemy stojące między członami wyrazu złożonego, np. *-o-* w wyrazach *żywopłot*, *północnowschodni*, *listonosz*.

Ze względu na **stopień samodzielności** morfemów gramatycznych w tekście dzieli się je na:

- **morfemy wolne** – niełączące się w całość z morfemem leksykalnym – takimi morfemami są na przykład w języku polskim przyimki;

- **morfemy związane** – łączące się w jeden wyraz z morfemami leksykalnymi – do tej klasy należą afiksy i końcówki fleksyjne;

Ze względu na **stopień spójności** fonemów tworzących morfem wyróżnia się:

- **morfemy ciągłe**, czyli takie, że między tworzące je fonemy nie można wstawić żadnego innego morfemu;

- **morfemy nieciągłe**, które charakteryzują się tym, że między tworzące je elementy można wstawić inny morfem. Charakter nieciągły mają te morfemy, w których środek można wstawić infiksy (problem omówimy dokładniej przy omawianiu relacji syntagmatycznych).

Jak już powiedzieliśmy, nie wszyscy teoretycy języków informacyjno-wyszukiwawczych wyróżniają w tych językach poziom morfologiczny, wydaje się jednak, że w językach tych można przyjąć istnienie prawie wszystkich powyższych typów morfemów. I tak morfemami leksykalnymi są np. deskryptory, symbole główne UKD, tzw. czynniki semantyczne z kodu semantycznego Perry'ego i Kenta. Morfemami gramatycznymi są na przykład wskaźniki więzi i wskaźniki roli języków deskryptorowych, symbole poddziałów wspólnych UKD, wykładniki relacji analitycznych kodu semantycznego Perry'ego i Kenta.

Wśród morfemów gramatycznych tych języków, podobnie jak w języku naturalnym, można wyróżnić prefiksy, np. wskaźniki roli w kodzie semantycznym Perry'ego i Kenta, takie jak: KEJ – afiks oznaczający materiał przetwarzany, poddany ponownemu procesowi, obecny np. w wyrażeniu KEJ.SAHP.003, sufiksy, np. wskaźniki więzi w językach deskryptorowych, np. 1, 2 w wyrażeniach MIEDŹ₁, OŁÓW₂, POKRYCIE₂, RURY₁, infiksy, np. wykładniki tzw. relacji analitycznych: A, E, I, O, U, Q, W, X, Y, Z w kodzie semantycznym Perry'ego i Kenta, np. w wyrażeniu PAST.004 – mucha, BITN.013 – kwiat.

Morfemami wolnymi są w językach informacyjno-wyszukiwawczych deskryptory, symbole główne UKD i określniki języka haseł przedmiotowych. Morfemami związanymi są np. wskaźniki roli i wskaźniki więzi języków deskryptorowych, symbole poddziałów analitycznych UKD, wszystkie morfemy kodu semantycznego Perry'ego i Kenta.

W językach informacyjno-wyszukiwawczych istnieją także morfemy nieciągłe, są nimi na przykład morfemy leksykalne w kodzie semantycznym Perry'ego

i Kenta (np.: P-AST oznaczający klasę dokumentów mówiących o owadach, czy B-TN oznaczający klasę dokumentów mówiących o roślinach), a także pewne morfemy UKD: [...] (nawias kwadratowy) i (...) (nawias okrągły), np.

364.4:550.348](498):[061.1(100):61WHO] Pomoc Międzynarodowej Organizacji Zdrowia dla ofiar trzęsienia ziemi w Rumunii

Jak już powiedzieliśmy, dokładnej definicji ani wyrazu, ani zdania nie podamy, spróbujmy jednak określić najważniejsze między nimi różnice. Najważniejsza różnica między morfemem a wyrazem polega na tym, że morfem składa się tylko z jednostek niemających swojego własnego znaczenia (z fonemów), których skład jest w morfemie zupełnie niezależny od tego, co dany morfem oznacza (jest przypadkowy), natomiast **wyraz** może się składać z jednostek znaczących, czyli morfemów, przy czym może się składać tylko z jednego morfemu, np. *wóz*, *dom*, *las*, *pieśń*, *nos* (są to tzw. **wyrazy rdzenne**) i wtedy równa się morfemowi, może jednak składać się z większej liczby morfemów, np. *wozu* (*woz*+*u*), *domu* (*dom*+*u*), *lasu* (*las*+*u*), *nosa* (*nos*+*a*), gdzie morfemy *-u*, *-a*, będące końcówkami fleksyjnymi, wskazują, że wyrazy te są formami dopełniacza liczby pojedynczej, lub *wózek* (*wóz*+*ek*), *domek* (*dom*+*ek*), *lasek* (*las*+*ek*), *nosek* (*nos*+*ek*), gdzie morfem słowotwórczy *-ek* wskazuje, że cały wyraz oznacza mały przedmiot: *wózek* to mały wóz, *domek* – mały dom, *lasek* – mały las, *nosek* – mały nos.

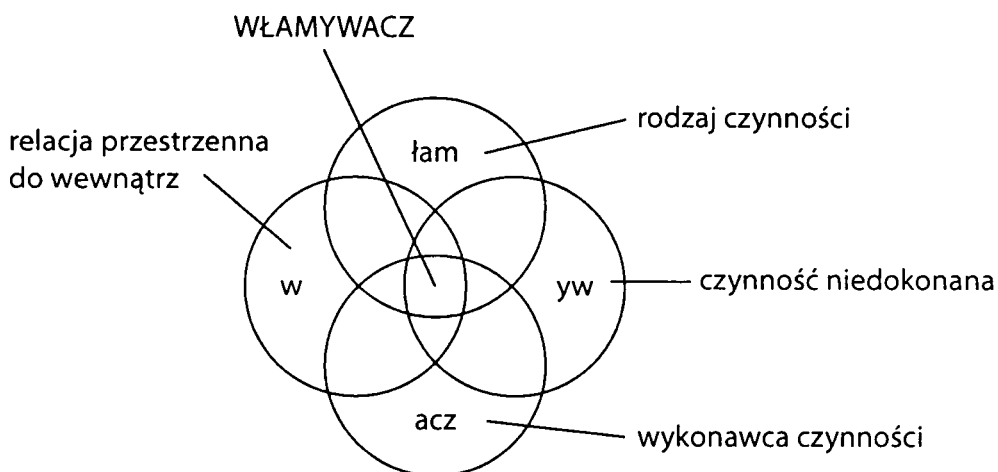
Wyrazy podane w tym ostatnim przykładzie tworzą jedną wspólną klasę – klasę przedmiotów małych, której wykładnikiem jest właśnie powtarzający się w nich morfem słowotwórczy *-ek*. Taką klasę, mającą swój wykładnik w postaci morfemu gramatycznego, nazywamy **kategorią semantyczno-gramatyczną**.

A oto inny przykład takiej klasy: *ladowacz*, *pakowacz*, *krajacz*, *rozdzielacz*, *włamywacz*, *kopacz*, które to wyrazy tworzą klasę semantyczno-gramatyczną nazw wykonawców czynności, której wykładnikiem jest formant słowotwórczy *-acz*, a sama czynność oznaczana jest przez podstawę słowotwórczą będącą częścią czasownika, od którego dany wyraz został utworzony.

Czasowniki będące podstawą dla utworzenia nowego wyrazu często składają się również z wielu morfemów, np. *włamywać się* (*w*+*lam*+*yw*+*ać*+*się*), gdzie rodzaj czynności wskazany jest przez morfem będący rdzeniem całej struktury *-lam-* (powtarza się on również w takich wyrazach jak: *lamać*, *rozłamać*, *przełamać*, *złamanie*, *lamy*, tworzących rodzinę wyrazów), przedrostek *w-* wskazuje, że czynność jest skierowana do środka jakiegoś odgraniczonego miejsca – w takiej samej funkcji przedrostek ten występuje na przykład w czasownikach: *wchodzić*, *wnosić*, *wpisywać*, *wsypać*, morfem *-yw-* wskazuje, że czynność jest niedokonana (por. *wpisywać*, *nagrywać*, *odczytywać*), natomiast przyrostek *-ać* jest wykładnikiem bezokolicznika.

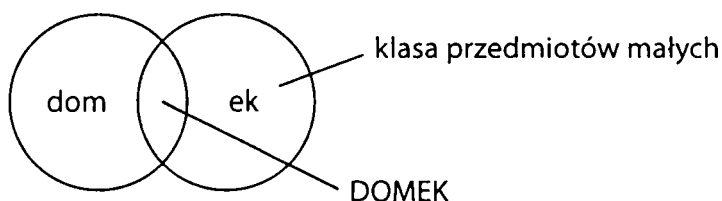
W tym wypadku jeden wyraz składał się już z kilku morfemów, przy czym skład ich w wyrazie nie był zupełnie przypadkowy, lecz uwarunkowany tym, do jakiej klasy semantycznej dany wyraz należy. Zgodnie z powyższą analizą, wyraz *włamywacz* należy do klasy nazw wykonawców czynności, klasy wyrazów nazywających wszystko, co jest związane z czynnością łamania, klasy wyrazów

nazywających relację przestrzenną „do wewnątrz” oraz klasy wyrazów oznaczających czynności niedokonane. Graficznie możemy to przedstawić w następujący sposób:



W wypadku wyrazów składających się z więcej niż jednego morfemu – morfemu leksykalnego i morfemów słowotwórczych – skład morfemowy wyrazu nie jest przypadkowy, ale, jak widzieliśmy, uwarunkowany względami semantycznymi, zaś samo składanie poszczególnych morfemów w wyraz, a często także wybór jednego z wielu morfemów synonimicznych, uwarunkowane są prawami składania morfemów (gramatyki) obowiązującymi na danym poziomie języka – reguły te nazywamy **regułami słowotwórczymi**. One to decydują również o kolejności, morfemów w wyrazie i nakładają pewne ograniczenia na łączące się ze sobą morfemy. I tak, na przykład, za pomocą formantu *-acz* można tworzyć w języku polskim nazwy wykonawców czynności tylko od czasowników zakończonych na *-ać* (jedynym chyba wyjątkiem jest tu wyraz *palacz*, utworzony od czasownika *pałiĆ* zakończonego na *-ić*), budowa wyrazu podstawowego decydowała więc o wyborze z dwóch formantów synonimicznych oznaczających nazwy wykonawców czynności: *-acz* i *-iciel/-yiciel* właśnie formantu *-acz*.

Wyraz *domek* można graficznie przedstawić następująco:



Podobnie jest w wypadku wyrazów złożonych z morfemu leksykalnego i morfemu będącego wykładnikiem relacji syntagmatycznej, gdzie o wyborze tego ostatniego decyduje rodzaj relacji, w jakich dany wyraz pozostaje z innymi wyrazami tworzącymi wyrażenie złożone wyższego poziomu struktury języka, a więc syntagmę lub zdanie.

Wyraz, podobnie jak morfem, nie niesie żadnej pełnej informacji o rzeczywistości, nazywa tylko pewne jej elementy, ale często, wtedy gdy składa się z kilku morfemów, robi to inaczej niż morfem, pokazuje przy okazji pewne relacje pomiędzy nim a innymi wyrazami, niesie więc informacje o strukturze słownika, a tym samym o pewnych relacjach, w jakie może wchodzić desygnat tego wyrazu z innymi elementami rzeczywistości. Na przykład, ten sam przedmiot może być w języku polskim oznaczony przez zawierający tylko jeden morfem leksykalny niemotywowany wyraz *pipeta* i przez wielomorfemowy wyraz *kropłomierz*. O tym, co oznacza wyraz *pipeta*, trzeba po prostu wiedzieć, tego, co oznacza wyraz *kropłomierz*, można się domyślić na podstawie jego struktury (znaczenie strukturalne), a więc znaczenia morfemów *kropl-* i *-mierz* oraz znajomości takich wyrazów, jak *światłomierz*, *kątomierz*, *amperomierz*, *wodomierz*. Wyraz *kropłomierz* ma więc wpisaną przynależność do klasy nazw przedmiotów będących przyrządami pomiarowymi, do której należą takie wyrazy, jak: *amperomierz*, *woltomierz*, *kątomierz*, *dalmierz*, *światłomierz*, *szybkościomierz*, *wysokościomierz*, *ciśnieniomierz*, czego nie ma w wyrazie *pipeta*.

Język odwzorowujący elementy rzeczywistości za pomocą samych tylko morfemów, w wypadku, gdy zbiór tych elementów rzeczywistości jest stosunkowo liczny (a tak jest w przypadku języka naturalnego, odwzorowującego całe uniwersum), musiałby mieć bardzo duży słownik, co znacznie utrudniałoby opanowanie takiego słownika, a więc i posługiwanie się językiem.

W języku sztucznym, a więc również w języku informacyjno-wyszukiwawczym, także wygodniej jest mieć do czynienia z takimi jednostkami leksykalnymi, które wskazywałyby od razu swoim kształtem przynależność wyrazu do odpowiednich klas semantycznych (a więc mających gramatyczne wykładniki takiej przynależności), w wielu przypadkach bowiem pozwoli to na wyszukiwanie pełnej interesującej nas informacji, co przy wyrazach niemotywowanych bardzo często byłoby w ogóle niemożliwe.

A oto przykład. Załóżmy, że użytkownika systemu informacyjnego interesują wszelkie informacje o wszystkich narzędziach i urządzeniach używanych w danej gałęzi przemysłu (przy danym procesie technologicznym, itp. – temat ten możemy oczywiście dowolnie zawęzić). Jeżeli ich nazwy nie będą niczym umotywowane (typu *pipeta*, *młot*, *łopata*), to pełną informację będzie można uzyskać tylko przez zadanie pełnej ich listy, a przecież o istnieniu jakichś nowych narzędzi i urządzeń, a więc często najlepszych, użytkownik może nie wiedzieć i co za tym idzie nie umieścić ich na liście. Jeżeli natomiast „bycie narzędziem” lub „bycie urządzeniem” będzie wpisane w strukturę wyrazu za pomocą oddzielnego morfemu (jest tak na przykład w wyrazach: *imadło*, *szydło*, *mydło*, *liczydło*, gdzie na „bycie narzędziem” wskazuje morfem *-dło*, oraz w wyrazach *drukarka*, *koparka*, *ładowarka*, *wiertarka*, *tokarka*, gdzie na „bycie urządzeniem” wskazuje formant *-arka*), pożądane informacje można uzyskać bez potrzeby wymieniania wszystkich narzędzi, po podaniu jedynie odpowiedniego formantu. Stąd wniosek, że w słowniku języka informacyjno-wyszukiwawczego powinny się znaleźć wyrazy o wyraźnym znaczenia strukturalnym, spełniające proporcje:

$$\frac{\text{znaczenie wyrazu 1}}{\text{znaczenie wyrazu 2}} = \frac{\text{struktura wyrazu 1}}{\text{struktura wyrazu 2}}$$

a więc np.:

(1)

$$\frac{\text{narzędzie do szycia}}{\text{narzędzie do mycia}} = \frac{\textit{szydło}}{\textit{mydło}}$$

$$\frac{\text{ten, kto spawa}}{\text{ten, kto kopie}} = \frac{\textit{spawacz}}{\textit{kopacz}}$$

$$\frac{\text{ten, kto myśli}}{\text{ten, kto uczy}} = \frac{\textit{myśliciel}}{\textit{nauczyciel}}$$

ale:

(2)

$$\frac{\text{narzędzie do szycia}}{\text{narzędzie do mycia}} \neq \frac{\textit{igła}}{\textit{gąbka}}$$

$$\frac{\text{ten, kto spawa}}{\text{ten, kto uczy}} \neq \frac{\textit{spawacz}}{\textit{nauczyciel}}$$

Przykład 2 jest ilustracją braku i w tym przypadku jedno-jednoznacznego przyporządkowania w języku naturalnym między funkcją a jej wykładnikiem, formanty *-acz* i *-iciel/-yciel* są bowiem formantami synonimicznymi (choć we współczesnym języku polskim nazwy stałych wykonawców czynności można tworzyć już tylko za pomocą formantu *-acz*).

W niektórych językach informacyjno-wyszukiwawczych takie kategorie gramatyczne oczywiście istnieją. Tak jest, na przykład, w UKD – tworzą je klasy wyrażenń stanowiących elementy poddziałów wspólnych, np.: poddział wspólny języka z wykładnikiem = ... (znak równania), poddział wspólny formy z wykładnikiem (0 ...) (nawias okrągły zero), poddział wspólny czasu z wykładnikiem "... (cudzysłów). W języku tym istnieją również gramatyczne kategorie semantyczne, np. kategoria semantyczna Nauki stosowane. Medycyna. Technika. Rolnictwo z wykładnikiem 6 ... (stojącym na początku wyrażenia).

Tak więc również wyrazy (elementarne jednostki leksykalne) niektórych języków informacyjno-wyszukiwawczych odznaczają się budową semantyczną motywowaną. W tych językach, których leksyka jest równoznaczna z wyrażeniami języka naturalnego, owa motywacja semantyczna należy do reguł systemowych nie języka informacyjno-wyszukiwawczego, lecz języka naturalnego – tak jest na przykład w języku haseł przedmiotowych, czego przykładem mogą być chociażby następujące jednostki leksykalne: Ptaki, Ptaki łowne, Ptaki wędrownie, Robotnicy, Robotnicy budowlani, Robotnicy portowi, ale w innych językach informacyjno-wyszukiwawczych taka motywacja semantyczna struktury jednostki leksykalnej jest ich cechą immanentną. Tak jest na przykład w UKD, czego przykładem może być tu istnienie tak zwanej rozbudowy równoległej symboli powo-

dującej, że jednostki leksykalne należące do różnych kategorii semantycznych, lecz zawierające pewien zespół takich samych komponentów semantycznych (mające wspólną część znaczenia) w swej formie różnią się jedynie wykładnikiem macierzystej kategorii semantycznej (reguła ta występuje w postaci „dalszy podział jak...” lub zapisów typu 669...1, 669...2). Przykładem mogą tu być trzy następujące gałęzie klasyfikacyjne:

- gałąź symboli wyrażających język dokumentu,
- gałąź symboli dokumentów, których tematem jest dany język,
- gałąź symboli dokumentów dotyczących literatury w danym języku lub należących do tej literatury, np.:

=20 Język angielski	802.0 Język angielski	820 Literatura angielska
=40 Język francuski	804.0 Język francuski	840 Literatura francuska
=84 Język polski	808.4 Język polski	884 Literatura polska
=945.11 Język węgierski	809.451.1 Język węgierski	894.511 Literatura węgierska

Budowę umotywowaną semantycznie mają oczywiście języki typu kodów semantycznych. Językom tym poświęcimy oddzielny rozdział skryptu.

W języku naturalnym spotykamy i takie wyrazy, których obecne znaczenie (**znaczenie realne**) odbiega od ich **znaczenia strukturalnego (etymologicznego)**. Do takich wyrazów należą, na przykład:

miednica, – kiedyś: przedmiot zrobiony z miedzi, dziś: naczynie w kształcie, dużej misy do mycia, prania itp.,

bielizna – znaczenie strukturalne „coś białego”, dawniej: białe ubranie, dziś: odzież spodnia albo wyroby z tkaniny używane do potrzeb domowych (bielizna pościelowa, stołowa), niekoniecznie białe.

Wszystkie morfemy, za pomocą których można tworzyć nowe wyrazy od wyrazów już w języku istniejących, to **formanty**. Formanty te mogą pełnić różne funkcje. Najważniejszą z nich, i najbardziej nas z punktu widzenia dobrego języka informacyjno-wyszukiwawczego interesującą, jest **funkcja realno-znaczeniowa (semantyczna)**. Taką funkcję pełnią na przykład wymienione już formanty: *-acz*, tworzący nazwy stałych wykonawców czynności, *-dło*, tworzący nazwy narzędzi, *-arka*, tworzący nazwy maszyn (choć nie jest to jedyna funkcja realno-znaczeniowa tego formantu), np.: *kopiarka*, *ładowarka*, *drukarka*, *dziurkarka*, *-arnia*, tworzący nazwy pomieszczeń (choć i ten formant pełni również inne funkcje), np.: *drukarnia*, *powielarnia*, *wyświetlarnia*, *suszarnia*.

Formant może pełnić funkcję realno-znaczeniową, ale nie musi jej pełnić, natomiast często pełni **funkcję strukturalną**, wskazując, do której kategorii gramatycznej (syntaktycznej, czyli tzw. części mowy), dany wyraz należy, Funkcję tylko strukturalną pełni na przykład w języku polskim formant *-ak*, tworzący rzeczowniki od przymiotników: np. *siwak*, *pustak*, *szarak*. Jedyna informacja, jaką możemy uzyskać z tego formantu, to taka, że wyrazy utworzone za jego pomocą są rzeczownikami rodzaju męskiego, formant ten nic bowiem nie mówi o ich

realnym znaczeniu. Jest to zarazem **funkcja metajęzykowa**, wyraża bowiem informację o samym kodzie, często niezbędną dla użytkownika języka do wyboru wykładników relacji syntagmatycznych przy tworzeniu syntagm i zdań języka. Tak jest na przykład w języku polskim, w którym pewne relacje syntagmatyczne wyrażane są przez związek zgody (dobór końcówek fleksyjnych wyrażających ten sam rodzaj gramatyczny, przypadek i liczbę).

Formanty tylko strukturalne są w języku naturalnym potrzebne, gdyż pozwalają na przenoszenie wyrazów z jednej kategorii syntaktycznej do innej kategorii, co jest często warunkiem użycia ich w odpowiedniej funkcji syntaktycznej. Tak jest na przykład w wypadku formantu *-ość*, tworzącego rzeczowniki od przymiotników, np. *mały – małość, znikomy – znikomość*, gdy istnieje potrzeba wprowadzenia danej informacji w pozycji podmiotu (por. np.: *Znikome efekty tej morderczej pracy zniechęcały do wszelkich działań. Znikomość efektów tej morderczej pracy zniechęcała do wszelkich działań.*). Tak jest w wypadku formantów *-anie, -enie, -cie*, tworzących rzeczowniki od czasowników, np.: *powlekać – powlekanie, spawać – spawanie, czyścić – czyszczenie, myć – mycie*, co wykorzystujemy tworząc charakterystyki słowne dokumentów, gdzie wszystkie czasowniki zostają zamienione na odpowiadającą im formę rzeczownikową, która jest standardową formą słów kluczowych.

Formanty mogą pełnić **funkcję ekspresywną**, pełnią ją na przykład w języku polskim przyrostki tworzące zdrobnienia i zgrubienia, np.: *-eczek – koteczek, -uś – kotuś, -unio – kotunio, -ątko – kociątko, -ur – kocur, -isko – kocisko*, wskazujące na stosunek nadawcy informacji do przedmiotów, których informacja dotyczy.

Przykładem świadomego wykorzystania takich formantów w wypowiedzi poetyckiej jest wiersz Tadeusza Borowskiego:

Kościółeczek

*Kościółeczek – dla Niemeczek,
w kościółeczku – puściuteńko.
Na srodeczku kościółeczka
Maluteńka amboneczka.*

*W okieneczkach – witrażyki,
na ścianeczkach – obrazeczki,
przy ścianeczkach zaś – grobiki:
doniczeczki i krzyżyki.*

*W doniczeczkach – lilijeczki,
konwalijki i dzwoneczki,
rózyczeczki, fuksyjeczki,
maluteńkie, bieluteńkie,
czyściuteńkie, świeżuteńkie,
ze chrzcielniczki pokrapiane,
wypieszczone, święciuteńkie
pamiąteczki po SS-manach...*

Takie formanty w języku informacyjno-wyszukiwawczym nie są potrzebne, bo jednostki tego języka nie powinny zawierać żadnej subiektywnej oceny.

Do wyrazów zalicza się w języku polskim także pojedyncze morfemy typu *stół* dlatego, że reguły gramatyki poziomu morfologicznego mówią, że wyraz może się składać tylko z jednego morfemu, pod warunkiem, że jest to morfem leksykalny, ale także dlatego, że w systemie języka stoją one w opozycji do innych form, takich jak: *stolu*, *stoły*, *stolami*, już wielomorfemowych. W dalszym ciągu jednak otwarty jest problem, które połączenia morfemów w tekście należy uznać za jeden wyraz, a których nie, i co za tym idzie, w jaki sposób wyrazy są w tekście oddzielane. W subkodzie graficznym języka polskiego brak tu niejednokrotnie konsekwencji, na przykład *niewykonanie* to jeden wyraz, ale *nie wykonać* to dwa wyrazy, natomiast *niewykonalny* jeden; *mógłby* to jeden wyraz, natomiast *można by* dwa wyrazy. Dlatego też w subkodzie pisanym wyróżnia się tzw. **wyrazy ortograficzne**, w wypadku których o ich granicach decydują reguły ortograficzne danego języka. Dla subkodu akustycznego języka naturalnego wprowadzono pojęcie wyrazu fonologicznego, a o wydzielaniu takich wyrazów w zdaniu decydują akustyczne środki delimitacji (segmentacji tekstu), takie jak akcent wyrazowy. Podział tekstu zgodnie z tym kryterium daje inne wyrazy niż podział oparty na kryteriach omówionych wyżej, za jeden wyraz należałoby bowiem uznać ciągi morfemów zawierające tzw. **klityki**, czyli morfemy niemające własnego akcentu, a łączące się w jeden ciąg (zestrój akcentowy) albo z morfemami poprzedzającymi – są to tzw. **enklityki**, np. w języku polskim formy krótkie zaimków takie jak *mi*, *ci*, a także zaimek *się*, albo z morfemami następującymi – tzw. **proklityki**, np. w języku polskim przyimki niestanowiące sylaby, takie jak *w*, *z*, a także niektóre inne przyimki i spójniki.

Również dla języków informacyjno-wyszukiwawczych nie zawsze łatwo jest sformułować definicję wyrazu. W świetle naszych poprzednich rozważań o problemie istnienia w tych językach morfemów nie jest możliwa do przyjęcia definicja mówiąca, że elementarną jednostką leksykalną języków informacyjno-wyszukiwawczych jest najprostszy znak semantycznie pełny.

Niemożliwa jest do przyjęcia także popularna definicja utożsamiająca elementarne jednostki leksykalne z tzw. wyrazami ortograficznymi (gdybyśmy za wyraz ortograficzny, tak jak to się czasami przyjmuje dla języka naturalnego, uznali ciąg liter od spacji do spacji), istnieją bowiem, na przykład w językach deskryptorowych i w językach haseł przedmiotowych, wyrażenia takie, jak Robotnicy budowlani, Ptaki wędrowne, KARA OGRANICZENIA WOLNOŚCI, KODEKS POSTĘPOWANIA W SPRAWACH O WYKROCZENIA, które składają się z kilku takich wyrazów ortograficznych (języka naturalnego), a w języku informacyjno-wyszukiwawczym stanowią jedną elementarną jednostkę leksykalną.

Nie można także przyjąć definicji mówiącej, że elementarna jednostka leksykalna języka informacyjnego jest to wyrażenie stanowiące odrębne hasło w leksykonie tego języka (słowniku), istnieją bowiem leksykony, które zawierają nie jednostki leksykalne, lecz całe zdania (np. słownik haseł przedmiotowych z zakresu medycyny Konopki), istnieją także takie, które zawierają i jednostki leksykalne,

i czasami, zdania, np. symbol znajdujący się w tablicach UKD 0611(100):327.56 Rada Bezpieczeństwa ONZ lub wyrażenie w słowniku języka haseł przedmiotowych Gardło – choroby, będące hasłem, a więc zdaniem tego języka składającym się z tematu Gardło i określnika choroby. Ponadto nie wszystkie elementarne jednostki leksykalne stanowią odrębne hasło w słowniku danego języka – tak jest na przykład w UKD, gdzie w tablicach brak elementarnych jednostek leksykalnych tworzonych zgodnie z regułą składniową „Dalszy podział jak...”.

O tym, jak należy łączyć ze sobą morfemy tak, aby powstał wyraz, decydują w każdym języku reguły składania morfemów – **reguły morfologiczne**, a wśród nich **reguły słowotwórcze**.

A oto przykłady reguł składniowych poziomu morfologicznego języka polskiego (reguł takich jest oczywiście o wiele więcej niż w wypadku systemu fonologicznego, więcej jest również jednostek stanowiących słownik tego poziomu):

- Wyraz może się składać tylko z jednego morfemu, np. *dom, las, nóż*.
- Jeżeli wyraz składa się tylko z jednego morfemu, to morfemem tym musi być tzw. morfem leksykalny (czyli morfemy gramatyczne – słowotwórcze i fleksyjne – nie mogą samodzielnie stanowić wyrazu; morfemy leksykalne są tu więc kategorią autosyntaktycznych morfemów obligatoryjnych, pozostałe morfemy tworzą klasę morfemów synsynstaktycznych).

- Morfem fleksyjny występuje zawsze w tzw. postpozycji, a więc po ostatnim morfemie słowotwórczym z klasy przyrostków lub po morfemie leksykalnym w wypadku braku przyrostków (reguły tej nie można zapisać w postaci: morfem fleksyjny występuje jako ostatni w wyrazie, mogą bowiem po nim wystąpić postfiksy, np. *któryś, któregoś, któremuś*, może też pojawić się w wyrazie w dwóch miejscach, gdy jest to wyraz składający się z dwóch morfemów leksykalnych, zaliczany do klasy zrostów, np. *Rzeczypospolitej, Białegostoku, Zielonejgóry*).

- Wyraz może zawierać więcej niż jeden morfem leksykalny – tak jest w tzw. wyrazach złożonych, takich jak: *listopad, zlewozmywak, meblościanka, woziwoda, królobójca, szaroniebieski, błękitnooka* lub w tzw. zrostach, np.: *krajobraz, psubrat, Wielkanoc* (te ostatnie nie są strukturami produktywnymi, a budowa ich da się wytłumaczyć jedynie historycznie – reguły ich tworzenia nie należą zatem do reguł współczesnego systemu morfologicznego języka polskiego). Wyrazy zawierające więcej niż dwa morfemy leksykalne, na przykład występujący w tytule książki Konwickiego *zwierzoczekoupiór*, zdarzają się w języku polskim niezmiernie rzadko.

- Jeżeli wyraz zawiera więcej niż jeden morfem leksykalny, to między te morfemy należy wstawić morfem gramatyczny (słowotwórczy) *-o-* lub *-i/-y-*.

- Jeżeli w wyrazie złożonym pierwszy człon leksykalny jest morfemem imiennym (a więc np. rzeczownikowym lub przymiotnikowym), to po nim musi występować morfem gramatyczny *-o-*, np.: *listonosz, białogłowa, drewnopodobny, niebieskozielony, białolicy, różanopalcza, kwasorodny, kwasoodporny, światłoczuły*.

- Jeżeli w wyrazie złożonym pierwszy człon leksykalny jest morfemem czasownikowym, to po nim musi wystąpić morfem *-i/-y-*, np.: *woziwoda, liczyrzepa, wierzypięta*.

- W wyrazie może wystąpić więcej niż jeden morfem z klasy przedrostków, np.: *nawypisywać, porozlepić, ponawypisywać, porozdzielać*. Kolejność przedrostków określają odpowiednie reguły słowotwórcze.

- W wyrazie może wystąpić więcej niż jeden morfem z klasy przyrostków, np.: *matusieńka, tatusieczek*. Kolejność przyrostków określają odpowiednie reguły słowotwórcze.

- W wyrazie może wystąpić zarówno przedrostek jak i przyrostek, np. *powozik, zalesić, rozanielić, napisać, najmilszy*.

- W wyrazie może wystąpić tylko jedna końcówka fleksyjna (wyjątek stanowią tzw. zrosty, np. *Białystok, Białegostoku*).

Reguły gramatyki składania morfemów w wyrazy często są tzw. regułami kontekstowymi, muszą bowiem uwzględniać sąsiedztwo innych morfemów współtworzących wyraz lub syntagmę. Takie właśnie reguły określają wybór **allomorfów**, czyli morfemów mających to samo znaczenie, ale różniących się formą (morfemów synonimicznych) realizowanych w danym kontekście. Takimi wariantami językowymi poziomu morfologicznego języka polskiego są oboczne postaci morfemów leksykalnych, np.,: *ręk-/rąk-/ręc-/ręcz-/rącz-*; *nóg-/nóż-/nóż-/noż-*; *móg-/moż-/móż-/móc*, których wybór zależy od morfemów następujących po nich (*ręka, rąk, ręce, ręcznik, rączka; noga, nóg, nóżka, nożna; mogę, możesz, móc*), przy czym wybór ten warunkują zarówno morfemy słowotwórcze, jak i morfemy fleksyjne (kończówki), a w tym ostatnim wypadku takie kontekstowe reguły muszą uwzględniać informacje pochodzące z poziomu syntaktycznego, a więc wyższego poziomu struktury języka. Dlatego to w takim wypadku często mówi się o **regułach morfosyntaktycznych** i o **systemie morfosyntaktycznym** języka naturalnego. Reguły kontekstowe dotyczą także tzw. **morfemów supletywnych**, takich jak polskie warianty *człowiek-/ludz-* czy *rok-/lat* (*człowiek, ludzie; rok, lata*). Różnica między tymi wariantami polega na tym, że współczesne różne formy morfemów obocznych są wynikiem fonetycznych zmian, jakie zachodziły w morfemach języka polskiego (morfemy oboczne wywodzą się więc z jednego morfemu), natomiast formy morfemów supletywnych były po prostu różne.

Nie wszystkie połączenia morfemów dozwolone przez system języka są faktycznie w języku realizowane, gdyż nie wszystkie są potrzebne. Zawsze jednak stanowią one **struktury potencjalne**, które w miarę potrzeby mogą być aktualizowane. I tak, zgodnie z prawami słowotwórstwa, możemy w języku polskim utworzyć na przykład od nieistniejącego, ale utworzonego zgodnie z jego regułami, rzeczownika *karborynka*, przymiotnik *karborynkowy*, stąd rzeczownik *karborynkowość*, przysłówek *karborynkowo*, czasownik *karborynkować*, od niego rzeczownik *karborynkowanie*, od czasownika nazwę wykonawcy czynności *karborynkowacz*, nazwę maszyny *karborynkowarka*, nazwę pomieszczenia, w którym wykonuje się tę czynność *karborynkowiarnia*, a także przymiotnik oznaczający cechę potencjalności *karborynkowalny* itd.

Takie niezrealizowane struktury słowotwórcze istnieją obok faktycznie istniejących jednostek systemu słownikowego, stanowiąc puste miejsca w systemie, które zawsze w miarę potrzeby mogą zostać zapełnione. W języku polskim mamy

na przykład takie puste miejsce obok słowa *żyrafa* w serii: *koń – konina, wieprz – wieprzowina, wół – wołowina, baran – baranina, gęś – gęsina, jagnię – jagnięcina*, w które moglibyśmy, gdyby była taka potrzeba, wstawić wyraz *żyrafina*. Istnienie takich pustych miejsc w systemie jest podstawą **produktywności systemu słowotwórczego** języka naturalnego. Jest również przyczyną tego, że słownika języka naturalnego nie można uznać za zbiór zamknięty, czego konsekwencją jest niemożność wyliczenia wszystkich jego jednostek. Nie robią tego nawet największe słowniki (leksykony) żadnego języka naturalnego.

Literatura

- Grzegorzczkova R.: *Zarys słowotwórstwa polskiego: słowotwórstwo opisowe*. Warszawa 1982.
- Jadacka H.: *System słowotwórczy polszczyzny (1945-2000)*. Warszawa 2001.
- Satkiewicz H.: *Produktywne typy słowotwórcze współczesnego języka polskiego*. Warszawa 1969.
- Grzegorzczkova R. i inni (red.): *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa 1984 (rozdz.: *Podstawowe pojęcia morfologii. Problemy ogólne słowotwórstwa*).

IX

System syntaktyczny. Relacje syntagmatyczne. Typologia syntaktyczna języków

W komunikacie muszą zostać odwzorowane odpowiednie elementy rzeczywistości i odpowiednie relacje zachodzące między nimi. O tym, w jaki sposób język odwzorowuje elementy rzeczywistości, których dotyczy informacja, powiemy dokładniej w następnych rozdziałach, teraz przedstawimy możliwość wyrażania w tekście relacji zachodzących między nimi.

Z postulatów odnoszących się do kodu optymalnego wiemy, że kod optymalny powinien być izomorficzny w stosunku do odwzorowywanej przez ten kod rzeczywistości, a komunikat wyrażony w takim kodzie powinien być izomorficzny w stosunku do odwzorowywanego w nim wycinka rzeczywistości (sytuacji zachodzącej w tej rzeczywistości). Żeby postulat ten mógł być spełniony, w systemie powinny być odwzorowane w sposób jedno-jednoznaczny wyróżnione w rzeczywistości elementy – umożliwiłoby to jednoznaczne odwzorowanie w komunikacie elementów tworzących daną sytuację, o której ma informować komunikat, system powinien też posiadać wykładniki relacji zachodzących między elementami rzeczywistości (również przyporządkowane jedno-jednoznacznie), co umożliwiłoby jednoznaczne odwzorowanie w komunikacie relewantnych dla użytkownika komunikatu relacji między elementami tworzącymi sytuację. Spełnienie postulatu izomorfizmu kodu względem rzeczywistości byłoby zatem tylko wtedy możliwe, gdyby przed skonstruowaniem takiego kodu z góry ustalić zarówno listę relewantnych elementów odwzorowywanej przez ten kod rzeczywistości, jak i listę zachodzących między nimi relacji relewantnych dla użytkowników kodu.

Język składa się ze słownika i gramatyki, teoretycznie optymalnym językiem byłby więc język, w którym elementy rzeczywistości byłyby odwzorowywane przez elementy słownika, a relacje między tymi elementami przez relacje zachodzące między wyrażeniami występującymi w utworzonych przez nie wyrażeniach złożonych (zdaniach). Relacje wyrażane w takim języku miałyby wyłącznie wykładniki gramatyczne, choć wykładniki te mogłyby mieć różny charakter: mogłyby to być pozycje w strukturze wyrażenia złożonego (miejsce, a w wypadku ciągu linearnego szyk, a właściwie miejsce w szyku) lub wykładniki w postaci odpowiednich morfemów. Zauważmy jednak, że do pełnego określenia relacji nie wystarczy wskazanie, jakie elementy są daną relacją połączone (a więc jakie ma ona

argumenty), konieczne jest także określenie rodzaju relacji łączącej te argumenty. Teoretycznie można to zrobić na dwa sposoby: albo wprowadzić wykładniki relacji, albo wykładniki ról połączonych nią argumentów. Jeśli między elementami odwzorowywanej rzeczywistości zachodzi więcej niż jedna relewantna relacja, to konieczne jest ponadto wskazanie, które elementy są argumentami jednej relacji, trzeba więc wprowadzić wykładniki wiążące funktor (wykładnik relacji) w jedną całość z jej argumentami (a więc tzw. wskaźniki więzi). W wypadku języka naturalnego liczba relacji zachodzących między elementami rzeczywistości jest jednak tak duża, że znaczna ich część ma swoje wykładniki leksykalne, np.:

Marysia Kowalska jest matką Ani.

Ania jest córką Marysi Kowalskiej.

Kowalski jest wychowawcą klasy piątej.

Ania jest starsza od Marysi, ale Marysia jest wyższa (od Ani).

Zwróćmy jednak uwagę na to, że w zdaniach tych relacje są współokreślane przez inne wykładniki, o charakterze gramatycznym (tu przez końcówki fleksyjne). Tym rodzajem wykładników relacji między elementami tworzącymi wyrażenie złożone (zdanie) zajmiemy się w tym rozdziale. Relacje takie, zachodzące między jednostkami języka współwystępującymi w wyrażeniu złożonym (zdaniu) nazywa się **relacjami syntagmatycznymi**, w odróżnieniu od relacji paradygmatycznych, które zachodzą między izolowanymi elementami systemu językowego. Termin relacje syntagmatyczne pochodzi od twórcy strukturalizmu w lingwistyce, Ferdinanda de Saussure'a i określa wszelkie relacje zachodzące między wyrażeniami w utworzonym przez nie wyrażeniu złożonym (tekście), niezależnie od poziomu struktury języka: na poziomie fonologicznym są to więc relacje między fonemami w morfemie, na poziomie morfologicznym – relacje między morfemami tworzącymi wyraz, na poziomie syntaktycznym – relacje między wyrazami tworzącymi syntagmy i zdania. W węższym znaczeniu termin ten bywa używany tylko do określania tych ostatnich relacji, zwanych także **relacjami syntaktycznymi** i w tym znaczeniu będziemy go dalej używać.

Nim przejdziemy do dalszych rozważań, jeszcze raz chcielibyśmy podkreślić, że w wypadku języka naturalnego nie ma izomorfizmu pomiędzy strukturą wyrażenia a strukturą odwzorowywanych przez nie sytuacji w rzeczywistości, inna jest więc struktura relacyjna tych dwóch planów – planu sytuacyjnego i planu syntaktycznego wyrażenia (zdania), a wykładniki relacji syntagmatycznych często współokreślają wraz z wykładnikami leksykalnymi relacje zachodzące w rzeczywistości lub tylko sygnalizują, że między wyróżnionymi jej elementami jakies relacje zachodzą, nie wskazując, jakie to są relacje (tego odbiorca komunikatu będzie musiał się domyślić na podstawie znajomości rzeczywistości pozatekstowej lub na podstawie sytuacji komunikacyjnej). Żaden język naturalny nie realizuje więc zasady, że relewantne (ze względu na ten język) elementy rzeczywistości odwzorowywane są przez morfemy, a relewantne relacje zachodzące między nazywanymi przez nie elementami przez zachodzące między tymi morfemami relacje syntagmatyczne w tworzonych przez nie wyrażeniach złożonych.

Zacznijmy od rozpatrywania możliwości teoretycznych, które poprzemy przykładami z języków naturalnych. W językach naturalnych najczęściej spotyka się wiele różnych sposobów wyrażania relacji, nie stanowią więc one czystych typów syntaktycznych, toteż zazwyczaj zalicza się je do takiego typu syntaktycznego, który w danym języku przeważa. Będziemy więc rozpatrywać tu takie idealne, „czyste” typy syntaktyczne.

Każda wypowiedź ma przebieg linearny, a skoro tak, to teoretycznie rzecz biorąc najbardziej ekonomiczny mógłby się wydawać taki język, w którym miejsce elementu w wypowiedzeniu wskazywałoby na to, w jakich relacjach znajduje się odwzorowywany przezeń element rzeczywistości w stosunku do wszystkich innych elementów odwzorowywanych w danym wypowiedzeniu. Taki sposób oznaczania relacji mógłby się wydawać najbardziej funkcjonalny, gdyż nie wymagałby dodatkowych leksykalnych wykładników relacji, których wprowadzenie powodowałoby wydłużenie tekstu. Najprostszym przykładem takiego właśnie sposobu wyrażania relacji między poszczególnymi elementami tekstu może być konwencja umieszczania na karcie tytułowej książki nazwiska autora, tytułu dzieła i wydawcy, wiadomo bowiem, że na pierwszym miejscu (u góry) umieszcza się (w obowiązującej w Polsce konwencji) imię i nazwisko autora (relacja bycia autorem zachodząca między człowiekiem a tekstem), na drugim miejscu (w środku) tytuł książki (relacja bycia tytułem tekstu), a na ostatnim (u dołu) nazwę wydawcy (relacja bycia wydawcą tekstu), np.:

Thomas Stearns Eliot
Szkice krytyczne
Państwowy Instytut Wydawniczy

Eliza Orzeszkowa
Meir Ezofowicz
Michał Arct

W pierwszym przykładzie w ustaleniu relacji dopomaga nam wiedza o języku, wiemy bowiem, że *Thomas Stearns Eliot* to imiona i nazwisko, *Szkice krytyczne* to pewien gatunek dzieła literackiego, mogący pełnić funkcję tytułu, a *Państwowy Instytut Wydawniczy* to nazwa instytucji wydawniczej. Ale w drugim przykładzie nie ma żadnych takich dodatkowych punktów zaczepienia dla ustalenia rodzaju relacji, mamy tu bowiem do czynienia trzy razy z imieniem i nazwiskiem i tylko pozycja pozwala nam na ustalenie odpowiedniej relacji między nimi. Taką samą sytuację mamy również w przypadku karty katalogowej książki.

Pozycja pozwala również określić, co jest imieniem, a co nazwiskiem, konwencjonalnie bowiem imiona pisze się w języku polskim na miejscu pierwszym (ewentualnie na pierwszym i drugim), nazwisko zaś na ostatnim, mamy więc:

Marek Kowalski,

Marek Stanisław Kowalski,

ale i:

Andrzej Marek,

Andrzej Grzegorz Marek,

gdzie w pierwszej grupie przykładów *Marek* jest imieniem, natomiast w drugiej *Marek* to nazwisko.

Takich idealnych **języków pozycyjnych**, bez żadnych dodatkowych wykładników relacji syntagmatycznych, nie spotykamy wśród języków naturalnych po prostu dlatego, że język taki przestaje być funkcjonalny wtedy, gdy trzeba wyrażać wiele różnych relacji, w języku takim bowiem liczba relacji musi być z góry przewidziana i dla argumentu każdej z nich musi być z góry przewidziane miejsce. W rzeczywistości sprowadza się to do tego, że w języku naturalnym, którego teksty mają charakter linearny, z góry musiałaby być określona długość zdania (liczba elementów, z których musi się zdanie składać), nawet, jeśli w danej informacji nie wszystkie elementy występują. Jakiegokolwiek skrócenie zdania jest tu niemożliwe, gdyż prowadziłoby to do zmiany pozycji jego elementów, a przecież pozycja jest wykładnikiem relacji i zmieniona być nie może. Ponieważ rzadko jednak w wypadku komunikatów nadawanych w języku naturalnym istnieje potrzeba wyrażania jednocześnie relacji pomiędzy więcej niż kilkoma elementami, istnieje natomiast potrzeba wyrażania wielu różnych relacji (a więc zarezerwowania na nie odpowiedniej liczby miejsc w wypadku przyjęcia pozycyjnych wykładników relacji syntagmatycznych), w języku takim trzeba by przewidzieć bardzo długie zdania, a do jego słownika wprowadzić jakiś element pusty, który zajmowałby miejsce w wypadku, gdy w danej informacji brak danego argumentu przewidzianej relacji.

Rozpatrzmy to na przykładzie:

Załóżmy, że w uniwersum, którego odzwierciedleniu służy nasz język, istnieje tylko dziewięć relewantnych (interesujących nas) relacji jednoargumentowych. Każde zdanie takiego języka będzie się więc składało z dziewięciu jednostek elementarnych, słownik naszego języka będzie się składał z małych liter alfabetu łacińskiego, którym są przyporządkowane jakies elementy rzeczywistości (uniwersum), będących znakami pełnymi języka oraz z elementu pustego – 0, wypełniającego puste miejsca w zdaniu, w wypadku, gdy argumenty danej relacji nie występują w danym komunikacie. Taki język mógłby na przykład przekazywać informacje o produkcji jakichś stosunkowo nieskomplikowanych związków chemicznych, mógłby więc służyć do zapisywania receptur. W zdaniach tego języka kolejne pozycje ich elementów byłyby więc na przykład wykładnikami następujących relacji:

- 1 – bycie wyjściowym składnikiem głównym,
- 2 – bycie wyjściowym składnikiem dodatkowym,
- 3 – bycie katalizatorem wyjściowym,
- 4 – bycie produktem pierwszego etapu procesu,
- 5 – bycie składnikiem dodatkowym w drugim etapie,
- 6 – bycie pierwszym katalizatorem w drugim etapie,

7 – bycie drugim katalizatorem w drugim etapie,

8 – bycie produktem zamierzonym,

9 – bycie produktem dodatkowym.

Moglibyśmy więc mieć w tym języku następujące zdania:

z k 0 0 0 0 0 n 0

l c f 0 0 0 0 a 0

n d a z k m r w y

a m 0 i u r 0 z x

Chcąc umożliwić sobie skracanie zdań, co jest szczególnie ważne z punktu widzenia automatyzacji operacji tekstowych, mogliśmy nasz minijęzyk przebudować, likwidując element pusty, natomiast przy każdym pełnym elemencie dopisywać numer miejsca tego elementu w ciągu całego zdania będącego wykładnikiem relacji. Tak zmodyfikowane powyższe zdania miałyby obecnie postać:

z1 k2 n8

l1 c2 f3 a8

n1 d2 a3 z4 k5 m6 r7 w8 y9

a1 m2 i4 u5 r6 z8 x9

Ale teraz zmieniła się struktura języka. W przypadku języków pozycyjnych bowiem wykładniki relacji nie były elementami tekstu, nie były więc elementami słownika morfemów, natomiast po takim przekształceniu stały się elementami tekstu, a po to musiały się oczywiście stać także elementami słownika. Teoretycznie mielibyśmy więc dwa główne **typy syntaktyczne języków**:

I. Języki, w których relacje syntagmatyczne nie mają wykładników morfemowych, a funkcję ich pełni pozycja (szyk) morfemów będących wykładnikami argumentów tych relacji. Takie języki nazywamy **językami pozycyjnymi**.

II. Języki, w których wykładniki relacji są morfemami, stanowią więc elementy tekstu (i słownika). Są to **języki niepozycyjne**.

Podział na języki pozycyjne i niepozycyjne opiera się na stosunkowo prostym kryterium, bardziej skomplikowany jest podział na poszczególne podtypy języków niepozycyjnych, jest on bowiem uwarunkowany strukturą języków naturalnych, a kryterium stanowi tu **stopień samodzielności wykładników relacji** w stosunku do elementów tekstu będących wykładnikami argumentów tych relacji. Tu teoretycznie możemy mieć dwie sytuacje:

II.1. Wykładnikami relacji są oddzielne, samodzielne morfemy niełączące się w całości wyrazowe z morfemami stanowiącymi ich argumenty (porównaj rozważania o wyrazie).

II.2. Wykładnikami relacji są morfemy łączące się w całości wyrazowe z morfemami stanowiącymi ich argumenty.

Pierwszy z wymienionych tu typów (II.1) to języki, w których morfemy leksykalne nigdy nie zmieniają swej postaci, a stosunki między nimi są wyrażane za pomocą luźnych morfemów gramatycznych. Takie języki nazywa się **językami izolującymi** albo **językami analitycznymi**. Drugi z wyróżnionych tu typów

języków, w którym wykładniki relacji wraz z wykładnikami argumentów tworzą jeden wyraz, nazywa się **językami nieizolującymi (językami syntetycznymi)**.

W tych językach morfemy będące wykładnikami relacji, ponieważ są elementami tekstu, muszą w tym tekście zajmować określone miejsce w stosunku do morfemów stanowiących ich argumenty. Tak jak w wypadku łączenia morfemów w wyrazy, tak i tutaj możliwe są następujące pozycje:

1. Morfem będący wykładnikiem relacji zajmuje miejsce przed swoimi argumentami, znajduje się więc w prepozycji.
2. Morfem będący wykładnikiem relacji zajmuje miejsce między swoimi argumentami, znajduje się więc w interpozycji.
3. Morfem będący wykładnikiem relacji zajmuje miejsce po swoich argumentach, a więc znajduje się w postpozycji.
4. Morfem będący wykładnikiem relacji zajmuje miejsce w środku jednego ze swoich argumentów (między tworzącymi go wyrażeniami elementarnymi), a więc w impozycji.

A oto przykłady takich syntagm z języka logiki i matematyki z wykładnikami relacji zajmującymi wyżej wymienione miejsca:

1. Prepozycja: Kxy – zapis koniunkcji w beznawiasowej notacji Łukasiewicza, gdzie K jest wykładnikiem koniunkcji.
2. Interpozycja: $x = y$ zapis równości w matematyce, gdzie wykładnikiem relacji jest znak równania $=$.
3. Postpozycja: x^n – zapis potęgowania, gdzie n jest wykładnikiem relacji między x -ami, które należy pomnożyć przez siebie n razy.

Przykład impozycji podamy przy omawianiu języków naturalnych.

Języki nieizolujące (II.2), czyli syntetyczne, a więc te, w których wykładniki relacji syntagmatycznych łączą się w całości wyrazowe z morfemami stanowiącymi wykładniki ich argumentów, można z kolei podzielić na dwa podtypy zależnie od tego, czy możliwy jest **podział wyrazu na elementy realno-znaczeniowe i funkcjonalne** (wykładniki relacji), czy też nie. Mamy więc:

II.2.1. Języki, w których nie można przeprowadzić takiej granicy – są to **języki fleksyjne**.

II.2.2. Języki, w których można przeprowadzić taką granicę – są to **języki niefleksyjne**.

Języki niefleksyjne, a więc takie, w których można wyraźnie oddzielić wykładniki semantyczne od wykładników relacji syntagmatycznych, dzielimy z kolei na dwie grupy zależnie od tego, czy są to **wykładniki ciągłe**, czy też **wykładniki nieciągłe**. Mamy więc:

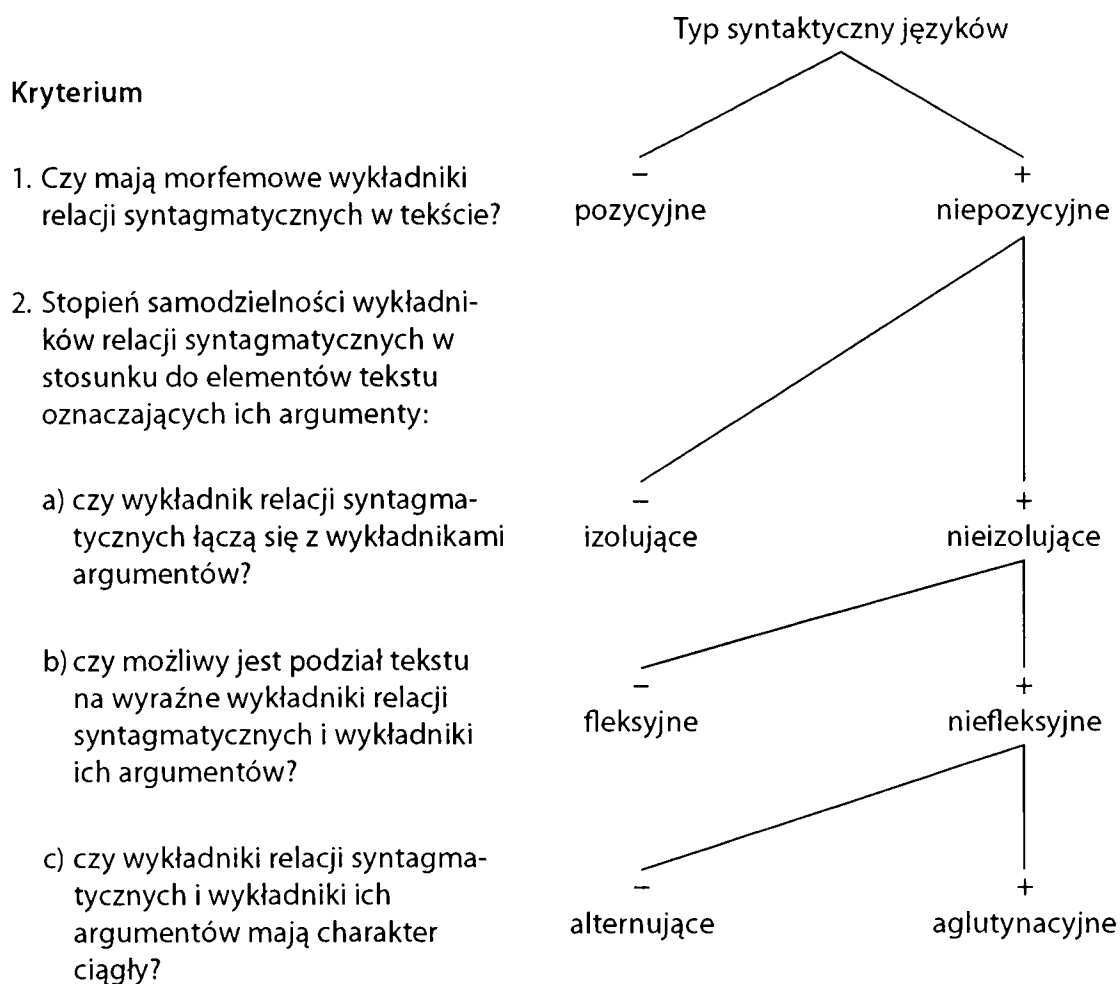
II.2.2.1. Języki z **ciągłymi wykładnikami** semantycznymi i ciągłymi wykładnikami relacji syntagmatycznych – są to **języki aglutynacyjne**.

II.2.2.2. Języki z **nieciągłymi wykładnikami** semantycznymi (morfemami leksykalnymi) – są to **języki alternacyjne**, przy czym wykładniki relacji syntagmatycznych mogą tu mieć charakter ciągły lub nieciągły.

Chcielibyśmy jeszcze zwrócić uwagę na to, że terminy język analityczny i język syntetyczny bywają używane także w innym znaczeniu: **językiem analitycznym**

nazywa się taki język, gdzie w jednym morfemie nie występuje kumulacja funkcji składniowej i semantycznej, **językiem syntetycznym** taki, gdzie jeden morfem jest wykładnikiem i funkcji syntaktycznej i funkcji semantycznej. W takim rozumieniu analityczne byłyby wyróżnione przez nas języki niefleksyjne (alternacyjne i aglutynacyjne), syntetycznymi byłyby języki fleksyjne.

Wyróżnione tu typy syntaktyczne języków przedstawiamy na wykresie:



Po tych teoretycznych rozważaniach przejdźmy do przykładów językowych, na razie jeszcze jednak z języków sztucznych, języki naturalne bowiem nie stanowią, jak wiemy, typów syntaktycznie czystych.

Wyżej podaliśmy już przykłady syntagm takich języków dla języka pozycyjnego i języka izolującego, bez egzemplifikacji pozostały jeszcze języki aglutynujące (aglutynacyjne), alternujące (alternacyjne) i fleksyjne. W tym celu będziemy musieli nasz przykładowy minijęzyk nieco skomplikować, wprowadzając wieloelementowe morfemy leksykalne (w poprzednich przykładach były one jednoelementowe) i wieloelementowe wykładniki relacji. Weźmy dla przykładu dwa morfemy leksykalne: abcdef, ghijkl i dwa wykładniki relacji (morfemy grama-

tyczne o funkcji syntaktycznej): 12345 i 56789. Zgodnie z naszymi kryteriami, w językach aglutynacyjnych i językach alternacyjnych wykładniki relacji syntagmatycznych mają się łączyć w całości wyrazowe przynajmniej z jednym z argumentów tej relacji. Teoretycznie mamy tu do dyspozycji następujące możliwości takich połączeń:

1. Bez rozbijania ich ciągłości, a więc np.:

12345abcdef
abcdef12345
abcdef56789
56789ghijkl
ghijkl56789

– takie połączenie jest typowe dla aglutynacji, mamy tu bowiem ciągle wykładniki zarówno funkcji semantycznej, jak i funkcji syntaktycznej, możliwe jest również przeprowadzenie ostrej granicy między tymi dwoma typami wykładników.

2. Rozbijając ich ciągłość, a więc np.:

abc12345def
a12345bcdef
a1b2c3d4e5f
g1h2i3j4k5l
1a2b3c4d5ef
123abcdef45

– takie połączenie jest typowe dla alternacji. Mamy tu wprowadzić możliwość dokładnego oddzielenia wykładników semantycznych od wykładników relacji syntagmatycznych – morfemy leksykalne składają się bowiem z małych liter alfabetu łacińskiego, natomiast morfemy gramatyczne o funkcji syntaktycznej z cyfr arabskich, to jednak jedne i drugie mogą występować w tekście w postaci nieciągłej.

Z powyższego widać, że również w wypadku języków nieizolujących wykładnik relacji syntagmatycznej występować może w jednej z trzech wyżej wymienionych pozycji, a więc w prepozycji – przed morfemem leksykalnym będącym wykładnikiem jednego z jego argumentów, w postpozycji – po morfemem leksykalnym będącym wykładnikiem jednego z jego argumentów i w impozycji – między elementami morfemu leksykalnego.

W opisie języków naturalnych **terminy języki aglutynacyjne, języki fleksyjne, języki alternacyjne** używane są często nie tylko w stosunku do sposobu wyrażania w tych językach relacji między wyrażeniami tworzącymi zdania, ale także w stosunku do sposobów tworzenia innych form gramatycznych – jest to więc znaczenie szersze, obejmujące także kategorie słowotwórcze. Użycie terminów języki aglutynacyjne i języki fleksyjne często ulega zawężeniu związanemu z pozycją wykładników relacji syntagmatycznych do języków, w których morfemy te występują w postpozycji w stosunku do morfemu leksykalnego, mówi się więc o końcówkach aglutynacyjnych i końcówkach fleksyjnych – definicja sformułowana wyżej pozycji takiej nie przesądza.

Odwzorowanie sytuacji pozajęzykowej w tekstach języka naturalnego nigdy nie jest odwzorowaniem jedno-jednoznaczny także dlatego, że liczba potencjalnie możliwych różnorodnych typów sytuacji nie jest niczym ograniczona – nie jest więc ograniczona liczba możliwych, faktycznie istniejących w otaczającym nas świecie relacji, natomiast repertuar służących do ich odwzorowania struktur językowych jest z góry przez gramatykę danego języka określony, stąd w ten sam sposób – za pomocą tego samego wzorca zdaniowego – odwzorowywane bywają sytuacje o zupełnie różnej strukturze. Przykładem mogą tu być chociażby polskie zdania:

Jan jedzie tramwajem.

Jan jedzie Marszałkowską.

będące realizacją jednej struktury zdaniowej:

$Y_{\text{Nom}} V_{\text{Pers}} Y_{\text{Instr}}$

a odwzorowujące zupełnie różne sytuacje, lub też będące realizacją tegoż schematu zdania:

Jan maluje pędzlem.

Jan maluje farbami olejnymi.

Jan maluje wieczorami.

Jan maluje całymi godzinami.

We wszystkich tych przykładach współwyznacznikiem relacji istniejącej pomiędzy elementami rzeczywistości odwzorowanymi w tekście przez X i Y jest oprócz ich formy gramatycznej (morfemu gramatycznego o funkcji syntaktycznej – tu końcówki fleksyjnej) użyty w tych zdaniach czasownik. I tak, na przykład, czasownik *jechać* wskazuje między innymi istnienie pewnej relacji przestrzennej, polegającej na zmianie położenia jakiegoś przedmiotu w jakiejś przestrzeni, natomiast czasownik *malować* takiej relacji nie komunikuje.

W tekście relacje syntagmatyczne między elementami tekstu są więc niejako zależne od naturalnych powiązań ich desygnatów, a struktura syntagmatyczna jest zazwyczaj prostsza niż struktura rzeczywistości.

Języki naturalne nie stanowią na ogół czystych typów syntaktycznych i w każdym z nich istnieje bardzo wiele różnorodnych środków służących wyrażaniu wielu różnorodnych relacji zachodzących w rzeczywistości pozatekstowej. Na przykład w języku polskim celowi temu służą czasowniki, przyimki, końcówki fleksyjne, niektóre przedrostki, spójniki oraz niektóre rzeczowniki, przy czym ta sama relacja może być wyrażona na kilka różnych sposobów (synonimicznie). Niżej podajemy takie przykłady z języka polskiego:

ojcowa chałupa, chałupa ojca, chałupa należąca do ojca;

miedziana rura, rura z miedzi, rura zrobiona z miedzi;

wiejski chłopak, chłopak ze wsi;

jechać konno, jechać na koniu;

*jechać motorem, jechać na motorze;
kamienne serce, serce z kamienia, serce jak kamień;
pociąg jedzie przez wsie i miasteczka, pociąg przejeżdża przez wsie i miasteczka;
spacerować Marszałkowską, spacerować po Marszałkowskiej;
Odrobiwszy lekcje poszedł grać w piłkę; Po odrobieniu lekcji poszedł grać w piłkę; Najpierw odrobił lekcje, a potem poszedł grać w piłkę; Gdy odrobił lekcje poszedł grać w piłkę; Kiedy odrobił lekcje, poszedł grać w piłkę; Odrobił lekcje i poszedł grać w piłkę.*

Do języków pozycyjnych zalicza się spośród języków naturalnych zazwyczaj język angielski (czasami także francuski), podając przy tym przykłady typu:

*John loves Mary.
Jean aime Marie.*

gdzie o tym, który z wyrazów jest podmiotem, a który dopełnieniem decyduje jedynie miejsce w zdaniu – podmiot stoi przed orzeczeniem, dopełnienie po orzeczeniu. Zamiana tych wyrazów miejscami powoduje zmianę funkcji, a zatem i zmianę znaczenia:

*Mary loves John.
Marie aime Jean.*

Zamiana taka jest natomiast możliwa w języku polskim, gdzie nie powoduje zmiany informacji, gdyż w tym przypadku wykładnikiem stosunków syntagmatycznych nie jest pozycja wyrazu w zdaniu, lecz jego końcówka: dla podmiotu forma mianownika, dla dopełnienia biernik:

*Jan kocha Marię.
Marię kocha Jan.
Maria kocha Jana.
Jana kocha Maria.*

Ani język angielski, ani francuski nie stanowią jednak czystego typu pozycyjnego, takiego o jakim mówiliśmy wyżej, gdyż w obu tych językach relacje mają w tekście także wykładniki morfemowe, na przykład bardzo liczne w tych językach przyimki:

*John is living in the mountains.
Jean entre dans la maison.*

Języki te należą więc raczej do języków izolujących, choć mamy w nich także czysto pozycyjne wykładniki relacji.

Językiem fleksyjnym jest język polski, choć i w nim spotykamy inne sposoby wyrażania relacji, gdyż nie w każdym wypadku jest tak, że nie można przeprowadzić wyraźnej granicy między elementami tekstu składającymi się na morfemy leksykalne i morfemy będące wykładnikami relacji syntaktycznych. Język polski jest, zgodnie ze sformułowaną wyżej definicją, językiem fleksyjnym dlatego, że **koń-**

cówki fleksyjne kumulują dwie funkcje: są wykładnikami relacji syntagmatycznej, w jakiej pozostaje argument wyrażany przez tzw. temat fleksyjny (z którym tworzą jeden wyraz) z innymi wyrazami w zdaniu (lub syntagmie), ale jednocześnie podają informację o rzeczywistości pozatekstowej dotyczącej elementu tej rzeczywistości odwzorowanego przez argument relacji syntagmatycznej, a mianowicie mówią o liczbie elementów zbioru denotowanego (w języku polskim informują, czy jest to jeden element, czy więcej niż jeden, trzeba bowiem wybrać – obligatoryjnie – między końcówką fleksyjną odpowiedniego przypadku liczby pojedynczej i liczby mnogiej). Za typowe współwykładniki relacji syntagmatycznych (obok końcówek fleksyjnych) w językach fleksyjnych uważa się tzw. **wymiany tematyczne**, a więc warunkowaną przez końcówkę fleksyjną postać obocznych tematów fleksyjnych różniących się składem tworzących je fonemów (zjawisko to nazywa się nawet **fleksją wewnętrzną**, w opozycji do fleksji właściwej, czyli **fleksji końcówkowej**). Trzeba jednak pamiętać o tym, że **alternacje głoskowe** w temacie są wynikiem historycznych procesów fonetycznych (upodobnień artykulacyjnych) i prymarnie nie pełniły funkcji wykładników relacji syntagmatycznych. Gdyby tak było, głoskowe wymiany tematyczne (samogłoskowe i spółgłoskowe) musiałyby mieć charakter systemowy na poziomie morfosyntaktycznym (a nie fonologicznym), a wtedy język polski obok fleksyjnych wykładników relacji syntagmatycznych miałby również wykładniki alternacyjne. Alternacje głoskowe we współczesnym języku polskim dodatkowo świadczą o tym, że w słowoformach występujących w zdaniach nie da się dokonać podziału na te elementy, które są wykładnikami informacji tylko o rzeczywistości pozatekstowej (argumentach) i te, które są wykładnikami informacji tylko o rodzaju relacji syntagmatycznej, a więc wzmacniają tezę stanowiącą sformułowane w typologii syntaktycznej języków kryterium, na mocy którego wydzielono typ języków fleksyjnych, ale nie one rozstrzygają o zaliczeniu języka polskiego (i innych języków z takimi alternacjami tematycznymi) do języków fleksyjnych. Taka struktura fleksyjnych języków naturalnych, a przede wszystkim oboczności tematyczne, powoduje trudności w formalizacji ich systemów morfosyntaktycznych oraz formalizacji operacji przetwarzania tekstów tych języków. Ma to istotne znaczenie w komputerowej obróbce tekstów, na przykład w automatycznym wybieraniu z tekstu słów kluczowych dla potrzeb wyszukiwarek internetowych oraz dla wyszukiwania w nich informacji za pomocą słów kluczowych w tych językach. Nie wystarcza tu bowiem zastosowanie zwykłej operacji maskowania, konieczne jest także uwzględnianie wszystkich wariantów obocznych tematów fleksyjnych.

Typowymi przykładami ilustrującymi powyższe są takie formy języka polskiego jak *ra*k w zdaniu:

*Maria nigdy nie za*lamuje *ra*k.

gdzie wykładnikiem dopełniacza, będącego tutaj wykładnikiem dopełnienia, jest zarówno brak końcówki, jak i samogłoska *a*, oboczna do *ę* (występującego na przykład w formie *ręka*), a także spółgłoska *k*, oboczna do *c* (na przykład w formie *ręce*). Zarówno samogłoska *a*, jak i spółgłoska *k* są tu jednak także wykładnikami

(składnikami) morfemu, któremu przyporządkowany jest w rzeczywistości pozatekstowej pewien określony przedmiot, są więc wykładnikami funkcji semantycznej, w połączeniu ze spółgłoską *r* tworząc morfem leksykalny. Ponadto cała ta forma jest także wykładnikiem liczby mnogiej w opozycji, na przykład, do analogicznej formy *ręki*, będącej wykładnikiem liczby pojedynczej, a więc informacji niemającej prymarnego charakteru wykładnika relacji syntagmatycznych (choć wykładnik liczby może wtórnie pełnić funkcję wykładnika relacji syntagmatycznych, stanowiąc wykładnik związku zgody – wskaźnika więzi, na przykład w formie przymiotnika lub czasownika – por. *chuda ręka – chude ręce; jakaś ręka ścisnęła mnie mocno – jakieś ręce ścisnęły mnie mocno*). Tak więc, w przytoczonym tu przykładzie formy *rąk* nie można oddzielić tych elementów wyrażenia, które są wykładnikami funkcji semantycznej – morfemu leksykalnego (nie może nim być samo *r* – tylko ono się tu nie zmienia – nie niesie bowiem tego samego znaczenia we wszystkich innych użyciach) od tych elementów, które pełnią rolę wykładników relacji syntagmatycznych. Podobnie dzieje się w formie *drodze*, np.:

Przy drodze rosły stare lipy.

gdzie we współczesnym języku polskim wykładnikiem miejscownika jest zarówno końcówka *-e*, jak i spółgłoska *dz*, oboczna do *g* na przykład w formie *droga*, która jest jednocześnie elementem morfemu leksykalnego, przy czym oba te elementy tekstu są równocześnie wykładnikiem liczby pojedynczej. Takie wymiany w temacie są bardzo charakterystyczne dla języka polskiego i innych języków fleksyjnych i, jak pamiętamy, noszą miano **fleksji wewnętrznej**, w opozycji do **fleksji końcówkowej** nazywanej **fleksją zewnętrzną**, ale nie są jednak warunkiem do uznania języka polskiego za język fleksyjny. Język polski byłby językiem fleksyjnym (zgodnie z przyjętą przez nas definicją języka fleksyjnego) nawet wówczas, gdyby istniała w nim tylko fleksja końcówkowa, gdyż i tak nie można by było dokonać w nim podziału tekstu na elementy pełniące wyłącznie funkcję wykładników relacji syntagmatycznych i elementy pełniące tylko funkcję wykładników argumentów tych relacji, każda końcówka fleksyjna pełni bowiem oprócz funkcji wykładnika relacji syntagmatycznych funkcję wykładnika liczby, która, przynajmniej w rzeczownikach, nie może być uznana za wykładnik relacji syntagmatycznych. Jako wykładnik takiej relacji można uznać liczbę wyrażaną przez końcówki części mowy występujących w zdaniu z rzeczownikiem w związku zgody, gdzie wybór końcówki warunkowany jest przez liczbę (i rodzaj gramatyczny) określanego rzeczownika, prymarną funkcją tych końcówek jest tu bowiem funkcja syntaktyczna (wskazują, które wyrażenia zdania stanowią jeden człon syntaktyczny i wspólnie niosą informacje o elemencie rzeczywistości odwzorowywanym przez argument relacji syntagmatycznej), natomiast funkcja wykładnika informacji o liczbie zbioru denotowanego w rzeczywistości pozatekstowej przez wykładnik argumentu tej relacji jest funkcją wtórną, a w dodatku jest to wykładnik redundantny. Dobrze widać to w przykładzie:

Przy polnej drodze rosły stare lipy

gdzie mamy do czynienia ze złożonym (nieciągłym) wykładnikiem relacji syntagmatycznej (między drogą a lipami) – główną jego część stanowi przyimek *przy*, określający rodzaj relacji przestrzennej, będący tu luźnym morfemem, nie tworzy bowiem formy wyrazowej z innymi morfemami tekstu (można powiedzieć: *przy polnej drodze*) – a więc mamy tu taką sytuację jak w językach izolujących, ale tylko z pozoru, gdyż drugą część wykładnika relacji syntagmatycznej stanowi morfem związany w postaci końcówki fleksyjnej *-e*, przy czym końcówka ta stanowi część obligatoryjną tego wykładnika (jest rządzona przez przyimek *przy*). W całości wykładnik ten wskazuje, który z elementów rzeczywistości występujący w określonej relacji przestrzennej jest elementem sytuującym, wykładnikiem elementu sytuowanego jest bowiem forma mianownika. Kończące fleksyjne przymiotników i czasownika (tu: *polnej, stare, lipy*) uwarunkowane są końcówkami fleksyjnymi określanych przez nie rzeczowników i powtarzają zawarte w końcówkach rzeczowników informacje o liczbie elementów denotowanych przez te rzeczowniki zbiorów w odwzorowywanej sytuacji w rzeczywistości pozatekstowej: *polnej* mówi, że droga jest jedna, *stare* i *rosły*, że lip było więcej niż jedna. Informacja o liczbie wyrażana przez te końcówki jest z punktu widzenia informacji o rzeczywistości zbędna, jest jednak obligatoryjna, bo narzucana regułami składni języka polskiego. Taka redundancja systemowa nie jest przejawem braku ekonomii w systemie języka, pełni bowiem ważną funkcję właśnie jako wykładnik relacji syntagmatycznej i pozwala odbiorcy komunikatu na oszczędzenie wysiłku włożonego w ustalanie relacji syntagmatycznych w zdaniu i argumentów tych relacji. Dzięki istnieniu tego rodzaju wykładników relacji syntagmatycznych możliwe jest dokonanie poprawnych przyporządkowań przy zmianie szyku wyrazów tworzących zdanie, na przykład w zdaniu:

Stare przy drodze rosły lipy polnej.

Dokładne oddzielanie elementów funkcjonalnych od semantycznych możliwe jest w językach aglutynacyjnych i alternacyjnych. Do **języków aglutynacyjnych**, a więc mających ciągle wykładniki relacji i ciągle wykładniki semantyczne, należą na przykład języki węgierski i turecki. Najlepszą ilustracją tego jest omówiony niżej przykład z języka tureckiego.

W tureckim morfem *ev* odpowiada polskiemu morfemowi *dom*. Wykładnikiem relacji wyrażanej przez polski dopełniacz jest morfem *in*, relacji wyrażanej przez polski celownik morfem *e*, a relacji wyrażanej przez polski biernik morfem *i* (w językach aglutynacyjnych nie można oczywiście mówić o istnieniu przypadków fleksyjnych w takim sensie, jak w językach fleksyjnych, można tylko mówić o ich ekwiwalentach). Po dołączeniu tych morfemów relacyjnych do morfemu semantycznego otrzymamy formy:

ev (dom)
evin (domu)
eve (domowi)
evi (dom).

Te same morfemy są wykładnikami tych samych relacji w liczbie mnogiej, mającej własny wykładnik, w językach aglutynacyjnych, jak wiemy, nie ma bowiem kumulacji funkcji semantycznej i syntaktycznej. Wykładnikiem liczby mnogiej jest w tureckim morfem *ler*. Mamy więc:

evler (domy)

evlerin (domów) – wykładnik mnogości ma pierwszeństwo przed wykładnikiem relacji syntagmatycznej

evlere (domom)

evleri (domy).

Wykładnikiem relacji przynależności przedmiotu do nadawcy informacji jest w tureckim morfem *im*, mający pierwszeństwo przed innymi wykładnikami relacji syntagmatycznych (w języku polskim takim wykładnikiem jest forma zaimka *mój*, będąca morfemem luźnym, choć i w polskim wykładnikiem przynależności, ale nie do nadawcy i nie do odbiorcy informacji, może być morfemem związany, np. *-owy, -owa, -owe*, gdy rzeczownik oznaczający właściciela jest rodzaju męskiego, np. *ojcowa chalupa, soltysowe pole, ojcowy płaszcz, Jasiowe wolki* lub *-in, -ina, -ine*, gdy rzeczownik oznaczający właściciela jest rodzaju żeńskiego, np. *Zosi-ne buty, lalczyzna sukienka* – jak widzimy, w języku polskim wykładnik relacji przynależności łączy się z innym niż w języku tureckim argumentem tej relacji). Mamy więc odpowiednie formy tureckie:

evim (mój dom)

evimin (mojego domu)

evime (mojemu domowi)

evimi (mój dom)

i odpowiednio dla liczby mnogiej:

evlerim (moje domy)

evlerimin (moich domów)

evlerime (moim domom)

evlerimi (moje domy).

Konstrukcje aglutynacyjne stanowią więc struktury zlepkowe (stąd nazwa, od łac. *agglutinare* – sklejać) i choć daje to dość długie teksty, to jednak pozwala na uzyskanie jedno-jednoznacznego przyporządkowania między funkcją a jej wykładnikiem. Zasada aglutynacji, sklejanania morfemów, wykorzystana jest także w języku polskim, tak właśnie bowiem, oczywiście z punktu widzenia opisowego nie zaś historycznego, traktować można na przykład formy trybu przypuszczającego, którego wykładnik *-by* jest po prostu dołączany do również aglutynacyjnych form imiesłowu przeszłego, np.:

czytałbym, czytałbyś, czytałby, czytałyby, czytałybyśmy...

gdzie:

czyt – morfem określający rodzaj czynności,

a – morfem wskazujący, że chodzi o czynność,

l/l – wykładnik czasu przeszłego (imiesłowu przeszłego),

\emptyset (morfem zerowy), *a, o*, – morfemy będące wykładnikami rodzaju gramatycznego rzeczownika będącego wykładnikiem wykonawcy czynności, odpowiednio: męskiego, żeńskiego, nijakiego (w liczbie pojedynczej),

i, y – morfemy będące wykładnikami rodzaju męskoosobowego lub niemęskoosobowego (w liczbie mnogiej),

by – wykładnik trybu przypuszczającego,

m, ś, θ – wykładniki osoby czasownika dla liczby pojedynczej

-śmy, -ście, θ – wykładniki osoby dla liczby mnogiej.

Takie morfemy aglutynacyjne w języku polskim – w przeciwieństwie do morfemów fleksyjnych – są elementami ruchomymi i mogą się łączyć z różnymi wyrazami zdania, por.:

Ty nie chodziłabyś do lasu i nie zbierałabyś grzybów.

Ty byś nie chodziła do lasu i nie zbierała grzybów.

Ty do lasu byś nie chodziła i nie zbierała grzybów.

Ty do lasu byś nie chodziła i grzybów byś nie zbierała.

ale

Ty nie chodzisz do lasu i nie zbierasz grzybów.

Językami alternacyjnymi są języki semickie, na przykład arabski, w którym wykładnikami morfemów semantycznych są spółgłoski, natomiast wykładnikami relacji (uzupełniając je, bo stanowiące ośrodek sylaby) – samogłoski. Mamy tu więc wykładniki nieciągłe, lecz wyraźnie różniące się od siebie pod względem fonetycznym i funkcjonalnym. Najlepiej zilustrują to różne formy uzyskane przez rozwinięcie pierwiastka leksykalnego *g-t-l-*, który wskazuje, że forma wiąże się ze znaczeniem zabijania:

gatila – został zabity,

gatala – zabił,

ya-gtulu – zabijał,

ya-gtalu – był zabijany,

ugtul – zabij,

gatl – zabójstwo,

gatil – zabójca

Alternacje te, jak widzimy, służą nie tylko do wyrażania relacji syntagmatycznych, ale również do modyfikacji znaczenia, pełnią więc funkcję słowotwórczą. Poza alternacjami w językach naturalnych zaliczanych do tego typu syntaktycznego wykładnikami relacji i modyfikacji znaczenia bywają nie tylko decydujące o typie syntaktycznym infiksy, ale także prefiksy i sufiksy. Teoretycznie największą zaletą naturalnych języków alternacyjnych jest to, że dają stosunkowo krótkie teksty, co umożliwia podwójne wykorzystanie samogłosek: jako ośrodka sylaby i jako wykładnika relacji. Zaleta ta dlatego jest zaletą przede wszystkim języków

naturalnych, bo ich głównym subkodem jest subkod akustyczny, wymagający samogłosek jako nośników fali głosowej, co nie ma oczywiście żadnego znaczenia w wypadku języków istniejących tylko w postaci graficznej (a takimi, jak dotąd, są języki informacyjno-wyszukiwawcze).

Aby odwzorować w komunikacie jakąś sytuację, należy w danym wycinku rzeczywistości wyróżnić interesujące nas obiekty oraz relacje zachodzące między tymi elementami. Obiekty połączone daną relacją to argumenty tej relacji. Argumenty te w wypadku relacji wieloargumentowej pełnią różną rolę. Taka relacja wieloargumentowa może być rozłożona na relacje dwuargumentowe, zależnie od tego, które z członów relacji są dla nadawcy komunikatu relewantne. Na przykład sytuacja rysowania może być rozłożona na następujące relacje dwuargumentowe, komunikowane na przykład w zdaniach:

Jan rysuje portret.

Jan rysuje piórkiem.

Jan rysuje tuszem.

Jan rysuje Marię.

Jan rysuje na brystolu.

Relacja rysowania jest więc w rzeczywistości relacją wieloargumentową, ale komunikujący ją predykat (czasownik *rysować*) jest predykatem składniowo dwuargumentowym, bo wszystkie powyższe zdania są w języku polskim poprawne (a być może tylko jednoargumentowym, jeśli za poprawne uznamy zdanie *Jan rysuje*).

Struktury syntaktyczne języka naturalnego są prostsze niż struktura odwzorowywanej przez nie rzeczywistości, a zdanie odwzorowujące wszystkie elementy powyższej rzeczywistości byłoby niepoprawne:

* *Jan rysuje portret Marię piórkiem tuszem na brystolu.*

W językowym komunikacie powyższa relacja musiałaby być rozłożona na pewne składowe, zależnie od tego, jakie informacje uzna nadawca tego komunikatu za relewantne dla jego adresata. Pełne przedstawienie takich wieloargumentowych relacji wymaga określenia roli jej argumentów, dla potrzeb formalizacji systemu języka naturalnego konieczne jest także przypisanie im realizujących je w danym języku struktur składniowych, a więc podanie wykładników odpowiednich relacji syntagmatycznych. W ten sposób tworzy się całe **scenariusze sytuacji** i realizujących je wykładników językowych. W rozpatrywanej wyżej sytuacji rysowania argumenty tej relacji pełnią następujące role:

- Jan jest wykonawcą czynności zwanej rysowaniem,
- Portret jest wynikiem czynności rysowania,
- Piórko jest narzędziem, za pomocą którego wykonywana jest czynność rysowania,
- Tusz jest materia pomocniczą wykorzystywaną w czynności rysowania,
- Brystol jest materia, na której rysowany jest portret,
- Maria jest obiektem odwzorowywanym w procesie rysowania.

Do pełnego określenia sytuacji konieczna jest jeszcze jej lokalizacja w czasie i przestrzeni, np.:

Jan rysuje w ogrodzie.

Jan rysuje rano.

Językowe wykładniki tej lokalizacji nie są jednak argumentami specyficznymi dla predykatu *rysować*, mogą być bowiem jako lokalizatory sytuacji umieszczone w każdym komunikacie (jeśli taka informacja jest relewantna) i dołączone do każdej struktury predykatowo-argumentowej.

Choć sytuacji w rzeczywistości jest nieskończenie wiele, to wyróżnić w nich można powtarzające się schematy relacji i relewantnych ról argumentów tych relacji. Zgodnie z zasadami konstrukcji optymalnego systemu językowego schematy takie powinny być odwzorowywane przez język za pomocą wykładników gramatycznych tworzących schematy składniowe (**struktury predykatowo-argumentowe**). I choć wykładniki ról i relacji są zależne od konkretnego języka naturalnego, to ich **struktura głęboka** powinna być taka sama, bo takie same są schematy postrzegania rzeczywistości przez ludzi, a więc i takie same są relewantne elementy strukturalizujące informację. Te niezmienniki sytuacyjne odwzorowywane w strukturach syntaktycznych każdego języka naturalnego (choć w różny sposób) w teorii lingwistycznej nazwano **rolami semantycznymi** lub **przypadkami głębokimi** (termin amerykańskiego językoznawcy Charlesa J. Fillmore'a). Ich liczba nie została do końca ustalona, ale najważniejsze z nich to:

- *agentivus* – wyraża agensa, czyli wykonawcę czynności,
- *benefactivus* – wyraża istotę żywą, ze względu na którą odbywa się czynność,
- *comitativus* – wyraża stosunek współuczestnictwa (towarzyszenia),
- *dativus* – wyraża istotę żywą, która znajduje się w danej sytuacji w jakimś stanie,
- *factitivus* – wyraża rezultat czynności (przedmiot),
- *instrumentalis* – wyraża narzędzie,
- *locativus* – wyraża miejsce,
- *pacjens* – wyraża obiekt, na którym dokonywana jest czynność.

Listę tę można uzupełnić jeszcze o takie role jak przyczyna, cel lub istotne w lokalizacji punkt wyjścia i punkt dojścia.

Językowe wykładniki przypadków głębokich mogą mieć różny charakter: może to być szyk, mogą to być morfemy izolowane, a więc przyimki lub związane: cząstki aglutynacyjne, końcówki fleksyjne, morfemy alternacyjne.

W językach informacyjno-wyszukiwawczych przypadkom głębokim odpowiadają wskaźniki roli, mające wykładniki w postaci oddzielnych wyrażen, na przykład w niektórych językach deskryptorowych lub wskaźniki wyrażane za pomocą szyku i przewidziane w tzw. formule zdaniowej, jak to jest w Klasyfikacji Dwukropkowej Ranganathana, gdzie tzw. formuła fasetowa:

PMEST

przewiduje układ wyrażeń reprezentujących poszczególne kategorie (fasety):

P – Personality (osoba)

M – Mater (materia)

E – Energy (energia)

S – Space (przestrzeń)

T – Time (czas).

Wśród syntaktycznych typów języków zwykle wymienia się jeszcze **języki polisyntetyczne**, których cechą charakterystyczną jest ogromna liczba kombinacji morfemów gramatycznych, co daje w wyniku bardzo długie ciągi-wyraży, często odpowiadające całym zdaniom innych języków, a więc mające wartość informacyjną (wykładnik predykatywności w wyrazie). Do takich języków należą języki eskimoskie, na przykład eskimoskie:

Imeramiutartinginakoratlerkamignayagtut

wyrażenie-zdanie znaczy: poszli tam, gdzie zwykli łowić norki (przykład z książki *Ognista kula. Legendy, baśnie i bajki eskimoskie*. oprac. Jacek Machowski),

takusariartorumagaluarnerpa

znaczy: czy myślicie, że naprawdę ma on zamiar iść się tym zająć?

auksautiesarsiniarpuna

znaczy: szukam, by pozyskać coś zdatnego na sznur rybacki, co dosłownie można by przetłumaczyć jako:

rybanarzędziestosownepozyskanie szukaniemoje.

Takie połączenie dużej liczby morfemów w jedną jednostkę wyrazową może występować również w językach innych typów, na przykład w polskim:

Konstantynopolitańczykowieczka,

ale w tym wypadku oznaczają one tylko jeden przedmiot, nie mają więc wartości informacyjnej (brak elementu predykacji), podczas gdy w językach polisyntetycznych niosą pełną informację.

Jeden opis sytuacji, jedno zdanie, przedstawia często więcej niż tylko jedną relację, zawiera więc kilka wykładników relacji i odpowiednią liczbę powiązanych nimi argumentów. Równie często bywa tak, że w tej samej relacji R, na przykład do przedmiotu x, znajduje się kilka różnych przedmiotów: a, b, c, ..., n, np.:

Na stole leżą książki, zeszyty, notatki i wykresy.

gdzie:

x – stół

a, b, c, d, – książki, zeszyty, notatki, wykresy

wykładnik relacji R – *na ...-e leżą.*

W takich sytuacjach trzeba koniecznie w jakiś sposób oznaczyć, które wykładniki wskazują argumenty danej relacji.

W wypadku języków informacyjno-wyszukiwawczych służą do tego tzw. **wskaźniki więzi**. Mechanizm działania wskaźników więzi przedstawia Černyj na następującym przykładzie:

„Założmy, że musimy zaindeksować dokument traktujący o eksporcie z Bułgarii jarzyn i owoców oraz o imporcie do tego kraju stali, żeliwa, maszyn i ropy naftowej. Charakterystyka wyszukiwawcza takiego dokumentu z zastosowaniem wskaźników więzi będzie miała postać następującą:

BUŁGARIA – A, B
EKSPORT – A
JARZYNY – A
OWOCE – A
IMPORT – B
STAL – B
ŻELIWO – B
MASZYNY – B
ROPA NAFTOWA – B

W wyniku zastosowania wskaźników więzi tekst tej charakterystyki wyszukiwawczej rozpada się na dwa quasi-zdania:

A-BUŁGARIA, EKSPORT, JARZYNY, OWOCE
B-BUŁGARIA, IMPORT, STAL, ŻELIWO, MASZYNY, ROPA NAFTOWA.

Zapobiega to tworzeniu się fałszywych związków współrzędnych między deskryptorami grupy A i B, powodujących wydawanie nieodpowiednich dokumentów”.

W innych językach takie wskaźniki więzi mogą mieć inny charakter. W językach sztucznych, na przykład w matematyce, rolę taką pełnią nawiasy:

$$(a + b) \cdot (c - d) = x$$

podobnie jest w języku logiki, np.:

$$[p \Rightarrow (q \Rightarrow r)] \Leftrightarrow [q \Rightarrow (p \Rightarrow r)]$$

W językach naturalnych takie wskaźniki więzi mogą mieć najróżniejszą postać. Omówimy je na przykładzie wskaźników więzi występujących w języku polskim. Mamy tu następujące podstawowe możliwości:

- I. Wskaźniki więzi nie stanowią elementów tekstu i wykorzystuje się w tym celu szyk zdania, umieszczając bezpośrednio wiążące się ze sobą wyrazy obok siebie (wykładniki pozycyjne).
- II. Wskaźniki więzi stanowią elementy tekstu. Tu mamy z kolei następujące możliwości:
 - II.1. Wskaźniki więzi dołączone są do wszystkich łączących się ze sobą wyrazów zdania, typ $R_1 a_1 b_1 c_1 \dots n_1$ gdzie:
 - R – wykładnik relacji
 - a, b, c, ..., n – wykładniki argumentów,
 - 1 – wskaźnik więzi.

II.2. Wskaźnikami oznacza się tylko argumenty danej relacji, np.:

R'a₁
R"a₂

Pierwszy z wymienionych wyżej sposobów wskazywania połączeń składniowych w języku polskim występuje stosunkowo rzadko. Przykładami są tu połączenia określeń oceniających *ekstra*, *super* z wykładnikami leksykalnymi przedmiotów lub faktów ocenianych, np.:

super modna, ekstra modna
super sukienka, ekstra sukienka
super dziewczyna, ekstra dziewczyna
wygląda super, wygląda ekstra,

co wyraźnie widać przy porównaniu użycia tego typu związków z użyciem związków, w których mamy niepozycyjne wskaźniki więzi, przy których można zmieniać szyk wypowiedzenia bez zmiany jego treści. Weźmy dla przykładu zdania:

W ekstra sukience dziewczyna wygląda interesująco.
W super sukience dziewczyna wygląda interesująco.
W modnej sukience dziewczyna wygląda interesująco.

W tym ostatnim zdaniu wyrazy *modna* i *sukienka* połączone są ze sobą tzw. **związkiem zgody** (typ II.1), występują więc w tej samej formie: liczbie (por. *w modnej sukience* – *w modnych sukienkach*), przypadku (por. *w modnej sukience* – *modna sukienka*), rodzaju gramatycznym (por. *modnej sukience* – *modnym sweterku*), dzięki czemu wyrazy te nie muszą stać obok siebie, żeby było wiadomo, że *modnej* jest określeniem *sukience*, możemy więc dowolnie zmieniać ich wzajemne położenie, bez zmiany znaczenia i naruszenia związków składniowych, np.:

W modnej sukience dziewczyna wygląda interesująco.
W sukience modnej dziewczyna wygląda interesująco.
W sukience dziewczyna modnej wygląda interesująco.
W sukience dziewczyna wygląda modnej interesująco.

W dwóch pierwszych zdaniach tego typu zmiana miejsca określeń *super* i *ekstra* spowoduje oczywiście zmianę znaczenia zdania i zmianę występujących w nim związków składniowych, gdyż wyrazy te stanowią określenie wyrazów, przed którymi stoją, por.:

W ekstra sukience dziewczyna wygląda interesująco.
W super sukience dziewczyna wygląda interesująco.
W sukience ekstra dziewczyna wygląda interesująco.
W sukience super dziewczyna wygląda interesująco.
W sukience dziewczyna wygląda ekstra interesująco.
W sukience dziewczyna wygląda super interesująco.

W analogiczny sposób łączą się z innymi członami zdania wyrazy modalne. Przykłady wraz z interpretacją cytuję z *Praktycznej stylistyki* A. i P. Wierzbickich:

- „1.A. *Ojciec nie wyjechał do Krakowa.*
B. *Ojciec wyjechał nie do Krakowa.*
C. *Nie ojciec wyjechał do Krakowa.*
- 2.A. *Jaś dał Małgosi tylko kwiaty.*
B. *Jaś dał tylko Małgosi kwiaty.*
C. *Tylko Jaś dał Małgosi kwiaty.*
- 3.A. *Nawet Franio spóźnił się wtedy na kolację.*
B. *Franio spóźnił się nawet wtedy na kolację.*
C. *Franio spóźnił się wtedy nawet na kolację.*

Zdanie 1.A znaczy: nie jest prawdą, że wyjazd ojca do Krakowa miał miejsce. Zdanie 1.B znaczy: wyjazd ojca miał miejsce, ale nie jest prawdą, że punktem docelowym podróży był Kraków. Zdanie 1.C znaczy: wyjazd do Krakowa miał miejsce, ale nie jest prawdą, że osobą, która wyjechała, był ojciec.

Zdanie 2.A znaczy: Jaś dał Małgosi kwiaty, nie dał jej nic innego (dziwię się, że nie dał jej nic innego). Zdanie 2.B znaczy: Jaś dał kwiaty Małgosi, nie dał kwiatów nikomu innemu (dziwię się, że nie dał nikomu innemu). Zdanie 2.C znaczy: Małgosia dostała kwiaty od Jasia, nie dostała od nikogo innego (dziwię się, że nie dostała od nikogo innego).

Zdanie 3.A znaczy: inne osoby spóźniły się wtedy na kolację i, co mnie dziwi, Franio spóźnił się także. Zdanie 3.B znaczy: Franio często spóźnia się na kolację i, co mnie dziwi, spóźnił się także wtedy. Zdanie 3.C znaczy: Franio spóźnił się wtedy kilka razy i, co mnie dziwi, spóźnił się także na kolację.

Wyrazy takie, jak: *nie*, *tylko*, *nawet* to wyrazy informujące o stosunku osoby mówiącej do pewnych faktów: mówiący przeczy (nie wierzy), dziwi się, że coś się stało, dziwi się, że coś się nie stało. Wyrazy te nie tyle wzbogacają informację o jakiś dodatkowy element, ile zmieniają w ogóle kierunek interpretacji całego zdania. Z przykładów widać: miejsce takich wyrazów w zdaniu jest szczególnie ważne; zmiana pozycji w zdaniu może spowodować istotną zmianę przekazywanej informacji, bowiem właśnie miejsce zajmowane w zdaniu przez wyrazy tego typu wskazuje, co mianowicie jest przedmiotem wątpienia, niedowierzenia, zdziwienia”.

Przykład rozpatrywany przez Černego należy do typu II.1. W ten sposób w języku polskim wskazuje się związek podmiotu (argumentu) z orzeczeniem (wykładnik relacji), które przy orzeczeniu w czasie teraźniejszym muszą posiadać wykładniki tej samej kategorii liczby, a w czasie przeszłym mają, a w przyszłym niedokonanym mogą mieć ponadto wykładnik tej samej kategorii rodzaju gramatycznego – jest to tzw. **związek zgody**, o którym już wspominaliśmy, np.:

Mój kolega studiuje na uniwersytecie.

Moja koleżanka studiuje na uniwersytecie.

Moi koledzy studiują na uniwersytecie.

Moje koleżanki studiują na uniwersytecie.

gdzie mamy zgodność co do liczby, oraz:

Moja koleżanka studiowała na uniwersytecie.

Mój kolega studiował na uniwersytecie.

Moje koleżanki studiowały na uniwersytecie.

Moje koleżanki będą studiowały na uniwersytecie.

Moje koleżanki będą studiować na uniwersytecie.

gdzie mamy zgodność także rodzaju gramatycznego.

Związek zgody łączy w zdaniu również rzeczownik z niektórymi typami jego określeń (przydawek) – przykłady ilustrujące ten typ wykładników więzi podaliśmy wyżej.

Dzięki tego rodzaju wskaźnikom więzi możemy dokonać poprawnych połączeń w zdaniach typu:

Spotkałam męża koleżanki, która studiowała ze mną na uniwersytecie.

Spotkałam męża koleżanki, który studiował ze mną na uniwersytecie.

W języku polskim można stosunkowo swobodnie operować szykiem, a mimo to dokonywać poprawnych połączeń wyrazów w zdania właśnie dzięki temu, że istnieje w nim kategoria rodzaju gramatycznego, która powoduje rozbitcie wszystkich rzeczowników na trzy klasy w liczbie pojedynczej (męska, żeńska i nijaka) oraz na dwie w liczbie mnogiej (męskoosobowa i niemęskoosobowa – w tym wypadku jest to oczywiście nie tylko kategoria formalna, ale i semantyczna), których wykładniki dołączane są w zależności od rodzaju gramatycznego rzeczownika do łączących się z nimi określeń: przymiotników, liczebników porządkowych, imiesłówów przymiotnikowych, niektórych zaimków, np. dzierżawnych, względnych i imiesłówów przeszłych. Należy przy tym pamiętać, że czym innym jest kategoria liczby rzeczownika, gdzie pełni ona funkcję semantyczną, wskazując, czy mówimy o jednym tylko desygnacie rzeczownika, czy o kilku, czym innym zaś kategoria liczby czasownika czy przymiotnika, przy których jest ona jedynie wskaźnikiem więzi. Gdyby nie ta funkcja, wszystkie wykładniki zależnych związków zgody, które powtarzają tę samą informację (o liczbie desygnatów określanego rzeczownika), byłyby oczywiście redundantne i można by je uznać za przejaw nieekonomiczności systemu językowego, np.:

Moja niestrudzona, stara już, sterana życiem, ale zawsze promienna i uśmiechnięta matka do późna w nocy siedziała schylona nad robotą.

Wykładnikiem więzi może być wreszcie określona postać wyrazu łączącego się z innym wyrazem (II.2), konotowana (narzucana) przez jeden z połączonych daną relacją wyrazów. Tak jest w wypadku tzw. **związku rządu**, a więc, gdy na przykład określony czasownik konotuje określoną formę łączącego się z nim rzeczownika: *oglądać* coś, ale *przyglądać się* czemuś, *spoglądać* na coś, *rozglądać się* za czymś, np.:

Anna oglądała książkę.
Anna przyglądała się książce.
Anna spoglądała na książkę.
Anna rozgląda się za książką.

Podobnie jest w przypadku nieodmiennych przysłówków stanowiących określenie czasowników, co dobrze widać przy porównaniu z przymiotnikowymi, a więc odmiennymi określeniami rzeczowników, np.:

Janek pilnie się uczy.
Janek jest dobrym uczniem.

co umożliwia dokonanie poprawnych połączeń, na przykład w zdaniu:

Janek pilnie jako dobry uczeń się uczy.

Przeprowadzona w tym rozdziale klasyfikacja syntaktyczna języków może być zastosowana przy opisie języków informacyjno-wyszukiwawczych, choć, podobnie jak języki naturalne, często nie są one pod tym względem typami jednorodnymi. A oto przykłady poszczególnych **typów syntaktycznych języków informacyjno-wyszukiwawczych**.

Typowym **językiem pozycyjnym** jest język haseł przedmiotowych, w którym wykładnikiem kategorii tematu i określnika jest ich pozycja w haśle przedmiotowym, będącym zdaniem tego języka: temat stoi na pozycji pierwszej, określnik lub określniki na pozycjach dalszych, np.:

Filozofia – historia	co znaczy: historia filozofii,
Historia – filozofia	co znaczy: filozofia historii,
Nauka – organizacja	co znaczy: organizacja nauki,
Organizacja – nauka	co znaczy: nauka o organizacji.

Przykłady wykładników izolujących można znaleźć choćby w UKD – są nimi wskaźniki relacji „:” (dwukropek) i „+” (plus).

Charakter **aglutynacyjny** mają w językach informacyjno-wyszukiwawczych na przykład wskaźniki roli stosowane w językach deskryptorowych:

TRUCIZNA (B)	gdzie B wskaźnik roli „być produktem”,
ROŚLINY (E)	gdzie E wskaźnik roli „być środowiskiem, otoczeniem, podłożem”

lub używane w tych językach wskaźniki więzi, np.:

MIEDŹ₁ OŁÓW₂ POKRYCIE₂ RURY₁

Wśród znanych języków informacyjno-wyszukiwawczych istnieją również **języki alternacyjne**. Jako przykład wymieniłem kod semantyczny Perry’ego i Kenta, gdzie morfemy leksykalne o charakterze nieciągłym modyfikowane są przez infiksy będące wykładnikami dziesięciu kategorii gramatycznych (A, E, I, O, U, Q, W, X, Y, Z), np. A – wykładnik relacji inkluzji – PAST – owad, PAST.004 – mucha, I – wykładnik „być częścią” – BITN – część rośliny, BITN.013 – kwiat,

BITN.026 – korzeń, E – wykładnik „być produktem z materiału...” – BETN.00 – cygara.

Jak widać z tego przykładu, w językach informacyjno-wyszukiwawczych – tak samo jak i w językach naturalnych – alternacja służy do różnego typu modyfikacji znaczenia leksykalnego, nie tylko do oznaczania relacji syntagmatycznych.

Wśród istniejących języków informacyjno-wyszukiwawczych za **języki fleksyjne** można by uznać klasyfikacje fasetowe, w których wykładniki kategorii semantycznych i kategorii syntaktycznych są nierozdzielne.

Językiem polisyntetycznym, a więc takim, w którym w jednym wyrazie realizowany jest informacyjny schemat całego zdania, jest wśród języków informacyjno-wyszukiwawczych Uniwersalna Klasyfikacja Dziesiętna. Na przykład: 669.1.002.2(47+57)(083.74)=161.1 symbol rozwinięty o odpowiedniku słownym: Normy obowiązujące przy produkcji stali stopowej w ZSRR, dokument w języku rosyjskim; symbol ten utworzony jest z następujących morfemów:

669.1	stale stopowe
.002.2	produkcja wytwarzanie – podział wspólny punktu widzenia
(47+57)	ZSRR – podział wspólny miejsca
(083.74)	normy – podział wspólny formy
=161.1	w języku rosyjskim – podział wspólny języka.

Literatura

- Chmielewska-Gorczyca E.: *Relacje syntagmatyczne w językach informacyjno-wyszukiwawczych*. „Zagadnienia Informatyki Naukowej” 1983, nr 2.
- Grzegorzczak R.: *Wykłady z polskiej składni*. Warszawa 1998.
- Karolak S.: *Zarys składni ogólnej*. Warszawa 1972.
- Majewicz A. T.: *Języki świata i ich klasyfikowanie*. Warszawa 1989 (rozdz. 3.3. *Typologia syntaktyczna*).
- Milewski T.: *Językoznawstwo*. Warszawa 1965.
- Poetyło M. (red.): *Gramatyka języków deskryptorowych na tle gramatyk innych języków informacyjno-wyszukiwawczych*. Warszawa 1972.
- Sadowska J.: *Gramatyka pozycyjna w języku hasel przedmiotowych*. „Zagadnienia Informatyki Naukowej” 1983, nr 1.
- Saloni Z., Świdziński M.: *Składnia współczesnego języka polskiego*. Warszawa 1985 (rozdz. II. *Co to jest zdanie?*).

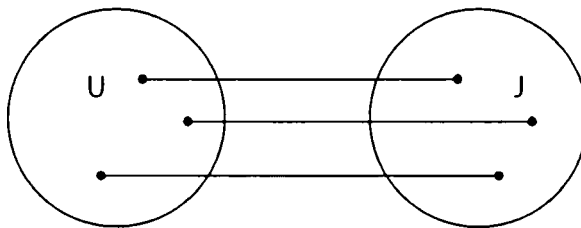
X

System semantyczny. Odwzorowanie rzeczywistości w języku

Ze stosunkowo niedużej liczby fonemów każdy język naturalny tworzy dużą liczbę morfemów, z których każdemu na zasadzie czystego przypadku, mającego już jednak moc uzusu, przyporządkowany jest jakiś wycinek rzeczywistości pozajęzykowej, i jeszcze większą liczbę wyrazów, często mających już znaczenie motywowane. Przyporządkowanie to nie jest jednak przyporządkowaniem jedno-jednoznacznym, bardzo często bywa przyporządkowaniem wielo-wieloznacznym. Wyjaśnijmy to sobie na przykładach.

Otoczająca nas rzeczywistość, czyli cały świat, stanowi pewien zbiór elementów U (uniwersum). Zbiór ten jest przedstawiany (komunikowany) za pomocą pewnego swojego podzbioru J będącego językiem. Aby taka reprezentacja, czyli zastępowanie elementów U przez elementy J była w ogóle możliwa, elementom U muszą być przyporządkowane elementy J , czyli zbiór U musi być odwzorowany w zbiór J . Możliwe tu są następujące warianty przyporządkowania: (zbiory U i J narysujemy oddzielnie, pamiętając jednak, że J zawarte jest w U).

1. Przyporządkowanie jedno-jednoznaczne



Każdemu elementowi należącemu do U przyporządkowany jest tylko jeden element należący do J , i odwrotnie – każdemu elementowi należącemu do J przyporządkowany jest tylko jeden element należący do U . Takie przyporządkowanie (odwzorowanie zbioru U w zbiór J) nazywamy jedno-jednoznacznym (wzajemnie jednoznacznym).

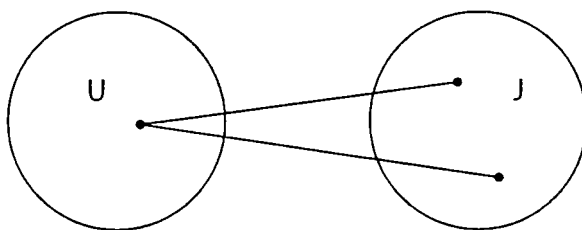
A oto przykłady, gdy zbiór J jest językiem polskim:

- przedmiotowi $\text{okulary} \in U$ przyporządkowany jest znak $\text{okulary} \in J$, a znakowi $\text{okulary} \in J$ przyporządkowany jest przedmiot $\text{okulary} \in U$;

- przedmiotowi $\text{olówek} \in U$ przyporządkowany jest znak $\text{olówek} \in J$, a znakowi $\text{olówek} \in J$ przyporządkowany jest przedmiot $\text{olówek} \in U$.

Uwaga: zapis taki należy traktować jako skrótowy, pamiętamy bowiem, że chodzi tu o klasę obiektów rzeczywistości pozajęzykowej (zbiór w sensie dystrybucyjnym), a nie o pojedynczy obiekt – te bowiem mają swoje wykładniki językowe (nazwy własne) tylko wyjątkowo; podobnie znak należy tu rozumieć jako element systemu językowego, któremu w tekście odpowiadają różne realizacje.

2. Przyporządkowanie jedno-wieloznaczne



Jednemu elementowi należącemu do U przyporządkowany jest więcej niż jeden element J , podczas gdy jednemu elementowi J przyporządkowany jest tylko jeden element U . Takie przyporządkowanie (odwzorowanie) zbioru U w zbiór J nazywamy jedno-wieloznacznym.

A oto przykłady, gdy J jest językiem polskim:

- przedmiotowi $\text{łódź} \in U$ przyporządkowany jest znak $\text{sztandar} \in J$ oraz znak $\text{chorągiew} \in J$, natomiast znakowi $\text{sztandar} \in J$ przyporządkowany jest tylko przedmiot $\text{łódź} \in U$, a znakowi $\text{chorągiew} \in J$ przyporządkowany tylko przedmiot $\text{łódź} \in U$;
- przedmiotowi $\text{samochód} \in U$ przyporządkowany jest znak samochód oraz znak $\text{auto} \in J$, natomiast znakowi $\text{samochód} \in J$ przyporządkowany jest tylko przedmiot $\text{samochód} \in U$, a znakowi $\text{auto} \in J$ przyporządkowany jest tylko przedmiot $\text{samochód} \in U$.

W wypadku takiego przyporządkowania stosunek między elementami J (wyrażeniami) nazywamy **synonimią**, natomiast elementy J (wyrażenia) **synonimami** w języku J . Synonimia jest relacją, która zachodzi między jednostkami znaczącymi języka – mamy więc synonimiczne morfemy, wyrazy, zdania, teksty, a także synonimiczne struktury składniowe.

Synonimy są przejawem pewnej nadmiarowości (redundancji) w systemie języka, **wariantami językowymi** na poziomie jednostek znaczących. W wypadku synonimii, chcąc poinformować o danym elemencie rzeczywistości, trzeba wybrać jedno z wielu odwzorowujących go wyrażen języka. Nie zawsze jednak wyrażenia te mogą być przedmiotem arbitralnego wyboru, czasami bowiem bywa tak, że poprzednio dokonany przez nadawcę informacji wybór jakiegoś wyrażenia pociąga za sobą konieczność wyboru właśnie tego, a nie innego spośród synonimów lub też nie pozwala jakiegoś synonimu wybrać, zmniejszając tym samym repertuar wyrażen, spośród których można wybierać. Synonimy, spośród których można dowolnie wybierać, nazywane są **synonimami fakultatywnymi**, takie, których

wybór narzucony jest kontekstem, w którym mają być umieszczone, **synonimami kombinatorycznymi** (czasami nazywane są synonimami kontekstowymi, my jednak ten termin zarezerwujemy dla innego jeszcze rodzaju synonimii).

Synonimami fakultatywnymi są na przykład wyrażenia: *bratowa* = *żona brata*, *żona* = *małżonka*, *samochód* = *auto*, natomiast **synonimami kombinatorycznymi** morfemy *-a* i *-u* będące końcówkami dopełniacza liczby pojedynczej rzeczowników rodzaju męskiego, np. *człowieka*, *konia*, ale *stołu*, *zeszytu*. Czyli: jeżeli wybraliśmy *człowiek-* lub *koń-*, to musimy wybrać *-a*, jeżeli *stół-*, *zeszyt-*, to musimy wybrać *-u*, nie są bowiem poprawnymi formami dopełniacza liczby pojedynczej wyrazy **człowieku*, **koniu*, **stola*, **zeszyta*. Synonimami kombinatorycznymi są też morfemy *człowiek-* i *ludź-* oraz *rok-* i *lat-*. *Człowiek-* oraz *rok-* uwarunkowane są użyciem liczby pojedynczej, natomiast *ludź-* i *lat-* – użyciem liczby mnogiej. Takie synonimiczne morfemy nazywane bywają **allomorfami** albo **morfemami supletywnymi**. A oto przykłady **synonimii kombinatorycznej**, z ograniczoną przez kontekst możliwością wyboru: *jasne włosy* = *blond włosy*, ale tylko *jasne oczy*; *czarny koń* = *kary koń*, ale tylko *czarne futro*; *brązowe oczy* = *piwne oczy*, ale tylko *brązowe włosy*.

Oprócz synonimii kombinatorycznej należałoby wyróżnić jeszcze jeden wypadek synonimii – czystej **synonimii kontekstowej**, bywa bowiem tak, że dwa elementy języka, które w izolacji nie są synonimami, stają się nimi w sąsiedztwie pewnych innych elementów języka. I tak, na przykład, nie są synonimami wyrazy *dużo* i *słono*, ale stają się nimi w połączeniu z wyrazem *zapłacić*, gdyż *słono zapłacić* znaczy tyle, co *dużo zapłacić*. A oto inne przykłady tego rodzaju synonimii:

- synonimy kontekstowe czasownika *znajdować się*: *rzeka płynie pod lasem*, *dom stoi na wzgórzu*, *miasto leży nad rzeką*;
- synonimy przymiotnika *duży*: *ulewny deszcz*, *ciężka choroba*, *słona cena*, *wysoka cena*, *głęboka czerń*.

Zarówno warianty kombinatoryczne, jak i fakultatywne występują w języku naturalnym na wszystkich poziomach jednostek złożonych. Stosunek synonimiczny łączy w języku nie tylko jednostki tego samego poziomu, ale również jednostki różnych poziomów zawierających jednostki znaczące, a także jednostki różnych kategorii gramatycznych. A oto przykłady takiej synonimii:

tyfus = *dur brzuszny*;

rodzice = *ojciec i matka*;

wuj = *brat matki*;

zięć = *mąż córki*;

Janek lubi czytać kryminały = *Janek chętnie czyta kryminały*;

Nauczyciel powtórzył pytanie = *Nauczyciel zapytał po raz drugi*;

Anna cały wieczór spacerowała Marszałkowską = *Anna cały wieczór spacerowała po Marszałkowskiej*

Janek odrobił lekcje i poszedł na sanki = *Janek po odrobieniu lekcji poszedł na sanki* = *Janek odrobiwszy lekcje poszedł na sanki*.

Synonimami są zdania w stronie czynnej i zdania w stronie biernej, np.:

Kolumb odkrył Amerykę = Ameryka została odkryta przez Kolumba,

gdyż stanowią opisy jednej i tej samej sytuacji, choć zdania te różnią się od siebie tzw. **strukturą tematyczno-rematyczną**, co innego jest w nich bowiem nową informacją, a mówiąc bardziej po prostu są one zdaniami o czym innym: pierwsze jest zdaniem o Kolumbie i może stanowić odpowiedź na pytanie: *Co zrobił (czego dokonał) Kolumb?*, drugie natomiast jest zdaniem o Ameryce. O takich synonimicznych zdaniach mówi się, że mają one jednakową **strukturę głęboką** i różnią się tylko **strukturą powierzchniową**, czyli mają tę samą treść, lecz różną formę.

Synonimy, choć znacznie zwiększają liczbę jednostek języka, nie stanowią wyłącznie zbytecznego balastu obciążającego pamięć użytkownika, one to bowiem umożliwiają pełnienie przez język funkcji ekspresywnej, impresywnej i estetycznej (poetyckiej), one też stanowią często dodatkowe źródło informacji o nadawcy. Klasyfikacja synonimów jest sprawą trudną, zazwyczaj dokonywane podziały krzyżują się, a granice między wyróżnionymi grupami synonimów są płynne. Taką ramową **klasyfikację synonimów** przytaczamy za *Stylistyką polską* Haliny Kurkowskiej i Stanisława Skorupki.

„Synonimami stylistycznymi sensu stricto są wyrazy bliskoznaczne różniące się przede wszystkim barwą (...) W grę wchodzi tu najrozmaitsze rodzaje barw. Można więc mówić:

- o synonimach chronologicznych, czyli czasowych, wtedy gdy zestawiamy jakiś wyraz przestarzały i słowo mające w języku współczesnym pełną żywotność (np. *pryncypał* i *szef*; *ongi* i *niegdyś*);

- o synonimach zróżnicowanych terytorialnie, gdy zestawiamy np. prowincjonalizm i jego odpowiednik ogólnopolski (wielkopolski *tytka* – ogólnopolski *torebka* (papierowa), wschodniopolski *światły* – ogólnopolski *jasny* itp.);

- o synonimach środowiskowych, tzn. o bliskoznacznych odpowiednikach zaczerpniętych z poszczególnych gwar środowiskowych (np. gwarowe warszawskie *posuwać* – ogólnopolskie *jechać*, *biegać*, podobnie *opylić* – *sprzedać*)

- o synonimach stylowych (np. bardziej książkowe *wieść* i bardziej potoczne *prowadzić*, neutralne stylistycznie *granice* i archaicznie podniosłe a przez to ograniczone do stylu artystycznego czy też stylu patetycznych przemówień *rubieżę*; obojętne pod względem zabarwienia *mąż* i oficjalne *małżonek*);

- wreszcie o synonimach zróżnicowanych pod względem emocjonalnym (np. *łazik*, *włóczęga*, *włóczykij*, *obieżyświat*; *żoldak*, *wojak*; *ateusz*, *ateista*; *kmiotek*, *chłop*)”.

Anna i Piotr Wierzbiccy w *Praktycznej stylistyce* podają szereg par wyrazów, których znaczenie (odniesienie przedmiotowe) jest takie samo, ale które różnią się z punktu widzenia niesionej przez nie informacji o stosunku mówiącego do opisywanej przez niego sytuacji. A oto tego rodzaju synonimy: *dziecko* = *bachor*, *staruszka* = *starucha*, *kobieta* = *baba*, *powolny* = *ślamazarny*, *spokojny* = *niemrawy*, *kontakty* = *konszachty*, *zespół* = *klika* = *kolektyw*, *mężczyzna* = *facet*, *klamać* = *łgać*, *ukraść* = *przywłaszczyć sobie*.

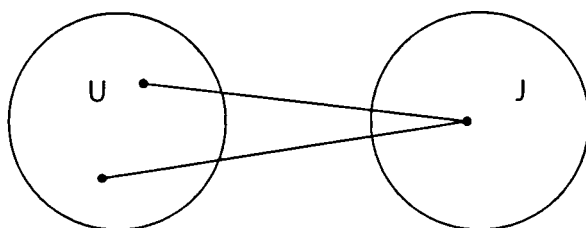
Synonimy mogą pełnić poza tym **funkcję ewokacyjną** (ewokacja – łac. *evocatio* – wywoływanie), będąc często źródłem informacji o przynależności społecznej nadawcy informacji – występujące w tekście (a więc wybrane przez nadawcę spośród innych wyrażen synonimicznych) wyrażenie wywołuje bowiem w pamięci odbiorcy warstwę osób, które zwykły tego wyrażenia używać. Przykładem mogą tu być synonimy *samochód*, *wóz* i *taksówka* – wyrazu *wóz* używają właściciele samochodów, wyrazu *taksówka* na oznaczenie samochodu osobowego – mieszkańcy niektórych wsi.

Kiedy się mówi o tym, że dwa wyrażenia jakiegoś języka są synonimami, ma się na myśli tylko tyle, że wyrażeniom tym przyporządkowany jest jeden i ten sam wycinek rzeczywistości odwzorowywanej przez język (że mają one takie samo odniesienie przedmiotowe, czyli taką samą denotację), nic się jednak nie orzeka o ich zawartości informacyjnej, która w wypadku synonimów wcale nie musi być taka sama. Synonimami są w języku chemii takie wyrażenia, jak:

woda – tlenek wodoru – H_2O
sól kuchenna – chlorek sodu – $NaCl$
siny kamień – siarczan miedzi – $CuSO_4$
siarczek sodu – Na_2S
siarkowodór – H_2S
ozon – O_3
gliceryna – 1,2,3-propanotriol – $C_3H_5(OH)_3$
alkohol etylowy – etanol – C_2H_5OH
aceton – keton dimetylowy – propanon – CH_3COCH_3

ale wyrażenia te, choć równoznaczne, podają inną informację o denotowanym przez siebie elemencie rzeczywistości pozajęzykowej, odwzorowują w planie wyrażania inny zestaw cech konotacyjnych, przy czym wyrażenie niemotywowane – np. *woda*, *ozon* – nie odwzorowuje ich wcale. Z analogicznymi sytuacjami spotykamy się również, w języku naturalnym, a przykładem mogą być chociażby pary wyrazów: *kropplomierz* i *pipeta*, *biustonosz* i *stanik*, *podomka* i *szlafrok*, *flamaster* i *pisak*.

3. Przyporządkowanie wielo-jednoznaczne



Jednemu elementowi U przyporządkowany jest tylko jeden element J, podczas gdy jednemu elementowi J przyporządkowanych jest kilka elementów U. Takie przyporządkowanie (odwzorowanie zbioru U w J) nazywamy wielo-jednoznacznym.

A oto przykłady, gdy J jest językiem polskim:

- przedmiotowi $\text{wóz} \in U$ przyporządkowany jest znak $\text{wóz} \in J$, przedmiotowi $\text{wóz} \in U$ przyporządkowany jest znak $\text{wóz} \in J$, natomiast znakowi $\text{wóz} \in J$ przyporządkowany jest przedmiot $\text{wóz} \in U$ oraz przedmiot $\text{wóz} \in U$.

W przypadku takiego przyporządkowania mamy do czynienia z **homonimią** jednostek należących do J – wyrażenia te nazywamy **homonimami**, a zjawisko występowania w języku homonimów nazywamy homonimią. Homonimia może przejawiać się w następujący sposób:

- może obejmować wszystkie formy (cały paradygmat) danej jednostki języka, tak jak w podanym przykładzie wyżej;
- może przejawiać się tylko na przecięciu paradygmatów, jak na przykład w wypadku wyrazu *wydrze*, który może stanowić formę przypadkową rzeczownika *wydra* oraz osobową czasownika *wydrzeć*, podczas gdy niektóre inne słowoformy tych wyrazów nie są homonimiczne;
- może być wreszcie i tak, że jedna z form paradygmatu jakiejś jednostki leksykalnej jest zbieżna z inną jednostką leksykalną nieodmienną, np. wyraz *bez*, który jest formą przypadkową rzeczownika *bez* oraz przyimkiem *bez*.

Formy typu *wydrze* i *bez* są homonimiczne również dlatego, że mogą być wykładnikami różnych form przypadkowych rzeczowników *wydra* i *bez*. Homonimia tych dwóch ostatnich typów jest szczególnie częsta w języku polskim.

O tym, jaki przedmiot należący do U należy przyporządkować danemu homonimowi (lub jaką relację), decyduje najczęściej kontekst, otoczenie, w jakim dana forma występuje. Dlatego mówimy, że zarówno gramatyka, jak i semantyka języka naturalnego, są zależne od kontekstu. I tak, na przykład, jeśli forma *wydrze* wystąpi w kontekście:

Czytałam dzisiaj o wydrze pana Paska,

wiemy, że jest to forma rzeczownika *wydra*, jeśli natomiast wyraz ten wystąpi w kontekście:

Człowiek wydrze kiedyś przyrodzie i tę tajemnicę

wiemy, że jest to forma czasownika *wydrzeć*. W rzadkich wypadkach nie pomoże jednak nawet kontekst i wyrażenia pozostają homonimiczne, a ich wartość znaczeniowa nieoznaczona. Przykładem mogą być następujące wyrażenia języka polskiego:

Samochód Jana uszkodził samochód Piotra.

(gdzie nie wiadomo, czyj samochód był sprawcą),

Ostatni wydrze koniec lata.

(to ostatnie zdanie ma aż kilka różnych znaczeń).

Homonimia może występować zarówno w języku mówionym, jak i w pisanim – np. w wypadku formy *bez*; może wystąpić tylko w języku mówionym – np. *może, ważyć*, którym w języku pisanym odpowiadają dwie formy niehomo-

nimiczne: *może i morze, ważyć i warzyć* – taki przypadek homonimii nazywamy **homofonią**, może wreszcie wystąpić tylko w języku pisanim np. *rozmarzać*, któremu w języku mówionym odpowiadają dwie różne formy. Taki wypadek nazywamy **homografią**.

Oprócz takiej typowej homonimii spotykamy się w języku naturalnym z **polisemią (wieloznacznością)**, która nie jest wcale (choć tak by się mogło z pozoru wydawać) czymś sporadycznym i wiele wyrazów języka naturalnego można, z pewnego punktu widzenia, uznać za wieloznaczne. Wieloznaczność tę pokazuje Jan Tokarski w pracy *Słownictwo (teoria wyrazu)* na przykładzie wyrazu *kot*, który oznacza:

- indywiduum z gatunku kotów domowych – kot indywidualny,
- każdy z osobna egzemplarz gatunku kot domowy, jeden ze zbioru tych egzemplarzy, desygnat z zakresu objętego nazwą *kot* – kot pojedynczy,
- kot ujęty abstrakcyjnie, jako zbiór cech gatunkowych *felis domestica* – kot uogólniony,
- kot z obrazka – obraz ikoniczny kota,
- kot jako wyraz – obraz arbitralny kota (użycie metajęzykowe).

Oczywiście wyraz *kot* może poza tym oznaczać dowolny egzemplarz z rodziny kotowatych, a więc także tygrysa, lwa czy lamparta.

A oto inne omawiane przez Tokarskiego przykłady wieloznaczności wyrazowej:

- *wiśnie, orzechy, czereśnie* to nazwy zarówno drzew, jak i owoców (choć język polski odróżnia *jabłka i jabłonie, gruszki i grusze, śliwki i śliwy*),

- *biblioteka* to: instytucja, księgozbiór, szafa na książki, budynek, a nawet seria wydawnicza, np. „*Biblioteka Narodowa*”,

- *administracja* to zarówno czynność (administrowanie), jak i wykonawcy tej czynności (administratorzy),

- *ręka* oznaczać może całą kończynę górną, ale też dłoń. *Skaleczył się w rękę* to raczej w dłoń, w przeciwstawieniu do ramienia. Ale *złamał rękę* odnosi się do ramienia lub przedramienia. Zresztą i *dłoń* może oznaczać: 1) rękę od nadgarstka do palców włącznie, 2) to samo, ale z pominięciem palców, 3) stronę wewnętrzną części ręki od nadgarstka do palców w przeciwstawieniu jej grzbietowi (na tej części dłoni ma wyrastać przysłowiowy kaktus), 4) całą kończynę górną.

Właśnie z powodu tej wieloznaczności biologowie wprowadzają terminy *kończyna górna* i *kończyna dolna*, choć są odwieczne *ręka* i *noga* (a zresztą kończyn górnych zwierząt – z wyjątkiem może małp – nie nazywa się *rękami*).

O ile synonimia jest pewną nadmiarowością systemu językowego, o tyle homonimia pewnym niedomiarem. Z punktu widzenia optymalności systemu językowego zjawiska te jak gdyby równoważą się. Synonimię należałoby ocenić negatywnie, bo obciąża pamięć użytkownika, ale tylko wtedy, gdyby język miał pełnić tylko funkcję semantyczną. Bez synonimii pełnienie na przykład funkcji ekspresywnej nie byłoby możliwe, a funkcji estetycznej (artystycznej) znacznie, jeśli chodzi o środki wyrazu, ograniczone. Homonimia leksykalna to mniej liczny słownik języka, ale to oszczędność tylko pozorna, więcej bowiem wysiłku

wkłada się w odkodowywanie komunikatu wieloznacznego i jego ujednoznaczenie. Niejednoznaczność komunikatu spowodowana przez homonim zazwyczaj usuwana jest przez kontekst i nie sprawia odbiorcy większego kłopotu. Znacznie większym problemem jest **homonimia składniowa** (struktur zdaniowych), której powodem najczęstszym jest homonimia wykładników relacji syntagmatycznych, na przykład końcówek fleksyjnych, jak w zdaniach:

*Samochód Jana uszkodził samochód Piotra.
Byt określa świadomość.
Konie mogą zastąpić traktory.*

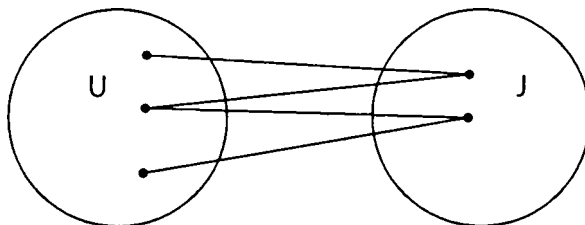
gdzie mianownik i biernik rzeczownika mają tę samą formę – tu tzw. końcówkę zerową. Powodem niejednoznaczności składniowej bywa także brak w języku polskim wykładników wskazujących zasięg łączliwości składniowej wyrażzeń, co widać na przykład w hasle reklamowym:

*Peugeot
z klimatyzacją
za darmo*

– przydałyby się w takich wypadkach nawiasy, jak w zapisie działań matematycznych, które pozwoliłyby na ujednoznaczenie struktury składniowej:

*(Peugeot z klimatyzacją) za darmo.
Peugeot z (klimatyzacją za darmo).*

4. Przyporządkowanie wielo-wieloznaczne



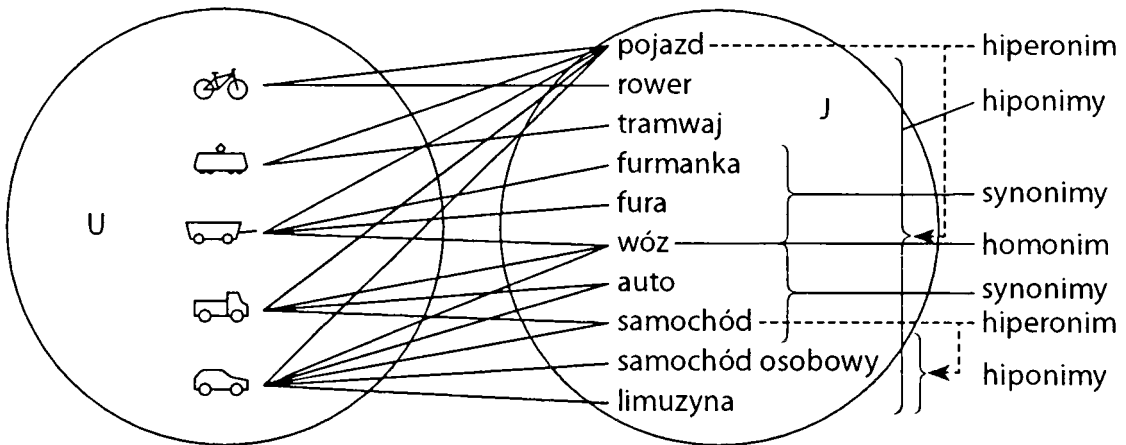
Jednemu elementowi należącemu do U przyporządkowanych jest kilka elementów należących do J i jednemu elementowi należącemu do J przyporządkowanych jest kilka elementów należących do U. Takie przyporządkowanie (odwzorowanie U w J) nazywamy wielo-wieloznacznym.

A oto przykłady, gdy J jest językiem polskim:

- przedmiotowi $\text{☞} \in U$ przyporządkowany jest znak $\text{wóz} \in J$, znak $\text{samochód} \in J$, znak $\text{auto} \in J$, a także znak $\text{samochód osobowy} \in J$, znak $\text{limuzyna} \in J$ oraz znak $\text{pojazd} \in J$, natomiast znakowi $\text{wóz} \in J$ przyporządkowany jest przedmiot $\text{☞} \in U$, przedmiot $\text{☞} \in U$ oraz przedmiot $\text{☞} \in U$, któremu z kolei przyporządkowany jest znak $\text{furmanka} \in J$, a także znak $\text{fura} \in J$. Z kolei znakowi $\text{samochód} \in J$ przyporządkowany jest przedmiot $\text{☞} \in U$, ale też przedmiot $\text{☞} \in U$, natomiast znakowi $\text{pojazd} \in J$ przyporządkowane są oprócz wszystkich powyż-

szych także na przykład przedmioty $\text{tramwaj} \in U$ oraz $\text{rower} \in U$, którym w J przyporządkowane są znaki *tramwaj* i *rower*.

Przyporządkowania te ilustruje poniższy wykres:



Każdy język naturalny odwzorowuje rzeczywistość w sposób wielo-wieloznaczny, a przyporządkowania między elementami języka i elementami odwzorowywanej przezeń rzeczywistości są oczywiście o wiele bardziej skomplikowane niż to ukazano na rysunku, nie oddano bowiem na nim wszystkich relacji paradygmatycznych. Elementom każdego języka naturalnego przyporządkowane są ponadto elementy należące do tego języka, wtedy gdy pełni on funkcję metajęzykową, co jeszcze bardziej komplikuje relacje między językiem a rzeczywistością pozajęzykową.

Najlepszą ilustracją braku jedno-jednoznaczności przyporządkowania między elementami języka a elementami odwzorowywanej przezeń rzeczywistości jest następujący przykład podany przez Stanisława Lema w *Filozofii przypadku*, uzmysławiający jednocześnie trudności w formalizacji języka naturalnego – nie byłoby z tym takich kłopotów, gdyby jedno-jednoznaczne przyporządkowanie istniało. (W tekście Lema wyrażenia użyte w supozycji materialnej występują w cudzysłowach, cytując tekst zaznaczyliśmy je, analogicznie jak w całym tekście podręcznika, kursywą, pozostawiając cudzysłowy tam, gdzie pełnią one inną funkcję, na przykład oznaczają znaczenie.)

„Powiedzmy, że chcemy zaprogramować maszynę cyfrową tak, żeby potrafiła tłumaczyć z polskiego na jakiś inny język i zajmujemy się na razie korzeniem słowa *woda*. Mamy go w słowach *nadwodny* i *podwodny*, ale *zawodny* nie znaczy „taki, co jest za wodą”, tylko „taki, co zawodzi”, *dowodny* nie znaczy „taki, co do wody się ma”, tylko „dowodzący”. *Wodnik* jest w sąsiedztwie *wodniaka*, jako tego, co robi coś na wodzie, np. pływa po niej łódka, ale *rozwodnik*, nie ma już nic z *wodnikiem* wspólnego. *Podwodnik* to na upartego taki, co jest marynarzem łodzi podwodnej (nie powiadam, że takie słowo istnieje, ale tylko, że większość ludzi, dobrze znających polski, na ogół tak by je mogło rozumieć) (słowo to mogłoby w języku polskim oznaczać człowieka, który dostarcza podwoły – BB), ale znów

przewodnik wcale się już nie odnosi do *wody*; *zawodny*, jak już wiemy, „zawodzi”, ale *zawodnik* nie zawodzi, gdyż oznacza sportowca. *Owodnia* jest błoną płodową, ale *przewodnia* nie jest nawet rzeczownikiem (*myśl przewodnia*). Nie jest też rzeczownikiem *udowodni* – i nie oznacza to słowo „uda” jakichś „wodni”. Również *zwody*, *zawody*, *wywody*, *przewody*, *odwody*, *rozwody*, *dowody* nie mają z wodą nic wspólnego, przy czym *dowody* i *odwody*, wbrew podobieństwu strukturalnemu, semantycznie są bardzo odległe od siebie, a *przewód*, tak jak *przewodnik* zresztą, ma rozmaite znaczenia (każde z tych słów z osobna): „przewód sądowy” i „elektryczny”, „przewodnik” jako człowiek i jako obiekt martwy, np. kabel prądowy. Mamy też *wzwód* (jako erekcję), *wywód*, mamy *wody* jako „uzdrowiska” („pojechał do wód”), *wodowanie* (jako „ładowanie na wodzie”) i *powodowanie*, przy czym podobieństwo obu tych słów do pary *wożenie* i *powożenie* albo *siekanie* i *posiekanie* jest mylące; jest i *wódka*, nie stanowiąca „małej wody”, chociaż *bródka* jest małą brodą, i znów z wodą spokrewniony *niewód*; *podwód* nie istnieje, ale jest *podwoda*, znowuż z tym, co pod wodą, nawet nie spowinowacona; *rozwodniony* to nie taki, co się rozwiódł, bo ten będzie *rozwiedziony*, a chociaż *zwodniczy* jest kształt, którym będzie ktoś „zwiedziony”; *rozwiedziony* i *zwiedziony* niekoniecznie mają się do siebie tak, jak *złączony* i *rozłączony*, bo *łączący* i *wodzący* stoją w innych do siebie stosunkach aniżeli dane sufiksami czy prefiksami strukturalne odpowiedniki tych par w obrębie relacji strukturalnych. *Wodzić* nie ma się do *wodować* jak *głodzić* do *głodować* ani jak *słodzić* do *słodować* ani jak *poszkodzić* do *poszkodować*; a chociaż *odwodzący* lub (nie istnieje) *odwodowujący* są bardziej jakby powiązane z *wodą*, aniżeli *odwadniający* – właśnie to ostatnie słowo należy do „akwaticznego korzenia” a nie tamte. Otóż nasza maszyna, gdyby sobie nie połamala dotąd wszystkich tranzystorów na tym materiale, już byłaby spokojna przynajmniej o jedno – „wiedziałyby”, że *o* przemienia się w korzeniu *wod* w *a* – w formie częstotliwej (*odwodnić* – *odwadniać*). Toteż spokojnie przystąpiłaby do tłumaczenia *udowadniać*, jako że jest w nim owo *wod*, od wody pochodzące – najzupełniej fałszywie znowu, bo niczego nie „udowadnia” ten, kto „nawadnia”. Maszyna mogłaby mieć do nas pretensje o to, że *obwodnica* nie jest z wody, ale *wodnica* jest (a to, że *obwodnica* od *obwodu* pochodzi, nic jej nie pomoże, bo mamy nadto *zwodnicę*) oraz, że *nawodnica* jest z wody, a *nawodnik* nie – itd., bez końca”

Gdybyśmy chcieli zbiór U, czyli całą rzeczywistość odwzorować w zbiór J (język) w ten sposób, że jednemu elementowi J przyporządkowany byłby tylko jeden element U, to słownik takiego języka rozrósłby się wprost monstrualnie i językiem takim w ogóle nie można by się było posługiwać. Dlatego postępujemy w ten sposób stosunkowo rzadko i zazwyczaj obiekty informacji wskazujemy w inny sposób: przyporządkowując jednemu wyrażeniu językowemu nie jeden element rzeczywistości, ale wszystkie elementy w jakiś sposób do siebie podobne, i to nie tylko takie, które istnieją w danym momencie, ale również takie, które kiedykolwiek istniały lub będą istniały, przy czym stopień podobieństwa może tu być w przypadku różnych nazw różny. I tak na przykład, wszystkie elementy rzeczywistości nazywane wyrażeniem *pies* mają więcej cech wspólnych niż ele-

menty rzeczywistości nazywane wyrażeniem *zwierzę*. Taka nazwa wydziela więc z całej rzeczywistości klasę przedmiotów, których dotyczy dana informacja, ale jeszcze nie wskazuje w tym zbiorze dokładnie tego właśnie jedyne elementu będącego obiektem zainteresowania nadawcy informacji. Takie uściślenie w niektórych językach naturalnych jest obligatoryjne, a wykładnikami tej informacji są na przykład rodzajniki określone i nieokreślone. Nie zawsze jest to potrzebne, gdyż często jest tak, że obiektem informacji może być tylko jeden jedyny element danego zbioru (o tym, że tak jest, wie i nadawca i adresat informacji), a wtedy dokładne wskazywanie go w klasie nie jest już potrzebne. Tak na przykład jest w wypadku zdań:

Dziecko chore, więc nie mogę iść z tobą do teatru.

Pies jest dziś jakiś dziwnie niespokojny.

Przewodniczący zarządził głosowanie.

w zdaniach tych chodzi bowiem nie o jakieś dziecko, jakiegoś psa, ale o jedno jedyne dziecko i jednego psa, których mówiący i słuchający mają na myśli. Elementy zbioru można również identyfikować przez podanie jakiejś cechy przysługującej temu jednemu elementowi danego zbioru (mamy tu więc do czynienia właściwie z iloczynem kilku zbiorów), na przykład *moje dziecko, mój pies, najmłodsze dziecko Anny, największa rzeka Polski, obecny przewodniczący naszego stowarzyszenia*. Bardzo często bywa i tak, że obiektem informacji jest sam nadawca informacji lub jej odbiorca lub też zbiór, w którym jednym z elementów jest nadawca lub odbiorca – w tym przypadku taką funkcję identyfikującą pełnią zaimki *ja, ty, my, wy* (w języku polskim także ich synonimy w postaci odpowiednich końcówek czasownika 1 i 2 osoby).

Jednoznaczne powiązanie odpowiedniego elementu tekstu z jego desygnatem jest więc warunkiem koniecznym i najczęściej wystarczającym do porozumienia, nawet jeśli istnieje pewna różnica w rozumieniu samej treści wyrazów. Jak pisze Jan Tokarski w *Słownictwie*: „Pod względem stosunku do desygnatu wyraz *kot* może być jednoznacznie użyty przez zoologa i przez przedszkolaka, porozumiewają się oni w razie potrzeby, nieporównywalna jest jednak wiedza, jaką obaj mają o desygnacie”.

Język pełni zarazem funkcję **klasyfikowania elementów rzeczywistości** (a często służy również do wydzielenia tych elementów) – dzielenia ich na klasy – i nadawania tym klasom nazw. W samej rzeczywistości brak wyraźnych granic między tymi klasami, a ustalenie ich jest każdorazowo wynikiem arbitralnej decyzji klasyfikującego (użytkownika języka). Elementy rzeczywistości można grupować w bardzo różne klasy, zależnie od konkretnej potrzeby. Każdy język naturalny jest klasyfikacją rzeczywistości, ale każdy język dzieli tę rzeczywistość według swoich własnych kryteriów i dlatego języki naturalne są w swym zewnętrznym kształcie właściwie nieporównywalne. Wiele świadczących o tym przykładów podaje Anna Wierzbicka w książce *O języku dla wszystkich*. Przytoczymy je tu za autorką jako niebagatelny argument świadczący o tym, że zbudowanie wielojęzycznych słowników jest sprawą trudną, jeżeli nie wręcz niemożliwą.

„Zulusi z Południowej Afryki, znający przecież krowy, nie posiadają w swym języku wyrazu odpowiadającego naszemu „krowa”, angielskiemu *cow* i francuskiemu *vache*. Owszem, mają wyraz oznaczający białą krowę i inny, oznaczający czerwoną krowę, ale „krowa w ogóle” – nie, takiego wyrazu nie ma.

W języku *lushei*, którym mówi część ludności Indii, znaleziono 10 wyrazów oznaczających różne rodzaje mrówek, 20 – oznaczających różne rodzaje koszy, kilkanaście – dla różnych odmian jeleni, ale wyrazów odpowiadających po prostu naszym „mrówka”, „kosz”, „jeleń” – nie ma.

Indianie Klamath z Oregonu używają dziesiątków wyrazów na oznaczenie różnych odmian lisów, wiewiórek, motyli; ogólnych wyrazów o znaczeniu ‘lis’, ‘wiewiórka’, ‘motyl’ – brak.

Lapończycy posługują się mnóstwem wyrazów na oznaczenie rena; innym słowem nazywa się rena jednolatka, innym dwu-, trzy-, cztero-, pięcio-, sześć-, i siedmiolatki; istnieje około 20 słów na oznaczenie lodu, jednaście – na zimno, 41 – na postacię śniegu, ale próżno szukalibyśmy w tym języku wyrazów odpowiadających naszym „mróz”; „śnieg”, „zimno”. „Śnieg”, ale jaki? – zapytałby Lapończyk. „Padający czy leżący, suchy czy mokry? Jak można określać jednym wyrazem zjawiska tak różne?”

Maorysi z Nowej Zelandii mają słownictwo, którego bogactwo wprawiało w podziw podróżników europejskich. Znają oni płeć drzew... Mają odrębne nazwy dla drzew męskich i żeńskich pewnych gatunków. Mają różne nazwy dla drzew, których liście zmieniają formę, stosownie do różnych momentów wzrostu. (...) Mają odrębne wyrazy dla oznaczenia ogona ptaka, ssaka lub ryby.

Indianie Koroadow z Brazylii mają odrębne nazwy dla każdej odmiany małp, palm, papug, tak szczegółowe, że język ich mógłby służyć za przewodnik botaniczny i zoologiczny, ale znów – na próżno szukalibyśmy w ich języku wyrazów ogólnych odpowiadających naszym „małpa”, „papuga”, „palma”.

Indianie Kuna, zamieszkujący Panamę i Kolumbię, mają podwójne nazwy dla większości zwierząt i roślin: jeden szereg nazw na określenie zwierząt i roślin w dzień, drugi – na określenie tych samych zwierząt i roślin w nocy.

I jeszcze jeden przykład, tym razem z języka Tasmańczyków. Nie mają oni wyrazu, który odpowiadałby znaczeniem naszemu „czarny”. Mają natomiast szereg wyrazów oznaczających różne odmiany tego koloru. I tak *kotkot* – to czarny jak kruk; *toworo* – to czarny kolor zwęglonych orzechów aleurytów; *luluba* – to czarny określony kolor błota; *dep* – to kolor czarny spalonej żywicy drzewa kanarii; *likutan* – to czerń szczególnie ciemna.

(...) w języku *hopi*, którym mówią Indianie z Arizony, istnieje wyraz *masa'yta-ka*, oznaczający wszystko co lata, czy to będzie owad, samolot czy lotnik.

Podobnie w sudańskim języku *anuk* jest wyraz oznaczający każdy przedmiot zrobiony z metalu, czy to będzie igła czy samolot. U Azteków w Meksyku – jeden wyraz odpowiada naszemu „mróz”, „śnieg” i „lód”.

W języku indiańskich Siouksów nie istnieją czasowniki, które odpowiadałyby naszym „odłamać”, „oderwać”, „odciąć”, istnieje natomiast jeden czasownik o znaczeniu bardziej ogólnym ‘oddzielić’:

(...) W języku kaka w Kamerunie (zachodnia Afryka) istnieje wiele szczegółowych nazw oznaczających różne rodzaje owoców.

Wyrazu odpowiadającego naszemu ogólnemu „owoc” – brak, są jednak dwa inne: jeden z nich obejmuje banany i ananasy (takiego znów brak w angielskim czy polskim), drugi – wszystkie inne rodzaje owoców, a także nasiona różnych roślin, galkę oczną, gruczoły, pigułki, piłki.

W jednym z języków indiańskich nie ma wyrazu, który odpowiadałby znaczeniem naszemu wyrazowi „drzewo”, istnieje natomiast inny wyraz, oznaczający ‘długą rzecz stojącą’

W indiańskim języku navaho nie ma wyrazu, który odpowiadałby naszemu „ptak”. Istnieje wprawdzie wyraz *cidiî*, ale oznacza on tylko małe ptaki; wyraz *nààtʔà.i* natomiast może się odnosić do małych i do dużych ptaków, zarazem jednak oznacza on nietoperze i inne latające zwierzęta, nie będące ptakami.

W tymże navaho jest wyraz oznaczający nasiona dyni (*ʔálásciî*), inny oznaczający nasiona owoców, nie ma natomiast wyrazu, który odpowiadałby ściśle naszemu „pestka”; najbliższy mu znaczeniem wyraz *-nààʔ*. to w navaho nie tylko ‘pestka’, ale i ‘oko’.

(...) „Drzewo” – to po francusku *arbre*, po niemiecku *Baum*, po duńsku *træ*, „las” po francusku *bois*, po niemiecku *Wald*, po duńsku *skov*. Kiedy jednak mówimy, że coś zrobione jest z drzewa, Francuz użyje tu wyrazu nie *arbre*, ale *bois*, Niemiec nie *Baum*, ale *Holz*. Z drugiej strony francuskiemu *bois* i niemieckiemu *Holz* odpowiada w wielu kontekstach polski wyraz „drewno”, w duńskim – w jednych wypadkach *træ*, w innych *skov*. Ten sam wyraz francuski *bois* znaczy bardzo często tyle, co polski „las”, a niemiecki *Wald*. Jednakże i tu nie ma odpowiedniości zupełnej, bo niemiecki *Wald* – to nie tylko francuskie *bois*, ale i *forêt*, polskie „puszcza”. Schematycznie można to przedstawić tak:

polski	francuski	niemiecki	duński
<i>drzewo</i>	<i>arbre</i>	<i>Baum</i>	<i>træ</i>
<i>drewno</i>	<i>bois</i>	<i>Holz</i>	
<i>las</i>		<i>forêt</i>	<i>Wald</i>
<i>puszcza</i>			

Jak różne języki analizują rzeczywistość, widać szczególnie wyraźnie w dziedzinie nazw kolorów. Barwy tworzą, jak wiadomo, pewien ciąg, w widmie słonecznym przechodzą niepostrzeżenie jedne w drugie, bez żadnych ostrych granic. Języki jednak wyodrębniają z tego ciągu pewne pasy – w sposób całkowicie z punktu widzenia optyki dowolny – i nadają im nazwy, każdy język w trochę inny sposób. I tak, na przykład, angielskiemu wyrazowi *green* (zielony) odpowiada

w języku walijskim *gwyrd*. W niektórych jednak wypadkach angielskie *green* oddałoby się walijskim wyrazem *glas*, który odnosi się jednak i do odcieni, określanych w angielskim jako *blue* (niebieski) lub *grey* (szary). Ale niektóre rodzaje szarości oznacza się w walijskim nie wyrazem *glas*, lecz *llwyd*, który z kolei trzeba by czasem tłumaczyć na angielski jako *brown* (brązowy). Oto schemat tych różnic:

angielski	walijski
(zielony) <i>green</i>	<i>gwyrd</i>
(niebieski) <i>blue</i>	<i>glas</i>
(szary) <i>grey</i>	
(brązowy) <i>brown</i>	<i>llwyd</i>

Jeżeli porównamy nazwy barw w języku polskim i w języku navaho, którym mówią Indianie z Arizony i Nowego Meksyku, okaże się, że polskiemu wyrazowi „czarny” odpowiadają w navaho dwa wyrazy, odróżniające rodzaje czerni, natomiast jeden tylko wyraz odpowiada polskiemu „zielony” i „niebieski”. Oto tabelka:

navaho	polski
<i>ligàì</i>	<i>biały</i>
<i>dìxìl</i>	<i>czarny</i>
<i>łìžìn</i>	
<i>łìčì?</i>	<i>czerwony</i>
<i>do ǰìž</i>	<i>niebieski</i>
	<i>zielony</i>

Wiele języków afrykańskich ma tylko trzy nazwy kolorów, które dzielą między siebie całe widmo słoneczne.

I jeszcze jeden ciekawy wypadek różnic między językami, tym razem między angielskim a chińskim. W języku angielskim istnieje wyraz *strawberries* (‘tru-

skawki'); w języku chińskim dokładnego odpowiednika nie ma, są natomiast dwa wyrazy o znaczeniu węższym, konkretniejszym: *yángmèi* i *tsāumèi*; nazwa ogólniejsza niż *strawberries* – to w angielskim *berries*, wyraz oznaczający drobne owoce, jak truskawki, poziomki, maliny, jagody; odpowiednika chińskiego brak; jeszcze szersza, ogólniejsza nazwa – to *fruit* ('owoce'); w chińskim odpowiada jej wyraz *shwěigwǒ*; zarówno wyraz angielski *fruit*, jak chiński *shwěigwǒ* nie obejmują swym znaczeniem orzechów. W chińskim istnieje nazwa jeszcze ogólniejsza, obejmująca i owoce i orzechy: *gwǒ* – tu z kolei brak odpowiednika angielskiego (przy okazji warto wspomnieć, że język chiński nie ma wyrazu, który by odpowiadał ściśle angielskiemu *nuts* ('orzechy'); istnieje wprawdzie w chińskim wyraz *gǎngwǒ*, ale oznacza on – oprócz orzechów – także suszone owoce).

(...) „Brat” po polsku znaczy „syn tych samych rodziców”. W języku malajskim wyraz *sudara* odnosi się i do braci, i do sióstr, znaczy po prostu „dziecko tych samych rodziców”(...) Język węgierski odróżnia za pomocą osobnych wyrazów i braci od sióstr, i braci (siostry) młodsze od braci (sióstr) starszych. Tek więc malajski wyraz *sudara* ma w języku polskim dwa odpowiedniki, a w węgierskim cztery, choć nie zachodzi tu zjawisko wieloznaczności.

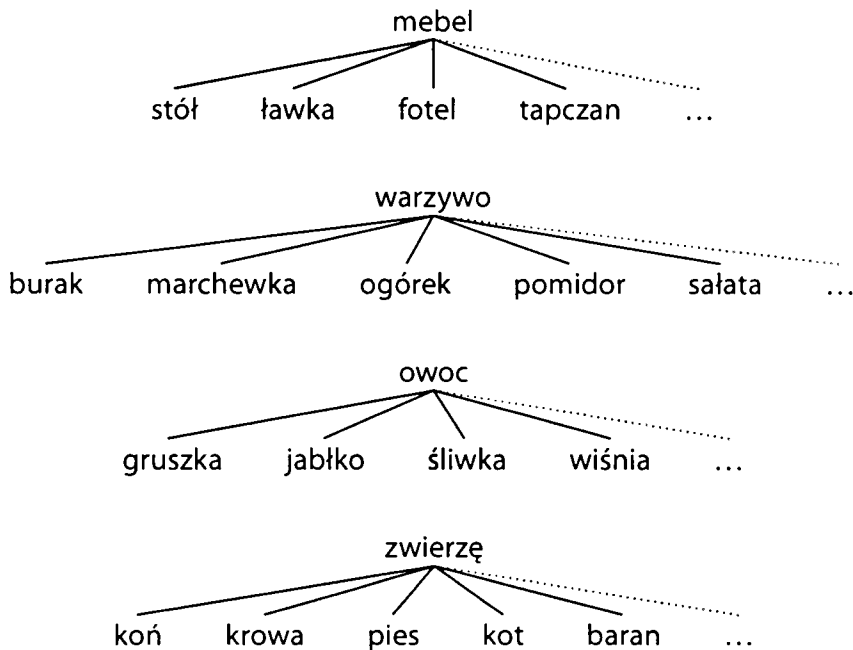
Znaczenie	węgierski	polski	malajski
starszy brat	<i>batya</i>	<i>brat</i>	<i>sudarā</i>
młodszy brat	<i>öccs</i>		
starsza siostra	<i>néne</i>	<i>siostra</i>	
młodsza siostra	<i>híg</i>		

(...) I myliłby się ten, kto by sądził, że takie niepokrywanie się znaczeń wyrazów dotyczy tylko jakiegoś marginesu słownictwa, jakichś wyjątków. Uczony amerykański Maurice Swadesh, twórca glottochronologii starał się ustalić listę 100 elementarnych jednostek wyrazowych, które odpowiadałyby sobie ściśle w większości języków świata. Wybrał jednostki takie, jak nazwy części ciała, podstawowych zjawisk przyrody (np. słońce, wiatr), stopni pokrewieństwa (ojciec, matka, brat), elementarnych czynności (śmiać się, krzyczeć, śpiewać) itp., wychodząc z założenia, że są to pojęcia wspólne wszystkim ludziom, a więc zapewne wyodrębnione za pomocą wyrazów we wszystkich językach świata.

I cóż się okazało? Ze wszystkich stron posypały się na Swadesha głosy krytyki: językoznawcy różnych specjalności przytaczali wciąż nowe przykłady, przeczące propozycjom Swadesha. Okazało się, że nie można ustalić nawet 100 jednostek wyrazowych, które ściśle odpowiadałyby sobie znaczeniem w większości języków świata”. (Anna Wierzbicka: *O języku dla wszystkich*).

System słownikowy każdego języka naturalnego ma budowę luźną i choć każdy język jest zarazem jakąś **klasyfikacją rzeczywistości**, którą odwzorowuje,

to jednak jednostki leksykalne języka naturalnego nie dałyby się ułożyć w jedno drzewo klasyfikacyjne. Klasyfikacja ta nie spełnia warunków stawianych poprawnej klasyfikacji chociażby dlatego, że nie przeprowadza konsekwentnie podziału rozłącznego (każdy przedmiot może być zaliczany do wielu różnych klas). W każdym języku naturalnym najwięcej jest jednostek leksykalnych nazywających konkretne przedmioty, a więc jednostek nazywających najmniejsze wyróżnione podzbiory uniwersum, takich jak: *stół, ławka, fotel, tapczan, marchewka, ogórek, sałata, koń, krowa, pies, kot*. W każdym języku istnieją również pewne klasy, grupujące z kolei klasy najniższego rzędu, np.:



Stosunkowo rzadko jednak język naturalny rozbudowuje tę klasyfikację dalej, rzadko bowiem istnieje potrzeba posługiwania się jednostkami o wysokim stopniu abstrakcji, a więc niosącymi mało informacji. Klasyfikacja ta może być oczywiście w miarę potrzeby rozbudowana poprzez łączenie lub dzielenie zakresów, co powoduje tworzenie się znacznie rozgałęzionych wielopoziomowych drzew klasyfikacyjnych. Tak dzieje się w języku nauki, a najlepszym przykładem takiej klasyfikacji może być systematyka roślin i zwierząt.

Podział systematyczny świata roślinnego (według *Małej encyklopedii przyrodniczej*)

I. Plechowce

1. Glony (...)
2. Grzyby (...)
3. Porosty (...)

II. Mszaki

1. Mchy (...)
2. ...

III. Paprotniki

1. Paprocie (...)
2. Skrzypy (...)
3. Widłaki (...)
4. ...

IV. Nasionne

1. Nagonasienne

- a) ...
- b) ...
- c) Szpilkowe
 - cisowate (cis)
 - sosnowate (świerk, jodła, modrzew, sosna)
 - cyprysowate (jałowiec, ...)
 - ...

2. Okrytonasienne

a) Dwuliścienne

- brzozowate (brzoza, olcha, grab, leszczyna)
- bukowate (buk, dąb, kasztan)
- orzechowate (orzech włoski)
- wierzbowate (wierzba, topola)
- morwowate (figa, ...)
- konopiowate (konopie, chmiel)
- pokrzywowate (pokrzywa)
- wiązowate (wiąz)
- gązewnikowate (jemiola)
- rdestowate (rdest, szczaw, rabarbar, gryka, ...)
- kaktusowate
- komosowate (burak, szpinak)
- goździkowate (goździk, firletka, kąkol)
- jaskrowate (jaskier, kaczeniec, zawilec, sasanka, ...)
- krzyżowe (lewkonía, kapusta, chrzan, rzodkiew, ...)
- fiołkowate (fiołek, bratek)
- dziurawcowate (dziurawiec)
- skalnicowate (porzeczka, jaśmin, ...)
- różowate (róża, głóg, grusza, jabłoń, malina, poziomka, śliwa, tarnina, morela, wiśnia)
- ...

(wyróżniono 84 takie grupy)

b) Jednoliścienne

- liliowate (lilia, czosnek, tulipan, konwalia, ...)
- amarylkowate (narcyz, przebiśneg, agawa)
- trawy (ryż, żyto, pszenica, kukurydza, proso, owies, jęczmień, ...)
- palmy
- ...

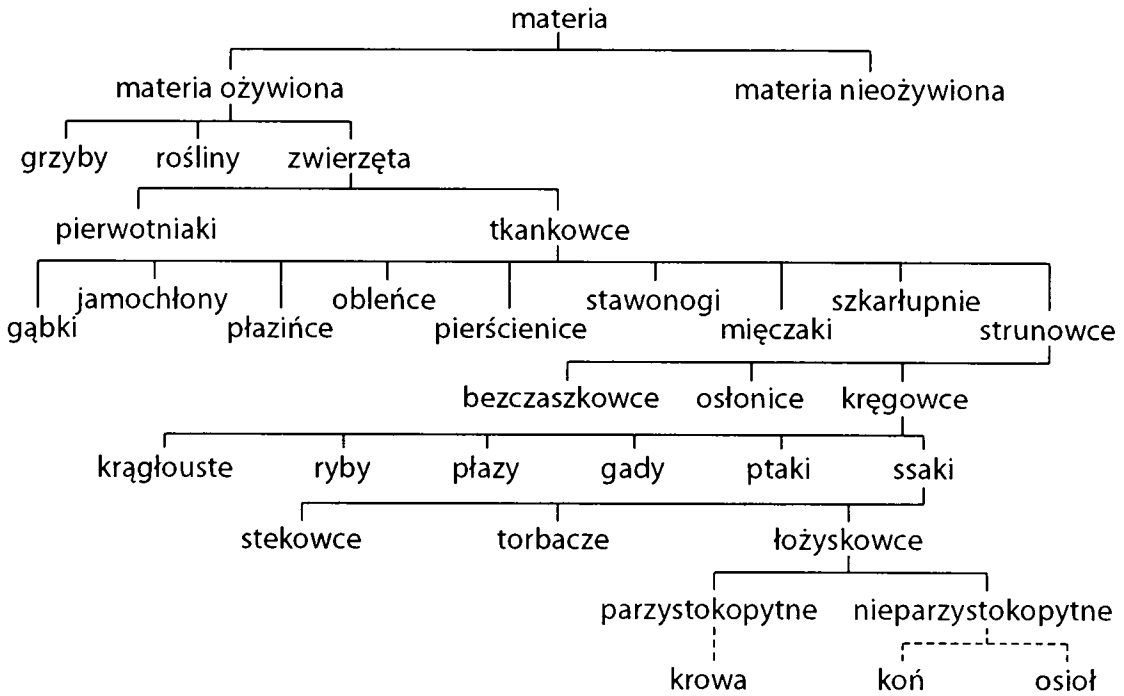
Język naturalny tę samą rzeczywistość dzieli inaczej. Podział ten nie spełnia logicznych warunków klasyfikacji i przez językoznawców nazywany jest **kategoryzacją**. Cechuje go przewaga słownictwa z tak zwanego poziomu podstawowego, wyznaczanego przez relewancję dla użytkowników danego języka informacji konstytuującej treść wyrażen, uboga rozbudowa hierarchiczna i mieszanie elementów z różnych poziomów paradygmatycznych. Przykład takiej kategoryzacji roślin podaję za Grażyną Habrajską (*Kategoryzacja a klasyfikacja – potoczne i naukowe widzenie świata* w: R. Grzegorzczkova i A. Pajdziński (red.): *Językowa kategoryzacja świata*).

Należy zwrócić uwagę na fakt, że zakres wyrazów jest w języku naturalnym w pewnej mierze przypadkowy, wyrazy te włączone w drzewo klasyfikacyjne języka nauki nabierają często innej treści. I tak, na przykład, wyraz *owoc* obejmuje w języku naturalnym swym zakresem zakresy takich wyrazów, jak: *jablko, gruszka, śliwka, truskawka, porzeczka*, ale nie obejmuje zakresów wyrazów: *pomidor, ogórek, orzech czy strąk*, które obejmuje zakres wyrazu *owoc* przyjęty przez język nauki. W języku polskim istnieją ponadto wyrazy *warzywo* i *jarzyna*, których zakres obejmuje zakres nazw bardzo różnych części różnych roślin, będących ich owocami (w znaczeniu przyjętym przez język nauki) – jak np. pomidory, ogórki; korzeniami – jak buraki, marchew; czy liśćmi – jak np. sałata. Są one łączone w jedną klasę na tej podstawie, że wszystkie są jadalne i sadzone na wiosnę i nie rosną na drzewach ani na krzewach (wyraz *jarzyna* nie obejmuje jednak swym zasięgiem kartofli, gdyż te zjawily się w Polsce już po ustabilizowaniu się znaczenia tego wyrazu). Mamy również w języku polskim wyraz *włoszczyzna* obejmujący tylko niektóre warzywa i to tylko wtedy, gdy występują w małych ilościach łącznie jako przyprawa do zupy (są to zazwyczaj: marchew, pietruszka, seler, por, kapusta). W tych ostatnich przykładach mamy więc objęte wspólną nazwą przedmioty o zupełnie różnych cechach fizycznych, które zostały jednak połączone w jedną klasę na podstawie ważnej z punktu widzenia użytkownika języka (twórcy i użytkownika klasyfikacji) cechy funkcjonalnej.

Przykłady te są dobrą ilustracją względności jednego z praw teorii poprawnej klasyfikacji, która wymaga, aby klasyfikacja była naturalna, to znaczy, aby przedmioty należące do zakresu jednego członu podziału były do siebie bardziej podobne niż przedmioty należące do zakresów dwóch różnych członów.

Zakresy istniejących w języku naturalnym jednostek najniższego rzędu w razie potrzeby mogą być dalej dzielone, tak jak jest na przykład w wypadku nazwy *pies*, która jest podzielona na poszczególne rasy. Często bywa tak, że w języku potocznym inny jest zakres wyrażen niż w języku naukowym. Na przykład potoczna nazwa *zwierzę*, nie obejmuje owadów i robaków, a *robaki* to także potocznie niektóre owady, na przykład karaluchy. Język potoczny grzyby zalicza do roślin, podczas gdy w biologii są to dwa różne królestwa. W języku naukowym wyraz *zwierzę* jest jednostką na dużo wyższym poziomie abstrakcji, a jego zakres podlega wielostopniowemu podziałowi.

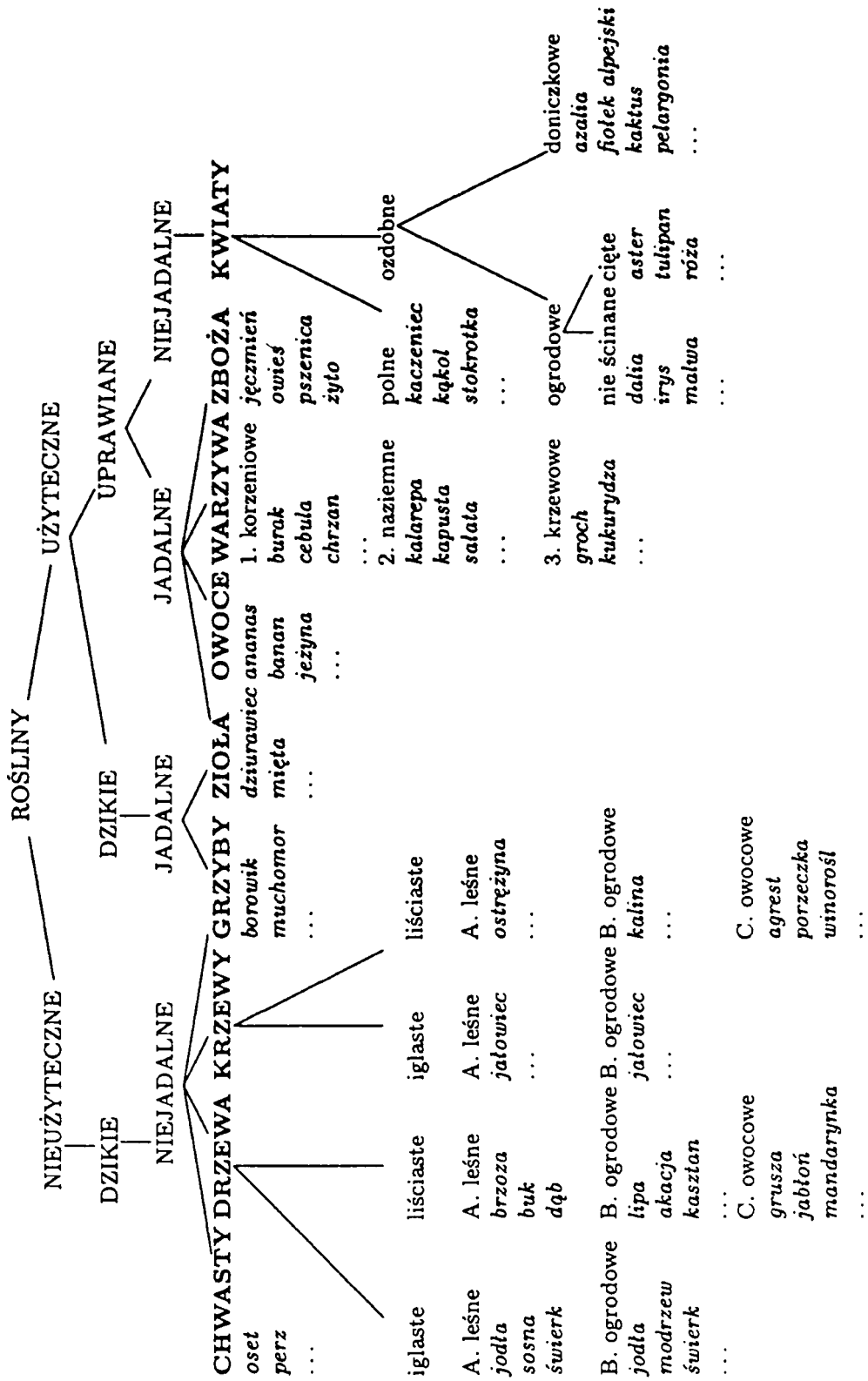
Dla ilustracji przytaczamy wycinek tej klasyfikacji, oczywiście opuszczając wiele jej poziomów:



Jednostki języka naturalnego bardzo często nie wyznaczają zakresu odwzorowywanych przez nie zbiorów w sposób ostry i jednoznaczny. Przykładem mogą tu być chociażby wyrazy *rzeka*, *potok*, *strumień*, którym wcale niełatwo przyporządkować jednoznacznie zbiory ich desygnatów. I tak, płynącą wodę szerokości kilkudziesięciu metrów nazwiemy na pewno *rzeką*, przy szerokości kilkudziesięciu centymetrów nazwiemy *strumieniem* (lub *strumykiem*), ale nie da się ściśle określić, kiedy mamy „jeszcze” rzekę, a kiedy „już” potok czy strumień. Podobnie nieostry zakres mają rzeczowniki *pagórek*, *wzgórze*, *góra*. Takie sytuacje są dla języka naturalnego bardzo typowe, ostre sprecyzowanie granic zakresu poszczególnych wyrażen jest możliwe i potrzebne tylko w języku sztucznym.

W wypadku **kategorii języka naturalnego** nie tylko granice kategorii bywają trudne do określenia, ale okazuje się, że równie trudno sprecyzować kryteria, jakie muszą spełniać elementy rzeczywistości, żeby być do danej kategorii zaliczone. W wyniku badań użytkowników języka okazało się, że pewne przedmioty są lepszymi niż inne reprezentantami danej kategorii – takie, które pierwsze przychodzą na myśl, jako najbardziej typowe, nazywa się **prototypami**. I tak, lepszym reprezentantem kategorii ptak będzie dla mówiących po polsku wróbel lub gołąb niż kura, gęś, struś czy pingwin, lepszym reprezentantem kategorii ssak pies, koty, koń, krowa niż wieloryb czy delfin. Okazuje się, że wszystkie elementy należące do danej kategorii semantycznej można by opatrzyć odpowiednimi wskaźnikami pokazującymi ich „odległość semantyczną” od centrum kategorii wyznaczonej przez prototyp tej kategorii. Przykładem może tu być wynik eksperymentu przeprowadzonego wśród 200 ankietowanych studentów angielskich odnoszący się do kategorii MEBEL.

Potoczna klasyfikacja świata roślinnego



Okaz	Ranga	Ocena trafności	Okaz	Ranga	Ocena trafności
krzesło	1,5	1,04	lampa	31	2,94
sofa	1,5	1,04	taboret	32	3,13
kanapa	3,5	1,10	klęcznik	33	3,43
stół	3,5	1,10	szuflada	34	3,63
fotel	5	1,33	pianino	35	3,64
szafa	6,5	1,37	poduszka	36	3,70
fotel-bujak	6,5	1,37	stojak na gazety	37	4,14
stolik kawowy	8	1,38	hi-fi	38	4,25
kołyska	9	1,42	szafka na naczynia	39	4,27
fotel dwuosobowy	10	1,44	stereo	40	4,32
komoda	11	1,48	lustro	41	4,39
biurko	12	1,54	telewizor	42	4,41
łóżko	13	1,58	barek	43	4,46
sekretarzyk	14	1,59	półka	44	4,52
tapczan	15,5	1,61	chodnik	45	5,00
pomocnik (stolik)	15,5	1,61	poduszka do spania	46	5,03
materac na stelażu	17	1,70	kosz na śmieci	47	5,34
nocna szafka	18	1,83	radio	48	5,37
kufer	19	1,98	maszyna do szycia	49	5,39
skrzynia (kufer) cedrowa	20	2,11	kuchnia	50	5,40
serwantka	21	2,13	lada	51	5,44
regał książkowy	22	2,15	zegar	52	5,48
kozetka	23	2,17	kotary	53	5,67
szezłąg	24	2,26	lodówka	54	5,70
otomana	25	2,43	obraz	55	5,75
podnózek	26	2,45	szafa wnękowa	56	5,95
gablotka	27	2,49	wazon	57	6,23
szafka na porcelanę	28	2,59	popielniczka	58	6,35
ława	29	2,77	wentylator	59	6,49
bufet	30	2,89	telefon	60	6,68

Z podobnych eksperymentów wynika, że układ pola semantycznego języka naturalnego nie przypomina mozaiki, bo jego elementy nie są wyraźnie odgraniczone.

Jak z przytoczonych wyżej przykładów widać, język naturalny dokonuje w swym systemie semantycznym pewnej **kategoryzacji rzeczywistości**, którą opisuje. Kategoryzacja ta znajduje wyraz przede wszystkim w systemie leksykalnym języka. Również u podłoża wszelkich języków sztucznych pełniących funkcję reprezentowania jakiejś rzeczywistości leży zawsze jakaś kategoryzacja tej

rzeczywistości. Często modelem takiej kategoryzacji bywają znane kategoryzacje ontologiczne, szczególnie te o zakresie uniwersalnym. Jedną z najstarszych i najbardziej znanych jest klasyfikacja Arystotelesa, który wyróżnił dziesięć kategorii obejmujących, według niego, całe uniwersum: substancja, ilość, jakość, relacja, miejsce, czas, położenie, stan, działanie i doznanie. Liczbą i rodzajem wyróżnionych kategorii różnią się między sobą systemy filozoficzne – na przykład reizm jest przykładem systemu uboższego w kategorie, gdyż zakłada jedynie istnienie rzeczy.

Kategorie ontologiczne często stają się modelem kategorii semantycznych w językach sztucznych, a także modelem klasyfikacji wyrażen języka naturalnego. Przykładem może tu być kategoryzacja zastosowana w leksykonie M. Rogeta, gdzie wyróżniono sześć kategorii: stosunki abstrakcyjne, przestrzeń, materia, intelekt, wola, władze uczuciowe i moralne, czy też kategoryzacja zaproponowana przez R. Halliga i W. von Wartburga, gdzie wyróżniono tylko trzy kategorie: wszechświat, człowiek, człowiek we wszechświecie. Kategoryzacją o charakterze uniwersalnym jest też zaproponowany przez Ranganathana podział uniwersum na pięć kategorii: Indywidualium, Materia, Energia, Przestrzeń, Czas, które legły u podstaw stworzonego przez niego języka informacyjno-wyszukiwawczego, Klasyfikacji Dwukropkowej.

Klasyfikacja rzeczywistości znajduje swoje odbicie w kategoriach leksykalnych wyróżnionych w poszczególnych językach informacyjno-wyszukiwawczych. Odpowiednie przykłady podamy w rozdziale XVI.

Literatura

- Grodziński E.: *Językoznawcy i logicy o synonimach i synonimii*. Wrocław 1985.
- Grzegorzczkowska R., Pajdzińska A.: *Językowa kategoryzacja świata*. Lublin 1996.
- Levi-Straus C.: *Myśl nieoswojona*. Warszawa 1967.
- Sapir E.: *Kultura, język, osobowość*. Warszawa 1978.
- Schaff A.: *Język a poznanie*. Warszawa 1964.
- Taylor J. R.: *Kategoryzacja w języku*. Kraków 2001.
- Tokarski J.: *Słownictwo. Teoria wyrazu*. Warszawa 1971.
- Whorf B.L.: *Język, myśl i rzeczywistość*. Warszawa 1982.
- Wierzbicka A.: *O języku dla wszystkich*. Warszawa 1967.
- Woźniak J.: *Kategoryzacja. Studium z teorii języków informacyjno-wyszukiwawczych*. Warszawa 2000.

XI

System semantyczny. Relacje paradygmatyczne. Pole semantyczne

Każda jednostka języka wchodzi w systemie języka w pewne relacje z innymi jednostkami systemu, przede wszystkim z innymi elementami zbioru wyrażen danego poziomu struktury językowej, natomiast jednostki tworzące wyrażenie złożone wchodzi w relacje z innymi jednostkami tworzącymi to wyrażenie. Te pierwsze relacje, między jednostką systemu lub tekstu a innymi elementami systemu (relacje systemowe) nazywamy **relacjami paradygmatycznymi**, natomiast stosunki zachodzące między jednostkami wyrażen złożonych (relacje tekstowe) – **relacjami syntagmatycznymi** (od wyrazu **syntagma** oznaczającego zespolenie kilku wyrażen następujących po sobie w danym ciągu wypowiedzeniowym). Zarówno relacje syntagmatyczne, jak i relacje paradygmatyczne zachodzą na wszystkich poziomach struktury języka, nas jednak najbardziej będą interesowały **relacje paradygmatyczne w systemie semantycznym** języka. Poprzednio mówiliśmy już o nich, teraz omówimy je w sposób bardziej systematyczny.

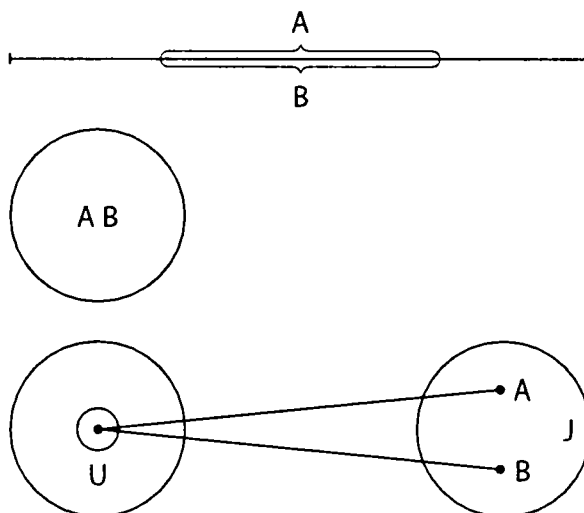
Przede wszystkim należy pamiętać o tym, że relacje paradygmatyczne są stosunkami między jednostkami języka, a nie między oznaczanymi przez nie desygnatami (rzeczami), mówią więc jedynie o strukturze języka, nie zaś o strukturze rzeczywistości. Te dwie płaszczyzny są bardzo często mylone, co może się stać przyczyną nieporozumień. Termin relacje (stosunki) paradygmatyczne bywa rozumiany szeroko i oznacza relacje między tymi elementami w systemie języka, które mogą występować w tym samym lub podobnym kontekście, a więc między wyrażeniami tworzącymi tę sama kategorię syntaktyczną, na przykład między przymiotnikami mogącymi wystąpić w funkcji przydawki w kontekście

Widziałam ... film

a więc takimi jak: *dobry, interesujący, ciekawy, najnowszy, zabawny, nudny* itd. Najczęściej jednak termin ten używany jest w węższym sensie i oznacza **relacje zachodzące między denotacją (zakresem) wyrażen**, a w logice między zakresami nazw (pełnych).

Między zakresami dwóch wyrażen może istnieć tylko jeden z poniższych stosunków (dla lepszej ilustracji przedstawiamy je graficznie w trzech różnych wersjach):

I. Równoważność (synonimia)



Zakres wyrażenia A równa się zakresowi wyrażenia B.

Taki stosunek zachodzi między parami *auto* i *samochód* oraz *zięć* i *mąż córki*, nie ma bowiem w rzeczywistości takich przedmiotów, o których można orzec, że są *autami*, a o których nie można orzec, że są *samochodami*, czyli że każde auto jest samochodem i każdy samochód jest autem. Analogicznie: każdy zięć jest mężem córki i każdy mąż córki jest zięciem.

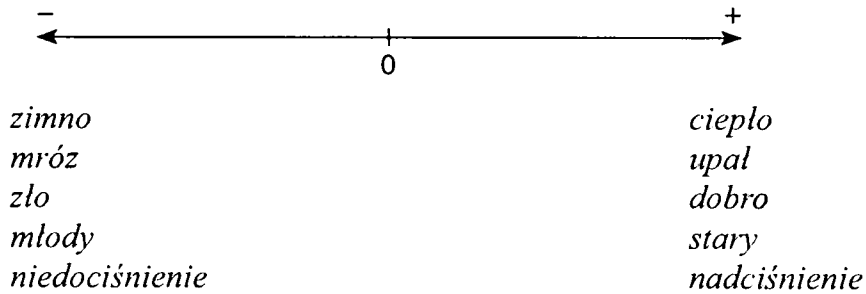
Zakres wyrażenia A równa się więc zakresowi wyrażenia B wtedy i tylko wtedy, gdy każdy obiekt, który jest nazwany A, może być nazwany B i każdy obiekt, który jest nazwany B, może być nazwany A.

Relację synonimii omówiliśmy dokładniej w poprzednim rozdziale.

W przypadku takiej relacji między terminami tworząc paranaturalny język informacyjno-wyszukiwawczy, wybieramy na deskryptor tylko jeden z synonimów, kierując się takimi względami, jak na przykład: stopień rozpowszechnienia terminu, przejrzystość jego budowy słowotwórczej (znaczenie strukturalne), poprawność językowa, jednoznaczność (eliminując ewentualny synonim homonimiczny – np. homonimiczny *wóz* jako synonim terminów *auto*, *samochód*), czy wreszcie długość terminu (lepszy krótszy, o ile spełnia wymogi redundancji).

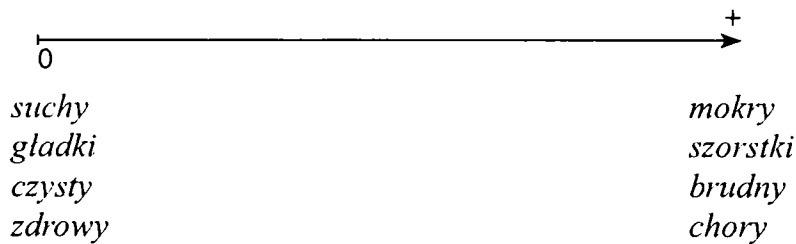
W językoznawstwie za przeciwieństwo synonimii uważa się **antonimie**, czyli, mówiąc najogólniej, relację zachodzącą między wyrażeniami o **przeciwstawnym znaczeniu**, takimi jak: *slaby – silny*, *jasny – ciemny*, *szlachetność – podłość*, *upal – mróz*, *dobro – zło*. Antonimami nazywa się więc wyrażenia połączone różnymi relacjami semantycznymi (choć zawsze jest to relacja wykluczania się zakresów) i dlatego, na przykład dla potrzeb wyszukiwania informacji, nie można ich traktować jednakowo. Wyróżnić tu można następujące typy:

1. **Wyrażenia parametryczne**, określające własności stopniowalne na skali wielowartościowej, przy czym wyrażenia te nazywają tylko punkty biegunowe skali lub biegunowe przedziały wartości, np.:



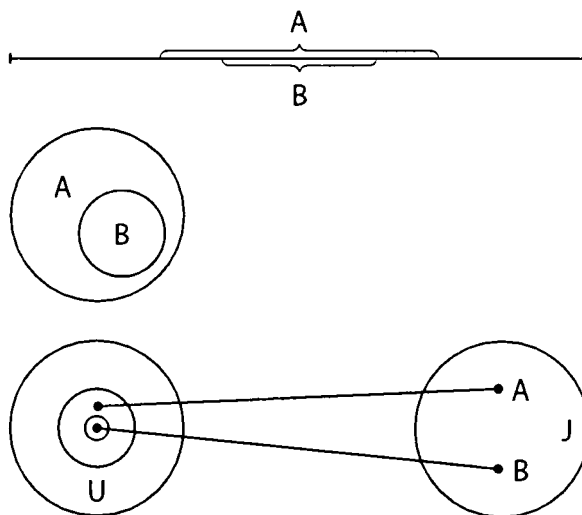
2. **Wyrażenia nieimplikujące stopniowalności cech**, np.: *start – meta, ślub – rozwód, analiza – synteza, przyływ – odpływ, rufa – dziób, parowanie – skraplanie, dedukcja – indukcja, reformacja – kontrreformacja, faszyzm – antyfaszyzm*, w których nie tak łatwo określić, na czym polega owa przeciwstawność znaczenia – są tu nazwy procesów odwrotnych, wyrażenia zawierające znaczenia początku i końca itp.

3. **Wyrażenia o znaczeniu komplementarnym**, a więc takie, że znaczenie jednego wyrażenia jest takie samo jak znaczenie negacji drugiego. Wyrażenia te nazywają cechy określone na skali binarnej – jedno komunikuje brak cechy (stopień zerowy), a drugie jej obecność, np.:



4. **Wyrażenia o znaczeniu konwersywnym**, a więc połączone relacją odwrotną, np.: *kupić – sprzedać, rodzice – dzieci, mąż – żona, nauczyciel – uczeń, poprzedzać – następować, większy – mniejszy*.

II. Nadrzędność (termin szerszy, hiperonimia)



Zakres wyrażenia A jest nadrzędny w stosunku do zakresu wyrażenia B wtedy i tylko wtedy, gdy nie każdy obiekt, który może być nazwany A, może być również nazwany B (nie każde A jest B), natomiast każdy obiekt, który może być nazwany B, może być także nazwany A (każde B jest A).

Taki stosunek zachodzi między następującymi parami: *kobieta – studentka* (nie każda kobieta jest studentką, ale każda studentka jest kobietą), *związek organiczny – aminokwas* (nie każdy związek organiczny jest aminokwasem, ale każdy aminokwas jest związkiem organicznym). Relacja taka nosi również nazwę **relacji rodzaj-gatunek** (termin szerszy jest nazwą rodzajową, termin węższy – nazwą gatunkową).

III. Podrzędność (termin węższy, hiponimia)

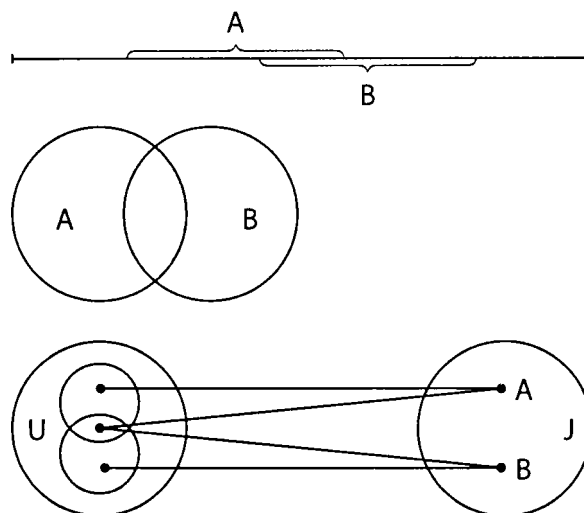
Podrzędność jest odwróceniem stosunku nadrzędności.

Zakres wyrażenia A jest podrzędny w stosunku do zakresu wyrażenie B wtedy i tylko wtedy, gdy zakres wyrażenia A jest zawarty w zakresie wyrażenia B, czyli, gdy każdy obiekt, który można nazwać A, można również nazwać B (każde A jest B), ale nie każdy obiekt, który można nazwać B, można również nazwać A (nie każde B jest A).

Taki stosunek zachodzi między parami *studentka – kobieta*, (każda studentka jest kobietą, ale nie każda kobieta jest studentką), *aminokwas – związek organiczny* (każdy aminokwas jest związkiem organicznym, ale nie każdy związek organiczny jest aminokwasem). Relacja taka nosi również nazwę **relacji gatunek – rodzaj**.

Relacje podrzędności i nadrzędności zakresowej, nazywane relacjami generycznymi, występują w wielu językach informacyjno-wyszukiwawczych – omówimy je dokładniej w rozdziale XVI.

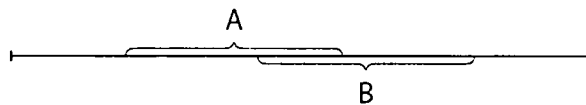
IV. Krzyżowanie się zakresów



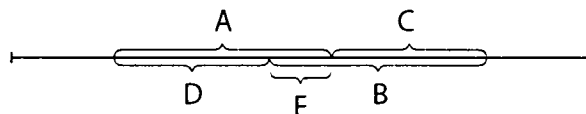
Zakresy wyrażen A i B krzyżują się wtedy i tylko wtedy, gdy pewne obiekty nazwane A można również nazwać B, pewne obiekty nazwane B można również nazwać A, ale istnieją również obiekty nazwane A, których nie można nazwać B, oraz obiekty nazwane B, których nie można nazwać A.

Relacja taka zachodzi na przykład między parami: *studentka* – *panna* oraz *inżynier* – *pracownik informacji*, bo istnieją studentki będące pannami, istnieją panny będące studentkami, ale istnieją studentki niebędące pannami i istnieją panny niebędące studentkami, oraz: niektórzy inżynierowie są pracownikami informacji, ale nie wszyscy, i niektórzy pracownicy informacji są inżynierami, ale nie wszyscy.

Stosunek krzyżowania się zakresów może niekiedy spełniać taki warunek, że krzyżujące się zakresy obejmują w sumie cały zakres uniwersalny, czyli że wtedy każdy przedmiot uniwersum należy przynajmniej do jednego z krzyżujących się zakresów, co ilustruje rysunek:



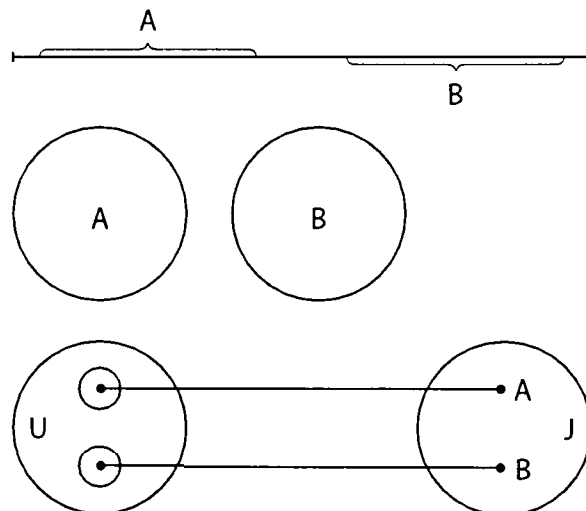
Przykładem takiej relacji może być para *ptak* i *niewróbek*, co odpowiada następującej sytuacji



gdzie: A – ptak, B – niewróbek, C – nieptak, D – wróbek, E – niewróbek będący ptakiem (czyli ptak inny niż wróbek).

Relacje krzyżowania się zakresów nie są relacjami par excellence językowymi – do ustalenia zbioru wyrażen połączonych tymi relacjami często nie wystarcza sama znajomość systemu językowego – konieczna jest także wiedza o świecie (o rzeczywistości pozajęzykowej).

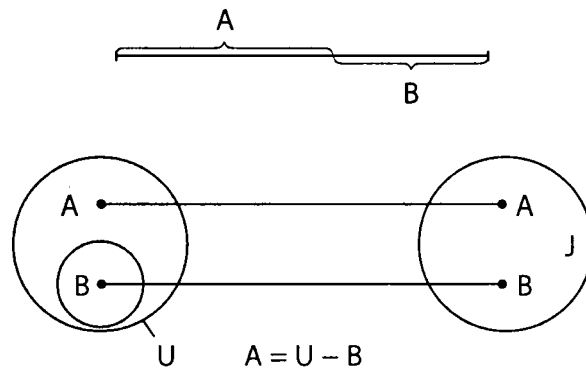
V. Wykluczanie się (rozłączność) zakresów



Zakres wyrażenia A wyklucza się z zakresem wyrażenia B wtedy i tylko wtedy, gdy żadnego obiektu nazwanego A nie można nazwać B i żadnego obiektu nazwanego B nie można nazwać A.

Tak jest zawsze w wypadku jedno-jednoznacznego przyporządkowania między elementami uniwersum i elementami odwzorowującego je języka. Przykłady: *wróbel* – *kukułka*, *samochód* – *motocykl*, *wróbel* – *motocykl*, bo żaden wróbel nie jest kukułką i żadna kukułka nie jest wróblem, żaden samochód nie jest motocyklem i żaden motocykl nie jest samochodem, żaden wróbel nie jest motocyklem i żaden motocykl nie jest wróblem.

Przy tworzeniu klasyfikacji jednym z najważniejszych wymogów poprawności jest rozłączność zakresów wyodrębnionych klas. Taką rozłączność zapewnia **podział dychotomiczny**, dokonywany według cechy A – nie-A (np. podział materii na ożywioną i nieożywioną). Jest to szczególny przypadek wykluczania się zakresów, gdy wykluczające się zakresy dopełniają się, dając w sumie zakres uniwersum. Wtedy każdy przedmiot musi należeć do jednego z wykluczających się zakresów, co pozwala na spełnienie jeszcze jednego warunku stawianego poprawnej klasyfikacji, która powinna być pełna (kompletna, wyczerpująca). Ten wypadek wykluczania się można przedstawić graficznie w sposób następujący:



Zgodnie z założeniami logiki o stosunkach tych można mówić właściwie tylko w wypadku nazw i to nazw niepustych, tylko im bowiem przysługuje własność posiadania zakresów rozumianych jako zbiór ich desygnatów. Pomiedzy zakresami nazw pustych zachodziłby więc zawsze stosunek równoważności, np. pomiędzy zakresami nazw *syrena*, *pegaz*, *centaur*, bo w rzeczywistości nie istnieją ani syreny, ani pegazy, ani centaury, a więc nazwy te desygnują tym samym zbiór pusty, który jest tylko jeden. Nazwy te byłyby więc synonimiczne, a co więcej – byłyby podrzędne w stosunku do każdej nazwy niepustej (bo zbiór pusty mieści się w każdym zbiorze niepustym). Z punktu widzenia języka naturalnego nie chcielibyśmy jednak przyjmować takiej synonimii, gdyż nazwy te używane są tak, jak gdyby istniały syreny, pegazy, centaury. Stwierdzenie, czy one naprawdę istnieją, czy nie istnieją, nie zależy od wiedzy o języku, ale od wiedzy o świecie, która jest zmienna w czasie.

Analiza komunikatu pokazuje, że wyrażenia należące do różnych poziomów drzewa klasyfikacyjnego odzwierciedlające relacje paradygmatyczne używane są z różną częstością. W każdym takim drzewie wyróżnić można poziom wyrażen używanych najczęściej, przede wszystkim ze względu na ich wartość informacyjną.

Na przykład normalnie mówimy:

To jest krzesło. a nie To jest mebel. To jest przedmiot.

Kowalski hoduje rybki. a nie Kowalski hoduje zwierzęta.

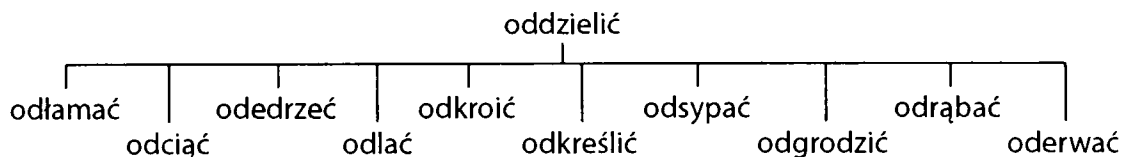
Kupiłam sobie żółtą sukienkę. a nie Kupiłam sobie żółtą sztukę odzieży.

Powiemy jednak:

Ugryzł mnie jakiś owad. (jeśli nie wiemy jaki),

Dostałam mnóstwo kwiatów. (bo nieważne jakie).

Przy okazji rozważań poświęconych stosunkom paradygmatycznym chcielibyśmy jeszcze zwrócić uwagę na związek między **zakresem** jakiejś jednostki leksykalnej a jej zawartością informacyjną (**treścią**). Obowiązuje tu stosunek odwrotnie proporcjonalny: im większy zakres, tym mniejsza pojemność treściowa, tym mniej informacji niesie dane wyrażenie (i oczywiście, im większa treść tym mniejszy zakres). Tak więc, polski czasownik *oddzielić* będzie miał większy zakres niż np.: *odłamać*, *oderwać*, *odedrzeć*, *odciąć*, *odkroić*, *odlać*, *odsypać*, *odkreślić*, *odrąbać*, ale mniejszą treść, gdyż te ostatnie czasowniki niosą ponadto informację o tym, w jaki sposób coś zostaje oddzielone. Hierarchiczne drzewa klasyfikacyjne mogą więc tworzyć nie tylko rzeczowniki, ale również inne części mowy, dla wyżej podanych czasowników mamy więc drzewo:



Hierarchiczna budowa systemu semantycznego języka naturalnego pozwala nadawcy komunikatu na dostosowanie szczegółowości zawartych w przekazie informacji do potrzeb adresata. Na przykład do różnych adresatów skierowane są komunikaty:

Kowalski odziedziczył nieruchomość.

Kowalski odziedziczył dom.

Kowalski odziedziczył kamienicę.

Nie byłoby to możliwe przy płaskiej strukturze systemu semantycznego.

Znajomość **struktury paradygmatycznej języka naturalnego** niezbędna jest przy tworzeniu i dekodowaniu **tekstu spójnego**, którego reguły podporządkowane są nie tylko funkcji informacyjnej, ale także pewnym wymogom natury estetycznej. Jedną z takich reguł stanowi, że jeżeli dany element rzeczywistości staje się przedmiotem kilku pozostających w bezpośrednim sąsiedztwie informacji, to w każdej z nich powinien być odwzorowywany przez inny wykładnik językowy. Unikanie powtarzania tych samych wyrazów umożliwia, między innymi, synonimia i hiperonimia – synonimy i odpowiednio uszeregowane w tekście hiperonimy pozwalają uczynić zadość zasadom stylistyki, jednocześnie gwarantu-

jąc tożsamość oznaczanego nimi obiektu. Przykładem wykorzystania synonimii w tekście spójnym jest następujący tekst:

Lubię to moje stare auto, ale przydałby mi się lepszy wóz. Mam jednak ważniejsze potrzeby, więc nowy samochód będzie musiał jeszcze poczekać.

Przykładem posłużenia się hiperonimiami może być tekst podany w *Praktycznej stylistyce* Wierzbickich:

„Na podwórzu zanosił się wprost ujadaniem mały kundelek. Patrzcie jaki śmieszny! – pisnęła Agatka. Pozwoliłem jej pogłaskać pieska. Zwierzak nastroszył się, szczerknął jeszcze głośniejsze, a potem ucichł nagle i rozwalil się brzuszkiem do góry. Spodobało mu się widać głaskanie Agatki”.

Przy okazji chcielibyśmy zwrócić uwagę na to, że bardzo często o znaczeniu wielu wyrazów wiemy tylko to, do jakiej szerszej kategorii należą ich desygnaty, znamy więc relacje paradygmatyczne, nie znając ich znaczenia: ani denotacji, ani dokładnej konotacji, choć zazwyczaj znamy niektóre cechy konotacyjne. I tak, na przykład, wiele osób wie, że *tuberoza* to pachnący kwiat, *raniuszka* to mały ptak, *janowiec* to krzew, *preferans* to gra w karty, *prunelki* to pantofle, ale obiektów tych nie umiałoby wskazać (nie zna więc denotacji tych nazw), nie zna też przysługujących im cech wyznaczających konotacje tych wyrażeń.

Relacje paradygmatyczne należą do szeroko rozumianych **relacji asocjacyjnych**, czyli **relacji skojarzeniowych**. W umyśle użytkownika języka to właśnie te wyrażenia najczęściej się ze sobą kojarzą, struktura paradygmatyczna języka jest bowiem ważną częścią kompetencji językowej, a do orzeczenia o tym, jakiego typu relacje paradygmatyczne łączą dane wyrażenia, wystarcza tylko znajomość systemu językowego. Dlatego relacje te nazywa się **relacjami systemowymi**. Z danym wyrażeniem kojarzą się użytkownikowi języka także wyrażenia niepołączone relacjami systemowymi. Dobrą ilustracją takich skojarzeń jest opowiadanie Czechowa zatytułowane *Końskie nazwisko*. Bohater opowiadania zapomniał, jak nazywa się jego znajomy, pamiętał tylko tyle, że jest to „końskie nazwisko”. Próby podpowiedzi utworzyły następujący łańcuch asocjacyjny (podaję według polskiego przekładu):

Kobyłski, Ogierski, Żrebięcki, Kobylecki, Kobyliński, Żróbczyński, Koniński, Szkapowski, Żrebakiewicz, Klaczkowski, Klaczkiewicz, Zaprzężny, Zaprzęgowski, Orczykowski, Żrebacki, Ogierowicz, Ogierkowski, Szkapowski, Chabeciki, Stadnikiewicz, Klusowski, Rumacki, Wierzgalski, Tabuński, Kopytowski, Brykalski, Konienko, Konikowski, Truchciński, Kobyłkowski, Parokoński, Uzdowski, Gniadowski, Cwaliński, Rozpędowicz, Wałachowski, Bulański, Siodelkiewicz, Chomątowicz, Czapracki, Owsianko.

Relacje asocjacyjne w języku są więc relacjami zachodzącymi między wyrażeniami podobnymi pod względem formy lub treści albo między wyrażeniami występującymi często razem w jednym kontekście. Wyrażenia powiązane tymi relacjami tworzą tzw. **pole asocjacyjne**. Teoria pól asocjacyjnych rozwijana była

przez wielu językoznawców (np. F. de Saussure, Ch. Bally, J. Trier, G. Ipsen, W. Porzig), miała ona jednak swoich poprzedników wśród leksykografów, którzy sporządzali słowniki w tzw. układzie rzeczowym, grupującym obok siebie wyrazy języka naturalnego mające wspólny element znaczeniowy. Słowniki takie miały być pomocne tym, którzy dobrze wiedzieli, co chcą wyrazić, ale nie zawsze wiedzieli, jakie słowa do tego celu najlepiej się nadają. Były to więc pierwsze tezaury, a najbardziej znany spośród nich jest tezaurus Petera M. Rogeta, ułożony dla języka angielskiego, który powstał w roku 1852.

Teorię budowy takich słowników próbował stworzyć w roku 1910 R.M. Meyer, który przeprowadził klasyfikację **pól semantycznych (znaczeniowych)**, a więc klas wyrażen wyróżnianych ze względu na wspólny element znaczenia, na naturalne, sztuczne i półsztuczne. Naturalne odpowiadały różnicom i podobieństwom istniejącym w naturze – przykładem mogą tu być pola semantyczne grupujące nazwy zwierząt, roślin, minerałów; sztuczne grupowały wyrazy nazywające dysfunkcje sztucznie wytworzone przez człowieka, na przykład rangi wojskowe; półsztuczne grupowały nazwy twórców ludzkich, ale wykorzystujących element naturalny – na przykład terminologia myśliwska. Meyer próbował też określić cechę semantyczną wspólną dla wszystkich wyrazów tworzących jedno pole, na przykład dla wyrazów: *iść, biec, lecieć, maszerować, płynąć* cechą taką jest ‘poruszać się’ dla wyrazów: *obfitość, bujność, dużo* cechą taką jest ‘ilość większa niż normalnie’.

Za twórcę pierwszej pełnej **koncepcji pola znaczeniowego** uważa się Josta Triera (1931). Najważniejsze jej tezy głosiły:

- wszystkie treści, jakie można wyrazić w danym języku naturalnym, dzielą się na pewne strefy, a te z kolei na mniejsze pola;
- tak podzielonym treściom odpowiadają w języku wyrazy służące do ich wyrażania, które tworzą odpowiednie pola wyrazowe;
- wyrazy te pokrywają pole treści na zasadzie mozaiki;
- znaczenie każdego wyrazu zależy od całego układu pola i tylko w opozycji do pozostałych – na przykład w układzie świetlnym sygnalizacji ulicznej czerwone światło znaczy tylko dlatego, że istnieje światło zielone (i ewentualnie żółte). Izolowane słowo nic by nie znaczyło, gdyby człowiek, który je nadał lub odebrał, nie miał w pamięci całego układu pól semantycznych.

Twórcą samego terminu pole znaczeniowe jest Gunther Ipsen. Termin ten powstał przez analogię do pola grawitacyjnego – tak jak pole grawitacyjne skupia wokół centrum elementy rzeczywistości, przyciąga je z siłą zależną od natężenia i ustala układ elementów znajdujących się w jednym polu, tak w polu semantycznym wokół wyrazu stanowiącego jego centrum grupują się wyrazy mocniej, a im dalej od centrum, tym słabiej z nim skojarzone.. Ipsen uważał, że pole znaczeniowe obejmuje zespół wyrazów skojarzonych nadrzędną treścią znaczeniową mających określone formalne wykładniki owej cechy semantycznej łączności. Przykładem takiego pola znaczeniowego mogłaby być tzw. rodzina wyrazów, czyli klasa wyrazów zawierających formalny wykładnik łączności semantycznej w postaci wspólnego rdzenia, np.:

*ręka, ręcznik, naręcze, podręczny, podręcznik, zaręczyny, wyreka, wyręczyć, ręczyć, obrączka, rączka, doręczyć, doręczyciel, poręczyciel...
czytać, czytnik, czytno, czytanka, przeczytać, wyczytać, zaczytać, odczyt, doczytać, czytacz, czytelnia, czytelnik, czytelnictwo...*

F. de Saussure uważał, że wyraz stanowiący centrum pola kojarzy się przede wszystkim z czterema ciągami słownymi:

- z wyrazami genetycznie pokrewnymi (przykłady jak wyżej),
- ze słowami o analogicznej strukturze (na przykład na takiej zasadzie kojarzą się wyrazy: *biegacz, pakowacz, kopacz, wybielacz, utrwalacz, zjadacz...*),
- z synonimami,
- z wyrazami podobnymi pod względem dźwiękowym (na takiej zasadzie kojarzą się wyrazy: *koliszki, kieliszki, Koluszki; pryncypalny i pryncypialny; adoptować, adaptować.*

Ch. Bally twierdził, że w skład pola wchodzi wyrazy skojarzone ze słowem-centrum różnorodnymi związkami semantycznymi, na przykład słowo *wół* jest punktem przecięcia następujących szeregów asocjacyjnych:

- *krowa, byczek, cielę, rogi, przeżuwacz*, a więc należą tu wyrazy połączone relacją „bycia bezpośrednio w naturalny sposób związanym z cechą wołowości”,
- *orka, plug, jarzmo, mięso, rzeźnia*, grupujący wyrazy połączone z wyrazem stanowiącym centrum pola relacją „bycia obiektem lub dziedziną życia, w której wykorzystuje się woły lub do korzystania z której konieczne są woły”,
- *sila, wytrzymałość, cierpliwość, pracowitość*, a więc wyrazy związane z wołem relacją znaczeniową „bycie cechą przypisywaną zwykle wołowi”.

Interesującą koncepcję pola znaczeniowego stworzył Walter Porzig (1950). Według niego przynależność wyrazów do jednego pola semantycznego uwarunkowana jest ich występowaniem we wspólnych kontekstach, gdyż użycie jednego wyrazu implikuje użycie innych wyrazów związanych z nim semantycznie różnego typu relacjami semantycznymi asocjacyjnymi – na przykład dla nazwy procesu są to wyrazy o znaczeniu „bycie instrumentem procesu”, „bycie agensem procesu”. Centrum takiego pola stanowi czasownik, który otwiera w tekście miejsce dla wyrazów nazywających elementy wyrażanego przez ten czasownik procesu (konotuje je) np.:

szczekać: pies, hiena, wilk...

jeździć: jeździec, amazonka, dżokej, koń, siodło, strzemień, wodze, kłus, galop...

Podobnie dzieje się z przymiotnikami, konotującymi odpowiednie rzeczowniki, które mogą określać, np.:

piwne: oczy,

gniady: koń,

blond: włosy.

Pole semantyczne tworzą, według Porziga, także wyrazy zawierające ten sam komponent semantyczny, np.:

dla cechy „wierchowiec”: *koń, muł, osioł, wielbłąd...*

dla cechy „poruszanie się”: *iść, biec, płynąć, frunąć.*

Do ciekawych wniosków na temat pola asocjacyjnego prowadzą eksperymenty. Wyniki takiego eksperymentu przytoczone zostały w pracy *Slovar asociativnykh norm russkogo jazyka* (pod red. A.A. Leonteva). Eksperyment ten, przeprowadzony wśród Amerykanów, Niemców, Francuzów, Polaków, Rosjan, Kazachów, Kirgizów i Uzbeków, miał na celu ustalenie najczęstszych skojarzeń ze słowami: *głodny, chleb, masło, owoc* (ros. *plod*). Wyniki tego eksperymentu przytaczam poniżej.

Ze słowem *głodny* kojarzy się:

- Amerykanom: jedzenie, jeść, trawiony głodem, syty, umierający z głodu, żołądek;
Niemcom: trawiony głodem, jeść, syty, chleb, żołądek;
Francuzom: głód, chleb, wilczy głód, jeść, żołądek, nędza;
Polakom: (jak) wilk, (jak) pies, człowiek zły, syty;
Rosjanom: wilk, syty, rok, zimny, pies;
Kazachom: syty, bardzo głodny, (jak) wilk, człowiek, głód;
Kirgizom: syty, wilk, łakomy, goły, lisica;
Uzbekom: syty, wilk, łakomy, bardzo, lisica, obdarty;

Ze słowem *chleb* kojarzy się:

- Amerykanom: masło, jedzenie, jeść, woda, biały;
Niemcom: jeść, masło, bochenek, jedzenie, głód;
Francuzom: wino, biały, jeść, głód, miękisz;
Polakom: powszedni, razowy, masło, czerstwy, świeży;
Rosjanom: biały, sól, masło, powszedni, czerstwy, bułka, smaczny, świeży;
Kazachom: masło, woda, mąka, cukier, jeść, trójnóg (jako podstawa do naczynia);
Kirgizom: biały, pszenica, masło, chleb domowy, herbata, świeży, cukier, gorący, sól;
Uzbekom: głodny, herbata, woda, cukier, jeść.

Ze słowem *masło* kojarzy się:

- Amerykanom: chleb, żółte, miękkie, tłusty, jedzenie;
Niemcom: chleb, miękkie, mleko, margaryna, ser;
Francuzom: miękkie, żółte, świeże, chleb, tłuste, mleko;
Polakom: chleb, margaryna, maślany, tłuszcz, świeże;
Rosjanom: śmietankowe, maślane, okrągły, sadło, chleb, tłuszcz, ogień, kasza, mięso;

Kazachom:	sadło, śmietankowe, twaróg, bryndza, bawełniane, jeść;
Kirgizom:	topione, mleko, słodkie, ajran, jeść, pokarm, jedzenie;
Uzbekom:	ogień, śmietankowe, mięso, mleko, woda, sadło.

Ze słowem *owoc* kojarzy się:

Amerykanom:	jabłko, warzywo, pomarańcze, mucha, jagoda;
Niemcom:	warzywa, jabłko, drzewo, gruszka;
Francuzom:	jabłko, dojrzałe, pomarańcza, gruszka, brzoskwinia;
Polakom:	jabłko, słodkie, smaczne, dojrzałe, soczyste, zakazany;
Rosjanom:	jabłko, miłości, pracy, smaczny, zakazany, warzywo, fantazji;
Kazachom:	jagoda, jabłko, suszone morele, płody, roślina, winogrona, drzewo, rodzynki;
Kirgizom:	jabłko, sad, warzywa, słodkie, dużo;
Uzbekom:	różne owoce, jabłko, winogrona, sady.

Z analizy powyższego materiału widać, że do jednego pola skojarzeniowego mogą wchodzić:

- wyrazy oznaczające rodzaj, np.:

masło – tłuszcz, jedzenie (pokarm);

chleb – jedzenie (pokarm);

owoc – płody;

- wyrazy oznaczające gatunek, np.:

owoc – pomarańcza, jabłko, brzoskwinia, gruszka, jagoda, winogrona, rodzynki;

- wyrazy oznaczające cechę gatunkową, np.:

chleb – razowy;

masło – śmietankowe, topione;

- wyrazy oznaczające cechę właściwą, np.:

chleb – biały, czerstwy, świeży, smaczny, gorący;

masło – żółte, miękkie, świeże, tłuste, okrągłe, słodkie;

owoc – dojrzałe, słodkie, smaczne, soczyste;

- wyrazy oznaczające typowy kształt (ukształtowanie), np.:

chleb – bochenek;

- wyrazy oznaczające obiekty pełniące taką samą funkcję (terminy współrzędne), np.:

chleb – bułka;

masło – margaryna, sadło;

owoc – warzywo;

- wyrazy oznaczające obiekty mające takie samo pochodzenie, np.:
masło – ser, twaróg, mleko, ajran, bryndza;
- wyrazy oznaczające surowiec lub źródło, np.:
chleb – pszenica, mąka;
masło – mleko;
owoc – drzewo, sad;
- wyrazy oznaczające inne przedmioty często współwystępujące w sytuacji, np.:
chleb – masło, wino, sól, woda, cukier, herbata;
masło – chleb, kasza, mięso, woda;
owoc – mucha;
- wyrazy oznaczające przeznaczenie, np.:
chleb – jeść;
- wyrazy oznaczające sytuację, w jakiej występuje dane zjawisko, np.:
głodny – nędza;
- wyrazy o znaczeniu przeciwnym (antonimy), np.:
głodny – syty;
- wyrazy szczególnie często współwystępujące w kontekście (najczęściej w stałych związkach frazeologicznych), np.:
głodny – trawiony głodem, wilczy głód, jak wilk, jak pies, wilk, lisica, człowiek zły (por. polskie przysłowie: Polak jak głodny, to zły);
chleb – powszedni;
owoc – zakazany, miłości, pracy, fantazji.

Oczywiście powyższe relacje nie wyczerpują wszystkich możliwości, pokazują jednak, że skojarzenia takie mogą być uwarunkowane:

- strukturą języka – relacje paradygmatyczne,
- strukturą tekstu – kojarzenie się wyrazów współwystępujących w związkach frazeologicznych,
- względami kulturowymi, a więc pozajęzykowymi i pozatekstowymi, czego przykładem mogą być występujące w eksperymencie skojarzenia:

chleb – wino dla Francuzów,
chleb – sól dla Rosjan,
chleb – herbata dla Kirgizów i Uzbeków.

Te ostatnie względy mają również wpływ na pewną część uwarunkowań kontekstowych.

Literatura

- Apresjan J.D.: *Antonimy leksykalne i przekształcenia wyrażen antonimicznych*, w: Mayenowa M.R. (red.) *Tekst i język. Problemy semantyczne*. Wrocław 1974.
- Apresjan J.D.: *Semantyka leksykalna. Synonimiczne środki języka*. Wrocław 2000.
- Bojar B.: *Pole informacyjne, dokumentacyjne, tematyczne, semantyczne czy leksykalne?* „Zagadnienia Informacji Naukowej” 2004, nr 4.
- Buttler D.: *Koncepcje pola znaczeniowego*. „Przegląd Humanistyczny” 1967, nr 2.
- Marciszewski W.: *Relacje tematyczno-hierarchiczne w językach deskryptorowych*. „Zagadnienia Informacji Naukowej” 1973, nr 2.
- Miodunka W.: *Teoria pól językowych. Społeczne i indywidualne ich uwarunkowanie*. Warszawa-Kraków 1980.
- Sosińska B.: *Typologia relacji pomiędzy jednostkami leksykalnymi języków informacyjnych*. „Zagadnienia Informacji Naukowej” 1979, nr 2.
- Tokarski R.: *Struktura pola znaczeniowego*. Warszawa 1984.
- Wereszczyńska-Cisło B.: *Relacje kojarzeniowe w językach informacyjnych*. „Zagadnienia Informacji Naukowej” 1985, nr 1.

XII

Znaczenie

Jak widać z naszych dotychczasowych rozważań, termin znak używany bywa w literaturze w wielu różnych znaczeniach. Taka wieloznaczność nie jest cechą tylko tego terminu, jeszcze więcej znaczeń ma sam termin znaczenie, przy czym zdefiniowanie terminu znaczenie wydaje się zadaniem teoretycznie niemożliwym, gdyż aby ustalić znaczenie poszczególnych „znaczeń”, trzeba wprawdzie wiedzieć, czym jest znaczenie, a w tym wypadku nie wystarczy już tylko intuicyjne rozumienie tego terminu. C.K. Ogden i I.A. Richards różnym znaczeniom terminu znaczenie poświęcili specjalną monografię *The Meaning of Meaning* (Londyn 1923) i po gruntownej analizie wyróżnili aż 23 różne znaczenia „znaczenia” (nie miejsce tu na ich analizę). Adam Schaff we *Wstępie do semantyki* podał następującą listę możliwości interpretacji **znaczenia**:

- znaczenie to przedmiot, którego nazwą jest znak,
- znaczenie to właściwość przedmiotów,
- znaczenie to przedmiot idealny, względnie inherentna właściwość myśli,
- znaczenie to relacja:
 - między znakami,
 - między znakiem a przedmiotem,
 - między znakiem a myślą o przedmiocie,
 - między znakiem a działaniem ludzi,
 - między ludźmi porozumiewającymi się za pomocą znaków.

Mała encyklopedia logiki dokonuje analizy znaczenia zgodnie z dziedziną, w jakiej jest ono umiejscawiane, podając następujące możliwości:

- sfera zjawisk psychicznych,
- sfera przedmiotów fizycznych,
- sfera przedmiotów abstrakcyjnych,
- sfera języka,
- sfera stosunków między ludźmi, językiem a rzeczywistością zewnętrzną (po-zajęzykową).

Z różnych rozstrzygnięć dotyczących istoty znaku, bilateralnej i unilateralnej koncepcji znaku wynika ponadto, że znaczenie można traktować albo jako część znaku, albo jako coś pozostającego poza znakiem.

Już z powyższego zestawienia widać, że listę wątpliwości, które należy rozstrzygnąć, rozpatrując znaczenie terminu, należałoby uzupełnić o jeszcze jeden przynajmniej problem – rozważania dotyczące statusu bytów, którym przysługuje cecha posiadania znaczenia, i tu bowiem możliwe są różne ustalenia. Mówiąc najogólniej, wyróżnić tu można dwa stanowiska:

- pierwsze, według którego znaczenie przysługuje tylko zdaniom lub znakom użytym w funkcji zdań, a więc komunikatom znakowym,
- drugie, które przypisuje znaczenie także wyrażeniom niebędącym zdaniami, a więc wyrazom i morfemom (przypomnijmy, że morfem definiowany jest jako najmniejsza jednostka systemu językowego mająca znaczenie).

Nie możemy tu omówić wszystkich koncepcji znaczenia, zatrzymamy się tylko na kilku najważniejszych, najczęściej przyjmowanych w literaturze, najbardziej przydatnych dla naszych dalszych rozważań.

Znaczenie jako przedmiot oznaczany – denotacja

Potocznie mówiąc, zrozumieć znaczenie jakiegoś znaku, to tyle, co umieć przyporządkować mu (wskazać) element rzeczywistości, o którym ten znak informuje. Przy takim rozumieniu znaczenia, znaczeniem znaku jest zbiór przedmiotów przez ten znak oznaczany, czyli jego **denotacja**. Zgodnie z tym, dwa wyrażenia *gwiazda poranna* i *gwiazda wieczorna* mają takie samo znaczenie, ponieważ denotują ten sam element rzeczywistości (jest nim planeta Wenus). Gottlob Frege, który ten przykład rozpatrywał (nawiązując do koncepcji denotacji J.S. Milla), ten rodzaj znaczenia nazywa **oznaczaniem**, inni, jak Russel **denotacją**, a jeszcze inni **referencją**.

Przy takim rozumieniu znaczenia przysługiwałoby ono tylko nazwom, i to **nazwom pełnym**, mającym swój **desygnat**, czyli realnie istniejący przedmiot oznaczany przez daną nazwę. Nie można wskazać (w sensie dosłownym) elementów rzeczywistości, które można by przyporządkować takim wyrażeniom jak: *syrena*, *centaur*, *pegaz*, bo one nie istnieją realnie, ani takim wyrazom jak: *przyjaźń*, *miłość*, *wojna*, bo oznaczają relacje lub sytuacje, ani takim jak: *pisać*, *czytać*, *uczyć się*, bo oznaczają czynności lub procesy. Wyrażenia języka naturalnego nie przesądzają o istnieniu desygnatów, ale używane są tak, jak gdyby desygnaty te istniały. To, jak jest w rzeczywistości, zależy nie od wiedzy o języku, ale od wiedzy o świecie, a ta jest zmienna.

Sprawa nabiera szczególnej wagi w przypadku bogatej przecież literatury naukowej o bytach hipotetycznych, na przykład różnego typu cząsteczkach materii, których istnienie lub nieistnienie przyjmuje się w teorii. Doświadczalne potwierdzenie ich istnienia jest bowiem często w danym momencie rozwoju nauki niemożliwe. Aby uniknąć tych konsekwencji, nazwy puste bywają niekiedy traktowane jako nazwy tworów złożonych (na mocy definicji: z **nazw niepustych**), np.:

centaur = nazwa oznaczająca twór złożony, w którym przednią połową jest mężczyzna, a tylną połową tylna połowa konia,

syrena = nazwa oznaczająca twór złożony, w którym górną połową jest górna połowa kobiety, a dolną połową tylna połowa ryby,
pegaz = nazwa oznaczająca twór złożony z konia i skrzydeł.

W takich wypadkach zakres nazwy wyznaczany jest przez zbiór cech przypisywanych jakimś bytom uniwersum – to, czy tak wyróżniony podzbiór uniwersum zawiera jakieś elementy, czy nie, użytkownika języka nie musi interesować. W takim wypadku desygnat wyrażenia, choć nie istnieje w świecie realnym, istnieje w tzw. **świecie możliwym** (termin teorii semantycznej Kripke’go).

Podobnemu zabiegowi przekształcenia w nazwy niepuste poddaje się również wszystkie **nazwy pozorne**, takie jak np.: *miłość*, *strzał*, *wojna*, *odwaga*, *zwycięstwo*, *obróbka*. Wyrażenia te również nie mają desygnatów w zwykłym rozumieniu tego słowa, nie ma bowiem takich rzeczy jak miłość, strzał, wojna, odwaga, zwycięstwo, obróbka – są tylko ludzie (lub inne obiekty), którzy kochają, strzelają, walczą, którzy są odważni, zwyciężają lub obrabiają jakieś przedmioty.

Wskazania **denotatu** wyrażenia nigdy zresztą nie można uznać za przesadzające o znajomości znaczenia, struktura paradygmatyczna języka naturalnego powoduje bowiem, że jeden przedmiot może należeć do wielu różnych klas, ponadto może być denotatem każdej przysługującej mu cechy, niekoniecznie konstytuującej konotację nazywającego go wyrażenia.

Znaczenie jako właściwość przedmiotów – konotacja

Znaczenie rozumiane jako denotacja przeciwstawia się zazwyczaj, tak jak to robili J.S. Mill, G. Frege i B. Russel, koncepcji znaczenia jako **konotacji**, czyli zbioru cech przedmiotów denotowanych przez dany znak. Zgodnie z tą koncepcją, wyrażenia *gwiazda poranna* i *gwiazda wieczorna* mają różne znaczenie, różną treść, bo wyrażają różne cechy denotowanych przedmiotów (poranna i wieczorna). O tych dwóch różnych znaczeniach terminu znaczenie G. Frege tak pisze (cyt. za: A. Schaffem, *Wstęp do semantyki*):

„Nazwa właściwa (słowo, znak, kombinacja znaków, wyrażenie) wyraża swe znaczenie, zastępuje zaś lub oznacza to, do czego się odnosi. Za pomocą znaku wyrażamy jego znaczenie i oznaczamy jego przedmiot”.

Określenie tak rozumianego znaczenia nie jest wcale proste, wymaga bowiem ustalenia, które spośród cech denotatów konstytuują tzw. **treść charakterystyczną**, a więc są cechami przysługującymi wszystkim przedmiotom oznaczanym przez dany znak. Ponadto zaś niektórzy uważają, że do znaczenia danego znaku wchodzi nie wszystkie cechy charakterystyczne, lecz tylko niektóre spośród nich, składając się na tzw. **treść językową**, a więc te, których znajomość wystarcza użytkownikowi znaku do rozstrzygnięcia, czy przedmiot je posiadający może być danym znakiem oznaczany. Aby lepiej wyjaśnić to wcale nieproste zagadnienie, przytoczymy podany w *Małej encyklopedii logiki* następujący przykład:

„Ktoś, komu zakomunikowano, że pewna figura da się wpisać w koło, ma boki parami równoległe i dwie prostopadłe przekątne, nie potrafi odpowiedzieć, czy ta figura jest kwadratem. Widać z tego, że nie zna on treści charakterystycznej nazwy „kwadrat”, nie wynika stąd jednak, że nie zna treści językowej, czyli

konotacji tej nazwy. Gdyby potrafił zidentyfikować kwadrat np. na podstawie cech równoboczności i prostokątności, dałby tym dowód, że rozumie słowo „kwadrat” zgodnie z jego znaczeniem w języku polskim, a więc zna konotację tego słowa”.

Konotacja, jak sądzą niektórzy językoznawcy (np. Jan Tokarski), zależna jest nie tylko od wiedzy językowej użytkownika, ale także od jego wiedzy pozajęzykowej (granice między wiedzą językową a wiedzą pozajęzykową jest zresztą bardzo trudno wytyczyć) i zmienia się w czasie – na przykład inna jest konotacja słowa *kot* dla małego dziecka (często jedyną cechą charakterystyczną jest, że „robi miau”) inna dla ucznia, a jeszcze inna dla zoologa, a o tym, że mimo to potrafią się oni ze sobą na temat kotów porozumiewać, decyduje znajomość denotacji.

W obu wymienionych wyżej znaczeniach terminu znaczenie — jako denotacji i jako konotacji, samo znaczenie umieszcza się w sferze przedmiotów fizycznych – desygnatów i ich cech, a więc w sferze rzeczywistości pozajęzykowej (poza-znakowej).

Znaczenie jako przedmiot idealny

Twórcą takiej teorii znaczenia jest Edmund Husserl. Znaczenie umieszcza on poza językiem i poza umysłem, przypisując mu cechy realnie istniejącego **przedmiotu idealnego**, istniejącego podobnie jak liczby i jakości. Pisze on o tym tak (cyt. za: A. Schaffem):

„Nie jest naturalnie naszym zamiarem istnienie przedmiotów idealnych stawiać w jednym rzędzie z istnieniem w myśli przedmiotów fikcyjnych lub w sobie sprzecznych (...). Przedmioty idealne (...) istnieją naprawdę. Jest rzeczą ewidentną, że nie tylko sensownie mówimy o takich przedmiotach (np. o jednej i tej samej liczbie 2, jednej i tej samej jakości: czerwień, jednej i tej samej zasadzie sprzeczności itp.) i przedstawiamy je sobie jako wyposażone w cechy, lecz także uchwytnujemy z oczywistością pewne prawdy kategoryczne odnoszące się do takich przedmiotów idealnych. Jeżeli te prawdy mają walor, to musi istnieć to wszystko, co jest przez ich walor obiektywnie złożone (...)”. Co więcej, według Husserla, owe przedmioty idealne, a więc i znaczenia, istnieją niezależnie od tego, czy istnieją oznaczające je znaki – Husserl pisze dalej:

„Każdy przypadek tworzenia jakiegoś nowego pojęcia pokazuje nam, jak realizuje się znaczenie, które przedtem nigdy jeszcze nie było zrealizowane. Tak jak liczby – w sensie idealnym, tym, które zakłada arytmetyka – nie powstają i nie giną wraz z aktem liczenia i jak – co za tym idzie – nieskończony szereg liczb przedstawia zespół przedmiotów ogólnych (rodzajowych), ostro odgraniczony przez idealną prawidłowość, tak samo ma się rzecz również z idealnymi, czysto logicznymi jednościami: pojęciami, zdaniem, prawdami, krótko mówiąc: ze znaczeniami logicznymi. Tworzą one idealnie zamknięty zespół przedmiotów ogólnych (rodzajowych), dla których jest czymś przypadkowym to, że bywają one myślane i wyrażane. Istnieją więc niezliczone znaczenia, które w zwykłym, względnym rozumieniu tego słowa są znaczeniami tylko możliwymi, jako że nigdy nie zostały wyrażone i na skutek ograniczoności władzy poznawczej człowieka nigdy nie mogą być wyrażone”.

Tak więc, według Husserla, znaczenia są realnie istniejącymi bytami idealnymi i to istniejącymi niezależnie od wyrażającego je znaku. Husserl określa więc klasę przedmiotów, do których należy znaczenie, jego cech charakterystycznych jednak nie podaje, gdyż znaczenie jest, według niego, bytem niedefiniowalnym, co jasno wynika z następującego stwierdzenia (cyt. za: A. Schaffem):

„Czym jest „znaczenie” to może być nam dane równie bezpośrednio, jak dane nam jest, czym jest barwa i ton. Dalej nie możemy już tego definiować, to jest już deskryptywny ekstrem”.

W sferze przedmiotów abstrakcyjnych umieszcza także znaczenie Roman Ingarden – według niego znaczenia należą do przedmiotów zwanych przez niego **przedmiotami intencjonalnymi**, czyli do wytworów ludzkiej psychiki. Różnią się one od przedmiotów idealnych, istniejących zawsze i niezależnie tym, że mają swój początek i swego twórcę – zostały kiedyś przez ludzi wytworzone.

Znaczenie jako relacja między znakami

Definiowanie znaczenia jako pewnej relacji zachodzącej między znakami lokuje znaczenie w sferze języka. Przy takim pojmowaniu zrozumieć znaczenie jakiegoś wyrażenia (znaku) to tyle, co umieć podać wyrażenia pozostające z nim w odpowiednich relacjach, przede wszystkim umieć odpowiedzieć na pytanie, czy dwa wyrażenia znaczą to samo, czy nie, przy czym wyraźnej definicji terminu znaczenie najczęściej się tu nie podaje. W takim znaczeniu, jako relacji między znakami, używa się słowa znaczenie na przykład w zdaniach:

Co to znaczy book? – Book znaczy tyle, co książka.

Co to znaczy desygnat? – Desygnatem nazwy nazywamy każdy przedmiot, o którym można ją zgodnie z prawdą orzec.

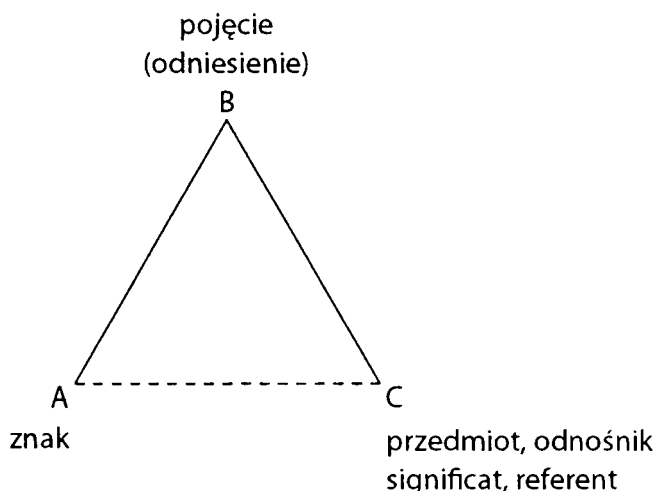
Pojmowanie znaczenia jako relacji między znakami jest cechą charakterystyczną strukturalizmu, przy czym pojawiają się tu następujące jego określenia:

- znaczeniem wyrażenia (znaku) jest **zbiór wszystkich kontekstów**, w jakim ono występuje (**dystrybucja**);
- znaczeniem wyrażenia (znaku) jest **zbiór wszystkich jego parafraz**;
- znaczeniem wyrażenia (znaku) jest **zbiór wszystkich opozycji**, w jakich dane wyrażenie występuje względem innych wyrażen języka;
- znaczeniem wyrażenia (znaku) jest uporządkowany **zbiór cech semantycznych** (np. Apresjan);
- znaczeniem zdania (wypowiedzenia) jest **zbiór konsekwencji** bezpośrednio z niego wynikających (zwróćmy uwagę na to, że w myśl tej definicji znaczenie przysługuje dopiero zdaniom, a nie ich składnikom).

Niektórymi spośród tych koncepcji znaczenia zajmiemy się bardziej szczegółowo, tu chcielibyśmy jednak zwrócić uwagę na to, że słowo znaczenie w powyższych użyciach często nie występuje w swoim właściwym sensie, lecz stanowi skrót wyrażenia eksplikacja znaczenia, a to już nie jest oczywiście to samo (takie użycie nie wynika oczywiście z niedostrzegania tej różnicy lub próby jej zatarcia, jest po prostu przeniesieniem użycia słowa znaczenie z rozważań semantyków).

Znaczenie jako relacja między znakiem a przedmiotem

Jako ilustrację relacji zachodzących między znakiem a oznaczanym przezeń elementem rzeczywistości przytacza się często tzw. **trójkąt semantyczny** Ogdena i Richardsa:



Znaczenie rozumiane jako relacja między znakiem a przedmiotem – ilustrowane w trójkącie linią przerywaną, byłoby więc w tej koncepcji równoważne z omówionym już wcześniej, przy okazji denotacji, oznaczaniem.

Znaczenie jako relacja między znakiem a myślą o przedmiocie

Interpretatorzy relacji zachodzących między znakiem a przedmiotem bardzo często traktują tę relację jako relację pośrednią, uważając, że „wyraz oznacza rzecz za pośrednictwem pojęć” (Ullmann – cyt. za J. Lyons: *Semantyka*). Czyli inaczej, według tego samego autora, między znakiem a przedmiotem odniesienia pośredniczy „odniesienie”, które nie jest niczym innym jak „informacją, jaką nazwa przekazuje słuchaczowi” (Ullmann – cyt. za U. Eco: *Pejzaż semiotyczny*). Odniesienie to w różnych teoriach bywa różnie nazywane: **pojęciem**, obrazem myślowym, ideą, przedstawieniem. Takie pojmowanie znaczenia charakterystyczne jest dla teorii asocjacionistycznych, a ilustruje je najlepiej następujący cytat:

„Przez ustawiczne powtarzanie powstaje tak ścisłe powiązanie między określonymi dźwiękami i odpowiadającymi im ideami, że nazwy, skoro tylko się je słyszy, prawie tak samo szybko wywołują pewne idee, jak gdyby faktycznie pobudzały zmysły te właśnie rzeczy, które zdolne są je wywołać” (Locke, cyt. za *Małą encyklopedią logiki*). Tym tropem idzie również Ullmann, który pisze: „Jeśli słyszysz nazwę stół, to pomyślę o stole; jeżeli pomyślę o stole, to w razie potrzeby wymówię nazwę” (cyt. za J. Lyons: *Semantyka*). Znaczenie jest więc w tej koncepcji relacją między znakiem a pojęciem (myślą), która „pozwala im wywoływać się nawzajem”.

Pogląd ten podzielał również Kazimierz Ajdukiewicz, który wprowadził określenie „typ myśli” uważając, że nie można tu przecież mówić o czyichś indywidual-

alnych, jednorazowych przedstawieniach, ale jedynie o tym, co jest wspólne całej klasie indywidualnych przedstawień wywołanych przez dźwięk lub napis określonego rodzaju. Klasę takich przedstawień nazywa on **typem myśli**, zaś klasę podobnych co do kształtu wyrażen **typem wyrażenia**. Definicja znaczenia podana przez Ajdukiewicza brzmi: „Typ myśli M stanowi znaczenie językowe wyrażen typu W w języku J wtedy i tylko, gdy: 1° wyrażenia o typie W należą do zasobu języka J, 2° dowolna osoba X używa wyrażenia A będącego typu W jako wyrażenia języka pod tym i tylko pod tym warunkiem, że a) typ myśli, polegających na doznawaniu treści zmysłowej właściwej wyrażeniom typu W, jest w umyśle osoby X skojarzony (...) z typem myśli M, b) myśl będącą doznaniem treści zmysłowej właściwej wyrażeniu A jest zarazem myślą typu M” (cyt. za *Małą encyklopedią logiki*).

Koncepcja Ajdukiewicza, w którego definicji znaczenia pojawia się stwierdzenie, że znaczenie jest typem myśli, umieszcza znaczenie w sferze zjawisk psychicznych, zaś użycie w definicji wyrażenia „osoba X używa wyrażenia A będącego typu W jako wyrażenia języka pod tym i tylko pod tym warunkiem...” wiąże je z koncepcją traktowania znaczenia jako użycia wyrażen, a więc pewnego rodzaju ludzkiego działania.

Znaczenie jako relacja między znakiem a działaniem ludzi

Pojmowanie znaczenia jako relacji między znakiem a działaniem ludzi pojawia się w koncepcjach związanych z **behawioryzmem**. Twórcy tego kierunku uważali, że obserwacji dostępne są tylko zachowania ludzi, dlatego też w badaniach nad językiem jedynym, co możemy obserwować, jest językowe zachowanie się ludzi oraz związane z tym zachowaniem zachowanie (reakcje) pozajęzykowe, świat myśli i pojęć jest bowiem badaniu naukowemu absolutnie niedostępny.

Podobne poglądy na temat znaczenia wyrażał już wcześniej Charles S. Peirce, który pisał, że znaczenie to nic innego, jak praktyczne skutki myśli w działaniu. Podobne poglądy, upatrujące znaczenie w przejawiających się w działaniu ludzkim skutkach komunikowania myśli za pomocą znaków, wyrażali i inni pragmatyści, na przykład George H. Mead, który pisał: „Istoty znaczenia nie można przeto pojmować jako stanu świadomości lub jako zbioru zorganizowanych stosunków, które istnieją czy też przetrwały w sferze myśli, poza dziedziną doświadczenia, w skład którego wchodzi; przeciwnie, znaczenie należy w istocie swej pojmować obiektywnie, jako istniejące całkowicie w granicach owego doświadczenia. Odpowiedź jednego osobnika na gest drugiego w dowolnym akcie społecznym jest znaczeniem tego gestu” (cyt. za A. Schaffem).

Najdalej w praktycznej interpretacji znaczenia posunął się **operacjonizm**, którego twórca, Bridgman, głosił, iż: „...prawdziwe znaczenie terminu da się stwierdzić przez obserwację tego, co człowiek z nim robi, a nie na podstawie tego, co o nim mówi” (cyt. za A. Schaffem). Tak też rozumiane jest przez niego znaczenie pytań, pisał bowiem: „Jeśli jakieś pytanie ma znaczenie, to musi istnieć możliwość znalezienia takich operacji, za pomocą których można na nie odpowiedzieć...” (cyt. za A. Schaffem).

Zwolennikiem behawiorystycznej koncepcji znaczenia spośród językoznawców był Bloomfield, który twierdził, że znaczenie polega na obserwowalnych bodźcowo-reakcyjnych cechach wypowiedzi.

W sferze relacji między znakiem a działaniem ludzi umieszcza znaczenie Ludwig Wittgenstein, który formułuje następującą tezę: „W odniesieniu do wielkiej klasy przypadków użycia słowa «znaczenie» – choć nie do wszystkich – można to słowo wytłumaczyć następująco: «znaczeniem słowa jest użycie go w języku»” (cyt. za A. Schaffem). Znaczenie jest dla niego relacją znaku do działania, działania językowego, które polega na używaniu znaku w procesie mówienia. Wittgensteinowska teoria użycia wyrażeń tak jest przedstawiona w *Małej encyklopedii logiki* (w haśle „znaczenie”):

„Wittgenstein przyrównuje wyrażenia do narzędzi, a znaczenie wyrażen do funkcji pełnionych przez narzędzia. Znać znaczenie zdania, to wiedzieć, jak go używać, czyli wiedzieć, w jakich okolicznościach jego użycie jest poprawne, a w jakich niepoprawne. Zdanie coś znaczy, jeżeli istnieje sposób posłużenia się nim. Uczenie się znaczeń odbywa się metodą treningu prowadzącego do powstania pożądanej dyspozycji, nie zaś metodą wyjaśnień: ten nauczył się słowa «cegła», kto reaguje na to słowo w sposób właściwy, np. podając cegłę na żądanie. Nie jest natomiast konieczna (wbrew twierdzeniom asocjacionizmu), by na słowo «cegła» reagować przedstawieniem sobie cegły. Wszystkie narzędzia służą do wywoływania zmian w przedmiotach; słowa są narzędziami służącymi do wywoływania zachowań, czyli pewnych zmian w organizmach (...) Słowo «znaczenie» nie posiada desygnatu w żadnej dziedzinie (ani psychicznej, ani obiektywno-realnej, ani obiektywno-idealnej), ponieważ, wbrew swej gramatycznej formie, nie jest nazwą: jest wyrwanym z kontekstu niesamodzielnym fragmentem takich zwrotów jak: «x ma znaczenie» lub «x oraz y mają to samo znaczenie» (tzn. wywołują takie same zachowania)”.

Na marginesie należy zauważyć, że według Wittgensteina nie wszystkie wyrażenia językowe mają znaczenie. Także wiele innych koncepcji znaczenia przypisuje tę wartość dopiero zdaniom.

Na koncepcji Wittgensteina opiera się neopozytywistyczna koncepcja znaczenia Koła Wiedeńskiego, którego przedstawiciel Moritz Schlick pisze: „Ilekroć zatem zapytujemy, «co znaczy» jakieś zdanie, tylekroć oczekujemy wskazówki do tego, w jakich okolicznościach należy tego zdania używać (...) Stwierdzenie znaczenia zdania sprowadza się do stwierdzenia reguł jego używania, a to jest to samo, co stwierdzenie, w jaki sposób można zdanie zweryfikować (lub sfalsyfikować). Znaczeniem wyrażenia jest metoda jego weryfikacji.” (cyt. za A. Schaffem).

Bliska powyższej jest także koncepcja Rudolfa Carnapa, który pisał: „Gdybyśmy, wiedzieli, czego potrzeba, aby dane zdanie uznać za prawdziwe, wiedzieliśmy wówczas, jakie jest jego znaczenie. A jeśli dla dwóch zdań warunki, pod którymi mielibyśmy wziąć je za prawdziwe, są te same, wtedy zdania te mają te same znaczenia (...) Wyrażenie języka jest znaczące (w tym rozumieniu), jeśli wiemy, jak go użyć mówiąc o faktach empirycznych.” (cyt. za *Małą encyklopedią logiki*).

Podobnie u Ajdukiewicza (jak już widzieliśmy) podstawowym pojęciem teorii znaczenia jest pojęcie **językowych dyrektyw uznawania zdań**, a więc stosowania zdań w określonych okolicznościach. Teorie takie często, jak to jest na przykład u Ajdukiewicza, ograniczają klasę wyrażzeń, którym przysługuje cecha posiadania znaczenia – u Ajdukiewicza są to tylko wyrażenia będące zdaniami w sensie logicznym (lub ich składniki), czyli zdania oznajmujące, bo tylko wobec takich zdań można zajmować postawę uznawania.

Znaczenie jako zbiór konsekwencji

Znaczenie, w tym wypadku co najmniej zdania, może być też traktowane jako zbiór konsekwencji, czyli wniosków, które można wyciągnąć na podstawie informacji zawartej w tym samym zdaniu. Koncepcja taka pochodzi od Alfreda Tarskiego. Sformułował on następującą definicję **konsekwencji**: „Zdanie A jest konsekwencją zbioru X wtedy, i tylko wtedy, gdy istnieje taki skończony ciąg napisów D, że D jest dowodem zdania A w oparciu o zbiór zdań X”. Tak więc konsekwencją (lub wnioskiem) pewnego zbioru zdań X jest każde zdanie, którego można dowieść w oparciu o zbiór X, czyli wyprowadzić z pewnej liczby zdań zbioru X, posługując się regułami wnioskowania. Irena Bellert w artykule *O pewnym warunku spójności tekstu* rozważa taki przykład:

„Jeśli ktoś powie:

(1) Najstarszy syn Anny wyjechał z Warszawy na stypendium do Sorbony. wówczas, zakładając, że mówiący użył danego wypowiedzenia właściwie, odbiorca ma pewne podstawy do następujących wniosków:

- (a) Anna ma syna,
- (b) Anna ma więcej niż dwoje dzieci,
- (c) Najstarszy syn Anny był poprzednio w Warszawie,
- (d) Najstarszy syn Anny pojechał do Francji,
- (e) Najstarszy syn Anny jest studentem, pracownikiem naukowym, artystą lub literatem,
- (f) Najstarszy syn Anny skończył szkołę średnią”.

Wniosków tych może być oczywiście o wiele więcej. Uzupełnić je można choćby o następujące:

- Najstarszy syn Anny wyjechał do Paryża,
- Najstarszy syn Anny otrzymał stypendium na Uniwersytecie,
- Najstarszy syn Anny jest dorosły,
- Najstarszy syn Anny wyjechał do stolicy Francji,
- Najstarszego syna Anny nie ma obecnie w Warszawie.

Wnioski te można wyciągnąć na podstawie:

- znajomości systemu językowego,
- na podstawie wiedzy o rzeczywistości.

Do tych pierwszych na pewno należą wnioski a, b, c, do tych drugich chyba wnioski d, e, f. Piszemy „chyba”, ponieważ nie zawsze łatwo jest przeprowadzić ostrą granicę pomiędzy wiedzą o świecie a wiedzą o języku, przykładem czego mógłby być wniosek: „Najstarszy syn Anny wyjechał z Polski”, trudno bowiem

rozstrzygnąć, czy wiedza o tym, że Warszawa jest nazwą miasta leżącego w Polsce, należy do wiedzy językowej, czy do wiedzy pozajęzykowej. A oto inne przykłady wypowiedzeń i ich konsekwencji:

- konsekwencją zdania:

Marysia zdała właśnie maturę.

są zdania:

Marysia skończyła szkołę średnią.

Marysia ma średnie wykształcenie.

- konsekwencją zdania:

Jan rozwiódł się z Anną.

są zdania:

Jan był mężem Anny.

Jan i Anna byli małżeństwem.

Anna nie jest już żoną Jana.

- konsekwencją zdania:

Anna nie lubi zwierząt.

jest zdanie:

Anna nie lubi psów i kotów.

- konsekwencją zdania:

Anna zgubiła podręcznik.

są zdania:

Anna zgubiła książkę.

Anna miała podręcznik.

Anna nie ma już podręcznika.

Konsekwencja jest więc informacją, którą można wyprowadzić na zasadzie wnioskowania z innych informacji, będących przesłankami rozumowania. W formalnym opisie funkcja konsekwencji zdaniu lub zbiorowi zdań przyporządkowuje zbiór wszystkich jego konsekwencji. Znajduje to praktyczne zastosowanie na przykład w systemach wyszukiwania informacji, przede wszystkim w systemach faktograficznych, ale także w systemach dokumentacyjnych – można, na przykład, przyjąć, z pewnym uproszczeniem, że tezaurus stanowi taki zbiór zapisów funkcji konsekwencji poszczególnych deskryptorów (byłoby to pewne rozszerzenie teorii znaczenia jako konsekwencji zdania).

Wśród tak rozumianych konsekwencji wyróżnia się ich pewien podzbiór, który nosi nazwę presupozycji. Zdanie a jest **presupozycją** zdania b, jeśli prawdziwość zdania a jest koniecznym warunkiem prawdziwości zdania b.

Na przykład:

- presupozycją zdania:

Jan wyjechał z Warszawy.

jest zdanie:

Jan był w Warszawie.

- presupozycją zdania:

Jan przestał palić papierosy.

jest zdanie:

Jan palił papierosy.

- presupozycją zdania:

Anna znalazła wreszcie swoje notatki.

są zdania:

Anna zagubiła swoje notatki.

Anna szukała swoich notatek.

Szukanie notatek trwało długo.

Presupozycja danego zdania to zbiór założeń, które muszą być spełnione, żeby danego zdania można było sensownie użyć, na przykład, nie byłoby sensownie użyte zdanie *Jan rozwiódł się z Anną*, gdyby nie czyniło się założenia, że Jan i Anna byli małżeństwem, przy czym nie jest tu ważne, czy Jan i Anna byli naprawdę małżeństwem – nie jest ważny stan faktyczny, tylko pewne założenia co do istnienia takiego stanu.

Wykładnikami presupozycji są pewne wyrażenia, np. *przestać, skończyć, obudzić się, zasnąć, rozwieść się, pomagać, wyzdrowieć, wreszcie, nagle, aż*, itp., co nakłada pewne warunki na opis znaczenia jednostek leksykalnych.

Presupozycje, podobnie jak konsekwencje, wykorzystywane są w systemach wyszukiwania informacji.

Znaczenie jako relacja między ludźmi porozumiewającymi się za pomocą znaków

Analiza znaku i znaczenia powinna być przeprowadzana nie w oderwaniu, ale tylko w konkretnych aktach komunikacyjnych, a więc w **sytuacji znakowej**. Tak uważa, na przykład, Adam Schaff, który pisze: „Problem znaczenia występuje tam, gdzie mamy do czynienia ze znakiem w procesie komunikowania się ludzi. W tym sensie znaczenie jest określonym stosunkiem między porozumiewającymi się ludźmi (...). Są to swoiste stosunki między wszystkimi elementami sytuacji znakowej: między porozumiewającymi się ludźmi, między ludźmi a rzeczywistością, między ludźmi a znakiem, między znakiem a rzeczywistością, między znakiem a innymi znakami systemu. Są to stosunki różnego typu, leżące niejako

w różnych płaszczyznach, w płaszczyźnie psychologicznej i w płaszczyźnie ludzkich postaw i działań. Zawsze jednak decydujące w stosunkach jest porozumiewanie się ludzi. Tylko w tym kontekście zrozumiała jest sytuacja znakowa, zrozumiała jest znak i znaczenie. (...) Można inaczej powiedzieć, że znaczenie należy zarówno do sfery ludzkich działań, jak i ludzkiego myślenia. W rzeczywistości bowiem obie te sfery są ze sobą nierozzerwalnie związane.” (cyt. za *Wstęp do semantyki*).

Sam autor przyznaje jednak, że „Potraktowanie znaczenia jako określonych stosunków międzyludzkich (...) nie ratuje bynajmniej przed wieloznacznością terminu. Wychodząc bowiem z tego założenia teoretycznego można przez „znaczenie” rozumieć bądź całość stosunków składających się na sytuację znakową (proces semiozy), bądź jakiś ich wycinek (stosunek znaku do przedmiotu czy do myśli o przedmiocie), bądź desygnat czy denotat znaku (tzn. przedmiot porozumiewania się bez względu na realność jego bytu lub też taki przedmiot istniejący realnie), bądź stosunek znaku do systemu znaków (języka) lub do znaków innego języka itp.”

Powyższą prezentację najważniejszych teorii znaczenia najlepiej chyba podsumowuje stwierdzenie Charlesa W. Morrisa: „Termin znaczenie (...) wystarczająco precyzyjny dla analizy potocznej nie jest wystarczająco precyzyjny dla analizy naukowej. Opisy znaczenia zazwyczaj ciskają w tarczę zjawisk znakowych garść kłosa, podczas gdy techniczna semiotyka musi nam dostarczać słów ostrych jak strzały. «Znaczenie» oznacza każdą i wszystkie fazy procesów znakowych (status bycia znakiem, interpretację znaku przez kogoś, fakt oznaczania i significatum), a często sugeruje także procesy myślowe i wartościujące; stąd też jest dla semiotyki pożądane obejść się bez tego terminu i wprowadzić terminy specjalne dla różnych czynników, które w «znaczeniu» nie mogą być wyróżnione” (cyt. za A. Schaffem).

Rodzaje znaczeń

Staraliśmy się wyżej pokazać, jak różnie rozumiany jest termin znaczenie, obecnie powinniśmy poświęcić jeszcze nieco miejsca na przedstawienie różnych terminów złożonych, w których skład wchodzi termin znaczenie.

W logice wyróżnia się często tzw. **znaczenie wyraźne** i **znaczenie intuicyjne** wyrażen w zależności od tego, czy użytkownik danego wyrażenia potrafi wyjaśnić jego znaczenie, czy też potrafi tylko go używać w sposób poprawny, tzn. w stosunku do wszystkich jego denotatów i tylko do nich, a więc zna jego denotację, nie potrafi jednak podać jego konotacji. O problemie tym tak pisał Russel: „Nie jest bowiem konieczne dla «rozumienia» słowa, by «wiedzieć, co ono znaczy» w tym sensie, że jest się w stanie powiedzieć: «to słowo znaczy to a to». Słowo ma mniej lub bardziej niejasne znaczenie; odkryć znaczenie można jednak wyłącznie na podstawie obserwacji użycia słowa; użycie jest pierwotne, a znaczenie z niego się wyłania. Relacja między słowem a jego znaczeniem ma w istocie charakter przyczynowy, i nie ma więcej powodów po temu, aby osoba używająca

poprawnie danego słowa zdawała sobie sprawę z jego znaczenia, niż po temu, aby poprawnie poruszająca się planeta zdawała sobie sprawę z praw Keplera.” (cyt. za A. Schaffem).

W językoznawstwie mówi się także o innych rodzajach znaczenia. W rozważaniach nad zmianami zachodzącymi w sferze znaczenia leksyki języka wyróżnia się **znaczenie etymologiczne**, czyli znaczenie pierwotne wyrazu oraz **znaczenie realne (aktualne)**, a więc takie, jakie ma dany wyraz obecnie. Przykładem może tu być wyraz *sklep* oznaczający pierwotnie pomieszczenie sklepienie, a obecnie pomieszczenie, w którym dokonuje się zakupów.

Drugą opozycję, obecną w synchronicznych badaniach języka, stanowi znaczenie leksykalne oraz znaczenie strukturalne. **Znaczenie leksykalne** równe jest w tym wypadku znaczeniu realnemu, natomiast **znaczenie strukturalne** jest to znaczenie wynikające z budowy (struktury) danego wyrazu – na przykład znaczenie leksykalne wyrazu *szarak* jest takie samo, jak wyrazu *zajac*, natomiast jego znaczenie strukturalne to ‘coś szarego’.

W odniesieniu do wyrażen wieloznacznych mówi się także o **znaczeniu okazjonalnym**, które, w opozycji do **znaczenia podstawowego**, a więc najczęściej realizowanego, w tekstach jest realizowane rzadko. Takie znaczenie okazjonalne najczęściej jest **znaczeniem kontekstowym** i dotyczy znaczenia wariantu językowego realizowanego w określonym kontekście, na przykład znaczenia wyrazu *czarny*, które w kontekście *rozpacz* oznacza ‘wielki’, a w kontekście *charakter* ‘zły’.

W stylistyce mówi się o jeszcze jednym znaczeniu okazjonalnym, a mianowicie o **znaczeniu przenośnym** słowa, a więc znaczeniu, które ma ono wtedy, gdy używa się go w stosunku do przedmiotu niebędącego normalnie jego denotatem – na przykład słowo *maluch*, które normalnie znaczy ‘małe dziecko’ używane było przenośnie na oznaczenie Fiata 126p – podstawą takiego przeniesienia jest oczywiście znaczenie strukturalne tego słowa ‘coś małego’, wyrażające poza tym emocjonalny stosunek mówiącego (taką funkcję pełni bowiem formant *-uch*).

We współcześnie rozwijanej **lingwistyce kognitywnej**, zajmującej się takim opisem języka naturalnego, który odzwierciedla manifestujący się w języku sposób widzenia świata przez użytkowników tego języka, czyli rekonstrukcją uwarunkowanego kulturowo tzw. **językowego obrazu świata**, mówi się o **znaczeniu prototypowym i kategoriach prototypowych**. Uważa się, że zbiór cech konotacyjnych konstytuujących znaczenie danego wyrażenia nie musi się w takim samym stopniu odnosić do wszystkich denotatów tego wyrażenia – są wśród nich takie, które są lepszymi okazami denotacji i takie, które nie mają wszystkich cech wyznaczających jej zakres. Taki okaz, który jest najlepszym przykładem danej kategorii nazywany jest **prototypem**. Prototyp może być zarówno elementem rzeczywistości, jak i schematem konceptualnym, wyznaczanym przez zestaw cech, które powinny charakteryzować dobre elementy danej kategorii (tzw. **pojęcie prototypowe**). Prototyp stanowi centrum danej kategorii prototypowej (znaczeniowej), a pozostałe elementy umiejscawiane są w różnej od niego odległości, zależnie od tego, jak spełniają warunki przynależności do kategorii. I tak, na przykład

dla kategorii PTAK prototypem będzie u nas wróbel, dalej od centrum plasować się będzie gołąb, jeszcze dalej kura, a na obrzeżach kategorii struś i pingwin. Jak pokazały badania, w Stanach Zjednoczonych prototypem tej kategorii jest drozd. Wybór prototypu uwarunkowany jest kulturowo i środowiskowo, granice między kategoriami są nieostre, a poszczególne kategorie nakładają się na siebie.

Znaczenie wyrażen języków informacyjno-wyszukiwawczych omówimy w rozdziale XVI.

Literatura

- Ajdukiewicz A.: *Język i znaczenie*, w: *Język i poznanie*, t. 1. Warszawa 1960.
Grodziński E.: *Znaczenie słowa w języku naturalnym*. Warszawa 1964.
Grzegorzczkowska R.: *Wprowadzenie do semantyki językoznawczej*. Warszawa 1990.
Guiraud P.: *Semantyka*. Warszawa 1976.
Lyons J.: *Semantyka*, t. 1-2. Warszawa 1984-1989.
Pelc J.: *Wstęp do semiotyki*. Warszawa 1982.
Ryle G.: *Teoria znaczenia*, w: Pelc J. (red.): *Logika i język*. Warszawa 1967.
Schaff A.: *Wstęp do semantyki*. Warszawa 1960.

Prezentacja znaczenia – definicje

W poprzednim rozdziale zdaliśmy sprawę z różnych sposobów rozumienia znaczenia, obecnie omówimy najważniejsze sposoby **prezentacji znaczenia** poszczególnych znaków. Służą temu **definicje**, które najogólniej określić można jako każdy komunikat – choć najczęściej mają one postać komunikatu językowego, a więc wyrażenie prezentujące znaczenie jakiegoś znaku.

To, co ma być w definicji zdefiniowane, czyli człon definicji prezentujący wyrażenie definiowane, nazywa się **definiendum**, natomiast człon określający definicji, niezawierający wyrażenia definiowanego, nazywa się **definiens** (potocznie wyraz definicja używany bywa także tylko w stosunku do definiensa, np. podać definicję jakiegoś wyrażenia, to tyle co sformułować jego definiens).

Właściwie jedyną definicją niebędącą tylko wyrażeniem językowym, lecz mającą charakter tzw. wyrażenia mieszanego (sformułowanego w różnych kodach), jest tzw. **definicja ostensywna**, zwana też **definicją deiktyczną** lub definicją przez pokazanie. Definicja ta wiąże się z denotacyjną koncepcją znaczenia, a więc traktującą znaczenie jako zbiór elementów rzeczywistości pozajęzykowej, oznaczanych przez dany znak. Definicja taka stanowi więc regułę semantyczną ustalającą przyporządkowanie między elementami jakiejś rzeczywistości współwystępującej w procesie definiowania a rzeczywistością odwzorowujących ją znaków (jest to oczywiście współwystępowanie w przestrzeni, a najczęściej także w czasie, obu tych rzeczywistości). A oto przykłady komunikatów będących definicją ostensywną:

- definiujący pokazuje gruszkę i mówi: *To jest gruszka.*
- definiujący mówi: *To, co było wczoraj na deser, to był ananas.*
- ilustracja w książce:



Oczywiście, używając tego typu definicji, czyni się pewne założenia:

- niemożliwe jest przedstawienie całej denotacji danego znaku, ponieważ obejmuje ona wszystkie elementy tego wycinka rzeczywistości, który dany znak reprezentuje, a więc te, które istniały, istnieją i będą istniały – pokazuje się więc

tylko jeden lub kilka elementów klasy, które traktuje się jako wzorcowe egzemplarze reprezentujące zakres danego wyrażenia;

- w wypadku zastąpienia elementu rzeczywistości pozatekstowej znakiem ikonicznym (na przykład rysunkiem) zakłada się dodatkowo rozumienie przez odbiorcę zasady funkcjonowania takich znaków;

- każdy przedmiot należy do wielu różnych klas, oznaczanych przez różne wyrażenia języka – na przykład gruszka, jabłko, ananas mogą być denotatami wyrażen: *owoc*, *jadalne*, *deser*, dlatego też nie zawsze wiadomo, o co w danym akcie komunikacyjnym definiującemu chodzi. Powoduje to dodatkowe trudności w jednoznacznej interpretacji reguły semantycznej, a co za tym idzie konieczność naprowadzenia osoby informowanej, o jaką klasę definiującemu chodzi. Najprościej można to zrobić przez wskazanie klasy bezpośrednio szerszej, do jakiej dany element należy, oczywiście pod warunkiem, że osoba informowana zna już nazwę tej klasy. W takim wypadku podane wyżej definicje miałyby postać:

- *Ten owoc to gruszka.*
- *Ten owoc, który był wczoraj na deser, to ananas.*
- *owoce:*



jabłko

W przeciwnym wypadku zamierzony cel osiąga się przez pokazanie wielu wzorców przykładowych oraz wzorców negatywnych, ze słowami: „*To jest...*”, „*To nie jest...*”.

Oczywiście nie wszystkie wyrażenia języka jednakowo dobrze nadają się do definicji ostensywnych – stosunkowo łatwo definiować w ten sposób nazwy (pełne), trudno inne wyrażenia, jak choćby nazwy relacji, takie jak: *matka*, *ojciec*, *siostra*, *brat*, *córka*, *syn*, *przyjaźń*, *nienawiść*, a jeszcze trudniej nazwy puste, które nie mają z założenia desygnatów, a nie zawsze jest możliwe posłużenie się znakiem ikonicznym.

Inne rodzaje definicji wiążą się z innymi niż denotacyjna interpretacja znaczenia. Niżej omówimy najważniejsze ich typy, zaczynając od podziału definicji ze względu na funkcję, jaką pełnią względem wyrażen danego języka.

Najczęściej spotykanym typem definicji jest **definicja sprawozdawcza** zwana też **definicją analityczną**. Ma ona za zadanie przedstawiać (zdawać sprawę z) znaczenie danego wyrażenia przyjęte już w danym języku. A oto przykłady takich definicji:

Premier w języku polskim to prezes Rady Ministrów, szef rządu.

Nazwa pusta jest to nazwa niemająca desygnatów.

H₂O jest to symbol oznaczający wodę.

615.831.13 jest to symbol UKD oznaczający psychoanalizę jako metodę leczenia.

Definicje sprawozdawcze występują oczywiście we wszelkiego rodzaju słownikach jednojęzycznych.

Drugim ważnym typem definicji jest **definicja projektująca**, zwana też **definicją syntetyczną**. Funkcją jej jest wprowadzenie nowych wyrażeń do języka i określanie ich znaczeń lub też, czasami, tylko wprowadzenie zupełnie nowych znaczeń dla wyrażeń istniejących. A oto przykłady takich definicji:

Polon znaczy tyle co pierwiastek o ciężarze atomowym 210.

– definicja zaprojektowana przez Marię Curie-Skłodowską (obecnie jest to już definicja sprawozdawcza).

Orbitować znaczy tyle co poruszać się na zewnątrz pojazdu kosmicznego w czasie jego okrążeń dokola Ziemi w sferze zmniejszonego jej przyciągania.

– ta definicja projektująca powstała w czasie pierwszego wyjścia człowieka z poruszającego się pojazdu kosmicznego, kiedy to ogłoszono konkurs na wyraz nazywający taką właśnie sytuację – (definicja nie stała się definicją sprawozdawczą, ponieważ wyraz ten nie przyjął się w języku).

Palec znaczy tyle co zakończenie ruchomej dźwigni maszyny do pisania, na którym wytopiona jest czcionka jakiejś litery.

– definicja ta wprowadzała nowe znaczenie istniejącego już w języku polskim wyrażenia *palec*.

Definicje projektujące spotyka się często w literaturze naukowej oraz w normach terminologicznych.

Trzecim typem definicji jest **definicja regulująca**. Definicja ta najczęściej ma za zadanie określenie, ze względu na jakiś przyjęty cel, nicostrego zakresu znaczeniowego wyrażeń już w języku istniejących. Dzieje się tak często wtedy, gdy wyraz języka potocznego zaczyna funkcjonować jako termin naukowy. A oto przykłady takich definicji:

Noworodek jest to dziecko do ukończenia 1 miesiąca życia.

Niemowlę jest to dziecko w wieku od 1 miesiąca do 1 roku życia.

Wiek starczy jest to wiek powyżej 80 lat.

Wzrost wysoki jest to wzrost powyżej 160 cm dla kobiet i 170 cm dla mężczyzn.

Definicje bywają dzielone według różnych innych kryteriów, uzyskanymi w ten sposób typami nie będziemy się dokładniej zajmować. Omówimy tylko niektóre spośród nich, zasługujące na szczególną uwagę. Taką jest **definicja realna**, czyli **definicja rzeczowa**, która zgodnie z podanym wyżej określeniem definicji definicją właściwie nie jest, zadaniem jej nie jest bowiem podanie znaczenia jakiegoś wyrażenia, lecz podanie cech przedmiotu oznaczanego danym wyrażeniem (definicja taka ma postać definicji nominalnej w stylizacji przedmiotowej), np.:

Pierwiastek jest to substancja prosta, której drogą chemiczną nie można rozłożyć na inne ciała.

Wodniczka jest to przestrzeń w protoplazmie komórek wypełnionych sokiem komórkowym.

Definicji realnej przeciwstawia się zazwyczaj **definicję nominalną**, która ma za zadanie prezentację znaczenia wyrażen.

Ze względu na sposób formułowania definicji dzieli się je na definicje w stylizacji metajęzykowej oraz na definicje w stylizacji przedmiotowej. Definicja ma **stylizację metajęzykową**, jeżeli mówi się w niej o wyrażeniu definiowanym, w której więc definiendum występuje w supozycji materialnej, czego wykładnikiem bywa najczęściej cudzysłów (może być także inny rodzaj druku, najczęściej kursywa). Definicja ma **stylizację przedmiotową**, jeżeli pozornie mówi się w niej o przedmiotach, choć w rzeczywistości ma ona na celu prezentację znaczenia.

Ze względu na język, w którym formuluje się definicje, dzieli się je na definicje metajęzykowe i wewnątrzjęzykowe. **Definicja metajęzykowa** to definicja, w której człon określany ma postać wyrażenia cudzysłowowego, **definicja wewnątrzjęzykowa** to taka, w której wyrażenie definiowane nie jest cytowane przy pomocy cudzysłowu. W pierwszym wypadku wyrażenie będące definiendum może należeć do innego języka, na przykład języka sztucznego – język naturalny będzie w tym wypadku pełnił funkcję metajęzyka opisującego znaczenie. Definicja jest **definicją metajęzykową** również wtedy, gdy definiens formułowany jest w specjalnie skonstruowanym metajęzyku.

Z użyciem cudzysłowów wiąże się jeszcze jeden podział definicji – na definicje semantyczne oraz na definicje słownikowe. **Definicjami semantycznymi** nazywa się definicje, w których człon określany definicji występuje w cudzysłowie (a więc w supozycji, materialnej), człon zaś określający w supozycji zwykłej, np.:

„Jan jest w gorącej wodzie kąpany” znaczy, że Jan reaguje bez zastanowienia.
„Amalgamat” oznacza stop innego metalu z rtęcią.

Definicjami słownikowymi nazywa się natomiast definicje, w których oba człony są wyrażeniami cudzysłowowymi, np.:

„Amalgamat” znaczy tyle co „stop innego metalu z rtęcią”.

Ze względu na budowę najważniejsza jest **definicja klasyczna**. Ma ona postać: A jest B i C, gdzie A, B, C, są nazwami, „jest” zaś jest funktorem identyczności. Pierwsza nazwa stanowiąca część definiensa (B) jest nazwą klasy, do której należy (jest jej częścią właściwą) denotat wyrażenia określanego (definiendum), natomiast druga (C) określa cechę, która wyróżnia elementy klasy A spośród elementów nadrzędnej klasy B. Termin denotujący klasę nadrzędną nazywa się **terminem rodzajowym** (rodzaj, łac. *genus*), termin określający różnicę nazywa się **różnicą gatunkową** (łac. *differentia specifica*), ponieważ wyróżnia przedmioty tworzące pewien **gatunek** (łac. *species*) od innych przedmiotów należących do tego samego rodzaju (stąd często mówi się, że definiens to rodzaj plus różnica gatunkowa), np.:

Planeta jest to ciało niebieskie, które krąży dookoła Słońca.
gatunek rodzaj różnica gatunkowa

Definiując wyrażenia jakiegoś języka najczęściej formułuje się tzw. **definicję normalną** czyli **definicję równościową**. Jej formalne określenie ma postać: D jest definicją normalną wyrażenia W (na gruncie języka J) wtedy i tylko, gdy D ma postać równości lub równoważności, która pozwala przełożyć każdy zwrot językowy zawierający wyrażenie W na zwrot niezawierający tego wyrażenia (innymi słowy: pozwala wyeliminować W z dowolnego kontekstu). Definicje normalne dzielą się ze względu na budowę na wyraźne i niewyraźne (kontekstowe). W **definicji wyraźnej** definiendum tworzy jedynie wyrażenie będące przedmiotem definiowania, np.:

Centymetr jest to jedna setna część metra,

natomiast w **definicji niewyraźnej (kontekstowej)** wyrażenie, które ma być zdefiniowane nie stanowi całego definiendum, lecz tylko jego część umieszczoną w typowym dla niego kontekście, np. definicja logarytmu

$$x = \log_b a \Leftrightarrow b^x = a$$

Definicja, aby mogła być uznana za poprawną, musi spełniać pewne formalne **warunki poprawności definicyjnej**. Przede wszystkim wymaga się od niej, aby była **adekwatna**, tzn., aby zakres członu określającego był równy zakresowi członu określanego. Wymóg ten nie jest spełniony w tzw. **definicjach częściowych**, które nie są równoważnościami, lecz jedynie okresami warunkowymi i podają tylko niektóre warunki, jakie musi spełniać denotat wyrażenia definiowanego, aby można w stosunku do niego tego wyrażenia użyć: sam warunek wystarczający lub sam warunek konieczny (lub niektóre tylko warunki konieczne i tylko niektóre warunki wystarczające).

Definicje nieadekwatne mogą być albo **za wąskie**, np.

Polakiem jest człowiek, który ma obywatelstwo polskie.

albo **za szerokie**, np.

Proton jest to cząstka elementarna prosta.

Poprawna definicja powinna być **ostra**, tzn. formułować warunki jednoznacznie pozwalające rozstrzygnąć, czy dany element należy do danej klasy, czy też nie należy, np.

Pełnoletni jest człowiek, który ukończył 18 lat.

Przykładem definicji nieostrej może być następująca definicja:

Pediatra jest to lekarz leczący dzieci.

Przyczyną nieostrości jest tu oczywiście użycie w definiendum nieostrego wyrażenia *dzieci*, a wyrażen nieostrych jest w języku naturalnym bardzo dużo, czego przykładem mogą być takie ciągi wyrażen, jak: *pagórek, wzgórze, góra; noworodek, niemowlę, dziecko; mały, średni, duży, ogromny; młody, stary; gorąco, ciepło, chłodno, zimno*, czy też takie wyrazy, jak *młodzież* czy choćby *sprawiedliwy*,

który to wyraz, na przykład w wyrażeniu *sprawiedliwy podział*, miewa następujące różne interpretacje:

sprawiedliwy podział:

- ‘każdemu po równo’,
- ‘każdemu według potrzeb’,
- ‘każdemu według zasług’.

Oczywiście ostra definicja nicostrego wyrażenia byłaby definicją nieadekwatną.

Dobra definicja powinna być definicją **prostą**. Chodzi o to, aby wyrażenia prostego znaczeniowo nie definiować przez wyrażenie bardziej skomplikowane. Przykładem definicji niespełniającej warunku prostoty może być następująca definicja z *Malego słownika języka polskiego*:

koń – *equus caballus*, zwierzę ssące jednokopytne, roślinożerne, z rodziny o tej samej nazwie.

Warunek prostoty jest często bardzo trudny do spełnienia, a problem ten stanowi ważny problem metodologiczny semantyki – najlepszym wyjściem, choć oczywiście nie w każdym wypadku, wydają się być definicje ostensywne (sławna jest definicja autora pierwszej polskiej encyklopedii, księdza Chmielowskiego, „Koń, jaki jest, każdy widzi”).

Do najczęściej popełnianych **błędów** w formułowaniu definicji należą też:

- błąd **nieznane przez nieznane** (łac. *ignotum per ignotum*), polegający na tym, że w definiensie używa się wyrażen niezrozumiałych, nieznanych (nie zdefiniowanych w danym systemie),
- błąd **to samo przez to samo** (łac. *idem per idem*), zwany też **błędnym kołem** w definiowaniu (łac. *circulus in definiendo*).

Wyróżnia się tu dwa wypadki:

- błędne koło **bezpośrednie** – polega na użyciu wyrażenia definiowanego w członie definiującym, np.

dane – reprezentacja faktów lub pojęć w sformalizowanej postaci, umożliwiającej przekazywanie danych i wykonywanie na nich operacji. W znaczeniu informatycznym ma się na myśli automatyczne przekazywanie i automatyczne wykonywanie operacji na danych.

- błędne koło **pośrednie** – polega na tym, że termin A definiuje się przez zespół terminów zawierających termin B, a wśród terminów użytych do zdefiniowania terminu B występuje termin A, np.:

dane – reprezentacja faktów lub pojęć w sformalizowanej postaci umożliwiającej przekazywanie danych i wykonywanie na nich operacji.

operacja – czynność wykonywana na danych.

archiwum – (1) instytucja powołana do gromadzenia, opracowywania, przechowywania i udostępniania materiałów archiwalnych. (2) komórka organi-

zacyjna instytucji powołana do przejmowania, przechowywania, porządkowania i zabezpieczania materiałów archiwalnych tej instytucji materiały archiwalne, archiwalia – wszelkie dokumenty będące wytworem działalności instytucji (urzędu itp.) lub osób fizycznych, kwalifikujące się do trwałego lub czasowego przechowywania w archiwum

(definicje ze *Słownika terminologicznego informacji naukowej*, innymi błędami popełnionymi w tych definicjach nie będziemy się tu zajmować).

Ze *Słownika wyrazów obcych i zwrotów obcojęzycznych* Władysława Kopalińskiego pochodzi następujący przykład błędnego koła pośredniego:

kapituła – kolegium kanoników przy katedrze albo kolegiacie.

kanonik – duchowny świecki, członek kapituły przy katedrze a. kolegiacie.

kolegiata – kościół, przy którym istniała (a. istnieje) kapituła kanoników.

Błędne koło pośrednie jest najczęstszym błędem spotykanym w słownikach jednojęzycznych, a właściwie normalną praktyką leksykograficzną. Powodem jest to, że w słownikach tych język będący przedmiotem definiowania z założenia pełni funkcję metajęzyka względem samego siebie, przy czym nie wyróżnia się w nim zbioru wyrażen przyjętych w systemie definicyjnym jako terminy pierwotne (przyjmowane bez definicji). Przyczyną nie jest tu więc zła praktyka leksykograficzna, polegająca na niestaranności, co zawsze może się zdarzyć, ale przyjęta metodologia definiowania. Przykładów ilustrujących ten błąd nie podajemy – czytelnik może znaleźć je sam w każdym słowniku.

Przedstawione tu problemy związane z definiowaniem znaczenia od dawna były przedmiotem refleksji logików, definiowanie to jednak przede wszystkim problem dla językoznawców: teoretyczny – dla semantyków i leksykologów, praktyczny – dla leksykografów tworzących definicje słownikowe. Coraz powszechniejszy staje się pogląd, że **definicje słownikowe** powinny być czymś zasadniczo innym niż definicje encyklopedyczne, które do niedawna przeważały również w słownikach. Zasadniczym ich mankamentem było przede wszystkim to, że prezentowały wiedzę o rzeczywistości, często wiedzę naukową, podczas gdy powinny być odzwierciedleniem wyłącznie wiedzy o języku stanowiącej kompetencję leksykalną i semantyczną przeciętnego użytkownika języka, a więc utrwalony w języku potocznym „**naiwny**” obraz świata, zależny od kultury, przekonań i systemów wartościowania. Taki potoczny obraz rzeczywistości bywa uproszczony, zawiera bowiem tylko cechy relewantne dla użytkownika języka, które często różnią się od tych, które konstytuują definicje o charakterze terminologicznym. Na problem ten zwracaliśmy już uwagę, omawiając znaczenie rozumiane jako konotacja. Rekonstruowaniem takiego potocznego obrazu świata zawartego w semantycznej strukturze wyrażen językowych zajmują się przede wszystkim psycholingwiści, wykorzystując do tego informacje pozyskane od użytkowników języka w trakcie eksperymentów lub badań ankietowych.

Wyniki takiego eksperymentu przedstawia Nawoja Mikołajczak-Matyja w pracy *Definiowanie pojęć przez przeciętnych użytkowników języka i przez leksykografów*. Przytoczymy kilka definicji, „definicji skonstruowanych z elementów po-

dawanych przez przeciętnych użytkowników języka” (150 osób, uczniowie klas III i IV liceum ogólnokształcącego, studenci I i II roku), z których jasno wynika, że oddzielenie wiedzy o języku od wiedzy o rzeczywistości pozajęzykowej jest trudne.

„Czytanie i pisanie

- Czytać:** Rozszyfrowywać przy pomocy wzroku litery, tekst, słowa, zawierające informacje i zapisane np. w książce.
- Pisać:** Przelewać myśli na papier [na kartkę];; czynność mająca na celu wyrażanie myśli na papierze [na kartce]; kreślić znaki graficzne, litery na papierze [na kartce], przy pomocy np. długopisu.
- Rysować:** Kreślić obrazy [coś] za pomocą kredki [ołówka] na papierze [na kartce].

Zwierzęta

- Kot:** Zwierzę domowe, z miękkim, puszystym futerkiem, miauczące i mruczące.
- Pies:** Zwierzę domowe, które szczeka; najlepszy przyjaciel człowieka.
- Mysz:** Małe, obrzydliwe zwierzę [gryzoń, ssak] należące do gryzoni, ssaków, szare, posiadające ogon, żyjące na polu [w domu].
- Szczur:** Ohydne, obrzydliwe zwierzę [gryzoń] należące do gryzoni, myszowatych, większe niż mysz, posiadające długi ogon, żyjące w kanałach, miejscach zabudowanych.

Napoje

- Kawa:** Czarny, gorący napój z ziaren rośliny kawy, zawierający kofeinę, służący do picia; roślina, z której zmielonych i zalanych wrzątkiem ziaren robi się napój.
- Lemoniada:** słodki, zimny, orzeźwiający napój, zawierający CO₂, (= gazowany) i owoce.

Żywioty

- Woda:** Bezbarwna, bezwonna ciecz [płyn], niezbędna do życia, służąca do picia, mycia, produkcji; substancja [związek chemiczny] o płynnym stanie skupienia i wzorze H₂O, występująca na ziemi; H₂O.
- Ogień:** Żywiół wydzielający ciepło [ogrzewający]; płomień powstający w wyniku palenia się czegoś; PŁOMIEN.
- Powietrze:** Niezbędna mieszanina gazów [tlenu, azotu]; to, co umożliwia oddychanie; gazy, które nas otaczają; ATMOSFERA.”

Prezentacja znaczenia wyrażen języków informacyjno-wyszukiwawczych

W definiowaniu znaczenia wyrażen języków informacyjno-wyszukiwawczych funkcję metajęzyka pełni język naturalny, jego udział w prezentacji znaczenia jest jednak różny i zależy przede wszystkim od typu języka, ale także od stopnia „świadomości lingwistycznej” twórców języka oraz uwarunkowań natury technicznej.

Elementarne jednostki leksykalne języków informacyjno-wyszukiwawczych prezentowane są w słownikach tych języków, na przykład słowa kluczowe w słownikach słów kluczowych, symbole klasyfikacyjne w tablicach klasyfikacji, deskryptory w tezaurusach. Z typem języka wiąże się nie tylko tradycyjnie przyjęta nazwa słownika, ale także jego struktura, funkcja oraz zakres i sposób prezentowanych informacji o języku. Problem ten omówimy dokładniej w rozdziale XV.

Literatura

Grzegorzczkova R.: *Wprowadzenie do semantyki językoznawczej*. Warszawa 1990.

Kraszewski Z.: *Logika – nauka rozumowania*. Warszawa 1975.

Marciszewski W. (red.): *Mala encyklopedia logiki*. Warszawa 1970.

Mikołajczak-Matyja N.: *Definiowanie pojęć przez przeciętnych użytkowników języka i przez leksykografów*. Poznań 1998.

XIV

Prezentacja znaczenia – kody semantyczne

Prezentacja znaczenia jest chyba najważniejszym, choć zarazem najtrudniejszym, zadaniem w opisie każdego języka, bez względu na to, czy jest to język naturalny, czy sztuczny. Największym problemem jest tu wybór odpowiedniego metajęzyka opisu znaczenia, drugim, równie trudnym i ważnym, wypracowanie odpowiedniej koncepcji znaczenia.

Szczególnie płodna we współczesnym językoznawstwie okazała się koncepcja tzw. **semantyki składnikowej**, traktującej znaczenie wyrażen języka naturalnego jako strukturę znaczeń elementarnych (tzw. **znaczenie intensjonalne**). Koncepcje i metody wypracowane przez zwolenników tej teorii znalazły zastosowanie w semantyce, teorii języków informacyjno-wyszukiwawczych, psychologii i sztucznej inteligencji.

Pierwszym współczesnym językoznawcą, który przedstawił taką koncepcję opisu znaczenia wyrażen języka naturalnego, był **Luis Hjelmslev**, twórca kopenhaskiej szkoły strukturalistycznej. Twierdził on, że język zorganizowany jest tak, że ze stosunkowo niedużej liczby tzw. **figur (elementarnych znaczeń)** może być zbudowana duża liczba znaczeń złożonych, którym dopiero w planie wyrażenia konkretne języki naturalne przyporządkowują odpowiednie elementy (morfemy lub wyrazy). Figury takie odnajduje się poprzez porównywanie znaczeń wyrażen i wyodrębnianie **minimalnych opozycji semantycznych** metodą analogiczną do metody wyodrębniania minimalnych opozycji fonologicznych. A oto przykłady takiej analizy:

opozycja wyrażen

chłopiec – dziewczynka

kobieta – dziewczynka

jałówka – dziewczynka

opozycja figur

(1) męski – żeński

(2) dorosły – młody (dziecko)

(3) zwierzęcy – ludzki

Wyodrębnione figury potwierdzają się w opozycji znaczeń na przykład następujących wyrazów języka polskiego:

(1) *mężczyzna – kobieta, chłopiec – dziewczynka, ojciec – matka, syn – córka, brat – siostra, byk – krowa, ogier – klacz, kogut – kura,*

(2) *mężczyzna – chłopiec, kobieta – dziewczynka, dorosły – dziecko, krowa – jałówka, koń – źrebię, pies – szczenię, owca – jagnię,*

(3) *byk* – męczyzna, *szczenię* – dziecko, *jalówka* – dziewczynka, *klacz* – kobieta.

Tak wyodrębnione cechy znaczeniowe tworzą znaczenia złożone, którym w planie wyrażenia języka polskiego przyporządkowano wspólny wykładnik leksykalny, np.:

żeński + młody + ludzki = *dziewczynka*
żeński + dorosły + ludzki = *kobieta*
męski + młody + ludzki = *chłopiec*
męski + dorosły + ludzki = *męczyzna*
męski + dorosły + zwierzęcy (krowi) = *byk*
żeński + młody + zwierzęcy (krowi) = *jalówka*.

Konsekwentne zrealizowanie poszukiwań elementarnych opozycji znaczeniowych doprowadziłoby, zdaniem Hjelmsleva i jego następców, do wyodrębnienia **elementarnych znaczeń** (zwanych również w literaturze **semami**) i odkrycia systemu semantycznego języka. Pozwoliłoby to również na rozwiązanie problemu **metajęzyka** opisu znaczenia wyrażen języka naturalnego, w każdym języku naturalnym, jak głoszą niektórzy językoznawcy (np. Anna Wierzbicka), elementarne znaczenia znajdują bowiem wyraz w jego jednostkach leksykalnych. Problemem jest tylko ustalenie inwentarza tych elementarnych wyrażen oraz, dla potrzeb konkretnych zastosowań (np. zastosowań w systemach wyszukiwania informacji), ustalenie ekonomicznego poziomu opisu, należy bowiem i w tym wypadku pamiętać o zależności między liczbą jednostek poziomu elementarnego a stopniem skomplikowania opisu jednostek poziomu wyższego.

Dla potrzeb opisu wyrażen języka naturalnego **Anna Wierzbicka** pierwotnie postulowała zbiór 13 wyrażen elementarnych, tzw. **prymitywów semantycznych**, co miało wystarczyć do:

- przedstawienia każdego zdania każdego języka naturalnego w postaci intuicyjnie równoznacznego zdania w tym języku sformułowanego tylko z użyciem prymitywów,
- wyjaśnienia intuicyjnie wyczuwalnych relacji znaczeniowych między wyrażeniami: bliskoznaczności, przeciwstawności, zawierania się jednego znaczenia w drugim,
- wyjaśnienia bezsensowności czy dziwaczności wypowiedzi poprawnych składniowo, ale dewiacyjnych pod względem semantycznym.

A oto proponowana przez Wierzbicką **lista wyrażen elementarnych**:

1. ja,
2. ty,
3. świat,
4. ktoś,
3. coś,
6. ten,
7. chcę,

8. nie chcę,
9. myślę o,
10. mówię,
11. wyobrażam sobie,
12. część,
13. stawać się.

W dalszych badaniach zbiór ten został rozszerzony do około 60 pojęć uniwersalnych, a więc takich, które mają swoje proste wykładniki w każdym języku naturalnym. A oto ta lista:

TABLICA POJĘĆ UNIWERSALNYCH

1. JA, TY, KTOŚ, COŚ, LUDZIE, CIAŁO
2. TEN, TEN SAM, INNY
3. JEDEN, DWA, NIEKTÓRE, WSZYSTKO, DUŻO
4. DOBRY, ZŁY, MAŁY, DUŻY
5. MYŚLĘ, WIEM, CHCĘ, CZUJĘ, WIDZĘ. SŁYSZĘ
6. MÓWIĆ, SŁOWO, PRAWDA
7. ROBIĆ, DZIAĆ SIĘ, RUSZAĆ SIĘ
8. BYĆ (ISTNIEĆ), MIEĆ
9. ŻYĆ, UMRZEĆ
10. NIE, BYĆ MOŻE, MÓC, BO, JEŻELI
11. KIEDY, TERAZ, PO (POTEM), PRZED (PRZEDTEM), DŁUGO, KRÓTKO, PRZEZ PEWIEN CZAS
12. GDZIE, TUTAJ, POD, NAD, DALEKO, BLISKO, Z [TEJ] STRONY, WEWNĄTRZ, DOTYKAĆ (KONTAKT)
13. BARDZO, WIĘCEJ
14. RODZAJ, CZĘŚĆ
15. (TAK) JAK

Zbiór ten jest jednocześnie słownikiem tzw. **języka mentalnego** (*lingua mentalis*), który, według niektórych koncepcji, np. koncepcji **Noama Chomsky'ego**, jest językiem tzw. **struktur wrodzonych**, wspólnych umysłom wszystkich ludzi, które w różnych językach znajdują wyraz w różnych właściwych im strukturach powierzchniowych. Wierzbicka nie doprowadza swoich eksplikacji znaczenia aż do poziomu jednostek elementarnych, wiemy jednak, że jest to możliwe. A oto przykłady prób jej eksplikacji znaczenia, pochodzących zresztą z różnych etapów jej koncepcji semantycznych, które podamy bez komentarza, po prostu jako ilustrację zaprezentowanej tu metody opisu znaczenia:

dobry = taki, jak chcę, żeby był
kocha = pragnie powodować dobro
nienawiść = pragnienie powodowania zła
dobro = to, czego byśmy wszyscy chcieli
zło = to, czego byśmy wszyscy nie chcieli
wiedzieć = móc powiedzieć prawdę

umieć = móc zrobić dobrze
pragnę = czuję, że chcę
chce mi się = czuję, że chcę, nie dlatego, że chcę chcieć
mam wolę = chcę, dlatego, że chcę chcieć
ludzie = istoty takie jak my
on się boi = on się czuje tak, jak się zwykle czujemy, kiedy sądzimy, że może się nam stać coś złego i chcemy, żeby się to nie stało
trwoga = to, co czujemy wtedy, kiedy czujemy, że stanie się (nam) coś złego i pragniemy, żeby się nie stało
lękam się = czuję się tak, jak się czujemy wtedy, kiedy czujemy, że może nam się stać coś złego i pragniemy, żeby się nie stało
obawiam się = myślę, że może stać się coś, co pragnę, aby się nie stało
smutek = uczucie, którego doznajemy wtedy, kiedy stwierdzamy, że nie dzieje się to, co pragnęlibyśmy, aby się działo
radość = uczucie, którego doznajemy wtedy, kiedy stwierdzamy, że dzieje się to, co pragniemy, aby się działo
cholera! = czuję złość = czuję się tak, jak się czujemy zwykle, gdy dzieje się coś wbrew naszej woli
aż = sędzę, że wiesz, że to dużo
tylko = nie sędź, że więcej
zaledwie = sędzę, że rozumiesz, że to mało
jeszcze = sędzę, że wiesz, że można było sędzić, że później będzie (np. w zdaniu: *On jeszcze nie przyszedł*)
dość = nie sędź, że trzeba powiedzieć, że jest bardzo (więcej) (np. w zdaniu: *On jest dość chudy*)
po A było B, po B – C, po C – D ... = świat, którego częścią było A, stawał się światem, którego częścią było B, C, D...
działo się A, potem B, potem C ... = świat stawał się światami, których częścią były A, B, C ...
Budda żył przed Sokratesem (wcześniej) = żyjący Budda był częścią pewnego świata-kiedyś, który stawał się częścią tego świata-kiedyś, którego częścią był żyjący Sokrates
Chrystus żył po Sokratesie (później) = żyjący Chrystus był częścią pewnego świata-kiedyś, którym stawał się świat-kiedyś, którego częścią był żyjący Sokrates
Platon żył długo = można by wyróżnić wiele światów-kiedyś, które stawały się jedne drugimi i których częścią był żyjący Platon
Platon żył o 11 lat dłużej niż Sokrates = można by wyróżnić o 11 światów-pewnego-roku, które stawały się jedne drugimi i których częścią był żyjący Platon, więcej niż światów-pewnego-roku, które stawały się jedne drugimi i których częścią był żyjący Sokrates
X jest teraz w miejscu M = X, nie będąc częścią M, jest teraz jakby-częścią M
X się rusza = X zmienia miejsce = X staje się jakby częścią różnych części czegoś, czego częścią nie jest

A posuwa się w kierunku K = A staje się jakby-częścią części świata bliższych (niż poprzednie) K

Radość jest przed Miedzeszynem = cokolwiek się posuwa stąd do Miedzeszyna, staje się jakby-częścią Radości przedtem, jak (zanim) się staje jakby-częścią Miedzeszyna

A jest mniejsze niż B = A mogłoby być częścią B

A jest równe B = A nie mogłoby być częścią B ani B nie mogłoby być częścią A

od A jest bliżej do B niż do C = wszystkie te części świata, których jakby częścią staje się cokolwiek, co jest najpierw jakby-częścią A, a potem jakby-częścią B, mogłyby być częścią tych wszystkich części świata, których jakby-częścią staje się cokolwiek, co jest najpierw jakby-częścią A, a potem jakby częścią C

twarz = przednia, nie pokryta włosami część głowy

nos = wystająca środkowa część twarzy

uszy = wystające części głowy po obu stronach twarzy

broda = wystająca dolna część twarzy

czoło = płaska górna część twarzy

policzek = miękkie boczne części twarzy

oczy = części twarzy w górnej części twarzy, które zamykają się i otwierają

powieki = części twarzy, które opuszczając się, zakrywają oczy

rzęsy = włosy rosnące na brzegach powiek

*usta*¹ = część dolnej części twarzy, która się otwiera

*usta*² = otwór w dolnej części głowy za ustami¹

głowa = część ciała, która jest nad wszystkimi innymi

szyja = część ciała, której ruchy powodują ruchy głowy

*ręce*¹ = podłużne części ciała po bokach ciała

*ręce*² = płaskie części ciała na końcach rąk¹

nogi = podłużne dolne części ciała

tulów = ciało bez głowy, szyi, rąk i nóg

serce = część ciała wewnątrz górnej części tułowia z lewej strony, której ruch można słyszeć

żołądek = część ciała w górnej części ciała, w której części czegoś, co zostało wprowadzone przez usta do wnętrza ciała, stają się częściami ciała

żyły = podłużne części ciała pod skórą, wewnątrz których płynie krew

mózg = miękka część ciała w czaszce, bez której nie można myśleć

kości = twarde rzeczy wewnątrz ciała, które są częściami ciała

jeść = wkładać do ust i ruchem ust wprowadzać do wnętrza ciała coś, czego części mogą stawać się ciałem

trawi(ć) = części czegoś, co zostało wprowadzone z zewnątrz do wnętrza ciała, stają się ciałem

oddychać = ruchem części twarzy wprowadzać przez otwór w twarzy powietrze do wnętrza ciała i wydalać powietrze na zewnątrz ciała

łzy = krople słonej wody wydzielane przez oczy pod wpływem uczuć

krew = płyn, który płynie we wszystkich częściach żywego ciała

włosy = długie cienkie giętkie rzeczy, których początkowe części są w skórze i są częściami ciała

leb (łeppek, głowa zwierzęcia) = ta część ciała zwierzęcia, która jest w ciele zwierzęcia tym, czym w ciele człowieka jest głowa

lapy (nogi, odnóża zwierząt) = części ciała, które są w ciele zwierzęcia tym, czym są w ciele człowieka nogi

I jeszcze przykład eksplikacji znaczenia tzw. **wyrażeń meta-tekstowych**:

mianowicie = powiem, jakie:

następujące = powiem, jakie:

ozn., tj., czyli = powiem to samo inaczej

np. = powiem jedną taką rzecz po to, żebyś rozumiał, o jakich rzeczach myślę

powtarzam, że p = mówię po raz drugi: p

przypominam, że p = mówię po to, żebyś wiedział to, co wiedziałeś przedtem, a czego nie wiesz teraz

zwracam uwagę, że p = mówię po to, żebyś zaczął wiedzieć to, co postrzegasz

podkreślam, że p = mówię po to, żebyś nie mógł nie zauważyć: p

dodam, że p = powiem jeszcze coś o tym samym: p

reasumując, p = powiem to samo krótko: p

zawiadamiam, że p = mówię po to, żebyś wiedział: p

ogłaszam, że p = mówię po to, żeby wszyscy wiedzieli: p

zapowiadam, że p = mówię, co zrobię: p

przepowiadam, że p = mówię, co się stanie: p

przyznaję, że p = mówię: masz rację, że p; sądzę, że rozumiesz, że mógłbym pragnąć, żebyś nie miał racji

wyznaję, że p = mówię (o sobie) p; sądzę, że rozumiesz, że mógłbym pragnąć, żebyś nie wiedział

twierdzę, że p = mówię, że jest tak: p

odpowiadam, że p = mówię dlatego, że Y o to pytał: p

X obiecał Y-owi, że to zrobi = X powiedział do Y, że to zrobi, bo Y tego pragnie, po to, żeby Y mógł sądzić, że X to zrobi

X przyrzekł Y-owi, że to zrobi = X powiedział do Y-a, że to zrobi, po to, żeby Y mógł sądzić, że X musi to zrobić

X zgodził się z Y-em = X powiedział, że chce robić to, co Y powiedział, że chce żeby robili

X zaprotestował = X powiedział, że nie chce, żeby to robiono.

Eksplikacje semantyczne Anny Wierzbickiej nie mają nic wspólnego z definicjami encyklopedycznymi, do jakich przyzwyczaiły nas tradycyjne słowniki, nie jest bowiem ich zadaniem odwzorowywanie wiedzy naukowej, ale **kompetencji językowej** zwykłego użytkownika danego języka naturalnego, a więc pokazanie, jakie cechy denotatów konstytuują dla niego znaczenie danego wyrażenia, czyli tworzą jego konotację. zilustrujemy to przykładami, zestawiając eksplikacje semantyczne wybranych nazw gatunkowych i rodzajowych dokonane przez Wierzbicką (*Owoce i warzywa. Semantyka kategoryzacji ludzkiej*; w Wierzbicka A.:

Język – umysł – kultura, wyróżnienia graficzne według oryginału) z odpowiednimi definicjami podanymi w *Małym słowniku języka polskiego* (PWN, Warszawa 1968, w tekście MSJP) i w *Uniwersalnym Słowniku Języka Polskiego* (PWN, Warszawa 2003, w tekście USJP).

MSJP:

owoc – 1. organ rozmnażania się roślin okrytozalążkowych zawierający nasiona, u niektórych roślin jadalny. 2. zbiorowo: płody drzew i krzewów owocowych.

USJP:

owoc – 1. jadalna część rośliny powstała z kwiatu. 2. *bot.* wytwór zalążni słupka kwiatowego, organ roślin okrytozalążkowych osłaniający nasiona i rozsiewający je.

Wierzbicka:

OWOCE [FRUIT]

RZECZY RÓŻNYCH RODZAJÓW, KTÓRE LUDZIE JEDZĄ,
KTÓRE LUDZIE SPOŻYWAJĄ W TEN SAM SPOSÓB,
KTÓRE SĄ UŻYWANE W TEN SAM SPOSÓB I DO TEGO SAMEGO CELU
I KTÓRE SĄ UŻYWANE W ZESPOŁACH RZECZY RÓŻNYCH RODZAJÓW

MYŚLĄC O TAKICH RZECZACH LUDZIE MOGLIBY POWIEDZIEĆ O NICH TAK:

rosną one jako części pewnych rodzajów rzeczy wyrastających z ziemi
nie rosną w ziemi (POCHODZENIE)

chcąc wyobrazić sobie takie rzeczy

ludzie wyobrażaliby je sobie jako rosące na drzewach

[zaczynają wyrastać z pewnych części rzeczy wyrastających z ziemi o pewnej
porze roku

kiedy jest ciepło

i potem, kiedy było ciepło przez pewien czas

tak, że mogły zdarzyć się wszystkie te rzeczy w rzeczach wyrastających z ziemi
które w nich muszą się zdarzyć

nim mogą wyrosnąć na nich tamte rzeczy, które ludzie mogą jeść]

stają się one dobre do jedzenia, jeśli wystarczająco długo rosną na tych rze-
czach wyrastających z ziemi

zanim staną się dobre do jedzenia, są zielone albo zielonkawe z zewnątrz

ludzie powodują, że rzeczy tego rodzaju rosną w wielu miejscach

ponieważ chcą mieć te rzeczy, żeby je jeść

rzeczy, na których takie rzeczy mogą rosnąć, mogą także rosnąć w pewnych
miejscach, gdzie ludzie nie powodują, że one rosną

ale chcąc wyobrazić sobie takie rzeczy
ludzie wyobrażaliby sobie, że rosną one na rzeczach wyrastających z ziemi w
miejscach, gdzie ludzie spowodowali, że one rosną

mają one skórkę twardszą niż części wewnątrz (CZĘŚCI)
mają pewne małe części wewnątrz, oddzielne od innych części, niedobre
do jedzenia

te części włożone do ziemi mogą wyrosnąć na nowe rzeczy tego samego ro-
dzaju wyrastające z ziemi

mogą one być wystarczająco małe, żeby jedna osoba (ROZMIARY)
mogła łatwo włożyć więcej niż jedną rzecz tego rodzaju do ust i jeść je
wszystkie w tym samym czasie

albo za duże, żeby jedna osoba zjadła całą taką rzecz, po trochu, w tym samym
czasie

ale chcąc wyobrazić sobie takie rzeczy, ludzie wyobrażaliby je sobie jako
za duże, żeby jedna osoba łatwo włożyła ją sobie do ust i zjadła

i nie za duże, żeby jedna osoba mogła ją zjeść całą, po trochu, w jednym cza-
sie, trzymając ją w ręku

są one dobre do jedzenia bez gotowania (SPOSÓB JEDZENIA)
bez konieczności robienia z nimi czegoś jeszcze
bez innych rzeczy

i ludzie mogą jeść je z przyjemnością

zanim one staną się dobre do jedzenia, mogą być kwaśne
kiedy stały się dobre do jedzenia, są słodkie albo lekko słodkie, albo kwaśne,
ale dobre do jedzenia z czymś słodkim

chcąc wyobrazić sobie takie rzeczy, kiedy już stały się dobre do jedzenia
ludzie wyobrażaliby sobie rzeczy, które są lekko słodkie

mają one dużo soku

jedzenie ich bez gotowania sprawia, że ktoś czuje coś przyjemnego

są one także dobre do jedzenia gotowane z cukrem albo gotowane jako część
pewnych rzeczy, które zawierają cukier

ich sok jest dobry do picia

są one także dobre do jedzenia w postaci suszonej (bez soku)

są one myślane jako rzeczy, które są dobre i przyjemne do jedzenia
z którymi nie trzeba nic robić, żeby były gotowe do jedzenia,
i które ładnie wyglądają

MSJP:

warzywo – roślina zielna, jednoroczna lub wieloletnia, korzeniowa (np. bu-
rak), liściowa (np. sałata), cebulowa (np. cebula) lub owocowa (np. pomi-
dor), jadalna w stanie surowym lub jako potrawa.

USJP:

warzywo – roślina zielna, której jadalna część korzeniowa, liściowa, kwiatostanowa, cebulowa, nasienna lub owocowa używana jest jako pożywienie, jarzyna.

Wierzbicka:

WARZYWA [VEGETABLES]

RZECZY RÓŻNYCH RODZAJÓW, KTÓRE LUDZIE JEDZĄ
KTÓRE LUDZIE SPOŻYWAJĄ W TEN SAM SPOSÓB
I KTÓRE SĄ UŻYWANE W TEN SAM SPOSÓB I DO TEGO SAMEGO CELU

MYŚLĄC O TAKICH RZECZACH, LUDZIE MOGLIBY POWIEDZIEĆ O NICH TAK:

rosną one jako części pewnych rodzajów rzeczy (POCHODZENIE)
wyrastających z ziemi

mogą one rosnąć w ziemi albo nad ziemią
ale nie rosną na drzewach

ludzie powodują, że rzeczy, które mogą mieć takie części, rosną w wielu miejscach

ponieważ chcą mieć tamte rzeczy, żeby je jeść

są one dobre do jedzenia po ugotowaniu (SPOSÓB JEDZENIA)
z innymi rzeczami, które nie są słodkie
żeby nie być głodnym

niektóre takie rzeczy są dobre do jedzenia bez gotowania
ale po dodaniu do nich czegoś jeszcze

kiedy takie rzeczy są gotowe do jedzenia
można się spodziewać, że zostaną podzielone na części
i będą wyglądały inaczej niż wyglądają
kiedy rosną jako części rzeczy wyrastających z ziemi

są one myślane jako rzeczy, które są dobre do jedzenia dla ludzi

MSJP:

jabłko – 1. owoc jabłoni.

USJP:

jabłko – 1. owoc jabłoni.

Wierzbicka:

JABŁKA [APPLES]

RODZAJ RZECZY, KTÓRE LUDZIE JEDZĄ,

WYOBRAŻAJĄC SOBIE RZECZY TEGO RODZAJU, LUDZIE MOGLIBY POWIEDZIEĆ O NICH TAK:

rosną one na drzewach pewnego rodzaju przyłączone do gałęzi (POCHODZENIE)

wyrastają w pewnej porze roku z pewnych innych części tych drzew
ludzie powodują, że drzewa tego rodzaju rosną w wielu miejscach
ponieważ ludzie chcą mieć te rzeczy, żeby je jeść

są one dobre do jedzenia, kiedy rosną na drzewach wystarczająco długo
niedługo potem, jak stają się dobre do jedzenia oddzielają się od gałęzi i spadają na ziemię

są okrągłe (WYGLĄD – kształt)

z górną częścią, która wygląda trochę inaczej niż reszta
ponieważ jej środkowa część jest niżej niż przyległe części
i wystaje z niej krótka brązowa rzecz
za pomocą której cała rzecz jest przymocowana do gałęzi
zanim się od niej nie oddzieli

dolna część może także wyglądać trochę inaczej niż inne części strony zewnętrznej

jej część środkowa ma na środku ciemną kropkę
i także może być niżej niż przyległe części

zanim te rzeczy staną się dobre do jedzenia, są z zewnątrz zielonkawe (- kolor / dojrzałość)

kiedy staną się dobre do jedzenia, mogą być czerwone, żółte lub zielone
ale chcąc wyobrazić sobie rzeczy tego rodzaju, ludzie wyobrażaliby je sobie jako czerwone

są białawe wewnątrz

są za duże, żeby człowiek mógł trzymać je łatwo (- rozmiary)

w czasie jedzenia między palcem wskazującym a kciukiem

nie są za duże, żeby człowiek mógł wziąć jedną (taką rzecz) do ręki

nie są za duże, żeby spodziewać się, że człowiek zje taką jedną rzecz w jednym czasie

nie są za duże, żeby spodziewać się, że człowiek zje taką jedną rzecz jedną po drugiej

mają one gładką skórkę (SPOSÓB JEDZENIA – skórka)

która jest trudniejsza do zgryzienia niż części wewnątrz

mogą być jedzone ze skórką albo skórka może być odkrojona

są dobre do jedzenia bez gotowania (- przygotowanie i smak)

bez innych rzeczy

są twarde i przyjemne do gryzienia

wydają lekki odgłos, kiedy się je gryzie

są mało słodkie

są także dobre do jedzenia ugotowane, z cukrem
albo jako część pewnych innych rzeczy, które zawierają cukier

są one dobre do jedzenia pocięte na kawałki i wysuszone
ich sok albo inny płyn zrobiony częściowo z ich soku jest dobry do picia

między ich częścią górną a dolną jest część, (– części niejadalne)
której środek jest twardszy niż reszta
z kilkoma drobnymi częściami oddzielonymi od reszty
które mają twardą brązową skórę i są wewnątrz białawe
ludzie pozostawiają tę część nie zjedzoną

MSJP:

pomarańcza – *citrus sinensis*, drzewo z rodziny rutowatych, o białych pachnących kwiatach i złotych kulistych owocach; rośnie w strefie podzwrotnikowej i śródziemnomorskiej; owoc tego drzewa.

USJP:

pomarańcza – 1. *citrus sinensis*, wiecznie zielone drzewo cytrusowe o skórzastych liściach, białych pachnących kwiatach i dużych soczystych słodkich owocach, uprawiane w krajach tropikalnych i subtropikalnych. 2. owoc tego drzewa, okrągły o porowatej skórce koloru żółtego z czerwonym odcieniem, okrywającej soczysty, aromatyczny miąższ.

Wierzbicka:

POMARAŃCZE [ORANGES]

RODZAJ RZECZY, KTÓRE LUDZIE JEDZĄ

WYOBRAŻAJĄC SOBIE RZECZY TEGO RODZAJU, LUDZIE MOGLIBY POWIEDZIEĆ

O NICH TAK:

rosną one na drzewach pewnego rodzaju w ciepłych (POCHODZENIE)
częściach ziemi, przyczepione do gałęzi

wyrastają w pewnej porze roku z pewnych innych części tych drzew
ludzie powodują, że drzewa tego rodzaju rosną w wielu miejscach
ponieważ ludzie chcą mieć te rzeczy, żeby je jeść

są one dobre do jedzenia kiedy rosną na drzewach wystarczająco długo

są okrągłe (WYGLĄD – kształt)

zanim staną się dobre do jedzenia, są z zewnątrz (– kolor / dojrzałość)
zielonkawe

kiedy staną się dobre do jedzenia, są pomarańczowe z zewnątrz
są pomarańczowe wewnątrz

na zewnątrz skórka wygląda, jakby miała wiele drobnych otworów tak małych jak kropki

rzeczy tego rodzaju są za duże, żeby człowiek mógł trzymać (– rozmiary) łatwo jedną między palcem wskazującym a kciukiem kiedy ściąga skórę nie są za duże, żeby człowiek nie mógł łatwo jednej (takiej rzeczy) ująć w dłoń

nie są za duże, żeby spodziewać się, że człowiek zje całą taką jedną rzecz za jednym razem

są za duże, żeby spodziewać się, że człowiek zje więcej niż jedną w jednym czasie

mają one grubą skórę, białą od spodu (SPOSÓB JEDZENIA – skórka) którą można łatwo rozciąć i ściągnąć stopniowo palcami skórka nie jest dobra do jedzenia na surowo

wewnątrz mają one wiele części jednakowego rozmiaru (– części jadalne) i kształtu

które można oddzielić palcami jedną od drugiej aby jeść jedną po drugiej

kiedy się zdejmie skórę

można zobaczyć zaokrąglone grzbiety wszystkich tych części

które razem składają się na powierzchnię tej całej rzeczy

te grzbiety są białe

ponieważ części białej wewnętrznej części skórki są przyłączone do nich

między tymi częściami mogą także być, przyłączone do nich, białe kawałki, podobne do wewnętrznych części skórki

można je ściągnąć palcami

ich górne końce są wszystkie w tym samym miejscu, w środku górnej części a ich dolne części są wszystkie w tym samym miejscu, w środku dolnej części kiedy oddziela się je jedną od drugiej można zobaczyć, że część między dwoma końcami jest cienka

i że cienkie części wszystkich tych części są wszystkie razem w tym samym miejscu, w środku części dolnej

i że obie strony obu boków tej cienkiej części są płaskie

kiedy rozcina się rzeczy tego rodzaju na pół, od jednego boku do drugiego, można zobaczyć ładny wzorek

z wieloma częściami, które wszystkie wyglądają tak samo wszystkie one otaczają punkt znajdujący się w środku

każda z tych części wygląda, jakby składała się z wielu bardzo drobnych twardej części połączonych z sobą

z których każda ma wewnątrz tylko sok

i cienką gładką powierzchnię, przez którą wszystko widać

tak że można by połknąć małe części rzeczy tego rodzaju nie używając zębów

wewnątrz tych oddzielnych rzeczy są w różnych miejscach pewne drobne twarde części oddzielne od pozostałych zaokrąglone, z jednym końcem ostrym, te części nie są dobre do jedzenia ludzie wyjmują je albo wypluwają z ust

rzeczy tego rodzaju są dobre do jedzenia bez gotowania
bez innych rzeczy są miękkie i słodkie mają dużo soku, który jest dobry dla ludzi do picia ludzie mogą wyciskać sok rękami

(SPOSÓB JEDZENIA

– przygotowanie i smak)

rzeczy tego rodzaju mogą być także gotowane z cukrem, z małymi kawałkami skórki aby zrobić słodką substancję dobrą do jedzenia, którą nakłada się na coś, co ludzie jedzą na co dzień, żeby nie być głodnymi skórka gotowana z cukrem jest dobra do jedzenia bez innych rzeczy albo pocięta na małe kawałki i dodana do innych rzeczy, które zawierają cukier

MSJP:

pomidor – *solanum lycopersicum*, roślina z rodzaju psiankowatych, o dużych jadalnych owocach koloru czerwonego lub żółtego; owoc tej rośliny.

USJP:

1. *Lycopersicon esculentum*, jednoroczna roślina warzywna o pierzastych liściach, żółtych kwiatach, uprawiana ze względu na jadalne owoce.
2. owoc tej rośliny, zwykle duży, okrągły, koloru czerwonego, żółtego lub pomarańczowego, zawierający dużo witamin i soli mineralnych.

Wierzbicka:

POMIDORY [TOMATOES]

RODZAJ RZECZY, KTÓRE LUDZIE JEDZĄ

WYOBRAŻAJĄC SOBIE RZECZY TEGO RODZAJU, LUDZIE MOGLIBY POWIEDZIEĆ

O NICH TAK:

rosną ponad ziemią

(POCHODZENIE)

jako części rzeczy pewnego rodzaju, które wyrastają z ziemi

ludzie powodują, że rosną one w wielu miejscach

ponieważ chcą je mieć, żeby je jeść

są zaokrąglone

(WYGLĄD – kształt)

część, która jest w części górnej, kiedy rosną, jest spłaszczona

środek tej części jest trochę niżej niż części dokoła niej

i ma twarde miejsce w środku

w miejscu, gdzie ta rzecz jest przymocowana do innych części tej rzeczy, która wyrasta z ziemi

w środku części dolnej jest mała czarna kropka

zanim staną się dobre do jedzenia, są zielone z zewnątrz

(– kolor)

i zielonkawe wewnątrz

kiedy staną się dobre do jedzenia, są jasnoczerwone z zewnątrz i lekko czerwone wewnątrz

mają gładką cienką skórę

mają dużo miękkich części w środku

(– środek)

gdzie jest trochę drobnych żółtawych części, nie miękkich, które są oddzielne od innych części

i gdzie jest sok

te części są otoczone ze wszystkich stron przez części twardsze

kiedy się przetnie rzeczy tego rodzaju od jednego boku do drugiego

można zobaczyć te miękkie części przy wszystkich bokach

z tymi drobnymi żółtawymi częściami w nich

otoczonymi ze wszystkich stron przez te twardsze części

są one za duże, żeby człowiek mógł

(– rozmiary)

łatwo trzymać jedną taką rzecz między palcem wskazującym a kciukiem jedząc ją

nie są za duże, żeby jedna osoba nie mogła łatwo trzymać jednej takiej rzeczy w jednej ręce jedząc ją

są dobre do jedzenia bez gotowania

(sposób jedzenia)

z innymi rzeczami, które nie są słodkie

ich sok jest dobry do picia

rzeczy tego rodzaju można także gotować albo robić z nimi coś innego, żeby zrobić z nich coś, co można dodawać do innych rzeczy, które ludzie jedzą, do rzeczy nie słodkich po to, by lepiej smakowały

rzeczy tego rodzaju nie mają żadnych części, które są twarde albo których nie można jeść

Radzieccy twórcy tzw. **objaśniająco-kombinatorycznego słownika** języka rosyjskiego, **Jurij Apresjan**, **Igor Mielczuk** i **Aleksander Żółkowski** postawili sobie cel bardziej pragmatyczny. Chcieli oni stworzyć taki słownik, „który powinien (...) zapewniać zgrabne, idiomatyczne i sprawne wyrażanie zadanej treści; inaczej mówiąc – słownik powinien podawać w postaci jawnej i logicz-

nie uporządkowanej wszystkie najróżnorodniejsze wiadomości, które mogą być potrzebne, by prawidłowo wybrać i zastosować wyraz lub zwrot, konieczne dla wyrażania myśli w danym kontekście mownym”. Język ten może też pełnić rolę języka pośrednika przy tłumaczeniu z jednego języka naturalnego na drugi, a więc na przykład przy przekładzie maszynowym (komputerowym).

Rozważając problem **głębokości analizy semantycznej** wyrazów, autorzy doszli do wniosku, że zależy ona od tego, jaka część znaczenia ma dokładny odpowiednik jednowyrazowy w innym języku. Oczywiście, im znaczenie jest bardziej elementarne, tym większe jest prawdopodobieństwo tego, że będzie ono posiadać jednowyrazowy odpowiednik językowy i tym więcej jest języków, w których posiada ono taki sposób wyrażania. Możliwy więc jest rozkład znaczenia każdego wyrazu na składniki elementarne, ale głębokość analizy powinna być określona przez konkretne zadanie. Nie przedstawiamy tu całego opracowanego przez radzieckich lingwistów **modelu syntezy semantycznej**, ograniczymy się tylko do omówienia zastosowanej przez nich metody wyróżniania elementarnych znaczeń przez porównanie znaczeń złożonych, metoda ta ma bowiem wartość ogólnolingwistyczną.

Autorzy zastosowali w opisie wyrazów języka naturalnego tzw. **funkcje leksykalne**, czyli operacje przyporządkowujące danemu wyrazowi X inny wyraz Y, w pewien sposób związany z X pod względem znaczeniowym na paradygmatycznej lub syntagmatycznej osi języka. Wyróżnili przy tym dwa typy funkcji leksykalnych: zamiany leksykalne i parametry leksykalne. **Zamiany leksykalne** to funkcje przyporządkowujące danemu wyrazowi X taki wyraz Y, którego cechy semantyczne (w przekładzie na język semantyczny) w całości lub częściowo pokrywają się z cechami semantycznymi X. Zamiany leksykalne działają na paradygmatycznej osi języka. **Parametry leksykalne** działają na syntagmatycznej osi języka i przyporządkowują danemu wyrazowi X taki wyraz Y, który jest z nim związany semantycznie na osi syntagmatycznej, czyli występuje z nim razem w jednym kontekście składniowym. A oto przykłady (w oryginale symbole funkcji leksykalnych pisane są kursywą, wyrażenia języka naturalnego drukiem prostym – tu przykłady językowe podajemy kursywą, jak w całym tekście) :

Zamiany leksykalne:

1. Syn – synonim:
Syn (*tylko*) – *jedynie*, Syn (*lingwistyka*) – *językoznawstwo*, Syn (*auto*) – *samochód*,
2. Conv – konwers;
Conv (*kupić*) – *sprzedać*, Conv (*zwierzchnik*) – *podwładny*, Conv (*mąż*) – *żona*, Conv (*matka*) – *dziecko*, Conv (*dalej*) – *bliżej*, Conv (*trudniejszy*) – *łatwiejszy*,
3. Anti – antonim:
Anti (*ładny*) – *brzydki*, Anti (*dobro*) – *zło*, Anti (*przyjaciel*) – *wróg*,
4. Gener – wyrażenie nadrzędne:
Gener (*niebieski*) – *kolor*, Gener (*czolgać się*) – *poruszać się*, Gener (*ciecz*) – *substancja*,

5. S_0 – rzeczownik synonimiczny do X niebędącego rzeczownikiem:
 S_0 (*poruszać się*) – *ruch*, S_0 (*mlócić*) – *mlócka*, S_0 (*rodzić*) – *poród*, S_0 (*być białym*) – *bialość*, *biel*,
6. A_0 – przymiotnik synonimiczny do X niebędącego przymiotnikiem:
 A_0 (*Ziemia*) – *ziemski*, A_0 (*ziemia*) – *ziemny*, A_0 (*islam*) – *muzułmański*, A_0 (*Armenia*) – *ormiański*,
7. Adv_0 – przysłówek synonimiczny:
 Adv_0 (*sens*) – *sensownie*,
8. V_0 – czasownik synonimiczny:
 V_0 (*śmierć*) – *umierać*, V_0 (*razem z*) – *towarzyszyć*, V_0 (*napadać*) – *napad*.

Parametry leksykalne:

1. Magn – określenie o znaczeniu „wyższy stopień, bardzo”:
 Magn (*głupiec*) – *kompletny*, Magn (*głupi*) – *bezdennie*, *jak stołowe nogi*, *jak but*, Magn (*blondyn*) – *świński*, Magn (*deszcz*) – *ulewny*, Magn (*spać*) – *mocno*, *kamiennym snem*, *jak zabity*, Magn (*porządek*) – *idealny*, *wzorowy*,
2. Bon – określenie o znaczeniu „dobry”:
 Bon (*wpływ*) – *pozytywny*, Bon (*kupować*) – *tanio*, Bon (*propozycja*) – *kusząca*,
3. A_i – typowe określenie i-tego uczestnika sytuacji: A_1 – to taki, który ..., A_2 – takie, którego ...:
 A_1 (*nic nie umieć*) – *ignorant*, A_1 (*wąsy*) – *wąsaty*, A_1 (*oczekiwać dziecka*) – *w ciąży*, A_2 (*kierować, zarządzać*) – *pod zarządkiem, pod kierownictwem*,
4. $Able_i$ – typowe określenie i-tego potencjalnego uczestnika sytuacji: $Able_1$ – taki, który (łatwo) może, $Able_2$ – taki którego można:
 $Able_1$ (*odbywać się, mieć miejsce*) – *potencjalny*, $Able_2$ (*jeść*) – *jadalny*, $Able_2$ (*pić*) – *pitny*,
5. S_i – typowa nazwa i-tego uczestnika sytuacji:
 S_1 (*sprzedawać*) – *sprzedawca*, S_2 (*sprzedawać*) – *towar*, S_3 (*sprzedawać*) – *kupujący, klient*, S_4 (*sprzedawać*) – *cena*,
6. S_{instr} – rzeczownik nazywający instrument:
 S_{instr} (*walczyć*) – *broń*, S_{instr} (*przekonywać*) – *argument, dowód*,
7. S_{loc} – rzeczownik oznaczający miejsce:
 S_{loc} (*walczyć*) – *pole walki*, S_{loc} (*leczyć*) – *szpital*, S_{loc} (*strząły*) – *kolczan*,
8. Sing – jeden egzemplarz:
 Sing (*mak*) – *ziarnko*, Sing (*chleb*) – *bochenek*, Sing (*deszcz*) – *kropla*, Sing (*gniew*) – *atak, wybuch*,
9. Mult – zbiór:
 Mult (*owca*) – *stado*, Mult (*ryba*) – *ławica*, Mult (*człowiek*) – *grupa*,
10. Cap – naczelnik, „głowa”:
 Cap (*instytucja*) – *dyrektor*, Cap (*uniwersytet*) – *rektor*, Cap (*plemię, szczep*) – *wódz*,

11. Equip – personel, obsada:
Equip (*pułk*) – *żołnierze*, Equip (*instytucja*) – *pracownicy*, Equip (*klasztór*) – *mnisi, mniszki, zakonnicy, zakonnice, bracia, siostry*,
12. Figur – standardowe określenie figuralne sytuacji:
Figur (*oblężenie*) – *pierścień*, Figur (*brawa*) – *huragan* (ze znaczeniem Magn),
13. Centr – centralna, kulminacyjna część przedmiotu lub procesu:
Centr (*miasto*) – *centrum, śródmieście*, Centr (*szczęście*) – *szczyt*, Centr (*sila*) – *rozkwit, pełnia*,
14. Incep – czasownik o znaczeniu zaczynać:
Incep (*spać*) – *zasypiać*, Incep (*mówić*) – *zacząć, zagaić*,
15. Fin – czasownik o znaczeniu kończyć:
Fin (*przytomność*) – *tracić*, Fin (*spać*) – *budzić się*, Fin (*kochać się*) – *odkochać się*,
16. Caus – czasownik o znaczeniu „czynić tak, by X zaczęło mieć miejsce”, tzw. kauzacja:
Caus (*porządek*) – *zaprowadzić*, Caus (*rodzina*) – *zakładać*,
– Caus może być również zamianą leksykalną, np. Caus (*spać*) – *usypiać*,
17. Liqu – likwidować:
Liqu (*rozkaz*) – *odwołać*, Liqu (*spać*) – *budzić*,
18. Result – rezultat:
Result (*kłaść się*) – *leżeć*, Result (*uczyć się*) – *umieć*, Result (*umierać*) – *być martwym*,
19. Fact – realizować się, stać się faktem:
Fact (*marzenie*) – *ziszcza się*, Fact (*pokusa*) – *zwycięża*.

Oczywiście nie są to wszystkie funkcje leksykalne, lecz tylko ich przykłady. W opisie wyrażen języka naturalnego najczęściej występują funkcje złożone, np.:

Anti Magn (zwycięstwo) – *wątpliwe, problematyczne, wątpliwej wagi*,
Magn Anti (zwycięstwo) – *Pyrrusowe*
Anti Able₂ (zwycięzać) – *niezwyciężony*.

Omówione tu **języki semantyczne** są przykładem języków tworzonych dla potrzeb językoznawstwa – mają one pełnić rolę metajęzyków przy pewnych zadaniach lingwistycznych: opisie systemu języka naturalnego lub eksplikacji jego tekstów. Tworzenie takich języków poprzedzać musi wszechstronna analiza systemów języków naturalnych zmierzająca do ustalenia relewantnych dla tych języków cech opisywanej przez nie rzeczywistości. Analiza taka może być podstawą także do zbudowania języków, które oprócz funkcji metajęzykowej pełnić mogą także funkcję języków przedmiotowych, a więc języków służących do opisywania rzeczywistości pozatekstowej. W systemie takich języków muszą być odwzorowane w ich planie wyrażania wszystkie te cechy opisywanej rzeczywistości,

które są relewantne dla użytkowników tych języków. Analiza rzeczywistości poprzedzać więc musi proces budowy takiego języka.

Przykładem tego typu języków, nazywanych **kodami semantycznymi**, będą dwa języki. Pierwszy to kod semantyczny do opisu operacji transportowania, będący częścią kodu semantycznego projektowanego przeze mnie i Olgierda Wojtasiewicza dla systemu SINTO – prace nad tym kodem pozostały, niestety, w stadium projektowym. Projekt poprzedzała analiza polskich czasowników ruchu oraz zbudowany przeze mnie metajęzyk ich opisu, czemu poświęciłam rozprawę habilitacyjną. Drugi przykład to kod semantyczny pomyślany jako język opisujący wybrany wycinek rzeczywistości – jest to zaproponowany przeze mnie kod semantyczny do opisu dyscyplin sportowych.

Kodu semantycznego służącego do eksplikacji znaczenia polskich czasowników ruchu nie mogę tu zaprezentować w sposób pełny, ograniczę się tylko do przytoczenia wybranych w wyniku analizy czasowników komunikujących ruch ich cech dystynktywnych – elementarnych informacji o rzeczywistości sytuacji ruchu, które zostały zakodowane w planie wyrażenia języka polskiego, w jego systemie czasownikowym. Porównanie tych cech z cechami kodowanymi w języku służącym do opisu operacji transportowania najlepiej obrazuje, jak przypadkowe (wydawałoby się) informacje o rzeczywistości kodowane są w systemie leksykalnym języka naturalnego i jak bardzo jest on odległy od klasyfikacji w sensie logicznym.

Budowa kodu semantycznego do opisu sytuacji ruchu

Na podstawie analizy znaczeń polskich czasowników ruchu ustalono listę elementarnych wyróżnianych językowo informacji o sytuacji ruchu, przy pomocy których można by opisać w sposób możliwie najbardziej ekonomiczny dowolne sytuacje ruchu, nawet te, dla których brak w języku naturalnym leksykalnych wykładników. Oczywiście analiza jednego tylko języka nie może dać pełnego zbioru jednostek elementarnych potrzebnych do realizacji takiego zadania, wydaje się jednak, że w sposób wystarczający ukazuje metodę tworzenia takiego zbioru, który powinien być uzupełniany na podstawie analizy innych języków (w przypadku, gdyby zbiór ten miał spełniać funkcję języka pośrednika w procesie tłumaczenia), a także na podstawie analizy rzeczywistości pozajęzykowej.

Lista **dystynktywnych cech ruchu** kodowanych w czasownikach języka polskiego i przykłady ich językowej realizacji :

1. Informacje o poruszającym się ciele:

a) Rodzaj poruszającego się ciała:

- ciecz (*tryskać, strzykać, kapać, lać*),
- niemowlę poruszające się za pomocą czterech kończyn (*raczkować*),
- koń poruszający się za pomocą nóg lub jeździec (*galopować*),
- krew (*broczyć*),
- światło lub płomień (*pełgać*),
- ciało poruszające się za pomocą nóg, ruch nóg (*skakać, wierzgać*),
- pies lub ogon psa (*merdać*),

- dłoń lub głowa (*skinąć*),
 - szczęki (*ziewać*),
 - środek transportu lub ruch przy użyciu środka transportu (*jechać*),
 - powieki (*mrugać, mrużyć*).
- b) Kształt poruszającego się ciała:
- kanciasty (*kantować*),
 - kulisty (*kapać*),
 - strumień (*tryskać*).
- c) Pozycja poruszającego się ciała w momencie zakończenia ruchu:
- na kolanach (*klęknąć*),
 - leżąca (*położyć*),
 - stojąca (*postawić*).
- d) Liczba poruszających się ciał:
- więcej niż jedno (lub ciało złożone z odrębnie poruszających się elementów) (*sypać*).
2. Informacje o sile sprawczej ruchu:
- a) Lokalizacja siły sprawczej ruchu:
- wewnętrzna (ruch samoistny) (*biec, jechać, iść*),
 - zewnętrzna (ruch niesamoistny) (*toczyć, pchać, kołysać*).
- b) Lokalizacja zewnętrznej siły sprawczej ruchu:
- przed poruszającym się ciałem (*ciągnąć, wlec, holować*),
 - za poruszającym się ciałem (*pchać*),
- c) Kontakt pomiędzy ciałem poruszającym się a ciałem powodującym ruch:
- brak (*poganiać, popędzać*),
 - istnienie (*nieść*).
- d) Przemieszczanie się ciała powodującego ruch wraz z ciałem, którego zmianę lokalizacji powoduje:
- brak (*rzucić, cisnąć*),
 - istnienie (*nieść, wieźć*).
- e) Ochrona przed spadaniem ciała zmieniającego lokalizację przez ciało powodujące tę zmianę:
- brak (*strącać, podrzucać*),
 - istnienie (*zestawić, podsadzić*),
 - użycie środka transportu chroniącego przed spadaniem (*wieźć*).
- f) Czas działania siły sprawczej ruchu:
- w momencie początkowym ruchu (ruch inicjowany) (*pchnąć, rzucić*),
 - przez cały czas trwania ruchu (ruch podtrzymywany) (*pchać, nieść*).
- g) Wysiętek źródła siły sprawczej ruchu:
- duży (*taszczyć, dźwigać*).
3. Informacje o przestrzeni ruchu:
- a) Zmiana ośrodka w czasie trwania ruchu:
- brak zmiany (*plywać, fruwać*),
 - istnienie zmiany (*podskakiwać*),
- b) Jednoczesne ośrodki ruchu:

- jeden (*lecieć, płynąć, biec*),
 - więcej niż jeden – ruch po powierzchni (*sunąć, toczyć*),
 - więcej niż jeden – ruch nie po powierzchni (*koziolkować*).
- c) Kontakt pomiędzy powierzchnią ruchu a powierzchnią poruszającego się ciała:
- w momencie początkowym ruchu (*wyskoczyć*),
 - trwały (*sunąć*).
- d) Ściana kontaktująca z powierzchnią ruchu:
- różna (*toczyć, turlać, tarzać*),
 - ta sama cała (*przesuwać*),
 - różne wycinki tej samej ściany (*pelznąć*).
- e) Rodzaj ośrodka (stan skupienia):
- cieczer (*pływać*),
 - gaz lub próżnia (*latać, fruwać*).
- f) Opór ośrodka:
- duży (opór powierzchniowy) (*brnąć, przeciskać*),
 - mały (tarcie) (*ślizgać się*).
- g) Kształt toru ruchu:
- zakrzywiony (tak, że obejmuje co najmniej trzy ściany ciała, którego obecność w przestrzeni ruchu jest implikowana) (*okrążyć, opływać*),
 - po krzywej (*smuć się, szwendać się, kluczyć, błąkać się*),
 - po linii zbliżonej do okręgu (*krążyć*),
- h) Zagęszczenia przestrzeni ruchu przez poruszające się ciało:
- ekspansja przestrzenna (*rozlewać, rozwiewać*),
 - wypełnianie granic obszaru (*zaśmiecać, zakryć, zasypać*),
 - kurczenie się obszaru (*zjechać się, zsypać*).
4. Informacje o rodzaju ruchu:
- a) Zmiana lokalizacji w wyniku ruchu:
- brak zmiany (*wiercić się, wirować*),
 - istnienie zmiany (*przemieszczać, wyjść, dopłynąć*).
- b) Zmiana zwrotu (czyli ruch ciała wokół osi, ruch obrotowy):
- brak zmiany (*spadać, zapadać się*),
 - istnienie zmiany (*obrać się*).
- c) Wielkość zmiany:
- zwrot o kąt $\leq 180^\circ$ (*odwracać*),
 - zwrot o kąt $< 360^\circ$ (*zwracać*),
 - zwrot o kąt $\geq 360^\circ$ (*obrać, okręcać*).
- d) Powtarzalność (fazowość) elementarnych ruchów składowych:
- brak powtarzalności (*iść, płynąć*),
 - istnienie powtarzalności (*drgać, kołysać, trząść, huścić, bujać*).
- e) Częstotliwość zmian:
- duża (*drżeć*),
- f) Lokalizacja osi obrotu:
- wewnętrzna (*wirować, obracać*),

- zewnętrzna (*opływać*).
5. Informacja o kierunku ruchu:
- a) Stosunek kierunku ruchu do kierunku działania siły grawitacji:
 - przeciwny (*wzlatywać*),
 - zgodny (*zsuwać się, spuszczać, strącić*),
 - zgodny – ruch pod wpływem grawitacji (*spadać*).
 - b) Zmiana kierunku ruchu:
 - brak zmiany (*spadać*),
 - istnienie zmiany przy nierелеwantnej wartości zmiany (*skrecać*),
 - zmiana o małej wartości kąta (*zbaczać*),
 - zmiana kierunku o kąt = 180° (na kierunek przeciwny) (*zawracać*).
 - c) Stosunek kierunku ruchu do miejsca wcześniejszej lokalizacji poruszającego się ciała:
 - ruch w kierunku wcześniejszej lokalizacji (*zawracać*),
 - osiągnięcie wcześniejszej lokalizacji (*wrócić*).
 - ruch w kierunku miejsca wcześniejszej lokalizacji (zakładane osiągnięcie miejsca wcześniejszej lokalizacji) (*wracać*).
6. Informacje o prędkości ruchu:
- a) Wartość prędkości ruchu:
 - duża (*gnać, pędzić, mknąć*),
 - mała (*wlec się*).
 - b) Zmiana prędkości ruchu:
 - zwiększanie się prędkości (przyspieszanie) (*przyspieszać, rozpędzać się*),
 - zmniejszanie się prędkości do wartości = 0 (*zahamować*),
 - zmniejszanie się prędkości (zwalnianie) (*zwalniać, spowalniać*).
 - c) Harmonijność przebiegu ruchu:
 - gwałtowny (*potrząsnąć, tryskać, bryzgać*),
 - harmonijny (*kolebać*).
7. Informacje o zjawiskach akustycznych towarzyszących ruchowi:
- obecność zjawisk akustycznych (*człapać, łopotać*).
8. Informacje o relacjach między poruszającym się ciałem a innymi przedmiotami:
- a) Obecność innych przedmiotów w przestrzeni ruchu:
 - obecność ciał, których zachowanie przestrzenne oraz liczba jest irrelevantna (*mijać, zbliżyć się, odchodzić*),
 - obecność kilku ciał, których zachowanie przestrzenne jest irrelevantne (*lawiować*),
 - obecność ciał znajdujących się w ruchu (*wyprzedzić, dogonić, dopędzić, ścigać*),
 - obecność obserwatora ruchu (*cofać, nadchodzić, nadjechać*).
 - b) Stosunek między kierunkami poruszających się ciał:
 - zgodny (*gonić, wyprzedzać, dopędzać*).
 - c) Kierunek ruchu poruszającego się ciała w stosunku do ciała współwystępującego w tej samej przestrzeni:

- adlatywny (do) (*dochodzić, dopływać*)
 - ablatywny (od) (*odchodzić, odpływać*)
- d) Komunikowanie lokalizacji poruszającego się ciała:
- w momencie początkowym ruchu (*wychodzić, wypływać*),
 - w momencie końcowym ruchu (*wejść, wpłynąć*).
- e) Odległość poruszającego się ciała od lokalizatora w relewantnym momencie ruchu:
- co najmniej ogólna bliskość (*nadchodzić, podpływać, dobiegać, dopchnąć*),
 - co najmniej wspólna część powierzchni (*zjechać, najechać*).
- f) Zmiana odległości między ciałami leżącymi we wspólnej przestrzeni:
- zmniejszanie się (*doganiać, dochodzić*),
 - zmniejszanie się do odległości = 0 (*dogonić, zderzyć się, uderzyć*),
 - zwiększanie się (*odchodzić, odpływać*).
- g) Kolizyjność toru ruchu poruszającego się ciała z ciałem znajdującym się we wspólnej przestrzeni :
- bezkolizyjna (*mijać*),
 - kolizyjna (*zderzyć się*).
 - uniknięcie potencjalnej kolizji (*omijać, ominąć*).
- h) Usytuowanie w momencie początkowym ruchu poruszającego się ciała w stosunku do ciała, którego ruch jest implikowany:
- przed (*uciekać*),
 - za (*gonić*).
- i) Zakładane usytuowanie w momencie zakończenia ruchu ciała, którego ruch jest komunikowany w stosunku do ciała, którego ruch jest implikowany:
- przed (*wyprzedzić*),
 - przynajmniej w odległości = 0 (*dogonić*).
9. Możliwość określania czasownikiem przebiegu drogi (potencjalnego toru ruchu) (*biec, prowadzić, iść, ciągnąć się*).
10. Możliwość określania czasownikiem relacji „bycia początkowym punktem ruchu” (*cieknąć*).

Wszystkie te cechy ruchu wyrażane przez polskie czasowniki (będące komponentami ich znaczenia) mogą być także w strukturze znaczeniowej danego czasownika irrelewantne.

Projekt kodu semantycznego do opisu operacji transportowania

Na omówionej wyżej zasadzie zbudować można kod semantyczny służący do opisu już nie wyrażen języka naturalnego, lecz wszelkiego rodzaju zjawisk pozajęzykowych. Prace nad takim kodem prowadzono w Katedrze Lingwistyki Formalnej Uniwersytetu Warszawskiego. Dla ilustracji problemu podajemy charakterystyki wybranych sytuacji związanych z operacją transportowania jako operacją polegającą na ruchu pewnego rodzaju obiektów w pewnej wyróżnionej przestrzeni fizycznej.

Przyjmuje się przy tym następujące obiekty pierwotne podlegające operacji transportowania:

- informacja – I,
- energia – E,
- substancja (materia) – M.

Informacja rozumiana jest szeroko, także w sensie informacji biologicznej. Energia rozumiana jest w sensie najbliższym rozumieniu potocznemu, natomiast substancja (materia) jest pojęciem umownym, obejmującym obiekty fizyczne w potocznym sensie (terminologia ta nie zakłada niematerialności informacji i energii – substancją nazywa się te obiekty, które nie są ani informacją ani energią).

Dla informacji przyjęto dwa kryteria charakterystyki: semantyczność i znakowość, przy czym cecha semantyczności wyróżniona jest tu ze względu na odbiorcę – semantyczną będzie taka informacja, która odbierana jest przez świadomie interpretującego odbiorcę, dla którego odebrany komunikat pełni funkcję oznaki lub znaku, a więc zastępuje jakiś inny element rzeczywistości (inną sytuację), natomiast cecha znakowości wyróżniona jest ze względu na to, czy owo zastępowanie odbywa się na zasadzie arbitralności, czy też nie. Tak więc kryteria semantyczności i znakowości pozwalają wyróżnić następujące klasy informacji:

- informacja asemantyczna nieznakowa – I00
np. informacja strukturalna w biologii, muzyka,
- informacja asemantyczna znakowa – I0Z
np. zapis muzyczny,
- informacja semantyczna nieznakowa – IS0
np. informacja symptomowa (symptomy chorobowe),
- informacja semantyczna znakowa – ISZ
np. informacja w języku naturalnym.

Informacja może być utrwalona (U) lub nieutrwalona (0), może też być, w miarę potrzeby, scharakteryzowana przez podanie symbolu określającego rodzaj sygnału będącego nośnikiem informacji, np.:

- fala akustyczna – a,
- fala świetlna – l,
- fala termiczna – t,
- zjawisko elektryczne – e.

Informacja utrwalona może być scharakteryzowana przez określenie rodzaju nośnika informacji (jest nim na przykład taśma magnetofonowa, filmowa lub płyta gramofonowa w wypadku informacji akustycznej). Tą dalszą charakterystyką informacji nie będziemy się tu na razie zajmować.

Wstępnym kryterium charakterystyki substancji (materii) jest jej stan skupienia, mamy więc odpowiednio:

- substancję gazową – G,
- substancję płynną – C,
- substancję stałą – S.

Energia podlega na razie dalszej klasyfikacji jedynie na:

- energię elektryczną – E,

- energię mechaniczną – M
- energię termiczną – T.

Każda z tej postaci energii, teoretycznie, może być utrwalona (zmagazynowana) – U lub nie – 0.

Operacja transportowania rozumiana jest możliwie jak najszerszej, jako wszelkiego rodzaju przemieszczanie, przenoszenie, przekazywanie odbywające się dzięki działaniu siły zewnętrznej (jest to więc ruch wymuszany lub inicjowany). Obserwacja rzeczywistości każe tu przyjąć następujące kryteria charakterystyki tej operacji:

1. Ośrodek, w którym odbywa się ruch (transportowanie), charakteryzowany zgodnie ze stanem skupienia stanowiącej go materii. Mamy więc odpowiednio następujące ośrodki ruchu:

- ośrodek gazowy lub (quasi) próżniowy – G
- ośrodek płynny – C
- ośrodek stały – S.

2. Udział powierzchni – mamy tu następujące możliwości ruchu po powierzchni:

- styk ośrodka płynnego i gazowego – C/G
- styk ośrodka stałego i gazowego – S/G
- styk ośrodka stałego i płynnego – S/C

(znak stojący w zapisie symbolicznym na pierwszej pozycji wskazuje ośrodek, po którego powierzchni odbywa się ruch).

3. Stosowanie przewodnicy lub nie – p, – 0

– jest to więc opozycja prywatywna, a termin przewodnica rozumiany jest szeroko, jako wszelkie urządzenie, a nawet obiekt naturalny (np. rzeka w wypadku spławiania nią transportowanych obiektów) wyznaczające w sposób konieczny drogę (tor) transportu (dla przewozu samochodowego droga nie jest przewodnicą, gdyż przewóz taki może się odbywać po innym torze niż wyznaczony przebiegiem drogi). Oczywiście możliwa jest dalsza charakterystyka przewodnicy w zależności chociażby od jej stanu skupienia.

4. Stosowanie pojemnika (kontenera) lub nie – k, – 0

– jest to również opozycja prywatywna. Termin pojemnik także rozumiany jest szeroko. Kategoryzacji pojemników nie będziemy tu przeprowadzać.

Przyjęcie tylko tych kilku opozycji pozwala już opisać dość dużo różnych sytuacji transportowania. Niżej podamy przykłady takich charakterystyk, stosując następujący układ charakterystyk w ciągu:

obecność przewodnicy, obecność pojemnika, ośrodek, transportowany obiekt.

W wypadku przewodnicy oraz transportowanego obiektu czasami podawana będzie ich charakterystyka (takie wyrażenia złożone ujęte są w nawiasy). Wyrażenie charakteryzujące informację ma tu następujący układ:

semantyczność, znakowość, nośnik, utrwalenie, utrwalacz

(ten ostatni, będący nośnikiem nośnika informacji, scharakteryzowany jest jedynie ze względu na rodzaj substancji utrwalającej).

Przy zapisie charakterystyk operacji transportowania zrezygnowaliśmy z zapisu tablicowego, przyjmując w zamian zapis pozycyjny jako bardziej ekonomiczny.

A oto przykłady zapisów kodowych, obok których podajemy słownie przykład elementu należącego do danej klasy operacji transportowania:

- | | |
|------------------|--|
| [00]0G[I00a00] | np. przekazywanie muzyki, |
| [pS]0G[ISZa00] | np. nadawanie informacji językowej telefoniczną siecią napowietrzną, |
| [ps]0C[I00m00] | np. nadawanie muzyki telefonicznym kablem podwodnym, |
| [pS]0S[ISZa00] | np. nadawanie informacji językowej telefonicznym kablem podziemnym, |
| [00]0G[ISZl00] | np. tzw. gazeta świetlna, |
| [00]kG[I00aUS] | np. samolotowy transport powietrzny nagrań muzycznych, |
| [00]kG[ISZlUS] | np. samolotowy transport książek, |
| [00]kC[I00aUS] | np. transport wodny (np. statkami) nagrań muzycznych, |
| [00]kC[I00lUS] | np. transport wodny statkami nut, |
| [pC]00[I00aUS] | np. transport hydrauliczny (bezpojemnikowy) nagrań muzycznych |
| [pG]kG[ISZlUS] | np. podziemny transport pneumatyczny pojemników z listami, |
| [00]kS/G[I00aUS] | np. transport samochodowy nagrań muzycznych, |
| [00]kS/G[I0ZlUS] | np. transport samochodowy nut, |
| [pS]kS/G[ISZlUS] | np. transport kolejowy (naziemny) książek, |
| [pS]kG[ISZlUS] | np. transport listów napowietrzną kolejką linową, |
| [pS]0S[IS0e00] | np. przewodzenie informacji w kodzie doznaniowym przez nerwy, |
| [00]0G[ET0] | np. naświetlanie, nagrzewanie, |
| [00]0S[ET0] | np. przenoszenie energii cieplnej przez przewodzenie w ciele stałym, |
| [00]0C[ET0] | np. przewodzenie energii cieplnej przez ciecze, |
| [00]kC[EEU] | np. przewóz statkami wodnymi akumulatorów elektrycznych, |
| [pS]0G[EE0] | np. przesyłanie energii elektrycznej kablem powietrznym, |
| [pS]0S[EE0] | np. przesyłanie energii elektrycznej kablem podziemnym, |
| [pS]0S/C[EE0] | np. przesyłanie energii elektrycznej kablem leżącym na dnie morskim, |
| [00]kC[EEU] | np. przenoszenie ładunku elektrycznego przez cząsteczki elektrolitu, |
| [pC]00[EM0] | np. przenoszenie energii mechanicznej metodami hydraulicznymi, |
| [pG]00[EM0] | np. przenoszenie energii mechanicznej metodami pneumatycznymi, |
| [pS]0C[EM0] | np. przenoszenie energii mechanicznej pasem transmisyjnym w ośrodku ciekłym, |

[00]0G[MG]	np. dyfuzja gazów, wentylacja bezkanałowa,
[00]0G[MC]	np. zraszanie, polewanie,
[00]0G[MS]	np. rozpylanie (substancji stałej)
[00]0C[MG]	np. gazowanie wody (np. nasycanie CO ₂ wody mineralnej),
[00]0C[MC]	np. dyfuzja cieczy,
[00]0S[MG]	np. nasycanie gazem ciał stałych,
[00]0S[MC]	np. nasycanie płynem ciała stałego, nasączenie,
[00]0C[MS]	np. mineralizacja wody,
[00]kG[MG]	np. transport pojemników z gazem drogą powietrzną,
[00]kG[MC]	np. transport pojemników z płynem drogą powietrzną,
[00]kG[MS]	np. transport ciał stałych samolotami,
[00]kC[MC]	np. przewóz ropy łodziami podwodnymi,
[00]kC/G[MC]	np. przewóz ropy tankowcami,
[00]kS[MS]	np. transport ciał stałych urządzeniami typu „kret”,
[pS]kG[MS]	np. transport ciał stałych kolejką linową,
[pC]kC/G[MS]	np. spławianie drzewa rzeką,
[pG]0S[MG]	np. transport gazu gazociągiem podziemnym,
[pS]0G[MS]	np. transport transporterowy (np. schodami ruchomymi ciał stałych),
[pG]kS[MS]	np. transport pocztą pneumatyczną,
[pS]kS/G[MS]	np. transport kolejowy.

Takie opisy sytuacji można w miarę potrzeby uściślać, wprowadzając dalsze charakterystyki. I tak, można by wprowadzić określenie źródła siły sprawczej ruchu, co pozwoliłoby, na przykład, uściślić charakterystykę spławiania jako transportowania przewodnicowego bezpojemnikowego substancji stałej na styku środowiska gazowego i ciekłego po powierzchni cieczy przy wykorzystaniu naturalnej formy energii. Koniunkcja tych warunków pozwala, jak się wydaje, wystarczająco scharakteryzować spławianie. Wprowadzenie kryterium istnienia kontaktu pomiędzy powierzchnią ruchu a poruszającym się ciałem pozwoliłoby natomiast na scharakteryzowanie transportu przy pomocy takich środków, jakimi są poduszki (co nie było możliwe przy dotychczas wprowadzanych charakterystykach), byłoby to bowiem transportowanie bezprzewodnicowe pojemnikowe na styku środowiska stałego i gazowego lub płynnego i gazowego z udziałem powierzchni stałej lub płynnej bez kontaktu pomiędzy powierzchnią a pojemnikiem.

Budowa takiego języka informacyjnego zakłada wybór jego jednostek semantycznych na podstawie porównywania zarówno znaczeń jednostek leksykalnych języka naturalnego oraz wyrażeń z nich złożonych, jak i sytuacji pozajęzykowych, gdyż tylko wtedy można będzie za jego pomocą opisywać tak jednostki znaczące języków naturalnych, jak i sytuacje pozajęzykowe, których opis w języku naturalnym byłby często nieekonomiczny lub niewystarczająco precyzyjny.

Kod semantyczny do opisu dyscyplin sportowych

Przedstawiony tu kod semantyczny do opisu dyscyplin sportowych był pierwszą próbą zbudowania języka typu kodu semantycznego, stanowiąc swoiste poletko

doświadczalne, pomocne przy ewentualnych dalszych pracach związanych z kodami semantycznymi. Poletko to jest wygodne, gdyż obejmuje dziedzinę niewielką, o dobrze określonych elementach.

Intencją naszą było spojrzenie na dyscypliny sportowe z punktu widzenia sprawy dla każdej z dyscyplin najistotniejszej, a więc tego, na czym polega zwycięstwo w danej konkurencji. Określenie tego, co jest istotne czy najistotniejsze, zawsze nasuwa wiele trudności, wydaje się jednak, że sam tekst pracy wyjaśni, o co chodzi. Tworząc opis, nie interesowaliśmy się różnymi szczegółami regulaminowymi odnoszącymi się do różnych dyscyplin, jak np. rozmiary boisk, długość jednego okrążenia bieżni lub toru, waga przedmiotów, którymi dokonuje się rzutów. Znajduje to zresztą potwierdzenie w samych regulaminach: rozmiary boiska do gry w piłkę nożną nie są ściśle określone, muszą jedynie mieścić się w ustalonym limicie; w pchnięciu kulą waga kuli jest inna dla mężczyzn, a inna dla kobiet itd., co jednak nie wpływa na podstawowe zasady rozgrywania danej konkurencji.

Jak już powiedzieliśmy, proponowany kod semantyczny na pierwszy rzut oka może się wydawać podobny do postulowanych we współczesnym językoznawstwie języków typu mnożników czy też znaczników semantycznych, różnica, jak się wydaje istotna, polega na tym, że o ile postulowane przez lingwistów elementarne jednostki semantyczne służą do opisu znaczeń wyrażen języków naturalnych, o tyle postulowany przez nas kod semantyczny ma przede wszystkim służyć do bezpośredniego opisywania rzeczywistości, z pominięciem opisującego ją jakiegoś języka naturalnego. Tak więc nie będziemy w naszej pracy zajmować się opisami znaczenia słów i wyrażen nazywających poszczególne dyscypliny sportowe, takich jak, na przykład, *boks*, *bieg*, *chód*, *skok*, *slalom*, *maraton*, *trójskok*, będziemy natomiast opisywać poszczególne dyscypliny sportowe, wydobywając ich cechy dystynktywne. Będziemy przy tym dążyć do ustalenia najmniejszej liczby cech, przy pomocy których można poklasyfikować wszystkie dyscypliny sportowe, ustalając dla każdej dyscypliny najmniejszy możliwy zbiór cech odróżniający ją od innych dyscyplin, nie zaś zbiór maksymalny, w pełni tę dyscyplinę charakteryzujący (nie będziemy, na przykład, w opisie boksu zajmować się typami ciosów, bo nie jest to dystynktywne dla innych elementów opisywanej przez nas rzeczywistości, nie będziemy też opisywać, z ilu i jakich figur składa się jazda obowiązkowa w jeździe figurowej na łyżwach, ani też jakie przeszkody ustawia się w biegu z przeszkodami), choć oczywiście możliwość podania takich informacji dodatkowych należy również przewidzieć.

Na marginesie przypominamy, że proponowany przez nas zbiór cech dystynktywnych jest zbiorem cech opisujących tylko takie typy zachowań człowieka, które nazywane są tradycyjnie dyscyplinami sportu. Przy rozszerzeniu opisywanej rzeczywistości trzeba by oczywiście rozszerzyć także zbiór opisujących ją cech.

Cechą charakteryzującą wszystkie dyscypliny sportowe jest ruch, rozumiany jako zmienianie się w czasie położenia ciała materialnego względem wyróżnionego układu odniesienia lub zmienianie się wzajemnego położenia części tego ciała. Cecha ta jest dystynktywną dla wszystkich dyscyplin sportowych, nie może być

więc uznana za dystynktywną, czyli odróżniającą w obrębie opisywanej przez nas rzeczywistości. Ruch mógłby być uznany za cechę dystynktywną tylko w wypadku, gdybyśmy do dyscyplin sportowych włączyli brydż (jeżeli przyjmiemy, że operacji odłączania elementów od zbioru nie uznamy za szczególną postać ruchu), brydż stanowiłby bowiem jedyną dyscyplinę sportową nieposiadającą tej cechy (ani szachy, ani warcaby warunku tego nie spełniają, polegają one bowiem na zmienianiu położenia jakichś elementów w jakichś przestrzeniach, nawet, jeśli i elementy te, i przestrzeń mają charakter abstrakcyjny), byłyby więc członem nienacechowanym opozycji prywatywnej, choć byłyby to opozycja izolowana.

Za cechę dystynktywną, charakteryzującą poszczególne dyscypliny sportowe, może być uznany rodzaj poruszającego się ciała, przy czym chodzi tu o to ciało, którego bądź ruch, bądź też położenie w przestrzeni w momencie zakończenia ruchu bierze się pod uwagę przy ustalaniu wyniku danej konkurencji. Mamy tu dwie możliwości:

1. Dana konkurencja polega na ruchu zawodnika (a),
2. Dana konkurencja polega na ruchu ciała innego niż zawodnik, przemieszczanego przez zawodnika (b).

W pierwszej grupie mamy takie konkurencje, jak: biegi, skoki, pływanie, skoki do wody, skoki spadochronowe, jazda szybka na lodzie, jazda figurowa na lodzie, konkurencje narciarskie, kolarstwo, wyścigi motocyklowe, samochodowe i konne, regaty wioślarskie i żeglarskie, szybownictwo i gimnastyka. Do grupy drugiej należą: rzuty, strzelanie z broni palnej i łuku, piłka nożna, ręczna i wodna, hokej na lodzie i na trawie, koszykówka, siatkówka, rugby, szermierka, zapasy, dżudo, podnoszenie ciężarów, tenis, tenis stołowy, golf itp.

Jak sądzę, przedstawiony tu podział wymaga pewnych dodatkowych wyjaśnień. Do grupy, w której relewantny jest tylko ruch zawodnika, zaliczono takie konkurencje jak kolarstwo, wyścigi konne, motocyklowe i samochodowe, wioślarstwo, żeglarstwo i szybownictwo, gdzie zawodnik porusza się wraz z jakimś przedmiotem wspomagającym (instrumentem): pojazdem lub zwierzęciem. W konkurencjach tych ważny jest jednak ruch zawodnika, nie zaś ruch instrumentu – kolarz w wyścigu szosowym lub przełajowym zostanie sklasyfikowany nawet wtedy, gdy do mety dobiegnie z rowerem w rękę, natomiast koń, który dobiegnie do mety, zrzuciwszy jeźdźca, nie zostanie sklasyfikowany.

W grupie drugiej znalazły się takie konkurencje jak rzuty, w których chodzi jedynie o to, żeby w przestrzeni przemieścił się jakiś przedmiot (np. kula, oszczep, dysk) – sam ruch zawodnika nie jest tu relewantny; tu zaliczono także strzelanie, gdzie relewantne jest przemieszczenie się pocisku lub strzały, a także piłkę nożną i siatkówkę, gdzie ważny jest tylko (liczy się do wyniku) ruch piłki, natomiast poruszający się zawodnicy mają jedynie ten ruch spowodować. W tej grupie znalazły się również takie dyscypliny jak szermierka i podnoszenie ciężarów – w pierwszym wypadku chodzi bowiem o przemieszczenie w przestrzeni (stanowi ją wyznaczona część ciała przeciwnika) broni, w drugim o przemieszczenie ciężaru. W wypadku zapasów i dżudo ciałem, którego położenie przestrzenne jest relewantne przy ustalaniu wyniku, jest nie sam zawodnik, lecz jego przeciwnik

– zawodnik, aby zdobyć punkt, musi w określony sposób przemieścić ciało swego przeciwnika (np. w wypadku zapasów położyć przeciwnika na łopatki).

Jak widać z powyższych przykładów, zmienna b może mieć różne wartości, np.: oszczep, młot, kula, dysk, pocisk, strzała, piłka, krążek, floret, szpada, szabla, sztanga, ciało przeciwnika.

Drugą cechą, która może być uznana za dystynktywną przy opisie dyscyplin sportowych, jest obecność lub nieobecność instrumentu wpływającego na ruch ciała. Mamy tu następujące możliwości:

1. Brak instrumentu wpływającego na ruch ciała (0),
2. Instrument wpływający na ruch porusza się wraz z poruszającym się ciałem (obecny przez cały czas trwania ruchu – (a),
3. Instrument wpływający na ruch obecny jest tylko w momencie początkowym ruchu, wprowadzając w ruch poruszające się ciało – (b).

A oto przykłady dyscyplin sportowych zaliczanych do wymienianych tu klas:

1. Brak instrumentu wpływającego na ruch ciała w takich konkurencjach jak: biegi, skoki wzwyż i w dal, trójskok, pływanie, skoki do wody z wieży, rzuty, piłka nożna, koszykówka, siatkówka, rugby, gimnastyka artystyczna, szermierka, boks, zapasy, podnoszenie ciężarów.

Oczywiście zawodnika wprowadzającego w ruch poruszające się ciało (np. w wypadku rzutów, piłki nożnej, szermierki) nie traktujemy tu jako instrumentu, choć i takie stanowisko byłoby możliwe – udział zawodnika w danej dyscyplinie sportowej nie może być uznany za cechę dystynktywną, bo jest to cecha przysługująca wszystkim elementom opisywanej przez nas rzeczywistości. Nie jest to zresztą konieczne z punktu widzenia ekonomii opisu, gdyż liczba przyjętych cech w zupełności wystarcza do wydzielenia tych dyscyplin w oddzielną klasę.

2. Instrument towarzyszący przemieszczającemu się ciału zawodnika (a) obecny jest w takich dyscyplinach sportu jak: spadochroniarstwo, konkurencje łyżwiarskie i narciarskie, kolarstwo, sporty motocyklowe i samochodowe, wioślarstwo i żeglarstwo, hippika i szybownictwo, saneczkarstwo, a także gimnastyka przyrządowa (z wyjątkiem skoku przez konia, co należy raczej do grupy b). Zmienna a może więc mieć następujące wartości: spadochron, łyżwy, narty, motocykl, rower, samochód, łódź wiosłowa lub żaglowa, sanki, bobsleje, szybowiec, koń.

3. Instrument obecny tylko w początkowym momencie ruchu (b) mamy w takich dyscyplinach sportu jak: skok o tyczce, skoki do wody z trampoliny, strzelanie z broni palnej i łuku, hokej na lodzie i na trawie, tenis, tenis stołowy, polo i golf. Pod zmienną b można więc podstawiać następujące wartości: tyczka, trampolina, bron palna, łuk, laski, kije, rakiety.

Za cechę dystynktywną może być także uznany przy opisie dyscyplin sportowych ośrodek ruchu, czyli stan skupienia przestrzeni materialnej, w której przemieszcza się poruszające się ciało. Relewantne są tu trzy środowiska:

1. Powierzchnia ciała stałego (a),
np. biegi, łyżwiarstwo, narciarstwo (z wyjątkiem skoków narciarskich), wyścigi samochodowe, kolarskie, motocyklowe i konne, bojery, piłka nożna,

hokej, gimnastyka artystyczna, zapasy, dżudo. Zmienna a może w tym wypadku przybierać różne wartości, np.: śnieg, lód, żużel i inne.

2. Woda lub jej powierzchnia (b),

np.: konkurencje pływackie, wioślarskie i żeglarskie, piłka wodna.

3. Powietrze (c),

np.: rzuty, skoki, strzelanie, szybownictwo, spadochroniarstwo, koszykówka, siatkówka, piłka ręczna, gimnastyka przyrządowa, szermierka, boks, tenis.

Dystynktywną cechą w opisie dyscyplin sportowych jest również stosunek ruchu do siły grawitacji ziemskiej. W tym wypadku mamy następujące możliwości:

1. Gdy grawitacja jest nerelewantna (0),

np.: biegi, rzuty, skok w dal, trójskok, pływanie, wyścigi różnych typów, konkurencje strzeleckie, biegi narciarskie płaskie, łyżwiarstwo, piłka nożna, tenis.

2. Gdy ruch odbywa się z wykorzystaniem siły grawitacji (a),

np.: skoki do wody, skoki narciarskie, skoki spadochronowe, slalomy narciarskie i kajakowe, biegi zjazdowe w narciarstwie, saneczki i bobsleje.

3. Gdy ruch odbywa się w kierunku przeciwnym do działania siły grawitacji (antygrawitacyjne) – (b),

np.: skok wzwyż, skok o tyczce, podnoszenie ciężarów.

Następną opozycją, którą należy uwzględnić przy opisie dyscyplin sportowych, jest relewantność lub nerelewantność kształtu toru, rozumianego tutaj zawsze jako linia, po której dane ciało się porusza. Opozycja ta jest prywatywna, może więc przybierać tylko następujące wartości:

1. Gdy kształt toru jest nerelewantny (0),

np.: w biegach płaskich, rzutach, pływaniu, jeździe szybkiej na lodzie, różnego typu wyścigach, w strzelaniu, w piłce nożnej, hokeju, siatkówce, tenisie, koszykówce, szermierce.

2. Gdy kształt toru jest relewantny (1), tzn. gdy:

a) decyduje on o wyniku konkurencji,

np.: skoki do wody, skoki narciarskie, jazda figurowa na łyżwach, gimnastyka przyrządowa, akrobacje lotnicze,

b) jest on ściśle wyznaczony, tzn. gdy przemieszczenie się ciała w przestrzeni po innym niż wyznaczony torze powoduje wykluczenie zawodnika z konkurencji lub niezaliczenie boju,

np.: biegi przez płotki, biegi z przeszkodami, slalomy narciarskie i kajakowe, niektóre konkurencje hipiczne (parcours), podnoszenie ciężarów.

Każde poruszające się ciało przebywa w odcinku czasu $t_1 \dots t_n$ (czasie trwania ruchu) pewien dystans. Dystans ten może być w poszczególnych dyscyplinach sportowych relewantny (1) lub nerelewantny (0), wszystkie konkurencje sportowe rozpadają się więc na dwie klasy. Nim jednak podamy przykłady elementów tych klas, jeszcze kilka słów wyjaśnienia, jak należy tu rozumieć dystans, nie jest on bowiem równy długości drogi przebytej w czasie $t_1 \dots t_n$ przez poruszające się ciało. Mówiąc najogólniej, przez dystans (nazywany także w niektórych dyscyplinach sportu, np. w rzutach, odległością lub w skokach wysokością) rozumie się w sporcie:

1. Albo długość prostej łączącej rzut punktu początkowego (lub punktu umownie przyjętego za początkowy) ruchu i punktu końcowego na płaszczyznę – tak jest na przykład we wszystkich konkurencjach rzutowych i strzeleckich, w skoku w dal i w trójskoku, a także w biegach narciarskich i biegach przez płotki, choć w konkurencjach tych ciało porusza się po łuku łączącym rzeczywisty punkt początkowy ruchu i punkt końcowy (lub też części przebytej drogi stanowią takie łuki, jak w biegu przez płotki i w trójskoku) – droga przebyta przez poruszające się ciało jest więc w rzeczywistości dłuższa niż dystans decydujący o wyniku konkurencji.

2. Albo różnicę odległości pomiędzy najniższym położonym punktem, w którym znalazło się poruszające się ciało, a z góry wyznaczonym punktem, powyżej którego musiało znaleźć się poruszające się ciało – tak jest na przykład w skoku wzwyż i w skoku o tyczce, w których to wypadkach faktycznie przebyta droga jest co najmniej dwa razy dłuższa od dystansu decydującego o wyniku konkurencji (a w rzeczywistości jest o wiele dłuższa).

3. Albo długość wyznaczonego toru łączącego punkt początkowy i punkt docelowy ruchu – tak jest na przykład w biegach płaskich i biegach z przeszkodami oraz w innego typu wyścigach, w których to wypadkach droga przebyta przez ciało może być w rzeczywistości dłuższa od dystansu.

4. Albo długość uśrednionej krzywej łączącej punkt początkowy i punkt docelowy ruchu leżącej na wyznaczonym torze, po którym porusza się dane ciało – tak jest na przykład w slalomie narciarskim i kajakowym, gdzie długość drogi rzeczywiście przebytej przez ciało jest zawsze większa niż dystans określony dla danej konkurencji.

A oto przykłady. Dystans nierelevantny mamy na przykład w biegach i różnych wyścigach na dystans, a także w strzelaniu – tu wszędzie dystans jest nierelevantny, choć z góry wyznaczony. Nierelevantny jest także dystans w piłce nożnej, hokeju, koszykówce, siatkówce, tenisie, szermierce, zapasach itp., choć oczywiście nie jest on tu ustalony. Relevantny dystans mamy natomiast w takich konkurencjach jak: biegi i wyścigi na czas, rzuty, skoki w dal, wzwyż, o tyczce i trójskok oraz skoki narciarskie. W tych konkurencjach dystans decyduje o wyniku.

Bywa i tak, że w jakiejś konkurencji nieważne jest ani to, po jakim torze porusza się ciało, ani to, jaki dystans przebędzie, lecz ważne jest to, żeby w wyniku ruchu znalazło się ono w wyznaczonym z góry miejscu przestrzeni, w której się porusza. Tak jest na przykład w piłce nożnej i koszykówce, gdzie piłka w wyniku ruchu powinna znaleźć się w wycinku przestrzeni wyznaczonym przez bramkę lub kosz, tak jest w szermierce i boksie, gdzie chodzi o ułożenie poruszającego się ciała (broń szermierza lub pięść zawodnika) w określonej przestrzeni ciała przeciwnika, tak jest w tenisie i siatkówce, gdzie liczy się dotknięcie przez poruszające się ciało jakiegoś punktu przestrzeni boiska (pamiętajmy, że ośrodkiem ruchu jest w tych dyscyplinach powietrze), tak jest wreszcie w dźudo i zapasach, gdzie chodzi o to, aby określone punkty poruszającego się ciała znalazły się w wyniku ruchu w danym miejscu przestrzeni (w zapasach przeciwnik musi być położony na łopatkach).

Choć w dwóch ostatnich przykładach chodzi wprawdzie nie o osiągnięcie w wyniku ruchu określonego punktu przestrzeni, lecz o to, aby w wyniku ruchu poruszające się ciało zostało odpowiednio w tej przestrzeni usytuowane (zetrnęło się z nią wyznaczonymi punktami), to jednak z punktu widzenia ekonomiczności opisu nie ma potrzeby wydzielenia tu dwóch odrębnych kategorii, nie ma bowiem takich dyscyplin sportowych, w których w wyniku ruchu poruszający się przedmiot ma dotknąć ściśle wyznaczonego punktu powierzchni ruchu ściśle wyznaczonymi punktami swej powierzchni. Jest natomiast potrzeba wydzielenia tych konkurencji, gdzie o wyniku decyduje nie znalezienie się poruszającego się ciała w określonym wycinku przestrzeni, lecz to, w jakiej odległości znajdzie się ono od jednego punktu tej przestrzeni wyznaczonego jako punkt docelowy ruchu (im mniejsza ta odległość, tym lepiej; w szczególnym wypadku może to być odległość zerowa). Tak jest na przykład w skokach spadochronowych oraz w strzelaniu do tarczy z broni kulowej i z łuku; inaczej jest w strzelaniu do sylwetek i do rzutków oraz do krążków (w biathlonie narciarskim), gdzie możliwe jest tylko trafienie lub chybienie. Dlatego też, ze względu na relewantność punktu docelowego ruchu, wszystkie dyscypliny sportowe należy podzielić na trzy klasy:

1. Gdzie punkt docelowy ruchu jest nerelewantny (0),
2. Gdzie ważne jest osiągnięcie wyznaczonego miejsca przestrzeni ruchu (a),
3. Gdzie ważna jest odległość od wyznaczonego punktu docelowego (b).

Na marginesie należy zauważyć, że w opisywanej rzeczywistości nie istnieje takie połączenie cech dystynktywnych, żeby relewantny był i dystans, i punkt docelowy ruchu (choć oczywiście obie te cechy mogą być nerelewantne).

Ostatnią cechą uznaną przez nas za dystynktywną przy opisie dyscyplin sportowych jest czas trwania ruchu. Mamy tu dwie możliwości:

1. Czas trwania ruchu jest nerelewantny (0),
np.: biegi na czas, rzuty, skoki, strzelanie, tzw. gry zespołowe, zapasy, dżudo, boks, szermierka, podnoszenie ciężarów
(choć oczywiście może to być czas z góry ustalony lub limitowany).
2. Czas trwania ruchu jest relewantny (1),
np.: różnego rodzaju wyścigi
(we wszystkich tych wypadkach czas trwania ruchu decyduje o wyniku konkurencji).

Tu również należy zauważyć, że tam, gdzie relewantny jest czas trwania ruchu, nie jest relewantny dystans (i odwrotnie), co w połączeniu z wykluczeniem jednoczesnej relewantności dystansu i punktu docelowego ruchu oznacza, że nie ma takiej konkurencji, w której punktuje się, kto szybciej i krótszą drogą osiągnie docelowy punkt ruchu.

Wydaje się, że zaproponowanych wyżej osiem opozycji dystynktywnych wystarcza do opisanego dyscyplin sportowych, choć oczywiście przyjęta tu metoda opisu pozwala w razie potrzeby wprowadzić różnego typu informacje dodatkowe. Taką informacją dodatkową mogłaby być na przykład informacja o wadze poruszającego się ciała, która jest dotąd relewantna tylko w podnoszeniu ciężarów, stanowi zatem opozycję izolowaną i jako takiej nie warto jej wprowadzać

do systemu, tym bardziej, że jak się wydaje, zaproponowane tu cechy wystarczają do jednoznacznego scharakteryzowania tej konkurencji.

Na zakończenie podajemy tabelę obrazującą wartości członów opozycji dystynktywnych dla kilkudziesięciu wybranych przykładowo dyscyplin sportowych.

Matryca cech dystynktywnych dyscyplin sportowych

Lp.	Nazwa dyscypliny	I	II	III	IV	V	VI	VII	VIII
1.	Akrobacje lotnicze	a	a samolot	c	0	1*	0	0	0
2.	Biegi na czas	a	0	a	0	0	1*	0	0
3.	Biegi narciarskie	a	a narty	a śnieg	0	0	0	0	1*
4.	Biegi płaskie	a	0	a	0	0	0	0	1*
5.	Biegi przez płotki	a	0	a	b	1	0	0	1*
6.	Biegi z przeszkodami	a	0	a	b	1	0	0	1*
7.	Bobsleje	a	a bobslej	a śnieg	a	0	0	0	1*
8.	Bojery	a	a bojer	a lód	0	0	0	0	1*
9.	Boks	a	0	c	0	1*	0	a*	0
10.	Chód	a	0	a	0	0	0	0	1*
11.	Dzudo	b przeciwnik	0	a	0	1*	0	a*	0
12.	Gimnastyka artystyczna	a	0	a	0	1*	0	0	0
13.	Gimnastyka przyrządowa	a	a przyrz. gimn.	c	0	1*	0	0	0
14.	Golf	b piłka	b kij	a	0	0	0	a*	0
15.	Hokej na lodzie	b krążek	b kij	a lód	0	0	0	a*	0
16.	Hokej na trawie	b piłka	b łaska	a	0	0	0	a*	0
17.	Jazda figurowa na lodzie	a	a łyżwy	a lód	0	1*	0	0	0
18.	Jazda szybka na lodzie	a	a łyżwy	a lód	0	0	0	0	1*
19.	Koszykówka	b piłka	0	c	0	0	0	a*	0
20.	Maraton	a	0	a	0	0	0	0	1*
21.	Narciarstwo zjazdowe	a	a narty	a śnieg	a	0	0	0	1*
22.	Parcours	a	a koń	a	b	1*	0	0	1*
23.	Pchnięcie kulą	b kula	0	c	0	0	1*	0	0
24.	Piłka nożna	b piłka	0	a	0	0	0	a*	0
25.	Piłka ręczna	b piłka	0	c	0	0	0	a*	0
26.	Piłka wodna	b piłka	0	b	0	0	0	a*	0
27.	Pływanie	a	0	b	0	0	0	0	1*
28.	Podnoszenie ciężarów	b sztanga	0	c	b	1	0	a*	0
29.	Polo	b piłka	b kij	a	0	0	0	a*	0
30.	Rugby	b piłka	0	a	0	0	0	a*	0
31.	Rzut dyskiem	b dysk	0	c	0	0	1*	0	0
32.	Rzut młotem	b młot	0	c	0	0	1*	0	0
33.	Rzut oszczepem	b oszczep	0	c	0	0	1*	0	0
34.	Saneczkarstwo	a	a sanki	a śnieg	a	0	0	0	1*
35.	Siatkówka	b piłka	0	c	0	0	0	a*	0
36.	Skoki do wody z trampoliny	a	b trampolina	c	a	1*	0	0	0
37.	Skoki do wody z wieży	a	0	a	a	1*	0	0	0

Lp.	Nazwa dyscypliny	I	II	III	IV	V	VI	VII	VIII
38.	Skoki narciarskie	a	a narty	c	a	1*	1*	0	0
39.	Skoki spadochronowe	a	a spadochron	c	a	0	0	b*	0
40.	Skok o tyczce	a	b tyczka	c	b	0	1*	0	0
41.	Skok w dal	a	0	c	0	0	1*	0	0
42.	Skok wzwyż	a	0	c	b	0	1*	0	0
43.	Sialomy kajakowe	a	a kajak	b	a	1	0	0	1*
44.	Sialomy narciarskie	a	a narty	a śnieg	a	1	0	0	1*
45.	Strzelanie z broni palnej do rzutków	b pocisk	b broń	c	0	0	0	a*	0
46.	Strzelanie z broni palnej do sylwetek	b kula	b broń	c	0	0	0	a*	0
47.	Strzelanie z broni palnej do tarczy	b kula	b broń	c	0	0	0	b*	0
48.	Strzelanie z łuku	b strzała	b łuk	c	0	0	0	b*	0
49.	Szermierka	b broń	0	c	0	0	0	a*	0
50.	Szybownictwo ¹	a	a szybowiec	c	0	0	0	b*	
51.	Tenis	b piłka	b rakietka	c	0	0	0	a*	0
52.	Tenis stołowy	b piłka	b rakietka	c	0	0	0	a*	0
53.	Trójskok	a	0	c	0	1	1*	0	0
54.	Wyścigi kolarskie	a	a rower	a	0	0	0	0	1*
55.	Wyścigi kolarskie na czas	a	a rower	a	0	0	1*	0	0
56.	Wyścigi konne	a	a koń	a	0	0	0	0	1*
57.	Wyścigi motocyklowe	a	a motocykl	a	0	0	0	0	1*
58.	Wyścigi motocyklowe na żużlu	a	a motocykl	a żużel	0	0	0	0	1*
59.	Wyścigi samochodowe	a	a samochód	a	0	0	0	0	1*
60.	Wyścigi wioślarskie	a	a łódź	b	0	0	0	0	1*
61.	Wyścigi żeglarskie	a	a łódź żagl.	b	0	0	0	0	1
62.	Żapasy	b	0	a	0	1	0	a	0

¹ Szybownictwo należy rozbić na poszczególne konkurencje w zależności od sposobu ich rozgrywania.

W tabeli przyjęto następującą kolejność opozycji:

I. Rodzaj poruszającego się ciała:

- a – zawodnik,
- b – niezawodnik.

II. Instrument ruchu:

- 0 – brak instrumentu wpływającego na ruch ciała,
- a – instrument obecny przez cały czas trwania ruchu,
- b – instrument obecny tylko w momencie początkowym ruchu.

III. Ośrodek ruchu:

- a – powierzchnia ciała stałego,
- b – woda,
- c – powietrze.

IV. Grawitacja:

- 0 – nierelevantna,
- a – ruch odbywa się z wykorzystaniem siły grawitacji,
- b – ruch odbywa się w kierunku przeciwnym do działania siły grawitacji.

- V. Kształt toru ruchu ciała:
 - 0 – nierelevantny,
 - 1 – relevantny.
- VI. Dystans:
 - 0 – nierelevantny,
 - 1 – relevantny.
- VII. Punkt docelowy ruchu:
 - 0 – nierelevantny,
 - a – wyznaczone miejsce przestrzeni,
 - b – odległość od wyznaczonego punktu docelowego.
- VIII. Czas trwania ruchu:
 - 0 – nierelevantny,
 - 1 – relevantny.

Gwiazdką oznaczono te cechy, które brane są pod uwagę w ustalaniu wyniku danej konkurencji (brak jeszcze jednej gwiazdki przy podnoszeniu ciężarów, gdzie o wyniku decyduje również waga poruszającego się ciała).

Jeszcze raz należy podkreślić, że wybrane cechy dystynktywne należy uważać jedynie za ilustrację proponowanej metody opisu, być może sam opis w szczegółach powinien jeszcze ulec zmianie.

Literatura

- Apresjan J.D.: *Koncepcje i metody współczesnej lingwistyki strukturalnej*. Warszawa 1971.
- Apresjan J.D.: *Definiowanie znaczeń leksykalnych jako zagadnienie semantyki teoretycznej*, w: Wierzbicka A. (red.): *Semantyka i słownik*. Wrocław 1972.
- Apresjan J.D.: *Semantyka leksykalna. Synonimiczne środki języka*. Wrocław 1980.
- Apresjan J.D.: *O strukturze znaczeń jazykowych edeníc*, w: Dobrzyńska T., Janus E. (red.): *Tekst i zdanie*. Wrocław 1983.
- Apresjan J.D.: Mielczuk I.A., Żółkowski A.K.: *Próba objaśniająco-kombinatorycznego słownika języka rosyjskiego*, w: Wierzbicka A. (red.): *Semantyka i słownik*. Wrocław 1982.
- Bojar B.: *Opis semantyczny czasowników ruchu oraz pojęć związanych z ruchem*. Warszawa 1979.
- Hjelmslev L.: *Prolegomena teorii języka*, w: „Acta Linguistica”, t. 6, nr 2/3.
- Leibniz G.W.: *Tabulae Definitionum*, w: Janus E. (red.): *Słownik i semantyka*. Wrocław 1975.
- Mielczuk I.A.: *O określaniu większej/mniejszej złożoności semantycznej wyrazów przy relacjach słowotwórczych*, w: Wierzbicka A. (red.): *Semantyka i słownik*. Wrocław 1972.
- Mielczuk I.A.: *Struktura znaków językowych i zachodzące między nimi relacje formalno-znaczeniowe*, w: Wierzbicka A. (red.): *Semantyka i słownik*. Wrocław 1972.
- Sosińska B.: *Uniwersalny Kod Semantyczny W.W. Martynowa*. „Zagadnienia Informatyki Naukowej” 1981, nr 1.

- Wierzbicka A.: *W poszukiwaniu semantycznego modelu czasu i przestrzeni*, w: Wierzbicka A. (red.): *Semantyka i słownik*. Wrocław 1972.
- Wierzbicka A.: *Dociekania semantyczne*. Wrocław 1969.
- Wierzbicka A.: *Kocha, lubi, szanuje... Medytacje semantyczne*. Warszawa 1967.
- Wierzbicka A.: *Rozważania o częściach ciała*, w: Janus E. (red.): *Słownik i semantyka*. Warszawa 1975.
- Wierzbicka A.: *Język, umysł, kultura*. Warszawa 1999.
- Wojtasiewicz O.A.: *Sformalizowana interpretacja semantyczna czasowników*. Studia semiotyczne, t. 6. Wrocław 1975.
- Wójcik T.: *Prakseosemiotyka. Zarys teorii optymalnego znaku*. Warszawa 1969.

XV

System semantyczny języków informacyjno-wyszukiwawczych. Znaczenie. Prezentacja znaczenia

Języki informacyjno-wyszukiwawcze są językami sztucznymi przeznaczonymi do odwzorowywania informacji zawartej w dokumentach i wyszukiwania relewantnych dla użytkowników informacji w zbiorach informacyjnych systemów informacyjno-wyszukiwawczych.

Każdy system informacyjno-wyszukiwawczy ma swój zakres – granice tego zakresu wyznaczają potencjalny zbiór informacji, jakie system powinien gromadzić. Ten zbiór informacji to **pole informacyjne systemu**. Zakres systemu wydziela z obszaru wszystkich informacji tylko te, które są relewantne dla użytkowników danego systemu. Pole informacyjne systemu nie jest równe zakresowi informacji faktycznie znajdujących się w dokumentach gromadzonych w systemie, w każdym systemie znajdują się bowiem informacje nierlewantne dla użytkowników systemu, które towarzyszą w danym przekazie informacyjnym informacjom relewantnym.

Tak określone pole informacyjne systemu należy rozumieć jako zbiór informacji w sensie dystrybutywnym, abstrakcyjnym, wyznacza je bowiem zbiór cech informacji relewantnych dla użytkownika. Konkretny zbiór informacyjny jest praktyczną realizacją takiego zbioru idealnego i taki zbiór informacyjny jest zbiorem rozumianym w sensie kolektywnym, jako zespół informacji zgromadzonych w danym systemie.

W tradycyjnych systemach informacyjno-wyszukiwawczych były gromadzone (księgozbiór) lub opisywane (systemy informacji dokumentacyjnej) dokumenty zawierające relewantne dla użytkownika informacje. Dokumenty te stanowiły konkretny (materialny) **zbiór dokumentacyjny** i konkretny **zbiór wyszukiwawczy** danego systemu. Zbiór taki był materialną realizacją potencjalnego (abstrakcyjnego) zbioru dokumentacyjnego lub wyszukiwawczego systemu, czyli jego pola dokumentacyjnego, wyznaczanego przez relewantne dla użytkownika cechy dokumentów, przy czym cechy te mogą być zarówno cechami immanentnymi, dotyczącymi fizycznej postaci nośnika (materialnymi) lub zawartych o nich informacji, jak i cechami uwarunkowanymi relacjami zewnętrznymi (autor, wydawca, miejsce wydania), czyli, mówiąc inaczej, cechami ujmowanymi w **opisie formalnym** i w **opisie rzeczowym dokumentu**.

Zbiór relewantnych dla użytkownika systemu informacyjno-wyszukiwawczego cech informacyjnych – wyznaczających pole informacyjne systemu, lub cech dokumentów – wyznaczających pole dokumentacyjne systemu, reprezentowany jest (i opisywany) poprzez wyrażenia języka informacyjno-wyszukiwawczego danego systemu. Język informacyjno-wyszukiwawczy jest więc zbiorem odwzorowującym zarówno potencjalny zbiór informacyjny – pole informacyjne, jak i potencjalny zbiór dokumentacyjny systemu – pole dokumentacyjne, a więc zarazem cechy formalne informacji i dokumentów, jak i ich cechy semantyczne (treściowe). Czyni to **język opisu formalnego i język opisu rzeczowego**, dwa podsystemy języka informacyjno-wyszukiwawczego. Język informacyjno-wyszukiwawczy odwzorowuje więc nie tylko treść i cechy formalne konkretnych dokumentów (i informacji), ale także bezpośrednio odwzorowuje (i tym samym wyznacza) pole informacyjne i pole dokumentacyjne systemu informacyjnego. **Funkcja semantyczna języka informacyjno-wyszukiwawczego, a więc znaczenie wyrażen języka informacyjno-wyszukiwawczego, ich denotacja i konotacja, określona jest na tych właśnie polach.**

Znaczenie

Tak więc **denotacją** wyrażenia W języka informacyjno-wyszukiwawczego J jest zbiór dokumentów D, którym, zgodnie z regułami tego języka, można to wyrażenie przyporządkować. Na przykład, denotacją wyrażenia języka słów kluczowych ZIEMNIAK jest zbiór dokumentów, którym, zgodnie z regułami tego języka, można to wyrażenie przyporządkować, a nie zbiór ziemniaków, jak w wypadku równokształtnego wyrażenia *ziemniak* w języku naturalnym. Denotacją wyrażenia UKD 629.735.33 (według tablic UKD UDC-P022, Warszawa : BN 1997) jest zbiór dokumentów, którym to wyrażenie można przyporządkować – będzie to zbiór dokumentów mówiących o samolotach.

Tak określana strefa rzeczywistości reprezentowanej przez wyrażenia języka informacyjno-wyszukiwawczego wyznacza jednocześnie interpretację ich konotacji, przesądzając, że stanowią ją pewne zbiory cech ich denotatów, a więc zbiory własności dokumentów. Tak więc **konotacją** wyrażenia języka słów kluczowych ZIEMNIAK jest własność dokumentów „być o ziemniaku (mieć za przedmiot treści ziemniak)”, konotacją wyrażenia SAMOLOTY jest własność dokumentów „być o samolotach (mieć za przedmiot treści samoloty)”. Liczebność zbioru cech składających się na konotację wyrażen języków informacyjno-wyszukiwawczych uwarunkowana jest paradygmatyką tych języków. I tak, na przykład, konotacja tematu „Samoloty” będącego wyrażeniem języka haseł przedmiotowych Biblioteki Narodowej składa się z cech elementarnych, uporządkowanych hierarchicznie: „być o maszynach”, „być o statkach powietrznych”, „być o samolotach” oraz ma pewne cechy wspólne z konotacją wyrażenia „Pionowzloty”, co wynika z artykułu przedmiotowego ze słownika JHP JHP BN:

Samoloty
TS Maszyny

Statki powietrzne
TK Pionowzloty
TW Samoloty pasażerskie
Samoloty sportowe
Samoloty wojskowe
Wodnosamoloty
Samoloty – budowa i konstrukcje
TK Modelarstwo
Śmigło

Natomiast na konotację symbolu głównego UKD 629.7 mającego odpowiednik słowny „Technika lotnicza i astronautyczna” składają się, zgodnie z postulatami znaczeniowymi tego języka, cztery hierarchicznie uporządkowane własności: „być o naukach stosowanych”, „być o naukach technicznych”, „być o technice środków transportu” oraz „być o statkach powietrznych”.

Tak wyznaczona denotacja i konotacja wyrażeń języków informacyjno-wyszukiwawczych nazywana bywa (zob. B. Sosińska: *Relacje między planem treści i planem wyrażania w językach informacyjno-wyszukiwawczych*) **denotacją bezpośrednią i konotacją bezpośrednią**, określana jest ona bowiem na elementach rzeczywistości bezpośrednio odwzorowywanej przez te języki – rzeczywistością tą jest rzeczywistość złożona z dokumentów (**pole dokumentacyjne systemu**).

Denotacji bezpośredniej i konotacji bezpośredniej przeciwstawić można **denotację pośrednią i konotację pośrednią**, przyjmując, że użytkownik systemu informacyjnego chce najczęściej nie tylko zapoznać się z informacjami na interesujący go temat zawartymi w dokumentach zgromadzonych w danym systemie, lecz korzystając z tych informacji, a więc za ich pośrednictwem, poznać lepiej pewien interesujący go obiekt **rzeczywistości pozadokumentacyjnej**. Przy takim założeniu należałoby przyjąć, że obiekty te tworzą rzeczywistość, na której także można określić denotację i konotację wyrażeń języków informacyjno-wyszukiwawczych. I tak, na przykład, zakłada się, że jeśli użytkownik systemu informacyjnego zwraca się z pytaniem, które wyrazić można w postaci ciągu słów kluczowych:

Produkcja. Samoloty. Rolnictwo. Polska,

to w rzeczywistości interesuje go produkcja samolotów rolniczych w Polsce, a nie to, co na ten temat zostało napisane (choć oczywiście nie można wykluczyć z góry i tej drugiej możliwości, gdyż na przykład użytkownika zamierzającego napisać rozprawę doktorską na temat produkcji samolotów rolniczych w Polsce interesować będzie nie tylko sama produkcja, ale przede wszystkim właśnie to, co na ten temat napisano). Jeśli tak jest, to przytoczone tu wyrażenie języka informacyjno-wyszukiwawczego desygnuje takie obiekty jak: produkcja, samoloty, rolnictwo, Polska i one stanowią denotację wyrażeń składowych, zaś cechy tych obiektów są cechami konotacyjnymi.

O konieczności wprowadzenia w teorii języków informacyjno-wyszukiwawczych powyższych rozróżnień Barbara Sosińska tak pisze (w rozprawie doktor-

skiej pt.: *Relacje między planem treści i planem wyrażania w językach informacyjno-wyszukiwawczych*):

„Rozróżnienie bezpośrednich i pośrednich desygnatów wyrażen JIW oraz implikowane przez nie rozróżnienie bezpośrednich i pośrednich denotacji i konotacji tych wyrażen pozwala wyjaśnić zasady konstruowania systemu leksykalnego JIW oraz ustalenia i interpretowania jego reguł syntaktycznych. Zwykle przyjmuje się, że znaczenie wyrażen złożonych JIW należy interpretować teoriomnogościowo, tj. traktując wyrażenia proste jako nazwy pewnych zbiorów, a łączące je relatory (występujące *implicite* lub *explicite*) jako wykładniki operacji na zbiorach. Założenie to daje się jednak utrzymać tylko wówczas, gdy wyrażenia JIW odnosi się do dokumentów, traci natomiast sens wtedy, gdy za desygnaty tych wyrażen przyjmuje się obiekty czy aspekty treści. Np: sensowna jest teoriomnogościowa interpretacja wyrażenia „Produkcja. Samoloty”, gdy uznamy, że desygnatem wyrażenia prostego „Produkcja” jest dokument dotyczący produkcji oraz że desygnatem wyrażenia prostego „Samoloty” jest dokument dotyczący samolotów, a interpretacja połączenia tych dwóch wyrażen odpowiada wynikowi operacji iloczynu teoriomnogościowego, a więc wyrażenie złożone wskazuje taki dokument, który równocześnie dotyczy produkcji i samolotów, czyli jest dokumentem o produkcji samolotów. Jeśli jednak przyjmiemy, że desygnatami wyrażen prostych są dla podanego przykładu obiekty pozadokumentacyjne oznaczone w JIW tymi wyrażeniami, to teoriomnogościowa interpretacja wyrażenia złożonego prowadzi do fałszywego wniosku jakoby miało ono wskazywać taki obiekt pozadokumentacyjny, który jest równocześnie produkcją i samolotami.

O ile interpretacja związków syntaktycznych większości języków informacyjno-wyszukiwawczych odwołuje się do desygnatów bezpośrednich, o tyle reguły budowy systemu leksykalnego oraz konstrukcji wyrażen o strukturze semantycznie motywowanej w wielu JIW odwołują się do desygnatów pośrednich. Np. wyróżniana często – zwłaszcza w JIW o polu semantycznym odpowiadającym takim dziedzinom jak anatomia, geografia czy konstrukcja maszyn i urządzeń – relacja mereologiczna ma prawo łączyć jednostki leksykalne JIW tylko wtedy, gdy oznaczają one obiekty pozadokumentacyjne, nie zaś dokumenty, gdyż poprawne jest stwierdzenie, że obiekt treści desygnowany np. przez wyrażenie „Kręgosłup” jest częścią pewnej całości będącej obiektem treści desygnowanym przez wyrażenie „Kościec”, lecz niepoprawne byłoby stwierdzenie, że każdy dokument desygnowany przez pierwsze z tych wyrażen jest częścią każdego dokumentu desygnowanego przez drugie z nich.”

Relacje paradygmatyczne

Wyrażenia języka informacyjno-wyszukiwawczego denotują odpowiednie podpola informacyjne, przy czym w wypadku języków o strukturze hierarchicznej taki podział pola informacyjnego może być „wielowarstwowy” (wielokrotny). Podział ten wyznacza paradygmatykę systemu semantycznego – **relacje paradygmatyczne** zachodzące między wyrażeniami języków informacyjno-wyszukiwawczych omówimy poniżej.

Choć obiegowy sąd głosi, że języki informacyjno-wyszukiwawcze wolne są od synonimii (i na tym ma polegać między innymi ich przewaga nad językami naturalnymi), to jednak nie jest to sąd prawdziwy. Synonimy występujące w językach informacyjno-wyszukiwawczych noszą nazwę **synonimów wyszukiwawczych**, a relację synonimii występującą w tych językach zdefiniować można jako relację między wyrażeniami mającymi identyczną denotację bezpośrednią i konotację bezpośrednią oraz identyczną denotację pośrednią i konotację pośrednią. Relacja ta występować może na poziomie jednostek leksykalnych, na przykład w języku słów kluczowych synonimami wyszukiwawczymi mogą być wyrażenia: Językoznawstwo – Lingwistyka, Tranzystor – Trioda półprzewodnikowa, ale tylko wtedy, gdy jest to w słowniku tego języka wyraźnie zaznaczone. Synonimami w językach informacyjno-wyszukiwawczych mogą być również wyrażenia, które w języku naturalnym nie są synonimami, np. ANTYKLERYKALIZM – KLERYKALIZM, WATYKAN – PAPIESTWO, ATMOSFERA – POWIETRZE, RELIGIA – ATEIZM.

Częstym zjawiskiem jest występowanie **synonimii wyszukiwawczej na poziomie syntaktycznym** języka. Może to być na przykład synonimia między wyrażeniem prostym i wyrażeniem złożonym, np.:

Choroby układu krążenia (wyrażenie proste) i Układ krążenia. Choroby (wyrażenie złożone).

Przyczyną takiej synonimii są błędy popełnione przy projektowaniu języka i w dobrze zaprojektowanym języku synonimy takie nie powinny istnieć. Może to być również synonimia między wyrażeniami złożonymi, na przykład w UKD między symbolem złożonym a symbolem rozwiniętym:

338.4:669 Ekonomia przemysłu metalurgicznego –
669.003 Przemysł metalurgiczny z punktu widzenia ekonomicznego,

343.000.141.4 Prawo karne z ateistycznego punktu widzenia –
343:141.4 Prawo karne a ateizm,

327.36/37:355 Ruchy pacyfistyczne a wojskowość –
327.36/37.000.355 Ruchy pacyfistyczne z punktu widzenia wojskowości.

Systemowy charakter ma w UKD synonimia będąca wynikiem transformacji wyrażeń złożonych za pomocą relatora „+” oraz większości symboli złożonych z relatorem „:” w wyrażeniach o zmienionym szyku wyrażeń prostych wchodzących w skład takiego wyrażenia. Jest to zasada tzw. inwersji symboli (odwracalności), która powoduje, że to samo znaczenie można wyrazić w UKD za pomocą n! symboli złożonych (gdzie n oznacza liczbę symboli prostych lub rozwiniętych połączonych za pomocą jednego z tych relatorów), np.:

616.24-007.5-092:622.2:622.86/.87 Rozedma płuc. Patogeneza. Roboty górnicze. Zagrożenie zdrowia górników,
622.86/.87:622.2:616.24-007.5-092 Zagrożenie zdrowia górników. Roboty górnicze. Rozedma płuc. Patogeneza,

622.2:622.86/.87:616.24-007.5-092 Roboty górnicze. Zagrożenie zdrowia górników. Rozedma płuc. Patogeneza,

gdzie wszystkie te charakterystyki opisują dokument o przyczynach powstawania rozedmy płuc wskutek prowadzenia robót górniczych,
lub:

656.2 (492+493) Transport kolejowy Holandii i Belgii –

656.2 (493+492) Transport kolejowy Belgii i Holandii.

Przy wyrażeniach złożonych za pomocą dwukropka i plusa możliwa jest poza tym jeszcze jedna transformacja, na tzw. zapis pionowy z opuszczeniem wskaźnika relacji.

Synonimia występuje również – z założenia – we wszystkich tych językach informacyjno-wyszukiwawczych, które służą do indeksowania współrzędnego i w których porządek jednostek leksykalnych w charakterystykach wyszukiwawczych jest dowolny – synonimiczne są w nich bowiem wszystkie charakterystyki, w których w różnym porządku występują te same jednostki leksykalne, np.:

EKSPORT. KUWEJT. ROPA NAFTOWA;

KUWEJT. ROPA NAFTOWA. EKSPORT;

KUWEJT. EKSPORT. ROPA NAFTOWA;

EKSPORT. ROPA NAFTOWA. KUWEJT;

ROPA NAFTOWA. EKSPORT. KUWEJT;

ROPA NAFTOWA. KUWEJT. EKSPORT.

Synonimia wcale nie jest więc w językach informacyjno-wyszukiwawczych zjawiskiem rzadkim.

Wśród wskazówek mówiących o tym, jak budować języki informacyjno-wyszukiwawcze, często znajduje się zalecenie, aby **antonimy języka naturalnego**, takie jak *rewolucja* i *kontrrewolucja*, *reformacja* i *kontrreformacja*, traktować jako **umowne synonimy** i wybierać na deskryptor tylko jedno wyrażenie z pary, reprezentujące tym samym cały zakres. Takie rozwiązanie uzasadnione jest rozumieniem znaczenia wyrażen języków informacyjno-wyszukiwawczych jako denotacji bezpośredniej, a więc relacjami między zbiorami dokumentów desygnowanymi przez wyrażenia tych języków – na przykład w dokumentach traktujących o kontrreformacji przeważnie jest także mowa o reformacji, dokumenty mówiące o kontrrewolucji zawierają także informacje o rewolucji. Należy jednak pamiętać, że takie rozwiązanie może powodować i szum informacyjny, i straty informacji, zależność ta bowiem może zachodzić tylko w jedną stronę: wprawdzie w dokumentach o kontrreformacji będą informacje o reformacji, to dokumenty mówiące o reformacji informacji o kontrreformacji mogą nie podawać, może być także tak, że użytkownik interesujący się dokumentami o reformacji kontrreformacją się nie interesuje. Nie zawsze jest więc tak, że wyrażenia języków informacyjno-wyszukiwawczych pozostające w języku naturalnym w tego typu stosunku antonimii denotują te same zbiory dokumentów, dlatego też nie powinny być traktowane jak synonimy.

Z przytoczonych wyżej przykładów widać też, że inne rozwiązanie należy przyjąć w wypadku wyrażeń o znaczeniu komplementarnym, inne w wypadku wyrażeń parametrycznych.

W wypadku komplementarności mamy zawsze tylko dwie możliwości – coś może być suche albo mniej (wilgotne) lub więcej mokre, gładkie albo mniej lub bardziej szorstkie, czyste albo mniej lub bardziej brudne – natomiast w wypadku wyrażeń parametrycznych coś może nie być ani dobre, ani złe – może być „takie sobie”, ktoś może nie być ani stary, ani młody – może być w średnim wieku, pogoda nie musi być ani upalna, ani mroźna – może być ciepło lub chłodno. W językach informacyjno-wyszukiwawczych te dwa wypadki antonimiczności należy wyraźnie rozgraniczać, bo chyba tylko w wypadku wyrażeń komplementarnych można się zgodzić na uznanie za jednostkę leksykalną tylko jednego z antonimów. Analogiczne rozwiązanie w wypadku wyrażeń parametrycznych prowadziłoby do nieuchronnego zwiększenia szumu informacyjnego, jeśli bowiem użytkownika systemu informacyjnego interesuje, na przykład, sytuacja ludzi starych, to nie musi interesować go sytuacja ludzi młodych, jeśli kogoś interesuje zachowanie się materiałów w wysokich temperaturach, to nie interesuje go zachowanie się materiałów w niskich temperaturach, jeśli przedmiotem zainteresowania jest zabudowa wysoka, to nie musi być nim zabudowa niska. Przykłady można by tu zresztą mnożyć. Podobnie wygląda sprawa wyrażeń antonimicznych typu *ślub – rozwód*, *start – meta*, *analiza – synteza*, a także wyrażeń połączonych relacją konwersu.

W istniejących językach informacyjno-wyszukiwawczych spotyka się różne rozwiązania tego problemu. W niektórych występuje tylko jedno wyrażenie spośród połączonych w języku naturalnym relacją antonimii – na przykład w języku haseł przedmiotowych może być:

Kontrreformacja zob. Reformacja
Antyfaszyzm zob. Faszyzm

W innych bywają traktowane współrzędnie, na przykład w UKD:

314.742 Imigracja
314.743 Emigracja

W wypadku komplementarności w językach informacyjno-wyszukiwawczych zazwyczaj występują tylko wyrażenia nazywające nacechowany człon opozycji, np.:

WILGOTNOŚĆ WŁÓKNA
SZORSTKOŚĆ POWIERZCHNI

czasami jednak występują oba wyrażenia komplementarne, traktowane współrzędnie.

W wypadku wyrażeń połączonych relacją konwersu w językach informacyjno-wyszukiwawczych mogą występować oba wyrażenia. Obecność ich w systemie leksykalnym prowadzi często do synonimii wyszukiwawczej na poziomie syntaktycznym, np.:

EKSPORT. POLSKA – agens. HOLLANDIA – kontragens. WĘGIEL – obiekt
IMPORT. POLSKA – kontragens. HOLLANDIA – agens. WĘGIEL – obiekt

Zazwyczaj jednak konwersja występująca w języku naturalnym nie jest przenoszona do języków informacyjno-wyszukiwawczych.

Koncepcja bezpośredniej denotacji wyrażeń języków informacyjno-wyszukiwawczych decyduje również o rozumieniu **hierarchicznych relacji paradygmatycznych**.

W językach informacyjno-wyszukiwawczych **relacje nadrzędności i podrzędności zakresowej** określa się mianem **relacji generycznych**. Relację typu rodzaj – gatunek nazywa się relacją **nadrzędności generycznej**, zaś relację gatunek – rodzaj relacją **podrzędności generycznej**.

Relacje generyczne są relacjami przechodnimi, z czego wynika, że jeśli relacja nadrzędności generycznej zachodzi między jakimiś wyrażeniami języka informacyjno-wyszukiwawczego, np.:

PARA WODNA – PARA PRZEGRZANA

oraz między wyrażeniami

PARA PRZEGRZANA – PARA PRZEGRZANA PIERWOTNA

to związek nadrzędności generycznej zachodzi również między wyrażeniami

PARA WODNA – PARA PRZEGRZANA PIERWOTNA

(analogicznie w wypadku relacji podrzędności generycznej).

Dla odróżnienia relacji generycznej zachodzącej między wyrażeniami zajmującymi w strukturze hierarchicznej języka węzły bezpośrednio sąsiadujące w łańcuchu klasyfikacyjnym od relacji generycznej zachodzącej między wyrażeniami zajmującymi niesąsiadujące ze sobą węzły tego łańcucha wprowadza się terminy **relacji generycznej bezpośredniej** – odpowiednio podrzędności bezpośredniej i nadrzędności bezpośredniej, oraz **relacji generycznej pośredniej** – podrzędności pośredniej i nadrzędności pośredniej.

Relacje podrzędności – nadrzędności między jednostkami leksykalnymi języków informacyjno-wyszukiwawczych nie zawsze pokrywają się z relacjami między zakresami odpowiadających im wyrażeń języka naturalnego, języki informacyjno-wyszukiwawcze są bowiem językami sztucznymi mającymi własną semantykę i własną paradygmatykę, o czym trzeba pamiętać, aby uniknąć nieporozumień. Łatwo o nie szczególnie w tych językach informacyjno-wyszukiwawczych, których relacje paradygmatyczne nie zostały wyraźnie zadekretowane i pozostawione domyślności użytkowników. Dotyczy to przede wszystkim języków słów kluczowych, w których relacje paradygmatycznych się nie określa, w przeciwieństwie do języków deskryptorowych, których relacje paradygmatyczne określone są w artykule deskryptorowym oraz w części systematycznej tezauryusa, a także w przeciwieństwie do klasyfikacji, których relacje paradygmatyczne ustanawia struktura drzew klasyfikacyjnych, a często odzwierciedla także notacja strukturalna (np. w UKD).

Relacja nadrzędności – podrzędności nazywana bywa w językach informacyjno-wyszukiwawczych także **relacją hierarchii tematycznej** bądź **relacją hierarchii dokumentacyjnej**. Wynika ona z denotacji bezpośredniej jednostek leksykalnych tych języków, a zachodzi najczęściej między wyrażeniami reprezentującymi pewne dyscypliny wiedzy a wyrażeniami reprezentującymi obiekty lub problemy badawcze tych dyscyplin, np.:

semiotyka – znaki;
psychologia – osobowość;
prawo kościelne – kary kościelne,

albo między wyrażeniami nazywającymi tematy szersze a wyrażeniami oznaczającymi tematy węższe, zawierające się w nich, np.:

ptaki – migracja ptaków;
zioła – uprawa ziół;
nowotwory – inwazyjność nowotworów

(bo dokumenty traktujące o znakach są podzbiorem dokumentów zawierających informacje o semiotyce, dokumenty mówiące o migracji ptaków mówią o ptakach, a o uprawie ziół mówią o ziołach). Zauważmy, że relacja ta nie zachodzi między równoznacznymi z wyrażeniami języków informacyjno-wyszukiwawczych wyrażeniami języka naturalnego – między nimi zachodzi relacja wykluczania się zakresów (bo żaden znak nie jest semiotyką, osobowość nie jest psychologią, a kary kościelne nie są prawem kościelnym).

Relacja ta często pełni w językach informacyjno-wyszukiwawczych funkcję **relacji organizującej paradygmatykę** tych języków i odwzorowuje relacje łączące zarówno informacje w tekście dokumentów, jak i w zbiorze informacyjnym systemu.

Wśród relacji ważnych z punktu widzenia budowy języków informacyjno-wyszukiwawczych znajduje się także relacja nazywana **relacją skojarzeniową (relacją asocjacyjną)**. Nazwa ta jest o tyle umowna, że właściwie wszystkie relacje systemowe, czyli paradygmatyczne są relacjami skojarzeniowymi. Relacja ta łączy wyrażenia reprezentujące tematy często omawiane w tych samych dokumentach, ma więc dużą wartość wyszukiwawczą. Dokładniejsze zdefiniowanie tej relacji jest trudne, terminem tym bowiem określa się relacje bardzo różne, zazwyczaj wszystkie te, które są dla wyszukiwania ważne, a które nie są ani relacją podrzędności, ani nadrzędności, ani równoważności. Na przykład A.I. Černyj w pracy *Metodyka budowy tezaurusów* uwzględnia następujące relacje:

- współrzędność,
- podobieństwo funkcjonalne,
- gatunek – rodzaj (rodzaj – gatunek), nazywaną też relacją generyczną,
- przyczyna – skutek (skutek – przyczyna),
- część – całość (całość – część), nazywaną relacją mereologiczną.

W języku naturalnym nie wszystkie wymienione wyżej relacje są relacjami językowymi (relacjami systemowymi między wyrażeniami). Do systemowych ję-

zykowych relacji należą tu tylko **relacje gatunek – rodzaj i rodzaj – gatunek**, czyli po prostu relacje podrzędności i nadrzędności, gdyż rodzaj (najbliższy rodzaj – *genus proximum*) jest to klasa nadrzędna względem jakiejś innej rozważanej klasy, zwanej wówczas gatunkiem (*species*). Oczywiście nie musi to być rodzaj najbliższy, czyli najmniejszy zbiór zawierający dany gatunek, ale ten, który jest z jakichś względów ważny. Relacji współrzędności Černýj nie definiuje, wydaje się jednak, że chodzi tu o relację między dwoma wyrażeniami zdominowanymi w drzewie klasyfikacyjnym przez wspólny węzeł, czyli, mówiąc najprościej, między dwoma wyrażeniami mającymi wspólny najbliższy termin rodzajowy. Tak jest na przykład w wypadku wyrazów *foksterier* i *jamnik*, mających wspólny najbliższy termin nadrzędny *pies* (*pies myśliwski* lub *pies rasowy*). Tak rozumiana relacja współrzędności nazywana bywa także **relacją współpodporządkowania (relacją współpodrzędności)**.

W językach informacyjno-wyszukiwawczych, jak wiemy, relacje podrzędności – nadrzędności określa się ze względu na relacje między zbiorami dokumentów (a właściwie między utrwalonymi w tych dokumentach tekstami), ale przy określaniu ich rodzaju korzysta się często z odniesień do relacji zachodzących między wyrażeniami w systemie języka naturalnego oraz do relacji zachodzących w rzeczywistości pozajęzykowej pomiędzy denotatami tych wyrażeń. I tak, ze względu na rodzaj tych relacji wyróżnia się w językach informacyjno-wyszukiwawczych różne rodzaje relacji współrzędności. Może to więc być:

- **relacja współrzędności generycznej**, zachodząca między wyrażeniami oznaczającymi różne gatunki tego samego rodzaju, np. *świerk – jodła* (wyrażenie rodzajowe – *drzewo iglaste*),

- **relacja współrzędności mereologicznej** (o relacjach mereologicznych powiemy niżej), a więc relacja między wyrażeniami nazywającymi części tej samej całości, np. *silnik samochodowy – podwozie – karoseria* (wyrażenie oznaczające całość: *samochód*),

- **relacja współrzędności tematycznej**, łącząca wyrażenia podrzędne tematycznie wobec tego samego wyrażenia, np. *napięcie prądu elektrycznego – natężenie prądu elektrycznego* podrzędne wobec wyrażenia *prąd elektryczny*.

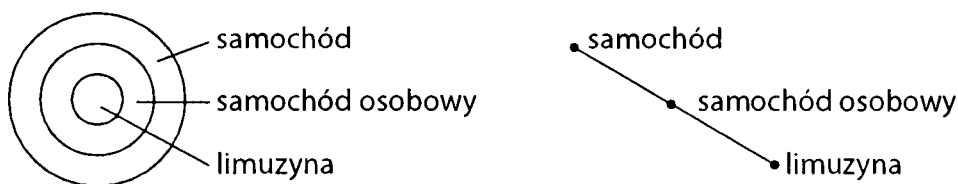
Relacja współrzędności bywa też rozumiana inaczej, jako relacja między wyrażeniami nazywającymi elementy wyodrębniane na tym samym stopniu podziału w drzewie klasyfikacyjnym, które nie mają wspólnego elementu bezpośrednio nadrzędnego – nie są więc współrzędne w pierwszym sensie, nie są bowiem połączone relacją bezpośredniego współpodporządkowania. Tak rozumiana współrzędność zachodzi na przykład między wyrażeniami *świerk* i *klon*, które mają różne terminy bezpośrednio nadrzędne, odpowiednio: *drzewo iglaste* i *drzewo liściaste*.

Pozostałe spośród wymienionych wyżej relacji asocjacyjnych uznawanych w językach informacyjno-wyszukiwawczych za relacje paradygmatyczne, a więc podobieństwo funkcjonalne, skutek – przyczyna oraz część – całość, nie są w wypadku języka naturalnego relacjami między wyrażeniami, lecz relacjami między desygnatami tych wyrażeń (nie są więc relacjami paradygmatycznymi języka

naturalnego, lecz relacjami zachodzącymi w rzeczywistości pozatekstowej). Jednakowe traktowanie w językach informacyjno-wyszukiwawczych relacji systemowych języka naturalnego i relacji zachodzących w odwzorowywanej w języku naturalnym rzeczywistości bywa przyczyną nieporozumień. Szczególnie często dotyczy to w językach deskryptorowych relacji deskryptor szerszy – deskryptor węższy, reprezentującej bądź relację między zakresami wyrażen języka naturalnego, bądź pozajęzykową relację całość – część, zwaną też relacją mereologiczną, która czasami oznaczana jest w tezaurusach jako relacja asocjacyjna.

Relacje między zakresami wyrażen odwołują się do rozumienia zbioru w sensie dystrybutywnym, natomiast **relacje mereologiczne** określane są na elementach zbioru rozumianego w sensie kolektywnym, traktowanego jako zespół wyróżnionych elementów – takimi zbiorami zajmuje się **mereologia** (tu synonimami terminu zbiór są terminy agregat, kompleks, konglomerat; zob. *Mała encyklopedia logiki*).

A oto przykłady obu rodzaju relacji i sposoby ich sprawdzania:



Limuzyna – samochód osobowy – samochód to terminy ułożone od najwęższego do najszerszego, bo każda limuzyna jest samochodem osobowym, a każdy samochód osobowy jest samochodem, natomiast nie każdy samochód jest samochodem osobowym i nie każdy samochód osobowy jest limuzyną. Mamy tu więc do czynienia z relacją między wyrażeniami, gdzie zbiór traktowany jest w sposób dystrybutywny, natomiast relacja między maską samochodu, karoserią samochodu i samochodem nie jest językową relacją podrzędności zakresowej, bo żadna maska samochodowa nie jest karoserią i żadna karoseria nie jest samochodem – między tymi wyrażeniami zachodzi tylko relacja wykluczania, natomiast między desygnatami tych terminów zachodzi relacja część – całość, maska jest bowiem częścią karoserii, a karoseria częścią samochodu. Zbiór jest tu więc traktowany w sensie kolektywnym (ze zdania *Kupiłam limuzynę* wynika, że kupiłam samochód, natomiast ze zdania *Kupiłam maskę samochodu* nie wynika, że kupiłam samochód).

O konieczności rozróżniania tych relacji należy pamiętać przy tworzeniu reguł wnioskowania dla tzw. systemów inteligentnych (ekspertowych).

Wśród **relacji mereologicznych** wyróżnić należy dwa podstawowe typy:

- tzw. **relację mereologiczną właściwą (stałą)**, zachodzącą między elementami pewnej całości stanowiącej strukturę stałą a jej częścią, np.:

drewno – konar, roślina – łodyga, krew – krwinka, kościec – kość, układ krwionośny – tętnica,

- tzw. **relację mereologiczną niewłaściwą (sytuacyjną)**, zachodzącą między strukturami niestałymi i ich częściami, np.:

samochód – silnik, silnik elektryczny – wirnik, samolot – śmigło.

Tylko relacje mereologiczne stałe mogą być wykorzystywane do organizowania systemu leksykalnego języków informacyjno-wyszukiwawczych i dlatego można je uważać w tych językach za relacje paradygmatyczne sensu stricto (relacje nadrzędności i podrzędności mereologicznej). Relacje mereologiczne niestałe zawsze są tylko relacjami asocjacyjnymi i należą do relacji pozajęzykowych (niesystemowych). Relacjami mereologicznymi są też niektóre relacje hierarchii tematycznej, na przykład relacje pomiędzy dyscypliną wiedzy a jej działem, np.:

BIBLIOTEKARSTWO – CZYTELNICTWO,
MATEMATYKA – GEOMETRIA,
FIZYKA – FIZYKA CIAŁA STAŁEGO,
LINGWISTYKA – FONOLOGIA.

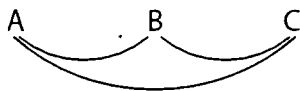
Relacje asocjacyjne mają w języku często charakter wtórny, są bowiem najczęściej odbiciem relacji zachodzących między elementami rzeczywistości, dlatego też nie zawsze bywają uważane za relacje paradygmatyczne języków naturalnych. W językach informacyjno-wyszukiwawczych, szczególnie zaś w językach deskryptorowych, relacje te, ze względu na dużą wartość wyszukiwawczą, włączone są do relacji paradygmatycznych, a wyrażenia nimi połączone umieszczone są w artykule deskryptorowym, przy czym czasami rodzaj relacji asocjacyjnej bywa sprecyfikowany, co znacznie zmniejsza szum informacyjny w systemie.

Najczęściej za relacje skojarzeniowe przyjmuje się tu relacje łączące następujące obiekty:

- obiekt – część, np. mikser – silnik,
- obiekt (agens lub instrument) – proces, np. konsultant – konsultacja, termometr – pomiar temperatury,
- obiekt – własność, np. wyszukiwanie informacji – efektywność techniczna, łącze telekomunikacyjne – przepustowość gwarantowana (CIR),
- własność – własność, np. barwa – ton,
- obiekt – obiekt jako atrybut, np. nafta – lampa naftowa,
- własność – własność jako atrybut, np. skośność – dźwigary skośne,
- proces – produkt, np. budowa okrętów – okręt,
- proces – lokalizacja, np. nauczanie – szkoła,
- obiekt, własność, stan – lokalizacja, np. języki obce – laboratoria językowe,
- przyczyna – skutek, np. korozja – perforacja,
- proces – proces współwystępujący w pewnej sekwencji, np. badanie lekarskie – diagnoza,
- obiekt – zastosowanie, np. tramwaje – komunikacja miejska,
- podobieństwo funkcjonalne, np. termometr – barometr – higrometr,
- zbieżność własności, np. mgła – gaz – para,
- obiekt – funkcja, np. strychnina – leki.

Zarówno relacja termin szerszy – termin węższy (nadrzędność – podrzędność), jak i relacja część – całość jest relacją przechodnią, czyli że spełnia odpowiednio warunki:

- jeżeli A jest B i B jest C, to A jest C
- jeżeli A jest częścią B i B jest częścią C, to A jest częścią C



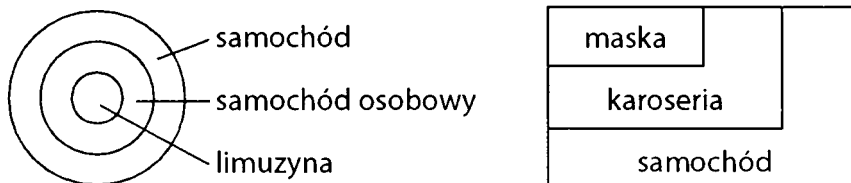
na przykład:

- jeżeli limuzyna jest samochodem osobowym, a samochód osobowy jest samochodem, to limuzyna jest samochodem,
- jeżeli jamnik jest psem, a pies jest zwierzęciem domowym, to jamnik jest zwierzęciem domowym,

oraz:

- jeżeli maska jest częścią karoserii, a karoseria jest częścią samochodu, to maska jest częścią samochodu.

Relacje te mogą więc tworzyć hierarchie, a więc:



stanowiące chociażby wycinki następujących drzew klasyfikacyjnych (w drzewach tych nie zaznaczono pełnej rozbudowy współrzędnej):



Relacja przechodnia zachodzi między dowolnymi (niekoniecznie kolejnymi) węzłami tej samej gałęzi drzewa klasyfikacyjnego. Należy jednak pamiętać, że nie jest relacją przechodnią każda relacja asocjacyjna, o czym się często zapomina, w tym przypadku więc, choć jest spełniony warunek:



to nie musi być tak, że:



Przykładem ilustrującym powyższe może być anegdota, którą podaje S. Lem w *Filozofii przypadku*:

„Stary, niewidomy Mordechaj przyszedł do Izaaka i spytał go:

- Co robisz?
- Piję mleko – odparł Izaak.
- Co to jest „mleko”?
- Taki biały płyn.
- Co to jest „biały”?
- Biały jest łabędź.
- Co to znaczy „łabędź”?
- To taki ptak z krzywą szyją.
- A co to jest „krzywe”?

Izaak zgiął rękę w łokciu i dał ją pomacać Mordechajowi.

- To jest krzywe.

Mordechaj uważnie obmacał jego rękę i rzekł z wdzięcznością: – Dziękuję ci, Izaak! teraz już wiem, jak wygląda mleko!”

gdzie:



ale nie



Doboru terminów skojarzeniowych (terminów pokrewnych) przy budowie języka informacyjno-wyszukiwawczego, na przykład języka deskryptorowego, dokonywać można trzema różnymi metodami, przy czym rezultaty takiego doboru mogą się znacznie od siebie różnić. Czynić to można, opierając się na:

- analizie tekstów dokumentów stanowiących bazę danych systemu informacyjnego;
- analizie potrzeb użytkowników systemu — na przykład na podstawie analizy zapytań informacyjnych wpływających do systemu, co jest metodą obiektywną, lub na podstawie eksperymentu podobnego do opisanego wyżej;

- intuicji i wiedzy projektantów języka, na co składa się zarówno ich wiedza o języku, jak i wiedza o rzeczywistości pozajęzykowej.

Wybór metody zależy od potrzeb konkretnego systemu i jego użytkowników.

Relacje między terminami pokrewnymi (skojarzonymi) zazwyczaj nie bywają w teaurusie dalej specyfikowane, co staje się przyczyną znacznego szumu informacyjnego i zmniejsza wskaźnik relewancji. Specyfikacja taka bywa trudna, co oczywiście nie oznacza, że należy z niej rezygnować. Jako przykład teaurusu z oznaczonym typem relacji między deskryptorami wchodzącymi w skład jednego artykułu deskryptorowego, służyć może radziecki *Tezaurus po teoretičeskoj i prikladnoj lingvistike* S.E. Nikitiny („Nauka”, Moskwa 1978), który w zamyśle autorki może być wykorzystany przy automatycznym przetwarzaniu tekstu dokumentów. Tezaurus ten może być uważany za pewien model teaurusu niezależnie od dziedziny wiedzy, którą reprezentuje i stanowić cenną pomoc metodyczną przy budowie języka deskryptorowego, dlatego też podamy schemat artykułu deskryptorowego z omówieniem rodzaju relacji oraz przykłady konkretnych artykułów.

Przy doborze terminów wchodzących w skład artykułu deskryptorowego zastosowano tu dwie metody:

- metodę symulacji pytań i analizy terminów powiązanych z terminami określającymi tematy pytań,
- eksperyment ankietowy ustalający zestawy terminów powiązanych relacjami asocjacyjnymi.

Wyniki porównano i w ten sposób otrzymano podzbiór terminów występujących w obu zbiorach. Również kontrolę wyboru relacji prowadzono metodami symulacji pytań i ankietowego eksperymentu kojarzeniowego. Wyodrębniono w ten sposób następujące relacje, oznaczone w artykule deskryptorowym podanymi niżej numerami, przy czym numerem 1 oznaczany jest deskryptor hasłowy (wszystkie przykłady ilustrujące wyróżnione relacje wyróżniam graficznie, chodzi tu bowiem o relacje między wyrażeniami języka deskryptorowego):

1. Deskryptor hasłowy
2. Synonimia

Za pełne synonimy uznano te terminy, które miałyby identyczne artykuły deskryptorowe (w tych wypadkach zastosowano odsyłacz do terminu uznanego za główny). Za synonimy uznano również terminy różnych nauk występujące w tym samym znaczeniu oraz pary różnojęzyczne, a także terminy żargonowe.

3. Korelacja

Relacje między terminami przeciwstawnymi, a więc antonimami lub terminami dopełniającymi się należącymi do jednej kategorii i razem stanowiącymi nazwę całego zbioru, np. zwykły szyk wyrazów – inwersja (antonimia), aspekt dokonany – aspekt niedokonany (terminy dopełniające się), temat – reumat (terminy wyczerpujące cały zbiór dzielony).

4. Relacja generyczna (nazwa rodzaju)
5. Relacja gatunkowa (nazwa gatunkowa)
6. Całość
7. Część

Analizując instrukcje wyszukiwawcze stwierdzono, że relacja całość zmienia się na relację część znacznie częściej niż część na całość, np. rzeczywisty związek implikacyjny zachodzi pomiędzy terminami paragraf (akapit) – tekst, a nie odwrotnie.

8. Klasa

Nazwa klasy, do której należy termin hasłowy, np. termin leksem pozostaje w relacji klasa do terminu słowoforma.

9. Człon klasy

Relacja odwrotna w stosunku do relacji klasa.

10. Parametr

Relacja ta wyraża istotne cechy pojęcia określonego przez deskryptor – powinna wskazywać istotne (dla danego systemu) cechy denotatu danego deskryptora, np. deskryptor części mowy jest związany relacją parametr z terminem kategorie gramatyczne.

11. Nośnik parametru

Relacja odwrotna do relacji parametr – np. termin czasownik jest w relacji nośnik parametru do terminu kategoria przechodniości – nieprzechodniości.

12. Podstawowa funkcja

Relacja ta wiąże deskryptor z terminem oznaczającym to, do czego denotat deskryptora służy, np. znaki przestankowe – delimitacja tekstu.

13. Sposób wyrażania funkcji

Relacja odwrotna do relacji podstawowa funkcja, np. podmiot – mianownik, wskazywanie – zaimek wskazujący.

14. Obiekt – sposób przedstawienia

Relacja ta wyraża związek między obiektami a ich reprezentacją (sposobem przedstawienia), np. struktura syntaktyczna – drzewo zależności, leksyka – słownik.

15. Sposób przedstawienia – obiekt

Relacja odwrotna do poprzedniej.

16. Podstawowa operacja (procedura)

Relacja operacja wskazuje:

a) procesy zachodzące w obiekcie – np. składanie słów,

b) działanic na obiekcie – np. segmentacja (w stosunku do obiektu tekst).

W relacji operacja denotat deskryptora może występować w następujących rolach:

a) jako wyjściowy lub końcowy obiekt operacji.

b) jako rezultat lub cel operacji,

c) jako instrument operacji,

d) jako pośredni przedmiot operacji,

e) jako cecha lub przedmiot, do którego kontroli lub wykrycia konieczna jest operacja.

Przykłady: dla afksu – afksacja, dla elipsa – uzupełnienie, dla walencja – nasylenie, dla homonimia – likwidacja (usunięcie),

Role szczególnie istotne specyfikowane są w następnych punktach (17 i 18).

17.

- a) obiekt wyjściowy,
- b) obiekt rezultat

Np. dla terminu automatyczne indeksowanie terminem określającym obiekt wyjściowy jest dokument, a obiekt rezultat (końcowy) – charakterystyka wyszukiwawcza dokumentu.

18. Instrument/metoda

Relacja ta obejmuje zjawiska różnego typu: instrumenty, metody, sposoby analizy, np. termin automatyczne indeksowanie powiązany jest tą relacją z terminami: słowo kluczowe, deskryptor, język indeksowania. Dla nazw nauk pozycja relacji instrument zawiera nazwy metod wykorzystywanych w danej nauce.

19. Poziom

Relacja ta wiąże nazwy poziomów z nazwami ich jednostek, np. afiks – poziom morfologiczny, grupa syntaktyczna – poziom syntaktyczny.

20. Jednostki poziom

Relacja odwrotna do relacji poziom, np. jednostkami dla poziomu morfologicznego są morfem i słowoforma.

21. Rodzaj stosunku

Relacja określająca rodzaj stosunku między dwoma obiektami, np. dla synonimów – synonimia.

22. Człony relacji

Relacja odwrotna do poprzedniej, np. dla deskryptora homonimia syntaktyczna terminami związanymi tą relacją są terminy: konstrukcje i struktury syntaktyczne.

23. Implikacja

Relacja ta wyraża ścisłą korelację typu: przyczyna – skutek, warunek – zjawisko uwarunkowane, np. konwersja – homonimia.

24. Kontekst/dyscyplina

Relacja wiąże deskryptor z nazwą dyscypliny albo z nazwą kontekstu, w którym dany termin się znajduje, np. subkategoryzacja – gramatyka generatywna, cecha dystynktywna fonemu – fonologia.

25. Aspekt

Relacja przyporządkowuje danemu deskryptorowi określony aspekt, w jakim jest on rozpatrywany, np. formant – plan wyrażania.

26. Obiekt nauki (dyscypliny)

Relacja wiąże nazwę działu wiedzy lub dyscypliny z nazwą obiektu, którym się zajmuje, np. lingwistyka matematyczna – gramatyki formalne.

27. Odnosi się do ...

Relacja wiąże deskryptor z przymiotnikiem wywodzącym się od niego znaczeniowo, przy czym związki semantyczne rzeczownika z przymiotnikiem mogą być różnorodne, np. alfabet – alfabetyczny.

28. Niestandardowe relacje asocjacyjne

Klasa ta obejmuje wszystkie te relacje, które nie zostały uwzględnione w poprzednich pozycjach, ale występują regularnie, są więc charakterystyczne dla wielu par terminów.

A oto przykłady artykułów deskryptorowych, przetłumaczonych na język polski (oczywiście nie wszystkie pozycje w każdym artykule zostały wypełnione):

1. Język informacyjno-wyszukiwawczy

2: JIW,

4: język informacyjny,

5: język deskryptorowy (język unitermów, JIW z gramatyką), języki indeksowania,

6: (system informacyjno-wyszukiwawczy),

7: (słownik deskryptorów), gramatyka 3, tezaurus informacyjno-wyszukiwawczy,

7 = 20: deskryptor, słowo kluczowe,

16: indeksowanie,

24: (informatyka).

1. Lingwistyka

2: językoznawstwo,

4: (nauka),

7: a) (fonetyka, fonologia), morfologia 2, syntaktyka 2, leksykologia, leksykografia;

b) językoznawstwo zewnętrzne, językoznawstwo wewnętrzne,

c) (językoznawstwo tradycyjne), lingwistyka strukturalna, lingwistyka formalna,

d) językoznawstwo statystyczne, lingwistyka stosowana,

16 = 18: modelowanie (lingwistyczne),

26: język naturalny,

27: lingwistyczny.

Pole semantyczne

Strukturalizacja pola semantycznego języka informacyjno-wyszukiwawczego wyznaczana jest przez zbiór wyrażeń tego języka, jego system leksykalny. To wyrażenia języka dzielą jego pole semantyczne na poszczególne podpola poprzez swoją denotację. Podpola te nie są równe swym zakresem elementarnym cechom znaczeniowym – stanowią one całe wiązki cech semantycznych skupionych przez dane wyrażenie w jednej strukturze. Tak wydzielone struktury semantyczne stanowią **konotację** wyrażeń danego języka. Poszczególne elementy konotacji (cechy semantyczne, elementarne znaczenia) mogą stanowić **zasady podziału** całego pola semantycznego danego języka i wyznaczać jego podpola, reprezentowane w systemie leksykalnym przez klasy wyrażeń zawierające w swoich konotacjach dany element. Tak więc podział pola semantycznego wyznacza odpowiedni podział leksyki danego języka – polom semantycznym odpowiadają **pola leksykalne**.

Choć rzeczywistość semantyczna danego języka (zbiór znaczeń odwzorowanych w danym języku) jest inną rzeczywistością niż rzeczywistość leksykalna

języka (słownik), to przy bilateralnej koncepcji znaku językowego, wywodzącej się od F. de Saussure'a, a więc traktującej znak językowy jako element dualny, składający się z zespolonych w jedną całość formy i treści (znaczenia), rzeczywistości te są nierozdzielalne ze sobą związane, a co za tym idzie mają wspólną strukturę. Jeśli rozdzielimy je, dla wygody rozważań, na plan wyrażania (formy wyrażania) i plan treści (formy treści), to otrzymamy struktury izomorficzne względem siebie. Orzekając o strukturze pola semantycznego, orzekamy tym samym o strukturze pola leksykalnego, a więc systemie leksykalnym języka, który nie jest zwykłym zbiorem wyrazów, ale zasobem wyrazów połączonych ze sobą relacjami paradygmatycznymi.

W wypadku języków informacyjno-wyszukiwawczych takie podpola w polu semantycznym całego języka wydzielane są ze względu na relewantne dla użytkowników danego języka cechy informacji i nazywane, zależnie od rodzaju języka, **kategoriami** lub **fasetami**. Na przykład polami (podpolami) semantycznymi w polu semantycznym Klasyfikacji Dwukropkowej Ranganathana są kategorie (fasety): Indywiduum, Materia, Energia, Przestrzeń, Czas, a wypełniające je wyrażenia tworzą odpowiednie kategorie leksykalne (podpola) w polu leksykalnym tego języka. W języku hasel przedmiotowych takimi polami semantycznymi są poszczególne kategorie tematów i określników – skupione w nich wyrażenia tworzą pole leksykalne tego języka. W Uniwersalnej Klasyfikacji Dziesiątej takimi polami semantycznymi będą na przykład poszczególne gałęzie klasyfikacyjne, a wypełniające je symbole klasyfikacyjne będą stanowiły pola (podpola) leksykalne klasyfikacji.

Kategorie leksykalne wyróżnione w poszczególnych językach informacyjno-wyszukiwawczych odzwierciedlają jakąś klasyfikację rzeczywistości (przyjętą **ontologię**).

A oto przykłady:

W UKD wyróżnia się kategorie leksykalne poddziałów wspólnych: języka, formy wydawniczej lub piśmienniczej dokumentu, czasu, miejsca, grupy etnicznej, punktu widzenia (sposobu ujęcia tematu), materiałów i osób (grup) oraz nazw własnych (osób, organizacji, miejsc, tzw. poddział A/Z). W języku tym kategorię leksykalną symboli głównych podzielono na dziesięć podkategorii, odpowiadających grupom dyscyplin wiedzy:

- 0 Dział ogólny
- 1 Filozofia. Psychologia
- 2 Religia. Teologia. Religioznawstwo
- 3 Nauki społeczne. Prawo. Administracja
- 4 (obecnie dział wolny)
- 5 Matematyka. Nauki przyrodnicze
- 6 Nauki stosowane. Medycyna. Nauki techniczne. Rolnictwo
- 7 Sztuka. Rozrywki. Sport
- 8 Językoznawstwo. Nauka o literaturze. Literatura piękna
- 9 Archeologia. Prehistoria. Geografia. Biografie. Historia.

W języku **hasel przedmiotowych** wyróżniano kategorie:

1) tematów, z podkategoriami: tematów klasowych, przedmiotowych i formalnych, kategorie tematów ogólnych i indywidualnych, podzielonych na podkategorie grupujące tematy: osobowe, korporatywne, językowo-etniczne, geograficzne, chronologiczne;

2) kategorie określników, podzielonych na kategorie określników szczegółowych z podkategoriami określników: jednostkowych, gatunkowych, językowo-etnicznych, określników ogólnych z podkategoriami określników: klasowych, formalnych oraz określników lokalizujących z podkategoriami określników geograficznych i chronologicznych.

Obecnie stosuje się inny, uproszczony, podział tych kategorii.

W **kodzie semantycznym Perry'ego i Kenta** leksyka podzielona została na sześć głównych kategorii podlegających dalszej systematyzacji:

- pojęcia ogólne, z podkategoriami: idee, stosunki, własności, pola badań, siły, klasyfikacje, podzielone ze względu na kompozycję, naturę, zastosowanie;

- stosunki, z podkategoriami: stosunki ekonomiczne, stosunki fizyczne, stosunki społeczne z podkategoriami: podstawowe pojęcia socjologiczne, pojęcia dotyczące społeczeństwa, procesy zachodzące w społeczeństwie, wskaźniki dotyczące społeczeństwa;

- stany, z podkategoriami: stany psychologiczne, stany wyzwajające działanie, stany fizyczne, podzielone na: odczuwalne zmysłowo, kondycjonalne i własności wywołujące stany;

- procesy, z podkategoriami: procesy fizyczne, procesy ogólne i główne, procesy materialne, podzielone na: procesy powodujące zmiany, procesy konstruktywne, procesy destruktywne, a także podkategorie procesy społeczne i przemysłowe;

- substancje, z podkategoriami: substancje ogólne i substancje szczegółowe, podzielone na: substancje organiczne w przyrodzie i substancje nieorganiczne wytworzone przez człowieka, a także podkategorie preparaty i przetwory;

- przedmioty, obiekty, z podkategoriami: przedmioty ogólne, podzielone na: naturalne (z włączeniem istot żywych), produkty (artefakty) oraz przedmioty specjalne, podzielone na: naturalne (z włączeniem istot żywych) oraz produkty (artefakty).

W omówionym w poprzednim rozdziale projekcie **Kodu Semantycznego dla SINTO** zaproponowano siedem głównych kategorii: obiekty, obiekty sekundarne, operacje na obiektach, operacje na operacjach, charakterystyki obiektów, charakterystyki operacji i role semantyczne.

Wśród obiektów wyróżniono trzy podkategorie obiektów prymarnych, odpowiadające trzem podstawowym kategoriom ontologicznym: informacja, energia i substancja (materia), podlegające dalszej kategoryzacji. I tak, w kategorii substancji wyróżniono podkategorie: substancja gazowa, substancja płynna, substancja stała, substancja plazmowa; w kategorii energii podkategorie: energia

mechaniczna, energia termiczna, energia elektryczna, energia nuklearna i energia elektromagnetyczna, zaś informację poddano dalszej kategoryzacji kombinatorycznej, przeprowadzonej według kryteriów: 1) semantyczność, gdzie wyróżniono: informację niesemantyczną, informację semantyczną nie interpretowaną świadomie, i informację semantyczną interpretowaną świadomie; 2) znakowość, gdzie wyróżniono informację nieznakową i informację znakową; 3) stosunek wobec systemu, gdzie wyróżniono: informację systemową (wytworzoną w obrębie systemu) i informację pozasystemową (pochodzącą spoza systemu).

Wśród obiektów sekundarnych wyróżniono wstępnie dwie podkategorie: instrument, z podkategoriami: obiekt nieożywiony bez napędu własnego, obiekt nieożywiony z napędem własnym, obiekt ożywiony niebędący istotą ludzką, istotą ludzką i instytucje oraz miejsce, z podkategoriami: lokalizacja w środowisku gazowym, lokalizacja w środowisku płynnym, lokalizacja w środowisku stałym, lokalizacja na styku środowiska gazowego i płynnego, lokalizacja na styku środowiska gazowego i stałego oraz lokalizacja na styku środowiska płynnego i stałego.

Wśród kategorii operacje na obiektach wyróżniono trzy główne podkategorie: transport, produkcja, konsumpcja, z możliwością ich dalszego podziału według kategoryzacji kombinatorycznej. I tak, kategorię transport podzielono według kryterium istnienia przewodnicy na transport bezprzewodnicowy i transport przewodnicowy, według kryterium istnienia kontenera na transport bezpojemnikowy i transport pojemnikowy, a także według wymienionych wyżej kryteriów rodzaju środowiska i rodzaju skali (której kategoryzację podamy niżej).

W kategorii produkcji wydzielono podkategorie według zasady kreatywności: produkcję eksploatacyjną (wykorzystywanie naturalnych form energii, substancji czy informacji) i produkcję fabrykacyjną, a według kryterium ingerencyjności: produkcję bezingerencyjną i produkcję ingerencyjną.

W kategorii konsumpcji wyróżniono podkategorie według kryterium zużywania: konsumpcję nieniszczącą i konsumpcję niszczącą.

W kategorii charakterystyk operacji wydzielono dwie podkategorie: środowisko fizyczne, podzielone dalej analogicznie jak kategoria lokalizacji, oraz kategorię skali zjawisk, gdzie wyróżniono podkategorie: skala mikrolokalna, skala lokalna, skala regionalna, skala krajowa, skala globalna i skala kosmiczna.

W kategorii ról semantycznych wyodrębniono 17 kategorii: agens, aktans, pacjens, aktant, surowiec, substancja pomocnicza, instrument, produkt, cel, rezultat, metoda działania, środowisko, okoliczność, miejsce (lokalizacja stała, początkowa, końcowa, trasa), czas (data, termin początkowy, termin końcowy, okres trwania), miara, proces.

O problemie kategoryzacji, przede wszystkim w językach informacyjno-wyszukiwawczych, traktuje książka Jadwigi Woźniak pt. *Kategoryzacja. Studium z teorii języków informacyjno-wyszukiwawczych*. Tam również znajdzie czytelnik obszerną bibliografię przedmiotu.

Wszystko, co powiedzieliśmy wyżej, dobrze tłumaczy się w terminach koncepcji pola semantycznego konstytuowanego na podstawie relacji paradygmatycz-

nych języka sensu stricto, tak jak to było w teorii Triera – Weisgerbera, gdzie relacją organizującą pole było posiadanie wspólnych elementów znaczeniowych przez wyrazy danego systemu leksykalnego. Do opisów pól semantycznych tych języków informacyjno-wyszukiwawczych, których wyrażenia zawierają w swojej strukturze wykładnik przynależności do danej kategorii, pasuje koncepcja pola semantycznego Ipsena, która wymagała, aby wyrazy tworzące pole semantyczne posiadały w swojej strukturze morfologicznej formalne wykładniki cechy semantycznej wspólnoty. Takie pola semantyczne tworzą wyrażenia języków informacyjno-wyszukiwawczych o notacji strukturalnej, na przykład Uniwersalnej Klasyfikacji Dziesiątej, gdzie przynależność do wspólnego łańcucha klasyfikacyjnego wyznaczana jest przez wspólną część symbolu klasyfikacyjnego, czy też wyrażenia kodów semantycznych, na przykład kodu semantycznego Perry'ego-Kenta, gdzie pole semantyczne ma swój wykładnik formalny w postaci tzw. mnożnika semantycznego utworzonego, żeby było łatwiej zapamiętać, z liter oznaczających spółgłoski w nazywających daną kategorię morfemach leksykalnych języka angielskiego, np. N-ML – zwierzę (*animal*). M-SR – miara (*measure*), T-MP – temperatura (*temperature*), M-CH – urządzenie (*machine*).

Koncepcja Porziga, włączająca do pola semantycznego również wyrażenia oparte na relacjach syntagmatycznych, oraz koncepcje pola asocjacyjnego (szeroko rozumianego) są zbieżnie ze sposobem prezentowania języków deskryptorowych – tezaurus jest właśnie obrazem takiego ustrukturalizowanego pola semantycznego języka informacyjno-wyszukiwawczego, pokazującym składowe pola semantyczne z centrum – deskryptorem tytułowym oraz innymi deskryptorami powiązanych z nim zarówno relacjami generycznymi (paradygmatycznymi sensu stricto), jak i relacjami asocjacyjnymi, najczęściej uwzględniającymi współwystępowanie tych deskryptorów w charakterystykach wyszukiwawczych dokumentów, a więc relacje kontekstowe (syntagmatyczne).

Hipertekstowe internetowe systemy informacyjne strukturalizują swoje pola informacyjne poprzez pola semantyczne tworzone przez słowa kluczowe wyszukiwarek połączone gęstą siecią odsyłaczy, odwzorowujących (i ustanawiających) relacje organizujące poszczególne podpola w sposób umożliwiający użytkownikowi odnalezienie w zbiorze informacyjnym potrzebnych mu informacji. Im lepiej zorganizowane takie pole – informacyjne, dokumentacyjne, tematyczne, semantyczne, leksykalne – tym droga do poszukiwanych informacji będzie krótsza.

Języki informacyjno-wyszukiwawcze odwzorowują informacje tworzące **pole informacyjne** systemu w pewien określony sposób. W każdej informacji wyróżnić można dwie jej części: część identyfikującą obiekt rzeczywistości pozainformacyjnej, której informacja dotyczy i część, w której o tym obiekcie coś się orzeka (informację o danym obiekcie). Ta pierwsza część to temat informacji (komunikatu), ta druga, to remat. Informacja ma więc tzw. **strukturę tematyczno-rematyczną**, realizowaną w komunikatach.

Zbiór potencjalnych tematów, tworzących potencjalne pole tematyczne systemu jest w danym zbiorze informacyjnym skończony, zbiór tematów nie. Dlatego tak naprawdę pole informacyjne systemu wyznaczone jest przez zbiór tematów

informacji, jest więc w istocie równoznaczne z polem tematycznym systemu informacyjnego. Bo to tematy informacji odwzorowane są w opisie rzeczowym w systemach informacyjno-wyszukiwawczych, to tematy są odwzorowane w języku informacyjno-wyszukiwawczym i to poprzez wyrażenia tych języków reprezentujące tematy użytkownik wyszukuje interesujące go informacje.

Tak więc i z punktu widzenia systemu informacyjno-wyszukiwawczego, i z punktu widzenia jego użytkownika istnieje równoznaczność między polem informacyjnym, polem dokumentacyjnym i polem tematycznym systemu informacyjnego, a także ich izomorfizm, bowiem i zakres tych pól i ich struktura wyznaczane są przez język informacyjno-wyszukiwawczy systemu – struktura ich pól jest więc izomorficzna ze strukturą pola semantycznego tego języka. Orzekając coś o strukturze pola semantycznego języka informacyjnego, orzekamy tym samym o strukturze pola tematycznego systemu informacyjnego, jego pola dokumentacyjnego i pola informacyjnego. Nie znaczy to jednak, że terminów tych można używać zamiennie – rzeczywistość semantyczna języka jest bowiem czymś innym niż rzeczywistość dokumentacyjna i rzeczywistość informacyjna, choć w wypadku systemów i języków informacyjno-wyszukiwawczych rzeczywistości te są względem siebie izomorficzne ze względu na dany język informacyjno-wyszukiwawczy. Każdy język informacyjno-wyszukiwawczy nakłada bowiem na całe uniwersum dokumentacyjne i informacyjne inną strukturę, tak jak każdy język naturalny nakłada inną siatkę semantyczną na rzeczywistość pozatekstową.

Prezentacja języka

Języki informacyjno-wyszukiwawcze są językami sztucznymi, formalnymi w tym sensie, że *explicite* powinien być zaprezentowany system języka, a więc jego słownik i gramatyka – pozwalające na generowanie wyrażen złożonych (zdań i tekstów spójnych), reguły semantyki – pozwalające na przyporządkowanie wyrażen języka elementom odwzorowywanej rzeczywistości pozajęzykowej oraz reguły pragmatyki – zapewniające poprawne użycie języka, pozwalające na tworzenie komunikatów adekwatnych do konkretnych sytuacji komunikacyjnych. Warunki poprawnej **prezentacji języka** nie zawsze są spełniane, z reguły twórcy tych języków nie formułowali reguł pragmatyki – czynili to zazwyczaj dopiero użytkownicy tworzący systemy informacyjne, formułując odpowiednie instrukcje indeksowania, czyli opracowania rzeczowego (wyjątkiem jest tu Adam Łysakowski, twórca polskiego języka haseł przedmiotowych, który wiele uwagi poświęcił problemowi odwzorowywania przedmiotu dokumentu), z reguły też częściowo tylko prezentowany był system semantyczny tych języków. Temu problemowi poświęcimy nieco więcej uwagi.

Funkcję metajęzyka w prezentowaniu znaczenia wyrażen języków informacyjno-wyszukiwawczych pełni język naturalny, nie zawsze jednak funkcja ta jest pełniona w sposób jawny. Zależy to przede wszystkim od typu języka informacyjno-wyszukiwawczego – funkcja ta jest w dużej mierze jawna w wypadku języków o notacji sztucznej, w wypadku języków paranaturalnych zazwyczaj bywa ukryta.

Najprostszą formą paranaturalnych języków informacyjno-wyszukiwawczych są **języki słów kluczowych**. Elementarne jednostki leksykalne tych języków, a więc słowa kluczowe są równokształtne z wyrażeniami języka naturalnego stanowiącego bazę do generowania słów kluczowych. Znaczenie słów kluczowych nie jest *explicite* podawane, zakłada się bowiem, że użytkownik języka słów kluczowych sam się go domyśli na podstawie swojej kompetencji językowej w zakresie odpowiedniego języka naturalnego. Zakłada się również, że użytkownik sam odkryje zasadę, że słowa kluczowe odwzorowują tematy dokumentów, a nie elementy rzeczywistości pozatekstowej, znaczenie ich odpowiada więc równokształtnym z nimi wyrażeniom języka naturalnego użytym w funkcji metainformacyjnej.

W najprostszej postaci języki słów kluczowych prezentowane są przez tzw. **słownik negatywny** (stop-listę), czyli spis wyrażeń odpowiedniego języka naturalnego, które nie stanowią wyrażeń języka informacyjno-wyszukiwawczego lub tylko przez spis kryteriów, jakie powinny spełniać wyrażenia, które nie mogą być słowami kluczowymi lub które mogą być użyte w tej funkcji. Są to tzw. **języki swobodnych słów kluczowych**. W takich wypadkach język informacyjno-wyszukiwawczy właściwie w ogóle nie jest prezentowany i stanowi część kompetencji w zakresie języka naturalnego poszczególnych użytkowników języka informacyjno-wyszukiwawczego, ich kompetencji leksykalnej i semantycznej, w tym kompetencji w zakresie paradygmatyki systemu semantycznego. Kompetencja użytkownika języka swobodnych słów kluczowych w zakresie tego języka kształtuje się więc w trakcie używania tego języka, a w wypadku użytkownika systemu informacyjnego posługującego się takim językiem głównie w procesie wyszukiwania informacji metodą prób i błędów. Różnice w kompetencji językowej poszczególnych użytkowników języków swobodnych słów kluczowych mogą być więc znaczne, a szczególnie duże w takich wypadkach, gdy bazą słownikową tego języka jest język niebędący ich językiem ojczystym, jak często bywa w poszukiwaniach internetowych. Jest to również istotne w wypadku, gdy stosowany w danym systemie język słów kluczowych jest inny niż język reprezentowanych przez te słowa kluczowe dokumentów, i tu bowiem przy rozumieniu znaczenia wyrażeń języka informacyjno-wyszukiwawczego nakładają się niejednoznaczności wynikające z nieprzystawania do siebie struktur semantycznych różnych języków naturalnych. Jest to zresztą istotne przy prezentacji i rozumieniu znaczenia wyrażeń wszystkich języków informacyjno-wyszukiwawczych, w wypadku języków swobodnych słów kluczowych o tyle bardziej istotne, że ich użytkownik nie dysponuje żadnymi informacjami dotyczącymi struktury semantycznej tych języków, które pozwoliłyby mu na korygowanie własnej kompetencji językowej w tym zakresie.

Informacji takich dostarczają **listy słów kluczowych** prezentujące systemy leksykalne języków słów kluczowych „nieswobodnych” (znormalizowanych), ale tylko w zakresie formy wyrażeń – ich treść i w tym wypadku pozostaje częścią kompetencji językowej użytkownika w zakresie semantyki odpowiedniego języka naturalnego.

Nieco więcej informacji o strukturze semantycznej systemu podają nowsze prezentacje języków słów kluczowych, gdzie często listom słów kluczowych towarzyszą **spisy słów kluczowych w układzie gniazdowym**. Explicite pokazują one kategoryzację semantyczną jednostek leksykalnych, a często także elementy struktury paradygmatycznej języka – tak przedstawiony język nie jest już jednak klasycznym językiem słów kluczowych, którego cechą charakterystyczną jest w założeniu płaska struktura leksyki. Jak wykazaliśmy wyżej, paradygmatyka tego języka tkwi implicite w kompetencji użytkownika w zakresie odpowiedniego języka naturalnego, co zawsze ma wpływ na indeksowanie, nieukazana explicite nie może być jednak wykorzystywana przy wyszukiwaniu – języki te w klasycznej postaci nie przewidują bowiem automatycznego wyszukiwania paradygmatycznego.

Jako przykład takiej podwójnej prezentacji systemu leksykalnego języka słów kluczowych podajemy fragmenty słownika słów kluczowych wybranych kategorii kultury (C. Robotycki, W. Babik: *Układ gniazdowy terminów i słownik słów kluczowych wybranych kategorii kultury. Etos, obrzędy, demonologia, magia*. Kraków 2002).

ETOS

Układ gniazdowy terminów

OBYCZAJ

. ZWYCZAJE GRUP RÓWIEŚNICZYCH

. . DZIECI

. . . ZASADA STARSZEŃSTWA

. . . PRACA DZIECI

. . . ZABAWY DZIECIĘCE

. . MŁODZIEŻ ŻEŃSKA [dziewczęta, panny, dziewczki, nastolatki]

. . . GRANICE WIEKOWE

. . . KOLEŻEŃSTWO

. . . . TAJNE JĘZYKI

. . . . ZNAKI PRZYJAŹNI

. . . SZKOŁA I INTERNAT

. . . . GRUPY KOLEŻEŃSKIE

. . . . GWARA UCZNIOWSKA

. . . . JĘZYK GESTÓW

. . . . PAMIĘTNIK

. . . . PELE-MELE

. . . . ZESZYT NASTOLATKI

. . . RYTUAŁY MŁODZIEŻY AKADEMICKIEJ

. . . . IMMATRYKULACJA

. . . . WRĘCZANIE DYPLOMÓW

. . . . JUVENALIA

. . . KONTAKTY Z CHŁOPCAMI

. . . . CHODZENIE Z PARTNEREM

- KONTAKTY SEKSUALNE
- NARZECZEŃSTWO
- . . . ZABAWY TANECZNE
- . . MŁODZIEŻ MĘSKA [kawalerka, parobcy]
- . . . GRANICE WIEKOWE
- . . . KOLEŻEŃSTWO
- . . . PRZYJMOWANIE DO GRUPY
- KONFLIKTY
- PRZEKRACZANIE PROGU DOJRZAŁOŚCI
- TAJEMNICE I ZASADY LOJALNOŚCI
- . . . SZKOŁA I INTERNAT
- . . . SPORT
- . . . RYTUAŁY MŁODZIEŻY AKADEMICKIEJ
- IMMATRYKULACJA
- WRĘCZANIE DYPLOMÓW
- JUVENALIA
- . . . KONTAKTY Z DZIEWCZĘTAMI
- CHODZENIE Z PARTNERKĄ
- KONTAKTY SEKSUALNE
- NARZECZEŃSTWO
- . . . ZABAWY TANECZNE
- . . . SŁUŻBA WOJSKOWA [pójście do wojska]
- POŻEGNANIE REKRUTA
- UNITARKA
- PRZYSIĘGA WOJSKOWA
- FALA
- CYWIL
- POWRÓT Z WOJSKA
- . . . PRYMICJE
- . . GRUPY KONTRKULTUROWE I SUBKULTUROWE
- . . . ANARCHIŚCI
- . . . BLOKERSI
- . . . GITOWCY
- . . . ZIELONI
- . . . HIPISI
- . . . METALOWCY
- . . . NEW AGE
- . . . PACYFIŚCI
- . . . PUNKI
- . . . RASTAFARIANIE
- . . . SKINHEDZI
- . . . ZNAKI GRUP KONTRKULTUROWYCH I SUBKULTUROWYCH
- MUZYKA
- UBIÓR

- ZACHOWANIE
- FANZINY
- . . . KIBICE, Fani, Szalikowcy, Kibole
- FOLKLOR
- OBSCENA
- PIOSENKA
- WIERSZYK
- NIECHĘCI I PRZYJAŹNIE
- WSPÓLNE WYJAZDY

Ukazany tu układ hierarchiczny wyraźnie pokazuje, że struktura paradygmatyczna tego języka słów kluczowych jest inna niż relacje paradygmatyczne zachodzące w języku naturalnym między równokształtnymi ze słowami kluczowymi wyrażeniami języka naturalnego, odzwierciedla bowiem relacje zachodzące w dokumentach (wyrażenia te często współwystępują w tekstach), relacje zachodzące między tematami dokumentów reprezentowanymi przez odpowiednie słowa kluczowe (denotacja bezpośrednia), a pośrednio także relacje zachodzące w rzeczywistości pozatekstowej (pozajęzykowej) między denotatami tych wyrażen (denotacja pośrednia). W wypadku tradycyjnie prezentowanych języków słów kluczowych do ustalenia takiej paradygmatyki języka, znajdującej odzwierciedlenie w charakterystykach dokumentów, użytkownikowi języka posługującemu się nim w procesie wyszukiwania informacji nie wystarczy kompetencja w zakresie języka naturalnego, konieczna jest także odpowiednia wiedza o rzeczywistości, prezentujących ją tekstach oraz strukturze opisującej ją wiedzy (odpowiedniej dziedziny nauki).

W prostych spisach słów kluczowych informacje o ich znaczeniu podawane są, jak zobaczymy w poniższym przykładzie, tylko w niektórych wypadkach. W wypadku, gdy odpowiednie wyrażenie języka naturalnego jest homonimem lub wyrażeniem wieloznacznym, słowo kluczowe opatrywane jest dodatkową informacją, zawierającą pewne wskazówki pozwalające na ustalenie znaczenia. Jednak tylko czasami znaczenie to podawane jest explicite – tak się dzieje, gdy informacja ta podaje wyrażenie języka naturalnego synonimiczne z wyrażeniem języka informacyjno-wyszukiwawczego w znaczeniu reprezentowanym przez dane słowo kluczowe, tak jak w poniższym przykładzie w wypadku słowa kluczowego BABKA : BABKA (akuszerka). Najczęściej podaje się tylko pewne wskazówki co do znaczenia słowa kluczowego, informujące na przykład o dziedzinie wiedzy, w jakiej wyrażenie to jest używane, typie kontekstów, w jakich występuje, czy wskazujące kategorię semantyczną, do jakiej należy. Przykładem może tu być informacja przy drugim słowie kluczowym BABKA, gdzie mamy BABKA (pokrewieństwo).

ETOS

Słownik słów kluczowych

AKADEMIA [ceremonie zawodowe]

AKCEPTACJA

ALKOHOL
ANARCHIŚCI
ANARCHIZM
ANIOŁ PAŃSKI
ANTROPOGONIA
APOKRYF
APOTROPEION
APROBATA
ARCHITEKTURA SAKRALNA
arenda zob. DZIERŻAWA
ATRYBUT
AUTORYTET
AUTORYTET FORMALNY
AUTORYTET WEWNĘTRZNY
AUTORYTET ZEWNĘTRZNY [w stosunku do społeczności]
AXIS MUNDI
BABKA (akuszerka)
BABKA (pokrewieństwo)
BAGNO
BAR
BĘBNISTA [wezwanie do mobilizacji wspólnoty]
BIBLIA DLA UBOGICH
BIERZMOWANIE
biesiada zob. UCZTA
BISKUP
BLOKERSI
BLUŻNIERSTWO
BOGACTWO
BOGACZ [stan społeczny]
boży grób zob. GRÓB CHRYSYTA
BÓG
BÓG – DIABEŁ
BÓJKA
BRACTWO
BUDOWA
BUDOWA DOMU
BUDOWA STUDNI
BUDOWANIE
BUDYNKI
BYTY NADPRZYRODZONE
CADYK
CECHA [znak wójta]
CECHY OSOBOWOŚCI [podstawa autorytetu]
CENTRUM

Szczególnie często takie informacje występują w języku haseł przedmiotowych, gdzie noszą nazwę **dopowiedzeń**. W *Słowniku encyklopedycznym informacji, języków i systemów informacyjno-wyszukiwawczych* w odpowiednim artykule hasłowym podaliśmy następującą definicję i przykłady dopowiedzeń:

„DOPOWIEDZENIE – symbol pomocniczy używany w języku haseł przedmiotowych jako wskaźnik polisemii, objaśnienie lub uzupełnienie tematu lub określnika; stanowi integralną część tematu, tworząc z nim nierozdzieloną całość. Np.: Hel (gaz), Hel (półwysep) i Hel (wieś).

Zwyczajowo wśród dopowiedzeń wyróżnia się **dopowiedzenia lokalizujące** i **dopowiedzenia objaśniające (identyfikujące)**. Dopowiedzenia lokalizujące realizują swą definicyjną funkcję poprzez podanie czasowej i/lub przestrzennej lokalizacji tematu, np.: Żeromski, Stefan (1864-1925), Kraków (Polska) – przykłady z języka haseł przedmiotowych KABA; Poznań (woj. wielkopolskie) – przykład z języka haseł przedmiotowych Biblioteki Narodowej. Dopowiedzenia objaśniające to wszystkie pozostałe dopowiedzenia; najczęściej są to nazwy klas semantycznych różnego typu i poziomu, także nazwy dziedzin wiedzy, w których wyrażenie używane jest w relewantnym znaczeniu, np.: Nihilizm (filozofia), Analiza harmoniczna (matematyka), Rak (choroba), Krakowiak (taniec), Citroën (samochody osobowe), Nowy Świat (ulica) – przykłady z języka haseł przedmiotowych KABA; Piastowie (ród), Kreda (okres geol.), Rehabilitacja (med.), Róg (muz.), Go (gra) – przykłady z języka haseł przedmiotowych BN. Dopowiedzenie może składać się z powtarzalnych elementów zarówno lokalizujących, jak i objaśniających, np.: Octavianus, Augustus (cesarz rzymski ; 63 a.C.-14), Zielony Balonik (Kraków, Polska ; kabaret ; 1905-1915) – język haseł przedmiotowych KABA; Bolesław Chrobry (król Polski, ca 967-1025), Maria Faustyna Kowalska (św. ; 1905-1938), Olimpiada (26 ; 1996 ; Atlanta), Międzynarodowy Kongres Eucharystyczny (46 ; 1997 ; Wrocław), Rogowo (woj. kujawsko-pomorskie, pow. Żnin) – język haseł przedmiotowych BN.”

Dopowiedzenia w paranaturalnych językach informacyjno-wyszukiwawczych, także w języku haseł przedmiotowych i w językach deskryptorowych, zapobiegają przyporządkowaniu przez użytkownika języka, korzystającego w ustalaniu reguł semantyki tego języka z własnej kompetencji językowej w zakresie języka naturalnego, danemu słowu kluczowemu znaczenia nieprzyjętego w danym języku słów kluczowych w wypadku, gdy równokształtne z nim wyrażenie języka naturalnego ma kilka znaczeń, pozwalając tym samym uniknąć przy wyszukiwaniu szumu informacyjnego. Dopowiedzenia są więc jednym ze sposobów eliminacji potencjalnej homonimii.

Homonimy, podobnie jak synonimy, występują w niektórych językach informacyjno-wyszukiwawczych zarówno na poziomie leksykalnym, jak i na poziomie składniowym. Homonimiami są na przykład w języku haseł przedmiotowych wyrażenia Historia, Filozofia, Polska, Encyklopedia, które mogą być albo tematami, albo określnikami, a o ich interpretacji decyduje pozycja w zdaniu, czyli hasle przedmiotowym. Ten typ homonimii jest więc, podobnie jak w języku naturalnym, rozwiązywany kontekstowo. Częstym zjawiskiem jest w językach

informacyjno-wyszukiwawczych homonimia zdań. Tak jest w językach deskryptorowych, gdzie na przykład zdanie:

OPAKOWANIA. POLIETYLEN. NARZĘDZIA

może charakteryzować dokumenty mówiące o: opakowaniach polietylenowych do narzędzi, opakowaniach do narzędzi polietylenowych, narzędziach do opakowań polietylenowych, narzędziach polietylenowych do opakowań. Taka homonimia jest na poziomie języka nierozwiązywalna i jest wysoce niepożądana ze względu na powodowanie szumu informacyjnego (pogorszenie parametru trafności wyszukiwania).

Języki haseł przedmiotowych w swej tradycyjnej postaci w zasadzie prezentowane były w katalogach kartkowych. Informacji o znaczeniu wyrażen było tu nieco więcej niż w tradycyjnych słownikach słów kluczowych, system odsyłaczy ustanawiał bowiem pewne relacje paradygmatyczne, interpretacja znaczenia zależała często od pozycji syntaktycznej wyrażenia w haśle przedmiotowym, ale przede wszystkim od kompetencji użytkownika w zakresie języka naturalnego i jego wiedzy pozajęzykowej. Inaczej się dzieje w wypadku języków haseł przedmiotowych używanych w nowoczesnych, komputerowych systemach informacyjnych, gdzie system języka prezentowany jest w tzw. **kartotekach wzorcowych**. Podają one pewne informacje o znaczeniu wyrażen oraz o ich użyciu przy tworzeniu opisu rzeczowego dokumentów.

Poniżej podajemy przykłady prezentowania znaczenia wyrażenia kultura pochodzące z języka haseł przedmiotowych Biblioteki Narodowej oraz języka haseł przedmiotowych KABA, używanego w bibliotekach uniwersyteckich.

Definicja terminu *kultura* przez postulaty znaczeniowe w artykule przedmiotowym Słownika Języka Haseł Przedmiotowych Biblioteki Narodowej, 2001

temat Kultura

NU Cywilizacja

TK Antropologia społeczna

TK Archeologia

TK Etnografia

TK Konkursy

TK Książka

TK Mecenat

TK Obrzędy

TK Obyczaje

TK Pedagogika kultury

TK Pluralizm społeczny

TK Regionalizm

TK Religia

TK Salon literacko-artystyczny

TK Społeczeństwo

TK Tradycja
TK Wiedza
TK Wychowanie fizyczne i kultura fizyczna
TW Kultura audiowizualna
TW Kultura chrześcijańska
TW Kultura islamu
TW Kultura masowa
TW Kultura materialna
TW Kultura socjalistyczna
TW Kultura życia codziennego
TW Literatura
TW Nauka
TW Oświata
TW Subkultura
TW Sztuka

zt. odpowiednie tematy z określnikiem – kultura, np.: Robotnicy – kultura, Polska – kultura, Żydzi – kultura

określnik kultura

Uwagi określnik – kultura – stosuje się po nazwach grup etnicznych, po nazwach geograficznych, po nazwach ciał zbiorowych i po tematach ogólnych

NU = nie używaj, termin (temat) odrzucony; TK = termin (temat) skojarzony; TW = termin (temat) węższy zakresowe, podrzędny; TS = termin (temat) szerszy zakresowo, nadrzędny.

Definicja terminu *kultura* przez postulaty znaczeniowe w artykule słownikowym w kartotece wzorcowej języka haseł przedmiotowych KABA, 2001

Rekord hasła wzorcowego wyrażenia używanego w podwójnej funkcji tematu i określnika

NP Kultura

Całokształt materialnego i duchowego dorobku ludzkości gromadzony, utrwalany i wzbogacany w ciągu jej dziejów, przekazywany z pokolenia na pokolenie.

TO Culture [f]

TO Culture [a]

zt. określnik – i kultura po nazwach osobowych oraz po nazwach odpowiednich instytucji i organizacji

zt. określnik – wkład do pojęcia kultury po nazwach osobowych

zt. określnik – cywilizacja po nazwach geograficznych

zt. określnik – życie intelektualne po nazwach geograficznych
zt. hasła zaczynające się od wyrazu Kultura
zt. hasła składające się w części z wyrazu Kultura, np. Język i kultura; Dobra kultury
zt. hasła składające się z wyrażenia z przymiotnikiem utworzonym od wyrazu kultura, np. Polityka kulturalna

TK Cywilizacja
TK Etnologia
TK i kultura
TK Język i kultura
TK Polityka kulturalna
TK wkład do pojęcia kultury
TK Życie intelektualne
TP Akulturacja
TP Antropologia i edukacja
TP Chrześcijaństwo i kultura
TP Dobra kultury
TP Edukacja
TP Edukacja międzykulturowa
TP Humanizm
TP Kino i kultura
TP Komunikacja i kultura
TP Mass media i kultura
TP Komunikacja międzykulturowa
TP Obiekty kulturalne
TP Polityka i kultura
TP Poznanie i kultura
TP Stosunki kulturalne
TP Subkultura
TP Wiedza i erudycja
Źródło: RAMEAU

NP = nazwa przyjęta; TO = termin (temat) odrzucony; TK = termin (temat) skojarzony; TP = termin (temat) podrzędny

Paradygmatyczna struktura systemu semantycznego języka informacyjnego z założenia prezentowana jest w wypadku języków deskryptorowych i języków typu klasyfikacji. W obu wypadkach znaczenie wyrażen rozumiane jest przede wszystkim jak relacja wewnątrzsystemowa, zachodząca między jednostkami tego systemu, ale sposób przedstawiania w słownikach pozostałych składników składających się na znaczenie elementarnych jednostek leksykalnych tych języków zależy przede wszystkim od przyjętej w tych językach notacji.

Słowniki języków informacyjno-wyszukiwawczych są słownikami dwujęzycznymi, umożliwiającymi tłumaczenie sformułowanych w języku naturalnym

tekstów charakterystyk słownych dokumentów i zapytań wyszukiwawczych na odpowiadające im teksty języka informacyjno-wyszukiwawczego, a więc charakterystyki wyszukiwawcze i instrukcje wyszukiwawcze. W wypadku języków paranaturalnych strona języka naturalnego pojawia się jednak tylko w określonych wypadkach, wtedy, gdy wyrażenie języka naturalnego, które potencjalnie mogłoby stać się wyrażeniem języka informacyjno-wyszukiwawczego, jest w tym języku reprezentowane przez inne wyrażenie, często synonimiczne z nim w języku naturalnym lub szersze albo węższe zakresowo. W języku słów kluczowych i w języku haseł przedmiotowych wyrażenia takie połączone są z odpowiednimi wyrażeniami języka informacyjno-wyszukiwawczego odsyłaczami całkowitymi (zob., – , stosuj lub używaj), np.:

Lingwistyka zob. Językoznawstwo,
Pies stos. Psy,
PAN – Polska Akademia Nauk.
Pistolet zob. Broń palna.

W językach deskryptorowych takie wyrażenia języka naturalnego nazywa się askryptorami i umieszcza na odpowiedniej pozycji w artykule deskryptorowym.

Wyrażenia języka naturalnego towarzyszą zawsze odpowiadającym im wyrażeniom języka informacyjno-wyszukiwawczego w językach typu klasyfikacji – tzw. odpowiedniki słowne towarzyszą symbolom klasyfikacyjnym, na przykład w UKD

821.111 Literatura angielska;
331.105.44: 371.12 Związek zawodowy nauczycieli [Organizacje pracowników. Nauczyciele];
159.9 Psychologia;
16 Logika. Teoria poznania. Metodologia;
629.5 Technika środków transportu wodnego. Technika okrętowa. Budownictwo okrętowe. Łodzie. Statki. Okręty.

Przykłady te wyraźnie pokazują, że wyrażenia języka naturalnego są nie tylko odpowiednikami słownymi symboli klasyfikacyjnych, ale pełnią swoją podstawową, metajęzykową funkcję względem wyrażeń języka informacyjnego, prezentują zakres klasy, czyli znaczenie symbolu klasyfikacyjnego.

W wypadku języków o strukturze hierarchicznej, takich jak klasyfikacje i języki deskryptorowe należy pamiętać, że istotna część informacji o znaczeniu ich wyrażeń zawarta jest w ich strukturze paradygmatycznej, układach prezentowanych w tezaurusach w artykułach deskryptorowych oraz w tablicach klasyfikacyjnych. Dlatego przy korzystaniu z języków informacyjno-wyszukiwawczych, a więc w procesie tworzenia charakterystyk wyszukiwawczych i instrukcji wyszukiwawczych nigdy nie można się posługiwać tylko informacją o znaczeniu zawartą w odpowiednich wyrażeniach języka naturalnego, bo zawsze będzie to tylko część znaczenia, wyznaczana przez znaczenie wyrażeń odpowiedniego języka naturalnego modyfikowane przez kompetencję semantyczną konkretnego

użytkownika w zakresie konkretnego języka naturalnego. Trzeba o tym pamiętać zwłaszcza w wypadku międzynarodowych języków informacyjno-wyszukiwawczych mających swoje wersje narodowe z odpowiednikami słownymi w różnych językach naturalnych (np. UKD) oraz w wypadku adaptacji języków informacyjno-wyszukiwawczych bazujących na językach naturalnych (np. polski język haseł przedmiotowych KABA, będący adaptacją francuskiego języka RAMEAU oraz amerykańskiego LCSH), bo jak wiemy, każdy język naturalny inaczej klasyfikuje rzeczywistość, a dotyczy to także terminologii naukowej, która mimo zbieżności form w różnych językach bywa nośnikiem różniących się treści.

Jako ilustrację prezentowania struktury paradygmatycznej języków informacyjno-wyszukiwawczych podajemy przykłady artykułów deskryptorowych z mikrotezauryusa *Nauka*, opracowanego w Bibliotece Sejmowej przez Ewę Chmielewską-Gorczycę oraz fragment tablic klasyfikacyjnych UKD prezentujący część klasy 81 Lingwistyka. Języki (fragment ten wybraliśmy ze względu na temat podręcznika).

W tezaurysie *Nauka* przyjęto następujące oznaczenia:

BT – termin szerszy (ang. Broader Term)

FT – termin obcojęzyczny – ang., fr., niem. (ang. Foreign Term)

NT – termin węższy (ang. Narrower Term)

RT – termin kojarzeniowy (ang. Related Term)

TT – termin będący nazwą działu (ang. Top Term)

UF – askryptor, tj. termin odsyłający do danego deskryptora (ang. Used For)

* – odpowiednik terminu (ang., fr., niem.) w Tezaurysie EUROVOC Parlamentu Europejskiego

nauka (70.0)

UF: dziedziny nauki

NT: nauki fizyczne

nauki humanistyczne

nauki o ziemi

nauki o życiu

nauki stosowane

polityka naukowa

technika

teoria

RT: badania naukowe

biblioteka naukowa

dokument dotyczący oświaty i nauki

futurologia

kultura

metodologia nauk

oświata

oświecenie

postęp naukowy

pracownicy naukowci
prasa naukowa
przedmioty szkolne
renesans
terminologia
wolność nauki
współpraca naukowa
własność intelektualna
FT: science (ang.)
sciences (fr.)
Wissenschaften (niem.)

informacja naukowa (76.0)

UF: nauka o informacji
BT: nauki stosowane
NT: bibliotekoznawstwo
dokument
dokumentacja
działalność informacyjna
polityka informacyjna
system dokumentacyjny
system informacyjny
RT: biblioteka naukowa
informacja
informatyka
informatyka stosowana
kadry informacyjne
przetwarzanie informacji
FT: information science * (ang.)
science de l'information * (fr.)
Informationswissenschaft * (niem.)

informatyka (75.0)

BT: nauki stosowane
NT: informatyka stosowana
prawo komputerowe
przemysł informatyczny
przetwarzanie danych
system informatyczny
zawód informatyka
RT: automatyzacja
elektronika
informacja
informacja naukowa

mikroelektronika
nauczanie wspomagane komputerem
programy edukacyjne
przemysł elektroniczny
technika

FT: computer systems * (ang.)
informatique * (fr.)
Informatik und Datenverarbeitung * (niem.)

językoznawstwo (77.0)

UF: lingwistyka
BT: nauki społeczne
NT: język
nazwy własne
socjolingwistyka
terminologia
transliteracja
tłumaczenie
wielojęzyczność

RT: język obcy
język ojczysty
literatura
nauka języka
semiotyka
tekst

FT: linguistics * (ang.)
linguistique * (fr.)
Sprachwissenschaft * (niem.)

metodologia nauk

TT: 70.2 FILOZOFIA
BT: filozofia
RT: badania
hermeneutyka
logika
metoda badawcza
nauka
prakseologia
statystyka
strukturalizm
teoria poznania

FT: methodology of sciences (ang.)
méthodologie d'une science (fr.)
Methodenlehre der Wissenschaften (niem.)

Tablice główne działu 8 Językoznawstwo. Literatura

- 8 JĘZYKOZNAWSTWO. NAUKA O LITERATURZE.
LITERATURA PIĘKNA**
- 80 OGÓLNE ZAGADNIENIA JĘZYKOZNAWSTWA I LITERATURY.
FILOLOGIA**
- 801 Prozodia. Nauki pomocnicze i źródła filologii
- 801.6 Prozodia: metryka, rytm, rym i wersyfikacja
Prozodię poszczególnych języków klasyfikuje się symbolami złożonymi, dodając po dwukropku symbol odpowiedniego języka lub literatury, np.:
801.6:811.162.1 Prozodia polska - badania językoznawcze
801.6:821.162.1 Prozodia polska - badania literackie
- 801.7 Nauki pomocnicze filologii
Hermeneutyka. Krytyka tekstów. Egzegeza. Tekstologia
- 801.8 Źródła językoznawcze
- 808.2 Edytorstwo. Wydawanie tekstu. Technika edytorska
Redaktorstwo
Ruch wydawniczy. Działalność wydawnicza zob. 655.41
- 808.5 Retoryka. Sztuka przemawiania
Sztuka oratorska. Sztuka i technika deklamacji i recytacji. Dykcja. Konwersacja
- 81 JĘZYKOZNAWSTWO. JĘZYKI**
Symbole poddziałów analitycznych -1 11/-26 oraz '0/'44 służą do klasyfikowania określonych zagadnień językoznawczych w ujęciu ogólnym lub w odniesieniu do określonego języka
- 81:1 Filozofia języka
- 81:39 Etnolingwistyka
- 81(091) Historia języka (ogólna)
Historia poszczególnych języków zob. 811
- 81-11 Szkoły i kierunki w językoznawstwie
Językoznawstwo diachroniczne (historyczne), językoznawstwo synchroniczne (opisowe), strukturalizm (językoznawstwo strukturalne)
- 81-115 Językoznawstwo porównawcze
Zob. też:
81 '362 Gramatyka porównawcza
- 81-2 Charakterystyczne cechy języka
- 81-21 Języki prymitywne (ludów pierwotnych)
- 81-23 Języki żywe
- 81-24 Języki martwe, wygasłe
- 81-25 Języki potoczne (mówione)
- 81-26 Języki pisane. Języki literackie

- Np.:
 811.162.1-26 Polski język literacki
- 81'0 Pochodzenie i okresy rozwoju języka
- 81'01 Okres archaiczny
 Prajęzyk, praformy języka, stary język
- 81'02 Okres klasyczny
- 81'04 Okres średniowieczny
- 81'06 Okres nowożytny, współczesny
- 81'1 Językoznawstwo ogólne
 Zagadnienia dotyczące poszczególnych języków klasyfikuje się symbolem danego języka
- 81'22 Ogólna teoria znaków w odniesieniu do językoznawstwa.
 Semiologia. Semiotyka
- 81'221 Niewerbalne komunikowanie się
 Język mimiki i gestykulacji, język głuchoniemych
 Zob. też:
 811.9 Języki sztuczne
- 81'23 Psycholingwistyka. Psychologia języka
 Język a myślenie. Neurolingwistyka. Patologia mowy. Afazja
- 81'24 Praktyczna znajomość języków
 Rozmówki językowe klasyfikuje się symbolem odpowiedniego języka, dodając symbol analityczny '24 i symbol podziału wspólnego formy (078), np.:
 811.111'24(078) Rozmówki angielskie
 Nauczanie języków w szkole zob. 372.881.1
- 81'246 Jednojęzyczność i wielojęzyczność. Poliglottyzm
- 81'25 Teoria przekładu (tłumaczenia)
 Translatoryka. Tłumaczenia ustne, tłumaczenia pisemne.
 Tłumaczenia literackie
- 81'26 Kontrola języka. Planowanie języka. Standaryzacja języka
- 81'27 Socjolingwistyka. Stosowanie języka
- 81'271 Poprawność językowa. Błędy językowe
- 81'276 Dialekty społeczne. Formy potoczne języka. Slang
 Język (idiomy, żargony) określonych środowisk i grup społecznych
 Dalszy podział przez dodanie symbolu podziału wspólnego osoby -05 lub po dwukropku odpowiedniego symbolu głównego, np.:
 81'276-053.5 Język (żargon) uczniowski
 81'276:34 Język prawniczy
 81'276:355.11 Język (slang) żołnierski
- 81'276.2 Język wulgarny
- 81'28 Dialektologia
- 81'282 Dialekty. Język lokalny i regionalny. Warianty językowe
 Gwary i narzecza

- Dialekty, gwary i narzecza poszczególnych języków klasyfikuje się, dodając symbol poddziału analitycznego '282 do symbolu odpowiedniego języka z działu 811 i ewentualnie poddziału wspólnego miejsca, np.:
811.162.1'282 (574) Język Polaków w Kazachstanie
- 81'282.8 Języki konwencjonalne. Języki mieszane
Gwara międzynarodowa
Np.: Pidgin-English
- 81'32 Językoznawstwo matematyczne
Językoznawstwo cyfrowe (numeryczne). Przekłady maszynowe. Maszynowa analiza i synteza mowy. Językoznawstwo statystyczne
- 81'34 Fonetyka. Fonologia
Fizjologia mowy zob. 612.78
- 81'34:534 Fonetyka akustyczna
- 81'342.8/.9 Akcent. Akcentowanie. Intonacja
Zob. też:
801.6 Prozodia: metryka, rytm, rym i wersyfikacja
- 81'35 Grafemika. Ortografia. Sylabizowanie
Interpunkcja
- 81'35'374 Słowniki ortograficzne
- 81'355 Wymowa. Wymawianie
- 81'36 Gramatyka
Gramatyka generatywna
- 81'36-112 Gramatyka historyczna
- 81'362 Gramatyka porównawcza
Zob. też:
81-115 Językoznawstwo porównawcze
- 81'366 Morfologia. Fleksja
- 81'367 Składnia. Syntaktyka
- 81'367.6 Części mowy
- 81'37 Semantyka
Zob. też:
510.27 Semantyka logiczna
- 81'373 Leksykologia
- 81'373.2 Nazewnictwo. Onomastyka
- 81'373.21 Nazwy geograficzno-topograficzne. Toponomastyka. Toponimia
- 81'373.22 Nazwy obiektów przyrodniczych (inne niż geograficzne i topograficzne)
Nazwy roślin, zwierząt, skał, minerałów, planet itd.
- 81'373.23 Nazwy osób i grup osób. Antroponimia
- 81'373.42 Wyrazy zbliżone brzmieniem lub znaczeniem
- 81'373.421 Synonimy
- 81'373.421'374 Słowniki synonimów
- 81'373.422 Antonimy
- 81'373.423 Homonimy

81'373.43	Neologizmy
81'373.44	Archaizmy
81'373.44'374	Słowniki archaizmów
81'373.45	Wyrazy obce. Wyrazy zapożyczone. Kalki językowe
81'373.45'374	Słowniki wyrazów obcych
81'373.46	Nazwy (terminy) oznaczające tematy specjalne Terminy (zwroty) techniczne
81'373.47/.49	Przekleństwa. Wyrazy ordynarne. Wyrazy modne. Eufemizmy
81'373.6	Etymologia
81'373.6'374	Słowniki etymologiczne
81'373.7	Frazeologia
81'373.7'374	Słowniki frazeologiczne
81'373.72	Idiomy. Wyrażenia idiomatyczne
81'374	Leksykografia. Słowniki: kompilacja i zawartość Słowniki językowe klasyfikuje się symbolem odpowiedniego języka Słowniki wielojęzyczne klasyfikuje się symbolem złożonym, stosując odpowiednio dwukropkę pojedynczy bądź podwójny, np.: 811.162.1::811.133.1]'374 Słownik polsko-francuski 811.162.1:811.133.1]'374 Słownik polsko-francuski i francusko-polski Słowniki dziedzinowe klasyfikuje się symbolem dziedziny, dodając symbol podziału wspólnego formy (038), np.: 33(038) Słownik ekonomiczny Słowniki wielojęzyczne dziedzinowe klasyfikuje się symbolem dziedziny, dodając symbole podziałów wspólnych języka, np.: 54(038)=00 Słownik chemiczny wielojęzyczny 621.3(038)=111=162.1 Słownik elektroniczny angielsko-polski
81'38	Stylistyka ogólna
81'42	Analiza mowy
81'44	Językoznawstwo typologiczne Typologia języków

Literatura:

- Artowicz E., Chmielewska-Gorczyca E.: *Wykorzystanie aparatu definicyjnego języków informacyjno-wyszukiwawczych w słownikach terminologicznych*. „Zagadnienia Informacji Naukowej” 1986, nr 1.
- Bojar B.: *Pole informacyjne, dokumentacyjne, tematyczne, semantyczne czy leksykalne?* „Zagadnienia Informacji Naukowej” 2004, nr 4.
- Chmielewska-Gorczyca E.: *Problem synonimii w teorii języków informacyjno-wyszukiwawczych*. „Zagadnienia Informacji Naukowej” 1975, nr 2.
- Marciszewski W.: *Relacje tematyczno-hierarchiczne w językach deskryptorowych*. „Zagadnienia Informacji Naukowej” 1973, nr 2.
- Sadowska J.: *Język hasel przedmiotowych Biblioteki Narodowej. Studium analityczno-porównawcze*. Warszawa 2003.

- Sosińska B.: *Struktura wyrażenia języków informacyjno-wyszukiwawczych a konotacyjna i denotacyjna interpretacja ich znaczenia*. „Zagadnienia Informatyki Naukowej” 1981, nr 2.
- Sosińska B.: *Typologia relacji pomiędzy jednostkami leksykalnymi języków informacyjnych*. „Zagadnienia Informatyki Naukowej” 1979, nr 2.
- Tokarski R.: *Struktura pola znaczeniowego*. Warszawa 1984.
- Sosińska-Kalata B.: *Podręcznik UKD dla bibliotekarzy i pracowników informacji*. Warszawa 1995.
- Wereszczyńska-Cisło B.: *Relacje kojarzeniowe w językach informacyjnych*. „Zagadnienia Informatyki Naukowej” 1985, nr 1.
- Wereszczyńska-Cisło B.: *Wykorzystanie relacji skojarzeniowych specyfikowanych w tezaurysie do określania hipertekstowych powiązań*. „Zagadnienia Informatyki Naukowej” 1997, nr 2.
- Woźniak J.: *Kategoryzacja. Studium z teorii języków informacyjno-wyszukiwawczych*. Warszawa 2000.

XVI

Tekst

Najmniejszą jednostką informacji jest zdanie, ale pojemność informacyjna jednego zdania zazwyczaj nie wystarcza nadawcy do przekazania odbiorcy całej informacji, jaką chciałby przekazać o danym wycinku rzeczywistości. Dlatego w procesie informacyjnym spotykamy się najczęściej nie z pojedynczymi zdaniami, lecz z całymi **tekstami** stanowiącymi połączenia wielu zdań. Oczywiście i na tym poziomie języka istnieją specjalne prawa składania zdań tak, aby można je było uznać za tekst spójny, nie każde bowiem połączenie zdań uznamy za poprawne w danym języku.

Nim przejdziemy do omawiania gramatyki tekstu spójnego, warto się może zastanowić nad tym, czym są **teksty spójne**. Okazuje się bowiem, że termin ten bywa rozmaicie rozumiany, często zresztą zupełnie intuicyjnie. Gdy przyjrzymy się bliżej różnym określeniom tekstów spójnych, to dojdziemy do wniosku, że właściwie w większości z nich występuje tylko jeden wspólny element – stwierdzenie, że tekst składa się ze zdań, co automatycznie wytycza kierunek poszukiwania reguł spójności tekstu jako reguł dotyczących łączenia zdań w większe całości. Stwierdzenie to występuje jednak nie we wszystkich definicjach tekstu spójnego, bo na przykład Olgierd A. Wojtasiewicz w artykule *O interpretacji pojęcia spójności tekstu* nie przyjmuje takiego założenia – według zaproponowanej przez niego interpretacji można by rozpatrywać spójność jednego zdania lub nawet sekwencji wyrazów niestanowiących zdania, a więc syntagmy.

W literaturze przedmiotu stawia się również pytanie (por. J. Wajszczuk: *Tekst spójny czy po prostu tekst?*), czy w ogóle termin tekst spójny jest terminem poprawnym, czy nie należałoby zamiast mówić o tekstach spójnych i tekstach niespójnych, mówić po prostu o jednym tekście i kilku tekstach. Odpowiedź na tak postawione pytanie wymaga rozstrzygnięcia, czy spójność tekstu jest jego wymogiem definicyjnym, a więc własnością, która decyduje, czy ciąg zdań nazwiemy tekstem, czy też nie. Próbując znaleźć odpowiedź na to pytanie, Jadwiga Wajszczuk tak pisze: „Konsekwentna redukcja wyrażenia tekst spójny do po prostu tekst wydaje się jednak zabiegiem sztucznym. Poprawność (sensowność) będącego w powszechnym użyciu wyrażenia „tekst spójny” w opozycji do – zauważmy – niesprzecznego wewnątrznie „tekst niespójny” znajduje być może uzasadnienie

w tym, że właściwości, o których mowa, są najwidoczniej w powszechnym odczuciu stopniowalne, przysługują tekstom w mniejszym lub większym stopniu (por. wyrażenie „tekst bardzo spójny”, „mało spójny”, „bardzo niespójny”). Negacja w wyrażeniu „tekst niespójny”, sądzić można, nie ma charakteru absolutnego, tak określane są teksty znacznie mniej spójne niż inne, teksty o „niskim poziomie spójności”. Biegun tej skali stanowi określenie „tekst całkowicie niespójny”. Czym tekst całkowicie niespójny różni się od nie-tekstu, tj. ciągu przypadkowo zestawionych zdań, zdań nie powiązanych? Czy tekst całkowicie niespójny to wciąż jeden tekst?”.

Rozstrzygnięcie tego problemu pociąga za sobą daleko idące konsekwencje, jeśli bowiem się uzna, że spójność nie jest cechą definicyjną tekstu, a więc, że mogą istnieć teksty spójne i teksty niespójne, to wtedy oddzielnego zdefiniowania wymaga sam tekst, oddzielnego zaś cecha jego spójności. W tym kierunku zmierza na przykład artykuł Andrzeja Bogusławskiego *Słowo o zdaniu i o tekście*, który, podając swoją definicję tekstu zrelatywizowaną do odbiorcy, a rozumianego jako pewien uporządkowany ciąg zdań (nie miejsce tu na jej przytoczenie i dyskusję nad nią), pisze, że „Teksty rozumiane zgodnie z przedłożonymi propozycjami mogą być, tak samo jak zdania, sprzeczne wewnątrznie, niespójne różnymi rodzajami niespójności, a także spójne w rozmaitych sensach”, podając przy tym przykład:

*Zemlja imeet formu šara
Priyatno videt' nos komara*

który według niego może być uznany za tekst, choć oczywiście byłby to tekst jakoś niespójny.

Jak więc widać, spójność może być rozpatrywana jako cecha przysługująca tylko pewnym tekstom i może to być cecha stopniowalna. Tak traktują ją również logicy, na przykład Witold Marciszewski, który w artykule *Spójność strukturalna a spójność semantyczna* pisze: „Chcę uzasadnić pogląd, że spójność należy traktować nie jako własność dzielącą rozłącznie zbiór tekstów na spójne i niespójne, ale jako własność stopniowalną. Stopniowalność zachodzi między pewnymi typami spójności, a w obrębie wyróżnionych tu typów można będzie mówić o tekstach mniej lub bardziej spójnych ze względu na stopień spełnienia kryteriów obowiązujących w danym etapie. Jednocześnie chcę wykazać (...), że stopniowalność cechuje zarówno strukturalne warunki spójności, jak i warunki semantyczne, które stanowią niezbędne uzupełnienie strukturalnych”. Marciszewski odróżnia więc **spójność syntaktyczną** od **spójności semantycznej**, postulując, jak zobaczymy dalej, różne typy spójności. Niektórzy autorzy zajmujący się tą dziedziną mówią jednak raczej tylko o jednym typie spójności, czego przykładem może być chociażby następujące zdanie Jeleny Paduczewej z artykułu *O strukturze akapitu*: „Ograniczenia łączliwości zdań w akapicie mogą być związane z odróżnieniem sensownego (spójnego) tekstu od bezsensownego nagromadzenia zdań (...)”.

Podobne do Marciszewskiego stanowisko reprezentuje Kazimierz Trzęsicki w artykule *Problem spójności tekstu w związku z aksjomatyczną teorią tekstu*,

w którym przedstawia definicję tekstu nicodwołującą się do definicji spójności, wyróżniając przy tym dwa rodzaje spójności: spójność syntaktyczną i semantyczną. U podstaw koncepcji Trzęsickiego leży teoria tekstu Romana Ingardena, a więc za Ingardenem odrzuca on koncepcję tekstu jako układu złożonego ze zbioru zdań i relacji liniowo porządkującej ten zbiór, pisząc: „może być tak, iż dany związek między zdaniami zachodzi niezależnie od kolejności, w jakiej te zdania zostały umieszczone w zapisie tekstu. Tzn. może być tak, że np. zdanie B jest znaczeniowo zależne od poprzedzającego go w fizycznie utrwalonym tekście zdania A. Zamiana miejscami tych zdań, a więc wpisanie w miejsce zdania A zdania B, i na odwrót, wcale nie narusza sposobu i typu zależności znaczeniowej tych zdań. W dalszym ciągu zdanie B jest tak samo i w ten sam sposób zależne znaczeniowo od zdania A. A więc zamiana miejscami tych zdań wcale nie zmieniła tekstu, jest to tylko zmiana zapisu tego samego tekstu. Stało się tak dlatego, że znaczenie przysługujące zbiorowi zdań tworzących tekst nie uległo zmianie (...). Tekst jest bowiem układem złożonym za zbioru zdań i znaczenia przyporządkowanego temu zbiorowi. Znaczenie przysługujące w danym tekście danemu zbiorowi zdań jest wyznaczane przez szereg zdań i związki pomiędzy tymi zdaniami”.

Tekst Trzęsicki definiuje, posługując się wprowadzonym uprzednio pojęciem operacji **tematyzacji**, rozumianej jako „operacja, która polega na uzupełnianiu zbioru zdań będącego jej argumentem o wszystkie te i tylko te zdania z najmniejszej całości znaczeniowej zawierającej ten zbiór – tak, aby w otrzymanym zbiorze sensy wszystkich jego elementów (zdań) były takie, jakie przysługują tym zdaniom w tej najmniejszej całości znaczeniowej. A więc tematyzacja polega na dodaniu do zbioru zdań argumentu operacji takiego możliwie najmniejszego kontekstu z najmniejszej całości znaczeniowej zawierającej ten zbiór, aby wszystkie zdania, wraz z tymi dołączonymi, miały taki sens, jaki im przysługuje w zawierającej ów zbiór najmniejszej całości znaczeniowej (...). Operacja tematyzacji polega na dołączeniu do danego zbioru zdań wszystkich i jedynie tych zdań z najmniejszej całości znaczeniowej zawierającej ten zbiór, od których są one znaczeniowo zależne, które to zdania same nie są zależne od żadnych innych zdań w tej całości znaczeniowej”. Założenie to prowadzi do definicji tekstu mówiącej, że „**tekst** jest zamkniętą całością znaczeniową, a znaczenie poszczególnych zdań jest ostatecznie wyznaczone w tekście jako całości i żadne ze zdań ze zbioru T nie jest znaczeniowo zależne, w danym tekście, od żadnych innych zdań (spoza zbioru T)”. Tekst jest więc rozumiany jako układ złożony ze zbioru zdań T i znaczenia przyporządkowanego temu zbiorowi, jest to według autora taki zbiór zdań, który jako całość nadaje się na cytaty. Dla tak określonego tekstu Trzęsicki określa cechę jego **spójności syntaktycznej**, rozumianej jako faktyczne zachodzenie związków między zdaniami – związków wyznaczonych przez operację **tematyzacji**, czyli tych, które są konieczne, aby tekst był uznany za skonstruowany zgodnie z zasadami budowy tekstu, mówiąc inaczej – zgodnie z zasadami **gramatyki tekstu**. Spójność syntaktyczna jest tu warunkiem **spójności semantycznej**, przy czym „za tekst spójny semantycznie uznawać się będzie taki i tylko taki tekst, któremu daje się przyporządkować pewien możliwy świat, w jakim ten tekst jest spełniony,

tn. zawierający wszystkie przedmioty, o których jako o istniejących mówi się w tekście; świat, w jakim z tymi przedmiotami zdarza się wszystko to, o czym mówi się w tekście, że się zdarza. W ten sposób pojęcie semantycznej spójności tekstu zostało zrelatywizowane do ontologii. Znaczy to więc, że może zdarzyć się tak, iż pewien tekst ze względu na jeden system ontologii jest spójny semantycznie, a ze względu na inny nie jest spójny semantycznie. W przypadku ontologii, w której przyjmuje się, że każdy niesprzeczny świat jest możliwy, pytanie o spójność semantyczną tekstu daje się sprowadzić do pytania tylko o pewne własności syntaktyczne tego tekstu”.

Koncepcji Trzęsickiego poświęciliśmy tu trochę więcej miejsca również dlatego, że jest ona przykładem traktowania w pewien sposób rozumianej spójności nie jako cechy samego tekstu pojmowanego jako struktura pewnego rodzaju wyrażen, a więc nie jako jego cechy immanentnej, wynikającej ze spełniania przezeń pewnych reguł językowych, lecz jako cechy, której przyznawanie tekstowi uzależnione jest od spełnienia przezeń pewnych warunków zewnętrznych – tu, poprzez odniesienie go do jakiejś innej niż on sam rzeczywistości, owego **świata możliwego**. Jest to jednak kryterium innego typu niż kryteria, które wystarczały do uznania wyrażen złożonych występujących na niższych poziomach struktury języka za wyrażenia danego języka, choć niektórzy językoznawcy, jak na przykład Chomsky w swej gramatyce generatywnej, postulowali wprowadzenie takich kryteriów już na poziomie zdań. Chomsky twierdził, iż gramatyka generatywna nie powinna dopuszczać tworzenia zdań bezsensownych typu „*Zielone bezbarwne idee śpią wściekle*” (logicy robili to już wcześniej, na jeszcze niższym poziomie struktury języka, wyróżniając tzw. nazwy puste).

Przydatność takich kryteriów do oceny spójności tekstu czy też poprawności zdań wydaje mi się wątpliwa i należałoby chyba dążyć do formułowania jedynie kryteriów formalnych, sądę bowiem, że uciekanie się do koncepcji światów możliwych jest zabiegiem czysto teoretycznym, przesuwającym pytanie o to, czym jest spójność tekstu na inną płaszczyznę. Tym bardziej, że na taki możliwy świat nakłada się jedynie warunek jego niesprzeczności, a przysługiwanie cechy niesprzeczności – i to nawet konkretnemu zdaniu – bywa trudne do stwierdzenia wobec możliwości rozumienia znaczenia występujących w nim wyrażen bądź jako ich denotacji, bądź jako konotacji. Przykładem może tu być chociażby ocena zdania *Atom można rozbić* (w chwili, gdy pojawiło się ono po raz pierwszy), jeżeli bowiem słowo *atom* rozumieć będziemy zgodnie z jego pierwotnymi cechami konotacyjnymi, wprowadzonymi definicją projektującą, jako najmniejszy element materii, zdanie to będzie nonsensem, choć nie będzie nim, jeżeli przy naszej ocenie weźmiemy pod uwagę jedynie denotację tego wyrażenia.

Na owych zewnętrznych kryteriach opierają się też inne definicje spójności tekstu, jak chociażby definicja podana przez Zygmunta Saloniego w artykule *Definicja spójności tekstu*, która brzmi: „Tekst jest spójny, o ile nie istnieje takie drugie, różne od autentycznego, uszeregowanie zdań, które nie różni się znaczeniowo od niego, oraz trzecie uszeregowanie, otrzymane z drugiego przez przesunięcie początkowego odcinka tekstu na jego koniec, również równoznaczne znaczenio-

wg”. Definicja ta nawiązuje do topologicznej definicji przestrzeni spójnej, a wynikają z niej następujące konsekwencje (przytaczam je z pewnym komentarzem):

- „O spójności można mówić tylko w wypadku, gdy mamy do czynienia z tekstem, a więc ciągiem poprawnych zdań. Jeżeli choćby jedno ze zdań w ciągu nie jest poprawne, nie mamy poprawnego tekstu i zastanawianie się nad jego poprawnością traci sens”. Jest to warunek dość mocny, postępując zgodnie z nim, nauczyciel sprawdzający wypracowanie ucznia mógłby, znajdując jedno niepoprawnie zbudowane zdanie, całość ocenić jako niespójną. Intuicyjnie wydaje się, że tak nie czynimy i skłonni jesteśmy uznać za spójny (w potocznym rozumieniu) tekst, w którym może być zdanie niepoprawne, na przykład:

Warszawa jest coraz piękniejsza. Serce rośnie, patrząc na to miasto. Miasto, które Hitler skazał na wymazanie z mapy świata.

Autor definicji nie definiuje terminu poprawne zdanie, co utrudnia interpretację wywiedzionych przez niego konsekwencji. W konsekwencji tej zawarta jest definicja tekstu, który rozumiany jest jako ciąg poprawnych zdań, spójność nie jest więc, w tym rozumieniu, cechą konstytutywną tekstu.

- „Każde zdanie jest spójne, bo istnieje tylko jedna permutacja zbioru jednoelementowego” – spójność jest więc cechą służącą do oceny wyrażen języka dopiero na poziomie tekstu (rozumianego jako ciąg zdań), nie można więc mówić o niespójnych zdaniach.

- „Fragment tekstu spójnego może być niespójny, np. zdania z *Popiołów* :

Jakież dziwny, niespodziewany żal!... Gdzież jest ten dzień, gdzie się na wieki został?...

są niespójne, bo bez zmiany znaczenia można przestawić ich kolejność, choć całe *Popioły* stanowią tekst spójny” – czym jest wobec tego spójność tekstu złożonego z niespójnych fragmentów? – na to pytanie brak odpowiedzi, choć jest to jakaś inna spójność niż wyżej zdefiniowana.

- „Do oceny spójności tekstu nie zawsze jest konieczne jego jasne rozumienie. Na przykład, jeśli przechodzień na ulicy powie do nas (przykład O.A. Wojtasiewicza – B.B.) :

Gdzie mogę dostać paratus na dwójkę. Przyjechałem specjalnie z Grójca. Tymczasem nie wiem, dokąd się mam udać.

uznamy, że wygłosił tekst spójny, choć nie wiemy, czym jest rzecz, o której mówi”.

- „Przy badaniu tekstu należy brać pod uwagę tylko jego znaczenie, nie zaś znaczenie nadawane przez elementy sytuacyjne. Musimy więc stwierdzić, że tekst:

Strzał. Co tam jest? Ale wrzask na stadionie. A Rumun leży.

jest niespójny, choć w sprawozdaniu z meczu może być dla słuchacza wystarczający i dobrze zrozumiały”.

Ten ostatni przykład byłby tekstem spójnym w koncepcji spójności Olgierda A. Wojtasiewicza, wyłożonej w artykule *O interpretacji pojęcia spójności tekstu*. Nie

będziemy tu przytaczać formalnego zapisu warunków spójności tekstu, mówią one w każdym bądź razie, że „jeżeli dla jakichś elementów zbioru X zachodzi to, że każdy z tych elementów ma swój opis w odpowiednim elemencie tekstu, i jeśli między tymi elementami zbioru X zachodzi relacja R, to ciąg elementów tekstowych t_1, \dots, t_n jest spójny”. Spójność tekstu jest tu więc uzależniona od spójności opisywanej przez ten tekst rzeczywistości, jest to więc koncepcja zbliżona do koncepcji oceniania spójności tekstu przez odniesienie go do tzw. światów możliwych. Według Wojtasiewicza tekst:

„...zaskakujący strzał Brychczego – Kostka z trudem piąstkuje – piłka wychodzi na róg...”

jest spójny, ponieważ taka sekwencja wydarzeń w meczu piłki nożnej jest możliwa (przy czym zgodność takiego opisu z faktycznym przebiegiem wydarzeń jest sprawą inną, a w pewnych wypadkach całkowicie irrelevantną). Koncepcja Wojtasiewicza nie zakłada przy tym, że zdanie jest *ex definitione* spójne i że o spójności lub niespójności można mówić dopiero wówczas, gdy ma się do czynienia z sekwencją zdań – można rozpatrywać spójność jednego zdania lub nawet sekwencji wyrazów niestanowiącej zdania.

Do kryterium **znaczenia tekstu** w ocenie jego spójności odwołuje się również Irena Bellert w artykule *O pewnym warunku spójności tekstu*, przyjmując „następującą definicję roboczą tekstu spójnego, czyli **dyskursu**, zgodną z intuicyjnym rozumieniem tego terminu. Dyskurs jest to taki ciąg wypowiedzeń S_1, \dots, S_n , w którym interpretacja semantyczna każdego wypowiedzenia S_i (dla $2 \leq i \leq n$) jest zależna od interpretacji wypowiedzeń w ciągu S_1, \dots, S_{i-1} . Innymi słowy, adekwatna interpretacja wypowiedzenia występującego w dyskursie wymaga znajomości poprzedzającego kontekstu”. Autorka przyjmuje również, że interpretacja semantyczna dowolnego wypowiedzenia to zbiór konsekwencji, czyli wniosków, które można wyciągnąć na podstawie tego wypowiedzenia.

Maria Renata Mayenowa w pracy *Spójność tekstu a postawa odbiorcy* również nakłada na tekst spójny warunki zewnętrzne. Píše ona:

„Tekst spójny opiera się na trzech jednościach.

Jest tekstem sformułowanym przez jedną osobę. Mówiąc tak, nie twierdzący bynajmniej, że tekst spójny musi być sformułowany przez jednostkę w sensie biologicznym. Idzie o tę jedność, która wynika z możliwości utożsamiania wszystkich „ja” ram modalnych tekstu. (Termin „rama modalna” wprowadziła A. Wierzbicka, jest to równoważnik terminu B. Russela *propositional attitude*. Chodzi o to, że w każdym zdaniu jest implicite zawarty stosunek mówiącego do powiadomienia, który można by sformułować tak: „wiedz, że...”, „chcę, żebyś wiedział”, „czuję, że...”; por. o tym rozważania A. Wierzbickiej w tomie pt. *Dociekania semantyczne*).

Jest tekstem adresowanym do jednego odbiorcy. Ów jeden odbiorca nie musi być jedną osobą. Jest jednym odbiorcą w sensie pewnego obowiązującego typu wiedzy. Jak zobaczymy dalej – nie tylko wiedzy językowej. Może to być zatem bardzo wąska i wyspecjalizowana grupa ludzi, może być grupa tak szeroka, jak

całe pokolenie, mówiące określonym językiem, może być grupa tak szeroka, jak cały naród.

I wreszcie jedność najtrudniejsza do względnie jednoznacznego określenia: tekst musi być tak zbudowany, by treść wszystkich zdań, które się nań składają, dawała w rezultacie opis jednego przedmiotu, opowieść o jednym przedmiocie lub rozumowanie dowodzące jednej tezy. Przedmiot ten może być dowolnie skomplikowany. Tym przedmiotem może być wygląd jednego drobnoustroju lub stosunki w obrębie nowoczesnego społeczeństwa; mogą to być dzieje motyla lub dzieje globu ziemskiego; rozumowanie dowodzące, że suma kątów w trójkącie równa się 180° lub że nieprawdą jest, jakoby to, co dobre było jednocześnie piękne i vice versa (...).

Dwa pierwsze warunki spójności tekstu nie wymagają chyba komentarzy i w równej mierze odnoszą się także do tekstów jednozdaniowych. Dla tekstów wielozdaniowych szczególnie ważny jest warunek trzeci, często bowiem bywa tak, że w tekście brak jakichkolwiek formalnych wykładników spójności składających się nań zdań (i tylko wiedza odbiorcy o sytuacji komunikowanej przez ten tekst pozwala potraktować dany zbiór zdań jako tekst spójny – zwracał na to uwagę O.A. Wojtasiewicz).

O funkcji odbiorcy informacji w wykrywaniu mechanizmów spójności tekstu pisze również Stanisław Lem w *Filozofii przypadku*:

„Z raz powstała sytuacją czytelnik nie bardzo chce się rozstać, jeśli tekst w sposób właściwy normatywnym stereotypom nie przeprowadza jej w następną (czyli kiedy ów tekst nie podlega takiej transformacji, jakiej reguły czytelnik ma dobrze wdrożone).

I znów najlepiej ujrzyć to na przykładzie. Niechaj pierwszym słowem tekstu będzie „pas”. Olbrzymia ilość kontekstów, w jakiej słowo to może występować, redukuje się gwałtownie przy następnym słowie „cnoty”; powstało bowiem nakierowanie semantyczne, najzupełniej odmienne od dawanych alternatywami w rodzaju: „przepuklinowy” czy „betonu”. (...)

(...) Jeśli dalszy ciąg zdania brzmi: „ofiarowany hrabinie przez małżonka, był...” – nie wiemy wprawdzie, jaki był ów pas, lecz antycypacja zwięzi poważnie pole dalszych możliwości. Najpierw, wiemy już, że pewne słowa pojawić się po „był” nie mogą (np. „gdyby”, „jeżeli”), bo na to nie pozwala składnia. Dalej, choć składnia fleksyjna zezwala na to, nie oczekujemy słów takich jak: „uśmiechnięty”, „sześciokołowy”, ponieważ raczej zabrania tego składnia semantyczna. Jeśli po „był” odczytamy „zamykany na kłódkę”, sytuacja ujednoznaczni się i niejako zaskoczy w pewien stereotyp; natomiast jeśli owym słowem będzie „niewidzialny”, poweźmiemy decyzję odmienną i zaszeregujemy rzecz w kontekst – stereotypu bajek. Albowiem w im wyższym stopniu jest nieobecna sytuacja ujednoznaczniająca wypowiedź (kosytuacja), tym bardziej jest wypowiedź zdana na zbiór schematów sytuacyjnych, jakimi dysponujemy dzięki nie tylko językowemu doświadczeniu. W odbiorze dążymy do tego, aby tak artykulację włączyć, jak tego wymaga strategia żądająca ujednoznacznienia sytuacyjnego całości wypowiedzi.

Jeśli po zdaniu: „Pas cnoty, ofiarowany hrabinie przez małżonka, był niewygodny” – pada następnie: „Hrabia wyjechał do Ziemi Świętej” – sytuacja unieruchamia się w swojej wyrazistości. Gdy kolejne zdanie brzmi: „Za oknem baszty stała noc bezgwiezdna” – nie możemy tego zdania zintegrować z poprzednim, jeśli nie podejmiemy innego rodzaju, nowej antycypacji interpretującej: spodziewamy się, iż od tego opisu, który obrysowuje sytuację na jej dalszym obwodzie, lecz pewno tę samą, tekst wróci do hrabiny i do pasa. Utwierdza nas w tym zdanie: „Młody giermek piął się po dzikim winie, ściskając w zębach wytrych” – gdyż wytrych pasuje do zamka, którego domyśliliśmy się w pasie cnoty, a giermek – do hrabiny. Lecz zdanie: „Brodawki leczy się najlepiej rosołem z czarnej kury” – „Wazelina jest rodzajem węglowodoru” – „Giermek jest to rzeczownik rodzaju męskiego” – zbiją nas z pantafelku. W jakich realnych okolicznościach mogłoby zdanie to przecież nastąpić po wymienionych powyżej? Chyba tylko, jeśli byśmy się znajdowali np. w antykwariacie i towarzysz nasz, na chybił trafił otworzył kolejno dwie książki leżące na ladzie, na głos odczytał wyrwane ze środka tekstów zdania. Wcale byśmy się wówczas sąsiedztwu pasa cnoty, hrabiny i wazeliny nie dziwili. Jednakże takich założeń nie jesteśmy skłonni czynić pod nieobecność realnej sytuacji; albowiem wypadałoby nad ową sytuacją, jaką zrazu rozbudował tekst, wnieść następną, wyższego rzędu, która zawiera w sobie, na prawach elementów, dwie od siebie względnie niezależne (jedną z nich daje historyjka o hrabinie, drugą – wiadomość o składzie chemicznym wazeliny)”.

Jak widać, w rozważaniach na temat spójności tekstu wyróżnić można dwa stanowiska:

- O spójności danego tekstu decydują jego cechy immanentne – w tym ujęciu dany tekst może być tylko albo spójny, albo niespójny dla każdego odbiorcy,
- O spójności danego tekstu decydują nie tylko jego cechy immanentne, ale także wiedza odbiorcy o komunikowanej przez ten tekst rzeczywistości, w tym i o innych tekstach (relacje intertekstualne) – tekst może być dla jednego odbiorcy spójny, a dla innego niespójny. Spójność tekstu jest więc stopniowalna, a stopień spójności tego samego tekstu dla tego samego odbiorcy może być różny w różnym czasie, zależy bowiem od wiedzy tego odbiorcy i umiejętności uspoźniania tekstu, a więc włączania do struktury systemu wiedzy danego odbiorcy. Bliskie temu stanowisku jest rozróżnianie dwóch rodzajów spójności tekstu: **kohezji**, czyli **spójności strukturalnej i koherencji**, czyli **spójności semantycznej**.

Szczególną rolę w uspoźnianiu tekstu odgrywa wspomniana wyżej wiedza o innych tekstach, do których dany tekst odsyła lub które przywołuje w sposób jawny (poprzez cytaty, powołania, odsyłacze) lub niejawnie. Takie jawnie wyrażone **relacje intertekstualne** cechują pewne gatunki tekstów, na przykład teksty ustaw prawnych, częste są także w tekstach naukowych. Umiejętność rozpoznania niejawnych relacji intertekstualnych, mających jedynie charakter relacji asocjacyjnych, może być warunkiem uspoźnienia tekstu, odkodowania zawartych w nim informacji lub odebrania intencji jego nadawcy, a więc rozpoznania na przykład ironii, aluzji, parodii czy pastiszu. Relacje intertekstualne mogą być relacjami organizującymi tekst literacki, a najlepszym przykładem takiego tekstu jest nowa

powieść Umberta Eco, pt. *La misteriosa fiamma della Regina Loana* (*Tajemniczy płomień królowej Loany*, 2004), o której Andrzej Litwornia tak pisze (Litwornia A.: *Koniec papierowego świata*. Gazeta Wyborcza, 19 lipca 2004):

„Jest bowiem najnowsza powieść intertekstualnym monologiem w poszukiwaniu utraconej pamięci. Czterystopięćdziesięciostronicowa narracja utkana jest z brawurowo nakładających się cytatów, powiedzonek, złotych myśli, pierwszych słów arcydzieł włoskiej głównie literatury (choć bywa, że i francuskiej), ostatnich słów wielkich ludzi, przemówień i podręczników życia duchowego. Jednym słowem jest to intertekstualizm w stanie wrzenia i erupcji, ale też w pełni czytelny chyba tylko dla Włocha, któremu dzieciństwo i młodość przypadały ba lata 1930 – 50.”

l Tekst omawianej powieści jest także dobrym przykładem **tekstu mieszane- go**, a więc będącego realizacją więcej niż jednego systemu znakowego, jak pisze cytowany wyżej autor „Bogate ilustracje w powieści to pieczołowicie wybrane przez Eco precyzyjnie wprowadzone do tekstu reprodukcje włoskich komiksów, tekstów pieśni masowych i piosenek, okładek pism kobiecych, kolorowych kalendarzy, ilustracji i kart tytułowych literatury popularnej i młodzieżowej, plakatów faszystowskiej propagandy, pierwszych stron gazet, starych, egzotycznych znaczków pocztowych, blaszanych pudełek, barwnych nalepek reklamowych, opakowań płyt gramofonowych oraz starych zdjęć.”

Cecha spójności jest więc nie tylko cechą tekstów jednego języka (jednego systemu znakowego), ale charakteryzuje także teksty będące realizacją kilku systemów znakowych, z tym że najczęściej w takim tekście mieszanym jeden system pełni w tekście funkcję organizującą – zazwyczaj jest to język naturalny. Takim tekstem mieszanym może też być tekst jednozdaniowy, realizujący struktury syntaktyczne języka naturalnego, w których niektóre pozycje syntaktyczne wypełnione są przez wyrażenia innego języka naturalnego lub sztucznego. Tekstem mieszanym jest na przykład podręcznik gramatyki języka obcego (innego niż język naturalny, który pełni funkcję metatekstową), tekst języka naturalnego z cytatami w języku obcym, tekst podręcznika matematyki, logiki, chemii lub opisujący jakiś język informacyjno-wyszukiwawczy. Szczególnym rodzajem tekstu mieszane- go jest komiks, w którym język naturalny nie pełni funkcji organizującej tekst. Oprócz tekstów mieszanych możemy też mieć do czynienia z **kodami mieszanymi**, których system zakłada współwystępowanie wyrażen pochodzących z różnych kodów. Przykładem takiego kodu jest system znaków drogowych, włączający elementy języka naturalnego i notacji dziesiętnej zapisu wartości liczbowych.

Jedną z najważniejszych cech formalnych dłuższych tekstów spójnych – co do tego zgodni są wszyscy badacze zajmujący się tą problematyką – są **powtórzenia**, każde zdanie tekstu składa się bowiem z części powtarzającej informację znaną odbiorcy, często z poprzednich partii tekstu oraz z części zawierającej informację nową, komunikowaną w tym zdaniu. Funkcję wykładników informacji, która była już zawarta w poprzednich zdaniach, pełnią oprócz powtórzeń pozwalające ich unikać zaimki anaforyczne oraz inne wyrażenia mające w danym kontekście

taką samą denotację, a więc synonimy, deskrypcje określone oraz wyrażenia szersze zakresowo. Okazuje się jednak, że aby tekst można było uznać za spójny, powtórzenia te muszą być w nim odpowiednio rozłożone, nie wystarczy bowiem, jeśli dane jednostki leksykalne powtarzają się tylko w sąsiednich zdaniach. Nie uznamy więc za tekst spójny następującego ciągu zdań (przykład z pracy I.P. Sevbó: *Struktura sviaznogo teksta i avtomatizacija referirovanija*):

„Spojrzałem dziś w niebo. Ciele było w czarnych chmurach. Chmury powstają z pary. Gorącą parę dobrze jest wdychać przy katarze. Katar jest pewną oznaką grypy. Grypę wywołują wirusy. Wirusy bada się w laboratoriach”

Z tego samego powodu nie będzie spójny tekst (przykład Stanisława Lema z *Filozofii przypadku*):

„Lokomotywa stanęła przed kupą piasku, pustynnego w swej samotności, która jest przeznaczeniem ludzi niewidomych, stukających łaskami w chodnik, pełen skrzyń, z napisami reklamowymi, co daje dużo zysku fabrykantom włókna i armat, strzelających do Murzynów w wilgotnych i piwnicznych dżunglach, lochach arabskiego kłamstwa...”

Z powyższych przykładów widać, że relacja spójności między zdaniami nie jest relacją przechodnią i nie wystarczy do uznania za spójny tekstu, w którym semantyczne czy gramatyczne nawiązania występują tylko w obrębie sąsiadujących jednostek zdaniowych. Będzie natomiast spójnym tekst (Henryk Worcell: *Wpisani w Giewont*):

„– Nie rozumiem – mówił tonem wzburzonym – dlaczego niektórym ludziom tak bardzo zależy na degradacji człowieka! Na zrównaniu go ze zwierzętami, gorzej, bo z automatem. Gruczoły! Gruczoły nakazywały pójść świętym na mękę, heretykom na stos! Gruczoły zmusiły Poniatowskiego do skoczenia w Elstere, Koziatulskiego do zdobycia Samosierry, Witosa do bezkompromisowej polityki, gruczoły każą nam ryzykować życie, narażać się na więzienie, na wygnanie. Ona by pewnie działaniem gruczołów wytłumaczyła cudowne nawrócenie świętego Pawła, według niej gruczoły pobudzały rewolucjonistów do działania, Słowackiego do pisania „Króla Ducha”, zdobywców bieguną do męczarni w arktycznych lodach. Gruczoły, gruczoły, wydzieliny gruczołów, jakieś związki chemiczne, jakieś...”

gdzie powtórzenia są równomiernie rozłożone (pełnią one tu również funkcję ekspresywną).

Spójność tekstu osiąga się również przez umieszczenie w następnych zdaniach jednostek leksykalnych znajdujących się o stopień (lub o kilka stopni) wyżej w tej samej gałęzi drzewa hierarchicznego (nazywanie tego samego elementu rzeczywistości terminem coraz szerszym). Pozwala to, jak już mówiliśmy, na uniknięcie dosłownych powtórzeń; co byłoby niezręczne ze względów stylistycznych. Przykład ilustrujący tę prawidłowość podajemy z *Praktycznej stylistyki* A. i P. Wierzbickich:

„Na podwórzu zanosił się wprost ujadaniem mały kundelek. Patrzcie jaki śmieszny! – pisnęła Agatka. Pozwoliłem jej pogłaskać pieska. Zwierzak nastroszył się, szczechnął jeszcze głośniej, a potem ucichł nagle i rozwalil się brzuszkiem do góry. Spodobało mu się widać głaskanie Agatki”.

Jelena Paduczewa nazywa takie wypadki **zdominowaniem nazwy** (mówi o tym w pracy *O strukturze akapitu* zamieszczonej w tomie *O spójności tekstu*):

„Nazwa (z danego zdania) jest dominowana, jeśli spełnia jeden z dwóch warunków: 1) oznacza przedmiot wspomniany już w poprzednich zdaniach danego tekstu (dominacja bezpośrednia); 2) oznacza przedmiot, w całości jeszcze nie wymieniany, którego część stanowi przedmiot, wymieniony poprzednio (dominacja pośrednia). Nazwa, dzięki której dana nazwa staje się zdominowana, nazywa się dominującą dla danej nazwy (nazw dominujących może być kilka, jeśli przedmiot był już kilkakrotnie wymieniany).

Por. przykład dominacji bezpośredniej:

«Wania i Tania wyszli ze szkoły. Dzieci spieszyły się do kina. Chłopiec biegł na przedzie, dziewczynka ledwie nadążała za nim».

Dla nazwy «dzieci» nazwą dominującą są «Wania» i «Tania»; dla nazwy «chłopiec» – «dzieci» i «Wania»; dla nazwy «dziewczynka» – «dzieci» i «Tania».

W poprzednim przykładzie nazwą dominującą dla nazwy „piesek” jest nazwa „kundelek”, dla nazwy „zwierzak” – nazwy „kundelek” i „piesek”.

A oto podane przez Paduczewą przykłady dominacji pośredniej:

„Wania i Tania wyszli z domu bardzo wcześnie. Szkoła, w której dzieci uczyły się, była daleko od domu.”

Nazwa „szkoła” jest zdominowana pośrednio przez nazwę „dzieci”.

„Bawialnię radcy stanu Szamarykina otula miły półmrok. Duża brązowa lampa z zielonym abażurem barwi na zielono a la „noc ukraińska” ściany, meble, twarze... Od czasu do czasu na dogasającym kominku pryska tlejąca głownia”.

Nazwy „lampa” i „kominek” są tu dominowane pośrednio przez nazwę „bawialnia”, przy czym predykat wyrażający wzajemny stosunek tych nazw (przynależność) jest jedynie domyślny.

Jedną z cech tekstów spójnych jest więc występowanie w nich jednostek leksykalnych należących do tego samego pola semantycznego (powiązanych między sobą relacjami skojarzeniowymi). Te właśnie prawidłowości tekstu spójnego pozwalają mieć nadzieję na automatyzację procesu sporządzania informacyjno-wyszukiwawczych charakterystyk dokumentów.

Powtarzanie się wyrażen w tekście jest, jak widzieliśmy, koniecznym warunkiem spójności tekstu. Powtórzenia te gwarantują spójność tekstu od zdania do zdania, w każdym zdaniu musi być bowiem zawarta jakaś minimalna informacja, która albo była już podana przez zdania poprzednie, albo była z innych źródeł znana odbiorcy tekstu. Najczęściej bywa tak, że to, co w jednym zdaniu jest informacją nową, w następnych zdaniach występuje jako informacja już znana,

do której następnie dołącza się nowe informacje. Taka **informacyjna struktura tekstu** pozwala dokonać tzw. **aktualnego rozczłonkowania tekstu**, wydobyć w nim te wyrażenia, które są wykładnikami informacji znanej – tzw. **datum** i informacji nowej – tzw. **novum**. Analizę taką przeprowadzić można w obrębie każdego wypowiedzenia, wyróżniając w nim tzw. **temat**, a więc informację znaną już odbiorcy, nazywającą ten element rzeczywistości, którego dotyczy nowa informacja, oraz tzw. **remat**, część podającą nową informację, to, co nadawca komunikuje odbiorcy o temacie. Wykładnikami poszczególnych członów może być szyk wyrazów w zdaniu, np. w zdaniach:

*Kobieta wyszła z domu,
Janek jest chory.
Jutro wyjeżdżam na wczasy.*

tematami są odpowiednio: *kobieta, Janek, jutro*, zaś rematami: *wyszła z domu, jest chory, wyjeżdżam na wczasy*, natomiast w zdaniach:

*Z domu wyszła kobieta.
Chory jest Janek.
Na wczasy wyjeżdżam jutro.*

tematami są: *z domu, chory, na wczasy*, zaś rematami: *wyszła kobieta, Janek, jutro*.

Istnienie morfemowych wykładników relacji syntagmatycznych pozwala na stosunkowo dowolne operowanie szykiem wyrazów w zdaniu, który, zwolniony od funkcji syntaktycznych, może być, na przykład, wykorzystany jako wykładnik hierarchii ważności dla nadawcy i odbiorcy poszczególnych elementów informacji lub też stopnia ich nowości, jak to jest w języku polskim. Dla przykładu zacytujemy tu wraz z komentarzem omówione w *Praktycznej stylistyce* (A. i P. Wierzbiccy) różniące się jedynie szykiem zdania:

*„Jamnika postanowiłem podarować teściowej.
Teściowej postanowiłem podarować jamnika.
Postanowiłem podarować teściowej jamnika.
Postanowiłem podarować jamnika teściowej.*

Jest chyba rzeczą jasną, że informacje zawarte w tych czterech zdaniach są różne. Z pierwszego zdania dowiadujemy się, że autor dysponuje jamnikiem, że zastanawiał się, komu go podarować i że wybór padł na teściową. Drugie zdanie powiadamia, że autor obmyślał prezenty dla kilku osób, że jedną z tych osób była teściowa, i że dla niej autor wybrał jamnika. Z trzeciego dowiadujemy się, że autor zastanawiał się, jak postąpić w stosunku do swojej teściowej, i zdecydował: podaruję jej jamnika. Wreszcie czwarte zdanie informuje: autor zastanawiał się, co zrobić z jamnikiem i doszedł do wniosku: Podaruję go teściowej”.

Zdania te różnią się więc tzw. **strukturą tematyczno-rematyczną**, a wykładnikiem tych różnic jest właśnie różny szyk wyrazów w zdaniu.

Czasami wykładnikami struktury tematyczno-rematycznej jest przebieg intonacyjny wypowiedzenia, tzw. akcent zdaniowy, na przykład w wypowiedzeniu:

Janek jest chory.

wyraz *Janek* wyróżniony akcentem jest rematem – taka struktura akcentowa odpowiada zdaniu: *To Janek jest chory*. Wyróżnieniu struktury informacyjnej zdania służy również w języku polskim opozycja strony czynnej i strony biernej, na przykład w zdaniu:

Kolumb odkrył Amerykę.

tematem jest *Kolumb*, natomiast w zdaniu:

Ameryka została odkryta przez Kolumba.

tematem jest *Ameryka*. Wykładnikami tej tzw. **perspektywy funkcjonalnej zdania** mogą też być odpowiednie wyrażenia metainformacyjne, np.: *to właśnie*, które jest wykładnikiem rematu, np.:

*To właśnie Janek zachorował.
jeżeli chodzi o..., co się tyczy..., co do...,*

które są wykładnikami tematu, np.:

*Jeżeli chodzi o Janka, to jest chory.
Co się tyczy Janka, to jest chory.
Co do Janka, to jest chory.*

Zdarza się też, że temat w ogóle nie ma swojego wykładnika w wypowiedzeniu, są bowiem takie wypowiedzenia, które w całości są rematami, np:

*Pada.
Grzmi.
Świta,
Leje.
Deszcz pada.
Idę do sklepu, więc może coś ci kupić.
Jutro mnie nie będzie w pracy.*

Oczywiście w różnych językach naturalnych takie wykładniki struktury informacyjnej wypowiedzenia też mogą być różne, na przykład język japoński ma gramatyczny wykładnik tematu. Jest nim partykuła *-wa*, dodawana do tematyzowanego wyrażenia. Na przykład w zdaniu:

Watasi wa itumo wa hiru no syokuzi wa itizi mae ni wa sinain' desu ga, kyoo wa gogo iku tokoro ga arimasu kara sukosi hayaku sitain' desu.

Co znaczy: Jeżeli o mnie chodzi, jeżeli o to chodzi co zawsze, jeżeli chodzi o południowy posiłek, jeżeli chodzi o to wszystko, to przed pierwszą godziną

nie robię tego, lecz jeżeli chodzi o dzisiaj, to ponieważ mam wyjść popołudniu, zrobię to trochę wcześniej.

Na koniec naszych krótkich rozważań na temat struktury tematyczno-rematycznej zdania chcielibyśmy zwrócić uwagę na to, że nie zawsze wyrażenia temat, datum oraz remat, novum używane są synonimicznie, choć najczęściej oznaczają one zjawiska językowe bardzo do siebie podobne.

Choć operację analizy aktualnego rozczłonkowania wypowiedzenia przeprowadza się w obrębie poszczególnych zdań, to w istocie jest to operacja dokonywana nie na strukturze syntaktycznej zdania, lecz na **strukturze informacyjnej wypowiedzenia** (komunikatu), a do przeprowadzenia jej niezbędne są często informacje pochodzące z innych zdań tekstu, dlatego problem ten omawiamy w rozdziale dotyczącym tekstu.

W każdym tekście spójnym istnieją ponadto różnego rodzaju formalne wykładniki spójności poszczególnych jego zdań niosące informację o relacjach między tymi zdaniami. Do takich wykładników należą przede wszystkim spójniki międzyzdaniowe i różnego typu wyrażenia nawiązujące, takie jak np. *ale, taki, jednak, tak więc, a mianowicie, na przykład, tj., czyli, reasumując, itd., itp.* Stopień nasycenia tekstu tymi wyrażeniami spajającymi zależny jest w dużej mierze od stylu i rodzaju tekstu. Największa różnorodność tego typu wyrażen występuje w prozie naukowej, najprostsze (i często najbardziej prymitywne) cechują teksty potoczne. Podamy tu dwa przykłady zastosowań takich najprostszych „spajaczy” tekstu: opowiadanie Nel z *W pustyni i w puszczy* Sienkiewicza:

„I, tatusiu! I nas porwali i wieźli na wielbłądach – i Gebhr mnie uderzył – i Staś mnie bronił – i przyjechaliśmy do Chartumu – i tam ludzie marli z głodu – i Staś pracował, żeby dostać dla mnie daktyli – i byliśmy u Mahdiego – i Staś nie chciał zmienić religii – i Mahdi wysłał nas do Faszody – i potem Staś zabił lwa i wszystkich – i mieszkaliśmy w wielkim drzewie, które się nazywa Kraków – i King był z nami – i miałam febrę – i Staś mnie wyleczył – zabił wobo – i zwyciężył Samburów – i był zawsze dla mnie bardzo dobry, tatusiu!...”

i opowiadanie starej kobiety o filmie Antonioniego z opowiadania Tadeusza Różewicza pt. *Na placówce dyplomatycznej*:

„A w tem kinie, to panie jedna panna chodziła po pokoju, a on nie patrzył. Ona się patrzy, a on nie patrzy, a potem ona chodzi koło ściany, a on siedzi, to znów panie, on koło ni chodzi, a ona stoi, potem ona chodzi albo leży na łóżku, potem nic, ino trzyma telefon, ale nic nie mówi, ino śmieje się i rozkłada na tem łóżku, a potem na innym łóżku z jakoms innom paniom siedzi i się śmieje, pokładają się na tym łóżku i znów telefonują, a ta to taka była smutna, umalowała się na Murzynkę i w ręczniku chodziła po mieszkaniu – i, panie, dupskiem to tak kręciła, jakby jom co ugryzło, tom się zaśmiała, ale synowa sykla, tom się już nie śmiała, a wieczorem przy stole nic, ino o tym Antonim gadali, że taki, że owaki, cudowny, bo samom esencyje poka-

zał, a potem ta, co się w tym ręczniku wierciła, to zaczęła po ulicy chodzić i przy beczce z wodom przystawała, stała przy ty beczce i stała, a w ty beczce to pewno deszczówka była, to ona do ty deszczówki patyk rzuciła i se poszła, a potem znów przyszła i do ty beczki jakby gadała, a w ty beczce już były różne śmieci, potem jakiś inny kawaler do ni chodził, ale ona mu uciekła, to on za nią chodził i pod domem wystawał, a ona się znów śmiała i tak przez cały film latali, jedne w te, drugie w te, i tak się kochali, potem siedzieli na łóżku, potem na kanapie leżeli i się śmieli, a potem przestali się śmiać i ona gdzieś poleciała i sobie włosy odgarniała bez przerwy, bo nie była uczesana, tylko wiatr ją potarł, ale jak się skończyło, czy się pożenili, to nie wiem, bom się trochę zdrzemła, ale się obudziłam, jak mi synowa do ucha sykla”.

Jak widać z powyższego, autorzy posługują się terminem „spójność”, rozumiejąc go raczej intuicyjnie. Witold Marciszewski, korzystając ze zgromadzonych tu przykładów, wykazał, że wchodzi w grę trzy nierównoważne pojęcia spójności, które nazwał: **zwartością**, **uporządkowaniem**, **monotematycznością**. Na zakończenie naszych rozważań o spójności tekstu przytoczymy tutaj tę interpretację in extenso.

„Przez **zwartość** rozumiem gęstość powiązań między terminami tekstu, przy czym powiązania traktuje się jako relacje dwuargumentowe wyrażane przez odpowiednie predykaty. Jeśli liczbę terminów tekstu (wszystkich lub tylko tych, które się wyróżnia jako słowa kluczowe) oznaczymy przez n i założymy, że dla każdej pary terminów zachodzi najwyżej jedna relacja (z tego uproszczenia można w razie potrzeby zrezygnować), to najmniejsza zwartość wyrazi się liczbą $n - 1$, największa zaś liczba $(n^2 - n) / 2$. Uzasadnienie tych wzorów podaję w pracy *Organizacja semantyczna tekstu* („Studia Semiotyczne”, t.VI), tutaj zwrócę tylko uwagę, że w przypadku granicznym, gdy $n = 2$, powinno być tak, że maximum = minimum, co istotnie jest konsekwencją zaproponowanych wzorów.

Tekstem zwartym minimalnie, czyli pozbawionym zwartości, jest taki oto cytowany w skrypcie fragment:

„Spojrzałem dziś w niebo. Ciele było w czarnych chmurach. Chmury powstają z pary. Gorącą parę dobrze jest wdychać przy katarze. Katar jest pewną oznaką grypy. Grypę wywołują wirusy. Wirusy bada się w laboratoriach”.

Oto występujące w tej wypowiedzi terminy i relacje:

1. ja – niebo. Relacja: patrzeć na.
2. niebo – chmury. Relacja: być w.
3. chmury – para. Relacja: powstawać z.
4. para – katar. Relacja: być pomocnym na.
5. katar – grypa. Relacja: być oznaką.
6. grypa – wirusy. Relacja: być skutkiem.
7. wirusy – laboratorium. Relacja: być badany w.

Mamy tu 7 powiązań przy 8 terminach, a więc zgodnie z wzorem $n - 1$ zwartość jest minimalna.

Ciekawym przykładem wysokiej zwartości jest tekst następujący.

„Wania i Tania wyszli ze szkoły. Dzieci spieszyły się do kina. Chłopiec biegł na przdzie, dziewczynka ledwie nadązała za nim”.

Terminami tego tekstu są: 1. Wania, 2. Tania, 3. Szkoła, 4. Kino. Relacjami są: wyjść z (oznaczmy ją przez Z), iść do (ozn. Do), wyprzedzać (W), podążać za (\overline{W}) – konwers relacji wyprzedzania, iść razem (R). Przedstawmy te związki w tabeli (ten rodzaj tabeli, bardzo pomocny w badaniu zwartości, da się przyporządkować każdemu tekstowi).

	1	2	3	4
1. Wania		R,W	Z	Do
2. Tania	$\overline{R},\overline{W}$		Z	Do
3. Szkoła				
4. Kino				

Jeśli pominąć konwersy relacji (nie wnoszące nowych informacji), to okazuje się, że dla 4 terminów istnieje tu 6 powiązań, co jest liczbą wysoką zważywszy, że w sytuacji, gdy każdy termin odniesiony jest do jakiegoś innego po razie (a nie po dwa razy, jak tutaj „Wania” do „Tania”), to maksymalna liczba powiązań (po odjęciu trywialnych równości na przekątnej i konwersów rozłożonych symetrycznie po drugiej stronie przekątnej) wynosi: $(n^2 - n) / 2$; dla 4 terminów daje to liczbę 6. Ten maksymalny wskaźnik realizuje się tu mimo braku relacji między „szkoła” i „kino”, dzięki podwójnej relacji między „Wania” i „Tania”.

Tenże tekst pozbawiony jest jednak spójności, jeśli rozumieć przez spójność, zgodnie z sugestią niektórych autorów, uszeregowanie zdań tekstu wedle jakiejś **relacji porządkującej**. Wyraża się to w tym, że zdania można by poprzestawiać bez szkody dla sensu opowiadania, np. przestawić zdanie drugie z trzecim. Przykładem tekstu całkowicie uporządkowanego, a jednocześnie całkiem pozbawionego zwartości, jest:

„... zaskakujący strzał Brychczego – Kostka z trudem piąstkuje – piłka wychodzi na róg...”

Relacją porządkującą zbiór zdań tekstu jest tutaj relacja odwzorowująca kolejność zdarzeń w czasie i, ewentualnie, pewien rodzaj stosunku sprawstwa. W obrębie tej sekwencji nie wolno dokonać żadnego przestawienia, jeśli nie chcemy dopuścić się naruszenia sensu i zniekształcenia przez to obrazu opisywanej rzeczywistości. Uporządkowanie jest tu przeto maksymalne. Zwartość tekstu jest jednak minimalna, wyraża się bowiem wskaźnikiem 5 przy 6 terminach i przy relacjach: robić (R), przytrafiać się (P), następować po (N). Oto odpowiednie zestawienie.

1. Brychczy – strzał ... R
2. Kostka – piąstkowanie ... R
3. Piąstkowanie – strzał ... N
4. Piłka – wyjście na róg ... P
5. Piąstkowanie – wyjście na ... N

Jak widać z porównania powyższych tekstów, spójność w sensie zwartości jest czymś innym od spójności w sensie uporządkowania.

Są też teksty, które odbieramy jako spójne, choć mają one niskie wskaźniki zarówno uporządkowania, jak i zwartości. Oto kolejny przykład ze skrytu:

„– Nie rozumiem – mówi tonem wzburzonym – dlaczego niektórym ludziom tak bardzo zależy na degradacji człowieka! Na zrównaniu go ze zwierzętami, gorzej, bo z automatem. Gruczoły! Gruczoły nakazywały pójść świętym na męki, heretykom na stos! Gruczoły zmusiły Poniatowskiego do skoczenia w Elsterę, Koziętulskiego do zdobycia Samosierry, Witosza do bezkompromisowej polityki, gruczoły każą nam ryzykować życie, narażać się na więzienie, na wygnanie. Ona by pewnie działaniem gruczołów wytłumaczyła cudowne nawrócenie świętego Pawła, według niej gruczoły pobudzały rewolucjonistów do działania, Słowackiego do pisania Króla Ducha, zdobywców bieguny do męczarni w arktycznych lodach. Gruczoły, gruczoły, wydzieliny gruczołów, jakieś związki chemiczne, jakieś...”

Wskaźnik zwartości, jak można policzyć, znajduje się tu pośrednio między minimum i maximum; uporządkowanie nie istnieje, bo te liczne przykłady (rzekomego) wpływu gruczołów na postępowanie można by dowolnie przedstawiać. Ale istnieje jeden termin powiązany relacjami z każdym prawie spośród pozostałych terminów, mianowicie słowo „gruczoły”. W tabeli relacji wyrazi się to przejrzystą konfiguracją, polegającą na nieprzerwanej sekwencji predykatów relacyjnych przy terminie „gruczoły”, co będzie wskazywać, że termin ten wyraża główny, wręcz jedyny, temat tekstu. Tę cechę tekstu można by określić jako **monotematyczność**.

W organizacji tekstu spójnego szczególną rolę pełnią omówione już wcześniej, w rozdziale *Metainformacja, metatekst, hipertekst*, **wyrażenia metatekstowe** strukturalizujące tekst i wskazujące relacje między zawartymi w tekście informacjami, a więc wykładniki **relacji intratekstualnych**, zachodzących między fragmentami tekstu, takimi jak na przykład akapity. Stopień nasycenia tekstu wykładnikami relacji intratekstualnych zależy od długości tekstu, ale przede wszystkim od typu tekstu, a szczególnie duży jest w tekstach naukowych (liczne odsyłacze do innych fragmentów tekstu lub między tekstem głównym a towarzyszącymi mu przypisami).

Ważną rolę w tekście spójnym pełnią również wyrażenia o **funkcji fatycznej**, przy pomocy których nadawca tekstu pilnuje, żeby tekst w całości dotarł do adresata, sygnalizując początek tekstu, sprawdzając, czy w trakcie jego nadawania kontakt między nim a adresatem nie został zerwany i ewentualnie sygnalizując koniec tekstu. Użycie wyrażeń o funkcji fatycznej jest często skorelowane z poszczególnymi gatunkami tekstu, na zasadzie przyjętej konwencji. Najlepszym przykładem są tu zwroty zaczynające i kończące listy lub pisma urzędowe, teksty o funkcji oficjalnej, np. przemówienia, ale także takie teksty jak zwykła rozmowa, która wymaga nawiązania kontaktu przez nadawcę poprzez zwrócenia się do adresata komunikatu poprzez zwrot adresatywny, którego wybór uwarunkowany jest

wymaganym przez osobiste lub społeczne relacje między uczestnikami procesu komunikacji stopniem grzeczności (do problemu tego wrócimy w rozdz. XVIII).

Wyraźne ramy delimitujące i strukturalizujące tekst związane są z konwencjonalnymi **gatunkami tekstu**, takimi jak bajki (baśnie), które często zaczynają się słowami:

Dawno, dawno temu żył sobie...
Za siedmioma górami, za siedmioma lasami...
Daleko, daleko, za siódma górą i rzeką...
Był sobie raz...

W wypadku **tekstu pisanego** spójność i struktura tekstu może być wskazywana przez specjalne rozmieszczenie tekstu na nośniku, wskazanie początku przez specjalne wyróżniki graficzne, na przykład wielkość i krój liter, zastosowanie światła wyróżniającego całości tworzące spójny tekst – na przykład akapity, a często także wyraźne wskazanie, że tekst jeszcze się nie skończył, np., napisy *c.d.n.*, *ciąg dalszy s....*, lub że się już skończył, np. napis **KONIEC** lub użycie specjalnego znaku graficznego. Dobrym przykładem takich wskaźników spójności tekstu jest strona gazety.

Te wszystkie wyróżniki struktury tekstu są bardzo ważne nie tylko dla pomyślnego przebiegu procesu odbierania komunikatu – wszystkie mają adresatowi komunikatu ułatwić tę operację, ale również pomagają w dokonywaniu na tekście wszelkich operacji transformacji, takich jak streszczanie czy wybieranie informacji relewantnych z jakiegoś punktu widzenia.

Znajomość struktury tekstu oraz wykładników organizujących tę strukturę, szczególnie zaś jego strukturę informacyjną, jest niezbędna w procesie analizy treściowej tekstów dla potrzeb systemów informacyjnych.

Literatura:

- Bartmiński J., Boniecka B. (red.): *Tekst: problemy teoretyczne*. Lublin 1998.
de Beaugrande R-A., Dressler W.U.: *Wstęp do lingwistyki tekstu*. Warszawa 1990
Dobrzyńska T. (red.): *Teoria tekstu*. Wrocław 1986.
Dobrzyńska T., Janus E. (red.): *Tekst i zdanie*. Wrocław 1983.
Duszak A.: *Tekst, dyskurs, komunikacja międzykulturowa*. Warszawa 1998.
Marciszewski W.: *Metody analizy tekstu naukowego*. Warszawa 1977.
Mayenowa M.R. (red.): *O spójności tekstu*. Wrocław 1971.
Wilkoń A.: *Spójność i struktura tekstu*. Kraków 2002.
Wojtasiewicz O.A.: *Wybrane pojęcia teorii tekstu*. „Zagadnienia Informatyki Naukowej” 1978, nr 2.

XVII

Statystyczne prawidłowości w tekście

Aby w pełni opisać język, nie wystarczy opisać tylko jego słownik, gramatykę, reguły transformacji i reguły interpretacji, trzeba jeszcze podać częstość występowania w tekstach poszczególnych jego elementów. Częstość ta jest, jak sądzą niektórzy, np. Zipf, Guiraud, tak istotna, że w każdym elemencie języka są oni skłonni wyróżniać obok jego formy i treści także częstość jako trzecią niezbędną dla jego definicji cechę. Badanie częstości występowania elementów języka należy więc traktować na równi z analizą ich funkcji komunikatywnej, ekspresywnej i impresywnej, użyciem języka rządzą bowiem prawa wielkich liczb, i choć każdy mówiący może w pewnym sensie wybierać między poszczególnymi elementami języka, to jednak w ogólnej masie użycie przjawia się już zdeterminowanie, które sprawia, że częstości tych elementów są stałe.

Badanie nad **statystycznymi prawidłowościami w tekstach** rozpoczęto jeszcze w XIX wieku, a poprzedzało je sporządzanie tzw. konkordancji tekstów, czyli spisów, przeważnie ułożonych według alfabetu, wszystkich wyrazów, haseł występujących w tekście z podaniem ich lokalizacji i kontekstów, np. konkordancja Biblii i sanskryckich ksiąg Rigwedy. Dopiero później zaczęły powstawać tzw. słowniki frekwencyjne. Pierwszym był opracowany w 1861 roku przez Geisbla słownik podający częstość znaków chińskich (opracowany dla potrzeb stenografii). W później opracowanych słownikach podano już nie tylko częstość występowania wyrazów, ale również ich przynależność do określonego stylu. Słowniki obecnie opracowywane bardzo często podają także informacje dotyczące morfologii.

Wszystkie **słowniki frekwencyjne** możemy podzielić następująco (wg J. Sambor: *Słowa i liczby*):

1. Ze względu na rodzaj elementu językowego, którego częstość słowniki podają, na:

- **słowniki leksykalne** – podstawą jest w nich wyraz jako jednostka leksykalna (np. *dom*, *głowa*),

- **słowniki leksykalno-morfologiczne** – uwzględniające zarówno częstość jednostek leksykalnych, jak i form fleksyjnych (słowniki te związane są ściśle z typem syntaktycznym języka),

- **słowniki semantyczne** – jednostką jest znaczenie wyrazu (podają one częstości znaczeń dla poszczególnych wyrazów, np. *dom* jako budynek, *dom* jako mieszkanie; *głowa* jako część ciała, *głowa* jako ktoś stojący na czele, kierujący, np. *głowa domu*, *głowa państwa*),

- **słowniki frazeologiczne** – zawierają częstości związków frazeologicznych.

II. Ze względu na zasięg badanego słownictwa, na:

- **słowniki ogólne** – podające słownictwo języka ogólnego,

- **słowniki specjalne** – zawierające słownictwo o zasięgu ograniczonym (do tych ostatnich należą np. słowniki środowiskowe oraz słowniki poszczególnych autorów – w Polsce np. słowniki Mickiewicza i Paska). Dla języka polskiego ukazały się następujące słowniki frekwencyjne:

- słownik opracowany przez Walerego Pisarka podający frekwencję wyrazów w prasie (*Frekwencja wyrazów w prasie. Wiadomości – komentarze – reportaże*. Kraków 1972) sporządzony na podstawie wylosowanych z roczników „Trybuny Ludu”, „Życia Warszawy” i „Expressu Wieczornego” prób tekstu reprezentujących trzy różne gatunki stylu dziennikarskiego, a obejmujących 130 tys. wyrazów;

- słownik opracowany przez Andrzeja Lewickiego, Władysława Masłowskiego, Jadwigę Sambor i Jerzego Woronczaka pt.: *Słownictwo współczesnej publicystyki polskiej. Listy frekwencyjne*. Warszawa 1971. Zawiera on pełny materiał wyrazowy, jaki uzyskano z próby polskich tekstów publicystycznych z lat 1963-1967 o łącznej długości 100 tys. wyrazów;

- słownik opracowany przez Idę Kurcz, Andrzeja Lewickiego, Jadwigę Sambor i Jerzego Woronczaka pt.: *Słownictwo współczesnego języka polskiego. Listy frekwencyjne*, tom I – *Teksty popularno-naukowe*. Warszawa 1974, tom II – *Drobne wiadomości prasowe*. Warszawa 1974, A. Lewickiego, M. Masłowskiego, J. Sambor, J. Woronczaka tom III – *Publicystyka*. Warszawa 1975, I. Kurcz, A. Lewickiego, J. Sambor, J. Woronczaka tom IV – *Proza artystyczna*. Warszawa 1976, tom V – *Dramat artystyczny*. Warszawa 1977.

Słowniki opracowano na podstawie próby pół miliona wyrazów reprezentujących pięć różnych stylów współczesnego języka polskiego – dla każdego po 100 tys. wyrazów: popularnonaukowy, publicystyki, drobnych wiadomości prasowych, prozy artystycznej i dramatu artystycznego. Słowniki te podają również wiele danych o charakterze morfologicznym. Wbrew zwyczajowi stosowanemu w niektórych słownikach frekwencyjnych, autorzy nie uwzględnili tekstów języka potocznego – jako że nie należą one do języka pisanego – ani tekstów poetyckich, które reprezentują na ogół zupełnie odrębny typ stylu artystycznego i nie są charakterystyczne dla języka ogólnonarodowego.

- *Słownik frekwencyjny polszczyzny współczesnej*. 1990, oprac. Ida Kurcz, Andrzej M. Lewicki, Jadwiga Sambor, Krzysztof Szafran, Jerzy Woronczak, Kraków;

- Zgółkowa Halina, 1983, *Słownictwo współczesnej polszczyzny mówionej*. Poznań;

- Zgółkowa Halina, Bułczyńska Katarzyna, 1987, *Słownictwo dzieci w wieku przedszkolnym*, Poznań.

Słowniki frekwencyjne układa się według malejącej frekwencji, a więc wyraz najczęściej występujący w tekście (o największej frekwencji f) otrzymuje rangę (r) 1 i pierwsze miejsce w słowniku.

Analiza tak ułożonego słownika pozwoliła na wykrycie pewnych **prawidłowości statystycznych**. Okazało się, że w każdym języku wyrazy występujące najczęściej są wyrazami najkrótszymi i najprostszymi morfologicznie, są to więc wyrazy niederywowane, najczęściej jednomorfemowe. Zależność między częstością wyrazów a ich długością badał Zipf i wyraził ją funkcją:

$$k = c \cdot \log r$$

gdzie: $c = \text{constans}$, k – długość wyrazu w fonemach (lub w sylabach), r – ranga wyrazu. Ilustracją tej prawidłowości jest następująca, ułożona przez Zipfa, tabela (podaję za J. Sambor: *Słowa i liczby*):

Język angielski Dane wg list Eldridge'a			Język łaciński – Plautus		
częstość	liczba wyrazów	przeciętna liczba fonemów	częstość	Liczba wyrazów	przeciętna liczba sylab
f	w	k	f	w	k
1	2976	6,656	1	5429	3,23
2	1079	6,151	2	1198	2,92
3	516	6,015	3	492	2,77
4	294	6,081	4	299	2,05
5	212	5,589	5	161	2,60
10	45	5,377	10	43	2,32
20	16	4,687	20	15	2,08
21 – 30	–	3,455	21 – 30	–	2,00
31 – 50	–	3,903	31 – 50	–	2,05
51 – 60	–	3,333	51 – 60	–	1,70
61 – 4290	–	2,666	62 – 514	–	1,40

Jak widać z porównania przytoczonych w tabeli liczb, długość wyrazów maleje w miarę, jak ich częstość rośnie, i to zarówno długość mierzona w fonemach, jak i w sylabach. Na zależność tę wskazują również dane z opartego na próbie 11 milionów wyrazów słownika frekwencyjnego języka niemieckiego, opracowanego jeszcze w 1868 r. przez F.W. Kaedinga (podaję za P. Guiraud: *Zagadnienia i metody statystyki językoznawczej*). I tak 11 milionów wyrazów tego słownika zawiera:

- 5 426 326 wyrazów jednosylabowych
- 3 156 448 wyrazów dwusylabowych
- 1 410 484 wyrazów trzysylabowych
- 643 971 wyrazów czterosylabowych.

W częstoci występowania poszczególnych elementów języka przejawia się **ekonomia systemu językowego**, obowiązujące również w języku **prawo najmniejszego wysiłku**, zgodnie z którym elementy łatwiejsze, mniej skomplikowane, występują częściej niż elementy trudniejsze, bardziej skomplikowane. Najlepszą ilustracją tego jest częstość występowania w języku polskim poszczególnych fonemów.

W poniższych tabelach zamiast pisowni fonetycznej fonemy oznaczono literami:

Wyniki badań Marii Steffen

samogłoski		spółgłoski (przykładowo)	
e	10,2%	p	3,1%
a	9,4%	t	4,4%
o	8,6%	s	3,0%
i/y	8,0%	r	3,6%
u	3,4%	m	3,5%
ę	3,4%	ł	2,2%
ą	0,6%	b	1,5%
		d	2,2%
		z	1,7%

Listy rangowe fonemów najczęściej występujących w badanych tekstach (podaję za J.Sambor: *Język polski w świetle statystyki*)

Ranga	Jassem i Łobacz (1971) język mówiony (bez pauzy)	Rocławski (1981) język mówiony	Steffen (1957) język pisany
1	e 111,3	e 107,8	e 102,0
2	a 101,1	a 98,0	a 94,0
3	o 83,7	o 90,4	o 86,0
4	t 50,7	t 51,4	j 45,0
5	j 45,8	j 45,7	t 44,0
6	n 42,2	y 40,5	y 41,0
7	y 40,3	n 38,9	n 40,0
8	m 36,7	m 36,0	r 36,0
9	i 35,6	i 35,8	m 35,0
10	w 33,1	w 32,7	w 35,0
11	p 31,3	p 31,7	i 34,0
12	r 30,8	r 31,3	u 34,0
13	u 29,7	u 31,0	p 31,0
14	dz 29,2	k 28,9	s 30,0
15	k 26,5	s 28,0	k 27,0

Widzimy więc, że spółgłoski bezdźwięczne mają frekwencję około dwa razy większą niż odpowiadające im spółgłoski dźwięczne, bardziej skomplikowane i trudniejsze w artykulacji, bo zawierające jedną więcej cechę dystynktywną.

We wszystkich tekstach wśród najczęstszych piętnastu znalazły się te same fonemy, co świadczy o tym, że częstość występowania fonemów w języku polskim jest cechą systemową niezależną od subkodu, czyli pisanej lub ustnej odmiany języka. Dane te wskazują na jeszcze jedną prawidłowość: struktura statystyczna inwentarza fonemów jest w języku polskim stała i nie zależy od subkodu, w jakim język jest używany.

Prawo najmniejszego wysiłku potwierdza również struktura najczęściej występujących wyrazów. Wiemy już, że najczęściej występującą grupę morfemów tworzą te, które są wykładnikami kategorialnymi, a wśród nich najczęstsze są wykładniki relacji syntagmatycznych. W języku polskim wykładnikami takich relacji są bądź końcówki fleksyjne, bądź przyimki – jedne i drugie krótkie, zazwyczaj jedno- lub dwufonemowe. Potwierdza to lista 15 najczęściej występujących we współczesnej publicystyce polskiej wyrazów:

- | | | | | |
|----------------|-----------------------|---------------|---------------|------------------|
| 1) <i>w/we</i> | 4) <i>na</i> | 7) <i>się</i> | 10) <i>on</i> | 13) <i>o</i> |
| 2) <i>i</i> | 5) <i>z/ze</i> | 8) <i>nie</i> | 11) <i>to</i> | 14) <i>który</i> |
| 3) <i>być</i> | 6) <i>ten, ta, to</i> | 9) <i>do</i> | 12) <i>że</i> | 15) <i>a</i> |

Z badań frekwencyjnych nad częstością występowania fonemów wynika jeszcze jeden wniosek: w kodowaniu informacji w języku polskim o wiele większą rolę odgrywają spółgłoski niż samogłoski, gdyż samogłoski są w języku polskim bardzo łatwo przewidywalne na podstawie kontekstu, mają więc bardzo dużą redundancję, a małą entropię (wartość informacyjną), natomiast spółgłoski przeciwnie – mają małą redundancję a dużą entropię. Przekonać się o tym łatwo – jeśli spróbujemy odczytać tekst pisany pozbawiony liter oznaczających samogłoski (przy dłuższym tekście można je z łatwością wstawić) i tekst pisany pozbawiony liter oznaczających spółgłoski (nie da się tego zrobić), np.

.Ę.Y.O..A....O ..A ..U.E..Ó. I..O..A..I .AU.O.E.
 J.Z.K.ZN.WSTW. DL. ST.D.NT.W .NF.RM.CJ. N..K.W.J
 JĘZYKOZNAWSTWO DLA STUDENTÓW INFORMACJI NAUKOWEJ

Przejawem wspomnianego wyżej prawa jest również fakt, że wyrazy potoczne, które jako neologizmy są wyrazami wielomorfemowymi, po zadomowieniu się w języku, jeżeli są często używane, ulegają skróceniu lub zostają zastąpione krótszymi synonimami, jednym z praw rządzących zmianami językowymi jest bowiem wynikające z powyższego prawo **tendencji do skrótu**. I tak na przykład wyrazy *kinematograf* został skrócony do postaci *kino*, *automobil* do postaci *auto*, a *radioodbiornik* i *radiofonia* do *radio*. Zamiast wyrazu *motocykl* często używa się wyrazu *motor*, a zamiast *samochód* mówimy *wóz*.

Bywa i tak, że zgodnie z tym prawem, w języku potocznym jeden wyraz zastępuje nazwy wielowyrazowe. W takim procesie uniwerbizacji specjalizują się nawet niektóre formanty, na przykład *-ówka*, *-ak*, *-ec*, *-anka*, np.:

szkoła podstawowa – podstawówka
liceum ogólnokształcące – ogólniak
szkoła zawodowa – zawodówka
samolot odrzutowy – odrzutowiec
statek do przewozu drobnicy – drobnicowiec
fabryka odzieży lub technikum odzieżowe – odzieżówka
technikum budowlane – budowlanka
technikum samochodowe – samochodówka
motor na ropę – ropniak
wychowawczyni w przedszkolu – przedszkolanka
buty z gumy – gumiaki
pracownik zajmujący się zaopatrzeniem – zaopatrzeniowiec.

Tu również należą gwarowe:

most Poniatowskiego – Poniatoszczak
kotlet schabowy – schaboszczak.

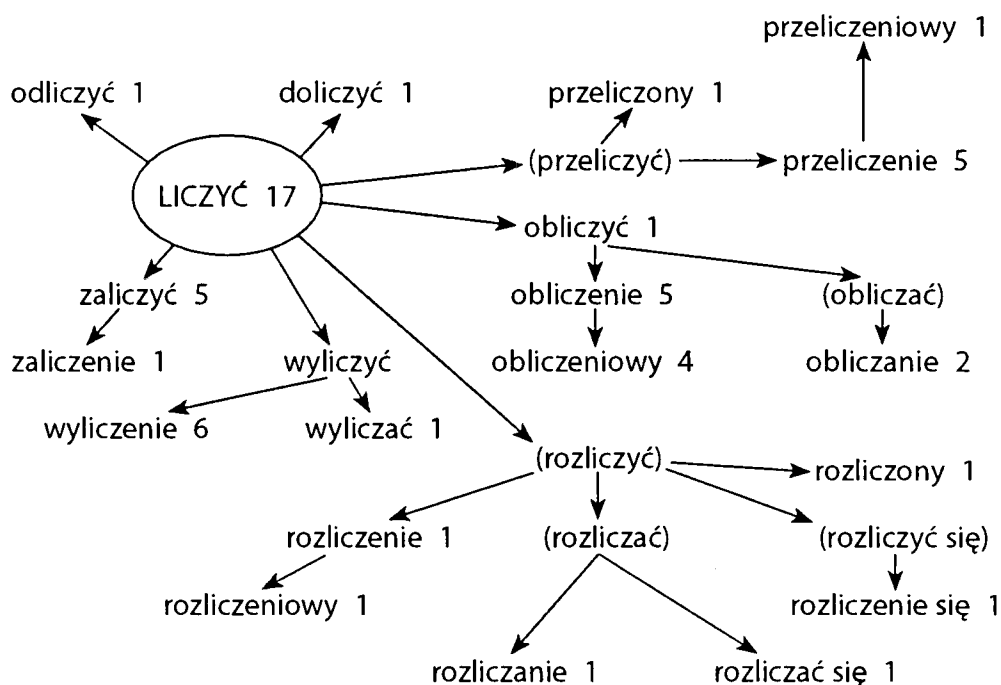
Przejawem tej samej tendencji są także liczne skróty i skrótowce, np.:

Państwowa Komunikacja Samochodowa – PKS – pekaes,
Centrala Przemysłu Ludowego i Artystycznego – CPLiA – cepelia.

Elementy językowe o wysokiej frekwencji działają ekspansywnie i często wypierają synonimiczne elementy o frekwencji niskiej. Na przykład, w potocznym języku polskim formant *-ak*, wypiera formant *-ę* tworzący nazwy zwierząt niedorosłych, co pozwala na przeniesienie tych wyrazów z trudnego i obejmującego niewiele rzeczowników wzorca deklinacyjnego do wzorca bardziej popularnego – *cielak, cielaka, kurczak, kurczaka* zamiast *cielę, cielęcina, kurczę, kurczęcia*. Na tej samej zasadzie również forma *warszawiak* zastępuje formę *warszawianin*, która daje dłuższe formy przypadków zależnych (*warszawiaka* ale *warszawianina*).

Przejawem tej prawidłowości jest również rozkład częstości wyrazów pochodnych w porównaniu z częstością wyrazów o prostej budowie słowotwórczej, a więc krótszych. Dla ilustracji przytaczamy na następnej stronie podany przez J. Sambor przykład bardzo rozbudowanego gniazda słowotwórczego, gdzie widać wyraźnie, że liczne wyrazy pochodne mają o wiele mniejsze częstości niż czasownik podstawowy, a równocześnie charakteryzują się o wiele bardziej złożoną budową, są więc dłuższe. Na wykresie tym przy wszystkich wyrazach podane są częstości, strzałki wskazują kierunki relacji „podstawa słowotwórcza → derywat”, hasła podane w nawiasach w badanym materiale nie wystąpiły.

O prawidłowości tej należy pamiętać przy tworzeniu języków sztucznych, jeśli mają one być językami ekonomicznymi, tzn. jeśli teksty ich mają być stosunkowo krótkie. Spośród języków informacyjno-wyszukiwawczych warunek ten spełnia jedynie Klasyfikacja Spawalnictwa opracowana przez Zygmunta Dobrowolskiego, w której zastosował system notacyjny tzw. znaków skróconych, nadając formy znaków najkrótszych jednostkom leksykalnym zajmującym pozycje klas końcowych.



Pewna zależność występuje również między częstością wyrazów a liczbą ich znaczeń. Okazuje się bowiem, że wyrazy najczęstsze są zarazem wyrazami o największej liczbie znaczeń, natomiast słownictwo rzadkie jest na ogół jednoznaczne. Słowników frekwencyjnych podających liczbę znaczeń poszczególnych wyrazów jest niewiele, badania takie sprawiają bowiem dużo trudności ze względu na brak ścisłych kryteriów wyodrębniania różnych znaczeń jednego wyrazu. Problemem tym zajmowała się Jadwiga Sambor, analizując wieloznaczność derywatów (wyrazów pochodnych), między innymi z przyrostkiem *-acja*, na materiale współczesnej publicystyki polskiej. Okazało się, że wśród najczęstszych wyrazów tego typu było wiele wyrazów wieloznacznych, np. *organizacja*, *reprezentacja* (oznaczające czynność i grupę ludzi), *administracja* (oznaczający czynność, urząd, budynek, grupę ludzi), *informacja* (oznaczający czynność, treść komunikatu), *kontrakcja* (oznaczający czynność, obiekt tej czynności), *ilustracja* (oznaczający czynność oraz pewien obiekt), natomiast wyrazy rzadko występujące o takiej samej strukturze, np. *degradacja*, *kontemplacja*, *kapitulacja*, *sublimacja*, *rezygnacja*, *wibracja*, *separacja*, *eksploatacja* są wyrazami jednoznacznymi.

Statystyka leksykalna wykryła jeszcze jedną zależność – zależność między zakresem i treścią wyrazów a ich częstością, okazało się bowiem, że wyrazy o szerokim zakresie a ubogiej treści występują w tekstach często, natomiast wyrazy o dużej treści (a więc małym zakresie) występują w tekstach rzadko. Wniosek stąd, że słownictwo bogate w informację jest słownictwem rzadkim. Ta prawidłowość może być wykorzystana przy budowie języków informacyjno-wyszukiwawczych oraz instrukcjach sporządzania charakterystyk dokumentów. Ilustracją tej prawidłowości niech będą następujące dane dotyczące częstości wystąpień (podane przez J. Sambor) pochodzące ze słownika frekwencyjnego tekstów popularnonaukowych:

zwierzę 66 – koń 15 – ogier 1,
urządzenie 82 – maszyna 57 – turbina 13,
roślina 34 – drzewo 18,
metal 35 – żelazo 23 – Fe_{59} 1,
choroba 64 – nowotwór 9 – rak 4.

Oczywiście tak dzieje się tylko w tekstach niespecjalistycznych. W konkretnych tekstach specjalistycznych, omawiających wąskie zagadnienia, wyrazy o małym zakresie i dużej treści występują częściej niż wynikałoby to z prawidłowości ogólnojęzykowych i one właśnie są w tych tekstach słowami kluczowymi.

Wyrazy najczęściej występujące stanowią również najstarszą warstwę słownictwa, co łatwo zobaczyć przeglądając listę frekwencyjną języka polskiego. Wśród słownictwa najczęstszego spotyka się najwięcej tzw. wyrazów gramatycznych (przyimki, spójniki, zaimki), które znamy przecież jeszcze z lektury tekstów staropolskich (tego typu słownictwo prawie się w języku nie zmienia), tu również znajduje się najwięcej wyrazów prawie tak samo brzmiących w różnych językach słowiańskich, a więc pochodzących jeszcze z języka prasłowiańskiego (por. załączone listy frekwencyjne dla języka polskiego i rosyjskiego).

Układając listy frekwencyjne dla różnego rodzaju tekstów zauważono powtarzającą się w nich prawidłowość, okazuje się bowiem, że w każdym tekście występuje bardzo mała grupa wyrazów o bardzo dużej frekwencji i bardzo duża grupa wyrazów o frekwencji małej. Ilustracją jest tu następujące zestawienie, wykonane dla próby tekstów naukowo-technicznych dotyczących zagadnień obserwacji sztucznych satelitów Ziemi, obejmujących 10018 wyrazów (charakterystyki ilościowej tego typu tekstów dotyczyła moja praca magisterska) – zob. tabela: Procentowy wykaz pokrycia słownika przez wyrazy o danej frekwencji f.)

Jak widać, wyrazy o częstości ponad 10 stanowią zaledwie 0,44% całego słownika, wyrazy o częstości ponad 50 tylko 0,93% całego słownika, natomiast wyrazy o częstości 1 stanowią aż 50,47% słownika, natomiast wyrazy o częstością ≤ 3 aż 73,39% całego słownika danej dziedziny wiedzy.

W zależności od częstości występowania poszczególnych wyrazów całą leksykę podzielić można na pewne strefy. Zazwyczaj wyróżnia się w dużych słownikach cztery takie strefy. Okazuje się, że w strefie I, w grupie 50-100 wyrazów najczęstszych, znajdują się wyrazy o funkcji gramatycznej: rodzajniki, spójniki, przyimki, zaimki osobowe i wskazujące, czasowniki posiłkowe, w strefie II, grupie wyrazów o rangach 101-1000, a więc wśród 1000 wyrazów najczęstszych znajduje się słownictwo podstawowe, wspólne wszystkim stylom języka, o znaczeniach bardzo ogólnych, wyrażających podstawowe zjawiska, czynności, stany, cechy (np. *chodzić, mówić, dobry, człowiek* itp.), w strefie III znajduje się słownictwo specjalne, odznaczające się dużą częstością tylko w jednym stylu, w strefie IV słownictwa rzadkiego znajduje się najwięcej wyrazów (strefa ta stanowi olbrzymią większość każdego słownika tekstu) – tu znalazły się, na przykład, takie wyrazy, jak: *angażowanie się, apartament, bankructwo, batalistyczny, biadać, biedzić się*. Ilustracją I strefy leksyki mogą być listy zawierające 10 najczęstszych wyrazów w ogólnych słownikach języka polskiego, angielskiego i rosyjskiego:

Dziesięć najczęstszych wyrazów w języku polskim, angielskim i rosyjskim

Ranga	Język polski (h - 500 000)		Język angielski (h - 5 000 000)		Język rosyjski (h - 1 000 000)	
1	w	16 316	and	106 064	v	42 854
2	i	12 386	of	100 551	i	36 266
3	się	10 046	a/an	89 572	nie	19 226
4	być	9 620	to	67 263	na	17 262
5	na	8 598	that	53 263	ja	13 839
6	nie	8 341	have	48 432	być	13 307
7	z	8 310	in	43 337	что	13 185
8	on	6 363	it	35 512	он	13 143
9	do	5 857	he	31 765	s	12 975
10	ten	5 743	has	27 789	a	10 719

Procentowy wykaz pokrycia słownika przez wyrazy o danej frekwencji f

Ranga r	Frekwencja f	% ogólnej ilości wyrazów różnych wyrazów o frekwencji	
		f	≥f
1	> 300	0,04	0,04
	200 - 300	0,18	0,22
	150 - 199	0,09	0,31
	100 - 149	0,13	0,44
10	80 - 99	0,13	0,57
	60 - 79	0,18	0,57
	50 - 59	0,18	0,93
	40 - 49	0,39	1,32
	30 - 39	0,30	1,62
	25 - 29	0,26	1,88
	20 - 24	0,64	2,52
	15 - 19	1,36	3,88
100	12 - 14	1,73	5,61
	1 - 11	1,43	7,04
200	10	0,82	7,86
	9	1,25	9,11
	8	1,81	10,92
	7	1,55	12,47
	6	2,11	14,58
	5	3,61	18,19
500	4	4,56	22,75
	3	6,57	29,32
1 000	2	16,35	45,67
	1	50,47	96,14
nazwy własne		3,86	100,00

Zajmując się badaniami statystycznymi języka, Zipf odkrył prawidłowość, że: iloczyn rangi wyrazu i jego frekwencji jest dla danego tekstu wielkością stałą, co zapisujemy:

$$\text{ranga} \cdot \text{częstość} = \text{constans}$$

Zależność tę ilustrują następujące przykłady:

- z wykazu Hemnona (podaję za P. Guiraudem):

Ranga r	Frekwencja f	Iloczyn r · f
50	698	34 900
100	314	31 400
200	158	31 600
500	64	32 000
1 000	31	31 000

- z *Ulissesa* Jamesa Joyce'a (dane Zipfa, podaję za A. Wierzbicką, *O języku dla wszystkich*):

Ranga r	Frekwencja f	Iloczyn r · f
10	2 653	26 530
20	1 311	26 220
30	926	27 780
40	717	26 680
50	556	27 800
100	265	26 500
200	133	26 600
300	84	25 200
400	62	24 800
500	50	25 000
1 000	26	26 000

- z tekstów dotyczących obserwacji sztucznych satelitów Ziemi:

Ranga r	Frekwencja f	Iloczyn f · r
10	103	1 030
12	91	1 092
20	58	1 160
50	22	1 100
60	20	1 200
70	19	1 330
80	17	1 560
90	15	1 350

Wzór podany przez Zipfa był później wielokrotnie poprawiany, okazało się bowiem, że jest on poprawny tylko dla dość ograniczonej liczby wyrazów, nie spełnia go bowiem ani słownictwo najczęstsze, ani słownictwo najrzadsze. Model Zipfa, choć obejmuje tylko część słownika tekstu, to jednak odzwierciedla statystyczną strukturę słownictwa ogromnej części każdego tekstu, gdyż wyrazy najczęstsze pokrywają największą jego część. I tak obliczono, że:

- 100 najczęstszych wyrazów pokrywa 60% tekstu,
- 1 000 najczęstszych wyrazów pokrywa 8% tekstu,
- 4 000 najczęstszych wyrazów pokrywa 97,5% tekstu

(oczywiście chodzi tu o frekwencję przewidzianą dla danego języka, nie zaś o frekwencję wyrazów w konkretnym tekście, który może być znacznie krótszy).

Dla lepszej ilustracji powyższej prawidłowości podajemy w oddzielnej tabeli procentowy wykaz pokrycia tekstu przez wyrazy o frekwencji f występujące w tekstach dotyczących obserwacji sztucznych satelitów Ziemi:

	Frekwencja	Procentowy wskaźnik pokrycia tekstu przez wyrazy o frekwencji	
		f	$\geq f$
1	> 300	3,46	3,46
	200 - 300	8,86	12,32
	150 - 199	3,35	15,67
10	100 - 149	3,60	19,27
	80 - 99	2,62	21,89
	60 - 79	2,59	24,48
	50 - 59	2,31	26,79
	40 - 49	4,11	30,90
	30 - 39	2,44	33,34
	25 - 29	1,65	34,99
	20 - 24	3,38	38,37
	15 - 19	5,59	43,96
100	12 - 14	5,61	49,57
	11	3,77	53,34
	10	1,90	55,24
200	9	2,60	57,84
	8	3,35	61,19
	7	2,51	63,70
	6	2,93	66,63
	5	4,18	70,81
500	4	4,22	75,03
	3	4,57	79,60
1 000	2	7,52	87,12
	1	11,66	98,78
nazwy własne		1,22	100,00

Dane tego typu mają szczególne znaczenie przy nauczaniu języków obcych, gdzie słowniki frekwencyjne są nieodzowne. Niestety, wyrazy najczęstsze niosą najmniej informacji (co jest zgodne ze sformułowanym przez Shannona wzorem na ilość informacji, uwzględniającym prawdopodobieństwo zdarzenia), im więc wyraz rzadszy, tym większa jest jego pojemność informacyjna. Jak obliczył w przybliżeniu Guiraud:

- 100 najczęstszych wyrazów zawiera 30% informacji,
- 1 000 najczystszych wyrazów zawiera 50% informacji,
- 4 000 najczęstszych wyrazów zawiera 70% informacji.

Z tego wynika, że znajomość 1 000 najczęstszych wyrazów danego języka, jakkolwiek nieodzowna dla zrozumienia jakiegokolwiek tekstu, nie wystarcza do tego, aby odebrać pełną zawartą w tym tekście informację, do zrozumienia każdego tekstu potrzebne jest bowiem opanowanie słownictwa podstawowego (najczęstszego) oraz słownictwa specjalnego z zakresu przedmiotu tekstu. Szacuje się, że słownik czynny osoby ze średnim wykształceniem obejmuje około 15-20 tysięcy wyrazów. Dla ilustracji zagadnienia podajemy dane obrazujące przynależność wyrazów do grup wyróżnionego słownictwa: potocznego, ogólnonaukowo-technicznego i specjalistycznego zawartego w tekstach dotyczących zagadnień obserwacji sztucznych satelitów Ziemi:

Rodzaj słownictwa	Liczba wyrazów różnych	% pokrycia słownika	Frekwencja w tekstach	% pokrycia tekstu
Potoczne	1 918	82,82	8 515	85,00
Ogólnonaukowo-techniczne	242	10,45	861	8,59
Specjalistyczne	66	2,85	520	5,19
Nazwy własne	90	3,88	122	1,22
Razem	2 316	100,00	10 018	100,00

Dla lepszej ilustracji wszystkich powyższych wywodów podajemy również listy 100 najczęściej występujących wyrazów podanych w słowniku opracowanym przez Walerego Pisarka, w słownikach opracowanych przez Jadwigę Sambor i innych, w ogólnym *Słowniku frekwencyjnym polszczyzny współczesnej* (oprac. Kurcz I., Lewicki A.M., Sambor J., Szafran K., Woronczak J. Kraków 1990), w słowniku sporządzonym przez autorkę dla tekstów naukowo-technicznych dotyczących zagadnień obserwacji sztucznych satelitów Ziemi (opartym na próbie obejmującej 10 018 wyrazów), oraz w trzech słownikach frekwencyjnych języka rosyjskiego:

- L.N. Zasoriny, obejmującym 120 474 słów (literatura piękna, dramaturgia, teksty popularnonaukowe, gazety i czasopisma),
- E.M. Stepanowej, obejmującym 400 000 słów (matematyka, fizyka, chemia, biologia, medycyna, geologia z geografą),
- E.A. Sztejnfeldta, słowniku frekwencyjnym współczesnego języka rosyjskiego.

Walery Pisarek *Frekwencja wyrazów w prasie*

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	w	4807	51	to	186
2	i	3144	52	czy	182
3	z	2141	53	człowiek	178
4	na	2132	54	jaki	175
5	ten	2055	55	problem	175
6	być	1568	56	więc	174
7	do	1492	57	dalszy	161
8	nie	1254	58	między	161
9	on	1172	59	czas	159
10	że	1010	60	odbywać się	154
11	który	939	61	siebie	154
12	o	796	62	nawet	147
13	a	654	63	ostatni	145
14	mieć	553	64	każdy	136
15	po	448	65	sytuacja	136
16	przez	418	66	cel	134
17	jak	409	67	obecnie	134
18	od	392	68	polityka	133
19	dla	388	69	też	133
20	za	378	70	my	132
21	ale	351	71	strona	132
22	swój	344	72	dobry	131
23	już	340	73	można	129
24	sprawa	326	74	aby	128
25	inny	307	75	droga	128
26	bardzo	302	76	około	128
27	kraj	302	77	tu	124
28	moc	301	78	miejsce	123
29	tylko	300	79	świat	123
30	co	282	80	państwo	122
31	nowy	266	81	właśnie	121
32	również	262	82	okres	117
33	praca	252	83	kilka	116
34	tak	244	84	produkcja	116
35	przed	243	85	przedstawiciel	114
36	przy	239	86	bez	113
37	jeszcze	238	87	dokonywać	113
38	jednak	235	88	prowadzić	113
39	oraz	235	89	musieć	111
40	nasz	233	90	należać	111
41	cały	226	91	trzeba	111
42	wiele	216	92	podkreślać	110
43	duży	214	93	polityczny	110
44	bieżący	208	94	stwierdzać	110
45	wielki	201	95	dzień	109
46	między	200	96	nad	109
47	taki	199	97	zadanie	106
48	także	199	98	znajdować się	106
49	wszystek	192	99	chcieć	105
50	mówić	186	100	życie	104

Andrzej Lewicki, Władysław Maślowski, Jadwiga Sambor, Jerzy Woronczak
Słownictwo współczesnej publicystyki polskiej.

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	w/we	3 819	51	jeszcze	163
2	i	3 061	52	plan	163
3	być	1 865	53	przykład	162
4	na	1 681	54	również	160
5	z/ze	1 527	55	polski	159
6	ten, ta, to	1 457	56	rozwój	158
7	się	1 287	57	dwa	157
8	nie	1 225	58	wysoki	155
9	do	1 088	59	problem	153
10	on	976	60	procent	153
11	to (zaimek rzecz.)	880	61	pierwszy	151
12	że	849	62	należać	148
13	o	724	63	warunek	148
14	który	722	64	cały	147
15	a	507	65	przedsiębiorstwo	145
16	praca	426	66	jaki	141
17	mieć	393	67	zostać	140
18	tysiąc	386	68	my	138
19	jak	343	69	powinien	134
20	dla	339	70	zadanie	134
21	sprawa	335	71	jako	131
22	przez	333	72	państwo	130
23	ale	330	73	szkoła	130
24	nasz	329	74	kultura	130
25	od/ode	310	75	można	129
26	tylko	295	76	musieć	129
27	swój	277	77	wiele	129
28	móc	269	78	to (spójnik)	127
29	bardzo	254	79	gospodarka	126
30	czy	249	80	partia	124
31	dziewięćset	247	81	przemysł	124
32	taki	240	82	siebie	124
33	inny	238	83	zakład	123
34	za	226	84	ludzie	122
35	wszystek	224	85	komisja	121
36	co	219	86	organizacja	120
37	tak	217	87	sam	120
38	lato	213	88	mówić	118
39	jeden (przymiot.)	206	89	drugi	117
40	kraj	205	90	nawet	117
41	nowy	195	91	miasto	116
42	po	195	92	także	114
43	produkcja	192	93	aby	113
44	już	191	94	naród	113
45	rada	188	95	też	113
46	przy	187	96	trzeba	113
47	jednak	173	97	tu	113
48	społeczny	172	98	gospodarczy	112
49	narodowy	165	99	sytuacja	112
50	Związek (im. własne)	165	100	związek	112

Ida Kurcz, Andrzej Lewicki, Jadwiga Sambor, Jerzy Woronczak
Słownictwo współczesnego języka polskiego. Teksty popularnonaukowe

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	w/we	3570	51	praca	161
2	i	2473	52	sam	159
3	być	2049	53	czy	157
4	na	1784	54	pod	157
5	z/ze	1744	55	duży	152
6	się	1510	56	jako	151
7	on	1423	57	okres	146
8	do	1245	58	też	146
9	ten	1196	59	większy	145
10	nie	895	60	lato	143
11	który	738	61	sposób	141
12	to	655	62	cały	140
13	że	654	63	siebie	136
14	a	555	64	liczba	135
15	o	533	65	różny	135
16	od	477	66	nowy	133
17	przez	424	67	człowiek	132
18	jak	375	68	należać	131
19	móc	364	69	punkt	128
20	mieć	341	70	zwany	125
21	tysiąc	314	71	nawet	123
22	jeden	304	72	wielki	123
23	dla	293	73	gdy	122
24	przy	286	74	woda	122
25	bardzo	285	75	jaki	121
26	po	282	76	pewny	120
27	taki	274	77	często	118
28	tak	273	78	trzy	118
29	lub	255	79	dziecko	116
30	tylko	254	80	dziewięćset	116
31	inny	236	81	życie	116
32	pierwszy	225	82	badanie	114
33	już	220	83	część	112
34	jednak	209	84	wiele	111
35	rok	209	85	zostać	111
36	swój	209	86	aby	110
37	co	207	87	przed	109
38	przykład	206	88	kraj	107
39	dwa	199	89	ziemia	107
40	więc	196	90	warunek	104
41	czas	191	91	u	103
42	ale	187	92	nasz	102
43	również	184	93	tu	102
44	można	183	94	między	99
45	drugi	175	95	układ	98
46	oraz	172	96	także	95
47	za	170	97	niż	94
48	wszystek	168	98	wiele	94
49	każdy	165	99	silnik	93
50	jeszcze	161	100	stopień	93

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	w Prp	16316	51	czas S	848
2	i Cnj	12385	52	chcieć V	839
3	być V	9621	53	praca S	821
4	się Prn	9302	54	zostać V	796
5	na Prp	8600	55	powiedzieć V	785
6	nie Prt	8341	56	to Cnj	784
7	z Prp	8310	57	tu Adv	779
8	on Prn	6650	58	oraz Cnj	757
9	do Prp	5854	59	lata S	755
10	ten Adj	5743	60	musieć V	747
11	to Prn	5113	61	cały Adj	745
12	że Cnj	4314	62	bo Cnj	736
13	a Cnj	3226	63	dzień S	718
14	o Prp	3072	64	mój Adj	706
15	ja Prn	3028	65	więc Cnj	692
16	który Adj	2997	66	nowy Adj	684
17	mieć V	2591	67	przed Prp	682
18	jak Cnj	2251	68	drugi Adj	673
19	co Prn	2224	69	jednak Prt	668
20	ale Cnj	1963	70	pod Prp	668
21	tak Prt	1903	71	pani S	656
22	pan S	1819	72	nawet Prt	637
23	od Prp	1780	73	jaki Adj	625
24	po Prp	1728	74	też Prt	613
25	móc V	1574	75	można V	606
26	przez Prp	1464	76	no Prt	600
27	taki Adj	1441	77	nic Prn	597
28	dla Prp	1361	78	ludzie S	585
29	tylko Prt	1340	79	każdy Adj	581
30	za Prp	1336	80	dziewięćset Num	572
31	ty Prn	1292	81	również Prt	572
32	tysiąc S	1207	82	trzy Num	568
33	swój Adj	1188	83	kraj S	561
34	jeden Adj	1152	84	żeby Cnj	536
35	bardzo Adv	1139	85	wszystko Prn	525
36	siebie Prn	1139	86	miejsce S	514
37	czy Prt	1119	87	tam Adv	503
38	rok S	1025	88	człowiek S	500
39	jeszcze Prt	1014	89	jakiś Adj	500
40	przy Prp	995	90	nad Prp	498
41	wiedzieć V	977	91	widzieć V	488
42	dwa Num	936	92	coś Prn	487
43	pierwszy Adj	936	93	teraz Adv	482
44	inny Adj	897	94	przecież Prt	481
45	mówić V	897	95	kiedy Adv	471
46	nasz Adj	890	96	może Prt	469
47	sprawa S	883	97	wielki Adj	465
48	sam Adj	882	98	życie S	462
49	my Prn	875	99	bez Prp	459
50	wszystek Adj	870	100	polski Adj	458

Bożenna Bojar *Słownictwo tekstów dotyczących obserwacji sztucznych satelitów Ziemi*

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	w/we	347	51	fotograficzny	22
2	być	261	52	pewien	22
3	i	218	53	po	22
4	na	206	54	współrzędna	22
5	z/ze	202	55	gdzie	21
6	satelita	170	56	tak	21
7	ten, ta, to	165	57	wzgląd	21
8	się	131	58	zmiana	21
9	do	126	59	co	20
10	ziemia	103	60	metoda	20
11	on, ona, ono	92	61	pole	20
12	obserwacja	91	62	położenie	20
13	od	79	63	obserwacyjny	20
14	orbita	68	64	układ	20
15	dla	65	65	wysokość	20
16	nie	64	66	ziemski	20
17	przy	62	67	inny	19
18	sztuczny	60	68	około	19
19	o	60	69	płaszczyzna	19
20	przez	57	70	podstawa	19
21	oraz	54	71	różny	19
22	stacja	49	72	elektron	18
23	a	48	73	pomoc	18
24	móc	48	74	radiowy	18
25	że	48	75	sygnał	18
26	który	46	76	wpływ	18
27	wartość	46	77	za	18
28	SSZ*	43	78	dwa	17
29	jak	42	79	jeden	17
30	czas	41	80	nachylenie	17
31	można	40	81	system	17
32	częstotliwość	37	82	elipsoida	16
33	odległość	37	83	kosmiczny	16
34	fala	34	84	nad	16
35	duży	32	85	obliczenie	16
36	punkt	32	86	obserwator	16
37	wzór	32	87	rok	16
38	okres	30	88	warunek	16
39	powierzchnia	29	89	wyznaczenie	16
40	dany	27	90	astronomiczny	15
41	kierunek	27	91	bardzo	15
42	badanie	26	92	cel	15
43	ruch	26	93	dokładność	15
44	element	25	94	geograficzny	15
45	lub	25	95	jonosfera	15
46	mieć	24	96	kąt	15
47	zakres	24	97	łączność	15
48	moment	23	98	pierwszy	15
49	praca	23	99	pomiar	15
50	atmosfera	22	100	przypadek	15

* SSZ – sztuczny satelita Ziemi – liczony jako jeden wyraz

Walery Pisarek

Listy 100 najczęstszych wyrazów w wiadomościach komentarzach i reportażach

(Listy zawierają łącznie 184 różne wyrazy. Z tej liczby 39 wyrazów występuje na wszystkich trzech listach, 42 wyrazy występują tylko na liście wiadomości, 27 – tylko na liście komentarzy, 38 – tylko na liście reportaży, 15 – tylko na listach wiadomości i komentarzy, 4 – tylko na listach wiadomości i reportaży oraz 19 – tylko na listach komentarzy i reportaży. Wyrazy, które znajdują się na wszystkich trzech listach, zostały podkreślone linią ciągłą; wyrazy, które występują tylko na jednej liście, są podkreślone linią przerywaną.)

Ranga r	WIADOMOŚCI	KOMENTARZE	REPORTAŻE
1	<u>w</u>	<u>w</u>	<u>w</u>
2	<u>i</u>	<u>i</u>	<u>i</u>
3	<u>z</u>	<u>ten, ta, to</u>	<u>ten, ta, to</u>
4	<u>na</u>	<u>na</u>	<u>z</u>
5	<u>do</u>	<u>z</u>	<u>na</u>
6	<u>być</u>	<u>do</u>	<u>być</u>
7	<u>który</u>	<u>być</u>	<u>on, ona, ono</u>
8	<u>nie</u>	<u>on, ona, ono</u>	<u>nie</u>
9	<u>ten, ta, to</u>	<u>że</u>	<u>do</u>
10	<u>o</u>	<u>nie</u>	<u>który</u>
11	<u>że</u>	<u>który</u>	<u>że</u>
12	<u>bieżący</u>	<u>o</u>	<u>o</u>
13	<u>również</u>	<u>a</u>	<u>a</u>
14	<u>sprawa</u>	<u>mieć</u>	<u>mieć</u>
15	<u>przez</u>	<u>przez</u>	<u>ale</u>
16	<u>a</u>	<u>swój</u>	<u>po</u>
17	<u>duży</u>	<u>jak</u>	<u>już</u>
18	<u>oraz</u>	<u>od</u>	<u>od</u>
19	<u>odbywać się</u>	<u>dla</u>	<u>za</u>
20	<u>po</u>	<u>kraj</u>	<u>wszystek</u>
21	<u>mieć</u>	<u>po</u>	<u>tylko</u>
22	<u>przed</u>	<u>sprawa</u>	<u>swój</u>
23	<u>jak</u>	<u>za</u>	<u>dla</u>
24	<u>dla</u>	<u>inny</u>	<u>co</u>
25	<u>praca</u>	<u>móc</u>	<u>tak</u>
26	<u>przy</u>	<u>to</u>	<u>jak</u>
27	<u>także</u>	<u>ale</u>	<u>człowiek</u>
28	<u>za</u>	<u>co</u>	<u>móc</u>
29	<u>problem</u>	<u>jednak</u>	<u>inny</u>
30	<u>dalszy</u>	<u>nasz</u>	<u>jeszcze</u>
31	<u>podkreślać</u>	<u>tylko</u>	<u>więc</u>
32	<u>bardzo</u>	<u>tak</u>	<u>bardzo</u>
33	<u>między</u>	<u>wszystek</u>	<u>przez</u>
34	<u>kraj</u>	<u>nowy</u>	<u>praca</u>
35	<u>prowadzić</u>	<u>polityka</u>	<u>ja</u>
36	<u>obecnie</u>	<u>już</u>	<u>taki</u>
37	<u>nowy</u>	<u>wiele</u>	<u>my</u>
38	<u>przewidywać</u>	<u>jaki</u>	<u>tu</u>
39	<u>dokonywać</u>	<u>polityczny</u>	<u>nowy</u>
40	<u>około</u>	<u>również</u>	<u>cały</u>
41	<u>jednak</u>	<u>oraz</u>	<u>każdy</u>
42	<u>przedstawiciel</u>	<u>cały</u>	<u>nasz</u>
43	<u>stwierdzać</u>	<u>wielki</u>	<u>mówić</u>
44	<u>od</u>	<u>między</u>	<u>czy</u>
45	<u>cel</u>	<u>jeszcze</u>	<u>siebie</u>

Ranga r	WIADOMOŚCI	KOMENTARZE	REPORTAŻE
46	oświadczać	taki	nawet
47	podejmować	sytuacja	jakiś
48	inny	wojna	przy
49	działalność	przy	można
50	przybywać	przed	duży
51	ubiegły	czy	przed
52	wiele	czas	tam
53	zadanie	świat	kraj
54	brać	dzień	wielki
55	mówić	naród	sprawa
56	ostatni	konferencja	trzeba
57	rozmowa	problem	dużo
58	szereg	ostatni	życie
59	pod	by	stary
60	podawać	także	chcieć
61	znajdować się	nawet	czas
62	produkcja	pod	jaki
63	przyjmować	siebie	dobry
64	tylko	praca	jednak
65	dotykać	aby	wiele
66	móc	strona	musieć
67	przyszły	stawać się	właśnie
68	otrzymywać	bez	mały
69	droga	granica	własny
70	już	właśnie	bo
71	okres	człowiek	też
72	wynik	iż	dobrze
73	kilka	walka	miejsce
74	realizacja	cel	pracować
75	temat	stanowisko	dom
76	teren	dalszy	nad
77	w ciągu	nad	pod
78	wszystek	należęć	wiedzieć
79	związek	prawo	gdy
80	w czasie	mówić	zakład
81	wielki	też	młody
82	on, ona, ono	wojskowy	także
83	poważny	obecnie	droga
84	strona	kilka	teraz
85	jeszcze	miejsce	również
86	aby	opinia	zaczynać
87	posiedzenie	sposób	żeby
88	rozwój	udział	kiedy
89	trwać	broń	przemysł
90	dziedzina	cena	może
91	udział	lecz	bez
92	rozpoczynać się	okres	dziś
93	dobry	pokojowy	pewien
94	liczba	produkcja	przecież
95	spotkanie	władza	przykład
96	sytuacja	fakt	wtedy
97	ważny	prasa	budowa
98	wzrost	bieżący	coś
99	zamierzać	dziś	mój
100	dotychczas	obaj	społeczny

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	и	5 261	51	если	243
2	в	3 759	52	или	239
3	он	2 340	53	говорить	236
4	на	2 111	54	до	236
5	не	1 949	55	страна	229
6	я	1 841	56	когда	228
7	что	1 558	57	кто	224
8	же	1 511	58	очень	220
9	это	1 497	59	первый	218
10	с	1 369	60	его	217
11	быть	1 342	61	советский	217
12	а	1 156	62	другой	214
13	тут	1 095	63	человек	213
14	как	857	64	идти	207
15	ты	782	65	себя	206
16	мы	696	66	уже	204
17	вы	683	67	большой	202
18	она	683	68	время	201
19	и	682	69	свет	200
20	весь	586	70	сам	194
21	по	565	71	год	191
22	они	554	72	где	186
23	из	543	73	стать	185
24	у	539	74	жизнь	179
25	о, об	516	75	новый	179
26	все	512	76	люди	176
27	то	504	77	глаз	175
28	но	503	78	день	170
29	за	502	79	дело	166
30	свой	491	80	ну	165
31	который	450	81	хотеть	163
32	так	434	82	надо	162
33	от	400	83	рука	162
34	для	396	84	под	158
35	такой	355	85	мир	152
36	наш	349	86	самый	151
37	только	333	87	видеть	149
38	один	326	88	народ	144
39	бы	303	89	ни	144
40	вот	300	90	просто	142
41	мочь	295	91	должен	141
42	еще	290	92	сила	136
43	знать	284	93	каждый	133
44	какой	284	94	наука	133
45	тот	282	95	теперь	133
46	да	272	96	земля	133
47	сказать	272	97	луч	132
48	чтобы	265	98	сейчас	132
49	мой	247	99	без	129
50	нет	244	100	стекло	129

Ranga r		Frekwencja f	Ranga r		Frekwencja f
1	в	14015	51	получить	744
2	и	11971	52	различный	728
3	этот	4468	53	определить	726
4	на	4179	54	некоторый	723
5	с	3876	55	о	723
6	при	3196	56	два	716
7	тот	3014	57	между	708
8	что	2974	58	вода	695
9	для	2581	59	равный	684
10	быть	2511	60	данный	680
11	который	2434	61	до	679
12	из	2270	62	например	678
13	не	2245	63	тело	677
14	к	2236	64	также	672
15	или	1968	65	движение	670
16	по	1946	66	но	665
17	число	1799	67	прямой	665
18	весь	1798	68	она	661
19	от	1788	69	через	657
20	как	1728	70	изменение	655
21	так	1609	71	более	635
22	иметь	1593	72	называться	631
23	мочь	1565	73	раствор	617
24	система	1507	74	общий	614
25	функция	1489	75	организм	605
26	а	1476	76	свойство	596
27	точка	1476	77	где	591
28	такой	1474	78	каждый	588
29	если	1408	79	первый	584
30	являться	1353	80	образом	577
31	они	1234	81	температура	574
32	мы	1233	82	большой	566
33	уравнение	1205	83	ряд	544
34	случай	1183	84	у	540
35	один	1147	85	область	553
36	значение	1122	86	его	549
37	другой	1116	87	поверхность	542
38	элемент	1093	88	скорость	536
39	можно	1051	89	вид	532
40	процесс	974	90	результат	529
41	множество	964	91	частица	526
42	часть	949	92	из	523
43	же	913	93	количество	515
44	он	908	94	определение	507
45	время	863	95	формула	500
46	вещество	860	96	ток	496
47	величина	844	97	должен	486
48	условие	836	98	основной	483
49	только	781	99	за	482
50	плоскость	746	100	развитие	482

Ranga r		Ranga r	
1	и	51	день
2	он (она, оно, они)	52	год
3	в (во)	53	кто
4	на	54	себя
5	не	55	дело
6	я	56	нет
7	что	57	рука
8	с (со)	58	очень
9	этот (это)	59	большой
10	быть (есть)	60	ну
11	а	61	новый
12	весь (все)	62	стать
13	тот (то)	63	школа
14	как	64	работа
15	мы	65	сейчас
16	к (ко)	66	время
17	у	67	идти (шел)
18	ты	68	если
19	за	69	два
20	но	70	мой
21	вы	71	жизнь
22	по	72	до
23	из (изо)	73	где
24	о (об, обо)	74	каждый
25	свой	75	хотеть
26	же (ж)	76	здесь
27	сказать	77	надо
28	так	78	теперь
29	один	79	дом
30	вот	80	пойти
31	сам (самый)	81	раз
32	который	82	товарищ
33	человек (люди)	83	ни
34	наш	84	или
35	только	85	ведь
36	еще	86	советский
37	от	87	работать
38	такой	88	город
39	мочь	89	там
40	говорить	90	слово
41	бы (б)	91	глаз
42	для	92	потом
43	уже	93	видеть
44	знать	94	под
45	да	95	даже
46	какой	96	думать
47	когда	97	хорошо
48	другой	98	можно
49	первый	99	тут
50	ребята	100	тысяча

Z porównania tych słowników widać, że grupa słownictwa najczęstszego jest we wszystkich tych słownikach wspólna, należą bowiem do niej przede wszystkim wyrazy o funkcji gramatycznej. Następną pod względem częstości występowania w tekstach grupę wyrazów stanowią słowa potoczne, natomiast wyrazy specjalistyczne mają stosunkowo małą frekwencję. I tak wśród słownictwa tekstów dotyczących zagadnień obserwacji sztucznych satelitów Ziemi tylko 10 wyrazów zaliczonych do grupy słownictwa specjalistycznego uzyskało częstość ≥ 10 (przy częstości przeciętnej 4,3), są to: *satelita* 170, *orbita* 68, *SSZ* (skrót *sztuczny satelita Ziemi* – traktowany jako jeden wyraz) 43, *astronomiczny* 15, *jonosfera* 15, *geocentryczny* 14, *półoś* 14, *perigeum* 19, *sputnik* 10.

Badania nad strukturą częstości słownictwa w tekstach przeprowadza się w celu:

- racjonalizacji procesu nauczania języków,
- optymalizacji systemów kodowania – elementy najczęstsze winny mieć jak najmniej skomplikowaną strukturę,
- wykorzystania danych liczbowych do badań językoznawczych i literaturoznawczych – na przykład rozstrzygnięcie o autorstwie danego dzieła literackiego, o chronologii powstania dzieł tego samego lub różnych autorów,
- optymalizacji procesów deszyfracji tekstów.

Znajomość prawidłowości statystycznych w tekstach może też być wykorzystana do **automatycznego sporządzania charakterystyk dokumentów** poprzez wybranie spośród wyrazów najczęstszych w danym dokumencie tych, których frekwencja jest relatywnie większa (oddzielną sprawą jest określenie, o ile ma być większa) od średniej występowania w tekstach należących do danej dziedziny wiedzy. Aby było to możliwe, konieczne jest opracowanie dla każdej dziedziny wiedzy i techniki oddzielnych słowników frekwencyjnych. Tego typu automatyzacja przy sporządzaniu charakterystyk dokumentów (wybór słów kluczowych) jest możliwa, wymaga jednak opracowania algorytmu umożliwiającego sprowadzenie do wspólnej formy hasłowej poszczególnych słowoform, w jakich wyrazy te reprezentowane są w tekście, co, jak wiemy, w wypadku języków fleksyjnych z fleksją wewnętrzną nie jest sprawą prostą. Znacznie trudniejsze, choć zapewne bardziej dokładne, byłoby obliczanie frekwencji dopiero po zastąpieniu w tekście wszelkiego rodzaju zaimków i uzupełnieniu elips pełnymi formami, jest to już jednak problem struktury tekstów spójnych.

Przykładem metody wykorzystania prawidłowości statystycznych w tekstach językowych przy automatyzacji procesów charakteryzowania dokumentów może być metoda zastosowana przez Pierre'a Guirauda w badaniach stylistycznych. Pokazał on, w jaki sposób przez porównywanie list frekwencyjnych słownictwa ustalić można zbiory tzw. **wyrazów tematycznych i wyrazów-kluczy**. Wyrazy tematyczne to słownictwo niegramatyczne mające najwyższe częstości w badanej klasie tekstów – dla potrzeb informacji naukowej byłyby to zbiory dokumentów obszaru tematycznego danego systemu informacyjnego. O tym, że słownictwo to zależy od rodzaju tekstów, najlepiej świadczy porównanie takich wyrazów tematycznych przytoczonych przez J. Sambor:

- dla tekstów współczesnej publicystyki polskiej są to: *kraj, produkcja, narodowy, polski, rozwój, przedsiębiorstwo, kultura, państwo, gospodarka, przemysł, partia, zakład, komisja,, organizacja, naród,*
- dla tekstów drobnych wiadomości prasowych są to:
 - wskaźniki czasu, a więc: rok, miesiąc, dzień, godzina,
 - nazwy wszystkich dni tygodnia,
 - nazwy uczestników wydarzeń, w tym nazwy osobowe: *minister, przewodniczący, prezydent, sekretarz, general, przedstawiciel, mieszkaniec, członek, osoba,*
 - nazwy zbiorowe: *urząd, partia, organizacja, rada, prezydium, komitet, komisja, zespół, grupa,*
 - nazwy wydarzeń: *spotkanie, konferencja, posiedzenie, obrady, wizyta,*
 - czasowniki: *zostać, przybyć, odbyć się, oświadczyć, otrzymać, zajmować się,*
 - przymiotniki: *bieżący i ubiegły,*
 - ogólne słownictwo abstrakcyjne: *czas, miejsce, problem, związek, część, stan, cel, wynik, warunek,*
- dla tekstów popularno-naukowych:
 - słownictwo abstrakcyjne, np.: *układ, stopień, przypadek, proces, zjawisko, zasada, metoda, zmiana, stan,*
 - nazwy obiektów badań: *życie, ziemia, powietrze, woda, ciało, urządzenie, material, powierzchnia, dziecko, zwierzę, silnik, samolot,*
 - terminy matematyczne: *liczba, punkt, zbiór,*
 - nazwy cech fizycznych, np.: *prędkość, wysokość, temperatura,*
 - nazwy zjawisk fizycznych: *prąd, fala, promieniowanie, choroba.*

Z przykładów tych widać, że badania statystyczne mogą też być pomocne przy tworzeniu języków informacyjno-wyszukiwawczych, szczególnie tych, których jednostki leksykalne są równokształtne z wyrazami języka naturalnego.

Dysponując takimi charakterystykami poszczególnych wydzielonych grup tekstów można się pokusić o wybór metodą statystyczną tzw. **słów-kluczy** (termin ten nie jest dokładnym odpowiednikiem terminu słowo kluczowe, ale, jak zobaczymy, jest mu bardzo bliski). Wyrazy-klucze wybiera się przez porównanie częstości słownictwa danego tekstu z częstością słownictwa przeciętnego, stanowiącego dla danej grupy tekstów normę. Dla wygody robi się to po prostu przez porównywanie rang wyrazów – wyrazy-klucze danego tekstu mają w nim rangi o wiele niższe (występują częściej) niż ich rangi w słowniku kontrolnym. Niestety dysponujemy tylko przykładami tak wybranych słów-kluczy dla utworów poetyckich, ale myślę, że stanowią one dobrą ilustrację metody możliwej do zastosowania przy automatyzacji procesu wyboru słów kluczowych w systemach dokumentacyjnych (podaję za J. Sambor).

Wyrazy-klucze *Kwiatów zła* Baudelaire'a

Wyrazy	Rangi w tekście	Rangi kontrolne	Wyrazy	Rangi w tekście	Rangi kontrolne
ange (anioł)	1	2 420	plein (pełny)	9	140
coeur (serce)	2	345	noir (czarny)	10	400
comme (jak)	3	11	soleil (słońce)	11	440
beaute (piękność)	4	935	ame (dusza)	12	330
oeil (oko)	5	120	amour (miłość)	13	331
ciel (niebo)	6	425	doux (cichy)	14	350
sein (piersź)	7	2 200	douleur (cierpienie)	15	860
parfum (zapach)	8	1 810	nuit (noc)	16	137

Wyrazy-klucze poezji Valery'ego

Wyrazy	Rangi w tekście	Rangi kontrolne	Wyrazy	Rangi w tekście	Rangi kontrolne
azur (błękit)	1	5 435	ombre (cień)	5	695
songe (sen)	2	3 550	oncle (fala)	6	-
pur (czysty)	3	705	or (złoto)	7	625
chair (ciało)	4	2 320	don (dar)	8	2 000

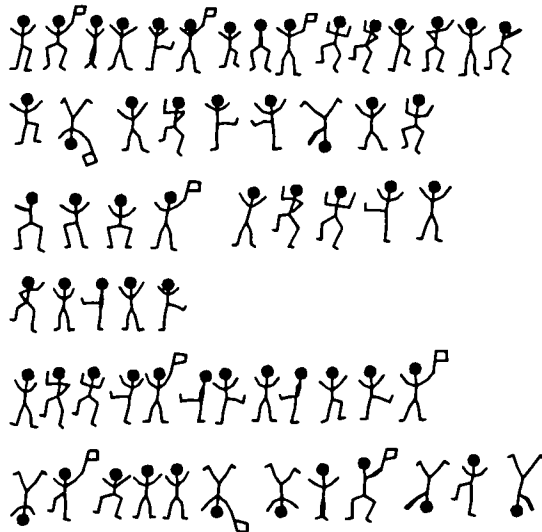
I jeszcze lista wyrazów-kluczy w *Weselu* Wyspiańskiego – pierwsza liczba oznacza rangę na liście zbiorowej *Wesela*, druga na liście kontrolnej – słownictwa polskiego dramatu współczesnego:

Wyrazy	Rangi w tekście	Rangi kontrolne	Wyrazy	Rangi w tekście	Rangi kontrolne
serce	5	323	poeta	112	-
dusza	9	590	poezja	113	-
kwę	16	594	siła	114	379
sen	20	782	dzwon	118	-
duch	21	582	lira	120	-
koń	24	395	przeklęty	125	-
złoty	26	-	zbroja	134	-
spać	40	292	moc	143	-
wicher	51	467	ogromny	144	069
słuch	69	-	osobliwy	146	-
głos	72	467	rozkaz	151	916
ogień	78	714	trup	153	-
skrzydło	81	-	czarny	162	751
łza	85	-	zaklęty	213	-
piekło	86	-	krwawy	275	-
złoto	101	-	okropny	287	-
ból	103	-			

Z punktu widzenia potrzeb automatyzacji operacji dokumentacyjnych i problemów konstruowania języków deskryptorowych konieczne są również badania statystyczne nad współwystępowaniem w bezpośrednim sąsiedztwie w ciągu syntagmatycznym wybranych par (ewentualnie trójek) wyrazów, terminy naukowe są bowiem często terminami wielowyrazowymi i tu nie wystarczą zwykłe słowniki frekwencyjne podające częstość występowania pojedynczych tylko wyrazów.

Wszyscy zainteresowani budową słowników frekwencyjnych powinni zaznajomić się z podaną literaturą przedmiotu, omawianie tych spraw wykracza bowiem poza ramy podręcznika.

Najlepszą ilustracją praktycznego wykorzystania wiadomości o statystycznej strukturze tekstu może być opowiadanie Arthura Conan Doyle'a pt. *Tańczące sylwetki*, gdzie Sherlock Holmes odkodował następujące teksty:



O tym, jak tego dokonał, stosując wiedzę lingwistyczną, dowiemy się z jego opowiadania:

„Zrobiłem jednak doświadczenie. Figurki mogły oznaczać przecież poszczególne litery. Zastosowałem wszystkie reguły, jakie rządzą tajnym pismem. I wówczas rozwiązanie systemu okazało się stosunkowo łatwe. Pierwsza zaszyfrowana wiadomość, jaką mi pokazano, była bardzo krótka. Niewiele mogłem się w niej zorientować, z wyjątkiem tego, że symbol:



oznacza literę E. Jak pan wie, E jest najczęściej używaną literą alfabetu angielskiego. Występuje bardzo często. Nawet w krótkim zdaniu powtarza się wielokrotnie. Z piętnastu figurek – symboli pierwszego zawiadomienia cztery były jednakowe. Istniała więc podstawa, aby uznać je za E. Co prawda część figurek E trzymała chorągiewki, a część nie. Z układu jednak chorągiewek można było wnioskować,

iż służą one do podziału zdania na poszczególne wyrazy. Przyjąłem to za hipotezę i zanotowałem, iż E wyrażone jest taką oto figurką:



W tym miejscu napotykamy jednak w badaniach na zasadniczą trudność. Otóż kolejność występowania w języku angielskim innych liter oprócz E w żadnym wypadku nie jest wyraźnie ustalona i jakakolwiek przewaga jednych liter nad drugimi w przekroju strony druku może ulec zupełnemu odwróceniu w pojedynczym krótkim zdaniu. Na ogół można przyjąć następującą kolejność: T, A, O, I, N, S, H, R, D i L. Ale T, A, O oraz I znajdują się bardzo blisko siebie, stojąc mniej więcej na jednym poziomie. Aby więc odkryć znaczenie szyfru, trzeba by badać każdą możliwą ich kombinację. To beznadziejna praca. Dlatego też czekałem na dalszy materiał. Podczas naszej drugiej rozmowy Mr. Hilton Cubitt dodał mi dwa dalsze, krótkie zaszyfrowane zdania i jedno zawiadomienie. Ujęte ono było, wobec braku chorągiewki w ręku figurki, w jednym słowie. Oto te rysunki. Teraz miałem już w pojedynczym, pięcioliterowym słowie dwa E, umieszczone na drugim i czwartym miejscu. Mogło to być słowo sever (zerwać), lever (dźwignia) lub never (nigdy). Bez wątpienia w grę wchodziło ostatnie słowo, które mogło stanowić odpowiedź na wezwanie. Okoliczności zaś wskazywały, że była to odpowiedź napisana przez lady Cubitt.

Uznając ten wniosek za słuszny, mogłem uznać znaki:



za odpowiadające literom N, V, R.

Jednak w dalszym ciągu piętrzyły się przede mną poważne trudności. Przypadkiem tylko szczęśliwa myśl dostarczyła mi kilku dalszych liter. Przypuszczałem, iż wezwanie otrzymuje lady od kogoś, kto znał ją dobrze już dawniej. Przyszło mi więc do głowy, czy słowo szyfru, zawierające trzy litery pomiędzy dwoma E, nie oznacza imienia Elsie. Przeprowadziłem badania i okazało się, iż słowo to stanowiło zakończenie zawiadomienia, trzykrotnie powtarzającego się. Było to więc z pewnością jakieś wezwanie skierowane do Elsie. W ten sposób rozszyfrowałem znaki odpowiadające literom L, S oraz I. Cóż jednak mogło zawierać samo wezwanie? Słowo stojące przed Elsie składało się tylko z czterech liter i kończyło na E. Musiało to więc być słowo come. Próbowałem wszelkich innych kombinacji czteroliterowych, nie mogłem jednak znaleźć żadnej, która by pasowała. W rezultacie otrzymałem symbole oznaczające C, O oraz M. Wobec tego powstała szansa na rozwiązanie pierwszego szyfrowego zawiadomienia. Podzieliłem je na słowa. W miejsce nie znanych mi jeszcze znaków powstawałem kropki. Powstał więc następujący tekst:

.M.ERE..E SL.NE.

Pierwszą literą mogło być chyba tylko A. Stanowiło to ważne odkrycie, ponieważ A występuje w tym krótkim zdaniu co najmniej trzykrotnie. W drugim słowie brakującą literą okazało się H.

W ten sposób zdanie przyjęło taką postać:

AM HERE A. E SLANE.

Wypełnienie pustego miejsca w nazwisku samo wprost się narzucało. Oto więc całe zdanie po wstawieniu B oraz Y: AM HERE ABE SLANEY (Jestem tutaj Abe Slaney). Oznaczyłem już tak wiele liter, iż mogłem przystąpić z dużą dozą pewności do rozwiązywania drugiego zawiadomienia. Wypisałem w nim znane mi litery:

A. ELRI. ES

Nadać sens temu zdaniu mogłem jedynie poprzez wstawienie T oraz G w miejsce brakujących liter. Zakładałem bowiem, iż w tym wypadku chodzi o nazwę domu lub gospody, w której zatrzymał się piszący.

Otrzymałem więc tekst:

AT ELRIGES (U Elrige'a)

(...) Hilton Cubitt przysłał mi ostatnie szyfrowane zawiadomienie Slaneya, wstawiając znane mi litery ułożyłem takie zdanie:

ELSIE. RE. ARE TO MEET THY GO.

Po dodaniu P i D uzyskałam pełny tekst:

ELSIE PREPARE TO MEET THY GOD!

(Elizo przygotuj się na spotkanie Twego Boga!)”

Znając już kod, mógł teraz Sherlock Holmes wysłać do przestępcy następujący list:



C O M E H E R E A T O N C E

(Przyjdź tu natychmiast!)

(tekst przytaczam za A.C. Doyle: *Tańczące sylwetki*, Poznań 1967.)

Literatura

Guiraud P.: *Zagadnienia i metody statystyki językoznawczej*. Warszawa 1966.

Sambor J.: *Słowa i liczby. Zagadnienia językoznawstwa statystycznego*. Wrocław 1972.

Sambor J.: *Językoznawstwo statystyczne dla pracowników informacji naukowej*. Warszawa 1978.

Hammerl R., Sambor J.: *Statystyka dla językoznawców*. Warszawa 1990.

Hammerl R., Sambor J.: *O statystycznych prawach językowych*. Warszawa 1992.

Sambor J.: *Aspekty ilościowe systemu języka*. w: *Encyklopedia kultury polskiej XX wieku*, t. 2: *Współczesny język polski* (red. Bartmiński J.) Wrocław 1993.

XVIII

Procesy informacyjne

W poprzednich rozważaniach o języku zajmowaliśmy się strukturą systemu językowego, teraz zajmiemy się problemami związanymi z **użyciem języka**, a więc realizacją systemu w konkretnych aktach komunikacyjnych.

Warunkiem koniecznym wszelkich procesów językowych jest **kompetencja językowa i komunikacyjna** uczestników tych procesów, czyli znajomość systemu językowego oraz zasad jego użycia, a więc wyboru środków językowych dostosowanych do konkretnych **sytuacji komunikacyjnych**. To dostosowanie umożliwia odpowiednią budowę języka naturalnego, a więc jego hierarchiczna struktura paradygmatyczna oraz obecność wyrażen i struktur synonimicznych. Na tym właśnie polega **bogactwo języka naturalnego**. Języki o strukturze płaskiej i nieposiadające synonimii są systemami ubogimi w tym sensie, że daną informację można w nich wyrazić tylko na jeden sposób, niezależnie od sytuacji komunikacyjnej. Takimi językami są przeważnie różnego typu języki sztuczne.

W opisie **kompetencji językowej** wyróżnia się **kompetencję gramatyczną**, polegającą na opanowaniu formalnej strony języka i umożliwiającej tworzenie i rozpoznawanie tekstów poprawnych składniowo (gramatycznie) oraz **kompetencję leksotaktyczną**, polegającą na opanowaniu także semantycznych, kontekstowych i stylistycznych ograniczeń łączliwości wyrażen. **Kompetencja semantyczna** pozwala użytkownikowi kodować i dekodować informację, ustalać relacje semantyczne między wyrażeniami, używać zwrotów idiomatycznych oraz dokonywać odpowiednich transformacji semantycznych tekstu. Kompetencja językowa użytkownika realizuje się w konkretnych sytuacjach komunikacyjnych – taką realizację nazywa się **performancją**. Warunkiem odpowiedniej realizacji jest odpowiednia **kompetencja komunikacyjna**. Zakres kompetencji komunikacyjnej i kompetencji językowej jest różny u różnych użytkowników języka, a odpowiedni stopień opanowania reguł jest warunkiem **powodzenia komunikacji**, a więc osiągnięcia przez nadawcę komunikatu **celu komunikacyjnego** oraz tzw. **fortunności komunikatu**.

Zasadami użycia języka w konkretnych sytuacjach komunikacyjnych zajmuje się stosunkowo nowy dział lingwistyki, **pragmatyka lingwistyczna**. Szczególną uwagę regułom rządzącym komunikacją poświęcili w swych badaniach

kognitywiści, reprezentujący naukę zajmującą się procesami poznawczymi. Reguły te, sformułowane na podstawie obserwacji językowej komunikacji ludzi, mają częściowo charakter dyrektyw i określają, w jaki sposób powinni się pod względem komunikacyjnym zachowywać uczestnicy konwersacji, aby ich zachowanie było efektywne, czyli zgodne z regułami racjonalnego działania.

Zwróćmy uwagę na to, że za podstawowy proces informacyjny przyjmuje się **konwersację**, a więc wzajemną wymianę komunikatów między uczestnikami, którzy na przemian pełnią rolę nadawcy i odbiorcy komunikatów, przy czym za kanoniczny proces komunikacji uznaje się konwersację między uczestnikami obecnymi w tej samej przestrzeni komunikacyjnej, tak że mogą się wzajemnie obserwować i nawzajem kontrolować proces nadawania i odbierania komunikatu, a dzięki uzyskanym w ten sposób informacjom o przebiegu tych procesów, na podstawie komunikatów językowych lub obserwacji partnera konwersacji, mogą modyfikować zakładane na wstępie i realizowane **strategie komunikacyjne**. Tak więc, między uczestnikami procesu komunikacyjnego dochodzi do **kooperacji**. Zasadę tę sformułował Herbert Paul **Grice** (1975, cyt. za *Współczesny język polski w: Encyklopedia kultury polskie*):

„Uczyń swój udział konwersacyjny w przewidywanym dla niego momencie takim, jakiego wymaga zaakceptowany cel lub kierunek rozmowy, w którym jesteś zaangażowany”.

Zasada ta rozpisana została przez Grice’a na tzw. **maksymy konwersacyjne**:

- **Maksyma ilości:**

Uczyń swoją wypowiedź na tyle informatywną, na ile jest to konieczne.
Nie czynj swojej wypowiedzi bardziej informatywnej, niż to konieczne.

- **Maksyma jakości:**

Nie mów tego, co uważasz za nieprawdziwe.
Nie mów tego, czego nie mógłbyś uzasadnić.

- **Maksyma odniesienia (relewancji):**

Mów na temat.

- **Maksyma sposobu:**

Mów jasno:
Unikaj niejasnego wyrażania się.
Unikaj niejednoznaczności.
Unikaj rozwlekłości (mów zwięźle).
Mów w sposób uporządkowany.

Zasady sformułowane przez Grice’a uzupełnił Geoffrey Leech (1980) o tzw. **maksymy grzeczności**:

taktu, wielkoduszności, aprobaty, skromności, zgodności, sympatii i zrozumienia.

Do zasad tych odniesiemy się w dalszych rozważaniach, choć nie będziemy ich szczegółowo oceniać. Zauważmy jednak, że choć zasady te nazwane zostały **zasadami zachowania konwersacyjnego**, to odnoszą się one właściwie tylko do nadawcy, dotyczą więc procesu nadawania komunikatu.

Nadawanie komunikatu

Nadawanie komunikatu to proces informacyjny, w którym nadawca produkuje sygnał będący materialnym nośnikiem informacji, którą chce przekazać danemu odbiorcy – adresatowi. Warunki, jakie muszą być spełnione to:

- środowisko, które może być odpowiednio przekształcone przez nadawcę, a więc takie, które pozwoli na produkowanie jednostek dyskretnych, wystarczająco różniących się od siebie, a zarazem takie, które umożliwi przemieszczenie przez nadawcę wyprodukowanego sygnału do odbiorcy, a więc zapewniające odpowiednią trwałość wprowadzonych przez nadawcę zmian. Środowisko to tworzy **kanal informacyjny**, a zmienione przez nadawcę jego elementy **sygnał informacyjny**, czyli materialną stronę komunikatu (formę wyrażenia). W komunikacji w języku naturalnym w subkodzie mówionym środowiskiem takim jest powietrze, sygnałem fala akustyczna wzbudzona przez nadawcę, a w subkodzie pisanym dowolna substancja pozwalająca na wytworzenie zmian będących źródłem zmian w fali świetlnej. Możliwe jest wykorzystanie także innych substancji, których ukształtowanie może być źródłem bodźców dotykowych (na przykład pismo niewidomych – alfabet Braile’a);

- **nadawca**, wyposażony w odpowiednie kompetencje umożliwiające wprowadzenie odpowiednich zmian w środowisku, a więc produkowanie sygnału informacyjnego. Niezbędne są do tego narzędzia służące do ukształtowania sygnału (odpowiedniej obróbki). Może to być aparat głosowy nadawcy – w wypadku tekstu mówionego, odpowiednie narzędzie w wypadku tekstu wizualnego, takie jak pióro, ołówek, rylec – w wypadku tekstu pisanego utrwalonego, źródło światła – w wypadku tekstu wizualnego, własne ciało – ręce, w języku głuchoniemych. "Narzędziem" pozwalającym na takie ukształtowania sygnału, aby stał się nośnikiem informacji, którą nadawca zamierza przekazać, jest **język (kod)**, który jest funkcją przyporządkowującą informacji sygnał. Nadawca musi też kontrolować stan kanału informacyjnego podczas nadawania komunikatu, sprawdzać, czy nie występuje w nim **szum informacyjny**, a więc, czy nie nakładają się na sygnał jego komunikatu sygnały, które mogłyby zniekształcać nadawany sygnał i tym samym powodować **straty informacji**. W razie wystąpienia takiego niebezpieczeństwa musi odpowiednio korygować swój sygnał, na przykład wzmacniać go (np. mówić głośniej) lub zapewnić mu odpowiednią **redundancję** (np. mówić wyraźniej, wolniej, powtarzać komunikat lub jego części) – przypomnijmy, że odpowiedni stopień redundancji powinien być zapewniony przez odpowiednią strukturę kodu. W przypadku gdy komunikat nadawany jest w sytuacji, gdy nadawca nie może bezpośrednio kontrolować procesu jego odbierania, powinien przewidzieć warunki, w jakich komunikat będzie znajdował się w czasie dzielącym jego nadawanie od odbioru oraz warunki, w jakich będzie odbierany, po to, aby odpowiednio zapobiec jego ewentualnemu zniekształceniu, a więc i stratom w informacji. W sytuacji bezpośredniego odbioru nadawca powinien kontrolować nie tylko warunki, w jakich komunikat jest przesyłany, a więc nie tylko kanał informacyjny, ale również proces odbierania informacji. Przede wszystkim powinien pilnować, żeby

w trakcie przesyłania sygnału nie został zerwany **kontakt** między nadawcą a adresatem komunikatu i kontrolować, czy kontakt taki istnieje, poprzez nadawanie sygnałów wymagających potwierdzenia odbioru przez adresata. Służą temu liczne środki językowe pełniące **funkcję fatyczną**. Nadawca powinien też sterować procesem odbierania komunikatu, tak aby komunikat odbierany był zgodnie z intencją nadawcy i przewidzianą przez niego strategią informacyjną. Służą do tego w języku naturalnym wyspecjalizowane środki – wykładniki funkcji metatekstowej i metainformacyjnej.

Przed rozpoczęciem nadawania komunikatu nadawca musi wiedzieć, czy **adresat komunikatu** jest w stanie komunikat odebrać, a więc:

- czy spełnione są odpowiednie warunki fizyczne dotyczące przesyłania sygnału informacyjnego,
- czy adresat może odebrać sygnał, a więc czy jest wyposażony w odpowiedni instrument informacyjny,
- czy adresat może sygnał odkodować, a więc czy zna język, w którym został nadany komunikat, i czy może dokonać odpowiedniej segmentacji sygnału (zapewniają to systemowe wykładniki delimitacji tekstu) oraz czy może wydzielonym fragmentom przyporządkować odpowiednie informacje.

I tak, na przykład, jeśli adresat nie może słyszeć, bo jest głuchy lub znajduje się zbyt daleko jak na donośność sygnału akustycznego, nadawca nie może posłużyć się przy nadawaniu komunikatu w języku naturalnym subkodem akustycznym, jeżeli adresat nie może widzieć, bo jest niewidomy lub znajduje się w warunkach, gdzie nie jest możliwe wyemitowanie fali świetlnej, nie można posłużyć się pismem. Nie można też posłużyć się subkodem wizualnym języka naturalnego, jeżeli adresat zna tylko subkod akustyczny danego języka naturalnego, nie zna więc reguł przyporządkowania między treścią znaku a tą właśnie jego formą (nie umie czytać).

W wypadku posłużenia się językiem naturalnym wiedza nadawcy o adresacie jego komunikatu musi obejmować także inne elementy. Przede wszystkim musi dysponować odpowiednią wiedzą umożliwiającą **nawiązanie kontaktu** komunikacyjnego, wybór bowiem odpowiednich zwrotów inicjujących kontakt we wszystkich językach naturalnych uzależniony jest od **relacji społecznych między uczestnikami sytuacji komunikacyjnej**. Relewantne dla komunikacji elementy tej wiedzy są różne dla różnych języków naturalnych, zawsze mają jednak wpływ nie tylko na wybór wykładników pełniących funkcję fatyczną, a więc na przykład inicjujących kontakt, podtrzymujących kontakt i kończących komunikat, ale również na wybór wykładników budujących tekst spójny i zapewniających tożsamość denotacji – na przykład form zaimków, a także na wybór wykładników funkcji impresywnej i ekspresywnej, a nawet wykładników obiektów będących przedmiotem komunikacji. W tym ostatnim wypadku w niektórych językach niezbędna dla nadawcy jest również wiedza o relacji łączącej obiekt komunikacji z adresatem komunikatu, a także z nadawcą komunikatu. Cała ta wiedza niezbędna jest nadawcy komunikatu po to, żeby mógł uczynić zadość **warunkom grzeczności**, które musi spełniać komunikat, co często stanowi o po-

wodzeniu komunikacji. Spełnieniu tych warunków służą tzw. wykładniki **honoryfikatywności**, w niektórych językach tworzące bardzo rozbudowane kategorie gramatyczne i leksykalne. Komunikując się, trzeba więc wybrać odpowiednią formę **adresatywną**, czyli sposób zwracania się do adresata. Na przykład w języku polskim relewantne są:

- przynależność nadawcy i odbiorcy do przedziałów wiekowych: dziecko – młody – dorosły:

- dzieci zwracają się do siebie przez formy drugiej osoby i zaimek *ty* i *wy*, gdy adresatów jest więcej niż jeden z tego samego przedziału wiekowego, lub gdy adresat jest jeden, ale osoby, o których się mówi, a do których należy również adresat, także są dziećmi;
- osoby młode coraz częściej tak jak wyżej, jeżeli komunikacja nie zachodzi w sytuacji oficjalnej, np. sprzedawca – klient, urzędnik – interesant;
- dzieci do osób dorosłych nienależących do kręgu rodzinnego zwracają się używając form *pan*, *pani*, *państwo*;
- dorośli zwracają się do dzieci przez formy drugiej osoby;
- dorośli zwracają się do siebie przez formy trzeciej osoby i *pan*, *pani*, *państwo*, chyba że zawarta przez nich umowa ustanawia formę *ty* wzajemną lub jednostronną (np. dorośli do dorosłych kolegów swoich dzieci, do których wcześniej, gdy byli dziećmi, mówili *ty*, mogą zwracać się dalej w tej samej formie, jeżeli relacje między nimi są bliskie lub jeżeli oni wyrażają taką wolę). W stosunku do osób należących do kręgu rodzinnego obowiązują inne zasady, zależne od rodzaju relacji pokrewieństwa lub powinowactwa. Inne zasady obowiązują również w pewnych sytuacjach oficjalnych;

- przynależność nadawcy i adresata do odpowiedniej grupy społecznej, w tym kręgu rodzinnego:

W rodzinie i kręgu osób bliskich rodzinie dla nadawcy relewantne są pewne relacje, które mają swój leksykalny wykładnik, np. *mama*, *tata*, *babcia*, *dziadek*, których adresatywne formy wołacza są nienacechowane emocjonalnie (odpowiedniki nacechowane emocjonalnie to np. formy *matko*, *ojcze*, *babko*, *mamusiu*, *mamuniu*, *mateńko*, *tatusiu*, *tatku*, *babuniu*, *babko*, *dziadziu*. Zwroty *ciociu*, *wujku* funkcjonują nie tylko jako wykładniki relacji rodzinnych, ale także używane są przez dzieci w stosunku do osób dorosłych bliskich rodzinie oraz przez osoby dorosłe (ale nie zawsze), jeśli tak się zwracały do tych osób będąc dziećmi. Wszystkim tym zwrotom adresatywnym towarzyszy forma 2 osoby liczby pojedynczej czasownika, obecnie prawie zawsze w stosunku do matki i ojca, albo 3 osoby liczby pojedynczej, np.:

Mamo, czy dasz mi na wycieczkę?

Mamo, czy mamusia da mi na wycieczkę?

W stosunku do innych osób z kręgu rodzinnego częściej używa się 3 osoby liczby pojedynczej, np.:

Czy była ciocia na tej wystawie?

To tylko przykłady reguł wyboru form adresatywnych w języku polskim – jest ich o wiele więcej. Na wybór form adresatywnych i zaimkowych wykładników adresata nakłada się jeszcze **rodzaj sytuacji komunikacyjnej**, na przykład: intymna, przy świadkach, oficjalna, a także **intencja nadawcy**: czy chce być grzeczny, uprzejmy, oficjalny, niegrzeczny, arogancki, czy też chce, na przykład, obrazić adresata. Naruszenie reguł wyboru wykładników adresata komunikatu traktowane jest jako poważne złamanie reguł komunikacyjnych i powoduje dyskomfort uczestników procesu komunikacyjnego, co ma wpływ na przebieg i powodzenie komunikacji.

W wypadku języka polskiego wiedza nadawcy komunikatu o adresacie niezbędna do wyboru form grzecznościowych używanych w komunikacie obejmuje stosunkowo niewiele elementów. Znacznie więcej jest ich w regułach komunikacji w innych językach, na przykład w japońskim, chińskim, wietnamskim czy koreańskim.

Dla uzmysłowienia czytelnikowi, jak różne elementy mogą tu być relewantne i niezbędne, bo przesądzają o wyborze odpowiednich form językowych, przytoczymy zaczerpnięty z książki Romana Huszczy p.t. *Honoryfikatywność* przykład reguł języka wietnamskiego. W wielu językach wybór ten bywa obligatoryjny, bo często, zgodnie z zasadami budowy kodu optymalnego, wykładniki takie mają charakter gramatyczny, przy czym bywa i tak, że w danym języku nie ma form nienacechowanych – grzecznych, ale neutralnych.

„Na tle języków wschodnioazjatyckich najbardziej oryginalny i uporządkowany system zaimków osobowych ukształtował się w języku wietnamskim, gdzie w roli zaimków osobowych używane są w bardzo szerokim zakresie stopnie pokrewieństwa. Nazwy te używane są nie tylko jako zaimki drugiej czy trzeciej osoby, ale także jako zaimki pierwszej osoby, gdyż opisywany w gramatykach jako właściwy zaimek pierwszej osoby *tôi* ‘ja’ wcale nie zawsze jest neutralny i w niektórych sytuacjach z uwagi na zbyt niską wartość honoryfikatywności nie może być używany przez nadawcę.

Sytuacja taka zachodzi w momencie, gdy mówiący zwraca się do swoich rodziców i wtedy zobowiązany jest używać w funkcji zaimka pierwszej osoby wyrazu *con* ‘dziecko’. (Warto dodać, że w katolickich modlitwach wietnamskich zaimek *con* jest również obligatoryjnym zaimkiem pierwszej osoby przy zwracaniu się do Boga.) Podobnie starsza osoba w rozmowie z dzieckiem, niekoniecznie przy tym spokrewnionym, gdy tytułowana jest przez nie zaimkiem *ông* dosł. ‘dziadek’ używa zaimkowo tego wyrazu w funkcji ‘ja’.

Dzieci w rozmowie z osobami sędziwymi używają w tej funkcji zaimka *cháu* o dosłownym znaczeniu ‘wnuk, wnuczka’. Ludzie dojrzały mówią o sobie używając zaimków *anh* dosł. ‘starszy brat’, *chị* ‘starsza siostra’, *em* ‘młodszy brat; młodsza siostra’ (w kontaktach osób młodszych wiekiem, ale dorosłych, z przestrzeganiem zasady starszeństwa) lub *cháu* ‘dziecko; siostrzeniec, siostrzenica; bratanek, bratanica’ (w kontaktach osoby dojrzałej z reprezentantami pokolenia dziadków lub wujków). Zaimek *tôi* traktowany w opisach jako zwykły zaimek pierwszej osoby jest honoryfikatywnie nienacechowany – co nie oznacza,

że w każdej sytuacji jest uważany za wystarczająco grzeczny – ale pozostaje w opozycji do zaimków *minh*, *tó* i *tao*, komunikujących bardzo poufały stosunek mówiącego do rozmówcy.

Jak widać, zasób zaimków pierwszej osoby w języku wietnamskim jest zdecydowanie najliczniejszy i jeżeli uwzględnimy fakt, że w funkcji tej używane są także imiona, to musimy tu mówić o klasie otwartej. Imiona wietnamskie są używane zaimkowo w kontaktach między osobami w jakiś sposób zżytymi, ale nie należy tego poziomu honoryfikatywności utożsamiać z polską formą *ty*, gdyż użycie takie nie ma charakteru poufałego i przypomina raczej polską formę *pan*. Używa się ich we wszystkich wartościach kategorii osoby, a więc także w odniesieniu do słuchacza rozmowy i tylko pod tym względem można zauważyć analogię z polszczyzną, chociaż polskie użycia imion są, jak powiedzieliśmy, połączone z zasadniczym obniżeniem rangi towarzyskiej ich nosiciela.

W opisie wietnamskich zaimków osobowych musimy wyraźnie oddzielić od siebie dwie sytuacje ich użycia, mianowicie użycie w obrębie rodziny oraz użycie w kontaktach pozarodzinnych. W pierwszej sytuacji zaimkowe użycie nazw pokrewieństwa zgodne jest z pozycją rozmówców zajmowaną w hierarchii rodzinnej i w zasadzie każda nazwa pokrewieństwa występuje wtedy w funkcji zaimka pierwszej i drugiej osoby (a także trzeciej w stosunku do słuchacza), np. *cha* ‘ojciec’ *mẹ* ‘matka’, *ông* ‘dziadek’, *bà* ‘babcia’, *bác* ‘starszy stryj, wuj lub ciocia’, *chú* ‘młodszy stryj’, *thím* ‘stryjenka’ *cậu* ‘wujek (młodszy brat matki)’ *mợ* ‘ciocia (żona młodszego brata matki)’ *anh* ‘starszy brat’, *chị* ‘starsza siostra’, *em* ‘młodszy brat lub siostra’, *chắt* ‘prawnuk, prawnuczka’, *cháu* ‘wnuczek, wnuczka; siostrzeniec, siostrzenica, bratanek, bratanica’, *con* ‘dziecko’. Tzw. właściwe zaimki osobowe mają tu dosyć ograniczony zakres użycia, gdyż np. użycie formy *tôi* w takim kontakcie jest znaczące jako złamanie reguł etykiety. Mówiący może użyć takiej formy np. w sytuacji konfliktowej, gdy chce zakomunikować adresatowi swoistą chwilową alienację od więzi rodzinnej, forma *tao* jest formą bardzo poufałą używaną przez starsze rodzeństwo do młodszego i w zwykłej sytuacji brzmi niegrzecznie. W sytuacjach konfliktowych formy tej używają rodzice w stosunku do dzieci, użycie *tao* w odwrotnym kierunku brzmi zdecydowanie wulgarnie i jest dla rodziców skrajnie obraźliwe. Forma *minh* jest formą bardzo poufałą między spokrewnionymi rówieśnikami.

W użyciu pozarodzinnych nazwy stopni pokrewieństwa tworzą niezwykle skomplikowany system zaimkowy z wieloma poziomami honoryfikatywności, w którym kryteria wyboru formy zaimkowej są oparte na różnicy wieku, płci, przynależności pokoleniowej, pozycji społecznej czy służbowej, hierarchii towarzyskiej, oraz na stopniu zażyłości między rozmówcami. Ponieważ formy orzeczenia w zdaniu wietnamskim nie posiadają gramatycznej kategorii honoryfikatywności, zaimki osobowe są jedynymi wykładnikami poszczególnych poziomów honoryfikatywności. Jeśli idzie o liczbę takich poziomów grzeczności to zaimki drugoosobowe w języku wietnamskim można ułożyć w następujący sposób, idąc w kolejności od poziomu najwyższego do najniższego w rezultacie otrzymamy siedem takich poziomów:

- cụ* forma używana w stosunku do osób sędziwych z pokolenia dziadków lub osób starszych wysoko usytuowanych w hierarchii
- ông* (dla r. m.) ; *bà* (dla r. ż.) formy używane w stosunku do osób dorosłych, nieco młodszych niż osoby tytułowane *cụ*, nieznajomych, żonaty i mężatek, powyżej 40 roku życia aż po wiek sędziwy
- bác* (dla obu rodzajów) forma używana w stosunku do osób obcych znacznie starszych od mówiącego w wieku dojrzałym, np. jeśli są rówieśnikami rodziców rozmówcy młodszych wiekiem, stanu wolnego, od lat kilkunastu aż po wiek dojrzały
- anh* (r. m.) ; *chị* (r. ż.) formy stosowane do młodzieży od lat 18 do około 30
- em* forma używana w stosunku do młodszych obcych dzieci i do młodych osób młodszych wiekiem od mówiącego, w jakiś sposób znanych
- mày* zaimek poufały, stosowany wobec osób bardzo blisko zaprzyjaźnionych, w stosunku do pozostałych brzmi zdecydowanie obraźliwie

Ponieważ w przeciwieństwie do polskich zaimków honoryfikatywnych *pan*, *pani* zaimki wietnamskie mogą być opuszczane i elipsa ta uwarunkowana jest jedynie regułami tematyczno-rematycznej struktury wypowiedzi, opozycja wartości honoryfikatywnej często ulega tu neutralizacji. Opisując teoretycznie możliwe sytuacje użycia wietnamskich zaimków osobowych poza rodziną musimy uwzględnić także i zaimki pierwszej osoby w rezultacie otrzymamy siedemnaście typów zróżnicowania form zaimkowych w dialogu (zaimki I os. podane są w nawiasie):

1. *tôi* : *tôi*
cụ ← : → *ông* dla r.m.
dla r.m. i r.ż. : → *bà* dla r.ż.

Takie formy zaimków używane są w sytuacji, gdy uczestnikiem rozmowy jest bardzo sędziwa osoba i osoba znacznie młodsza od niej, ale dojrzała, np. siedemdziesięcioletni mężczyzna i czterdziestolatka.

2. *tôi* : *tôi*
cụ ← : → *anh* dla r.m.
 : → *chị* dla r.ż.

Tak mogą zwracać się do siebie bardzo sędziwa osoba i osoba od niej młodsza, ale dorosła, np. siedemdziesięcioletni mężczyzna lub kobieta i osoba w wieku 20-50 lat.

3. *tôi* : *cháu*
cụ ← : → *cháu*

Takich zaimków używa np. bardzo sędziwa osoba i dziecko lub osoba bardzo młoda (18-30 lat).

4. *tôi* : *tôi*
 ông ← : → *ông* dla r.m.
 bà ← : → *bà* dla r.ż.

W taki sposób zwracają się do siebie starsze osoby lub wysoko postawione w hierarchii, albo osoby dorosłe całkowicie sobie obce. Może się tak również zwracać np. jakiś urzędnik do klienta nawet w młodszym wieku albo przygodni dorośli pasażerowie w pociągu itd.

5. *ông, bà* : *cháu*
 ông dla r.m. ← : → *cháu* dla obu rodzajów
 bà dla r.ż. ← :

Tak mogą mówić do siebie np. osoba w starszym wieku, z pokolenia dziadków i małe, nie spokrewnione z nią dziecko.

6. *tôi* : *tôi*
 ông ← dla r.m. : → *ông* dla r.m.
 bà ← dla r.ż. : → *chị* dla r.ż.

Takie formy zaimków używane są w rozmowie przez osobę starszą lub wyżej postawioną i osobę młodszą lub niższą pod względem pozycji społecznej, np. profesor i student.

7. *tôi* : *cháu*
 chú ← : → *cháu* dla obu rodzajów
 cô ← : → *cháu* dla obu rodzajów

Tak mówi osoba dojrzała ale młodszą od osób tytułowanych *ông* i *bà*, i osoba młoda, w jakiś sposób znajoma lub przypadkowo spotkana na ulicy.

8. *tôi* : *cháu*
 bác dla obu rodzajów ← : → *cháu* dla obu rodzajów

Są to zaimki używane w rozmowie między osobą w dojrzałym wieku i osobą od niej starszą, szanowaną w jakiś sposób znajomą.

9. *tôi* : *tôi*
 anh ← dla r.m. : → *anh* dla r.m.
 chị ← dla r.ż. : → *chị* dla r.ż.

W taki sposób zwracają się do siebie rówieśnicy, należący do młodszego pokolenia, pozostający w stosunkach niezbyt poufanych, bez bliskiej znajomości; dla pewnych środowisk taki sposób użycia zaimka jest neutralny, np. studenci, żołnierze, młodzi pracownicy itd.

10. *anh, chị* : *em*
 anh dla r.m. ← : *em* dla obu rodzajów
 chị dla r.ż. ← : →

Takich zaimków używają w rozmowie ludzie młodzi, których dzieli niewielka różnica wieku; rozmówcy młodszy mówiąc o sobie używają zaimka I osoby *em*, jeżeli jednak różnica jest niewielka zaimkiem takim może być także *tôi*.

11. *tôi* : *tôi*
chú dla r.m. ← : → *cô* dla r.ż.
: → *cháu* dla r.ż.

Takie formy zaimków występują w rozmowie dorosłego mężczyzny i młodszych wiekiem kobiet, jeżeli mężczyzna jest od nich starszy, ale nie osiągnął jeszcze wieku ich rodziców.

12. *tôi* : *tôi*
cô ← dla r.ż. : → *cô* dla r.ż.
cậu ← dla r.m. : → *cậu* dla r.m.

Takich zaimków używają młodzi ludzie, rówieśnicy, pozostający w pewnego rodzaju zażyłości, ale nie zaprzyjaźnieni ze sobą.

13. *mình, tớ* : *mình, tớ*
cậu ← dla obu rodzajów : → *cậu* dla obu rodzajów

W taki sposób zwracają się do siebie młodzi ludzie, rówieśnicy, np. koledzy i koleżanki w pracy; zaimek I osoby *mình* i jego synonim *tớ* używane przez mężczyzn i kobiety są bardziej poufale niż *tôi*.

14. *mình* : *em*
anh dla r.m. ←
chị dla r.ż. ← : → *em* dla obu rodzajów

Takich form zaimkowych używają młodzi ludzie, np. koledzy i koleżanki w pracy lub studenci przy niewielkiej różnicy wieku, gdy łączą ich stosunki mniej poufale niż w 13, ale bardziej poufale niż w 9.

15. *tao* : *tao*
mày ← : → *mày* dla obu rodzajów

Są to zaimki zdecydowanie poufale, używane przez osoby bardzo blisko zaprzyjaźnione.

16. *anh* : *em*
anh dla r.m. ← : → *em* dla r.ż.

Tak mogą mówić do siebie tylko osoby bardzo blisko związane ze sobą, np. zakochani, narzeczeni lub młodzi małżonkowie.

17. *tao* : *em*
anh dla r.m. ← : → *mày*

Takich form zaimkowych używają w rozmowie młody mężczyzna i młodszy od niego kolega żyjący z nim bardziej.

W obrębie omawianych tu zaimków wietnamskich używane poza rodziną nazwy pokrewieństwa pełnią, jak widać, rolę prototypowych wykładników pokoleniowej rangi adresata, gdyż przynależność pokoleniowa jest tutaj ważniejszym kryterium wyboru formy niż stopień znajomości.

(...) język wietnamski posiada nie tylko najliczniejszy zbiór zaimków pierwszej osoby, ale także i najwięcej poziomów honoryfikatywnych. Poza wymienionymi wyżej zaimkami pierwszej osoby używane są w wietnamszczyźnie także przestarzały nieco, poufały zaimek *qua*, autorytarny i uroczysty *đê*, poufałe formy *đây* oraz *ta* (jako wariant *tao*) i niektóre zaimki „tytułarne” typu *thây*, którego używają nauczyciele w stosunku do uczniów.”

Od relacji między nadawcą a adresatem zależy też **użycie wyrażen ekspresyjnych** dotyczących obiektów komunikacji, takich jak *babka*, *cizia*, *facet*, *gość*, *baba*, *babsztyl*, *babsko*, *babiszon*.

Warunki uprzejmości nałożone są także na sposób wykorzystywania materialnego nośnika komunikatu. Na przykład w wypadku tradycyjnego komunikatu pisanego tekst musi być odpowiednio rozmieszczony na stronie, w poczcie elektronicznej i bezpośredniej komunikacji komputerowej nie należy pisać tekstu wersalikami, bo uważa się to za formę niegrzeczną (odpowiednik krzyku). Przy komunikacji ustnej, a więc wtedy, gdy nadawca i adresat znajdują się w tej samej przestrzeni, ważne jest nie tylko napięcie głosu, ale też **dystans** (fizyczna odległość) między uczestnikami komunikacji. Dla różnych kultur i języków określony jest dość dokładnie dystans dla rozmowy intymnej, rozmowy między osobami bliskimi, obcymi, pozostającymi w hierarchicznej zależności oraz dla wystąpień oficjalnych i publicznych.

Pamiętać należy o tym, że w normalnej komunikacji werbalnej, gdy między uczestnikami procesu komunikacyjnego utrzymywany jest kontakt wzrokowy, dużo informacji przekazywanych jest równocześnie z komunikatem językowym za pomocą różnych środków niewerbalnych. Komunikat taki stanowi swojego rodzaju tekst mieszany, chociażby dlatego, że bardzo często funkcję denotacyjną, a więc identyfikującą elementy rzeczywistości, których dotyczy informacja, pełni **gesty**, np.

Daj mi to. Nie to, tamto.

gdzie denotaty zaimków muszą być wskazane pozawerbalnie. Odpowiednie gesty są skorelowane z pewnymi rodzajami komunikatu i należą do **językowego rytuału komunikacyjnego**, przy czym gesty te są uwarunkowane kulturowo (na przykład gest przysięgi, ślubowania, powitania).

Relevancja informacji

Nadawcy komunikatu niezbędna jest przede wszystkim wiedza o aktualnych **potrzebach informacyjnych adresata** oraz o jego wiedzy o elementach rzeczywistości wyznaczanych przez temat komunikatu. Komunikacja nigdy nie może się rozpocząć bez takich wstępnych założeń ze strony nadawcy, poczynionych na podstawie uzyskanych wcześniej informacji o adresacie, bezpośrednio od niego

lub pośrednio poprzez obserwację adresata lub znajomość jego środowiska informacyjnego. W tym sensie każdy komunikat pełniący podstawową funkcję komunikatu, czyli funkcję komunikacyjną, jest odpowiedzią skierowaną do kogoś i w założeniu ma być odpowiedzią na potrzeby informacyjne konkretnego adresata, nawet jeśli jest to adresat potencjalny. Dzięki tej wiedzy nadawca może uczynić zadość dyrektywie zawartej w pierwszej maksymie konwersacyjnej, maksymie ilości, a więc przekazać w komunikacie tyle i tylko tyle informacji (takich i tylko takich), ile jest potrzebne do osiągnięcia aktualnego celu informacyjnego.

I tak, na przykład, o swoim stanie zdrowia i samopoczuciu nie informujemy wszystkich, a tylko te osoby, które to interesuje, wyraziły *explicite* taką potrzebę, na przykład pytając *Jak się pani czuje?*, wiemy, że je to interesuje lub wynika z relacji między nadawcą komunikatu (np. z relacji bliskości emocjonalnej, rodzinnej lub relacji lekarz – pacjent). To, ile i jakich informacji na temat naszego samopoczucia udzielimy adresatowi naszego komunikatu, zależy od tego, ile i jakie informacje są potrzebne dla osiągnięcia **celu komunikacyjnego**. Skrajne sytuacje komunikacyjne w naszym przykładzie to:

- przypadkowo spotkany dalszy znajomy – odpowiadamy zazwyczaj *Dziękuję, dobrze*, bez względu na nasze faktyczne samopoczucie, bo takiej odpowiedzi wymaga **konwencja społeczna**, a same wypowiedzi pełnią przede wszystkim funkcję fatyczną,
- lekarz, któremu musimy szczegółowo opisać, co nam dolega.

Tak więc zawsze, w każdej sytuacji komunikacyjnej, powinniśmy wpierw poznać potrzeby informacyjne odbiorcy, dowiedzieć się, jakie informacje byłyby dla niego relewantne i przekazać mu te i tylko te informacje, chyba że nasza komunikacja ma pełnić inny cel niż cel informacyjny, na przykład ma podtrzymywać lub ustanawiać relacje społeczne.

Informacje przekazane adresatowi w komunikacie powinny być z jego punktu widzenia kompletne i nie zawierać szumu informacyjnego, mamy więc, zgodnie z maksymą odniesienia, **mówić na temat**. W przeciwnym razie wysiłek włożony w akt komunikacyjny będzie nieopłacalny i dla nadawcy i dla odbiorcy. Komunikat zawierający za dużo informacji kosztuje zbyt wiele i nadawcę i odbiorcę, kosztowne jest bowiem samo produkowanie i odbieranie sygnału informacyjnego. W wypadku odbiorcy dodatkowy koszt stanowi konieczna z jego punktu widzenia operacja **selekcji informacji**, do której nie zawsze jest odpowiednio pod względem informacyjnym przygotowany, wymaga ona bowiem zastosowania kryteriów wartościowania informacji, a więc skorzystania z innych informacji, którymi nie zawsze odbiorca dysponuje.

Przekazywanie informacji o danym elemencie rzeczywistości z odpowiednim do sytuacji komunikacyjnej stopniem dokładności umożliwia hierarchiczna struktura systemu semantycznego języka naturalnego, przy czym, jak pamiętamy, im węższy zakres wyrażenia, tym dokładniejsza identyfikacja i tym więcej informacji (treści) przyporządkowanej temu wyrażeniu.

Na przykład, odpowiedni urząd musimy poinformować o tym, że mamy psa, bo taka informacja jest potrzebna urzędnikom do celów podatkowych. To, ja-

kiej rasy jest pies, ich nie interesuje. Ale koło myśliwskie może interesować, czy mamy posokowca, zaś związek kynologiczny również jego rasa, na przykład to, że jest to wyżeł niemiecki. Wprawdzie w każdej z tych sytuacji, gdybyśmy udzielili informacji, że mamy wyżła niemieckiego, wynikałoby, że mamy psa (na zasadzie kompetencji językowej w zakresie relacji paradygmatycznych), to jednak w dwóch pierwszych wypadkach nasi adresaci uzyskaliby więcej informacji niż im potrzeba. Uzyskanie informacji relewantnej w wypadku automatycznego przetwarzania tekstu wymagałoby dodatkowego programu rozpoznającego relacje paradygmatyczne języka, a my musielibyśmy włożyć w komunikację więcej wysiłku – wyrażenie pies jest najkrótsze i dostęp do niego w naszej pamięci semantycznej jest najłatwiejszy.

Nadawca informacji musi też poczynić pewne założenia co do **wiedzy adresata** o rzeczywistości będącej przedmiotem aktualnej komunikacji, nie tylko po to, żeby nie nadawać informacji, które adresat już ma w swoim zasobie informacyjnym, ale także po to, żeby mógł odebrać przekazywane mu informacje. Adresat komunikatu nie będzie mógł bowiem odebrać i skorzystać z przekazanych mu informacji, jeśli nie będzie dysponował odpowiednią wiedzą o przedmiocie komunikacji, taką i tak ustrukturalizowaną, że umożliwi to włączenie nowej informacji w odpowiednie miejsce jego zasobów informacyjnych (systemu wiedzy). W przeciwnym razie komunikat do niego skierowany nie spowoduje żadnego przyrostu informacji w systemie odbiorcy – może być co najwyżej przez niego zapamiętany, ale nie przyswojony.

Powyższe zasady konwersacyjne odnoszą się także do takich sytuacji komunikacyjnych, w których w ogóle nie chodzi o przekazywanie informacji nowych dla uczestników procesu informacyjnego, a więc w sytuacjach, w których wartością jest nie informacja, tylko sam kontakt utrzymywany za pomocą przekazów językowych. W takich sytuacjach komunikaty pełnią w zasadzie przede wszystkim funkcję fatyczną i są do pewnego stopnia zrytualizowane.

Na przykład, przy przelotnym spotkaniu z sąsiadem na pytanie: *Co słychać?* Odpowiadamy: *Dziękuję, wszystko w porządku* albo *Jakoś leci*. Zasada ilości jest tutaj zachowana, bo w takiej sytuacji komunikacyjnej w ogóle nie chodzi o informację, a jedynie o potwierdzenie kontaktu i podtrzymanie relacji społecznych, co zostaje uczynione poprzez odpowiednie zachowanie językowe. A więc i w tym wypadku cel komunikacyjny zostaje osiągnięty.

Po ustaleniu celu konwersacji, rozpoznaniu zasobu wiedzy adresata, co jest niezbędne do wybrania strategii komunikacyjnej, ustaleniu, jakie informacje są dla adresata komunikatu relewantne, ustaleniu relacji społecznych niezbędnych do wyboru wykładników honoryfikatywności i form grzecznościowych, dzięki którym zostanie spełniona (lub świadomie naruszona) zasada grzeczności, nadawca komunikatu przystępuje do **formułowania komunikatu** – odwzorowywania w nim tych informacji o rzeczywistości, które zamierza przekazać. Wymaga to znalezienia w systemie języka, na podstawie własnej kompetencji językowej, ich wykładników i połączenia ich, zgodnie z regułami gramatycznymi danego języka w wyrażenia złożone, frazy, zdania i teksty.

Zwróćmy uwagę na to, że nadawca komunikatu, tworząc komunikat, musi dysponować nie tylko tą wiedzą o komunikowanej rzeczywistości, którą chce przekazać odbiorcy, bo uważa że takie informacje są dla niego relewantne, ale również **wiedzą relewantną językowo**, a więc taką, która jest relewantna ze względu na konieczność odwzorowania jej w obligatoryjnych w danym języku naturalnym wykładnikach kategorii gramatycznych.

I tak, na przykład, mówiąc w języku polskim, nadawca musi wiedzieć, czy informacja dotyczy jednego elementu rzeczywistości spośród elementów zbioru denotowanego przez dane wyrażenie, czy więcej niż jednego, w języku polskim obligatoryjne jest bowiem wyrażenie liczby, w opozycji jeden : więcej niż jeden, musi więc wybrać, na przykład, między formami

Przed domem rosło drzewo.

Przed domem rosły drzewa.

Przed domami rosło drzewo.

Przed domami rosły drzewa.

Podczas gdy na przykład w języku japońskim nie musi informować o liczbie elementów. Podobnie w języku polskim informacja musi być usytuowana w czasie, którego wykładniki są obligatoryjne i wskazują relacje między czasem zajścia zdarzenia, o którym się komunikuje a momentem nadawania komunikatu. Dodatkowo w formach liczby mnogiej czasu przeszłego musimy wiedzieć, czy wśród elementów zbioru, którego dotyczy informacja, znajdują się osoby płci męskiej, nadawca musi bowiem wybrać między formą męskoosobową i niemęskoosobową, np.

Wczoraj byliśmy w kinie.

Wczoraj byliście w kinie.

Odwiedziły nas znajome z wakacji.

Odwiedzili nas znajomi z wakacji.

Ważnym problemem jest takie odwzorowanie elementów rzeczywistości, których dotyczy informacja, żeby adresat informacji mógł je zidentyfikować (ustalić **odniesienie**). Jednoznaczność identyfikację umożliwiają tylko nazwy własne, ale, jak wiemy, język naturalny nie może składać się wyłącznie z nazw własnych. Używając nazw własnych, nadawca zawsze zakłada, że adresat zna ich odniesienie, a jeżeli w trakcie komunikacji okaże się inaczej, odniesienie to musi być ustalone, jak na przykład w poniższym dialogu:

Spotkałam wczoraj Ankę.

– Jaką Ankę?

– Ankę Kowalską.

– Nie znam jej.

– To siostra Janka. Tego, który był z nami na splotwie.

W wypadku, gdy element rzeczywistości nie ma nazwy własnej – a tak jest najczęściej – możemy go wskazać za pomocą wyrażenia oznaczającego zbiór,

do którego element ten należy oraz odpowiedniego wskaźnika identyfikującego dany element spośród wszystkich elementów należących do zbioru. Należy przy tym pamiętać o hierarchicznej strukturze pola semantycznego języka naturalnego i wybrać wyrażenie o odpowiedniej w danej sytuacji komunikacyjnej **informatywności**, na co zwracaliśmy uwagę, omawiając problem relewancji informacji. W bezpośredniej komunikacji, jeśli informacja dotyczy elementów rzeczywistości tworzących **konsytuację informacyjną**, możliwa jest identyfikacja przedmiotów komunikacji poprzez gest wskazujący, ale w takim wypadku pełny komunikat stanowi strukturę złożoną z tekstu języka naturalnego i **gestu**, będącego elementem innego kodu, np.

Podaj mi tę książkę. Stoi na tamtej półce.

Wskazanie elementu rzeczywistości, którego dotyczy komunikat, może być także dokonane przez wyrażenie nazywane **deskrypcją określoną**, np.

Rozmawiałam wczoraj z wychowawczynią Ani.

Opiekun naszego roku ma dyżur we czwartek.

Najstarszy syn Anny studiuje architekturę.

Tworząc dłuższy komunikat, nadawca do wskazania, że informacja dotyczy tych samych elementów rzeczywistości, o których już była mowa w komunikacie, i które wcześniej zostały zidentyfikowane, posługuje się przewidzianymi przez system języka zastępnikami uprzednio użytych w tej funkcji wyrażen. Mówiliśmy o tym, omawiając wykładniki spójności tekstu, tu tylko przypomnijmy, że mogą to być wyrażenia oznaczające rodzaj, do którego dany zbiór jest zaliczany, a więc szersze w stosunku do wyrażenia poprzednio użytego, np.:

Rozmawiałam wczoraj z wychowawczynią Ani. Nauczycielka bardzo ją chwaliła.

Najstarszy syn Anny studiuje architekturę. Chłopak jest naprawdę zdolny.

Do tego służą także zaimki anaforyczne, np.:

Rozmawiałam wczoraj z wychowawczynią Ani. Nauczycielka bardzo ją chwaliła.

Najstarszy syn Anny studiuje architekturę. Ten to ma szczęście w życiu.

Na ustalenie odniesienia zaimka w tekstach języka polskiego pozwala związek zgody, a więc odpowiednie wskaźnik więzi. Dzięki temu zaimek w ogóle może być pominięty, bo rolę taką mogą pełnić końcówki fleksyjne, np.:

Rozmawiałam wczoraj z wychowawczynią Ani. Bardzo ją chwaliła.

W innych językach naturalnych funkcję taką pełnią inne gramatyczne wykładniki.

Czasami nadawca komunikatu zapowiada, że odniesienie ustalone zostanie w dalszej części komunikatu, np.:

Cały wieczór tańczyłem z tą dziewczyną, którą wczoraj poznaliśmy na plaży.

Bywa i tak, że informacja dotyczy dowolnego elementu wskazanego zbioru, np.:

Muszę sobie kupić letnią sukienkę.

Czasami nadawca komunikatu nie może przekazać adresatowi informacji niezbędnej do identyfikacji elementu rzeczywistości, o którym mówi, bo sam jej nie ma. Brak tej wiedzy może zakomunikować adresatowi, np. w zdaniu

Przyniósł mi to jakiś student.

Może też nie przekazać informacji potrzebnej do identyfikacji, choć nią dysponuje, ale nie ma zamiaru jej udostępnić adresatowi, o czym go informuje, np.

Przyniósł mi to pewien student.

Zwróćmy uwagę na to, że mamy tu do czynienia z naruszeniem zasady grzeczności.

Maksyma sposobu głosi, że nadawca komunikatu powinien **mówić zrozumiale**. Ta maksyma zakłada, że nadawca komunikatu powinien z możliwości, jakie mu daje system języka, korzystać w sposób świadomy. Szczególnie dotyczy to takich zjawisk językowych jak synonimia i homonimia. Mówić zrozumiale, to używać takich wyrażen, które adresat komunikatu zna, elementem dyspozycji komunikacyjnej nadawcy musi być zatem wiedza o kompetencji językowej adresata, która pozwoli mu na przykład na wybór spośród synonimów takiego, co do którego może przyjąć sensowne założenie, że adresat zna jego znaczenie. Na przykład, mówiąc do osób niewykształconych, nie używa się raczej trudnych wyrazów obcego pochodzenia, jeśli istnieją polskie odpowiedniki, mówiąc do laików nie używa się terminów specjalistycznych, jeśli odpowiadają im wyrażenia należące do języka potocznego, a mówiąc do osób nieoczytanych, nie należy korzystać z relacji intertekstualnych. Na przykład, należy mówić raczej:

<i>stosunki dwustronne</i>	niż	<i>stosunki bilateralne</i>
<i>usunąć z uczelni</i>	niż	<i>relegować z uczelni</i>
<i>podpisać</i>	niż	<i>sygnować</i>
<i>szkarlatyna</i>	niż	<i>plonica</i>
<i>podrywacz</i>	niż	<i>Don Juan</i>

Nadawca powinien też zdawać sobie sprawę z homonimiczności pewnych wyrażen i struktur językowych, bo choć z jego punktu widzenia komunikaty zawierające takie wyrażenia nie są wieloznaczne – zawsze wie, jak jest i co chciał powiedzieć – to powinien takich wyrażen i struktur unikać, bo do ich ujednoznacznienia potrzebna jest odpowiednia informacja kontekstowa lub znajomość rzeczywistości, mogą więc zostać źle odczytane przez odbiorcę, a ich interpretacja zgodna z intencją nadawcy może w ogóle być niemożliwa lub może wymagać korygowania błędnej hipotezy co do znaczenia wyrażenia. Używanie struktur homonimicznych jest więc nieopłacalne i dla nadawcy – może nie zostać osiągnięty jego cel komunikacyjny lub osiągnięcie go może wymagać udzielenia dodatkowych informacji, i dla odbiorcy – dekodowanie wymaga większego wysiłku, a je-

żeli będzie niezgodne z intencją nadawcy, wysiłek ten może być w ogóle daremny, a w dodatku zamiast informacji użytecznej adresat może uzyskać informację wprowadzającą go w błąd.

Z homonimii leksykalnej częściej zdajemy sobie sprawę, homonimia struktur często uchodzi naszej uwadze, tym bardziej, że w języku polskim homonimia zdania bywa powodowana homonimią na przykład końcówek fleksyjnych tylko w wypadku jednego rodzaju gramatycznego, np.

*Samochód Pawła uszkodził samochód Piotra
Byt określa świadomość.*

Maksyma sposobu nakazuje również, aby mówić zwięźle i w sposób uporządkowany. **Mówić zwięźle** to znaczy nie wprowadzać szumu informacyjnego w postaci informacji nerelewantnych, należy jednak pamiętać o konieczności zachowania pewnej redundancji w przekazie, co wiąże się z wprowadzeniem niezbędnych powtórzeń informacji o największej relewancji. **Mówić w sposób uporządkowany**, to przede wszystkim panować nad strukturą tematyczno-rematyczną całego przekazu, a więc zmieniać temat wypowiedzi dopiero po wyczerpaniu wszystkich informacji wiążących się z nim, a relewantnych dla adresata. Należy też podawać informacje w optymalnym porządku – najczęściej jest nim porządek od najbardziej ogólnych do szczegółowych, co powinno znaleźć wyraz w odpowiednio dobranej hierarchicznej strukturze składniowej komunikatu.

Nadawca komunikatu w trosce o należyty odbiór zawartych w nim informacji powinien **sterować procesem odbierania komunikatu**, korzystając z odpowiednich wykładników metainformacyjnych i metatekstowych. W komunikacji bezpośredniej powinien w miarę możliwości kontrolować proces odbierania informacji, w miarę potrzeby uzupełniając treść lub zmieniając formę komunikatu. Wyrażenia metatekstowe ujawniają strukturę informacyjną komunikatu, explicite ukazując zależności między zawartymi w tekście informacjami, co dla odbiorcy ma ogromne znaczenie, pozwala mu bowiem na bardziej świadomy odbiór komunikatu i zwalnia go od samodzielnego strukturalizowania tekstu na podstawie zawartych w nim informacji cząstkowych, a strukturalizacja taka jest niezbędna do odpowiedniego włączenia informacji zawartych w tekście do zasobu wiedzy odbiorcy. Z kolei środki metainformacyjne zwracają uwagę adresata przede wszystkim na wartość poszczególnych informacji, wyrażają bowiem zakładany przez nadawcę stopień relewancji poszczególnych informacji dla adresata.

Maksyma jakości wymaga, aby nadawca komunikatu **mówił prawdę**, nie mówił tego, czego nie mógłby w należyty sposób uzasadnić, ani tego, o czym sądzi, że jest nieprawdziwe. Zalecenie to wymaga pewnego komentarza, nie zawsze bowiem jest tak, że nadawca przekazuje tylko informacje, o których sądzi, że są prawdziwe. Często jest tak, że przekazuje informacje, co do których prawdziwości ma uzasadnione wątpliwości, takie, co do których nie wie, czy są prawdziwe albo takie, które uzyskał od kogoś innego i za ich prawdziwość nie może wziąć odpowiedzialności. Wcale nie jest tak, że większą część naszej komunikacji wypełniają informacje, o których prawdziwości jesteśmy przekonani,

ważne jest jednak to, żeby informacja o tym, czy bierzemy odpowiedzialność za prawdziwość informacji, czy też nie, i jaka jest nasza ocena prawdziwości, była przekazana adresatowi komunikatu przez jego nadawcę. Jak wiemy służą do tego liczne środki metainformacyjne gramatyczne lub leksykalne.

Omawiając problemy związane z procesem nadawania komunikatu, skupiliśmy się na pewnej sytuacji modelowej (**komunikacja w czasie rzeczywistym we wspólnej przestrzeni komunikacyjnej**), jaką jest konwersacja w języku naturalnym między dwoma widzącymi się uczestnikami procesu komunikacyjnego, a więc sytuacji, gdzie na zmianę występują w roli nadawcy i odbiorcy komunikatu, gdzie nadawca, nadając komunikat do określonego adresata od razu otrzymuje od niego informację zwrotną o odbiorze komunikatu (w przekazie językowym lub innym, na przykład poprzez obserwację), może więc sterować procesem odbierania komunikatu i modyfikować w razie potrzeby własną strategię komunikacyjną. Inaczej przebiega proces nadawania komunikatu w sytuacji nieobecności adresata, przy czym możliwe są tu dwa wypadki, które omówimy poniżej.

Adresat może znajdować się w **innej przestrzeni informacyjnej** niż nadawca komunikatu, ale komunikacja odbywa się w czasie rzeczywistym i nadawca w czasie nadawania komunikatu utrzymuje z adresatem kontakt komunikacyjny. Tak jest na przykład w rozmowie telefonicznej lub w trakcie internetowego czatu. Taka komunikacja różni się od omówionej poprzednio tym, że komunikat językowy nie jest wspierany środkami pozajęzykowymi takimi jak gesty, czy mimika, wszystkie informacje relewantne muszą więc znaleźć w nim językowe wykładniki, co dotyczy na przykład posługiwania się zaimkami wskazującymi. Rola nadawcy jest tu trudniejsza, uboższa jest też informacja zwrotna od odbiorcy, mniejsza więc możliwość kontrolowania procesu odbierania komunikatu i reagowania ewentualnymi zmianami strategii informacyjnej. Mniejsze są też możliwości sprawdzania, czy kontakt między nadawcą i adresatem nie został przerwany, stąd więcej elementów fatycznych w komunikacie sprawdzających lub podtrzymujących kontakt. Sam komunikat musi zawierać więcej informacji wyrażanej językowo niż w wypadku, gdy uczestnicy procesu informacyjnego znajdują się w tej samej rzeczywistości, nie ma bowiem możliwości identyfikowania elementów rzeczywistości, których dotyczy informacja, za pomocą gestów wskazujących, jeśli elementy te są składnikami konsytuacji. Nadawca komunikatu musi się też liczyć z możliwością wystąpienia zakłóceń w kanale informacyjnym i ewentualnych strat informacji, musi więc starać się odpowiednio temu zapobiec poprzez wprowadzenie odpowiedniej redundancji do przekazu, na przykład mówić głośniej, wolniej, wyraźniej, ewentualnie powtarzać części komunikatu.

Inna jest sytuacja nadawcy komunikatu, gdy odbiór komunikatu jest przeniesiony w czasie, wówczas nadawca, konstruując komunikat, musi się oprzeć wyłącznie na wcześniejszej wiedzy o adresacie, nie ma bowiem możliwości jej uzupełniania, nie ma też możliwości wprowadzania żadnych korekt do przyjętej strategii komunikacyjnej. Rola nadawcy jest tu więc o wiele trudniejsza.

Omawiając problemy nadawania komunikatu, główną uwagę zwróciliśmy na subkod akustyczny języka naturalnego, większość z tych problemów jest

niezależna od kanału informacyjnego, pamiętajmy jednak, że przy **nadawaniu komunikatu w subkodzie wizualnym** występują pewne problemy specyficzne, choćby związane z tym, że język pisany rządzi się własnymi prawami dotyczącymi użycia pewnych konstrukcji. W wypadku, gdy teksty pisane mają postać tekstu utrwalonego, ich konstrukcja może się różnić także dlatego, że odbiorca takich komunikatów ma możliwość wielokrotnego przeglądania dowolnych partii tekstu, nie musi więc informacji pochodzących z wcześniejszych partii tekstu, a niezbędnych do dekodowania aktualnej sekwencji komunikatu, utrzymywać w pamięci, zawsze bowiem ma możliwość sprawdzenia jej i dokonania weryfikacji aktualnego odbioru komunikatu. Ma to wpływ przede wszystkim na stopień redundancji informacji zawartej w komunikacie. O innych problemach związanych z tekstami pisаныmi powiemy, omawiając proces utrwalania informacji.

Pzekazywanie komunikatu

Komunikat, jeśli ma być odebrany przez adresata, musi mu zostać przesłany, a więc musi odbyć pewną drogę w **kanale informacyjnym**. Przesyłanie komunikatu to proces znajdujący się w kompetencji nadawcy, który, nadając sygnał, musi uwzględnić właściwości kanału informacyjnego oraz to, że w kanale informacyjnym zawsze znajdują się inne sygnały, zakłócające sygnał komunikatu i powodujące **szum informacyjny**. Nadawca musi więc, żeby zminimalizować straty informacji, zapewnić sygnałowi komunikatu odpowiednią **redundancję**.

Często bywa i tak, że sygnał w swojej drodze od nadawcy do odbiorcy wielokrotnie zmienia kanał informacyjny, ponieważ warunki fizyczne jednego kanału nie zapewniają transportu komunikatu w tym sensie, że materiał tworzący sygnał albo ma zbyt małą moc przesyłową (zbyt dużą bezwładność lub zbyt małą pojemność), albo też istniejący w kanale szum informacyjny powodowałby zbyt dużą deformację sygnału, taką, której nie mogłaby zapobiec redundancja. Tak się dzieje na przykład przy komunikacji ustnej, która nie może odbywać się na zbyt dużą odległość, ponieważ wraz z odległością sygnał informacyjny stopniowo słabnie, aż w końcu zanika. Dlatego używa się różnego typu instrumentów informacyjnych umożliwiających przesyłanie sygnału – proces przesyłania komunikatu dotyczy zawsze tylko jego strony materialnej, a więc sygnału, który jest nośnikiem informacji. Czasami wystarczy tylko wzmocnienie sygnału – tak się dzieje na przykład przy użyciu tuby głosowej, zazwyczaj jednak to nie wystarcza i wówczas do przesyłania komunikatu konieczna jest zmiana rodzaju sygnału, a więc i kanału informacyjnego. Wówczas trzeba już dokonać transformacji komunikatu, przekodować jego sygnał pierwotny (właściwy) na inny sygnał, w innym materiale, w którym będzie mógł być przetransportowany. Czasami operacja przekodowywania sygnału połączona jest z procesem utrwalania nowego sygnału, co umożliwia przechowanie komunikatu i przesunięcie w czasie procesu jego odbioru względem nadawania. Problemy transportowania komunikatu jako procesu składającego się na przesyłanie komunikatu są często problemami technicznymi, nie stanowią więc przedmiotu naszych rozważań, zwróćmy uwagę jednak na fakt, że przy każdej takiej zmianie sygnału (i kanału informacyjnego) należy się liczyć

ze stratami informacji, urządzenia dokonujące kolejnych transformacji należy więc wyposażyć w odpowiednie mechanizmy zapewniające redundancję (tak jak np. w standardzie zapisu CD-ROM). Końcowym etapem kodowania sygnałów w procesie przesyłania komunikatu musi być doprowadzenie sygnału do takiej postaci, którą może odebrać adresat komunikatu i która zapewni możliwość odkodowania informacji bez strat informacji relewantnych.

Odbieranie komunikatu

Odbieranie komunikatu nie jest po prostu odwrotnością procesu nadawania komunikatu. Odbieraniem komunikatu rządzi zasada, że ilość informacji odebranej jest zawsze mniejsza od ilości informacji nadanej. Dzieje się tak z dwóch powodów. Po pierwsze, komunikat w trakcie przesyłania ulega pewnym deformacjom, po drugie, to, ile informacji odbierze odbiorca z komunikatu, zależy w dużej mierze od kompetencji odbiorcy, i to nie tylko od jego kompetencji językowej. Zależy również od tego, czy odbiorca komunikatu był jego adresatem, czy też nie – w pierwszym wypadku komunikat powinien być do potrzeb i możliwości odbiorczych jego odbiorcy odpowiednio przez nadawcę dostosowany (o czym już mówiliśmy), w drugim wypadku tak się nie dzieje, co może znacznie utrudniać i komplikować proces odbioru, czyniąc go o wiele bardziej pracochłonnym.

Składowe procesy odbierania komunikatu omówimy poniżej, należy jednak pamiętać, że nie następują one po sobie sekwencyjnie, ale nakładają się na siebie w czasie, a większość z nich przebiega równocześnie.

Warunki, jakie musi spełniać **odbiorca komunikatu**, wymieniliśmy poprzednio, kiedy zastanawialiśmy się nad tym, jaką wiedzą o adresacie musi dysponować nadawca komunikatu, aby komunikacja mogła przebiegać pomyślnie. Przypomnijmy, że najważniejszym jest możliwość odebrania sygnału komunikatu oraz kompetencja w zakresie języka komunikatu, a więc znajomość systemu językowego oraz reguł jego użycia w konkretnych sytuacjach komunikacyjnych.

Niezbędnym **warunkiem odebrania komunikatu** i zawartej w nim informacji jest **odebranie i identyfikacja sygnału informacyjnego**. To, co odbiorca musi robić w trakcie odbierania całego komunikatu, to starać się **zrekonstruować sygnał**, który do niego dotarł, do postaci kanonicznej, a więc postaci systemowej, zawierającej wszystkie jego cechy systemowo relewantne. Umożliwia to zawarta w komunikacie redundancja oraz kompetencja językowa odbiorcy, przede wszystkim w zakresie podsystemu fonologicznego, a więc znajomość systemowych odpowiedników głosek (fonemów) i reguł tworzenia z nich wyrażeń złożonych, szczególnie zaś reguł określających prawdopodobieństwo wystąpienia danej jednostki w danym ciągu (kontekście). Jednostki elementarne języka naturalnego układają się bowiem w ciągi będące tzw. **ciągami Markowa**, w których obowiązuje zasada, że im więcej znamy elementów poprzedzających, z tym większym prawdopodobieństwem możemy określić, jaki element pojawi się jako następny element ciągu. W języku naturalnym obliczenia te ułatwia także znajomość morfemów oraz reguł ich składania w większe całości, na przykład obowiązujący w składni polskiej związek zgody. Okazuje się, że potrafimy w procesie odbioru

komunikatu przewidzieć, często z dużym prawdopodobieństwem, co będzie dalej, a odebrany sygnał służy często tylko do weryfikowania hipotezy wstępnej. Pomocna jest tu także znajomość **reguł leksotaktyki**, czyli reguł określających współwystępowanie pewnych wyrazów, na przykład znajomość związków frazeologicznych stałych (na przykład, jeśli w tekście polskim występuje czasownik *broczyć*, to następnym wyrazem będzie *krew*).

Warunkiem poprawnego odebrania komunikatu, a więc odkodowania zawartej w nim informacji nie jest doprowadzenie sygnału do postaci kanonicznej (systemowej), a tylko do postaci uboższej od postaci systemowej o elementy redundantne, bo te, jak wiemy, stanowią nadmiar zapobiegający stratom informacji.

Na odbiór sygnału informacyjnego składa się również **identyfikacja elementów suprasegmentalnych** języka, takich jak intonacja, która nie tylko pozwala na delimitację sygnału, ale jest często wykładnikiem funkcji semantycznej, na przykład pytań i oznajmień, funkcji impresywnej, na przykład zakazów, prośb, czyli siły illokucyjnej wypowiedzenia, a zawsze jest wykładnikiem funkcji ekspresywnej komunikatu, niosąc dodatkowe informacje o jego nadawcy.

Do odkodowania zawartej w komunikacie informacji niezbędne jest dokonanie **delimitacji sygnału**, bo tylko w stosunku do sygnału podzielonego można zastosować reguły semantyki.

Akustyczne cechy sygnału komunikatu, takie jak jego dynamika i tempo mowy, są także wykładnikami innych informacji ważnych dla odbiorcy: mogą być wykładnikami struktury tematyczno-rematycznej, co jest ważne nie tylko przy odbieraniu informacji zakodowanej w komunikacie, ale również przy dokonywaniu pewnych operacji przetwarzania tej informacji. Mogą być wykładnikami metainformacyjnymi, za pomocą których nadawca celowo steruje procesem odbierania komunikatu, w tym wykładnikami zakładanej przez nadawcę relewancji informacji dla jej adresata. Dlatego też identyfikacja tych wykładników przez odbiorcę jest dla niego bardzo ważna, znacznie bowiem może oszczędzać wysiłek, jaki musi włożyć w proces odbierania komunikatu.

W normalnych warunkach odbierania komunikatu mówionego, a więc gdy proces odbierania przebiega prawie równocześnie z procesem nadawania (przesunięcie w czasie ze względu na czas potrzebny do przebycia przez sygnał drogi od nadawcy do odbiorcy możemy uznać tu za nieistotne), jeśli odbiorca spostrzeży, że sygnał dociera do niego uszkodzony lub zbyt słaby na to, żeby możliwe było odtworzenie jego postaci kanonicznej, może to zakomunikować nadawcy i poprosić o wprowadzenie odpowiednich korekt przy nadawaniu i powtórne nadanie komunikatu. Służą do tego na przykład takie zwroty jak:

Slucham?, Proszę?, Co mówisz?, Co?, Nie rozumiem!, Nie słyszę!, Powtórz, co powiedziałeś.

Identyfikacja i delimitacja sygnału to dopiero warunki wstępne odbioru komunikatu. Żeby odzyskać z komunikatu zakodowaną w nim informację, odbiorca musi, korzystając z kompetencji semantycznej, **przyporządkować sygnałowi znaczenie**. Do porozumienia się nadawcy z odbiorcą wystarczy, że tym samym

wykładnikom przyporządkują oni te same elementy rzeczywistości, a więc **tożsamość denotacji**, bo konotacja przypisywana wyrażeniom, jak wiemy, zależy nie tylko od kompetencji językowej, ale także od wiedzy o rzeczywistości. Dlatego też może się zdarzyć, że przyrost informacji u odbiorcy po odebraniu komunikatu będzie mniejszy lub większy, niż by to wynikało z intencji komunikacyjnej nadawcy.

Odkodowywanie informacji zawartej w tekście jest procesem złożonym, wymagającym od odbiorcy ciągłej antycypacji, a więc stawiania z wyprzedzeniem pewnych **hipotez semantycznych** na podstawie odkodowanej wcześniej informacji, a następnie sprawdzania ich w miarę napływu dalszych części komunikatu, potwierdzania ich lub odrzucania, a w tym ostatnim wypadku wprowadzania odpowiednich korekt. Wymaga to korzystania z **informacji kontekstowej** i utrzymywania w pamięci operacyjnej odpowiednich informacji. Tak się dzieje w wypadku tekstu wieloznacznego, którego **ujednoznacznienie** często możliwe jest dopiero po uzyskaniu dalszych informacji w komunikacie lub po odwołaniu się do **wiedzy o rzeczywistości**. Tak jest na przykład w wypadku dwuznacznego zdania *Byt określa świadomość* (w niemieckim oryginale zdanie to nie jest dwuznaczne). Zwróćmy uwagę na to, że w wypadku wieloznaczności mamy do czynienia z brakiem symetrii między sytuacją odbiorcy a sytuacją nadawcy: dla nadawcy komunikat jest zawsze jednoznaczny, nawet jeśli występują w nim wyrażenia lub struktury homonimiczne, bo nadawca wie, co chciał powiedzieć (chyba że specjalnie sformułował komunikat wieloznaczny) – komunikat może być wieloznaczny tylko dla odbiorcy.

Odkodowywanie informacji przez odbiorcę wymaga poprawnej, takiej samej jak u nadawcy, identyfikacji elementów rzeczywistości, których dotyczy informacja, dlatego też odbiorca musi znaleźć **odniesienie** wszystkich elementów wskazywanych przez wyrażenia okazjonalne, w tym zaimków występujących w każdym tekście spójnym, odnoszących się albo bezpośrednio do komunikowanej sytuacji, albo do denotacji wyrażen współwystępujących w tym samym tekście. Wymaga to więc od odbiorcy zgodnej z intencją nadawcy identyfikacji relacji zachodzących między elementami tekstu a sytuacją informacyjną, w tym konsytuacją komunikatu, oraz zgodnej z regułami budowy tekstu spójnego identyfikacji argumentów relacji wewnątrztekstowych.

Ilość informacji odebranej z komunikatu zależy również od odbiorcy w wypadku występujących w komunikacie odniesień do innych tekstów – żeby tekst zrozumieć odbiorca musi zidentyfikować odpowiednie **relacje intertekstualne**. Taka wiedza potrzebna jest również do interpretacji znaczenia wielu przenośni i frazeologizmów.

Wiedza o rzeczywistości pozatekstowej, w tym wiedza o nadawcy komunikatu, niezbędna jest odbiorcy komunikatu także do **uspójnienia odbieranego komunikatu** i uzupełnienia zawartych w nim informacji informacjami niezbędnymi do przetworzenia tych informacji i odpowiedniego ich ustrukturalizowania, niezbędnego do włączenia ich do własnego zasobu wiedzy. Oczywiście można tu postawić pytanie, czy jest to jeszcze odbieranie komunikatu, czy też jest to już inny

proces, ale faktem jest, że bez dokonania takich operacji wiele tekstów byłoby po prostu sekwencją niepowiązanych ze sobą informacji. Na przykład, tylko **wiedza o rzeczywistości** pozwala uzupełnić niezbędnymi informacjami i zrozumieć poniższe dialogi:

– *Może pójdziemy do kina?*

– *Nie mam pieniędzy.*

– *Ale się opaliłaś!*

– *Byłam na Mazurach.*

– *Strasznie chce mi się pić.*

– *Już wszystko zamknięte.*

– *Miło spędziłaś niedzielę?*

– *Byłam u Kowalskich.*

– *Zalutuję to jutro.*

– *Jutro niedziela.*

Żeby zrozumieć komunikat odbiorca musi również znać zasady użycia pewnego rodzaju komunikatów, a więc **reguły pragmatyki języka**. Taka znajomość jest niezbędna do interpretacji siły illokucyjnej tzw. **aktów mowy nie wprost (pośrednich aktów mowy)** – tym łatwiej o interpretację niezgodną z intencją nadawcy, im mniejszy stopień skonwencjonalizowania takich aktów mowy. Na przykład, wypowiedzenia w formie pytania skierowanego do adresata z odnoszącym się do niego predykatem *móc, zechcieć*, często w formie zaprzeczonej lub w trybie przypuszczającym, nie są w intencji mówiącego pytaniami, ale prośbami, mają więc spowodować odpowiednie niejęzykowe zachowanie adresata., np.:

Czy możesz zamknąć okno?

Czy mógłbyś zamknąć okno?

Czy nie mógłbyś zamknąć okna?

Czy nie zechciałbyś jednak spróbować?

Pytania takie są całkowicie skonwencjonalizowane i zachowanie adresata normalne w wypadku innych pytań (o rozstrzygnięcie), a więc odpowiedź „tak” bez towarzyszącego jej działania lub usprawiedliwienia się w wypadku odpowiedzi negatywnej są sytuacyjnie niestosowne.

Innego udziału adresata w odbieraniu, a właściwie w przetwarzaniu informacji zawartej w komunikacie, wymagają **akty mowy pośrednie o funkcji illokucyjnej**, które mają mniejszy stopień konwencjonalizacji, tu bowiem informacja uzupełniająca musi być przez odbiorcę uzyskana z **konsytuacji**. Na przykład, jeśli ktoś mówi *Zimno mi*, to komunikat taki powinien być zinterpretowany jako prośba o odpowiednie działanie adresata, który powinien, zależnie od sytuacji, w której komunikat taki został nadany, na przykład zamknąć okno, dać coś ciepłego

do picia lub do okrycia, choć możliwe jest też zachowanie językowe, np. *Już niedaleko.*

Znajomość reguł pragmatyki to również umiejętność **rozpoznania rejestru językowego**, stylu wypowiedzi oraz znajomość pewnych konwencjonalnych **scenariuszy konwersacyjnych**, niezbędna odbiorcy do tego, aby, włączając się do procesu przekazywania informacji, mógł zachować się językowo właściwie. Niewłaściwa będzie reakcja kolokwialna na komunikat w stylu oficjalnym, urzędowym lub podniosłym (a także odwrotnie), np.:

- *Nie wiem, czy mógłbym zwrócić się do pana z pewną prośbą.*
- *Spoko. Nawijaj!*

Niewłaściwe też będzie szczegółowe sprawozdanie w odpowiedzi na grzecznościowe *Co słyhać?*

Omawiając proces odbierania komunikatów języka naturalnego, uwagę zwróciliśmy bądź na te operacje, które dotyczą głównego subkodu tego języka, a więc języka mówionego, bądź na te, które nie zależą od tego, w jakim subkodzie komunikat został nadany. Na problemy związane z odbieraniem komunikatu pisanego zwrócimy uwagę przy omawianiu procesu utrwalania informacji.

Przetwarzanie komunikatu

Jak widzieliśmy, odbieranie komunikatu połączone jest z równoczesnym **przetwarzaniem informacji**. Zrozumieć treść komunikatu to, zgodnie z teorią znaczenia Alfreda Tarskiego, umieć wyprowadzić z niego wszystkie **konsekwencje**. Językoznawcy od konsekwencji w szerokim sensie (takim jak u Tarskiego) odróżniają tzw. **presupozycje**, czyli takie informacje, które muszą być prawdziwe, jeśli prawdziwy ma być cały komunikat. Są to więc takie informacje, o których prawdziwości w danym komunikacie się nie orzeka (jeśli zakładamy, że cały komunikat ma być prawdziwy), tylko się ją zakłada. Z takim założeniem mamy na przykład do czynienia w znanej anegdocie:

Rozprawa sądowa. Sędzia zadaje pytania, świadek odpowiada rozwlekle i niejasno. Sędzia, zirytowany, mówi: Proszę na pytania Sądu odpowiadać tylko „tak” albo „nie”. Świadek na to: Kiedy tak się nie da, proszę wysokiego sądu. Bo jeśli ja zapytam, czy Wysoki Sąd przestał brać łapówki?

Takie pytanie zakłada bowiem, że Wysoki Sąd bierze łapówki, choć tego nie orzeka, i w takiej sytuacji żadna z odpowiedzi : „tak” lub „nie” presupozycji tej nie uchyla.

W *Słowniku encyklopedycznym informacji, języków i systemów informacyjno-wyszukiwawczych* presupozycję zdefiniowaliśmy następująco:

„PRESUPOZYCJA – zdanie a (informacja a) jest presupozycją zdania b, jeżeli prawdziwość zdania a (informacji a) jest koniecznym warunkiem prawdziwości zdania b (informacji b), np. presupozycją zdania *Jan wyjechał z Warszawy* jest zdanie *Jan był w Warszawie*, presupozycją zdania *Jan przestał palić papierosy* jest zdanie *Jan palił papierosy*, presupozycją zdania *Anna odnalazła swoje notatki*

są zdania *Anna zgubiła swoje notatki* i *Anna szukała swoich notatek*, zaś zdania *Anna odnalazła wreszcie swoje notatki* ponadto zdanie *Anna długo szukała swoich notatek*. Z przykładów tych widać, że można mówić nie tylko o presupozycji zdań, ale również o presupozycjach wyrazów, gdyż one to najczęściej są jej wykładnikami. Funkcja presupozycji zdaniu lub zbiorowi zdań przyporządkowuje zbiór wszystkich jego presupozycji. Znajduje to praktyczne zastosowanie w tzw. systemach inteligentnych oraz w systemach informacji faktograficznej.”

Termin konsekwencja zazwyczaj rozumiany jest szerzej:

„KONSEKWENCJA – zdanie a (informacja) jest konsekwencją zdania b, jeżeli zdanie (informację) a można uzyskać w procesie transformacji wnioskowania ze zdania (informacji) b. Tak więc funkcja konsekwencji zdaniu (lub zbiorowi) zdań przyporządkowuje zbiór wszystkich jego konsekwencji, gwarantując ich prawdziwość. Np. zdania *Jan jest bratem Piotra*, *Piotr jest bratem Jana* są konsekwencją zdania *Jan i Piotr są braćmi*; zdanie *Jana nie ma w Warszawie* jest konsekwencją zdania *Jan wyjechał z Warszawy*; zdanie *Jan nie pali już papierosów* jest konsekwencją zdania *Jan przestał palić papierosy*. Funkcja konsekwencji znajduje praktyczne zastosowanie w tzw. systemach inteligentnych oraz w systemach informacyjno-wyszukiwawczych, przy czym wykorzystuje się w nich nie tyle relację konsekwencji logicznej, co relację konsekwencji semantycznej. W językoznawstwie relacja konsekwencji odróżniana jest od relacji presupozycji, czasami jednak termin rozumiany jest szerzej, obejmując także wypadki presupozycji.”

Takie przetwarzanie informacji jest operacją **wnioskowania**, którego wszystkie przesłanki zawarte są w tekście, jest więc **operacją wewnątrztekstową**, uwzględniającą tylko zasady semantyki języka komunikatu. Widzieliśmy już, że w procesie odbierania komunikatu bardzo często mamy do czynienia z przetwarzaniem treści komunikatu także w oparciu o przesłanki pochodzące spoza jego tekstu, z wiedzy odbiorcy o rzeczywistości, w tym o sytuacji komunikacyjnej, i o regułach użycia języka (pragmatyki). Tak było w wypadku aktów mowy pośrednich, tak jest w wypadku **uspójniania tekstu** przez odbiorcę – informacje do tego potrzebne czerpie on często spoza tekstu. Na przykład, żeby zrozumieć dialog:

– *Jak było wczoraj u Kowalskich? Dobrze się bawiłeś?*

– *Byli Nowakowie.*

niezbędna jest wiedza o tym, jacy są Nowakowie, a taka odpowiedź może w rzeczywistości znaczyć albo „świetnie”, albo „okropnie”.

W normalnej komunikacji częściej mamy do czynienia z takim właśnie przetwarzaniem informacji zawartej w komunikacie i często nawet nie zdajemy sobie sprawy z tego, jak wiele przesłanek potrzebnych do wniosków, jakie wyciągamy na podstawie tekstu, pochodzi od nas samych, dlatego niejednokrotnie skłonni jesteśmy przyjąć, że ktoś nam powiedział coś, czego wcale nie mówił.

Przetwarzanie, czyli transformowanie komunikatu (tekstu) to również inne operacje dokonywane na tekście. Omówimy je pokrótce.

Transformacja to taka operacja, która dokonuje przekształcenia jednej struktury w drugą, z zachowaniem relewantnych cech przekształcanej struktury. Przy-

porządkowanie elementów struktury przekształcającej (będącej na wejściu procesu informacyjnego) w strukturę będącą wynikiem transformacji (będącą na wyjściu procesu transformacji) zazwyczaj odbywa się zgodnie z przyjętymi regułami odpowiedniości i często ma postać algorytmu.

Każdy komunikat to materialny nośnik (sygnał przekazywany w kanale informacyjnym) i zakodowana w nim informacja, czyli treść. Transformacja może dotyczyć obu tych elementów.

Transformacją sygnału komunikatu językowego jest na przykład zamiana sygnału akustycznego na impuls elektryczny, a tego z kolei na falę radiową, i odwrotnie, jak to jest przy nadawaniu audycji radiowych. Uznawane za tego rodzaju transformację przekształcanie tekstu mówionego języka naturalnego w tekst pisany, lub odwrotnie, jest w rzeczywistości transformacją dokonywaną, w mniejszym lub większym stopniu, także na elementach treści komunikatu, co zależy przede wszystkim od rodzaju pisma. Za transformacje dokonywane na sygnale komunikatu uznaje się także **transliterację** i **transkrypcję**, w rzeczywistości jednak tylko transliteracja jest taką operacją, transkrypcja często wymaga także operacji na treści komunikatu.

Równie złożony charakter ma transformacja dokonywana na języku komunikatu, na przykład **tłumaczenie** (przekład). Jest to **operacja międzyjęzykowa**, mająca na celu zmianę formy znaków języka naturalnego, ale w językach naturalnych nie ma jedno-jednoznaczności przyporządkowania między wyrażeniami należącymi do różnych języków, dlatego człowiek dokonujący takiej transformacji musi odwoływać się także do ich treści. Tłumaczenie maszynowe (komputerowe) jest operacją dokonywaną na formie znaków językowych, wymaga jednak wielu skomplikowanych reguł kontekstowych. Tłumaczenie jest transformacją złożoną, wymaga odkodowania komunikatu, a więc przyporządkowania jego formie informacji i zakodowania tych informacji w innym języku, przy czym relewantna jest tu nie tylko treść komunikatu transformowanego (chodzi o jak najmniejsze straty informacji), ale także wykładniki funkcji komunikatu – komunikat powstały w wyniku transformacji tłumaczenia, a więc przekład, powinien zawierać wyrażenia pełniące taką samą funkcję w języku przekładu jak wyrażenia użyte w tekście oryginalnym (np. wyrażenia ekspresywne, gwarowe itp.). Dokładniejsze omówienie transformacji przekładu wykracza poza ramy naszych rozważań.

Tekst komunikatu poddawać można wielu różnym **transformacjom wewnątrzjęzykowym**, zmieniającym jego formę, ale zachowującym treść. Wymienić tu można:

- zamianę zdań w stronie czynnej na zdania w stronie biernej lub odwrotnie,
- zamianę wyrażen na wyrażenia synonimiczne,
- podział zdań złożonych na zdania składowe,
- łączenie zdań pojedynczych w zdania złożone,
- zmianę szyku wyrażen,
- zmianę językowego wykładnika danego elementu rzeczywistości na inny wykładnik zachowujący tożsamość denotacji (w tym na synonim, ale także, w tekście spójnym, na zaimek lub wyrażenie o szerszym zakresie),

- zmianę wyrażenia nacechowanego stylistycznie na nienacechowane,
- zmianę mowy niezależnej na mowę zależną.

Takich operacji na tekście dokonujemy na przykład w trakcie tzw. **redakcji tekstu**.

Transformacji dokonywać można również na informacjach zawartych w komunikacie. Wymienimy najważniejsze, pamiętając, że często są to transformacje złożone, chociażby dlatego, że transformacja treści najczęściej wymaga także transformacji formy komunikatu (a więc i sygnału).

- **Kondensacja informacji** – polega na usunięciu z tekstu wszystkich informacji redundantnych, poza tymi elementami redundantnymi, które są realizacją w tekście redundancji systemowej języka naturalnego;

- **Streszczenie** – polega na kondensacji informacji i równoczesnym **selekcjonowaniu informacji**, a więc pozostawieniu tych informacji, które są relewantne ze względu na temat komunikatu oraz założony stopień relewancji. Streszczenie zakłada przekształcenie treści komunikatu w strukturę hierarchiczną, porządkującą informacje w drzewie hierarchicznym – od najważniejszych ze względu na scenariusz tekstu do najmniej ważnych – a następnie na wybieraniu z niego informacji do założonego stopnia hierarchii i przyporządkowania im odpowiedniej formy z planu wyrażania języka naturalnego, niekoniecznie tego samego, w którym jest sformułowany tekst streszczany (streszczać można w innym języku, przy czym najczęściej operacja ta nie rozkłada się na streszczenie w języku tekstu wejściowego i tłumaczenie uzyskanego w wyniku tej pierwszej transformacji streszczenia). Tekst uzyskany w operacji streszczania musi być istotnie krótszy niż tekst wejściowy, przy czym długość tekstu streszczenia może być z góry założona. To długość streszczenia wyznacza poziom hierarchii informacji tekstu streszczanego odwzorowywanych w streszczeniu. Poziom hierarchii informacji podporządkowanych poszczególnym tematom tekstu streszczanego powinien być w zasadzie dla wszystkich tematów taki sam, streszczenie powinno być bowiem proporcjonalne. Czasami jednak o informacjach umieszczanych w streszczeniu decydować może stopień ich **relewancji dla adresata** (użytkownika) streszczenia, jak to się dzieje w wypadku tzw. **streszczeń dokumentacyjnych** przygotowywanych dla potrzeb użytkowników określonych systemów informacyjno-wyszukiwawczych. Taka operacja streszczania powiązana jest więc z operacją selekcjonowania informacji, o czym powiemy niżej. Tekst powstały w wyniku operacji streszczania zazwyczaj poprzedzany jest informacją dotyczącą tekstu streszczanego, przy czym cały tekst streszczenia może (ale nie musi) mieć postać metainformacji. A oto przykłady takich streszczeń:

Klan,

„Wtorek: Kinga złości się na Michała, że rodzina go nie interesuje i że myśli wyłącznie o karierze. Beatę dziwi związek Sebastiana z racjonalistką Barbarą. Agnieszka ma imieniny, pamiętali o tym także Tomek i Dominika. Beata i Kinga wpadają z dziećmi na Sadybę, żeby złożyć życzenia z okazji Dnia Babci i Dziadka. Zosia opowiada Ryśkowi o zimowisku na Mazurach,

na które Grażynka mogłaby jechać jako opiekunka. W domu czeka na niego list od Danusi i Zenka, w którym zapraszają całą rodzinę na ferie.”

(Przy okazji tego przykładu zwróćmy uwagę na to, że tekst streszczany jest realizacją zupełnie innego kodu niż streszczenie – film jest streszczany w języku naturalnym).

Mao Hayder

PTASZNIK

„Nad Tamizą w Greenwich zostaje odnalezionych pięć zamordowanych młodych kobiet. Podczas autopsji wychodzi na jaw, że każde z ciał nosi przerażający ślad – znak rozpoznawczy sprawcy. Młody, zdolny inspektor Jack Caffery podejmuje indywidualne śledztwo – to jego pierwsza sprawa w szeregach specjalnej grupy dochodzeniowej. Identyczne ślady na zwłokach skłaniają detektywa do wniosku, że sprawca upodabnia ofiary do jakiegoś własnego ideału. Idąc tym tropem namierza pewnego bogatego nekrofila. Ale wtedy...psychopata dopada kolejną ofiarę. Dla niewzruszonego dotąd inspektora zaczyna się teraz osobisty wyścig z czasem.”

„«Systemy informacyjne w zarządzaniu» – to podręcznik, w którym dokonano kompleksowego przeglądu uwarunkowań, w jakich działa informacyjny system zarządzania, spełniając stawiane przed nim w praktyce zadania. Opisano zależności tworzące się w momencie włączenia technologii komputerowej do posługiwania się informacjami oraz uwarunkowania wynikające z możliwości finansowych, organizacyjnych i kadrowych przedsiębiorstw czy urzędów.”

Jacek Otto

Marketing relacji. Koncepcja i stosowanie

„Podręcznik zawiera liczne przykłady zastosowań marketingu relacji przez firmy zachodnie oraz wskazuje nowe obszary analiz rynkowych. Obejmuje całokształt problematyki marketingu relacji (relationship marketing), bazując na literaturze światowej z dziedziny marketingu. Ukazuje związki i zależności, jakie zachodzą między dotychczasowym dorobkiem marketingu i dziedzin pokrewnych a teorią marketingu relacji. Prezentuje modele marketingu relacji, wyniki badań przeprowadzonych przez znaczące ośrodki nauko-badawcze.”

Dwa ostatnie teksty to przykłady streszczeń dokumentacyjnych. Takie teksty nie są właściwymi streszczeniami, a główną ich funkcją, w przeciwieństwie do streszczeń właściwych, jest funkcja metainformacyjna, dlatego też nie są skrótami tekstów streszczanych, ale metainformacjami wymieniającymi ich najważniejsze tematy. Taką samą funkcję pełnią teksty, również nazywane streszczeniami, które nie odnoszą się do tekstów już istniejących, nie są więc transformacjami innych tekstów, a zadaniem ich jest informowanie o treści tekstów, które dopiero mają powstać.

- **Selekcjonowanie informacji** – polega na wybieraniu z tekstu (zbioru informacyjnego) tych informacji, które spełniają **kryteria relewancji** określone ze względu na potrzeby informacyjne użytkownika, przy czym wcale nie muszą to być informacje uznane przez nadawcę komunikatu za najbardziej relewantne dla jego adresata. Wybierając te ostatnie należy pamiętać o ich charakterystycznym rozmieszczeniu w strukturze tekstu spójnego, tak bowiem jest, że w dłuższym tekście określone partie są bardziej nasycone informacją uznaną przez nadawcę komunikatu za relewantną dla adresata niż inne. Na przykład w artykułach naukowych takimi miejscami są początek tekstu, gdzie określa się temat komunikatu, i koniec, gdzie zawarte są wnioski, a więc informacje nowe, często najbardziej wartościowe dla adresata.

- **Wnioskowanie** – jest to **transformacja prawdziwościowa** informacji, a więc taka, że z informacji uznanych za prawdziwe (**przesłanek**) uzyskuje się inne informacje (**wnioski**), przy czym informacjom tym przysługuje taki sam stopień prawdziwości, jaki miały przesłanki – jeżeli mamy do czynienia z **wnioskowaniem logicznym**, lub też mniejszy – jeśli jest to **wnioskowanie uprawdopodobniające**. To ostatnie zachodzi wtedy, gdy przesłanki mają niedookreśloną prawdziwość, albo gdy nie dysponujemy wszystkimi przesłankami potrzebnymi do wyciągnięcia danego wniosku. Wnioskowanie logiczne, choć pewne, nie daje przyrostu informacji – nowe informacje, choć o niepewnej wartości, możemy uzyskać tylko przy wnioskowaniu uprawdopodobniającym, na przykład przy **wnioskowaniu statystycznym** lub **wnioskowaniu przez analogię**. Istotna ze względu na problemy przetwarzania komunikatu różnica polega na tym, że w pierwszym wypadku wszystkie przesłanki potrzebne do wyciągnięcia wniosku znajdują się w danym komunikacie – są elementami jego struktury znaczeniowej i stanowią zbiór jego **konsekwencji**, natomiast w drugim wypadku są uzupełniane przesłankami spoza komunikatu, Taka operacja nie jest więc transformacją danego komunikatu, ale operacją przeprowadzaną na wielu komunikatach.

Utrwalanie informacji

Utrwalanie komunikatu często jest obecnie procesem technicznym, polegającym na wielokrotnym przetwarzaniu sygnału, odwzorowywaniu jego cech dystynktywnych w innym materiale lub przekodowywaniem na inny język, zależnie od możliwości materiału, urządzeń kodujących i dekodujących, a także właściwości kanału informacyjnego. Tymi procesami nie będziemy się zajmować, zwróćmy jednak uwagę na to, że zależnie od przyjętej technologii innym transformacjom poddaje się na przykład komunikat dźwiękowy, utrwalając go na płycie analogowej, innym na płycie CD lub DVD – tu konieczne jest zakodowanie sygnału w kodzie dwójkowym, jest to więc operacja przekodowywania, a jeszcze innym przy utrwalaniu w systemie MP3, gdzie konieczne jest także selekcjonowanie informacji.

Nowe techniki utrwalania komunikatu, w tym komunikatów języka naturalnego, to możliwości dopiero XX wieku – wcześniej do utrwalania komunikatów języków naturalnych służyło **pismo**. Temu procesowi przyjrzymy się nieco dokładniej, zwracając uwagę przede wszystkim na różnice zależne od rodzaju pisma.

Zwróćmy uwagę jeszcze na to, że kiedy mówimy o procesie **utrwalania komunikatów języka naturalnego**, możemy mieć na myśli dwa różne procesy informacyjne. Pierwszy proces to proces **nadawania komunikatu w subkodzie pisanym języka naturalnego** – tu wykonawcą tego procesu jest nadawca komunikatu, przy czym jest to w zasadzie proces nadawania informacji w subkodzie trwałym. Drugi, to proces bardziej złożony, proces **utrwalania komunikatu przez jego odbiorcę** – z sytuacją taką mamy do czynienia podczas pisania pod czyjeś dyktando. I choć w obu wypadkach wynikiem procesu jest ten sam rezultat – utwalony sygnał wraz z materiałem, w którym został utwalony, a więc dokument (nadawanie komunikatu w subkodzie pisanym nietrwałym nie jest utrwalaniem informacji, np. w wypadku gazety świetlnej), to operacje składowe tych procesów są różne.

Czasami, mówiąc o utrwalaniu informacji, mamy na myśli jeszcze coś innego, a mianowicie przyporządkowywanie w procesie informacyjnym – najczęściej w procesie nadawania komunikatu – informacji (treści), rozumianej abstrakcyjnie, jej formy wyrażenia (jednostki planu wyrażania danego systemu znakowego). Tak rozumianym procesem utrwalania informacji nie będziemy się tu zajmować, przyjęliśmy bowiem, zgodnie z koncepcją Ferdinanda de Saussure'a, bilateralną koncepcję znaku jako całości składającej się z formy i treści.

Utrwalaniem informacji w sensie technicznym, a więc procesem tworzenia trwałej formy znaku w odpowiednim materiale, nie będziemy się zajmować, nie jest to bowiem proces informacyjny sensu stricto. To, co nas będzie interesowało, to proces zapisywania informacji nadawanej w subkodzie mówionym języka naturalnego przy użyciu systemu **pisma**, a więc proces **odwzorowywania tekstu mówionego danego języka naturalnego w tekst pisany**. Operacja ta zależy przede wszystkim od **rodzaju pisma** właściwego dla danego języka naturalnego. Zasadniczy problem zawiera się w pytaniu, jaki składnik komunikatu mówionego jest utwalany w tekście pisanym. Możliwe są tu dwie sytuacje: **odwzorowanie sygnału**, a więc materialnej, dźwiękowej strony komunikatu (**formy**) lub **odwzorowanie informacji** zakodowanej w tym sygnale (**treści** komunikatu). W pierwszym przypadku odbiorca mający utrwalić komunikat analizę komunikatu może zakończyć na identyfikacji sygnału i dokonaniu jego delimitacji na elementy składowe zależnie od tego, jakie elementy ciągu fonicznego odwzorowywane są w danym rodzaju pisma. Tak zidentyfikowanym wyrażeniom przyporządkowuje następnie jednostki planu wyrażania subkodu graficznego. Teoretycznie rzecz biorąc, nie musi więc przeprowadzać operacji dekodowania sygnału i odkodowywania ukrytej za nim informacji (nie musi tekstu rozumieć). W drugim wypadku analiza musi się zakończyć rozumieniem komunikatu, a więc uwzględnieniem również planu treści komunikatu.

W praktyce w utrwalaniu tekstów języków naturalnych wykorzystuje się dwie możliwości. W pierwszym wypadku mamy tzw. **pismo fonetyczne**, które mniej lub bardziej dokładnie odwzorowuje sygnał komunikatu mówionego (tak jak w języku polskim), w drugim wypadku tzw. **pismo ideograficzne**, które odwzorowuje bezpośrednio plan treści komunikatu (tak jak w języku chińskim). Możliwe są

również systemy mieszane – tak jest w jednym z rodzajów pisma japońskiego, gdzie znaczenie morfemów leksykalnych odwzorowywane jest przez ideogramy pisma chińskiego, natomiast morfemy o funkcji syntagmatycznej kodowane są w piśmie fonetycznym.

Pismo wykształciło się w sposób naturalny, nie spełnia więc warunków kodu optymalnego. I tak, **pismo fonetyczne** nie jest pismem fonologicznym, nie odwzorowuje więc systemu fonologicznego danego języka naturalnego: ani samych fonemów, ani ich ciągów, ani dystynktywnych cech fonemów. Nie ma też w nim jedno-jednoznacznego przyporządkowania między elementami pisma a elementami ciągu fonicznego (głoskami lub ich ciągami, np. sylabami). Różny też jest dla różnych pism fonetycznych (różnych języków) podział ciągu fonicznego na elementarne wyrażenia odwzorowywane w piśmie. Teoretycznie mamy dwie możliwości podziału:

- albo wydziela się głoski i im przyporządkowuje elementarne wyrażenia subkodu graficznego, czyli litery – tak jest w **piśmie głoskowym**, zwanym też **pismem alfabetycznym**;
- albo wydziela się grupy głosek (sylaby) i im przyporządkowuje się jednostki elementarne planu wyrażania – tak jest w **piśmie sylabicznym**, na przykład w piśmie języka hetyckiego lub w pewnych odmianach pisma japońskiego.

W pierwszym wypadku, a więc w piśmie alfabetycznym teoretycznie są dwie możliwości:

- w piśmie odwzorowuje się wszystkie głoski ciągu fonicznego,
- w piśmie odwzorowuje się tylko wybrane głoski ciągu fonicznego – w zasadzie tylko spółgłoski, a samogłoski się pomija. Tak jest na przykład w piśmie języka hebrajskiego, będącego w zasadzie **pismem spółgłoskowym**, w którym tylko dla celów dydaktycznych, na przykład w podręcznikach i w Biblii, wprowadza się oznaczenia samogłosek w postaci znaków diakrytycznych (jak pamiętamy, w języku hebrajskim, który jest językiem alternacyjnym, morfemy leksykalne mają charakter nieciągły i składają się wyłącznie ze spółgłosek).

Subkod graficzny języka polskiego korzysta ze zmodyfikowanego **alfabetu łacińskiego**, podobnie jak wiele innych języków europejskich. Alfabet łaciński teoretycznie odwzorowuje głoski tekstu mówionego, zwróćmy jednak uwagę na fakt, że w żadnym języku nie ma jedno-jednoznacznego przyporządkowania między głoskami a literami, a na przykład w wypadku języka angielskiego w wielu wypadkach można mówić o znacznym stopniu arbitralności przyporządkowania ciągów liter ciągom głosek. Dzieje się tak przede wszystkim dlatego, że zasady pisma angielskiego odwzorowują o wiele starsze stadium rozwojowe tego języka, a i sam alfabet nie był stworzony dla tego języka i o wiele lepiej nadawał się do odwzorowywania formy fonicznej klasycznej łaciny niż staroangielskiego. A i w łacinie przyporządkowanie między głoską a literą wymagało korzystania z reguł kontekstowych, na przykład przy odwzorowywaniu w piśmie głoski k, gdzie, zależnie od następujących elementów ciągu, mogła to być litera c, k lub q.

Podobnie jest w piśmie języka polskiego. Żeby odwzorować w piśmie tekst mówiony nadawca musi nie tylko znać reguły przyporządkowania między głoskami a odwzorowującymi je literami, znakami diakrytycznymi lub ciągami liter, ale także dysponować odpowiednią wiedzą o zestawie morfemów należących do jednej klasy abstrakcji, czyli znać wszystkie fonetyczne warianty kontekstowe jednego morfemu (allomorfy), gdyż pisownia polska w dużej mierze opiera się na zasadach morfologicznych. Do tego musi znać długie listy wyjątków. Zilustrujmy to przykładami.

- Głoski bezdźwięczne odwzorowywane są w piśmie na dwa sposoby, np.

p jako p (*pas*, bo *pasy*) i b (*babka*, bo *baba*),
s jako s (*pies*, bo *psa*) i z (*kózka*, bo *koza*),
t jako t (*kot*, bo *koty*) i d (*kod*, bo *kody*);

- Głoski miękkie odwzorowywane są w piśmie na różne sposoby:
 - przez znak diakrytyczny oznaczający miękkość (kreska ukośna nad literą oznaczającą twardy odpowiednik głoski miękkiej lub nad drugą z liter, w wypadku gdy głoska odwzorowywana jest przez dwie litery), np.

przyjaźń, śnieg, źdźbło,

- przez literę i następującą po literze oznaczającej twardy odpowiednik głoski miękkiej, , np.

ciasto, siano, zimno, miś.

Wybór wariantu jest tu uwarunkowany kontekstowo: znak diakrytyczny, jeśli spółgłoska miękka występuje w danym wyrazie przed spółgłoską, litera i, jeśli przed samogłoską. Zwróćmy uwagę na to że, w wyrazach *zimno*, *miś* litera i pełni dwie funkcje – oznacza miękkość głoski poprzedniej i jednocześnie jest znakiem samogłoski i występującej po głosce miękkiej,

- Jedna głoska może być oznaczana przez:

- jedną literę, np.:

a – a, e – e, m – m,

- dwie następujące po sobie litery, np.:

š – sz, (*szafa*), č – cz (*czas*), ž – dz (*dzwonek*), ć – ci (*ciasto*),

- może też mieć dwa różne sposoby odwzorowania, np.:

χ – ch (*chata*), h (*hala*),

u – u (*mur*), ó (*bór*),

t – t (*kot*), d (*kod*),

ž – ż (*żaba*), rz (*rzeka*)

Pamiętać należy, że reguły **odkodowywania pisma**, czyli **głośnego czytania** (transformacji sygnału subkodu wizualnego na subkod akustyczny) nie są symetryczne względem reguł kodowania, bo w języku polskim przyporządkowanie

między tekstem mówionym a pisanym jest wielo-wieloznaczne (ale to już nie jest problemem utrwalania informacji, tylko transformowania komunikatu z subkodu graficznego na subkod akustyczny).

Utrwalanie informacji w wypadku **pisma ideograficznego** wymaga innego rodzaju analizy komunikatu mówionego – identyfikacji w nim elementów ciągu fonicznego tożsamy z jednostkami planu treści języka, im bowiem jako całościom przyporządkowane są znaki pisma ideograficznego. W klasycznej postaci pismo ideograficzne koduje bowiem nie jednostki planu wyrażania języka mówionego, lecz bezpośrednio jednostki planu treści. Gdyby więc różne języki naturalne miały identyczny plan treści, a różniły się tylko planem wyrażania subkodu mówionego, to pismo ideograficzne nadawałoby się do utrwalania komunikatów w każdym z tych języków (zależnie od języka różnie by się te znaki czytało w subkodzie mówionym). I tak się stało z pismem chińskim, którego znaki były wykorzystywane do utrwalania tekstów innych niż chiński języków naturalnych, na przykład koreańskiego, dawniej wietnamskiego, a obecnie japońskiego (choć w języku japońskim istnieją również inne systemy pisma). W samych Chinach istnieje zresztą wiele różnych języków, ale wszystkie posługują się tym samym zbiorem **ideogramów** i tylko dzięki pismu Chińczycy z różnych prowincji mogą się ze sobą porozumiewać.

Utrwalanie informacji w językach posługujących się pismem ideograficznym jest jednak w praktyce o wiele bardziej skomplikowane, i to nie tylko ze względu na liczebność zbioru znaków (podstawowy zbiór zawiera ich kilka tysięcy). W praktyce w tekście pisanym koduje się również inne informacje, informacje o charakterze metajęzykowym. Dzieje się tak z kilku powodów. Tekst zapisany ideograficznie nie odwzorowuje sygnału tekstu mówionego, jeśli się więc nie zna tzw. czytania znaku, nie można tekstu pisanego przekształcić w tekst mówiony, tak jak można to zrobić w wypadku pisma alfabetycznego czy sylabicznego, poza tym, zbiór znaków takiego pisma, choć bardzo liczny, jest i tak mniejszy niż zbiór jednostek planu treści języka, ten sam znak koduje więc często nie jedną jednostkę planu treści, a kilka, zazwyczaj połączonych jakąś paradygmatyczną relacją semantyczną. W praktyce w piśmie tym istnieje zatem **homografia** i w stosunku do planu treści, i w stosunku do planu wyrażania jego subkodu akustycznego. Informacja kontekstowa zazwyczaj nie wystarczałaby do usunięcia takiej wieloznaczności (i wielofoniczności), dlatego też niezbędne jest zapisanie w tekście informacji umożliwiających jednoznaczne przyporządkowanie znakowi jednostki języka mówionego (sposobu czytania) lub jednostki planu treści (znaczenia). Problem komplikuje to, że pismo ideograficzne używane jest w językach o zupełnie odmiennej strukturze syntaktycznej: klasyczny język chiński jest monosylabiczny (jednostki leksykalne są jednosylabowe, w języku tym nie ma więc w zasadzie wyrazów wielomorfemowych) i izolujący, współczesny język chiński ma elementy aglutynacji, a język japoński jest językiem aglutynacyjnym. W tekstach japońskich znakami ideograficznymi zapisuje się więc tylko morfemy leksykalne, natomiast cząstki aglutynacyjne pismem sylabicznym – jest to więc pismo o charakterze mieszanym.

Z tego, co powiedzieliśmy o nadawaniu komunikatów języka naturalnego w subkodzie pisanym wynika, że od rodzaju pisma zależy, jakim operacjom w procesie odbioru poddawany jest komunikat pisany. Niezależny od rodzaju pisma jest właściwie tylko proces **identyfikacji znaku**, łatwiejszy lub trudniejszy zależnie od tego, ile i jakie cechy relewantne znaku zostały odwzorowane w sygnale. Sytuacje skrajne to wyraźny druk i niewyraźne pismo odręczne, a i w wypadku wyraźnego druku prawie zawsze mamy do czynienia z jakąś utratą informacji. Identyfikacja znaku odbywa się przez porównanie znaków tekstu pisanego ze zbiorem idealnych znaków systemowych (**grafemów**), stanowiących składnik kompetencji odbiorcy, przy mniejszym lub większym udziale informacji kontekstowej (wynikającej z otoczenia znaku), redundancji systemowej, ogólnej kompetencji językowej oraz znajomości sytuacji komunikacyjnej, a także znajomości rzeczywistości pozatekstowej odwzorowywanej w komunikacie. Te ostatnie czynniki nie są specyficzne dla odbioru tekstu pisanego, wspomagają bowiem odbiór wszelkich komunikatów języka naturalnego. Udział tych składników zależy głównie od „jakości” tekstu pisanego w odniesieniu do zakodowania w nim relewantnych cech pisma.

Identyfikacja formy znaków tekstu pisanego jest warunkiem koniecznym odbioru komunikatu, dalsze operacje zależą już od rodzaju pisma. Choć teoretycznie możliwa jest bezpośrednia identyfikacja znaków tworzących jednostki znaczące z reprezentowanym przez nie znaczeniem, a więc bezpośrednio przyporządkowanie wyrażeniom tekstu pisanego kodowanej przez nie informacji, to jednak zazwyczaj w językach kodujących foniczną stronę znaków **dekodowanie tekstu pisanego** odbywa się w praktyce za **pośrednictwem subkodu mówionego**. Ciąg znaków pisanych podlega więc transformacji do postaci sygnału fonicznego i dopiero na tej postaci tekstu odbywa się proces dekodowania komunikatu, choć oczywiście dekodowania nie musi poprzedzać faktyczna artykulacja. Odbieranie komunikatów w piśmie fonetycznym (alfabetycznym lub sylabicznym) wymaga więc kompetencji językowej w zakresie subkodu mówionego danego języka (kompetencję taką posiadają nawet osoby głuchonieme, jeśli posługują się językiem naturalnym o piśmie fonetycznym, i choć ich model systemu językowego związany jest z inną realizacją materialną niż w wypadku osób słyszących, to modele te możemy uznać za równoważne).

Pośrednictwo subkodu mówionego nie jest konieczne przy dekodowaniu tekstów zapisanych znakami ideograficznymi, choć w praktyce często odwzorowanie tekstu pisanego w subkod mówiony towarzyszy odbieraniu komunikatów pisanych. W wypadku pisma ideograficznego możliwe jest bowiem bezpośrednio przyporządkowanie znakowi graficznemu jego znaczenia, możliwe jest zatem rozumienie takich tekstów pisanych bez znajomości subkodu mówionego danego języka naturalnego. Tak właśnie się dzieje, kiedy Japończycy czytają teksty chińskie, i odwrotnie, a nawet kiedy Chińczycy z prowincji mówiących różnymi językami chińskimi czytają te same teksty. Można więc, teoretycznie, nauczyć się tylko znaczeń chińskich znaków i „czytać” teksty, nie umiając po chińsku mówić, co w wypadku języka polskiego byłoby co najmniej trudne.

Przetwarzanie informacji w systemach informacyjnych

Wszystkie operacje wykonywane w systemach informacyjno-wyszukiwawczych są operacjami dokonywanymi na tekstach i polegają na przetwarzaniu zawartych w nich informacji. Wymienimy je po kolei.

Na wejściu systemu mamy **dokument**, a więc utrwaloną informację. Dokument ten jest w systemie opracowywany. Opracowanie dotyczy strony formalnej dokumentu (opis formalny) i treści dokumentu (opis rzeczowy). **Opis formalny dokumentu** polega nie tylko na identyfikacji i selekcji informacji zgodnie z przyjętymi kryteriami relewancji (zasadami opisu formalnego), ale często wymaga transformacji tekstu z jednego pisma (pisma dokumentu) na pismo przyjęte w danym systemie informacyjnym, a więc transliteracji, np. z cyrylicy na alfabet łaciński lub transkrypcji, np. z pisma ideograficznego na pismo fonetyczne.

Transliteracja jest transformacją stosunkowo prostą, zazwyczaj reguły odwzorowania ustanawiają odpowiednie normy (związane z normami opisu bibliograficznego), operacja **transkrypcji** pisma ideograficznego jest znacznie trudniejsza. Tu po prostu trzeba znać język dokumentu oraz przyjęty w języku opisu odpowiednik znaku ideograficznego, przy czym należy pamiętać o tym, że odpowiedniki te mogą się zmieniać wraz ze zmianą **konwencji transkrypcji**. Dotyczy to na przykład języka chińskiego, gdzie taka zmiana nastąpiła w ostatnich latach – na przykład: dawniej *Mao Tse Tung*, obecnie *Mao Zedong*, tradycyjnie *Pekin*, obecnie *Beijing*. Dla opisu formalnego jest to istotny problem, problemem jest również wprowadzenie takiej informacji do zbioru wyszukiwawczego, jeśli bowiem nie wprowadzi się odpowiednich odsyłaczy, informacja zostanie w zbiorze rozproszona, a być może nawet stracona.

Przy opisie formalnym potrzebna jest także wiedza o języku oryginału dotycząca **konwencji zapisu imienia i nazwiska** – w wielu wypadkach decyduje o tym pozycja, w wielu językach europejskich imię stoi przed nazwiskiem (choć dla celów formalno-urzędowych bywa odwrotnie), ale w języku węgierskim lub chińskim jest odwrotnie. Problemem są języki, w których w ogóle nie ma podziału nazwy własnej osobowej na imię i nazwisko. Problemem może być również odróżnienie imienia od nazwiska w wypadku nazw własnych więcej niż dwuczłonowych, pomocne mogą tu być listy imion przyjętych w danym języku (jest ich zazwyczaj mniej niż nazwisk) lub listy nazwisk (na przykład w języku chińskim jest ich mniej niż imion i listę taką można uznać za pełną), co nie zawsze pozwala na rozstrzygnięcie wątpliwości, bo na przykład w kulturze amerykańskiej imię może mieć formę nazwiska (np. nazwiska przodka).

Innym problemem w opisie formalnym dokumentu może być **zapisanie roku** wydania dokumentu, przy czym nie zawsze problem ogranicza się do transformacji z jednego systemu zapisu na drugi, np. cyfr rzymskich na arabskie, bo przy różnych **systemach liczenia czasu** może wymagać nie tylko wiedzy o języku, ale również wiedzy o rzeczywistości pozajęzykowej.

Opracowanie rzeczowe dokumentu to wiele różnych procesów informacyjnych. Pierwszy proces informacyjny to odbiór komunikatu. Problemy związane

z tym procesem omówiliśmy wcześniej, tu tylko wypada zwrócić uwagę na to, że obecnie w systemach informacyjnych mamy do czynienia nie tylko z dokumentami pisanymi, ale również z innymi ich postaciami, na przykład z tzw. dokumentami multimedialnymi, w tym z dokumentami elektronicznymi, do których odbioru niezbędne są odpowiednie urządzenia transformujące sygnał.

Proces odbioru komunikatu bardzo często łączy się z procesem **selekcjonowania informacji** relewantnej albo ze względu na treść dokumentu, albo ze względu na potrzeby systemu informacyjnego (a właściwie na potrzeby użytkowników systemu). Przyjęte **kryteria relewancji** modyfikują proces odbierania komunikatów utrwalonych także w tym sensie, że wstępna eliminacja informacji nerelewantnej odbywa się równocześnie z odbieraniem komunikatu – wszak długich tekstów nie czytamy dokładnie, jeśli zależy nam tylko na ogólnej orientacji lub wyszukaniu w nich potrzebnej nam informacji. Przy takiej operacji przydatna jest **wiedza o strukturze tekstu**, a szczególnie o zależności między pozycją informacji w tekście a stopniem jej relewancji dla odbiorcy. Tekst utrwalony pozwala również na **nieliniarny odbiór informacji** zakodowanej w języku naturalnym, co w wypadku tekstu nieutrwalonego nie jest możliwe, jak pamiętamy bowiem linearność jest jedną z cech języka naturalnego.

Pierwszym tekstem pochodnym odwzorowującym treść dokumentu powinno być **streszczenie dokumentacyjne**, jeśli więc tekst dzieła nie został w streszczenie zaopatrzony przez wydawcę albo autora, które można by dla potrzeb systemu odpowiednio zmodyfikować (to też transformacja komunikatu – **kondensacja i selekcja**), powinniśmy takie streszczenie zrobić sami. Dokładniej o operacji streszczania mówiliśmy uprzednio, tu tylko dodamy, że streszczenie dokumentacyjne dla potrzeb konkretnego systemu informacyjnego nie musi być proporcjonalne. Tworząc streszczenie i selekcjonując informację, osoba opracowująca dokument korzysta ze swojej wiedzy pozajęzykowej (o potrzebach informacyjnych użytkownika, czyli adresata swojego komunikatu), a także z wiedzy o strukturze tekstów spójnych (informatywności poszczególnych partii tekstu) oraz z metainformacji zawartych w tekście. Formułowanie tekstu streszczenia dokumentacyjnego to nadawanie komunikatu w subkodzie trwałym.

Następnym etapem opracowania rzeczowego dokumentu jest tworzenie jego **charakterystyki słownej**, poprzedzone wyborem **słów kluczowych** (z tekstu dokumentu lub tylko ze streszczenia dokumentacyjnego). O słowach kluczowych mówiliśmy w rozdziale poświęconym statystycznym prawidłowościom w tekście, tu tylko dodamy, że przy ich wyborze często korzystamy z tekstów metainformacyjnych towarzyszących dokumentowi, takich jak indeksy, spisy treści oraz z metainformacji zawartej w samym tekście (np. różnego rodzaju wyróżniki graficzne: wytłuszczenia, druk rozstrzelony, żywa pagina czy żywy margines). Utworzenie charakterystyki słownej dokumentu przeważnie poprzedza transformacja wyrazów wybranych z tekstu do ich postaci kanonicznej (w wypadku języka polskiego jest to zazwyczaj – z określonymi wyjątkami – forma rzeczownika w mianowniku liczby pojedynczej). Operacja ta wymaga kompetencji w zakresie systemu morfologicznego języka naturalnego, wymaga bowiem sprowadzenia występujących

w tekście słowoform (w języku polskim różniących się postacią tematów i końcówkami fleksyjnymi) do wspólnej postaci systemowej. Pracochłonność tej operacji zależy więc od typu języka naturalnego – na przykład dla języka angielskiego jest stosunkowo prosta, dla języka polskiego skomplikowana, co ma znaczenie przy automatyzacji procesu wybierania słów kluczowych z tekstu.

Następną operacją jest **tłumaczenie charakterystyki słownej** na język informacyjno-wyszukiwawczy systemu, czego wynikiem jest **charakterystyka wyszukiwawcza dokumentu**. Jeśli językiem systemu jest język swobodnych słów kluczowych, wówczas operacja ta ma charakter tożsamościowy, jeśli język słów kluczowych kontrolowany tzw. stop-listą, należy sprawdzić, czy słowa kluczowe nie znajdują się na tej liście. Tłumaczenie słów kluczowych na inne języki informacyjne jest operacją znacznie bardziej skomplikowaną, przy czym zawsze należy się liczyć ze stratami informacji, różnymi przy różnych językach informacyjno-wyszukiwawczych, co często prowadzi do tzw. **indeksowania uogólniającego**.

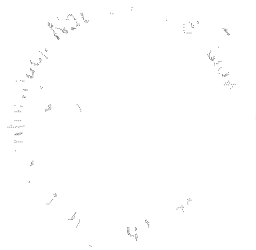
Utrwalenie komunikatu charakterystyki wyszukiwawczej dokumentu daje w wyniku **dokument pochodny** względem dokumentu opracowywanego. Następną operacją jest włączenie informacji pochodnej do **zbioru wyszukiwawczego** systemu, często, jak to się dzieje na przykład w systemach hipertekstowych, w określone miejsce struktury tego zbioru.

Analogicznym operacjom poddaje się **zapytanie informacyjne**, a więc komunikat wyrażający potrzeby informacyjne użytkownika systemu kierowane do systemu. Wynikiem tych operacji jest **instrukcja wyszukiwawcza**, sformułowana w języku wyszukiwawczym systemu, który w części odnoszącej się do treści dokumentu jest identyczny z językiem informacyjnym systemu. Różnica polega zazwyczaj na rozszerzeniu języka informacyjnego o tzw. **operatory logiczne** (boolowskie), określające relacje między zbiorami dokumentów denotowanymi przez wyrażenia charakterystyk wyszukiwawczych w zbiorze wyszukiwawczym systemu identyczne z wyrażeniami instrukcji wyszukiwawczej. Relacje te wyznaczają podzbiór dokumentów wyszukanych w trakcie selekcionowania dokumentów na podstawie ich cech relewantnych wyrażonych w instrukcji wyszukiwawczej (a ściśle biorąc, podzbiór charakterystyk wyszukiwawczych tych dokumentów). W najprostszym wypadku wyszukiwanie dokumentów w zbiorze wyszukiwawczym systemu polega na porównywaniu instrukcji wyszukiwawczej z charakterystykami wyszukiwawczymi zgromadzonymi w zbiorze. W **wyszukiwaniu paradygmatycznym** jest to operacja znacznie bardziej skomplikowana, porównywanie i selekcję poprzedza bowiem dopisywanie (w systemach komputerowych zazwyczaj automatyczne) do instrukcji wyszukiwawczej otrzymanej w wyniku indeksowania zapytania informacyjnego wyrażenń połączonych w systemie języka informacyjno-wyszukiwawczego paradygmatycznymi relacjami nadrzędności wyszukiwawczej lub relacjami asocjacyjnymi.

Pełna automatyzacja procesów informacyjnych w systemach informacyjno-wyszukiwawczych wymaga uprzedniej **formalizacji** tych wszystkich operacji, przy czym najtrudniejszym problemem jest formalizacja operacji transformacji tekstów języka naturalnego.

Literatura

- Majewicz A.F.: *Języki świata i ich klasyfikowanie*. Warszawa 1989, (rozdz. 3.5. *Typologia systemów graficznych*)
- Witold Marciszewski: *Metody analizy tekstu naukowego*. Warszawa 1981
- Diringer D.: *Alfabet*. Warszawa 1972
- Duszek A.: *Tekst, dyskurs, komunikacja międzykulturowa*. Warszawa 1998
- Huszczka R.: *Honoryfikatywność*. Warszawa 1996
- Artowicz E.: *Reprezentacja wiedzy w systemie informacyjno-wyszukiwawczym*. Warszawa 1997
- Sosińska-Kalata B.: *Modele organizacji wiedzy w systemach wyszukiwania informacji o dokumentach*. Warszawa 1999



Seria wydawana z inicjatywy
INSTYTUTU INFORMACJI NAUKOWEJ I
STUDIÓW BIBLIOLÓGICZNYCH
UNIWERSYTETU WARSZAWSKIEGO
oraz
WYDAWNICTWA SBP