

Téma 2 – Lingvistické a sémantické problémy obsahové analýzy

1. Co je to obsahová analýza (OA)?
2. **Ontologie OA**
Co je to obsah?
Kde je obsah?
3. **Gnozeologie OA**
Jak zjistit obsah?
Jak vyjádřit obsah?
4. **Praxe OA**
Jak chápat/modelovat obsah?

ÚISK – OBA T02

1

1

Lexikální subsystém jazyka – způsoby zaznamenání

Korpus
(parole)

syntagmatické

„přirozený jazyk
v počítači“
– Český národní korpus
<http://www.korpus.cz/>

Tezaurus
(langue)

paradigmatické



ÚISK – OBA T02

2

2

Výkladový slovník vs. jazykový tezaurus

slovník	tezaurus
způsob použití	
slovo (znak)	pojem (význam)
Co znamená? ↓	↓ Jak vyjádřit?
pojem (význam)	slovo (znak)
základní (primární) uspořádání	
abecední	systematické
základní (primární) proces	
sémaziologický	onomaziologický

ÚISK – OBA T02

3

3

Jak vyjádřit obsah / potřebu informace?

Lingvistické a sémantické problémy
organizace znalostí – asymetrické vztahy 1 : N

paradigmatické
syntagmatické

- 1) **synonymie** (1 význam – N lexémů)
- 2) **homonymie** (1 lexém – N významů)
XOR, buď / nebo, vylučující
- 3) **koordinace – složené pojmy / složená témata**
(N významů)
AND, a zároveň, partitivní
- 4) **slovní spojení** (1 význam – N lexémů)

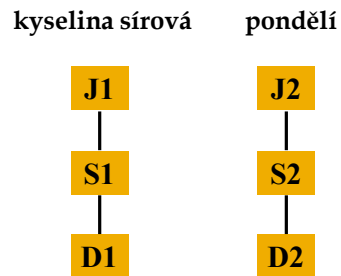
ÚISK – OBA T02

4

4

Ideál – vzájemně jednoznačná korespondence znak – smysl – denotát

Problém:
víceznačnost,
synonymie,
neurčitost
významu

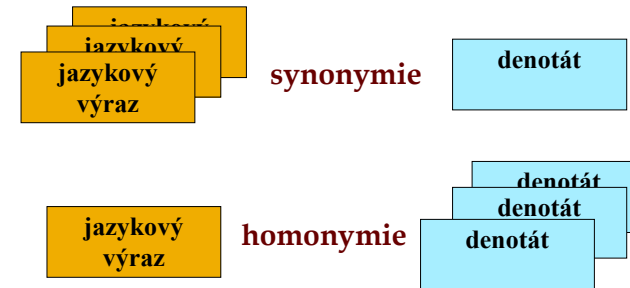


ÚISK – OBA T02

5

Problém – nejednoznačnost, neurčitost významu

Typy ekvivalence v jazyce

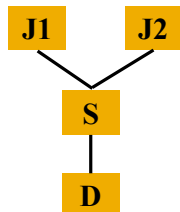


ÚISK – OBA T02

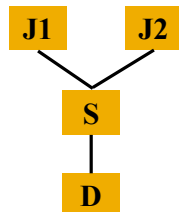
6

Synonymie

fotbal – kopaná



doktor – lékař



ÚISK – OBA T02

7

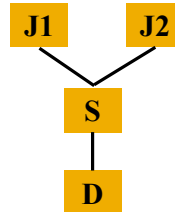
Typy synonymie

- a) **lexikální (věcná, sémantická)**
 - „pravá“ synonymie (ekvivalence)
cesta – silnice – dálnice – vozovka – stezka – pěšina
 - hierarchie
domácí zvíře – kočka – Macek
 - asociace
nevolnost – nemoc, plný – prázdný
- b) **gramatická (morfologická)**
1 slovo ve více tvarech – *cesta, cesty, cestě, cestu...*
- c) **mezijazyková ekvivalence**
kočka – die Katze – kouwka – cat
- d) **formální**
budova, budou, ubudou, budoucnost, vybudovat...

ÚISK – OBA T02

8

Ekvivalence v synonymii



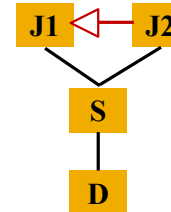
2406 **ZNAK** (logický) čeho (pojmu, druhu, jedince, pravdy ap.; z. základní, podstatný, nezbytný, nutný, odlišující, roz-, význačný, charakteristický, jistý, bezpečný, nemýlný, neo-), podstatná, příznačná, charakteristická známka, znamení čeho (vzdělanosti, poctivosti), význačná, rozlišující vlastnost, podstatný, charakteristický rys čeho, *kniž.* význak, **přívlastek** (p-y lidské; p-y hluboké povahy), **atribut** (pojmu, realismu), **charakteristikon** (národnosti, génia), **signum** (neomylné s. pravého umění), **pečeť** (individuality, pravdy), *zast.* **oznak** (hlavní o. pravdy), **přímět(ek)**, *zř.* **poznatek** (egoismu); *log. z.*, určující prvek při stanovení pojmu
viz též typ, typický

ÚISK – OBA T02

9

Hierarchie v synonymii

archilexém

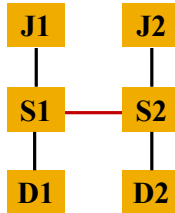


2406 **ZNAK** (logický) čeho (pojmu, druhu, jedince, pravdy ap.; z. základní, podstatný, nezbytný, nutný, odlišující, roz-, význačný, charakteristický, jistý, bezpečný, nemýlný, neo-), podstatná, příznačná, charakteristická známka, znamení čeho (vzdělanosti, poctivosti), význačná, rozlišující vlastnost, podstatný, charakteristický rys čeho, *kniž.* význak, **přívlastek** (p-y lidské; p-y hluboké povahy), **atribut** (pojmu, realismu), **charakteristikon** (národnosti, génia), **signum** (neomylné s. pravého umění), **pečeť** (individuality, pravdy), *zast.* **oznak** (hlavní o. pravdy), **přímět(ek)**, *zř.* **poznatek** (egoismu); *log. z.*, určující prvek při stanovení pojmu
viz též typ, typický

ÚISK – OBA T02

10

Asociace v synonymii



2406 **ZNAK** (logický) čeho (pojmu, druhu, jedince, pravdy ap.; z. základní, podstatný, nezbytný, nutný, odlišující, roz-, význačný, charakteristický, jistý, bezpečný, nemýlný, neo-), podstatná, příznačná, charakteristická známka, znamení čeho (vzdělanosti, poctivosti), význačná, rozlišující vlastnost, podstatný, charakteristický rys čeho, *kniž.* význak, **přívlastek** (p-y lidské; p-y hluboké povahy), **atribut** (pojmu, realismu), **charakteristikon** (národnosti, génia), **signum** (neomylné s. pravého umění), **pečeť** (individuality, pravdy), *zast.* **oznak** (hlavní o. pravdy), **přímět(ek)**, *zř.* **poznatek** (egoismu); *log. z.*, určující prvek při stanovení pojmu
viz též typ, typický

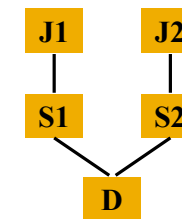
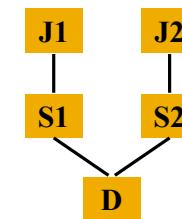
ÚISK – OBA T02

11

Kvazisynonymie, antonymie

fotbalista – Jan Kovář

tvrdý – měkký

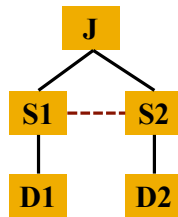


ÚISK – OBA T02

12

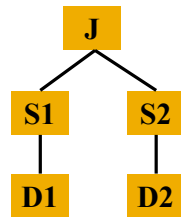
Homonymie / mnohovýznamovost (polysémie)

jazyk - jazyk



polysémie

los - los



homonymie

ÚISK - OBA T02

13

Typy homonymie

a) lexikální

stát - stát - stát - stát, bez - bez, Let let, léta letí, a - a

- zvuková (*mít - mýt, ve při - vepři, sním - s ním*)
- grafická (*email - email, . - . - .*)

b) syntaktická

vhazování domácích

církev chce vrátit ukradený majetek

moc slov - moc slov

c) sémantická - ironie, metafora, metonymie

to se ti povedlo

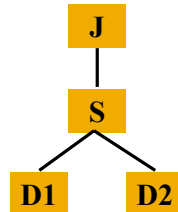
šli na skleničku

ÚISK - OBA T02

14

Homonymie - mnohoznačnost (dvojsmysl)

to je pěkné - to je pěkné



ÚISK - OBA T02

15

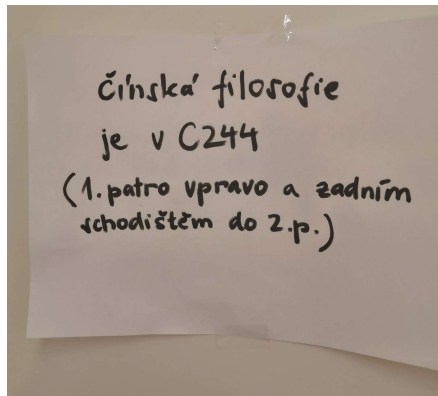
Homonymie - metafora



Nedělala bych si marné naděje. Vítr v peněžence
ještě nikdy žádnou vrtuli neroztočil...

16

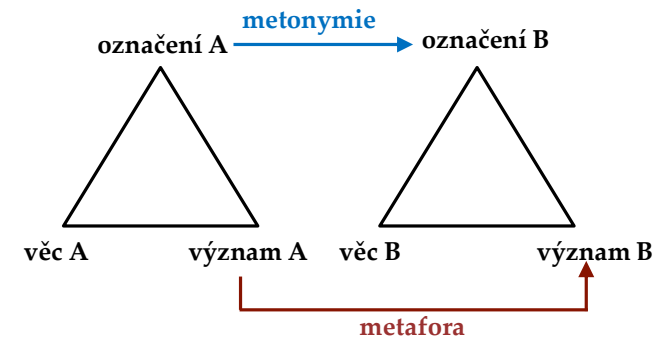
Homonymie – metonymie



17

17

Metafora a metonymie = přenos významu / označení

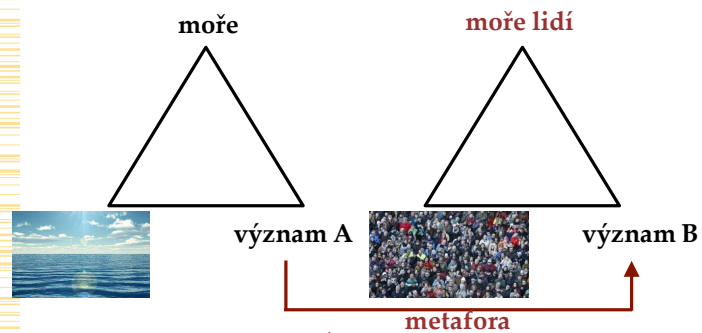


ÚISK – OBA T02

18

18

Metafora = přenos významu

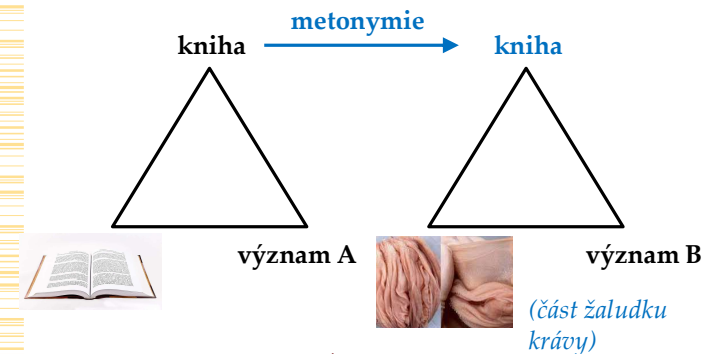


ÚISK – OBA T02

19

19

Metonymie = přenos označení



ÚISK – OBA T02

20

20

(část žaludku
krávy)

Složené téma / pojem

1 pojem – více jednotek obsahu

- elektrokolo
- husopaska
- teploměr kyber-egyptologie
- e-learning
- samoobsluha
- narkobaron, uhlobaron
- zahraničněpolitický
- alfanumerická notace
- aktivní meziknihovní výpůjční služba
- automatická tvorba tezauru
- daň z přidané hodnoty
- matematicko-statistická teorie informace

ÚISK – OBA T02

21

21

Slovní spojení

též sousloví, sdružené pojmenování, fráze, frazém, idiom (*phrase*), n-gram, víceslovný lexém, kolokace, koherentní víceslovný výraz

1 jednotka obsahu – více lexémů

*Jan Neruda, černí baroni, stavební úřad, pražské jaro, láry fáry,
New York City, postavení mimo hru, lže si do vlastní kapsy*

- Ustálené pojmenování ze dvou nebo více slov, které **označuje jednu skutečnost** a vstupuje do věty jako hotový celek
- **Nejčastější forma:**
substantivní spojení: lexikální jednotka, kterou lze rozčlenit na oddělené složky, přičemž každá z těchto složek může být vyjádřena substantivem a nezahrnuje sloveso ve tvaru určitém

ÚISK – OBA T02

22

22

Jak poznat složené téma / pojem

Zaměnitelnost

elektronická

kniha

přísluví

modlitební

cigareta

pro mládež

dětská

jízdenka

se slevou

ÚISK – OBA T02

23

23