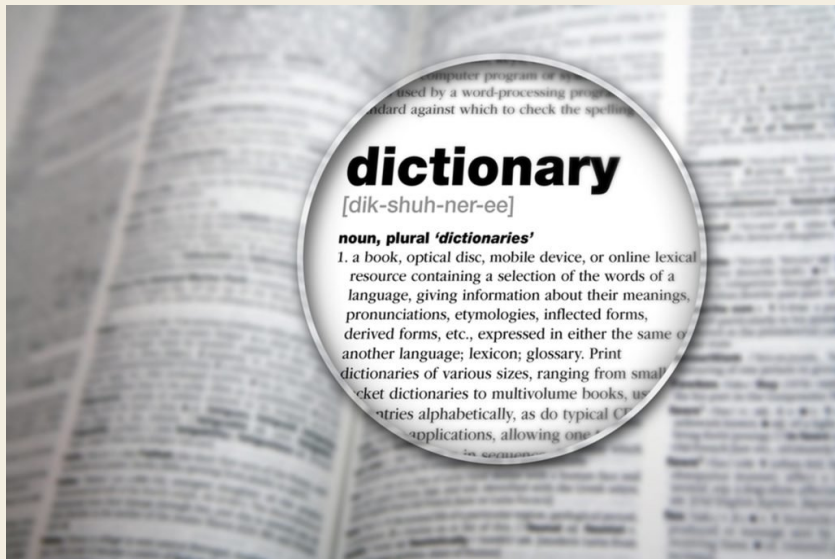**FACULTY
OF INFORMATICS**
Masaryk University

Automatically Created Noun Explanations for English

Marie Stará

Definition? Word Meaning Explanation?

Sketch Grammar

```
*STRUCTLIMIT s
*WSPOSLIST ",noun,-n,verb,-v,adjective,-j,adverb,-a,pronoun,-d"
#nouns
=and_other
1:Noun andOther 2:Noun

=hypo_hypero2
2:Noun (andOr Det? Adj? Noun)* andOr (Also)? Similar (Det? Adj{0,2} Noun)? Det? 1:Noun

=hypo_hypero
2:Noun (andOr Det? Noun)? toBe Det? Type Det? 1:Noun

=WORD_is
1:Noun upTo5words toBe upTo5words 2:Noun

=WORD_has
1:Noun upTo5words toHave upTo5words 2:Noun

=what_who_has_WORD
2:Noun upTo5words toHave upTo5words 1:Noun

=adj_modif
2:Adj Det? adjReflComma Det? 1:Noun
2:Adj andOr Det? adjReflComma Det? 1:Noun
adjReflComma andOr 2:Adj Det? 1:Noun
1:Noun 2:Adj
```

dictionary

(noun)
British National Corpus (TreeTagger) freq = 1,499 (13.26 per million)

adj_modif			is_subject			is_object			N_coord			gen		
		28.42			**18.48**			**15.41**			**5.34**			**4.34**
domain-specific	22	10.60	outperform	3	8.24	compile	13	9.08	thesaurus	9	11.36	slang	4	10.65
domain-specific dictionary			compile	6	8.05	compiled a dictionary			dictionaries and thesauri			quotation	3	9.30
machine-readable	20	10.18	index	3	8	consult	3	6.38	encyclopedia	4	10.49	language	7	5.94
machine-readable dictionaries			define	8	6.07	comprise	3	5.24	encyclopaedia	4	10.43	dictionary of the english language		
readable	9	8.99	my dictionary defines			speak	3	5.14	grammar	7	10.08	word	6	5.74
machine readable dictionaries			test	3	5.44	use	51	5.13	in dictionaries and grammars					
biographical	6	8.08				use standard english dictionaries			lexicon	3	9.89			
bilingual	5	7.92												

WORD_is			what_who_has_WORD		
		3.47			**1.13**
use	3	5.58	word	3	7.64

for_example			consists_of		
		1.87			**0.73**
source	4	7.28	contribution	4	9.34

Lexical Computing

2.36.7-SkE-2.163.4-3.116.13

Code → Result

dictionary
meanings similar to dictionary can also be expressed by (a/an)
thesaurus, encyclopedia, encyclopaedia, textbook, guide, lexicon, use
domain-specific, machine-readable, readable dictionary
for example (a/an) source
can have/contain (a/an) contribution
(a/an) word can have/contain (a/an) dictionary
is a subject of outperform, compile, index
is an object of compile, consult, comprise
of (a/an) slang, quotation, language

Evaluation

- 70 nouns (64 in BNC)
- frequency (from 2 to 67 826); abstract & concrete nouns; homonyms; synonyms; ad hoc selection
- "mostly good"

Somewhat Discouraging Results

mamluk (35)
is an object of exist

excavator (43)
meanings similar to excavator can also be expressed by (a/an) digger,
tractor, shaker, extractor, harvester, pride
mini, rotary, bulk excavator
for example (a/an) bow, hall, town
can have/contain (a/an) coin, finance, right
(a/an) other can have/contain (a/an) excavator
is a subject of sex, indemnify, assign
is an object of alert, prompt, decide
of (a/an) heart

Somewhat Encouraging Results

theatre (2)

meanings similar to theatre can also be expressed by (a/an) cinema, concert, opera, gallery, museum, play, life

open-air, Kuwaiti, provincial theatre

(a/an) experience can have/contain (a/an) theatre

is a subject of reopen, crowd, close

is an object of operate, tour, love

of (a/an) operation, war, dream

Somewhat Encouraging Results

bed (244)
similar meaning can have (a/an) breakfast, border, table, room, door, wall
double, twin, unmade bed
for example (a/an) doctor, child
(a/an) price, kitchen, room can have/contain (a/an) bed
is a subject of stare, knit, sleep
is an object of share, wet, strip
of (a/an) rose, lettuce, nail
with (a/an) flu, someone, sheet

More Encouraging Results

house (45 503)
meanings similar to house can also be expressed by (a/an) home,
sale, garden, flat, shop, building, room
public, upper, big house
for example (a/an) home, hotel
can have/contain (a/an) garden, bedroom, room
(a/an) family, time, people can have/contain (a/an) house
is a subject of build, stand, sell
is an object of terrace, buy, build
of (a/an) parliament, card, lord
with (a/an) garden, roof, wall

Correction

corpus. Three words are interpreted as a different part of speech (*an expose, (a) lead* are recognised/appearing in the corpus only as verbs, *a mammoth* as an adjective).

- *(a) lead*, *a mammoth* are present as nouns in BNC