

How to make a dictionary II

The Microstructure

Ruth Vatvedt Fjeld, University of Oslo



The Microstructure

= the internal structure of each entry of a dictionary

- Which information is to be given for each entry?

What does it mean to know a word?

1. the word's pronunciation
2. the word's spelling
3. the word's morphological components (if any)
4. the words that are morphologically related to it
5. the word's syntactic behaviour in a sentence
6. the full range of the word's meaning
7. the appropriate situations for using the word
8. the word's collocational restrictions
9. the word's distribution
10. the relation between the word and other words within a lexical set

Sense distinctions

Table 1 Corpus lines for *crawl**, cited in Atkins, Rundell & Weiner (1997: slide 6abc2)

You have to	crawl	along these tunnels.
Exhausted fugitives	crawl	from the lake.
Too tired even to read, he	crawled	into bed.
A two-mile tail-back	crawled	towards the Auditorium.
...as if a gigantic spider had just	crawled	across the table.
You've got little brown insects	crawling	about all over you.
The whole kitchen was	crawling	with ants.
East Germany is	crawling	with spies and traitors.
Angela Morgan's car was being	crawled	over inch by inch by a forensic team.
...to get women to support us by	crawling	to them.
Dark heavy clouds were	crawling	across the sky.
There was a little sheep trail	crawling	up the hillside.
She was having little chats as she	crawled	down the list.
The days before then seemed to	crawl	past.

DeepDict-analysis

(Bick 2008)

- **Subject + *krype* :**

PERS: jeg, hun, han, de, vi, du, alt

10.94:7 **PROP-hum** · 5.79:4 **kulde** · 5.46:3 **maur** ·

4.95:3 skygge · 4.98:2 ryggrad · 4.91:2 **skorpion** ·

5.77:1 **uhygge** · 4.65:2 **edderkopp** · 4.51:2

skumring · 4.27:2 kvikksølv · 3.98:2 **orm** · 4.96:1

hyllemeter · 3.69:2 **slange** · 2.54:3 **insekt** · 2.52:3

unge · 2.26:3 hånd · 4.16:1 **kakerlakk** · 2.96:2 katt

· 2.83:2 kveld · 1.82:3 hund · 2.61:2 **følelse** · 2.5:2

mørke · 3.48:1 **krabbe** · 1.46:3 menneske · 3.46:1

redsel · 2.77:1 PROP-org · 1.96:1 PROP

Collocations with *krype*

krype ...

7.9:8 inn · 9.07:6 rundt · 10.04:4 helt · 7.81:6 oppover · 8.28:4 bortover · 5.45:6 som · 6.22:5 tett · 4.85:6 igjen · 5.29:5 nær · 6.72:3 forsiktig · 5.34:4 nedover · 3.26:5 tilbake · 4.47:3 overalt · 4.44:3 godt · 3.43:4 litt · 3.26:4 lenge · 4.19:3 nede · 2.95:4 nå · 4.47:2 innover · 2.03:4 bare · 3.79:2 over · 3.79:2 langt · 3.79:2 snart · 4.77:1 formelig · 3.58:2 framover

krype under ...

9.07:5 dyne · 5.18:3 teppe · 5.87:1 sengeteppe · 4.57:2 gjerde · 5.43:1 laken · 2.83:3 hud · 3.55:2 seng · 2.55:2 bord · 2.76:1 sofa · 2.13:1 felle · 0.64:1 vindu

krype til ...

6.98:5 kors · 5.54:3 PROP-hum · 1.13:3 slutt · 2.93:1 høyre · 1.78:1 kilo · 1.45:1 vindu · 1.34:1 ball · 0.08:1 åpning

krype på ...

6.36:4 kne · 4.71:3 gulv · 4.58:3 mage · 3.6:3 fange · 3.62:2 kanin · 4.62:1 fang · 3.86:1 golv · 2.1:2 vegg · 2.43:1 kjøkken · 1.26:2 rom · 2.23:1 loft · 1.96:1 gradestokk · 1.86:1 seng · 1.8:1 buk · 0.56:2 måte · 1.41:1 innside · 1.4:1 lastebil · 1.25:1 skulder · 1.22:1 dekk · 1.04:1 mark · 1.02:1 rygg · 0.45:1 bakke · 0.41:1 kant · 0.06:1 hånd

krype gjennom ...

4.05:3 vindu · 3.38:3 hull · 4.96:1 søle · 3.45:2 gress · 3.19:1 gjerde · 0.46:2 kropp · 0.89:1 åpning · 0.77:1 hals · 0.77:1 munn

krype foran ...

4.9:2 peis

krype ved ...

5.87:1 bardisk · 1.61:3 side · 1.75:1 bål

krype fra ...

4.87:2 skjulested · 4.2:1 gjemmeded · 3.23:1 dyne · 1.73:1 skygge · 0.67:1 mørke · 0.41:1 kant

Collocation =

two lexemes in a grammatically well formed construction

noun-noun: *railway, chemin-de-fer, jernbane :*

noun-verb: *give a speech (*hold a speech),
make a decision (*take a decision)*

noun-adjective: *tall man (*high man)*

verb-adverb: *behave nicely/badly/well/cool*

NB! Such categories will be language specific.

What does the users want?

1. Translation or equivalent
2. Semi colons over numerals to separate senses
3. Frequency indication
4. No unknown symbols, like ~ as place-holder

Meaningful clusters

Also called 'multiword chunks or recurrent phrases:

komme	11.68:9 PROP- <u>hum</u> · 8.07:9 anvendelse · 7.49:9 uttrykk ·
*til ...	9.51:6 PROP-civ · 7.42:8 gode · 7.33:7 hekte · 5.84:8
	enighet · 6.53:7 unnsetning · 6.17:7 fradrag · 4.38:8
	stykke · 5.04:7 PROP-org · 4.89:7 nytte · 3.85:8 makt ·
	4.73:7 konklusjon · 3.66:8 resultat · 3.59:8 slutt · 3.06:8 hus ·
	3:8 verden · 3.68:7 kort · 3.67:7 overflate · 2.4:8 sted ·
	3.03:7 punkt · 2.6:7 skade · 3.6:6 bord · 1.53:8 by

- Komme til anvendelse – come into use
- Komme til uttrykk – be expressed
- Komme til gode – benefit
- Komme til enighet – agree
- Komme til unnsetning – help/assist
- Komme til fradrag – minus
- Komme til kort – fail
- Komme til verden – be born
- Komme til overflaten – appear/show
- Komme til skade – be injured

Concordance

- ib13-
5384926 Ikke før filmfolkene har satt sine bein i landsbyen begynner gamle stridstema å **komme til overflaten** , og det ene liket etter det andre dukker opp .
- ib13-
545788 Beruset av mitt eget mot svømmer jeg lenge under vann , og idet jeg **kommer opp til overflaten** legger jeg meg over på ryggen og svømmer mot land med blikket vendt mot den blå himmelen .
- ib13-
1869552 Traumer i barndommen legger grunnen for ubevisste konflikter som senere i livet kan **komme opp til overflaten** som psykiske sykdommer .
- ib13-
5301517 Under Karl den tredje ble de strenge lovene sigøynerne levde under mildnet , og deres undergrunnskultur **kom** endelig **til overflaten** og ut til folket .
- ib13-
1991386 Her har han ikke vært før , og han vet at nå gjelder det å sparke fra for å **komme seg opp til overflaten** igjen .
- ib13-
1782990 I Iran finnes det i dag salt-diapirer med salt som stiger opp fordi det har lavere tetthet enn bergartene over og på siden og samtidig er plastisk defor merbart i områder hvor det er så tørt at saltet etter at det har **kommet opp til overflaten** ikke blir skylt bort av regn vannet , men isteden flyter utover til siden i form av saltbreer .
- ib13-
5359612 Vi er tilfreds og bryr oss ikke om kritikken som er **kommet til overflaten** , sier pappa Hjalmar .
- ib13-
114985 Jeg kjenner at noe av nervøsiteten jeg føler nå er den samme som den gangen , som om den har ligget i skjul i meg helt siden da , og bare ventet på en anledning som denne til å **komme til overflaten** eller i hvert fall nærmere **overflaten** igjen .
- ib13-
5004892 De som følger med i avisene fikk med seg jødehatet som **kom til overflaten** under Gazademonstrasjonene i januar 2009 .

How to select the best examples

ankomme (arrive) Bokmålsordboka

- *han er nettopp ankommet fra Paris*
(he has just arrived from Paris)
- *hun ankom i shorts*
(she arrived in shorts)



From newspaper language (LBK)

- **Tusenvis** av talibanere har **ankommet** provinsene Helmand , Ghazni , Uruzgan , Kandahar, Kunar og Zabul
(thousands of talibans have arrived the provinces of Helmand , Ghazni , Uruzgan , Kandahar, Kunar og Zabul)

The sentence example shows

- The most typical/natural context
- The most frequent inflection
- The most frequent cluster
- The real life situation

