Collocational networks and their application to an E-Advanced Learner’s Dictionary of Verbs in Science (DicSci)

Araceli Alonso, Chrystel Millon, Geoffrey Williams
araceli.alonso; crystel.millon; geoffrey.williams@univ-ubs.fr

Equipe LiCoRN-HCTI, Université de Bretagne-Sud
Equipe InfoLex-IULA, Universitat Pompeu Fabra
Outline

- Starting Points
- The DicSci – An E-Advanced Learner‘s Dictionary of Verbs in Science
- Building-up an Organic Dictionary
- Collocational Networks: an example
- Conclusions and Future Remarks
Starting Points

- the dictionary as an ongoing learning tool for scientists
- the place of science in learner’s dictionaries
- the role of verbs in science
- the relationship between verbal and nominal forms
- the relationship between ‘specialised’ and ‘general’ words
The DicSci – An E-Advanced Learner’s Dictionary of Verbs in Science

- Virtual Dictionary
- Corpus-driven
- Dictionary for Non-native speakers of English
- Dictionary of Verbs
- Dictionary of Patterns
- Dictionary of Science
- Organic Dictionary

A bottom-up dictionary of verb patterns with corpus-driven thematic and conceptual groupings
Sinclairian Perspective

- Idiom principle (Sinclair 1991)
- Pattern grammar (Hunston & Francis 1999)
- Collocational phraseological patterning (Gledhill 2000)
- Collocational networks (Williams 1998)
- Collocational resonance (Williams, Hanks)
- Lexical primings (Hoey 2005)
- Norms and exploitaitons and Corpus Pattern Analysis (Hanks 2004)
Building-up an Organic Dictionary

Methods

- **Collocational Networks (Williams 1998)**
  Networks of statistically related collocates developed from a core lexical unit

- **Intertextual Collocational Resonance (Williams 2008)**
  Carrying over of aspects of meaning from one context to another, consciously and subconsciously

- **Corpus Pattern Analysis (Hanks 2004)**
  Work-in-progress corpus-driven methodology for mapping meaning onto use in texts
Building-up an Organic Dictionary

Collocational Network

Collocational resonance

Dictionary entry

CPA

Conceptual classification
Building-up an Organic Dictionary

BMC Corpus

SCIENTEXT INITIATIVE

- Open Source
- Size 33 million words
- Fully POS tagged and lemmatised
- Subcategorised for topic and genre

http://scientext.msh-alpes.fr/scientext-site/?article30
http://www.biomedcentral.com
Collocational Networks: an example

to treat
<table>
<thead>
<tr>
<th>No.</th>
<th>%</th>
<th>Pattern / Implicature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>69%</td>
<td>**[Human 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**([Human 1</td>
</tr>
<tr>
<td>2</td>
<td>17%</td>
<td>**[[Human 1 = Health Professional]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**([Human 1 = Health Professional]) applies a [Drug] or [Process = Medical] to [[Human 2 = Patient]] for the purpose of curing the patient’s [Disease</td>
</tr>
<tr>
<td>3</td>
<td>5%</td>
<td>**[[Human]] treat [[Inanimate]] (with [[Stuff]]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The chemical or other properties of [[Inanimate]] are improved or otherwise changed by [[Process]] or the application of [[Stuff]]</td>
</tr>
<tr>
<td>4</td>
<td>5%</td>
<td>**[[Human 1]] treat [[Human 2</td>
</tr>
</tbody>
</table>
|     |     | **[[Human 1]] gives or pays for [[Eventuality = Good]] as a benefit for [[Human 2 | Self]]**
Collocational Networks: an example

Concordances to treat

- A more consistent weight increase was achieved from week 8 to week 10.5 in mice treated with S. gordonii GP1294, with statistically significant differences compared to animals inoculated with the control strain.

- We observed a 16% increase in heart mass in both sexes after 3-days of isoproterenol treatment compared to mice treated with equivolume of saline.

- N-acetylcysteine prevents exacerbations of COPD because it is an anti-inflammatory agent and/or antioxidant, it may be difficult to see additional benefit in established exacerbations of COPD when the patients are also treated with prednisone, which has anti-inflammatory actions and the potential to reduce formation of reactive oxygen species from inflammatory cells.

- Control cells were treated with ethanol vehicle.
Collocational Networks: an example

DicSci

[[Human 1 | Human Group]] treat [[Human 2 = Patient | Laboratory Animal = Rat, Mouse | Organism = Cell]] (with [[Drug= Vehicle]])

Implicature

For the purpose of being cured

For getting a cure to a disease
Measurements were performed using a moving vehicle in some of the main streets of the city of Athens (Greece).

The vehicle was moving with a speed of approximately 60 km/h. The tests were performed in such a way and using the relevant equipment in order to simulate the data calls from an Ambulance vehicle.

Adsflt, AdLacZ, or vehicle was injected directly into the tumours.
Collocational Networks: an example

Conceptual classification

- ‘Giving Drugs’: dissolve, deliver, administer, receive, inject
- Laboratory Animal: rat, mouse, rabbit
- Drug: Adsflt, AdLacZ
- Word sense discrimination algorithm (Millon 2011)

Verb classes

Noun classes
Dictionary entry

INVESTIGATE

VFRAS

Article has cross-references to:
- examine
- explore
- prob
- involve
- isolate
- know
- lead
- mark
- may
- measure
- mediate
- might
- need
- observe
- obtain
- OBTAIN (Levi)
- occur
- participate
- perform
- POSITION

Synonyms: examination, view...
Conclusions and Future Remarks

- natural selection of the main cognitive nodes of scientific texts
- links between lexical units
- thematic patterns in texts
- differences between general and specialised uses
- sense disambiguation of polysemic words
- semasiologically and onomasiologically organisation
- organic dictionary

- conceptual classification
- microstructure
- collocational resonance
Thank you for your attention!!!

araceli.alonso@univ-ubs.fr
chrystel.millon@univ-ubs.fr
geoffrey.williams@univ-ubs.fr

http://www.licorn-ubs.com/index_eng.html