



UNIVERSITY
OF WARSAW

Corpus Linguistics Workshop, Day 3

Exercises in semantics and discourse analysis

Maciej Rosiński

Topic 1. Synonyms

Synonyms can be close in meaning, but never identical. How would you explain the difference between the following adjectives? What kind of research questions can we ask about these words to investigate their meaning in context?

ripe – mature

Choose another pair and investigate the difference in meaning using AntConc

antique - old
deep – profound
happy - joyful
big - large
fast - quick
smart - intelligent
easy - simple
begin - start
strong - powerful
beautiful - attractive
quiet - silent
rich - wealthy

Topic 2. Antonyms:

The relationship of antonymy includes different kinds of oppositions. An antonym of a word can be its binary opposite (dead-alive), it might profile the opposite end of a linear spectrum (long-short) or refer to an action which reverses a given process (plug-unplug).

smart – dumb

happy - sad
big - small
fast - slow
easy - difficult
begin - end
strong - weak
beautiful - ugly

quiet - loud
rich - poor
bright - dark
dead - alive

Topic 3. Metonymy

Consider the sentence, *The American public opinion fears another Vietnam*
What is the meaning of *Vietnam* in this context? How does it relate to the country of Vietnam?

Some common types of metonymy include:

PLACE FOR EVENT
INSTRUMENT FOR ACTION – *He hammered a nail*
CONTAINER FOR CONTENT – *I drank two bottles*
TRAIT FOR PERSON – *See that blonde?*
CONTROLLER FOR CONTROLLED – *Bush invaded Iraq*

What kind of search terms could be used to find other examples of these metonymy types?

Topic 4. Metaphor

Conceptual Metaphor Theory defines metaphor as a set of conceptual mappings between a source and a target domain. In simple terms it means “understanding one thing in terms of another” (see Lakoff & Johnson, 1980, 1999). Metaphorical expressions that appear in language use are a manifestation of the underlying patterns of thinking. To find metaphorical expressions in a corpus, we need a way to “operationalize” this theory, that is, adapt it to the tools at our disposal.

Let us consider the CAUSES ARE FORCES metaphor. Can you think of any words which in their most concrete sense express the use of a force? Here are some examples:

pull, push, hit, throw, and even force

We will call such expressions “source domain terms/keywords”. Our assumption here is that the conceptual metaphor CAUSES ARE FORCES will utilize source domain terminology, such as *push* to describe cause-effect relationships.

You’ll notice that some uses of *push* are quite literal. In (1) the verb is used to express physical causation involved in the act of opening a door:

- (1) my feet carried me to the green door and I **pushed** it open

But if you take a minute or two, you’ll find examples such as:

- (2) drought is **pushing** already endangered species towards extinction.
(3) People learn what they can, repeat and imitate. If **pushed**, invent.

In (2) *push* expresses causation as well, but the cause-effect relationship between drought and extinction is a long, complex process, which does not involve a direct act of applying a physical force. Similarly, the example in (3), can be paraphrased by saying that people experiencing stress may be creative. The force *pushing* people to do this, however, is mental rather than physical.

Use AntConc to find more examples of the CAUSES ARE FORCES metaphor.

Protocols for semantic analysis

You might ask: “How do I know that a word like *push* belongs to the conceptual domain of forces, rather than the domain of causation?”. The question is legitimate, and here we operationalize the answer for the purpose of corpus analysis. [Dictionaries like LDoCE](#) list all the common meanings of a word. Among them, it is often possible to find one which the Pragglejaz group (2007) calls “basic”. Such word senses are:

- more concrete; what they evoke is easier to imagine, see, hear, feel, smell, and taste;
 - related to bodily action;
 - more precise (as opposed to vague);
 - historically older.
- (p. 3)

According to the Metaphor Identification Procedure (ibid.), if the meaning of a given word in context differs from its basic sense, but can be motivated by mapping the basic sense onto the contextual sense, the word can be treated as a linguistic metaphor.

Exercise

Here are some common source and target domains in conceptual metaphors (Kövecses, 2010):

Sources: BODY, PLANTS, BUILDINGS, MACHINES, FIRE, JOURNEYS, FORCES, LIGHT

Choose a source domain, and consider possible source domain keywords.

Your task is to copy 20 concordance lines into a spreadsheet, and compare the basic and dictionary senses of your keyword in each concordance line. Mark each line as metaphorical or not metaphorical, and suggest a general ‘A is B’ label as well.

Discussion topic

Here are some common metaphorical targets.

Targets: EMOTIONS, MORALITY, MIND, SOCIETY, POLITICS, TIME, RELATIONSHIPS, RELIGION

How to find metaphorical expressions about these domains inside a corpus?

Derivational semantics

Consider the derivational suffix *-er*, as in *singer*, *player*, *fighter*

How would you characterize the meaning of the schema [V-er]_N?

Is it the only meaning?

The *-er* suffix can also combine with nouns as in *geographer*, *villager*

How would you go about establishing all the meanings of the *-er* suffix?

Queries for tagged corpus in AntConc:

_V - verb search

J - adjectives

N - nouns

RP - particles

R - adverbs

P - pronouns

I - prepositions

*_V*_be - will give all forms of be

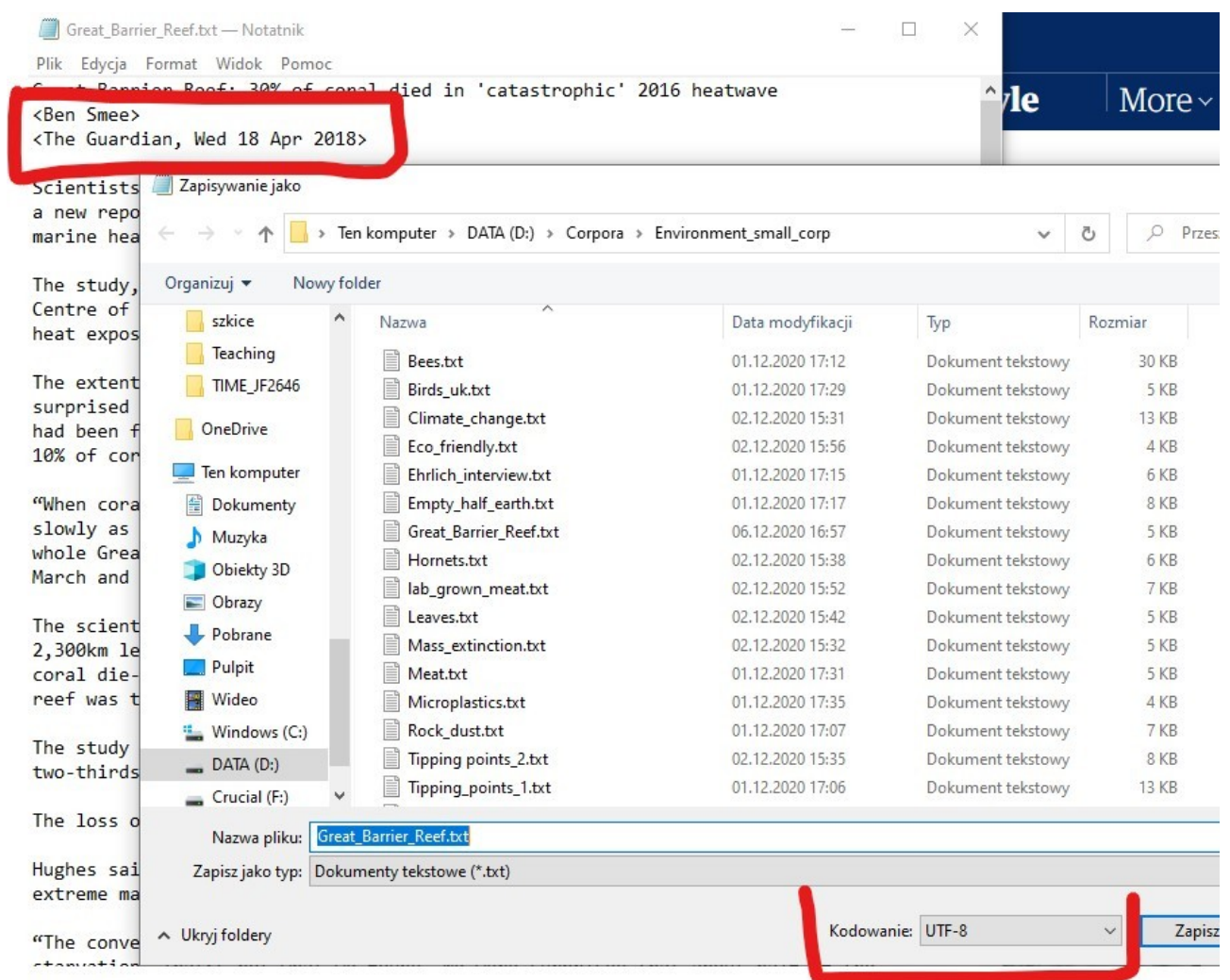
*_N*_er - will give all nouns that end with "er" in its lemma (both *fighter* and *fighters*) it won't always be a suffix.

Discourse analysis – Keywords and framing in a small thematic corpus

1. The first step is to construct a small thematic corpus. You need to find 1-2 web articles on the topic of the group's choice. E.g. environment, relationships, education...

It's enough to skim the articles, no need for reading! The goal is to extract the data into .txt files. This is called scraping, and it's actually more complicated than it sounds. Web pages are filled with lots of additional content you don't need: ads, dynamic headlines, etc. This means you won't always be able to 'select all', 'copy' and 'paste'. You'll have to copy the data selectively.

Scrape and save the texts using notepad, rather than Word. Add metadata in < > brackets (at least the source and date). Save your file using UTF-8 coding.



The screenshot shows a Notepad++ window titled 'Great_Barrier_Reef.txt' with the following content:

```
<Ben Smeee>  
<The Guardian, Wed 18 Apr 2018>
```

The background shows a Windows File Explorer window with the following table of files:

Nazwa	Data modyfikacji	Typ	Rozmiar
szkice			
Teaching			
TIME_JF2646			
OneDrive			
Ten komputer			
Dokumenty			
Muzyka			
Obiekty 3D			
Obrazy			
Pobrane			
Pulpit			
Wideo			
Windows (C:)			
DATA (D:)			
Crucial (F:)			
Bees.txt	01.12.2020 17:12	Dokument tekstowy	30 KB
Birds_uk.txt	01.12.2020 17:29	Dokument tekstowy	5 KB
Climate_change.txt	02.12.2020 15:31	Dokument tekstowy	13 KB
Eco_friendly.txt	02.12.2020 15:56	Dokument tekstowy	4 KB
Ehrlich_interview.txt	01.12.2020 17:15	Dokument tekstowy	6 KB
Empty_half_earth.txt	01.12.2020 17:17	Dokument tekstowy	8 KB
Great_Barrier_Reef.txt	06.12.2020 16:57	Dokument tekstowy	5 KB
Hornets.txt	02.12.2020 15:38	Dokument tekstowy	6 KB
lab_grown_meat.txt	02.12.2020 15:52	Dokument tekstowy	7 KB
Leaves.txt	02.12.2020 15:42	Dokument tekstowy	5 KB
Mass_extinction.txt	02.12.2020 15:32	Dokument tekstowy	5 KB
Meat.txt	01.12.2020 17:31	Dokument tekstowy	5 KB
Microplastics.txt	01.12.2020 17:35	Dokument tekstowy	4 KB
Rock_dust.txt	01.12.2020 17:07	Dokument tekstowy	7 KB
Tipping_points_2.txt	02.12.2020 15:35	Dokument tekstowy	8 KB
Tipping_points_1.txt	01.12.2020 17:06	Dokument tekstowy	13 KB

The File Explorer window also shows the file name 'Great_Barrier_Reef.txt' and the encoding 'UTF-8' in the bottom right corner.

2. Create a list of keywords with the British or American corpus as a reference.

3. Consider which of the keywords are issues that might be subject to metaphorical framing

4. Examine the selected keywords through concordance lines.

References

- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago University Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to the western thought*. Basic Books.
- Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor & Symbol*, 22(1), 1–39.

Further reading

- Charteris-Black, J. (2004). *Corpus Approaches to Critical Metaphor Analysis*. Palgrave Macmillan UK.
- Nacey, S., Dorst, A. G., Krennmayr, T., & Reijnierse, W. G. (Eds.). (2019). *Metaphor Identification in Multiple Languages: MIPVU around the world* (Vol. 22). John Benjamins Publishing Company. <https://doi.org/10.1075/celcr.22>
- Taylor, J. R. (2002). *Cognitive grammar*. Oxford University Press.