LL6033: Using Corpora in Translation studies

Session 1a: Corpora Evaluation

Choose two of the following corpora and explore their characteristics. Here you have a template that will help you evaluate their composition and think about their potential uses.

Name of the corpus:

Number of words:

What type of corpora is it?

Reference/Specialised

Synchronic/Diachronic

Dynamic/Static

Monolingual/Multilingual - Parallel/Comparable

Is the corpus. . . ?

Authentic?

Representative? How is its representativeness achieved?

Balanced?

What could this corpus be useful for?

Examples of corpora:

Catalan

Corpus Textual Informatitzat de la Llengua Catalana: https://ctilc.iec.cat

COVALT Catalan Translation Corpus: http://www.covalt.uji.es/en/

Spanish

Corpus del español: http://www.corpusdelespanol.org/x.asp

Chinese

Chinese Internet Corpus: http://corpus.leeds.ac.uk/query-zh.html

E-C Concord: http://ec-concord.ied.edu.hk/paraconc/index.htm

English Corpus of Contemporary American English: https://corpus.byu.edu/coca/

Collins Wordbanks Online English Corpus: https://collins.co.uk/pages/elt-cobuild-reference-the-collins-corpus

TEC: https://www.alc.manchester.ac.uk/translation-and-intercultural-studies/research/projects/translational-english-corpus-tec/

French Frantext: http://www.cnrtl.fr/corpus/frantext/

Corpus journalistique de l'Est Républicain: http://www.cnrtl.fr/corpus/estrepublicain/

German COSMAS II: http://www.ids-mannheim.de/cosmas2/

NEGRA: http://www.coli.uni-saarland.de/projects/sfb378/negra-corpus/

Italian CORIS/CODIS: http://corpora.dslo.unibo.it/coris_eng.html

BOnonia Legal Corpus - BOLC: http://corpora.dslo.unibo.it/bolc_eng.html

Various languages Corpus Tècnic: http://bwananet.iula.upf.edu

COMPARA: https://www.linguateca.pt/COMPARA/index.php

DCEP: Digital Corpus of the European Parliament: https://ec.europa.eu/jrc/en/language-technologies/dcep

Session 1b: Web as corpus

[Adapted from Gatto 2014]

A. Use the web to search for evidence of attested usage of:

Tall temperature vs high temperature

Now do the same for two different translation possibilities in your other language(s)

B. Devise a complex query to explore the following patterns in academic discourse:

Have a . . . impact on A . . . increase in

What are the most common terms for the blanks? How can you translate the expressions into your other language(s) (and make sure the patterns are equally common)?

C. Find out two regional or diachronic vocabulary differences in your language(s) with Google Trend, similar to the truck/lorry example that we looked at. How can Google Trends be useful for you when translating?

D. Explore the BYU iWeb corpus website: What are the benefits of having a tool like that? What can the different options be useful for?

E. Register for CQP web and explore the Europarl parallel corpora. How can they be useful for the translation of the text that you have been given? How can you combine the use of parallel corpora with the rest of things you have learned today for a fast and accurate translation of the terminology in the text?