

# Corpus linguistics in 8 minutes

Kat Gupta and Lorenzo Mastropierro

[alxkg2@nottingham.ac.uk](mailto:alxkg2@nottingham.ac.uk)

[Lorenzo.mastropierro@nottingham.ac.uk](mailto:Lorenzo.mastropierro@nottingham.ac.uk)

# What is corpus linguistics?

- o A corpus is a large collection of texts

# Differences between a text and a corpus

## Text

- o read whole
- o read horizontally
- o read for content
- o read as a unique event
- o read as an individual act of will
- o instance of *parole*
- o coherent communicative event

## Corpus

- o read fragmented
- o read vertically
- o read for formal patterning
- o read for repeated events
- o read as a sample of social practice
- o give insights into *langue*
- o not a coherent communicative event

N	Concordance
183	this had been a confession of another kind of love. Schomberg cudgelled his brains for a new
184	I never knew mine. I've never known any kind of love. There is something in the mere word.... Of
185	and of others she looked for love, any kind of love, as women will. And that confounded jail
186	to go round the cuddy with this labor of love. To him it had remained a marvel of painting,
187	discovery." "What's that?" "Why--my lack of love for...." "Oh! That's between ourselves," she
188	the degradation of her young ideal of life, of love, of work--all alone in the Treasure House of
189	people, a madness strong as the madness of love, the madness of battle, the desire to spill
190	amazing. To him his life--that cruel mirage of love and peace--seemed as real, as undeniable,
191	, she said, there was death in the mockery of love." Doña Rita moved slightly her beautiful

# What is corpus linguistics?

- o Corpus linguistics involves using computer programs to look for patterns in the texts
  - o Collocation
  - o Colligation
  - o Semantic preference
  - o Discourse prosody

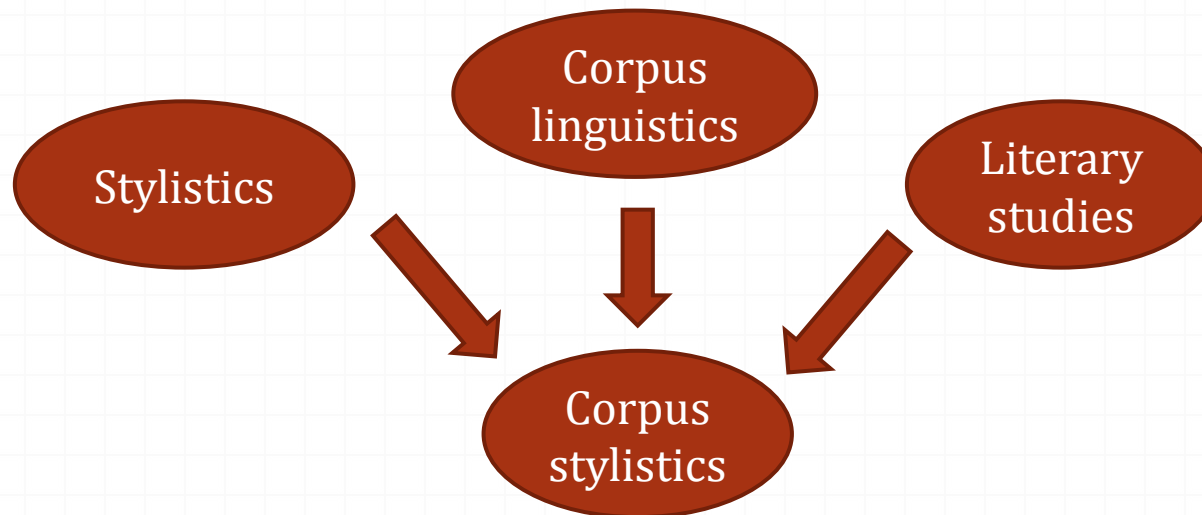
# What is corpus linguistics?

- Corpus linguistics involves using computer programs to look for patterns in the texts
  - Collocation
  - Colligation
  - Semantic preference
  - Discourse prosody
- Implications for grammar, language teaching, applied linguistics, translation, stylistics, historical linguistics, forensic linguistics and social discourses

# Corpus stylistics

What is corpus stylistics?

Corpus stylistics is a computer-aided approach to stylistics which applies methods from corpus linguistics to the study of (literary) style.



# Corpus stylistics

Why using a corpus approach to the study of literary language?

- o Entire text(s) VS text extracts;
- o Comparative approach;
- o Empirical data.

# Corpus stylistics

An example: adjectives and adverbs in Lovecraft's *At the Mountains of Madness*

- o Entire text VS text extract:  
critics' claims VS entire text analysis
- o Comparative approach:  
textual features compared with language norm
- o Empirical data:  
evidence for supporting interpretation



# Using corpora to examine social discourses

- Often used with approaches and frameworks from Critical Discourse Analysis
- Empirical approach
- Offers 'checks and balances'

# Using corpora to examine social discourses

## ○ Advantages:

- Reducing researcher bias
- Examining the incremental effect of discourse
- Examining resistant and changing discourses
- Triangulation

# Resources

## o Tools:

### o Antconc:

<http://www.antlab.sci.waseda.ac.jp/software.html>

### o WordSmith: <http://www.lexically.net/wordsmith>

### o CLAWS POS tagger: <http://ucrel.lancs.ac.uk/claws>

## o Corpora:

### o BNCweb: [bncweb.lancs.ac.uk/bncwebSignup](http://bncweb.lancs.ac.uk/bncwebSignup)

### o BYU corpora: <http://corpus.byu.edu>

### o WebCorp: <http://www.webcorp.org.uk>