How to encode prior knowledge for text processing has two components (obvious alert!):
- Figuring out what that prior knowledge is
- Figuring out how to encode it

These are dependent on each other.
- Parse trees lead to Pitman Yor processes (Wallach et al.)
- Groupings lead to DP clustering (Vlachos et al.)

I’d like to brainstorm other sources of prior knowledge.
Examples of other sources of prior knowledge

- Some ideas for sources of prior knowledge
  - Sentence level:
    *Mathematical Structures of Language* (Harris, 1967)
  - Document level:
    *Style: Toward Clarity and Grace* (Williams, 1995)
  - Corpus level:
    *The Structure of Scientific Revolutions* (Kuhn, 1962)

- What’s different about these levels of prior knowledge?
- How might we encode these?
- In which problems can knowledge at these different levels help?