Who was Zipf?
- Linguist with an emphasis of applying statistics to language
- Saw relation of frequency to length
  - Principle of least effort
What did Piantadosi try to Prove?
- Information content > Frequency
- Statistics between words is more important
How did he show it?
- 11 languages
- Coded information content into words
- Used Google’s Corpus (collection of written words)
- Analyzed with N-Gram test
  Ex: “I love a good DJ, especially DJ E-Trash and DJ Angst”
  2-gram = “I love”, “love a”, “a good”...
  3-gram = “I love a”, “love a good”, “a good DJ”...
  4-gram = “I love a good”, “love a good DJ”, “a good DJ, especially”...
- Equations
  - Accounted for context

Criticisms
- Better Title
- Some languages not included
- Not all showed a significance

Key Points and Takeaways
- Information content > Frequency
  - Semiotics!
- Context Matters
  - Semiotics!
- Zipf wasn’t quite right
  - Challenge Norms

Discussion Questions
- How can we incorporate a stronger emphasis on information’s context into what we do?
- Does all information have a context we should incorporate into what we study? Are there any cases where we should ignore context?