

# LING 5050 – Introduction

Marie-Catherine de Marneffe

Micha Elsner

The Ohio State University

Maysession 2014

# Linguistic questions necessitate data

---

E.g., are lexical subcategorization facts correct?

Some verbs must take objects

*\*Kim devoured.*

Others cannot:

*\*Dana's fist quivered Kim's lip.*

# But ...

---

Atkins and Levin (*IJLexicog* 1995) corpus study:

Transitive usages of intransitive verbs  
(in a 50 million word corpus!)

*The bird sat, quivering its wings*

# Factors of dative construction choice?

---

Verbs of instantaneous force allow dative alternation but not verbs of continuous imparting force like *push*

[Pinker 1981, Krifka 2001]

But ... [Bresnan et al. 2007]

“As Player A **pushed him the chips**, all hell broke loose at the table.”

# Factors of dative construction choice?

---

Verbs of instrument of communication allow dative shift but not verbs of manner of speaking

[Pinker 1981, Levin 1993, Krifka 2001]

“Hi baby.” Wade says as he stretches. You just **mumble him an answer**. You were comfy on that soft leather couch.

# Which factors affect verbal anaphors?

---

“Unexpectedly acceptable”

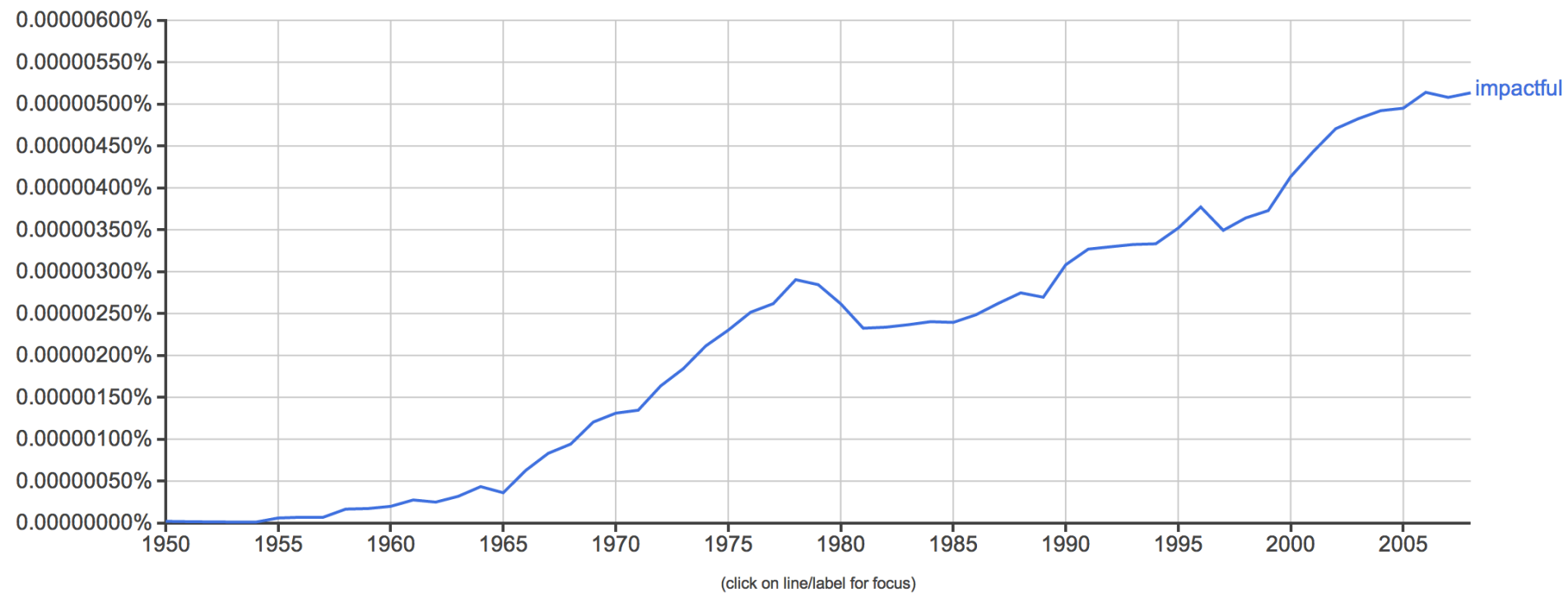
P. Miller at IvanFest, April 2013 (see also P. Miller 2011)

If I represent a wholly different Christianity, I **do so** only because I have found a way to help us recognize as Christians what extraordinary things we say when we worship God. (COCA)

Ungrammatical according to Huddleston and Pullum 2002, Culicover and Jackendoff 2005

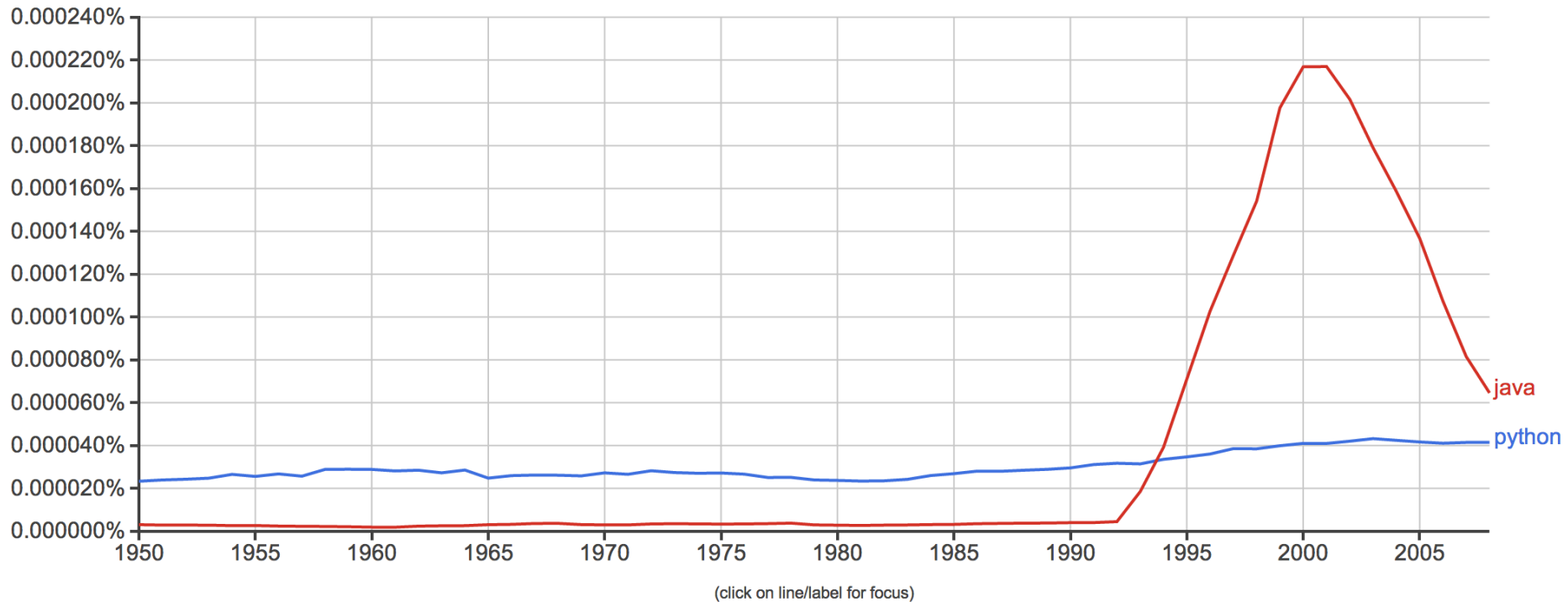
# Today: vast resources!

---



# Today: vast resources!

---







# But we don't want to do this by hand...

---

So we will learn to deal with tools that facilitate data processing!

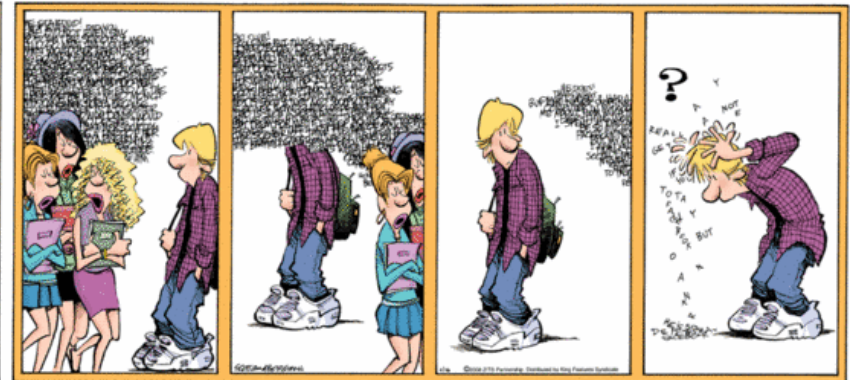
- Unix commands
- Python
- R (for structured datasets) 
- Praat (for audio files) 

First case study:

Do women talk more than men?

---

# First case study: Do women talk more than men?



[Thanks to Jason Grafmiller for the slide and idea]

# According to recent stories in the media, they do!

---

Previous research has shown that women talk almost three times as much as men. In fact, an average woman notches up 20,000 words in a day, which is about 13,000 more than the average man. (*Science World Report 2/20/2013*)

# According to recent stories in the media, they do!

---

Previous research has shown that women talk almost three times as much as men. In fact, an average woman notches up 20,000 words in a day, which is about 13,000 more than the average man. (*Science World Report* 2/20/2013)

New research indicates there's a biological reason why women talk so much more than men: 20,000 words a day spoken by the average woman, according to one study, versus about 7,000 words a day for the average man. (*Today Show* 2/21/2013)

# According to recent stories in the media, they do!

---

Previous research has shown that women talk almost three times as much as men. In fact, an average woman notches up 20,000 words in a day, which is about 13,000 more than the average man. (*Science World Report* 2/20/2013)

New research indicates there's a biological reason why women talk so much more than men: 20,000 words a day spoken by the average woman, according to one study, versus about 7,000 words a day for the average man. (*Today Show* 2/21/2013)

Researchers have claimed on average 20,000 words a day for a woman on average. Just 7,000 for a man. (*ABC News* 2/21/2013)

# According to recent stories in the media, they do!

---

Previous research has shown that women talk almost three times as much as men. In fact, an average woman notches up 20,000 words in a day, which is about 13,000 more than the average man. (*Science World Report* 2/20/2013)

New research indicates there's a biological reason why women talk so much more than men: 20,000 words a day spoken by the average woman, according to one study, versus about 7,000 words a day for the average man. (*Today Show* 2/21/2013)

Researchers have claimed on average 20,000 words a day for a woman on average. Just 7,000 for a man. (*ABC News* 2/21/2013)

Women speak an average of 20,000 words a day vs. 7,000 words for men, according to Louann Brizendine, a practicing physician at the University of California-San Francisco and author of *The Female Brain*. (*USA Today* 2/21/2013)

# How to test this?

---

Spoken conversational corpora seem like ideal testing grounds

- Switchboard, Fisher, CALLHOME, CALLFRIEND, etc.
- All of these have speaker info, including gender

It may not be *exactly* what we need, but it might give us an idea.



# Let's start!

---

- Download the “Fisher” file from Carmen
- We will take a look at the data using Unix commands
- And then start our way into Python