Getting Data

• Kinds of data for linguistics
  - Written
  - Spoken
  - Visual (ASL, body language)
    • Phonetics
      - Implosives-larynx lowering, rounding, x-ray movies
      - Judgments, reaction times, phonetic measurements, fMRI
Getting Data

- Don't reinvent the wheel, use available corpora
  - Linguistic Data Consortium (Audio)
    - Switchboard, CallHome, CallFriend (phone conversations)
    - Santa Barbara Corpus of Spoken English
Getting Data

• Don't reinvent the wheel, use available corpora
  - Linguistic Data Consortium (Audio)
    • Switchboard, CallHome, CallFriend (phone conversations)
    • Santa Barbara Corpus of Spoken English
      - How is /t/ pronounced across words?
        • “He read it aloud.” [t, ?, ɾ]
        • “The port of San Francisco” [t, ?, ɾ]
Getting Data

• Don't reinvent the wheel, use available corpora
  - Linguistic Data Consortium (Audio)
    • Switchboard, CallHome, CallFriend (phone conversations)
    • Santa Barbara Corpus of Spoken English
      - How is /t/ pronounced across words?
        • “He read it aloud.” [t, ?, r]
        • “The port of San Francisco” [t, ?, r]
        • [ɾ] is used more by older speakers and males
        • [ʔ] is used more by younger speakers and females
  - Podcasts
What are audio recordings good for?

- Pronunciation
  - Sociolinguistics
  - Learner speech
  - L2 acquisition
  - Phonetic questions
What are transcriptions of audio recordings good for?

- Vocabulary
- Idioms
- Conversational Analysis
  - What do “um, well, uh huh, mmm, yeah” mean?
  - How do you end your turn talking? Know the other person is done?
- Regionalisms
What are transcripts of audio recordings good for?

- Available transcripts
  - CNN
  - NPR
  - Movie Scripts
Getting your own recordings

- Practical considerations:
  - Audio only or audio and video?
  - Format for digital recorder?
  - How much time do you need?
  - External microphone
  - Making copies
  - Identifying speakers
  - Quiet place, no kids running around, traffic
Problems with recordings

- You may not get enough instances of what you are looking for
- Technical difficulties
- Very time consuming to go through and find things/transcribe
Ethical Questions

I. R. B.
Ethical Questions

I. R. B.
Institutional Review Board
Ethical Questions

I. R. B.
Institutional Review Board
YOU MUST GET THIS IF YOU ARE WORKING WITH PEOPLE
IRB

- BYU wants to protect people in studies
- BYU want to protect students doing studies
- BYU doesn't want its name associated with iffy, marginal, questionable studies
BYU wants to protect people in studies
BYU want to protect students doing studies
BYU doesn't want its name associated with iffy, marginal, questionable studies
BYU does not want to get sued
Why IRB?
Why IRB?
Ethical Questions

- Can you secretly record people?
Ethical Questions

- Can you secretly record people?
- Can you use broadcast information without permission of speakers?
Ethical Questions

- Can you secretly record people?
- Can you use broadcast information without permission of speakers?
- Can you use data from corpora like LDC without permission of speakers?
Ethical Questions

- Can you secretly record people?
- Can you use broadcast information without permission of speakers?
  - Oprah Winfrey and use of [a] vs. [aj]
  - The Queen's English
  - “Say yes to the dress”
- Can you use data from corpora like LDC without permission of speakers?
Ethical Questions

• Observation studies (non-linguistic)
  – Reporter joins Jerry Falwell's group, goes on mission trip, write book about experiences
  – Woman goes back to school at UNA, observes coed roommates, writes book about them, and her experiences (Cathy Small)
Observations

• Recordings are observations
  - "The collection of data without manipulating it."
  - "Simply observe ongoing activities, without making any attempt to control or determine them."

• People know they are being recorded
Observer's Paradox

- We want speech in natural, casual, unfiltered format.
Observer's Paradox

• We want speech in natural, casual, unfiltered format.
Observer's Paradox

- When people know they are observed they change their speech to more careful, standard usages.
Other kinds of observation

- **L1 acquisition**
  - Diaries
  - Wired house
- **Labov's department store**
- **Labov-French or English in Montreal**
  - Age, gender, place, topic of conversation
- **Give kids puppets to play with**
  - Girls use cooperative language, boys aggressive
Other kinds of observation

• Bring your friend to experiment
  – Experimenter called out, camera keeps filming
  – What to men and women talk about?

• Observe slips of the tongue
Problems with corpora, recordings, observations

- You may find few cases of what you are studying
  - Might could
  - She really nice lady
- What is age, origin, ethnicity, educational level of person?
Case Studies

- Study one person, or small group over long period of time.
  - L1 diaries
  - Aphasia patients, speech therapy
  - Investigator observes class over school year
    - Teachers interact less with minority students
  - Feral children (Genie)
Case Studies

• Problems
  - Generalizability
  - Subject retention
  - Researcher loses interest