Linguistic Data Consortium

# Linguistic Fieldwork and IRB Human Subjects Protocols: The LDC Experience

Denise DiPersio

dipersio@ldc.upenn.edu

# IRB History and Practice

- General principles for human subjects research
  - Respect for persons: autonomy, consent, truthfulness
  - Beneficence: do no harm, maximum research goals
  - Justice: fair, non-exploitative procedures
- Common Rule concerns
  - Will the study require the participation of vulnerable populations?
  - How will informed consent be obtained?
  - How will confidentiality be maintained?
- Social sciences research
  - IRB reviews geared for medical research
  - Lack of uniformity
  - Miscommunications

# LDC's Protocol

- In place for almost 20 years with University of Pennsylvania's IRB
- Covers speech, text, handwriting, language-related judgments
  - On-site at LDC, in the field, crowdsourcing
- Data collected distributed as corpora to support language research, education and technology development
- Umbrella protocol modified as needed to add new studies, approve new/revised consent forms, modify existing studies
- Largely successful
  - Challenges: new collection methods/technology, timing, increased interest/attention to social science research methods

# Protocol Features

- Record linguistic performance
  - Speech, writing, typing, dictation
  - In person, via phone, computer-mediated device, writing surface, no human/machine interlocutor
  - Optionally with headset transmitting silence/noise
- Collect judgments about linguistic behavior and decisions involving linguistic data
  - Auditing speech recordings
  - Judging handwriting legibility
  - Summarizing written text, reading comprehension
- Collect linguistic performance
  - Gaze tracking, strokes/minute

# Fieldwork Methods and Procedures, 1/2

- Non-remote field locations (Philadelphia, Seoul)
  - Speech recorded to digital recorders/computers; copied to LDC database as soon as practicable
- Remote field locations (Papua New Guinea)
  - Bilingual native speakers record participant speech to digital recorders
  - Uploaded to laptop; backed up on mass storage device; uploaded to LDC following each field trip
- Personal identifying information
  - Logbooks → spreadsheet →mass storage device → LDC
- Data
  - Secure storage; encrypted spreadsheet; fieldworker control
  - LDC: secured network, locked file cabinets

# Fieldwork Methods and Procedures, 2/2

- Consent
  - Written consent; informed consent form
  - Verbal consent, recorded (unwritten languages, speakers not literate in native language(s))
  - Consent through action (pushing button for telephone study; performing crowdsourcing task)
- Accommodations for IRB
  - Examples of questions that will be asked
  - “Script” for verbal consent

# Conclusions

- If “the way to do fieldwork is never to come up for air until it is all over” (Margaret Mead), getting the protocol is simple by comparison.
- Preparation – be able to articulate your plan
- Relationships – department, IRB
  - Fieldwork is consistent with federal guidelines
- Crib from/use available resources
  - Be sensitive to IRB independence