Legal, ethical and policy issues in recording and publishing primary linguistic material

Mark Liberman
University of Pennsylvania

Overview

- Relevant laws & regulations for PLM (U.S. University Version)
  - Recording
  - Copyright
  - Defamation
  - Human Subjects/IRB
  - Ethical Issues

Caveats: IANAL, YMMV

It's possible

- Since 1992, LDC has published spoken conversations involving
  - > 8,800 speakers in 16 languages
  - a lot of other language documentation
  - with
    - no legal problems
    - full IRB approvals (for LDC collections)
    - no ethical complaints or known problems

General advice

- Become informed
  - Background for this talk
  - Other web and print resources
- Focus on key issues
  - IRB review of “experimental protocols”
  - Relations with speakers/speech communities
  - Relations with data recipients
- Plan ahead
- Don’t let difficulties become excuses

Recording

- Voice recording: legal if all parties consent
  - Knowledge implies consent
  - view of device
  - announcement
  - 38 states & DC: only one party need consent
- Video-only recording: usually legal
  - Some new state “expectation of privacy” laws
- No U.S. legal issues here for linguists
  - since ethics require consent of all parties

Copyright

- Copyright law is complex and interesting
  - unexpected concepts and distinctions
  - multiple jurisdictions, subtypes of rights
  - tangled history, special interests
  - critical public policy debate now
    - see Lessig, "The Future of Ideas"
- Spend a day reading about it!
- PLM copyright: a simple recipe usually works
  - 2 sides: “author”/“publisher”, “publisher”/“user”
  - licenses structured to:
    - protect author, publisher;
    - control downstream uses
defamation

- [left out to save time]
- [but watch what you(r speakers) say!]

human subjects/IRB

- 45 CFR 46: all Federally-supported research involving human subjects must be reviewed and approved by an Institutional Review Board (IRB) at the responsible institution.
- Most universities: all human-subjects research must be reviewed, even if not government sponsored.
- Many types of linguistic research are ‘exempt’ — but exemption must be determined by IRB, not by researcher.
- Much non-exempt linguistic research is eligible for ‘expedited review’
- Review must precede data collection
- If not, regulations may require data to be destroyed (even if exempt!)
- Review is for permission to collect, not permission to publish!
- SO LEARN TO DEAL WITH YOUR IRB!

ethics

- Legal behavior may be unethical (and ethical behavior may be illegal)
  - potential issues are very diverse
  - so use sensitivity, empathy and common sense
- Politics of the speech community
  - linguistic documentation usually welcome
  - tendency to focus on (less common) problems!
  - special situations
    - language as local industry/export commodity
    - religious or cultural sensitivities
    - prescriptivist dislike of vernacular
    - nationalist dislike of minority languages

conclusions

- A stitch in time saves nine
  - Work out legal and regulatory issues before starting a project
  - Think about ethics and politics before, during and after
  - Teach students how to do this too
- IRB oversight looms
  - good in principle, can be a disaster in practice
  - LSA should get involved in national debate