Evolving Questions for Data-Coding Conventions

Christopher Cieri & Malcah Yaeger-Dror

ccieri@ldc.upenn.edu & malcah@email.arizona.edu

Intro to NSF-LDC Satellite
Three Coding Foci

Today’s main foci –

- Ethics
- Demographic coding: has been upgraded constantly to reflect linguists’ needs, but there are recurring problems where we under-differentiate.
- Attitude coding: towards one’s own and others’ groups

Tomorrow’s main focus

- Social Situation—
- interaction among speakers’ demographics
- Drawing some conclusions
Given

- that the questions we ask set the IV mood
- that insufficient information precludes sharing
- that often there are difficulties getting permission

A few do’s and don’ts of ethical review board interaction

- Natasha Warner—LSA’s ethics panel
- Denise DiPersio—LDC’s ethics proposal writer
Coding: Labov’s Rule #1

- Use more distinctions than you need!
- It is easier to merge groups later
- Than to recode from scratch.
- You can’t put back what you didn’t code…
- You can’t compare if the features aren’t there to compare – “Pollock/Hinton effect”
- So, even if there were little evidence for the importance of X in Y community, but more evidence for its importance in Z community: do
Three Coding Foci

- Today’s main focus –
- Demographic coding: has been upgraded constantly to reflect linguists’ needs, but there are recurring problems where we under-differentiate.
- *EMPHASIZING*
- The speaker’s expertise
- Research design adapted to permit sharing.
Demographics

- Region
- Sex (M/F)
- Racial or Heritage group(s)
- Years of education
- Age
Demographics

- **Socioeconomic background/place within the society** \{SES ML...\}
- **Sexuality** -- Eckert
- **Racial or Heritage group(s)**
  - \{how far back? How mixed? How strongly identified?\}
    - **Black** -- Blake
    - **Hispanic** -- Fought
    - **Asian** -- Wong/Hall-Lew
- **Age** -- Bowie
- **Religious affiliation** -- Bowie
Egypt: Copts, vs. Brotherhood
Religion \( \alpha \) Language

- The province of historical linguists or social psychologists
- Blanc’s work in the 60’s
  - Bagdad – entirely based on religion, with sociophonetic distinctions.
- Milroy’s work in the 70’s Belfast
- Giles’ early work in the 70’s
  - Belgian [French/Catholic] vs. Flemish [Dutch/Protestant]
  - Quebecois [French/Catholic] vs. Anglophone [Protestant]

More recent work on ‘language and religion’ in the Near East
  - /q, r, …/ varies with religion/ or degree of religiosity.

More recent work on ‘degree of religiosity’ in the US


- **Today’s main foci –**
  - Demographic coding
  - **Attitude coding:** towards one’s own and others’ groups
    - Llamas – North of England
    - Nagy, Hoffman – Toronto
    - Noels—conceptual
    - Poplack--Sociolinguistic importance of attitudes

- **Tomorrow’s main focus**
  - **Social Situation--Attitude Coding:** attitudes towards one’s own and others’ group
Social Situation

- Tagliamonte
- Rickford
- The interlocutors, and the relationship between them
- The overhearers – ratified and not. (Bell)
- Taking advantage of ‘serendipitous situational switching’
- Place
- Llamas
- Interaction of demographics among interlocutors
- Interaction of degree of commitment to each other or outsiders (Bell)

Discussion of how to archive factors

- Interaction between interlocutors’ demographics
- Attitudes towards one’s own and interlocutors’ demographics
Whatever groups we have they should all be represented. We want to suggest ways in which we can accommodate our coding conventions to new insights about 

Speaker demographics 

Speakers’ attitudes 

the influence of social attitudes on self-identification. 

As well as our need to attend to variability in self-identification perspectives, even when interacting with a single interlocutor. 

These talks focus on possible ways to determine these factors, And how best to unify our coding protocols to permit us to share data among ourselves.
Thanks to NSF & LDC for funding