Sociolinguistics and U.S. Census Bureau Language Research

Yuling Pan

U.S. Census Bureau

Interagency Language Roundtable Meeting
September 10, 2010
Roadmap of Today’s Talk

- Background
  - Census Bureau and language use in the U.S.
  - Census Bureau language-related research and programs
- New directions in language research
- Issues and challenges
- Three on-going projects
- Looking into the future
U.S. Census Bureau

One of the world’s largest and most comprehensive data-gathering organizations

**Mission:**
- To collect and provide timely, relevant, and quality data about the people and economy of the United States

**Conducts:**
- Population Census (taken every 10 years)
- Economic Census (taken every 5 years)
- Over 100 sample surveys

Provides information that mirrors the concerns of the country
Need for Language Research

- Increasing level of immigration and growing diversity in the U.S. population
- Respondents’ inability to speak or read English inhibits accurate data collection and affects data quality
- Research is needed on:
  - How to overcome language and cultural barriers in data collection
  - How to encourage diverse populations to participate in the Census and other surveys
American Community Survey data (Shin & Kominski, 2010)

- 20% of U.S. population age 5 and over spoke a language other than English at home
- 44% of them did not speak English very well
- About 9% of U.S. population age 5 and over did not speak English very well

- Spanish, Chinese, Russian and Vietnamese speakers
  - Higher proportion speaking English less than very well
American Community Survey (ACS)

- A monthly, nationally-representative household survey
- Collects information on household characteristics
- Replacement of the decennial Census long form
- Three sequential modes of data collection:
  1) Mail, 2) Phone, 3) Personal Interview
- Sample size:
  - 250,000 households per month
  - 3 million households annually
Language Question on the ACS

14. Does this person speak a language other than English at home?
   □ Yes
   □ No → SKIP to question 15a

b. What is this language?

   For example: Korean, Italian, Spanish, Vietnamese

c. How well does this person speak English?
   □ Very well
   □ Well
   □ Not well
   □ Not at all
Census Bureau
Language-related Research

- Identification of languages to translate the Census questionnaire and Census Bureau surveys
- Identification of number of languages to provide language assistance
- Areas to target mailing and partnership programs
- Focus on translation
Census Bureau Language Programs

- **2010 Decennial Census**
  - Census questionnaire in five languages
    - Spanish (English-Spanish bilingual form), Chinese, Korean, Russian, and Vietnamese
  - Language assistance guides in 59 languages
  - Advance letters and online materials in multiple languages

- **The American Community Survey (ACS)**
  - ACS questionnaire in three languages
    - Spanish, Chinese, and Korean
  - ACS survey letters and brochures in 10 languages
    - Spanish, Chinese, Korean, Russian, Vietnamese, Arabic, French, Haitian Creole, Polish, and Portuguese
New Directions in Language Research

Motivation
- Language barrier does not reside in language itself
- Need for a sociolinguistic paradigm

Interaction between language, culture, and society
- Functional or pragmatic approach to translation
- Strategies in reaching out to non-English speaking populations
Three Specific Research Projects

- 2010 Census form in five languages
- Chinese communication style and implications for their survey behavior
- Observing census enumeration of non-English speaking households in the 2010 U.S. Census
Challenges in Survey Translation

“Asks the same questions and offers the same response options” (Harkness & Schoua-Glusberg, 1998, p. 92)

Previous studies focus on word meaning and complex mappings between words in different languages
Traditional Translation Process

Translation → Review → Revision → Finalizing

Conducted by bilinguals (translators and reviewers)

Usually resolve translation issues at the **lexical and syntactic** levels

Issues at the **pragmatic** level are more subtle and difficult to overcome

These issues will affect data quality and will hinder survey participation
Census Bureau Translation Guidelines

- Census Bureau developed its translation guidelines (Pan and de la Puente 2005)
- Committee approach to translation
  - Translators, reviewers, subject matter experts, survey methodologists
- Pretesting of translations with target populations
  - Cognitive interviewing with monolingual speakers
  - A team (3-4) of language and cultural experts in each target language
  - Revision based on evidence from pretesting
- Sociolinguistic perspective of survey translation
The Sociolinguistics of Survey Translation

Linguistic forms

Cultural Norms

Social Practices
Coding Scheme for Translation Problems (Pan and Fond, 2010)

• 3 categories of codes:
  - UE: User Error
  - TE: Translation Error
  - LF: Linguistic Forms
  - CN: Cultural Norms
  - SP: Social Practices

• 2 additional categories:
  - UE: User Error
  - TE: Translation Error
Category 1: Linguistic Forms

- **Linguistic Forms**: classifies issues in the translation that are due to
  - Vocabulary
  - Grammar
  - Usage conventions, etc.
Category 2: Cultural Norms

- **Cultural Norms**: classifies issues in which concepts expressed one way in English are expressed differently in the target language
  - Address conventions
  - Numbering
  - Kinship terms
  - Conversational norms, etc.
Category 3: Social Practices

- **Social Practices**: classifies concepts that can be described in English but cannot be translated into the target language because
  - The concept does not exist in that culture
  - Respondents have no experience with the concept
Additional Categories

• 2 additional categories:
  • **User Errors**: actions by respondents that are easily corrected
    • Mistakes
    • Distractions, etc.
  • **Translation Errors**: translation problems that are easily corrected
    • Omissions
    • Typos, etc.
2010 Census Questionnaire in Five Languages

- 2010 Census questionnaire
  - 10 demographic questions and instructions
  - Self-administered questionnaire

- Languages
  - Chinese, Korean, Russian, Vietnamese, and English

- 112 cognitive interviews with monolingual speakers of the 5 languages

- To evaluate the translations based on respondents’ reaction and interpretation
2010 Census Questionnaire – Results

Types of issues in all four languages

<table>
<thead>
<tr>
<th>TOTALS</th>
<th>LF</th>
<th>CN</th>
<th>SP</th>
<th>TE</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>53</td>
<td>63</td>
<td>30</td>
<td>17</td>
<td>163</td>
</tr>
</tbody>
</table>

- LF: 33%
- CN: 39%
- SP: 18%
- TE: 10%

TOTAL: 163
Issues with Linguistic Forms

- Word-for-word translation
- Complex sentence structure
- Difficult wording
- Phonetic translation
- Not natural in target languages
Example of Cultural Norms: 
Korean Age Counting

7. What is Person 1’s age and what is Person 1’s date of birth? 
Please report babies as age 0 when the child is less than 1 year old. 
Print numbers in boxes.

Age on April 1, 2010  Month  Day  Year of birth

• Issues:
  • More than two thirds of the Korean respondents had difficulty writing their age
  • Korean’s age counting convention is different
• Revised version
  • Clarify that this question is asking for the American way of counting age
Examples of Social Practices

10. Does Person 1 sometimes live or stay somewhere else?
   - No
   - Yes — Mark \( \times \) all that apply.
   - In college housing
   - In the military
   - At a seasonal or second residence
   - For child custody
   - In jail or prison
   - In a nursing home
   - For another reason

2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1? Mark \( \times \) all that apply.
   - Children, such as newborn babies or foster children

3. Is this house, apartment, or mobile home — Mark \( \times \) ONE box.

- “Nursing home”
  - Chinese: mental hospital, recreational resort
  - Korean: resting place
  - Russian: medical establishment
  - Vietnamese: luxurious resort

- “Foster children”
  - Children under the temporary care of relatives or friends

- “Mobile home”
  - No clue at all
2010 Census Form – Results

Issues as a % of total (per language)

- Chinese
- Korean
- Vietnamese
- Russian

Legend:
- LF
- CN
- SP
- TE

U S C E N S U S B U R E A U
Communication Styles of Chinese Speakers: Implications on their Survey Response Behavior

Anna Yukyee Chan, Ph.D. 陳玉儀
Yuling Pan, Ph.D. 潘予翎
U.S. Census Bureau

Presentation for the
Annual Joint Statistical Meetings
July 31 to August 5th, 2010
Vancouver, Canada
Towards an Understanding on Response Behavior of Chinese Speakers

Cultural Norms & Schemata
- Harmony Oriented
- Face Concern
- Others Oriented
- Hierarchical

Politeness - Not deception
(Pan 2000, Pan & Kádár 2010)

Communication Style
- Indirect
- Contrary to Face Value (CFV)
(Ma, 1996)
Data

- Cognitive interviews with monolingual Chinese and English
  - Sites: Washington D.C. Metro, Chicago, IL & NC areas
- Non-random Sample: selected 79 cases
  - 46 Monolingual Chinese speakers
  - 33 with native English speakers
Personal Opinion Question

Survey-like question asking about respondents’ intention to participate in the American Community Survey.

Question:
“If you were selected, would you participate in the ACS?”

Response:

- Yes
- No
- Maybe

U S C E N S U S B U R E A U
Chart 1. Response Distribution: Will you participate in the ACS?

Face value responses

Significant $\chi^2=5.91$, $p<.05$
Chart 2. Overall Indirect Responses

Significant: $\chi^2=16.6, \ p<.001$
Chart 3. Contrary to Face Value Responses for Chinese

% CFV responses

Overall CFV

"Maybe is No"

"Yes is Maybe or No"

"No is Yes"

Face value and true intention
Conclusions

- Need to recognize the cultural differences in communication and response styles among different language speakers.
- Indirect responses to survey questions may pose a threat to the validity of the constructs being measured by the questions.
- Need to train interviewers on how to elicit and obtain appropriate and relevant responses from these speakers.
Observing Census Enumeration of Non-English-Speaking Households
Study Design

- Ethnographic study of observing Census interviews in the field
- Multilingual research teams in multiple sites
  - Spanish, Chinese, Korean, Russian, and Vietnamese, Arabic, Portuguese
  - English as a comparison group
- Research sites:
  - California, Illinois, Ohio, Michigan, Massachusetts, New York City, Greater Washington DC area
Research Goals

- To conduct systematic observation of Census interviews
- To investigate the challenges in field interviews with speakers of languages other than English
- To examine the interaction of language and sociocultural factors in this communicative event
Methodology

- A team of three bilingual ethnographers in each target language
- Each ethnographer observed census interviews in target locations for two weeks
- Each team observed 70 – 170 interviews
- Debriefing with respondents and enumerators
- Taking notes and taping interviews
- Analyze field notes and transcripts of census interviews and debriefings
Preliminary Results (1)

- Unfamiliarity with the purpose of U.S. Census
- No equivalent practice of Census taking in home country
- Fear/distrust of government
- Language barriers
  - No translation of census questions
  - Language variety and dialects
  - Literacy issue
Social and cultural barriers
- Gender
- Ethnicity
- Identity
- Immigration status
- More challenging!
Census Bureau language research:

- Departing from a heavy focus on translation
- Adopting a sociolinguistic perspective
- Investigating the interaction of language, culture, and society and its impact on data collection and data quality
- Utilizing research methods from multiple disciplines
Future Research

- Effective messages to diverse populations
- Discourse approach to translation of survey letters
- Appropriate questionnaires for the multicultural/multi-language environment
- Guidelines for the use of interpreters in field interviews
- Language proficiency assessment tool
References


Contact Information

Thank you!

Email: yuling.pan@census.gov