Avoiding Prescriptive Influence: Collaborative Fieldwork with Native Speakers

Helen Jeoung (Univ. of Pennsylvania)

Overview

- Collaborative Fieldwork with native speakers in East Java, Indonesia

- Goals of fieldwork:
  - Refining descriptions of ‘Object Voice’ in colloquial Indonesian
  - Avoiding prescriptive influence in data collection
  - Engagement of native speakers in linguistic fieldwork

- Outline:
  - Theoretical background: Object Voice in Austronesian languages
  - Research Challenges in obtaining data and judgments
  - Fieldwork Methodology
  - Results
  - Conclusions
Background: Object Voice

- In addition to Active voice and Passive voice, Indonesian and other related languages have a third voice called Object Voice.

- Object Voice word order in Indonesian: Patient–(Asp)–Agent–Verb

(1) Buku ini sudah saya baca.
   book this Perf 1sg read
   ‘I read this book.’

(2) Anak-anak nakal itu akan kamu tegur.
   child-Red naughty that will 2sg scold
   ‘You will scold those naughty children.’

In Object Voice, 1 and 2 pronouns are reported to be the only possible Agents (Chung 1976, Sneddon 1996).

(3) Buku ini sudah aku/ kita baca. ‘I/we read this book.’
    Buku ini sudah kamu baca. ‘You read this book.’

In contrast, 3 person pronouns, names and other nominals are reported to be ungrammatical.

(4) * Buku ini sudah dia baca. ‘S/he read this book.’
    * Buku ini sudah Siti baca. ‘Siti read this book.’
    * Buku ini sudah ibuku baca. ‘My mother read this book.’
    * Buku ini sudah anak-anak itu baca. ‘The children read this book.’
    * Buku ini sudah ribuan orang baca. ‘Thousands of people read this book.’
Background: Object Voice

- However, the restriction on the Object Voice Agent has been claimed to be a *prescriptive* rule only (Nomoto 2006, Sneddon 2006).

- Anecdotally, other Object Voice Agents do occur in natural speech. The class of possible Agents is unknown; the degree of variation between speakers is also unknown.

- Corpora:
  - Few corpora available/annotated for colloquial Indonesian
  - Object Voice is infrequent
  - Object Voice is ambiguous with topicalized Patient + no Aspect morpheme + bare Active verb
Goals of research

1. Which nominals occur as Object Voice Agents in colloquial Indonesian?
   - Acceptability judgments gathered from a larger number of speakers
   - Judgments in the context of informal, spoken Indonesian

2. Is there evidence of generational language change?
   - Difference in older speakers vs. younger speakers
East Java, Indonesia
A Collaborative Model

The research team:
5 native Indonesian speakers (‘Interns’)
   ○ University students (from UNISMA and Ma Chung Univ.)
   ○ Hired for 8-week research internship
1 Team leader / primary investigator
A Collaborative Model

- **Weeks 1-3: Training period**
  - Research principles, linguistic elicitation, interviewing techniques, sociolinguistic questionnaires, data compilation
  - All research materials collaboratively developed and checked
  - 4-5 Practice interviews by each intern, observed by team members
  - Standardization of techniques and judgments

- **Weeks 4-6: Collecting data**
  - Interns recruited subjects from campus, family, neighbors and social networks, and independently gathered data

- **Weeks 7-8: Data compilation, quantitative analysis, reporting**
A Collaborative Model

- Methodology
  - Each subject interviewed individually
  - Interview conducted completely in ‘everyday’ speech, in local variety of Indonesian
  - Emphasis on the subject’s input, not ‘right’/‘wrong’ answers
  - Subjects did not have contact with Team leader
  - Context provided for the sentences used (e.g. Imagine your sister was telling you a story and said...)
  - Interns read each sentence aloud, while subjects followed on printed copy
  - Sociolinguistic questionnaire (demographic information) was collected at the end of interview
  - Subjects were given small ‘prize’ for participation
Results

- 133 adult subjects were interviewed during 3 weeks
  - Judgments on Object Voice Agents
  - (Judgments on circumfix *ke- -an*)
  - (29 Audio recordings of narratives)
  - Sociolinguistic (demographic) data for each subject

- For Object Voice, 7 different sentence frames were used, with 19 different possible agents

- Judgments were recorded as:
  - yes / fine / no problem
  - no / not possible / can’t understand
  - ? / not sure
Results: Overall acceptability

(N=133)
Results: Overall acceptability

(N=133)

95% 92% 93% 93% 95% 92%

1 sg  2 sg  3 sg  1 pl  2 pl  3 pl  Name (simple) Title + Name Quantifier + N Indef. N Generic N N + relative WH-word
Results: Overall acceptability

(N=133)
Results: Overall acceptability

(N=133)
## Results

<table>
<thead>
<tr>
<th>Previously reported</th>
<th>Our findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ 1, 2 pronouns</td>
<td>✓ 1, 2 pronouns</td>
</tr>
<tr>
<td>X 3 pronouns</td>
<td>✓ 3 pronouns</td>
</tr>
<tr>
<td>X Names</td>
<td>✓ Names</td>
</tr>
<tr>
<td>X Other nominals</td>
<td>(variable) Indefinite N</td>
</tr>
<tr>
<td></td>
<td>(variable) Generic N</td>
</tr>
<tr>
<td></td>
<td>(variable) N with embedded relative</td>
</tr>
<tr>
<td></td>
<td>(variable) WH word</td>
</tr>
</tbody>
</table>
Subjects by Age Cohort

- < 18 (DOB after 1997): 9%
- 18-25 (DOB 1990-1997): 36%
- 26-49 (DOB 1966-1989): 32%
- > 50 (DOB before 1965): 23%
Acceptability by Age Cohort

(for nominals with overall acceptability <70%)

- Indefinite N 'a doctor'
- Generic N 'people'
- N with embedded relative 'that person who is smart'
- WH word 'who'

Cohort:
- <18
- 18-25
- 26-49
- >50
Subjects by Education level

- University: 39%
- Junior HS: 17%
- Senior HS: 33%
- Primary: 11%
Acceptability by Education Level
(for nominals with overall acceptability <70%)

Indefinite N 'a doctor'
Generic N 'people'
N with embedded relative 'that person who is smart'
WH word 'who'

Primary
Junior HS
Senior HS
University
Implications

- No evidence of ongoing or generational language change (between age cohorts)

- No evidence that age or education level is a statistical predictor

- The fact that the oldest speakers accept a variety of Agents in Object Voice indicates that the restriction had a prescriptive flavor even when the language was being standardized after Indonesian independence (1945 through 1960s)
Object Voice has a number of unusual characteristics:
- The restriction on the Agent is often cited as a selectional property of \( v \) (or Voice), which selects for 1 and 2 pronouns only.

However, when prescriptive influences are minimized during data collection, we find:
- 3 pronouns and names belong to the class of Agents;
- The restriction is not categorical but **variable** (between speakers); and
- Judgments may be **gradient** (for an individual speaker).
Conclusions – The Collaborative Model of Fieldwork

- **Advantages**
  - Large amount of data provided by many speakers, in a limited timeframe
  - Meaningful involvement of native speakers (cf. Austin 2014)
  - Improvement in research tools, based on collaboration with native speakers (cf. Thomason 1994)
  - Linguistic training, compensation, research experience and future opportunities for speakers of the language
  - All research conducted in the language
  - Prescriptive and ‘outsider’ influence minimized

- **Disadvantages**: Extra time and funds required for training, mentoring and ensuring consistency
References (1)

- Chung, Sandra. 2008. How much can understudied languages really tell us about how language works? Plenary Lecture given at the 2008 Annual meeting of LSA.
References (2)

Thanks...

- Thank you to the research team—terimakasih banyak kepada:
  - Gabriella Melody, Ma Chung University
  - Iline Megale, Ma Chung University
  - Nadhilah H. Semendawai, Ma Chung University
  - M. Khoirun Najib, Islamic Univ. of Malang (UNISMA)
  - Ilkham Hidayat, Islamic Univ. of Malang (UNISMA)

- Thanks to the English Letters Dept. at Ma Chung Univ. and the Linguistics Dept. at Univ. of Pennsylvania for support

- Funding granted by the GAPSA-Provost Award for Interdisciplinary Innovation 2015, Univ. of Pennsylvania.
Subjects interviewed N=133

Region where speaker was raised (up to age 15)

- East Java
- Other Java
- Sumatra
- Kalimantan/Maluku/Sulawesi
- Papua
Kalimat Pasif Semu

1. Buku ini sudah saya baca.
   Buku ini sudah aku baca.
   Buku ini sudah kubaca.
   Buku ini sudah kamu baca.
   Buku ini sudah kau baca.
   Buku ini sudah dia baca.
   Buku ini sudah kalian baca.
   Buku ini sudah kita baca.
   Buku ini sudah kami baca.
   Buku ini sudah mereka baca.

2. Buku ini sudah Siti baca.
   Buku ini sudah Pak Agung baca.
   Buku ini sudah orang baca.
   Buku ini sudah orang itu baca.
   Buku ini sudah kamu dan kamu baca.
   Buku ini sudah Siti dan Rini baca.
   Buku ini sudah ribuan orang baca.
   Buku ini sudah orang yang pintar itu baca.

3. Buku ini sudah siapa baca?
   Apakah buku ini sudah kamu baca?

4. Surat itu harus saya tulis.
   Surat itu harus aku tulis.
   Surat itu harus kutilulis.