Language Differences between Interviewers and Respondents in Sub-Saharan African Surveys

3MC Conference Presentation

Charles Q. Lau
Survey Methodologist
Acknowledgements

Co-authors (alphabetical)

- Melissa Baker
- Stephanie Eckman
- Ben Piper
Languages in Africa

Many countries are multi-lingual

Number of Languages Spoken

- Nigeria: 527
- Tanzania: 127
- Kenya: 68
- Senegal: 38
- South Africa: 34

Percent of Adults that Speak Three or More Languages

- Nigeria: 31%
- Tanzania: 23%
- Kenya: 61%
- Senegal: 26%
- South Africa: 30%

Source: Ethnologue (https://www.ethnologue.com/)

Source: Author analysis of Afrobarometer Wave 3 data
# Three Types of Languages in Africa

<table>
<thead>
<tr>
<th>Mother tongue</th>
<th>Local language of broader communication</th>
<th>International language of broader communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Language of local tribe or community</td>
<td>• Lingua franca</td>
<td>• Lingua franca</td>
</tr>
<tr>
<td></td>
<td>• African language</td>
<td>• Can be official language, even if population has limited proficiency</td>
</tr>
<tr>
<td></td>
<td>• Example: Swahili in Kenya</td>
<td>• Example: English in Nigeria</td>
</tr>
</tbody>
</table>
Three Types of Languages in Africa

- **Mother tongue**: Language of local tribe or community

- **Local language of broader communication**:
  - Lingua franca
  - African language
  - Example: Swahili in Kenya

- **International language of broader communication**:
  - Lingua franca
  - Can be official language, even if population has limited proficiency
  - Example: English in Nigeria

Issues with Using Mother Tongue in a Survey

1. Disconnect between written and spoken languages: Some may prefer to speak in mother tongue but read in English
2. Some mother tongues are never or rarely written
3. Mother tongues may have multiple dialects
4. Surveys are “formal” so mother tongue may be used less
5. Technical terms are more difficult to use in mother tongue
6. Interviewers may not speak mother tongue
Observations from the Field

- Respondent may shift from mother tongue to lingua franca
- Interviewer may shift to mother tongue of respondent
- Conversation may be in mother tongue with periodic interjections in lingua franca
- Translation on the fly for languages without formal translations.
Observations from the Field

- Respondent may shift from mother tongue to lingua franca
- Interviewer may shift to mother tongue of respondent
- Conversation may be in mother tongue with periodic interjections in lingua franca
- Translation on the fly for languages without formal translations.

But how often do these things occur?

Research Questions:

1. How often do the respondent and interviewer share the same first (i.e., native) language?
2. How often is the first language of the respondent and interviewer also the language of the conversation (interview)?
Possible Implications

- Lack of comprehension increases variance (random error), item non-response, acquiescence

- Using a lingua franca leads to “framing effects” and alters how respondents answer questions about democracy, West

*We will investigate these issues in a subsequent project, after the descriptive work is completed*
What Explains the Lack of Research?

- **Measurement of languages**
  - Sheer number of languages
  - How do you measure proficiency?
  - Need to measure multiple languages
  - People switch languages within the interview

- **Linguistic issues are tightly linked with logistics and personnel, and are harder to control** (contrast to mode/incentive/training)

- **Language is not randomly assigned to interviewers and respondents; language confounded with ethnicity, urban/rural, etc.**
Data: Afrobarometer Wave 4

27,713 respondents
810 interviewers
20 countries

Multi-stage stratified cluster samples;
random walk used to sample households

Paper-and-pencil survey about political attitudes and behavior

High levels of item non-response

Measures
1. Interviewer first language
2. Respondent first language
3. Language of conversation (i.e., interview)
4. Languages respondent speaks well
<table>
<thead>
<tr>
<th>NAME</th>
<th>DESCRIPTION</th>
<th>KENYA EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Respondent First Language Interviewer First Language Conversation</td>
<td></td>
</tr>
<tr>
<td>Common Language</td>
<td>Interviewer, Respondent, Conversation all in the same language (the ideal scenario)</td>
<td>Kikuyu Kikuyu Kikuyu</td>
</tr>
<tr>
<td>Opt-out of First</td>
<td>Interviewer and Respondent share language, but survey is conducted in different language</td>
<td>Kikuyu Kikuyu Swahili</td>
</tr>
<tr>
<td>Interviewer Compromises</td>
<td>Conversation is done in respondent's home language, but not interviewer's home language</td>
<td>Kikuyu Kikamba Kikuyu</td>
</tr>
<tr>
<td>Respondent Compromises</td>
<td>Conversation is done in interviewer's home language, but not respondent's home language</td>
<td>Kikamba Kikuyu Kikuyu</td>
</tr>
<tr>
<td>Third Language as Bridge</td>
<td>Respondent and Interview have different home languages, and conversation done in a 3rd language</td>
<td>Kikamba Kikuyu Swahili</td>
</tr>
</tbody>
</table>
Results: Language Combinations - Kenya

- Common: 58%
- Opt-Out of Native: 19%
- Intw. Compromises: 4%
- Resp. Compromises: 2%
- 3rd Language as Bridge: 17%
Results: Language Combinations – Nigeria

- Common: 41%
- Opt-Out of Native: 19%
- Intw. Compromises: 22%
- Resp. Compromises: 18%
- 3rd Language as Bridge: 0%

Legend:
- Common
- Opt-Out of Native
- Intw. Compromises
- Resp. Compromises
- 3rd Language as Bridge
### Results: Language Combinations – All countries

|-----------------|----------|------|---------|------------|-------|---------|-------------|--------|---------|--------|-------|--------|-------|---------|------------|-------------|---------|----------|-----------|--------|

**Legend:**
- Common
- Opt-Out of Native
- Intw. Compromises
- Resp. Compromises
- 3rd Language as Bridge
### Language of Conversation in Kenya (percentages)

<table>
<thead>
<tr>
<th>Language</th>
<th>(1) Common</th>
<th>(2) Opt-Out of First</th>
<th>(3) Interviewer Compromise</th>
<th>(4) Respondent Compromise</th>
<th>(5) Third Language as Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td>0</td>
<td>62</td>
<td>8</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td><strong>Kiswahili</strong></td>
<td>1</td>
<td>38</td>
<td>52</td>
<td>77</td>
<td>65</td>
</tr>
<tr>
<td><strong>Kikuyu</strong></td>
<td>30</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Luo</strong></td>
<td>28</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td><strong>Luhya</strong></td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Kamba</strong></td>
<td>13</td>
<td>0</td>
<td>12</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Kalenjin</strong></td>
<td>7</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Somali</strong></td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Common language uses local language
## Language of Conversation in Kenya (percentages)

<table>
<thead>
<tr>
<th>Language</th>
<th>(1) Common</th>
<th>(2) Opt-Out of First</th>
<th>(3) Interviewer Compromise</th>
<th>(4) Respondent Compromise</th>
<th>(5) Third Language as Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>0</td>
<td>62</td>
<td>8</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>Kiswahili</td>
<td>1</td>
<td>38</td>
<td>52</td>
<td>77</td>
<td>65</td>
</tr>
<tr>
<td>Kikuyu</td>
<td>30</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Luo</td>
<td>28</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Luhya</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Kamba</td>
<td>13</td>
<td>0</td>
<td>12</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Kalenjin</td>
<td>7</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Somali</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

- Common language uses local language
- People opt-out with lingua franca
## Language of Conversation in Kenya (percentages)

<table>
<thead>
<tr>
<th>Language</th>
<th>(1) Common</th>
<th>(2) Opt-Out of First</th>
<th>(3) Interviewer Compromise</th>
<th>(4) Respondent Compromise</th>
<th>(5) Third Language as Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td>0</td>
<td>62</td>
<td>8</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td><strong>Kiswahili</strong></td>
<td>1</td>
<td>38</td>
<td>52</td>
<td>77</td>
<td>65</td>
</tr>
<tr>
<td><strong>Kikuyu</strong></td>
<td>30</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Luo</strong></td>
<td>28</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td><strong>Luhya</strong></td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Kamba</strong></td>
<td>13</td>
<td>0</td>
<td>12</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Kalenjin</strong></td>
<td>7</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Somali</strong></td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

- Common language uses local language
- People opt-out with lingua franca
- Interviewer, respondent mostly compromise by using lingua franca
## Language of Conversation in Kenya (percentages)

<table>
<thead>
<tr>
<th>Language</th>
<th>(1) Common</th>
<th>(2) Opt-Out of First</th>
<th>(3) Interviewer Compromise</th>
<th>(4) Respondent Compromise</th>
<th>(5) Third Language as Bridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>0</td>
<td>62</td>
<td>8</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>Kiswahili</td>
<td>1</td>
<td>38</td>
<td>52</td>
<td>77</td>
<td>65</td>
</tr>
<tr>
<td>Kikuyu</td>
<td>30</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Luo</td>
<td>28</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Luhya</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Kamba</td>
<td>13</td>
<td>0</td>
<td>12</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Kalenjin</td>
<td>7</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Somali</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

- Common language uses local language
- People opt-out with lingua franca
- Interviewer, respondent mostly compromise by using lingua franca
- Lingua franca as a bridge
If an interviewer and a respondent both speak the same first language:

Why would they choose to Opt-Out of their first language and use a lingua franca?
If an interviewer and a respondent both speak the same first language:

Why would they choose to Opt-Out of their first language and use a lingua franca?

Logistic regression analysis: Higher education, urban, female, young more likely to opt-out
If an interviewer and a respondent both speak the same first language:

Why would they choose to Opt-Out of their first language and use a lingua franca?

Logistic regression analysis: Higher education, urban, female, young more likely to opt-out

Hypothesized reasons:
- Easier to speak about technical terms in lingua franca
- Social benefits to showing you can speak int’l language
- Discomfort with reading mother tongue language
Findings

- Linguistic issues are critical, understudied, and are hard to study
- Respondents often switch from mother tongue to lingua franca, but this varies by country
- When possible, respondents and interviewers choose to Opt-Out
- Respondents and interviewers compromise equally

Next Steps: What are the implications of language differences for data quality and response distributions?

Next Steps: Add dataset that contains information on language switching during interview.
Charles Q. Lau, PhD, MS
Survey Methodologist
+1.919.541.8798
clau@rti.org