Tutorial on Experimental work
Maria Polinsky & Stavros Skopeteas

Part II. experimental studies on Georgian
by Stavros Skopeteas

Paris, September 23, 2016
experimental methods in grammatical research

linguistic research based on two sources of evidence:

- observational research: possible linguistic entities in corpora
- introspective research: grammaticality judgments

recent focus on repeated-observation designs:

- quantitative corpus studies or production experiments
- questionnaire studies on speakers' intuitions

tutorial, part I: offline methods, not informative for cognitive processes, crucial for testing grammatical hypotheses

motivation for these developments: increase reliability, estimate gradience
contents of this part

experimental data in language comparison

- increase reliability
  illustrative study: focus in OV languages (Georgian, Armenian, Turkish)
  (questionnaire study)

speaking in the lab vs. speaking in the world

- same contrasts, different sources of variation
  illustrative study: prosodic effects of order and focus in Georgian
  (scripted speech vs. corpus data)
experimental data in language comparison
grammatical background

Georgian, Armenian and Turkish

are OV languages, but differ in details:

- local $\rho$-expression of focus (independent of order)
- optional V-fronting (creates word order flexibility)
grammatical background

Georgian, Armenian and Turkish

OV languages, differing in details:

- local \( p \)-expression of focus
  possible: \( X_PF \ Y_P \ V \)

- optional V-fronting
  possible: \( V \ X_PF \ V \)
method

illustration

Q  ა ვახლი  და ბურტი?  
   who    found     ball?  
A  კირა  და ბურტი  ვახლი.  
   Kira     ball     found  
          (does not fit)  1  2  3  4  5  6  7  (fits)  

factorial design

CONCEPT (question):
Sentence-focus|Subject-focus|Object-focus|Verb-focus

WORD ORDER (answer):
SVO|SOV|OVS|OSV|VSO|VOS

random factors

48 lexicalizations, 48 speakers
result

<table>
<thead>
<tr>
<th></th>
<th>Georgian</th>
<th></th>
<th>Armenian</th>
<th></th>
<th>Turkish</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S_F</strong></td>
<td>X_{f}V</td>
<td>VX_{f}</td>
<td>X_{f}V</td>
<td>VX_{f}</td>
<td>X_{f}V</td>
<td></td>
</tr>
<tr>
<td><strong>O_F</strong></td>
<td>VX_{f}</td>
<td>X_{f}V</td>
<td>VX_{f}</td>
<td>X_{f}V</td>
<td>X_{f}V</td>
<td></td>
</tr>
<tr>
<td><strong>VSO</strong></td>
<td>SOV</td>
<td>OVS</td>
<td>OSV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Graph of acceptability vs order](image1)

![Graph of acceptability vs order](image2)

![Graph of acceptability vs order](image3)
discussion

on the reliability of controlling sources of variance in a Between Languages design

tutorial, part I: acceptability vs corpus: what do we learn about modularity from the mapping between different data types?
speaking in the lab vs. speaking in the world
H-phrase tones are integrated within a single domain of downstep in SOV but not in SVO.
method

Context * Order
(all and only felicitous permutations)

<table>
<thead>
<tr>
<th></th>
<th>SOV</th>
<th>SVO</th>
<th>OSV</th>
<th>OVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>allF</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VP</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>VF</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OF</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Items (4)
nino mama-s e-loliav-eb-a.
Nino(NOM) father-DAT O3-care-THM-DEP.S3.SG
Nino cares about the father.

Context questions
- ALL: What happens?
- VP: What do we hear about Nino?
- S: Who cares about the father?

eight speakers
two repetitions per item

Total= 832 tokens
results

two sources of non-downstepped H-phrase tones:

- right-edge of a **non-final V**
- preceding a **focused XP**
theoretical relevance

reported phrase-languages

ALIGN(focus-\(i\varphi-i\))

Align the \(i\)-edge of the focus with the \(i\)-edge of a prosodic phrase.

(Büring 2010, Féry 2013)

hypothesis

In a subset of these languages (among else Georgian), the prosodic edges are mapped on nucleus (not on focus)

ALIGN(nucleus-\(i\varphi-i\))

Align the \(i\)-edge of the nucleus with the \(i\)-edge of a prosodic phrase.
syntax-phonology mapping

stress-based language
maximally embedded: stressed

phrase-based language
maximally embedded: separated
speaking in the lab vs. speaking in the world

external validity?

• is the inference from the sample to the world valid
• or is it an artefact of the lab situation?

background:

• current discussion in phonetics about scripted data.
  (Xu 2010; Wagner, Trouvain, & Zimmerer, 2014, among else)
# Georgian corpus

## Material

### Narratives
- Activity description
- Ancestor story
- Path description
- Event description
- Comparative description

### Style
- non-scripted data; not spontaneous, but less attention to speech; naturalistic behavior (laughter, spontaneous speech planning)

## Speakers

### 24 speakers
- residents of Tbilisi;
- speaking Georgian as a first language;
- generally high education level;
- age range: 20-58 (average 29.3).
phenomena of interest
- order: XP V YP vs. XP YP V
- focus: approximated through givenness

phenomena outside the scope of the intended generalization => excluded
- non-declaratives, embedded clauses
- laughter, disfluencies, etc.

phenomena for which the generalization is intended to apply (random factors)
- speakers
- syllable structures
- syntactic category of the XP

<table>
<thead>
<tr>
<th>speaker</th>
<th>XP V YP</th>
<th>XP YP V</th>
<th>dur (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>8</td>
<td>5</td>
<td>207</td>
</tr>
<tr>
<td>02</td>
<td>5</td>
<td>3</td>
<td>497</td>
</tr>
<tr>
<td>03</td>
<td>17</td>
<td>5</td>
<td>268</td>
</tr>
<tr>
<td>04</td>
<td>11</td>
<td>8</td>
<td>309</td>
</tr>
<tr>
<td>05</td>
<td>6</td>
<td>4</td>
<td>228</td>
</tr>
<tr>
<td>06</td>
<td>9</td>
<td>4</td>
<td>179</td>
</tr>
<tr>
<td>09</td>
<td>10</td>
<td>7</td>
<td>320</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>5</td>
<td>365</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>5</td>
<td>218</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>2</td>
<td>173</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>3</td>
<td>196</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>4</td>
<td>290</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>6</td>
<td>296</td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>10</td>
<td>389</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>8</td>
<td>361</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>2</td>
<td>197</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>3</td>
<td>215</td>
</tr>
<tr>
<td>22</td>
<td>11</td>
<td>3</td>
<td>483</td>
</tr>
<tr>
<td>23</td>
<td>13</td>
<td>2</td>
<td>167</td>
</tr>
<tr>
<td>24</td>
<td>7</td>
<td>4</td>
<td>99</td>
</tr>
<tr>
<td>total</td>
<td>182</td>
<td>93</td>
<td>91 min.</td>
</tr>
</tbody>
</table>
results

word order effect
final vs. non-final V

measurements
voiced part of the first and last syllable of the target constituent
\( F_0 \)-mean of three equal intervals per target \( \sigma \)

givenness
as predictor of focus
given = referent mentioned in the pre-text
comparison between data types

scripted speech

unscripted speech

rise in the medial word:
$F_0$-max (last syllable) – $F_0$-min (first syllable) of the medial constituent
final discussion

- on the external validity of experimental data
- controlling variation vs outbalancing variation
- are there lab grammars?

Continuing the discussion in tutorial, part I: limits of experimental designs, when should I come back to the traditional methods?
Tutorial on Experimental work
Maria Polinsky & Stavros Skopeteas

Part II. Georgian
by Stavros Skopeteas

Paris, September 23, 2016