Gender attribution of borrowings and native lexemes in Zilo (Andi)

Samira Verhees and George Moroz

Linguistic Convergence Laboratory, NRU HSE, Moscow, Russia

11 May 2018 г.

Anatolia-the Caucasus-Iran: Ethnic and Linguistic Contacts
Institute of Oriental Studies in Russian-Armenian State University
Andi and their neighbors

- Nakh
  - Chechen
- Avaro-Andic-Tsezic
  - Avar
  - Andi
    - Botlikh
    - Karata
    - Andi
Andi and their neighbors

- Nakh
  - Chechen
- Avaro-Andic-Tsezic
  - Avar
  - Andi
    - Botlikh
    - Karata
    - Andi

All Andi speakers are multilingual:

- Avar — old lingua franca of a region
- Russian — new lingua franca for the whole of Daghestan (2 centuries of active Russification)
Gender systems
Gender systems

- The impact the systems have most notably on verbal agreement, but also on some adjective, numeral, adverbial, pronominal, and postpositional agreement.
- Some POS agree with absolutive argument of the clause. Some POS agree with its head.
- In most cases gender is a covert category, but there are some exception, such as:
  - w-ots:i (M-sibling)
  - j-ots:i (F-sibling)
  - b-ots:i (A-N-sibling) — from the “Three Little Pigs” fairytale
# Andi gender inventory

<table>
<thead>
<tr>
<th>Reference</th>
<th>Language</th>
<th>SG</th>
<th>PL</th>
<th>M</th>
<th>F</th>
<th>AN</th>
<th>¬AN</th>
<th>¬AN 1</th>
<th>¬AN 2</th>
<th>¬AN 3</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Alekseev 1999]</td>
<td>Andi</td>
<td>w</td>
<td>j</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>r</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Salimov 1968]</td>
<td>Gagatli</td>
<td>w</td>
<td>j</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>r</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Suleymanov 1957]</td>
<td>Rikwani</td>
<td>w</td>
<td>j</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>r</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fieldwork</td>
<td>Zilo</td>
<td>w</td>
<td>j</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>r</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Suleymanov 1957]</td>
<td>Muni</td>
<td>w</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>b</td>
<td></td>
</tr>
</tbody>
</table>

- **M vs. F vs. ¬AN vs. ¬ANS**
- **SG = PL** (except AN and ¬AN 3)
- Rikwani’s ¬AN 3 gender contains only several words: *hotʃ* ‘scorpion’, *odoruk’a* ‘butterfly’, and other 5 words
- In “lower-group” (Muni and Kvanxidatli) there are only three genders (M vs. F vs. ¬H)

---

6
## Neighbour languages gender inventories

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>M</th>
<th>F</th>
<th>¬H</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Alekseyev and Atayev 1997]</td>
<td>Avar</td>
<td>SG</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Magomedbekova 1971]</td>
<td>Karata</td>
<td>SG</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>[Gudava 1962]</td>
<td>Botlikh</td>
<td>SG</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>[Nichols 1994]</td>
<td>Chechen</td>
<td>SG</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL</td>
<td>b</td>
<td>d</td>
</tr>
</tbody>
</table>

- M vs. F vs. ¬H
- SG ≠ PL (except Chechen)
Gender attribution experiment
Experiment: Set up

How speakers decide about new words?

- collect 25 native and 25 loan words of ¬AN 1 gender
- collect 25 native and 25 loan words of ¬AN 2 gender
- add some words of Rikwani’s ¬AN 3 gender
- pseudo random stimuli order
- 16 Zilo native speakers
- context: di-b / di-r X “my X”
Experiment: Selected Words

22  kots’i broom
48  aw?ara cap
4  bats’a grasshopper
13  ts’a fire
68  nitfo scythe
57  rok’o heart
92 flen water
29  hark’u eye
46  tɬ’ir bridge
25  q’en wall
85  refa tree
28  kʷeɬir sleeve
79  hunts’i honey
98  raketa rocket
106  kolxoz collective farm
39  kalendar’i calendar
83  sumka bag
65  masina car
58  pasport passport
71  simkarta sim-card
16  kirpitʃ brick
63  tarif tariff
105  rozjetka power outlet
45  patʃka packet
64  esemeska SMS
75  plastir’i plaster
### Experiment: Inter-rater Reliability

<table>
<thead>
<tr>
<th></th>
<th>% agreem.</th>
<th>Fleiss’ Kappa</th>
<th>z-score</th>
<th>[Fleiss 1971]</th>
</tr>
</thead>
<tbody>
<tr>
<td>all stimuli</td>
<td>74.5</td>
<td>0.849</td>
<td>95.7</td>
<td>Almost perfect agreement</td>
</tr>
<tr>
<td>native stimuli</td>
<td>84.2</td>
<td>0.937</td>
<td>77.5</td>
<td>Almost perfect agreement</td>
</tr>
<tr>
<td>borrowings</td>
<td>63.3</td>
<td>0.732</td>
<td>56.1</td>
<td>Substantial agreement</td>
</tr>
</tbody>
</table>

- **Almost perfect agreement**: very high but not ideal agreement among speakers in native lexicon
- **Small and high agreement among speakers in borrowing lexicon**: very high agreement among speakers
- **Almost perfect agreement**: smaller but still high agreement among speakers in borrowing lexicon
- **Substantial agreement**: very high agreement among speakers
## Experiment: Inter-rater Reliability

<table>
<thead>
<tr>
<th></th>
<th>% agreem.</th>
<th>Fleiss’ Kappa</th>
<th>z-score</th>
<th>[Fleiss 1971]</th>
</tr>
</thead>
<tbody>
<tr>
<td>all stimuli</td>
<td>74.5</td>
<td>0.849</td>
<td>95.7</td>
<td>Almost perfect agreement</td>
</tr>
<tr>
<td>native stimuli</td>
<td>84.2</td>
<td>0.937</td>
<td>77.5</td>
<td>Almost perfect agreement</td>
</tr>
<tr>
<td>borrowings</td>
<td>63.3</td>
<td>0.732</td>
<td>56.1</td>
<td>Substantial agreement</td>
</tr>
</tbody>
</table>

- very high agreement among speakers
- very high but not ideal agreement among speakers in native lexicon
- smaller but still high agreement among speakers in borrowing lexicon
Experiment: Variable Words

27 variable words with more than one speaker:

<table>
<thead>
<tr>
<th>id</th>
<th>Zilo</th>
<th>English</th>
<th>source</th>
<th>AN 1</th>
<th>AN 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>patśka</td>
<td>packet</td>
<td>loan</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>64</td>
<td>antena</td>
<td>antenna</td>
<td>loan</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>101</td>
<td>paket</td>
<td>packet</td>
<td>loan</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>χur</td>
<td>field, garden</td>
<td>native</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>44</td>
<td>provod</td>
<td>line</td>
<td>loan</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>70</td>
<td>gol</td>
<td>goal</td>
<td>loan</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>21</td>
<td>simkarta</td>
<td>SIM card</td>
<td>loan</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>24</td>
<td>zatśetka</td>
<td>students record-book</td>
<td>loan</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>wagon</td>
<td>train car</td>
<td>loan</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>20</td>
<td>odoruk’a</td>
<td>butterfly</td>
<td>native</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>39</td>
<td>kalendar:]</td>
<td>calendar</td>
<td>loan</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>75</td>
<td>plastir:]</td>
<td>plaster</td>
<td>loan</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>41</td>
<td>sputn:]ik</td>
<td>satellite</td>
<td>loan</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>98</td>
<td>raketa</td>
<td>rocket</td>
<td>loan</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>71</td>
<td>kometa</td>
<td>comet</td>
<td>loan</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>72</td>
<td>lift</td>
<td>lift</td>
<td>loan</td>
<td>14</td>
<td>2</td>
</tr>
</tbody>
</table>
Experiment: Speakers’ Hierarchical Clustering
Experiment: conclusion

- There is a variability in gender attribution even in native words
- Gender attribution in borrowings are more variable than native words
- Variability is symmetrical: there is no ¬AN 1 or ¬AN 2 gender preference
- Variability is somehow connected with the age, and not with the gender
Experiment: conclusion

- There is a variability in gender attribution even in native words.
- Gender attribution in borrowings are more variable then native words.
- Variability is symmetrical: there is no $\neg$AN 1 or $\neg$AN 2 gender preference.
- Variability is somehow connected with the age, and not with the gender.
- It is interesting to compare these results with results of non-words experiments (compare with [Dąbrowska 2005]).
Thank you!

Send us a mail!
agricolamz@gmail.com
jh.verhees@gmail.com
References


Gudava, T. E. (1962). Botlikhskiy yazik. AN GSSR.


