EXPERIMENTAL MORPHOSYNTAX OF SOUTH SLAVIC LANGUAGES – GENDER AGREEMENT

Anita Peti-Stantić, University of Zagreb/Tufts University

The Polinsky Language Science Lab
Harvard University
April 2015
COORDINATED RESEARCH IN THE EXPERIMENTAL MORPHOSYNTAX OF SOUTH SLAVIC LANGUAGES (EMSS)

Leverhulme project: http://www.ucl.ac.uk/pals/research/linguistics/research/leverhulme

Andrew Nevins, UCL, Research Head
Jana Willer Gold, Network Coordinator
Partners: Nova Gorica, Zadar, Zagreb, Sarajevo, Novi Sad, Niš
MOTIVATION AND BACKGROUND

• Some languages have rules which are basically syntactic, others rely on a semantic principle and yet others show interesting combinations of the two principles. … gender resolution rules are language specific.

• [Corbett, Resolution rules, CH8, Agreement pp 175]

• Greville Corbett: Hierarchies, Targets, and Controllers: Agreement Patterns in Slavic Languages; Gender; Number; Features; Agreement; The Syntax-morphology interface; Canonical Morhpology and Syntax
THEORETICAL ASSUMPTIONS

• In a variety of languages (in Slavic language family Bosnian, Croatian, Serbian, Slovenian, in other languages - Hindi, Ndebele), coordinated subjects (and in some cases, objects) consisting of two (and more) NPs can trigger verbal agreement with only one of these noun phrases, rather than with the coordination as a whole.

• Terminology:
  • non-coordinated/simple/non-conjoined subjects
  • coordinated/conjoined subjects
  • full/resolved agreement vs. partial agreement (agreement with a single conjunct)
**BASIC FACTS**

- **AGREEMENT – South Slavic languages**
  - Subjects obligatorily agree with the verb (both pre- and post-verbally)
  - Agreement with finite verbs (aux and main verbs) in person & number – without exception
    - Stol stoji nasred sobe. ‘The table is at the middle of the room.’
    - Stolovi stoje nasred sobe.
    - Nasred sobe stoji stol. ‘At the middle of the room there is a table.’
    - Nasred sobe stoje stolovi.
    - Ja stojim nasred sobe
    - Mi stojimo nasred sobe.

- Agreement with participles in number & gender – without exception with number, variable with gender
### BASIC DATA – FULL AGREEMENT

#### NON-COORDINATED SUBJECTS

<table>
<thead>
<tr>
<th>Sg</th>
<th>Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dječak je zaustavljjen.</td>
<td>Dječaci su zaustavljeni.</td>
</tr>
<tr>
<td>BoyMsg Auxsg StopPartMsg</td>
<td>Mpl pl Mpl</td>
</tr>
<tr>
<td>The boy has been stopped.</td>
<td>The boys have been stopped.</td>
</tr>
<tr>
<td>Pjesma je otpjevana.</td>
<td>Pjesme su otpjevane.</td>
</tr>
<tr>
<td>SongFsg Auxsg SingPartFsg</td>
<td>Fpl pl Fpl</td>
</tr>
<tr>
<td>The song has been sung.</td>
<td>The songs have been sung.</td>
</tr>
<tr>
<td>Pismo je poslano.</td>
<td>Pisma su poslana.</td>
</tr>
<tr>
<td>LetterNsg Auxsg SendPartNsg</td>
<td>Npl pl Npl</td>
</tr>
<tr>
<td>The letter has been sent.</td>
<td>The letters have been sent.</td>
</tr>
</tbody>
</table>
(NOT) FULL AGREEMENT

<table>
<thead>
<tr>
<th>COORDINATED SUBJECTS</th>
<th>COORDINATED SUBJECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MM/FF/NN</strong></td>
<td><strong>MM/FF/NN</strong></td>
</tr>
<tr>
<td><strong>Sg</strong></td>
<td><strong>Pl</strong></td>
</tr>
</tbody>
</table>

- **Dječak i pas su zaustavljeni.**
  - BoyMsg DogMsg Auxpl StopPartMpl
  - The boy and the dog have been stopped.

- **Pjesma i melodija su otpjevane.**
  - SongFsg MelodyFsg Auxpl SingPartFpl
  - The song and the melody have been sung.

- **Pismo i pero su poslani.**
  - LetterNsg FeatherNsg Auxpl SendPartMpl
  - The letter and the feather have been send.

- * **Pismo i pero su poslana.**
  - LetterNsg FeatherNsg Auxpl SendPartNpl
  - The letter and the feather have been send.

- **Dječaci i psi su zaustavljeni.**
  - Mpl Mpl pl Mpl
  - The boys and the dogs have been stopped.

- **Pjesme i melodije su otpjevane.**
  - Fpl Fpl pl Fpl
  - The songs and the melodies have been sung.

- **Pisma i pera su poslana.**
  - Npl Npl pl Npl
  - The letters and the feathers have been send.
BASIC FACTS – masculine

- Sg
  - Dječak je zaustavljen. Zaustavljen je dječak.
- Pl
  - Dječaci su zaustavljeni. Zaustavljeni su dječaci.
- Sg&Sg
  - Dječak i pas su zaustavljeni. Zaustavljeni su dječak i pas.
- Pl&Sg
  - Dječaci i pas su zaustavljeni. Zaustavljeni su dječaci i pas.
- Sg&Pl
  - Dječak i psi su zaustavljeni. Zaustavljeni su dječak i psi.
- Pl&Pl
  - Dječaci i psi su zaustavljeni. Zaustavljeni su dječaci i psi.

- ... Agreement pattern with Masculine NP’s is not repeated in Agreement patterns with Feminine and Neuter NP’s
BASIC FACTS – feminine

- Sg
  - Djevojčica je zaustavljena.
- Pl
  - Djevojčice su zaustavljene.
- Sg&Sg
  - Djevojčica i mačka su zaustavljene.
- Pl&Sg
  - Djevojčice i mačka su zaustavljene.
  - ? Djevojčice i mačka su zaustavljeni. DEF
- Sg&Pl
  - Djevojčica i mačke su zaustavljene.
  - ? Djevojčica i mačke su zaustavljene. DEF
- PI&PI
  - Djevojčice i mačke su zaustavljene.
  - ? Djevojčice i mačke su zaustavljene. DEF
**BASIC FACTS – neutrum**

- **Sg**
  - Pismo je poslano.
- **Pl**
  - Pisma su poslana.
- **Sg&Sg**
  - Pismo i pero su poslani.  **DEF**
  - *Pismo i pero su poslana.*
- **Pl&Sg**
  - Pisma i pero su poslani.  **DEF**
  - *Pisma i pero su poslana.*
- **Sg&Pl**
  - Pismo i pera su poslani.  **DEF**
  - *Pismo i pera su poslana.*
- **Pl&Pl**
  - Pisma i pera su poslana.
BASIC FACTS: PUZZLE

• What happens when the coordinated subjects consists of two nouns that differ in gender?

• Alltogether 9 combinations
  • 3 same-gender combinations: MM, FF, NN
  • 6 different-gender combinations: MF, MN, FM, FN, NM, NF
    • (directionality matters: MF and FM is not the same)
QUESTIONING THEORETICAL ASSUMPTIONS

• 1. Coordination between same-gender nouns > full agreement (M > M, F > F, N > N in pl&pl condition)
  • PUZZLE: sg+sg; pl+pl; sg+pl; pl+sg >pl, but no full gender agreement in all conditions

• 2. Coordination between different-gender nouns > ‘partial agreement’ (sg+sg/pl+pl.sg+pl/pl+sg >pl, but...
  • MF > M and/or F
  • MN > M and/or N
  • FM > F and/or M
  • FN > F and/or N and/or M
  • NM > N and/or M
  • NF > N and/or F and/or M
THEORETICAL ASSUMPTIONS

• Grammars treat Masc as a default
• If &P contains Masc, significantly higher proportion of Masc agreement
• If &P contains Masc, significantly faster
  • most probably because of Masc default, ‘consolation’ option
  • (OR SLOWER as ‘last resort’)
• The most interesting conditions are FN and NF (without Masc)

• MF > M and/or F
• MN > M and/or N
• FM > F and/or M
• FN > F and/or N and/or M
• NM > N and/or M
• NF > N and/or F and/or M
THEORETICAL ASSUMPTIONS

TWO BASICALLY DIFFERENT MECHANISMS

• 1. (Linearly) closest conjunct (Closest Conjunct Agr) – CCA
  • LCA in SV (lowest ranked conjunct)
  • FCA in VS (highest ranked conjunct (HCA))
• 2. (Hierarchically) highest/first conjunct (Highest Conjunct Agr) – FCA
  • In SV – first (FCA) and at the same time the furthest conjunct (DCA)
  • In VS – first (FCA), but not the furthest conjunct (DCA)

• Never agreement with conjunct that is neither closest, nor highest/first (conjunction with more than two conjuncts, second conjunct in VS)
THEORETICAL ASSUMPTIONS

• SV

• VS

FCA

\[
\begin{align*}
&\text{ConjP1} \\
&\text{& P} \\
&\text{& ConjP2} \\
\end{align*}
\]

\[
\begin{align*}
&\text{ConjP1} \\
&\text{& P} \\
&\text{& ConjP2} \\
\end{align*}
\]

CCA

\[
\begin{align*}
&\text{ConjP1} \\
&\text{& P} \\
&\text{& ConjP2} \\
\end{align*}
\]

\[
\begin{align*}
&\text{ConjP1} \\
&\text{& P} \\
&\text{& ConjP2} \\
\end{align*}
\]
PREVIOUS RESEARCH

• Bošković, 2009: CCA in SV positions blocked when M hierarchically highest/linearly first
  • ?*Svi gradovi i sva sela su (juče) uništena.
• Puškar and Murphy, 2014: CCA is an illusion
• Marušić et al. 2015: effects of mixing genders and directionality
  • Finding 1: Some default masculine agreement occurs even when both conjuncts are same gender (FF, NN)
  • Finding 2: More default agreement when conjuncts are neuter than feminine
  • Finding 3: Masc Agreement with Masc + non-Masc ambiguous: either default agreement or FCA/CCA
  • Finding 4: NF and FN cases reveal three distinct response types: highest, closest and default
THEORETICAL ASSUMPTIONS – SV and VS

EXAMPLES:

- **Pjesme, pisma i poruke su poslane.** FCA or CCA
  - F N F F
- **Pjesme, pisma i poruke su poslana.** neither FCA no CCA
  - F N F N
- **Pjesme, pisma i poruke su poslani.** DEF

- **Pjesme i pisma su poslana.** CCA
- **Pjesme i pisma su poslane.** FCA (possible, but rare)
- **Pjesme i pisma su poslani.** DEF

- **Poslane su pjesme i pisma.** CCA
  - F F N
- **Poslana su pjesme i pisma.** neither CCA nor FCA
  - N F N
- **Poslani su pjesme i pisma.** DEF
THEORETICAL ASSUMPTIONS

• A number of factors determine agreement with coordinated structures:

  • Preverbal (SV) or postverbal (VS)
  • **Number** (Bošković vs. Marušič et al. and others) – next step
  • Possibly: case, animacy, elipsis, information structure
EXPERIMENT: MATERIALS AND PARTICIPANTS

• Controlled experimental study
• 6 partner sites in 4 countries (Slovenia, Croatia 2x, Bosnia, Serbia 2x)
• Uniformity vs. Site specificity
  • uniform experimental testing battery at six universities in parallel
  • keep experimental settings as identical as possible
    • uniform stimulus design, presentation methods, number of participants
  • quantitative analysis
GENERAL GOALS OF OUR STUDY

• To extend the purview of psycholinguistic research of morphosyntactic variability to the South Slavic languages.

• Measure elicited production latencies as a function of the ‘ambiguity’ of the structural possibilities to be produced - a measure of the degree of inter-individual variation for given conditions.

• **HY0 :** The structures with the richest array of possibilities (CondFN and NF) - greatest amount of time needed to ‘decide’ on agreement (reflected in their latencies)
FURTHER GOALS OF OUR RESEARCH

• AIM of the First Phase of the Project - to chart the range of **possible and impossible variation** in the selection of agreement controllers in coordinate structures

• Two relevant factors:
  • CS preverbal or postverbal
  • conjuncts in CS all plural number

• HY1: **Preverbal coordination structures** allow a richer array of agreement controllers.
PILOT

- **Type**: On-line elicited production task with digital recording of response times. Participants are tested individually.
- Participants: n=min. 5 max.10
- **Goal**: Elicit spoken agreement in coordinated subjects with two, three and four conjuncts.
- **Data**: 12 conditions x 4 items per condition = 48 test examples + 48 fillers = 96 sentences. Matching and non-matching gender conditions.
  - MM, MMM, MMMM, FF, FFF, FFFF, MF, MMF, MMMF, FN, FFN, FFFN
  - 4 learning examples
  - Randomisation per site and participants

- IBEX automatically records the reaction times of the produced responses.
PILOT

• **Conditions:**
  • subjects with double, triple, quadruple conjunct
  • 2 gender possibilities
  • Preverbal

• **Fillers:**
  • non-conjoined plural subjects
  • matching and mixed genders (M, N, F)

• **Task:** Participant sees a model sentence on the screen, with a Msg NP as the subject.
  • The participant then sees a replacement NP in the center of the screen.
  • Their task is to produce an utterance in which they replace the subject of the model sentence with a new NP.
  • Based on materials developed in Marušič et al.
PILOT : Reaction Time
PILOT: Latency
Lokot je spremljen u kutiju.
Čekići i vijci
PILOT – general results

• 1. RESULT
  • AvgCondMM : 3334 ms
  • AvgCondMF : 3745 ms
  • AvgCondFN : 4693 ms

• 2. RESULT
  • Significantly longer reaction times when 3 or 4 conjuncts in CS (up to 7000 ms)
CURRENT EXPERIMENT

• **GOAL:** ELICIT SPOKEN AGREEMENT WITH THE PREVERBAL AND POSTVERBAL COORDINATED SUBJECTS.

• **TASK:** preverbal (1a) and postverbal (1b) coordinated subjects matched and non-matched in gender

• **TYPE OF EXPERIMENT:**
  • self-paced reading and sentence completion task
  • on-line elicited production task with digital recording of responses

• **RECORDING AND CODING:**
  • Audacity, Praat, IBEX
  • coded according to agreement endings and agreement features

• **PARTICIPANTS:** n=30 [1st year student, attended the local secondary school, non linguist, not students of B/C/S/S Language, F/M 50/50]

• **LANGUAGE/DIALECT/VARIETY:** Neutral dia-standard.
CURRENT EXPERIMENT

• EXPERIMENTAL ITEMS:
  • 9 conditions (matched: MM, FF, NN, non-matched: MF, MN, FM, FN, NM, NF)
  • 6 items per condition
    • 54 test examples
    • 54 fillers
    • total of 108 sentences
  • 2 subexperiments for preverbal and postverbal condition

• Model sentences used as primes for the test examples contained a simple non-conjoined singular subjects in Masculine.
CURRENT EXPERIMENT

• FILLERS:
  • 6 general conditions: simple non-conjoined subjects in all genders and numbers (M<sub>sg</sub>, M<sub>pl</sub>, F<sub>sg</sub>, F<sub>pl</sub>, N<sub>sg</sub>, N<sub>pl</sub>)
  • 3 additional filler conditions correlating with each of the 6 model sentence conditions consisted of 3 examples, where the 3 conditions were: Paucal (2, 3, 4) with a head noun in M<sub>sg</sub>, Hybrid noun in F<sub>sg</sub> (gospoda, poglavica, janjad etc.) and Object Relative clause with a head noun in N<sub>sg</sub>
CURRENT EXPERIMENT: GENERAL CONDITIONS FOR THE STIMULI

1: INTERNAL STRUCTURE

- all model sentences comply with the following word order and internal structures:
  - COND SV: [S Aux V Adv], *[S Aux V PP]
  - COND VS: [Adv Aux V S], *[PP Aux V S]
- all nouns in subject and non-subject positions are inanimate and non-human
CURRENT EXPERIMENT: GENERAL CONDITIONS FOR THE STIMULI

• 2. BALANCING AVERAGE LENGTH
  • model sentences priming the stimuli, all replacement phrases in the stimuli and each conjunct in the stimuli replacement phrase are of the similar length (measured in the number of characters)
  • Model Sentence Length : Characters Cca. 28, Syllables Cca. 10.
  • Stimulus Replacement Phrase Length : Characters Cca. 17, Syllables Cca. 7.
  • Conjunct Lenght : Characters Cca. 7, Syllables Cca. 3.
CURRENT EXPERIMENT: GENERAL CONDITIONS FOR THE STIMULI

• 3. SV-VS CONSISTENCY MEANS SV-VS COMPARABILITY (information structure concern)
  • all model sentence examples converted to the VS order have to maintain the neutral sentence interpretation, the subject should not be interpreted contrastively
CURRENT EXPERIMENT: EXPERIMENTAL SENTENCES

1. MM
- SV Ugovor je potpisan u utorak.
- VS U utorak je potpisan ugovor. Dokumenti i indeksi

2. FF
- SV Zakon je donesen na sjednici.
- VS Na sjednici je donesen zakon. Odluke i presude

3. NN
- SV Kvar je popravljen u garaži.
- VS U garaži je popravljen kvar. Jedra i vesla
CURRENT EXPERIMENT: EXPERIMENTAL SENTENCES

- **4. MF**
  - SV *Arhiv* je otvoren za javnost.
  - VS Za javnost je otvoren *arhiv*.

- **5. MN** *Ručnik* je obješen u kupaoni.

- **6. FM** *Sastav* je utvrđen na sastanku.

- **7. FN** *Dogovor* je pomaknut na petak.

- **8. NM** *Desert* je poslužen na terasi.

- **9. NF** *Doručak* je naručen u sobu.
CURRENT EXPERIMENT: FILLERS

- Matched for gender and number (Msg, Mp, Fsg, Fpl, Nsg, Npl) + 3 conditions (Paucal, Hybrid noun and Object Relative clause)

- Ulica je zatvorena za promet.
- Kćerka je naslijedila bogatstvo.
- Reklama je puštena na radiju.
- Tri trga
- Vojvoda
- Pitanje koje su postavili slušatelji
EXPERIMENT DESIGN

- INTRO
- PRACTICE
- EXPERIMENT
- OUTRO

- MODEL SENTENCE
- STIMULUS
- BLANK SCREEN
PRELIMINARY RESULTS

ZAGREB: DIFFERENT GENDER: variability according to the position (SV)
PRELIMINARY RESULTS

ZAGREB: DIFFERENT GENDER: variability according to the position (VS)
VARIATION ACROSS SUBJECTS – SV
VARIATION ACROSS SUBJECTS -VS

FN VS

NF VS
PRELIMINARY RESULTS – all

Agreement combined (SV)
PRELIMINARY RESULTS – all

Agreement combined (VS)
PRELIMINARY REACTION TIME RESULTS

**ZG - Non-Mixed Conjuncts (SV)**

**ZG - Non-Mixed Conjuncts (VS)**

**ZG - Mixed Conjuncts: Variation by Strategy (SV)**

**ZG - Mixed Conjuncts: Variation by Strategy (VS)**
## Croatian National Corpus: Frequencies

<table>
<thead>
<tr>
<th>Condition</th>
<th>Plural, NominaBve</th>
<th>Series1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>27896</td>
<td>3000</td>
</tr>
<tr>
<td>FF</td>
<td>10685</td>
<td>2500</td>
</tr>
<tr>
<td>NN</td>
<td>23</td>
<td>2000</td>
</tr>
<tr>
<td>MF</td>
<td>5667</td>
<td>1500</td>
</tr>
<tr>
<td>MN</td>
<td>700</td>
<td>1000</td>
</tr>
<tr>
<td>FM</td>
<td>4196</td>
<td>750</td>
</tr>
<tr>
<td>NM</td>
<td>722</td>
<td>500</td>
</tr>
<tr>
<td>FN</td>
<td>797</td>
<td>50</td>
</tr>
<tr>
<td>NF</td>
<td>2462</td>
<td>0</td>
</tr>
</tbody>
</table>

MM [msd="Ncmpn"] [word="i"] [msd="Ncmpn"] 27896 (128.7 per million)  
FF [msd="Ncfpn"] [word="i"] [msd="Ncfpn"] 10685 (49.3 per million)  
NN [msd="Ncnpn"] [word="i"] [msd="Ncnpn"] 23 (0.1 per million)  
MF [msd="Ncmpn"] [word="i"] [msd="Ncfpn"] 5667 (26.1 per million)  
MN [msd="Ncmpn"] [word="i"] [msd="Ncnpn"] 700 (3.2 per million)  
FM [msd="Ncfpn"] [word="i"] [msd="Ncmpn"] 4196 (19.4 per million)  
NM [msd="Ncnpn"] [word="i"] [msd="Ncmpn"] 722 (3.3 per million)  
FN [msd="Ncfpn"] [word="i"] [msd="Ncnpn"] 797 (3.7 per million)  
NF [msd="Ncnpn"] [word="i"] [msd="Ncfpn"] 2462 (11.4 per million)
H1: WORD ORDER AS A PREDICTOR – SV vs. VS

Fixed effects:

| Estimate | Std. Error | z value | Pr(>|z|) |
|----------|------------|---------|---------|
| (Intercept) | 4.06975    | 0.55642 | 7.314   | 2.59e-13*** |
| wordorderVS | -1.43660   | 0.05293 | -27.141 | < 2e-16***  |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**H1: Preverbal conjoined structures** open richer array of possibilities for agreement.
NEXT EXPERIMENT ALL: GRAMMATICALITY JUDGMENT
NEXT EXPERIMENT ALL: GRAMMATICALITY JUDGMENT

• Conjoined SV and VS and non-conjoined
• BOTH GOOD
  • Jaja i salate su naručena u sobu.
  • Npl Fpl Aux Npl FCA
  • U sobu su naručena jaja i salate.
  • Aux Npl Npl Fpl CCA
• BOTH BAD
  • Muzeji i galerije su otvorena za javnost.
  • Mpl Fpl Aux Npl
  • Za javnost su otvorena muzeji i galerije.
  • Aux Npl Mpl Fpl
• GOOD / BAD
  • Sjednice i vijeća su pomaknuta na petak.
  • Fpl Npl Aux Npl CCA
  • Na petak su pomaknuta sjednice i vijeća.
  • Aux Npl Fpl Npl neither FCA, neither CCA
• BAD / GOOD
  • Jedrilice i gliseri su uplovile u zaljev.
  • Fpl Mpl Aux Fpl FCA, but quite rare
  • U zaljev su uplovile jedrilice i gliseri.
  • Aux Fpl Fpl Mpl CCA
NEXT EXPERIMENT ZAGREB: NON-WORDS

- **GOAL:** To rule out semantic effects and effects of familiarity, frequency, neighborhood density that might play a role in previous experiment.

- **TYPE:** Controlled psycholinguistic self-paced reading experiment with measured reaction time and latency.

- **LINGUISTIC MATERIAL:**
  - Prime sentences with the same word-type (Msg) and stimuli consisting of non-words with attached morphological markers (Mpl, Fpl, Npl).
  - Only relevant combinations will be tested (MF, MN, FM, FN, NM, NF).
  - 10 experimental examples for each type
  - Plan to split participants into two groups – counterbalance data distribution
NEXT EXPERIMENT ZAGREB: NON-WORDS

• MATERIALS:
  • Plural forms
  • Synchretism
  • Long adverbs (3+ sillables)

• Reban je odjednom izgorio.
  • Brukani i polake su odjednom.
• Dubak je nezavisno stisnut.
  • Bobini i fredana su nezavisno
• Znalek je izuzetno očišćen.
  • Ramine i bubini su izuzetno