Horn strategies and optimization in Russian aspect

Atle Grønn
University of Oslo
Language, Games & Evolution
ESSLLI 2007
Outline

- Partial blocking in Bidirectional OT (and GT)
- A diachronic view on Russian aspect
- A "diachronic" explanation of aspectual competition and implicatures in the synchronic system
Partial blocking in the lexicon

- "kill" (f1) > "cause to die" (f2)
- direct killing (m1) > indirect killing (m2)

(1) Black Bart caused the sheriff to die.
Partial blocking (Horn strategies)

- Graphical representation (Dekker & van Rooy)
- \( <f1,m1> \) is (weakly) optimal
- \( <f2,m2> \) is weakly optimal
Definition 1
Bidirectional Optimality

A form-meaning pair \(\langle F, M \rangle\) is weakly optimal iff

1. \(\langle F, M \rangle\) is a member of GEN.
2. There is no pair \(\langle F', M \rangle\) in GEN such that
   (2.1) \(\langle F', M \rangle > \langle F, M \rangle\) and
   (2.2) \(\langle F', M \rangle\) is weakly optimal.
3. There is no pair \(\langle F, M' \rangle\) in GEN such that
   (3.1) \(\langle F, M' \rangle > \langle F, M \rangle\) and
   (3.2) \(\langle F, M' \rangle\) is weakly optimal.
Russian aspect

Every Russian verb form is

- Perfective (Pf) or
- Imperfective (Ipf)
Key notions in Slavic aspectology

- markedness
- competition
- underspecification
- context sensitivity
- pragmatic implicatures
Truth-conditional approaches to Russian aspect fail because they don’t consider alternative forms which the speaker could have used.

The competition perspective suggests an OT/GT analysis.

Claim
The perfective ("complete event interpretation"): $e \subseteq t$
$$[[Pf]] = \lambda P \lambda t \exists e[P(e) \land eCt]$$

The progressive imperfective ("incomplete event interpretation"): $t \subseteq e$
$$[[lpf\_prog]] = \lambda P \lambda t \exists e[P(e) \land tCe]$$
The imperfective aspect (progressive reading)

(2)  *Kogda my vstretilis’,*  
when we met_PAST.PF  
*on chital "Vojnu i Mir".*  
he read_PAST.IPF “War and Peace”

“When we met, he was reading “W&P”.”

- i.e. *the time of our meeting is temporally included in an event of him reading W&P*
Aspectual competition

(3) 

Ja chital "Vojnu i Mir" v šestom klasse,
I read_PAST.IPF “W&P” in sixth grade
prochital polnost’ju za 6 dnej. (internet)
read_PAST.PF completely in 6 days

“I read “W&P” in the 6th grade, read it through in 6 days.”
Aspectual competition (cont.)

- The perfective *prochital* in (3):
  *the event of reading W&P is temporally included in an interval of 6 days: e C t*

- The imperfective in (3):
  *the event of reading W&P is temporally included in an interval of 1 year: e C t*
The emergence of aspect

- Many synchronic semantic and syntactic facts can be analyzed from an evolutionary perspective as *frozen pragmatics* (Blutner 2006).
Towards a 1-1 mapping (Church Slavonic/Old Russian)

- atelic activities ↔ chitaj’ (to read)
- progressive accomplishments
  (incomplete events) ↔ chitaj’ pis’mo (to read the letter)
- non-progressive accomplishments
  (complete events) ↔ prochitaj’ pis’mo (to ”through-read” the letter)
Partial blocking in BiOT requires underspecification

Was prochitat’ pis’mo (”through-read” the letter) originally compatible with an incomplete event interpretation?

(4) Als ich den Brief durchlas, den meine Freundin Katja für ihre Tochter Anna schrieb, musste ich weinen. (google)
“As I was reading the letter (lit.: `the letter through-read’), which my friend Katja had written to her daughter, I had to cry”
The pre-aspectual Russian system as a 2x2 game

- $f_1 = \text{(in)transitive simplex V}$
- $f_2 = \text{prep + V}$
- $m_1 = \text{incomplete events}$
- $m_2 = \text{complete events}$
- $\text{GEN} = M \times F$
- Rankings: $f_1 > f_2$ (complexity);
  $m_1 > m_2$ (stereotypical)
Partial blocking in the pre-aspectual system of Russian

<table>
<thead>
<tr>
<th></th>
<th>Incomplete event (m1)</th>
<th>Complete event (m2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplex V (f1)</td>
<td>🌿 ← •</td>
<td></td>
</tr>
<tr>
<td>Prep + V (f2)</td>
<td>• ← 🌿</td>
<td></td>
</tr>
</tbody>
</table>
Given a situation corresponding to $m_1$, the preferred form of $S$ will be $f_1$, hence the alternative $<f_2,m_1>$ is \textit{blocked}.

Given the input form $f_1$, $H$ will choose the interpretation $m_1$, hence the alternative $<f_1,m_2>$ is also blocked.

The pair $<lpf,m_1>$ is considered optimal from both perspectives.

Remove the pairs which are blocked.

The only remaining pair $<f_2,m_2>$, itself not being blocked, is \textit{weakly} optimal.
Explaining partial blocking in terms of associative learning (Benz 2006)

- At stage 1, the 1-1 mapping between intransitive simplex verbs and atelic activities is the external factor which triggers f1 to be associated with incomplete events (progressivity, the subinterval property etc.).
- Through associative learning, the pair <f1,m1> gets strengthened at stage 2.
If the speaker then, at stage 3, wants to emphasize the complete event interpretation, he should choose the marked form f2.

At stage 4, this invites a strengthening of the pair <f2,m2>.

Finally, at stage 5, prefixation develops into perfectivization, giving rise to a new aspectual system.
Grammaticalization of Pf ...

- Pragmatics → frozen pragmatics → semantics
- GEN = M X F – {<Pf, tCe>}
- Secondary imperfectivization (16\textsuperscript{th} century): chitat‘_IPF (f1) > prochitat‘_PF (f2) > prochityvat‘_IPF (f3)
- Complexity of form does not produce a linear ranking of Pf and Ipf!
Deblocking in modern Russian

- Not obvious that the new system can be modeled as a game
- A 2x2 game with three solutions, preserving ambiguity, cannot be modeled in terms of bidirectional optimization (two-solutions games)…
Modern Russian (cont.)

- What are the rankings in modern Russian?
- How and when does the complete event reading of Ipf survive?
An excursus: Deblocking and Conceptual grinding

<table>
<thead>
<tr>
<th>&quot;cow&quot;</th>
<th>&quot;beef&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>animal</td>
<td>cow-meat</td>
</tr>
<tr>
<td>🐄</td>
<td>🐄</td>
</tr>
</tbody>
</table>

Hindus are not allowed to eat cow.
Three classes of contexts

Basic idea:

- Aspect is a temporal phenomenon
- Global reasoning – without compositionality

- Class 1: Small reference time
- Class 2: Big reference time
- Class 3: No reference time
Class 1: small reference time

Example:
\[ f_1 = \text{lpf+past+punctual\_temporal\_adverbial} \]
\[ f_2 = \text{Pf+past+punctual\_temporal\_adverbial} \]

\[ m_1 = \text{tCe} \land Tadv(t) \]
\[ m_2 = \text{eCt} \land t = \text{the interval preceding Tadv} \]
Class 1: Progressive vs. "past perfect" interpretation

(2’) Kogda my vstrelis’, when we met_PAST.PF on ?? "Vojnu i Mir". he ?? “War and Peace”

- chital_IPF ("was reading/had read") or
- prochital_Pf ("had read")?
Ranking in class 1

- Ranking on F (harmony, frequency, salience...):
  \[ f_1 > f_2 \]

- Ranking on M (stereotypical interpretation):
  \[ m_1 > m_2 \]
A complete event interpretation is not available for Ipf whenever a progressive interpretation is possible.
(2’’) … on uže pročital "Vojnu i Mir".

… he already read_PAST.PF “W&P”

“(when we met), he had already read “W&P””

- <lpf_when_clause, “the time of meeting C e>}
- <Pf_when_clause, e C “the whole past preceding the time of our meeting”>
Class 2: big reference time

Example:

- $f_1 = lpf + \text{past}$
- $f_2 = Pf + \text{past}$

- $m_1 = eCt \land t = \text{the whole interval preceding the utterance time.}$
- $m_2 = ??$
Ranking in class 2

- Ranking on F (harmony, frequency, salience...):
  \[ f_2 > f_1 \]

- Ranking on M (stereotypical interpretation):
  \[ m_1 > m_2 \]
# Class 2: Partial blocking

<table>
<thead>
<tr>
<th>f2 (Pf_past)</th>
<th>m1 (eCt ∧ ( t = \text{past} ))</th>
<th>m2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \odot )</td>
<td>( \bullet )</td>
</tr>
<tr>
<td></td>
<td>( \uparrow )</td>
<td>( \uparrow )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f1 (lpf_past)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \bullet )</td>
</tr>
<tr>
<td></td>
<td>( \leftarrow )</td>
</tr>
<tr>
<td></td>
<td>( \odot )</td>
</tr>
</tbody>
</table>
Strengthening of Pf in class 2

By associative learning (Benz 2006) it is expected that the interpretation of Pf gets strengthened to include an implicature of the *current relevance of the result state*.
Strengthening of Pf in class 2 (cont.)

(5) *Kto otkryl okno?*  
who open_PAST.PF window.

“Who has opened the window?”
*(the window is currently open)*
From deblocking to partial blocking

- What about m2?
- “The unemployed form may soon find a new job, generally expressing something closely related to but subtly different from the canonical interpretation that one might have expected” (Beaver & Lee, 2003:140).
Ipf and the convention of annulled result

(5’)  *Kto otkryval okno?*  
who open_PAST.IPF window.

“Who had the window open?”  
*(the window is currently closed)*
## Strategies in game theory

<table>
<thead>
<tr>
<th></th>
<th>open</th>
<th>closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Pf</td>
<td>Ipf</td>
</tr>
<tr>
<td>S2</td>
<td>Ipf</td>
<td>Pf</td>
</tr>
<tr>
<td>S3</td>
<td>Pf</td>
<td>Pf</td>
</tr>
<tr>
<td>S4</td>
<td>Ipf</td>
<td>Ipf</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Pf</th>
<th>Ipf</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>open</td>
<td>closed</td>
</tr>
<tr>
<td>H2</td>
<td>closed</td>
<td>open</td>
</tr>
<tr>
<td>H3</td>
<td>open</td>
<td>open</td>
</tr>
<tr>
<td>H4</td>
<td>closed</td>
<td>closed</td>
</tr>
</tbody>
</table>
Determine expected utility for each profile \((S,H)\) given the hearer’s probability function:

\[
EU(S,H) = \sum_{m} P(m) \times U(m,S,H)
\]

Utility function of successful communication:

\[
U(m,S,H) = \begin{cases} 
1, & \text{if } H(S(m)) = m \\ 
0, & \text{otherwise}
\end{cases}
\]
P. Parikh’s GT-model (example cont.)

Assumptions (part of the common ground):
- Probability of window being open at the utterance time: 0.6
- Probability of window being closed at the utterance time: 0.4
Combining strategies (without complexity of form)

<table>
<thead>
<tr>
<th>Partial</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1</td>
<td>0</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>S2</td>
<td>0</td>
<td>1</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>S3</td>
<td>0.6</td>
<td>0.4</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>S4</td>
<td>0.4</td>
<td>0.6</td>
<td>0.6</td>
<td>0.4</td>
</tr>
</tbody>
</table>
P. Parikh’s GT-model (cont.)

- $U(m,S,H) = 1/\text{Complexity}(S(m)),$
  - if $H(S(m)) = m$
  - $= 0$ otherwise

Assumptions (part of the common ground):
- Complexity of "Pf+past": 1
- Complexity of "lpf+past": 2
## Combining strategies (with complexity of form)

<table>
<thead>
<tr>
<th>Partial</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.8</td>
<td>0</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>S2</td>
<td>0</td>
<td>0.7</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>S3</td>
<td>0.6</td>
<td>0.2</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>S4</td>
<td>0.2</td>
<td>0.6</td>
<td>0.6</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Blutner vs. P. Parikh

- \(<S1,H1>\) (Horn) and \(<S2,H2>\) (anti-Horn) are strict Nash equilibria.

- \(<S1,H1>\) is Pareto optimal (highest expected utility) – Parikh’s unique solution.

- Blutner’s BiOT and Parikh’s GT model both account for partial blocking in these simple cases.
Class 3: no reference time

Example: imperatives under negation

- \( f_1 = \text{lpf\_neg\_imperative} \)
- \( f_2 = \text{Pf\_neg\_imperative} \)

- \( m_1 = S \text{ wants } H \text{ not to perform action } a \)
- \( m_2 = S \text{ warns } H \text{ against accidentally performing action } a \)
Ranking in class 3

- Ranking on F (harmony, frequency, salience…):
  \[ f_1 > f_2 \]

- Ranking on M (stereotypical interpretation):
  \[ m_1 > m_2 \]
Class 3
(unmarked form/meaning)

(6) *Ty, požalujsta, ne opazdyvaj.*

you please not be _late_ IMP.IPF

“Please don’t be late.” (internet)
Class 3
(marketed form/meaning)

(6’) [Čerez 10 minut budet uže pozdno.]
Smotri, ne opazdaj!
lookIMP.IP not become_late IMP.PF

“In 10 minutes it’ll already be too late. Be careful not to be late.” (internet)
Partial blocking in class 3

<table>
<thead>
<tr>
<th></th>
<th>m1</th>
<th>m2</th>
</tr>
</thead>
<tbody>
<tr>
<td>f1 (lpf_imp)</td>
<td>![Symbol]</td>
<td>![Symbol]</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>f2 (Pf_imp)</td>
<td>![Symbol]</td>
<td>![Symbol]</td>
</tr>
</tbody>
</table>
Conclusion

- GT and BiOT approaches can provide new insights into aspectual competition in Russian
- A principled explanation for various implicatures associated with Ipf and Pf
- The temporal dimension of aspect plays a crucial role in the rankings of F and M.
- GT, BiOT or associative learning?
References

References (cont.)


References (cont.)