

# Horn strategies and optimization in Russian aspect\*

Atle Grønn  
Dept. of European Languages (ILOS)  
University of Oslo

atle.gronn@ilos.uio.no

## Abstract

I propose to explain the diachronic development of the aspectual system in Russian in terms of Horn strategies (partial blocking). The analysis is based on Blutner's bidirectional OT, which has a strong diachronic dimension. I also suggest that a context-sensitive version of bidirectional OT can play a role in explaining the synchronic situation for Russian aspect.

## 1 Horn strategies in the lexicon

Standard examples of Horn's iconicity principle – (un)marked forms are mapped with (un)marked meanings – come from lexical pragmatics (Horn, 1984), (Blutner, 1998). In the spirit of Blutner (2000) and his implementation of partial blocking in bidirectional OT, Dekker and van Rooy (2000) came up with graphical representations like the following:

	m <sub>1</sub>		m <sub>2</sub>
f <sub>1</sub>	●	←	○
	↑		↑
f <sub>2</sub>	○	←	●

Table 1: (Weakly) optimal pairs in a 2x2 game

The speaker – i.e. the vertical arrows – has a preference for short, unmarked forms (e.g. ‘kill’ (f<sub>1</sub>) > ‘cause to die’ (f<sub>2</sub>)) and the hearer – i.e. the horizontal arrows – prefers stereotypical, unmarked meanings (e.g. direct killing (m<sub>1</sub>) > indirect killing (m<sub>2</sub>)). Horn's division of pragmatic labor predicts that from two forms with underspecified meanings, the simple form ‘kill’ is mapped with the meaning of direct/canonical killing, while the more complex form ‘cause to die’ will be used to denote an indirect killing. The latter is demonstrated in (1), which suggests that the sheriff's death came about after some unusual event, perhaps an accident.

- (1) Black Bart caused the sheriff to die.

The algorithm of weakly bidirectional OT (Jäger, 2002) starts with the optimal pair <f<sub>1</sub>,m<sub>1</sub>> (●). Both pairs which are beaten by this pair in *one* direction (○) are then removed from the tableau, and this leaves us with the *weakly* optimal pair <f<sub>2</sub>,m<sub>2</sub>> (●). The weakly optimal

---

\*I would like to thank Hanne Eckhoff and Dag Haug for their comments on the diachronic data.

pair survives despite the presence of the optimal pair. True, there is a better form ( $f_1$ ), but *not* given meaning  $m_2$ . Similarly, there is a better meaning ( $m_1$ ), but *not* given form  $f_2$ .

The Horn strategy (partial blocking) is arguably rational for the speaker and hearer, and represents the optimal way of resolving two conflicting interests in communication: economy (the speaker's I/R principle in neo-Gricean pragmatics) vs. diversification (the hearer's Q-principle). In this paper, I will discuss whether this kind of rationality plays any role in the development of a grammatical category: viewpoint aspect in Russian, that is, the opposition between imperfective aspect (IPF) and perfective aspect (PF). Aspect in Russian is morphologically independent of tense and finiteness and occurs obligatorily in all tense/mood configurations. In this respect, Russian aspect is a lexical phenomenon.

## 2 The meaning of aspect

There is a close relationship between (a)telicity, i.e. lexical aspect, and grammatical aspect in Russian. PF is indeed almost exclusively restricted to telic predicates (Vendlerian accomplishments/achievements). At the same time, the Russian IPF:PF distinction, like the French aspectually loaded past tenses Imparfait:Passé simple, is semantically speaking a viewpoint aspect. For instance, accomplishments can be denoted by imperfective verbs with a progressive interpretation (giving rise to the famous imperfective paradox). For the purposes of the following discussion, I distinguish between three meanings, which form a continuum:

- $m_0$ : atelic activities
- $m_1$ : (telic accomplishments, where) the speaker's viewpoint is restricted to the preparatory process
- $m_2$ : (telic accomplishments, where) the speaker's viewpoint includes the end point of the event

For short, I will refer to  $m_0$  and  $m_1$  – which both, in a certain sense, satisfy the subinterval property – as *incomplete event interpretations*, while  $m_2$  represents a *complete event interpretation*. Thus, so far, we have two forms, IPF and PF, which are to be mapped with three meanings  $m_0$ ,  $m_1$  and  $m_2$ . But let's start with the beginning.

## 3 The emergence of aspect

Bidirectional OT is a powerful explanatory principle in diachronic linguistics, since pragmatic bidirectionality creates special interpretations that can become conventionalized (Blutner and Zeevat, 2004). Many synchronic semantic and syntactic facts can be analyzed from an evolutionary perspective as 'frozen pragmatics' (Blutner, 2006). This strong diachronic dimension of BiOT suggests that we should start with the emergence of the aspectual system itself, i.e. the PF:IPF opposition which came to replace the old Indo-European tenses Aorist:Imperfect.

The earliest Russian texts (Old Church Slavonic, Old Russian) contain morphologically simplex verbs, such as *čitat'* – *to read*. Used intransitively, simplex verbs ( $f_0$ ) denote atelic

activities. In its transitive use, with a quantized object, the VP ( $f_1$ ) denotes an accomplishment, just like its Germanic counterparts: *čitat' pis'mo* – *to read the letter*. However, accomplishments can also be referred to by a prefixed verb ( $f_2$ ), which in this case is formed by adding the preposition/prefix *pro* – *through* to the simplex verb: *pročitat' pis'mo* – *to read through the letter* (lit. ‘through-read’). In hindsight, we know, of course, that the emerging form-meaning pairs are the following:

$\text{čitat' } (f_0)$	$\Leftrightarrow$	atelic activities ( $m_0$ )
$\text{čitat' pis'mo } (f_1)$	$\Leftrightarrow$	‘progressive’ accomplishments ( $m_1$ )
$\text{pročitat' pis'mo } (f_2)$	$\Leftrightarrow$	‘non-progressive’ accomplishments ( $m_2$ )

Table 2: The interpretation of the Russian VPs ‘read (through) (the letter)’

Is this 1-1 mapping a result of bidirectional optimization?

In terms of complexity of forms, the following ranking suggests itself:  $f_0 > f_1 > f_2$ . Since  $f_1$  and  $f_2$  are obviously excluded from being paired with  $m_0$ , and  $m_0$  is presumably the simplest meaning from a conceptual point of view, the pairing  $\langle f_0, m_0 \rangle$  is straightforward semantics and requires no further pragmatic/optimalty theoretic reasoning. The question is whether we can find a principled explanation for why  $f_1$  is mapped to  $m_1$ , and  $f_2$  to  $m_2$ .

### 3.1 Underspecification

In BiOT, a prerequisite for partial blocking to obtain, is that the candidate forms start out with an underspecified semantics (but see (Benz, 2006) for a critical discussion of this point). I assume here that both  $f_1$  and  $f_2$  were originally compatible with both  $m_1$  and  $m_2$ . The pair  $\langle f_1, m_2 \rangle$  is still viable in contemporary Russian, where it is known as the *factual IPF* (Grønn, 2004):

- (2) Ne vyzyvaet somnenij i to, čto Stalin sam **čital**<sup>IPF</sup> **pis'mo** Bulgakova. (google)

There is also no doubt that Stalin himself **read the letter** from Bulgakov.

On the other hand, after the process whereby prefixation turned into perfectivization (PF), the pair  $\langle f_2, m_1 \rangle$  was no longer grammatical in Russian. However, initially, there is no reason to believe that the Russian VP *pročitat' pis'mo* behaved differently from its direct counterpart in contemporary German:

- (3) Als ich **den Brief durchlas**, den meine Freundin Katja für ihre Tochter Anna schrieb, musste ich weinen. Dieser Brief hat mich so zu Tränen gerührt, ich konnte es kaum aushalten! (google)

As I **was reading the letter** (lit.: ‘the letter through-read’), which my friend Katja had written to her daughter Anna, I had to cry. This letter moved me to tears, such that I could hardly bear it!

Although the German VP in (3) typically denotes a complete event, this particular example shows that the prefixed transitive accomplishment predicate is compatible with an incomplete event interpretation. Why would the Russian equivalent *pročitat' pis'mo* develop

into perfective aspect, i.e. loose its compatibility with an incomplete event interpretation? There are two conceivable lines of explanation – the Gricean and the neo-Gricean, which I will briefly present below.

### 3.2 Gricean reasoning

It has been argued that telic predicates by default have a ‘perfective’, complete event interpretation (Bohnemeyer and Swift, 2004). This gives rise to an implicature, which, at **stage 1**, can be exploited by the speaker, who relies on the hearer to infer the complete event interpretation of *pročítat’ pis’mo*. If communication is successful (repeatedly), the implicature can turn into an entailment at **stage 2**.<sup>1</sup>

This is Gricean, but not neo-Gricean reasoning. Nothing here hinges on the presence of  $f_1$ . The BiOT perspective could still play a role, but only at a later stage: Given the optimal pair  $\langle f_2, m_2 \rangle$ , the pair  $\langle f_1, m_1 \rangle$  becomes weakly optimal at **stage 3**. One possible problem with this kind of analysis is that it blurs the intuitive relationship between  $f_0$  and  $f_1$ . If there is no link between  $f_0$  and  $f_1$ , and no interaction between  $f_1$  and  $f_2$  at stage 1 and 2, one may ask why the telic VP  $f_1$  did not develop into a purely perfective marker.

On the other hand, if we assume that  $f_1$  indeed inherits a feature, say [+subinterval property], from its close relative  $f_0$ , then we can apply the Gricean reasoning once more and abandon the OT perspective:  $f_1$  is prototypically used with the implicature of  $m_1$ , which develops into an entailment. However, this is not satisfactory for the simple reason that IPF is not a grammaticalization of atelicity and/or the subinterval property, and  $f_1$  does not entail  $m_1$  (cf. example (2) above).

### 3.3 Neo-Gricean reasoning

Let’s now have a look at a neo-Gricean approach, where the diachronic change of *pročítat’ pis’mo* from a complete event implicature to an entailment is explained by looking at the prefixed verb not in isolation but in the competitive environment of forms and meanings.

According to (Benz, 2006), Horn’s derivation scheme proceeds roughly as follows:

1. The speaker used a marked expression  $f'$  containing ‘extra’ material ... when a corresponding unmarked expression  $f$ , essentially coextensive with it, was available.
2. The ‘extra’ material must have been necessary, i.e.  $f$  could not have been appropriately used.
3. ... Therefore, the unmarked alternative  $f$  tends to become associated (by use or – through conventionalization – by meaning) with the unmarked situation  $m$ .
4. The marked alternative  $f'$  tends to be associated with the complement of  $m$  with respect to the original extension of  $f/f'$  ...

Benz points out some problems with Horn’s use of his own principle. Most importantly, it remains unclear why  $f'$  had to be used in the first place, given that  $f$  and  $f'$  were essentially

---

<sup>1</sup>Thanks to Dag Haug on this point.

coextensive. Benz argues that the procedure should start with the unmarked form being associated with the stereotypical situation  $m$  through *learning* – producing the optimal pair  $\langle f, m \rangle$  at stage 1. In order to achieve this, Benz introduces associative learning as a third component (besides Zipf’s principles of economy and diversification) in diachronic BiOT. Then, at the next stage,  $f'$  is paired with  $m'$  (weakly optimal pair).

Back to Russian aspect, there is one complication: why should  $m_1$  be ranked as more stereotypical/normal than  $m_2$ ? In fact, the above-mentioned generalization in (Bohnemeyer and Swift, 2004) suggests the opposite ranking for accomplishments:  $m_2 > m_1$ . If we keep the straightforward ranking on forms suggested by morphological complexity,  $f_1 > f_2$ , this seems to imply that Russian aspect develops from an anti-Horn strategy:

	$m_2$		$m_1$
$f_1$	○	←	●
	↑		↑
$f_2$	●	←	○

Table 3: Anti-Horn strategy

The two winning form-meaning combinations in table 3 are marked with ●. Although the anti-Horn strategy is evolutionary stable (van Rooy, 2004a), it is hardly attested in natural language. This is not surprising since this equilibrium is clearly more costly than the Horn strategy, where the shortest form is associated with the most frequent meaning. I think the problem with the above line of reasoning – which leads to the anti-Horn strategy – is the tacit assumption that competition at this stage is restricted to telic predicates with their progressive/non-progressive interpretations. A more appropriate ranking on forms and meanings would be the following, where we do not distinguish between  $f_0/f_1$  and  $m_0/m_1$ :

- (in)transitive simplex verbs > prefixed verbs
- incomplete event interpretations > complete event interpretations

The outcome is now the expected Horn strategy in table 4:

	$m_{0,1}$		$m_2$
$f_{0,1}$	●	←	○
	↑		↑
$f_2$	○	←	●

Table 4: Horn strategy for Russian aspect

Also from the point of view of associative learning (Benz, 2006), table 4 makes sense: At **stage 1**, the form  $f_1$  is not perceived as aspectually different from  $f_0$ . The 1-1 mapping between  $f_0$  and atelic activities is the external factor which triggers  $f_1$  to be associated with incomplete events (progressivity, the subinterval property etc.). Through associative learning, the pair  $\langle f_1, m_1 \rangle$  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  $f_2$ . At **stage 4**, this invites a strengthening of the pair  $\langle f_2, m_2 \rangle$ . Finally, at **stage 5**, prefixation develops into *perfectivization*, giving rise to a new aspectual system.

#### 4 Markedness and secondary imperfectivization

As a result of the first round optimization described above, PF is grammaticalized. This changes the status of *pročitat' pis'mo*, which is no longer underspecified, but gets a uniform, invariant semantics:  $\langle \text{PF}, \text{complete event} \rangle$ . In BiOT, perfectivization leads to a reduction of GEN:

$$\text{GEN} = \text{F} \times \text{M} - \{ \langle \text{PF} (f_2), \text{incomplete event} (m_1) \rangle \}$$

This creates a new situation. Even if we consider the emergence of the aspectual system to be bidirectionally optimal, it is far from clear that BiOT can play any role *after* grammaticalization.<sup>2</sup> What happens to the status of  $f_2$  (now: PF) in terms of ranking on forms? From the application of BiOT in lexical pragmatics, we are used to treating lexicalized items as unmarked forms (e.g. 'kill'). Indeed, markedness theory is tricky (and controversial) in the domain of Russian aspect. I think we have to distinguish between a broad perspective and a narrow perspective: IPF is unmarked in the grammar of Russian (broad perspective), but PF is the default, most salient and frequent choice in the domain of accomplishments/achievements (narrow perspective).

The grammaticalization of PF creates a series of morphological gaps in the verbal paradigms. To take one example: the imperfective simplex verb *kryt'* – *to cover* can be prefixed with *ot*, which has the basic meaning of motion/action away from a given point. The result is the perfective accomplishment verb *otkryt'* – *to open (uncover)*. Clearly, in this case, the prefixed verb does not form an aspectual correlate to the simplex verb: their lexical semantics are not compatible. This gap is filled at **stage 6** by the productive morphological device of secondary imperfective suffixes, which in this case produces the following aspectual pair: *otkryt'<sup>PF</sup>* – *otkryvat'<sup>IPF</sup>*.

Note that morphological complexity cannot any longer be the crucial factor for ranking of forms at stage 6 since this would not produce a linearly ordered ranking of PF and IPF: *čitat'<sup>IPF</sup>* ( $f_0$ ) > *pročitat'<sup>PF</sup>* ( $f_2$ ) > *pročityvat'<sup>IPF</sup>* ( $f_3$ ).

#### 5 Deblocking and context-sensitivity in modern Russian

In the optimality theoretic reasoning in the previous section, the pair  $\langle \text{IPF}, \text{incomplete event} \rangle$  came out optimal or weakly optimal, depending on the choice of forms/meanings under consideration and the corresponding constraints and rankings. Is this situation confirmed by the synchronic data?

As argued in (Grønn, to appear), a complete event interpretation is never available for IPF in a context which licenses a progressive interpretation. Hence, the possibility of an incomplete event interpretation will effectively block the factual IPF. In examples like (2) above, the complete event interpretation of IPF is *deblocked* since no misunderstanding can

<sup>2</sup>van Rooy (2004b) notes some shortcomings of BiOT which in principle could be problematic for an application to Russian aspect. Consider the following situation:  $f$  is a lighter expression than  $f'$ ,  $f > f'$ ; and  $m'$  is more stereotypical than  $m$ ,  $m' > m$ . If the meaning of  $f$  is underspecified, while  $f'$  can only mean  $m'$ , BiOT predicts that  $m$  cannot be expressed. It turns out that the pair  $\langle f, m \rangle$  is *not* weakly optimal. van Rooy argues that certain game-theoretical approaches to communication are better suited to handle such cases.

occur: The sentence and utterance context of (2) do not provide any reference times which could be used to ‘zoom in’ on the event. The interaction of past tense and aspect produces a temporal configuration, whereby the complete event is located at some proper subinterval of the whole past preceding the utterance time.<sup>3</sup>

Consider also the following example of aspectual competition and deblocking of the complete event interpretation of IPF:

- (4) Kto  $\left\{ \begin{array}{l} \text{otkryl}^{PF} \\ \text{otkryval}^{IPF} \end{array} \right\}$  okno?  
Who **opened** the window?

PF is the default and expected choice in (4). Importantly, however, the use of IPF in (4) still gives rise to a complete event interpretation since no competing incomplete/progressive interpretation is available for similar reasons as in (2). What we observe in these contexts, is in fact a second round of partial blocking following the deblocking of the factual IPF. If we isolate contexts which only allow for complete event interpretations<sup>4</sup>, the partial blocking in table 5 emerges:

	canonical complete events		non-canonical complete events
PF	●	←	○
	↑		↑
IPF	○	←	●

Table 5: Partial blocking in contexts of deblocking of the factual IPF

A canonical complete event interpretation is arguably a complete event which produces a relevant *result*. Thus, if the window is open at the utterance time, the speaker should use PF in (4). As a result of pragmatic strengthening and polarization, the factual IPF can have any complete event interpretation where the result state is not relevant or cancelled. A nice example is the use of IPF in (4), from which the hearer typically infers that the window is closed at the utterance time.

## 6 Conclusion

This paper is a first step towards an analysis of Russian aspect in bidirectional OT. The emphasis here has been on the diachronic process which leads to the grammaticalization of perfective aspect. Many issues are still open, such as the ranking on forms and meanings at various stages, and the question as to which forms/meanings should be considered competitors in each round of optimization. The synchronic situation also requires further research, notably concerning deblocking of the complete event interpretation of imperfective aspect.

The phenomenon of partial blocking raises the issue of diachronic vs. synchronic explanations. The two approaches may coexist in BiOT, as argued recently by Blutner (2006). A

<sup>3</sup>See (Grønn, 2004) for details.

<sup>4</sup>In order to model this properly we would need a context-sensitive version of BiOT, perhaps along the lines of (Benz, 2001). See also (Grønn, to appear) for some suggestions and discussion.

difference is still worth pointing out: When the Horn strategy applies in cases of deblocking, the coordination game of the speaker and hearer does not seem to be fully conventionalized. For instance, in (4), different non-canonical complete event interpretations of IPF may be possible. On the contrary, the division of labor which brought about perfective aspect became completely grammaticalized.

## References

- Anton Benz. Towards a framework for bidirectional optimality theory in dynamic contexts, *ms.*, 2001.
- Anton Benz. Partial blocking and associative learning. *Linguistics and Philosophy*, 29: 587–615, 2006.
- Reinhard Blutner. Lexical pragmatics. *Journal of Semantics*, 15:115–62, 1998.
- Reinhard Blutner. Some aspects of optimality in natural language interpretation. *Journal of Semantics*, 17:189–216, 2000.
- Reinhard Blutner. Embedded implicatures and optimality theoretic pragmatics. In T. Solstad, A. Grønn and D. Haug, editors, *A Festschrift for Kjell Johan Sæbø*, pages 11–29. Oslo, 2006.
- Reinhard Blutner and Henk Zeevat. Editors’ introduction: Pragmatics and optimality theory. In R. Blutner and H. Zeevat, editors, *Optimality Theory and Pragmatics*, New York, 2004. Palgrave Macmillan.
- Jürgen Bohnemeyer and Mary Swift. Event realization and default aspect. *Linguistics and Philosophy*, 27(3):263–296, 2004.
- Paul Dekker and Robert van Rooy. Bi-directional optimality theory: An application of game theory. *Journal of Semantics*, 17:217–242, 2000.
- Atle Grønn. *The Semantics and Pragmatics of the Russian Factual Imperfective*, volume 199 of *Acta Humaniora*. Unipub, dr.art thesis, Oslo, 2004.
- Atle Grønn. Russian aspect as bidirectional optimization, to appear.
- Lawrence Horn. Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature. In D. Shiffrin, editor, *Meaning, Form, and Use in Context: Linguistic Applications*, pages 11–42. Georgetown University Press, Washington, 1984.
- Gerhard Jäger. Some notes on the formal properties of bidirectional optimality theory. *Journal of Logic, Language and Information*, 11(4):427–451, 2002.
- Robert van Rooy. Evolution of conventional meaning and conversational principles. *Synthese (Knowledge, Rationality and Action)*, 139:331–366, 2004a.
- Robert van Rooy. Signalling games select Horn strategies. *Linguistics and Philosophy*, 27: 493–527, 2004b.