Negative Russian perfectives both presuppose and evoke explanatory influences

Bridget Copley (SFL, CNRS/Paris 8) and
Olga Kagan (Ben-Gurion University of the Negev)

June 2023

Abstract

Russian negative past perfectives have unexpected construals not predicted by existing analyses. These construals evoke the existence of additional *influences*: eventualities that potentially affect the state of affairs. These influences seem to explain why the course of events was expected. We argue that the perfective itself contributes one such (presupposed) influence to the causal model, which can correspond either to the agent's intention or to a particular circumstance. Further, in the case of negative perfectives, we argue that one or more accommodated influences must be evoked to explain the otherwise inexplicable outcome that an expected result does not occur.

1 Introduction

As has been extensively discussed in the linguistic literature, Slavic perfective predicates, in contrast to their imperfective counterparts, convey that an event is bounded, temporally delimited, or maximalized (e.g. Forsyth, 1970; Smith, 1991; Krifka, 1992; Filip, 2000, 2008; Timberlake, 2004; Filip and Rothstein, 2006). For instance, Filip (2008) argues that perfective verbs contribute a covert maximalization operator on events MAX_E, which "picks out the unique largest event at a given situation" (p. 217). In the prototypical case, this will be an event that reaches its inherent natural endpoint, but this could also be an activity that is maximal with respect to some ordering criterion.

However, certain Russian data demonstrate that such concepts as telicity or temporal delimitation are insufficient for the purposes of capturing the entire distribution of past perfective predicates. Specifically, negative past perfectives evoke the existence of additional influences, seemingly to explain why the course of events was expected or feasible. We argue that the perfective itself contributes one (presupposed) additional influence, and one or more accommodated influences are evoked to explain otherwise inexplicable outcomes generated by the combination of negative and perfective semantics. To argue for this picture, we consider data discussed by Dickey (2000) on temporal definiteness in East

Slavic, as well as Russian data recently introduced by Kagan (2020), and use causal models (Pearl, 2000), which easily represent such inexplicable outcomes and multiple causal influences.

2 The perfective / imperfective distinction

In Russian, every verb is specified for aspect: perfective or imperfective. As pointed out above, perfective predicates are generally analyzed as telic, bounded or temporally delimited. For the sake of illustration, consider the example in (1):[1]

(1) Vasja pomyl mašinu.
Vasja washed.PERF car
'Vasja washed a / the car.'

This sentence entails that the event of washing the car reached its normal completion. The reported event is temporally delimited and maximal, in the sense that all of the car has been washed. The imperfective counterpart of this sentence, shown in (2), does not carry the same entailment.

(2) Vasja myl mašinu.
Vasja washed.IMP car
'Vasja washed / was washing / used to wash a / the car.'

Imperfective aspect is compatible with a wide range of readings. For instance, (2) may be interpreted as 'Vasja was washing the car' (a progressive reading) or as 'Vasja used to wash the car' (habitual reading). In addition, it can receive an interpretation analogous to the Experiential Perfect in English. Under this reading, the sentence roughly means that Vasja has had the experience of washing the car, in the sense that he has been engaged in an event of this kind at least once. This is the so-called Statement of Fact convention of usage (see e.g. Smith, 1991; Grønn, 2004; Altshuler, 2010, and references therein).

In this way, an imperfective verb can be appropriate even when the event predicate is telic and an event of the type denoted by this predicate is realized only once. For instance, a single completed car-washing event makes (2) true under the Statement of Fact reading. It thus follows that a single event that reaches its inherent natural endpoint can be described with either a perfective or an imperfective verb, and the question emerges whether and how the meaning of the resulting sentences differs. Such a comparison reveals the existence of pragmatic meaning components which systematically characterize sentences with perfective verbs and which should be set apart from concepts such as telicity, temporal delimitation, and boundedness.

[1] In this example and throughout the article, we specify in the glosses only those grammatical features that are relevant to the analysis at hand.

3 Perfective verbs: Event specificity and feasibility

3.1 Event Specificity

There exists a strong intuition that sentences with past perfective verbs report events that are in some intuitive sense specific. Consider, for example, the following minimal pair:

(3) a. Lena ne pozvonila.
Lena NEG phoned.PERF
'Lena hasn't phoned.'
b. Lena ne zvonila.
Lena NEG phoned.IMP
'Lena hasn't phoned.'

Intuitively, (3a) conveys that Lena did not call in a particular situation. Plausibly, she was expected to call, at a particular time or in order to discuss a particular topic, and this expected eventuality was not realized. In turn, (3b), which contains an imperfective verb, is much more neutral in this respect. It is compatible with a context in which a particular phoning event has been expected. But this is not a necessary condition. (3b) can be also understood as a neutral statement that there has not been any event of Lena calling the speaker. Such an event need not have been previously expected, scheduled or discussed.[2]

The same kind of contrast can be observed in interrogatives:

(4) a. Lena pozvonila?
Lena phoned.PERF
'Has Lena phoned? / Did Lena phone?'
b. Lena zvonila?
Lena phoned.IMP
'Has Lena phoned / Did Lena phone?'

[2]Of course, negative sentences even with imperfective verbs cannot be used entirely out of the blue: there must be a reason why the speaker chooses to report that something is NOT the case. Still, the restrictions on the use of perfective aspect are much stronger. For instance, consider the following minimal pair:

(i) a. Ja ne čitala *Vojnu i mir.*
I NEG read.IMP *War and Peace*
'I haven't read *War and Peace.*'
b. Ja ne pročitala *Vojnu i mir.*
I NEG read.PERF *War and Peace*
'I haven't read *War and Peace.*'

The first sentence, with an imperfective verb, can be uttered in the context of a discussion of Russian literature and one's familiarity with it. In contrast, the second sentence, which contains a perfective verb, is only appropriate in a more restricted context whereby the speaker has been assigned to read the book or has been expected to read it for another contextually specified reason.

(4a), similarly to (3a), presupposes a particular expected event. Lena has been supposed / expected to phone, the option that she would phone specifically today or at another reference time has been judged as plausible or has been previously discussed. The speaker asks whether this specific event has been instantiated. (4b) is less restricted in this respect. Of course, there must exist some reason for the speaker to ask this kind of question to begin with. But if Lena merely calls once in a while, and the speaker is curious whether such an event happened to have occurred today, this is sufficient to license (4b), in contrast to (4a).

The observation that the Russian imperfective constitutes the less restricted, more default-like aspect is not surprising. The same has been claimed about imperfectivity as far as its purely aspectual function is concerned. According to Jakobson (1984), perfective / imperfective in Russian is a binary opposition which contains one marked member (the perfective) and one unmarked member (the imperfective). It follows from this approach that the Russian perfective is associated with a particular semantic feature, whereas the Russian imperfective is a kind of elsewhere aspect. The same approach is argued for and formalized by Kagan (2008, 2010).

The event specificity meaning component accompanying perfective verbs is particularly striking in downward-entailing environments, but it can also be observed in simple affirmative declarative sentences, especially if these are compared to Statement of Fact imperfectives. Consider the following example:

(5) a. Ivan vstrečal Lenu.
Ivan met.IMP Lena
'Ivan has met with Lena (at least once)'.
b. Ivan vstretil Lenu
Ivan met.PERF Lena
'Ivan met Lena (under particular circumstances)'

Intuitively, (5a) under the relevant use asserts that the event kind described by *Ivan meet Lena* was instantiated at least once. This is a rather weak statement, which is made true by any event of Ivan meeting Lena. In contrast, (5b) reports a specific meeting, e.g., one that took place at a particular time or location.

The view that past perfectives contribute this kind of event specificity is supported by the fact that past perfective verbs strongly tend to be incompatible with temporal adjuncts such as *nikogda* 'never' (but see the special construction in (7) below), *kogda-libo* 'ever' and *kogda-nibud'* 'ever, some time', 'any time':

(6) a. *Dima nikogda ne vstretil Mašu.
Dima never NEG met.PERF Masha
'Dima has never met Masha.'
b. *Esli Dima kogda-libo/kogda-nibud' vstretil Mašu, on ejo
if Dima ever/ever met.PERF Masha he her
ne zabudet.
NEG forget.FUT

'If Dima has ever met Masha, he won't forget her.'

These sentences become grammatical if the perfective verbs are substituted by their imperfective counterparts. Here, inherently non-specific time (a product of the temporal adjuncts) does not allow for a specific event reading. Indeed, it has been argued that perfective aspect in East Slavic languages expresses temporal definiteness (Leinonen, 1982; Thelin, 1990, 1991; Dickey, 2018, 2000). Perfective event predicates have a specific flavor because they are mapped to definite / specific temporal intervals. Such approaches successfully capture the data in (3) - (6), including the pragmatic nuances discussed above.

Still, while a correlation between specific times and specific events is indeed quite strong, we can also find statements about specific events that lack a specific (or even fixed) temporal location. This is illustrated in (7) below. This sentence explicitly asserts that a particular phoning event which was expected, promised or scheduled, never took place, i.e., it did not take place at any time. It is thus possible to conceive of an (unrealized) specific event without linking it to a specific temporal location.[3]

(7) On tak nikogda ej i ne pozvonil.
he so never her and NEG called.PERF
'Ultimately/in the end, he never called her.'

Such examples suggest that perfectivity conveys event specificity, whereas temporal specificity (or definiteness) is a meaning component that often, but not always, arises when a specific event is reported. Prototypically, a specific event is one that takes place at a specific temporal interval. However, as (7) shows, this is not obligatory. We discuss the way to formalize the specificity intuition below. However, first, an additional meaning component associated with perfectivity has to be considered.

3.2 Feasibility and defeasibility

A further striking restriction on the use of the past perfective is exhibited in non-veridical environments, with negative contexts providing especially salient examples. Consider, for instance, the contrast in (8), also described by Kagan (2020). Suppose that Ivan is found murdered and Anna gets accused of the crime. She denies the guilt and says "I didn't kill Ivan". For this purpose, only imperfective aspect, as in (8a), is appropriate. Even though we are discussing a single, specific event whose description here is associated with a natural endpoint, the perfective is a bad choice, because in essence, the perfective sentence (8b) sounds like a confession of something at least, if not of Ivan's killing. The sentence informs the addressee that although the killing of Ivan by Anna did not successfully take place, it was reasonable to expect such a murder. For instance,

[3]For a modal analysis of a related phenomenon, namely *end by V(-ing)*, see Amaral and Del Prete (2020). Here we eschew a modal analysis in favor of a causal analysis. To be clear, we don't necessarily claim that English *ultimately/in the end ... Neg* has the same contribution as the Russian negative perfective but we find the gloss to be appropriate enough.

it is possible that Anna tried to kill Ivan but failed as he was stronger. Alternatively, she may have planned the murder but ultimately decided not to perform it (because that would be too risky). Or at least, she seriously considered the possibility of committing the murder but then, again, decided to backtrack.

(8) a. Ja ne ubivala Ivana
I NEG killed.IMP Ivan
'I didn't kill Ivan.'

b. Ja ne ubila Ivana.
I NEG killed.PERF Ivan
'(Ultimately,) I didn't kill Ivan.'

In order for the perfective to be used, telicity and even event specificity is apparently insufficient. Intuitively, the choice of perfective aspect means that something happened in the world that made an instantiation of the event described by the verb phrase plausible, expected, or feasible.[4]

Sometimes, this results from the fact that the event in question was actually taking place, although it did not reach completion, e.g. (9) may be uttered if Misha wrote a part of the letter but didn't complete the action (although in this context the use of the verb *dopisal* 'finished writing' may be preferable). But it is also possible to use negative perfectives when the event described by the verb phrase did not even begin. The fact that sentences with perfective accomplishment verbs do not presuppose that the event in question has started is shown convincingly by Zinova and Filip (2014). The sentence in (9) may therefore be acceptable even if the writing event did not start; still, crucially, some other eventuality in the actual world must have taken place which made the occurrence of letter-writing expected or, in some sense, realistic. Thus, (9) is appropriate if Misha did not even begin writing the letter, but at some point promised or intended to do so. Crucially, an instantiation of the event is not required, even a partial instantiation; but some kind of feasibility or expectedness is necessary.

(9) Miša ne napisal pis'mo.
Misha NEG wrote.PERF letter
'Misha didn't write / hasn't written the letter.'

Despite the plausibility, expectedness or feasibility of Misha writing a letter, whatever conditions or events which made it thus were not able to make the writing actually happen. Something else got in the way of, or "defeated" the manifestation of the expected outcome. We can speak then about the "defeasibility" of the outcomes that the conditions or events would have tended to make happen. In the case of (8), Anna, her intention or her circumstances were such that that the normal course of events would have eventually led to Ivan's death, but Anna did not carry out the killing to the end, because, for whatever reason,

[4] In this sense, negative perfectives are reminiscent of another aspectual phenomenon in Russian, so-called *bylo*-sentences, which report that an event failed to receive an expected continuation (Kagan, 2011).

it was thwarted. (N.B.: *defeasibility* is 'defeat-ability', not 'de-feasability').

4 Modeling event specificity, feasibility, and defeasibility

We propose to build our semantic model in such a way as to easily represent the most prominent components of meaning in Russian negative perfectives as observed above: event specificity, feasibility, and defeasibility.

Event specificity tells us that we are dealing with a specific event of some kind, anchored to some established situation.[5] Feasibility tells us that this kind of event tends to lead to the maximal outcome spelled out by the verb phrase, and defeasibility tells us that in this case this event did not lead to such an outcome; that is, some other condition or event interfered, and so the event was not "efficacious" in the sense of Copley and Harley (2015), for whom a situation is efficacious just in case its *ceteris paribus* successor outcome occurs.

Ever since Dowty's (1977, 1979) work on the English progressive, notions of feasibility and defeasibility have canonically been modeled using quantification over a set of possible worlds, in Dowty's case a set of "inertia worlds". The idea is that an outcome is feasible if on these worlds, where things proceed in accordance with their properties in the actual world—that is, without interruption—the outcome occurs. But the actual world could either be in the set of inertia worlds, or outside of this set. A defeated outcome means that the actual world is outside of the set of inertia worlds.

We will not use quantification over possible worlds, however; instead we will define explicitly defeasible causal relations between eventualities, which will allow us to view eventualities, both circumstances and events, as *influences*. This framework gives us a simple analysis that we believe is explanatory. We discuss the issues around the choice of explicit causal relations versus quantification over possible worlds below in section 7.

What we need specifically is the ability to represent multiple potential—that is, defeasible—influences on an eventuality. The causal relations we use must be fit for this purpose. In the most prominent linguistic treatment of causation at the moment, namely one where causal relations are taken to relate Davidsonian events, multiple influences are not represented. This is not because multiple influences are incompatible with Davidsonian event arguments; rather, it is only because the relation of causation has not been defined in that framework to admit multiple opposing influences, or even, indeed, multiple congruent influences. Typically, something identified as "the cause" is given pride of place, and following Dowty (1979), any other influencing conditions are relegated to expression through quantification over the preferred set of worlds (inertia, closest, etc.). Moreover, in a case where an event occurs and we are tempted to think that that event normally tends toward a certain outcome, if that outcome doesn't in

[5] See Von Heusinger (2002) on the role of anchoring in specificity within the nominal domain.

fact occur, Dowty's theory is forced to deny any causal relation at all between that particular event and the actual outcome. This is exactly because Dowty's definition of causation entails that the outcome occurs. This fact is what necessitates the use of possible worlds, where the causal relation holds with an expected but defeated outcome. This outcome is caused only in worlds that are not the actual world (and see also discussion in Copley and Wolff (2014) about "result-entailing" and "non-result-entailing" theories of causation).

There exists a different formalism of causation that does allow multiple influences, namely causal models, most famously explicated by Pearl (2000) (and see work within linguistics such as Kaufmann, 2013; Nadathur, 2019; Nadathur and Lauer, 2020; Baglini and Bar-Asher Siegal, 2020; Copley, 2021; Nadathur, 2021; Nadathur and Filip, 2021; Copley and Mari, 2022). However, we also find that the way in which "classical" causal models formalize multiple influences might not be quite appropriate for accounting for feasibility and defeasibility in the meanings of Russian negative perfectives. We will adjust the definition of causal models accordingly.

4.1 Classical causal models

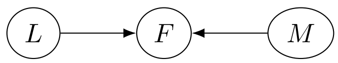
A causal model is a formal representation of the structure that causal relations give to a conceptual model of the world.[6] Causal models are formally represented by means of directed acyclic graphs (DAGs). In a DAG, there is a set of variables that are the vertices (or nodes) of the graph. These variables are allowed to have various values. The variables are connected by a set of edges (or arrows). The causal model conveys that the values of some variables influence the values of other variables, according to the arrows of the graph (which show the *direction* of causation). For instance, Ⓐ → Ⓑ represents that the value of Ⓑ is dependent in some way on the value of Ⓐ, without specifying which values go together. Functions associated with the model give more information about which values go together.

Information about which values go together is something that comes from world knowledge. Consider, for instance, a scenario where either lightning or arson may cause a forest fire. We can model this scenario with the causal model in Figure 1. The nodes correspond to *whether* clauses[7]; in this case, whether there is fire depends on both whether there is lightning and also whether there is a lit match.

The entire structure is a "collider" structure, since two influences in a sense collide. The possible values of each node are the truth values of the associated proposition. We can represent the value of Ⓕ given the values of Ⓛ and Ⓜ: so the value of Ⓕ is given by a function over the possible values of Ⓛ and Ⓜ. In this case, that function happens to be the *OR* function. In principle, we could set it to be any function we like that takes the values of the other variables and returns the value of Ⓕ, but it should faithfully reflect world knowledge.

[6]For a gentle introduction to causal models, see Chapter 1 of Pearl and Mackenzie (2018).

[7]This is true in classical causal models as in Pearl (2000), but we innovate the use of *whether* clauses to explicate them instead of "Ⓐ=1 iff . . . ".

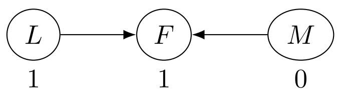


Ⓛ = whether there is lightning
Ⓕ = whether there is fire
Ⓜ = whether there is a lit match

Ⓛ	Ⓜ	Ⓕ
1	1	1
1	0	1
0	1	1
0	0	0

Figure 1: A collider with no values specified

To represent a certain (perhaps actual) state of affairs, we can pick from among the lines of the truth table to populate the nodes with values, as shown in Figure 2.



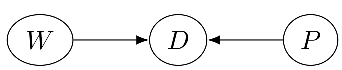
Ⓛ = whether there is lightning
Ⓕ = whether there is fire
Ⓜ = whether there is a lit match

Ⓛ	Ⓜ	Ⓕ
1	1	1
1	0	1
0	1	1
0	0	0

Figure 2: The same collider with actual values specified

Causal models are useful for us in that they allow us to represent the directionality of causation (given by the variables and the arrows) without a particular variable necessarily being given a particular value. This means that a node's presence in a model $\mathcal{M}$ does not mean that the situation it describes actually exists in the world under consideration.

Note that the same DAG, e.g. a collider structure, can be used with a different function, as shown in Figure 3. The function need not even be a familiar one from propositional logic. In Figure 3, for instance, the plug prevents the water from draining, and the function involved is not a familiar one.



Ⓦ = whether there is water in the tub
Ⓓ = whether water drains from the tub
Ⓟ = whether there is a plug in the drain

Ⓦ	Ⓓ	Ⓟ
1	1	0
1	0	1
0	1	0
0	0	0

Figure 3: Same structure, different function

Cases of prevention like these will be of interest to us in explaining the case of negative perfectives in Russian, because these are the cases we want to use to represent the "defeat" of one influence by another.

4.2 Modeling defeasibility

To model defeasibility, we are interested in treating the cases where an outcome is defeated as examples of "anti-efficacy": i.e., Ⓐ → Ⓑ is in the model, but

Ⓑ does not have the value that Ⓐ's influence would have been expected to make it have. But note already that this kind of talk is not congruent with a classical causal model, where, as we saw above, one multivariable function globally determines what happens. "Determine" is a good word here, since such a model represents the world as deterministic. In a deterministic framework, it's difficult to know what to do with intuitions such as those surrounding the negative perfective example (8b) above:[8] maybe Anna wanted to kill Ivan, and *normally* her action would have sufficed to kill him, but *exceptionally* that's not how things turned out.

What's missing in classical causal models is a sense of what might be normal and what might be exceptional. There are no doubt a number of ways to add a formalized sense of normality—feasibility, expectation, and the like—to classical causal models. The way we propose to do it, following Copley (2021), is to write it into the definition of what it means for an arrow to be present in a causal structure, as in (10):

(10) Local listening (Copley, 2021)
The presence of an arrow Ⓐ → Ⓑ in a model denotes that in a closed world containing only Ⓐ → Ⓑ, the value of Ⓑ "listens" to the value of Ⓐ.

Note that in (10), we reference the so-called "closed world assumption"—the assumption that any condition or event not represented makes no difference to the value of Ⓑ. "World" here is not a possible world but just the local state of affairs, the one that corresponds to just Ⓐ → Ⓑ. "Local" refers to the idea that this may not be the only arrow in our model; it is a locally closed model. (and see Copley & Harley 2015 on the topic of *"ceteris paribus"*). That is, if Ⓐ → Ⓑ is in the model, we assume background conditions such that the value of Ⓑ listens to the value of Ⓐ, provided that we erase any other arrows in our model that might alter the value of Ⓑ. This means that to calculate local listening, we also ignore any actual values of Ⓐ and Ⓑ given by the valuation $\mathcal{V}$, and just consider the arrow function.

We also explicitly require the value of Ⓑ to listen to the value of Ⓐ in every DAG in which there is an arrow Ⓐ → Ⓑ. This requirement will in turn lead us to rethink the role of functions in causal models.

4.3 Flavors of influence and arrow functions

We can think of an arrow from Ⓐ to Ⓑ in the DAG of a causal model $\mathcal{M}$ as representing that the value of Ⓐ influences the value of Ⓑ in a closed world, where any other nodes and arrows in the DAG were to be erased. Since we are thus interested in whether the value of Ⓑ listens to the value of Ⓐ, we need a *local* function—a function that is associated not with the whole model, but

[8] As always, we're not concerned with metaphysics here, just in trying to work out the causal calculus for the meaning. If you want to have a deterministic metaphysics underpinning our non-deterministic causal calculus, we are fine with that.

only with the arrow from Ⓐ to Ⓑ. This is a departure from classical causal models, where as we saw above, the function that codes the relationship between influences and outcomes is a global function, and can take as many arguments as there are influencing nodes. But our formalization of normality as "local listening" in (10) requires the use of local, arrow-associated functions, because local listening is expressed with a function, about a single arrow.

So, given a simple causal structure like Ⓐ → Ⓑ, we can think of the arrow itself as being associated with an arrow function representing which values of Ⓐ go with which values of Ⓑ (Copley, 2021; Copley and Mari, 2022). Crucially, again, the arrow function must be such that the value of Ⓑ must "listen" to the value of Ⓐ; that is, a difference in the value of Ⓐ must make a difference in the value of Ⓑ. But this requirement can be filled in two different ways.

We use the set of truth values as both the domain and the range of arrow functions. There are only four possible one-place functions from {0,1} to {0,1}, shown in Figure 4.

Figure 4: *All possible one-place functions from truth values to truth values*

x	y	x	y	x	y	x	y
1	1	1	0	1	1	1	0
0	0	0	1	0	1	0	0

But only the first two functions in Figure 4 can serve as arrow functions. The reason is that, in order to represent influence, whatever the flavor, the value of the effect must "listen" to the value of the cause. This means that any difference in the value of the cause (here represented by x) must make a difference in the value of the effect (here represented by y). This is the case, for instance, in the first function in Figure 4: When $x = 1$, $y = 1$; and if we were to change the value of x to 0, the value of y would also change, namely, to 0. Likewise, in the second function, changing the value of x leads to a change to the value of y. However, in the last two functions, changing the value of x does not lead to a change in the value of y. This being the case, only the first two functions in Table 4 can serve as arrow functions. These two functions are familiar: the first is the identity function and the second is negation. However, here they will be playing different roles due to arrow functions' association with causation.

In our causal models either of these functions can be associated with an arrow from Ⓐ to Ⓑ, and they correspond to two "flavors of influence" (Copley and Mari, 2022), namely *stimulation* and *inhibition*.

Note as well that the presence of Ⓐ → Ⓑ in the model does not guarantee that Ⓑ = 1 is likely or even plausible given that Ⓐ = 1, because an additional efficacious inhibitory influence may be part of the most plausible scenario, as we will see now.

Recall that in examples like (11) (= (8b) above), the intuition is that there was an influence that normally would have caused the event of the speaker killing Ivan, but something prevented them from doing so.

Figure 5: *Stimulatory and inhibitory arrow functions*

Stimulatory arrow function (f^+) Inhibitory arrow function (f^-)

Ⓐ	$\xrightarrow{f^+}$	Ⓑ
1		1
0		0

Ⓐ	$\xrightarrow{f^-}$	Ⓑ
1		0
0		1

(11) Ja ne ubila Ivana.
I NEG killed.PERF Ivan
'(Ultimately,) I didn't kill Ivan.'

To be more specific, we can identify (at least) three contexts that would make (11) true (and see Martin (2020) and references therein for antecedents to these in the literature):

- a "failed-event" context where the speaker either wanted to initiate an event that would feasibly kill Ivan, or was in a circumstance that would have feasibly led to them initiating such an event; either way, the event did not occur, and Ivan remains alive.
- a "failed-result" context where the speaker either wanted to initiate an event that would feasibly kill Ivan, or was in a circumstance that would have feasibly led to them initiating such an event; either way, the event occurred but did not result in Ivan's death, and Ivan remains alive.
- a "failed-event but result (= Ivan is dead) anyway" ("febra") context where the speaker either wanted to initiate an event that would feasibly kill Ivan, or was in a circumstance that would have feasibly led to them initiating such an event; either way, the event did not occur, but some other event resulted in Ivan's death. This context is equivalent to the cooccurrence of a failed-event context with a failed-result context; the event in question didn't occur but nonetheless Ivan is dead.

To preview our analysis of (11), we will propose that the use of the perfective requires the speaker to presuppose an additional influence on the event under discussion. This additional influence could be an intention or a circumstance. Consider, for example, the "failed-event" context where Anna intended to kill Ivan, but the killing event does not begin (we'll consider the other contexts later). Negation is what ensures the non-occurrence of the event. This means that we have a model as in 6 below, where Ⓘ represents whether Anna has an intention to kill Ivan.[9] We use a gray background to indicate what is presupposed, so here Ⓘ falls within the gray background. Ⓔ represents whether there occurs an event, with Anna as the agent, and which is expected to kill Ivan; and

[9]Philosophically-minded readers may at this point be skeptical, as for many philosophical views, intentions are not causal (Setiya, 2022). We note that the causal models here could be viewed as merely providing explanations, which might be enough to overcome such objections.

Ⓡ represents whether Ivan is dead. In a locally closed model, the value of Ⓔ, i.e., whether the event occurs, depends only on the value of Ⓘ, in a stimulatory flavor—if Ⓘ = 1, then Ⓔ = 1. Likewise, if we locally close the model Ⓔ → Ⓡ, the value of Ⓡ depends only on the value of Ⓔ, also in a stimulatory flavor.

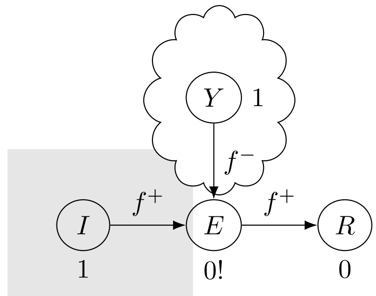


Figure 6: Model for failed-event context

However, we also know, because of negation, that at least one of Ⓔ and Ⓡ has the value zero. For the failed-event context, in fact, they both have the value zero. Thus, the value of Ⓔ is unexpected—in fact, inexplicable—in terms of the value of Ⓘ together with the stimulatory arrow function characterizing the arrow between Ⓘ and Ⓡ. We mark that inexplicable quality of the value of Ⓔ with an exclamation point. (In contrast, the value of Ⓡ *is* expected in light of the value of Ⓔ together with the stimulatory arrow function between Ⓔ and Ⓡ, so there is no exclamation point on the value of Ⓡ.) Because of the inexplicability of the value of Ⓔ, this model evokes, or invites the inference, of a new, additional influence Ⓨ to explain the value of Ⓔ. We could equally have given Ⓨ a value of 0 and the Ⓨ to Ⓔ arrow a stimulatory flavor to model an *absence* of an expected influence. (This, incidentally, is how causation by omission is to be modeled.) To collect both of these cases, we will heretofore use the notation ":1" as shorthand for "f(Ⓐ) = 1" on the arrow.

This is the account we will develop now. The first thing we must do is to explain how such inexplicable values are to be dealt with in the causal model, and then, in section 5, we will see how all of this fits together compositionally.

4.4 Inexplicable values and the causal principle

Our proposal, recall, is that inexplicable values can appear in causal models, even though they would in classical models require the erasure of the arrow. For us the arrow can remain, since it reflects only local listening, which says only that Ⓘ gets to be efficacious in the absence of another influence. But here we have another influence, namely Ⓨ, so all bets are off with respect to the efficacy of Ⓘ. We can view the arising of Ⓨ as due to what is commonly called the "causal principle", stated casually in (12a), and more precisely for our particular use of it in (12b):

(12) The causal principle

a. (pithily) Events have causes.

b. (in our framework) For all well-formed models $\mathcal{M}$, for any endogenous node (node with at least one arrow pointing to it) Ⓑ in $\mathcal{M}$, there exists a node Ⓐ such that:
Ⓐ $\xrightarrow{f}$ Ⓑ is in $\mathcal{M}$ and f(Ⓐ) = Ⓑ. In that case we can say that the value of Ⓐ *explains* the value of Ⓑ.

(13) Inexplicable values: Where the causal principle is violated, i.e., where there is a Ⓑ such that for all Ⓐ such that Ⓐ $\xrightarrow{f}$ Ⓑ is in $\mathcal{M}$, f(Ⓐ) $\neq$ Ⓑ, the value of Ⓑ is *inexplicable* in $\mathcal{M}$

Being inexplicable is different from merely being unexplained. For instance, the value of Ⓐ in Figure 7 is unexplained by anything in the model but is not inexplicable because it has no influences on it in the model that would "want" it to have a different value. The value of Ⓑ, on the other hand, is inexplicable (as well as unexplained).

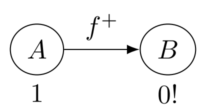


Figure 7: Model showing the difference between an unexplained value (value of Ⓐ)and an inexplicable value (value of Ⓑ)

We also need to be intolerant of inexplicable values, in the follow sense:

(14) Intolerance of inexplicable values: If $\mathcal{M}$ contains an inexplicable value, it cannot be a closed world. M must be altered so as to explain the value.

So, an exception to a causal generalization is where there is an arrow (the value of Ⓑ normally listens to the value of Ⓐ) but the value of Ⓑ is not given by f(Ⓐ). In this case, Ⓐ is anti-efficacious; the arrow can remain, because it *would* be efficacious in a locally closed model, without the actual values, and considering the function f^+. As it is, however, Ⓐ is not efficacious, and another influence on Ⓑ must be evoked to explain the value of Ⓑ.

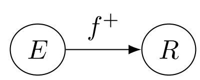
5 Building the denotation

To build the denotation of Russian negative perfectives, we will address the contributions to the model by the verb stem, by the perfective operator, and by negation, and then show how the several construals discussed above are derived on this model. Afterwards, we will relate these models to a more familiar situation semantics for compositionality.

5.1 Events and results

We consider telic verb phrases to introduce a model which corresponds to the occurrence in the world of a (perhaps partial) event or process (Ⓔ), which itself influences whether a result occurs (Ⓡ). We do not present an example of an atelic verb phrase here (within the perfective domain, these are arguably represented by verbs with the prefixes *po-* and *pro-*; cf. Borik (2002) for this claim). We note however that a recognition of the existence of entrainment causation, where the cause occurs at the same time as the effect (e.g. Talmy, 2000; Copley and Harley, 2015), is an underused possibility that would be helpful here. So, if we understand atelic dynamic predicates such as *dance* as cases of entrainment, they would contribute the same Ⓔ $\rightarrow$ Ⓡ structure as telic verbs. Ⓔ would correspond to whether an energetic process occurs (e.g. whether the dancer puts energy into dancing) and Ⓡ would correspond to the cotemporal existence of the abstract object (e.g. the dancing motion) that is created by the energetic process. Note that the result need not be a state; nor does it have to occur after the event. Ultimately it will be the verb phrase that tells us whether the event is temporally prior ($\prec$) to the result, even though the event is always causally prior ($\rightarrow$) to the result.

For the case we are interested in, ANNA KILL IVAN contributes an event node Ⓔ and a result node Ⓡ, as well as a causal relation between them. The model for ANNA KILL IVAN is given in 8 below:



Ⓔ = whether there exists an Anna-initiated event
that feasibly results in Ivan being dead
Ⓡ = whether Ivan is dead

Figure 8: Model for ANNA KILL IVAN

Anna kill Ivan contributes the model above, but it does not contribute the values of Ⓔ or Ⓡ. Rather, these values are contributed either by negation (see section 5.3 below), or affirmation; in the case of affirmation, both nodes get a value of 1.

5.2 The contribution of the perfective in Russian

We propose as well that the perfective in Russian contributes a specific not-at-issue influence (we could also call it an "aspectual" or a "perfective" influence) with an arrow pointing toward the event node Ⓔ. This new influence we notate as Ⓧ. We believe that the presence of Ⓧ is presupposed because it is maintained under negation (3a), in interrogatives (4a), and in the antecedent of conditional sentences as in (15):

(15) a. Esli Dima čital Vojnu i mir, on pojmët
if Dima read.IMP War and Peace he understand.FUT
obsuždenie.
discussion
'If Dima has read War and Peace, he will understand the discussion.'

b. Esli Dima pročital Vojnu i mir, on pojmët
if Dima read.PERF War and Peace he understand.FUT
obsuždenie.
discussion
'If Dima has read War and Peace, he will understand the discussion.'

Unlike (15a), (15b) requires a context in which Dima was expected or supposed to read War and Peace.

What is the nature of this event-influencing Ⓧ? An event may be influenced by the subject's intention, which makes the model in Figure 9 a representation for intentional action; or alternatively, an event can be influenced by ad-hoc circumstance(s), which makes Figure 9 a representation for mere (accidental) actions and non-agentive events (and see Martin 2020 for an argument for treating the latter two similarly). This influence (Ⓧ) will be contributed as presupposed/not-at-issue meaning by perfective aspect. Here we represent not-at-issue meaning on a gray background, with at-issue meaning on a white background.[10]

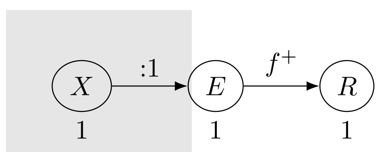


Figure 9: Model for PERF(*Anna kill Ivan*)

(16) Two possibilities for Figure 9
a. Ⓧ is Ⓘ, where Ⓘ is subject's intention
b. Ⓧ is Ⓒ, where Ⓒ is a circumstance; for accidental (unintentional) or non-agentive event readings

5.3 Negation

Negation on this account acts on the at-issue meaning, namely the valuation of the model that is provided by the verb stem. We get three construals, shown in

[10]This means that, strictly speaking, Ⓔ $\xrightarrow{f^+}$ Ⓡ should have a gray background, since the influence between whether the event occurs and whether the result occurs is not at issue, while the *valuation* of Ⓔ $\xrightarrow{f^+}$ Ⓡ is what is at-issue.

Figure 10: there can be neither event nor result occurring, or just the event, or just the result. We'll assume that the reason these possibilities are all available is because negation is focus-sensitive, and there can be focus on any (possibly improper) part of the at-issue portion of the model, and negation takes the focused node(s) and sets its/their value to 0.

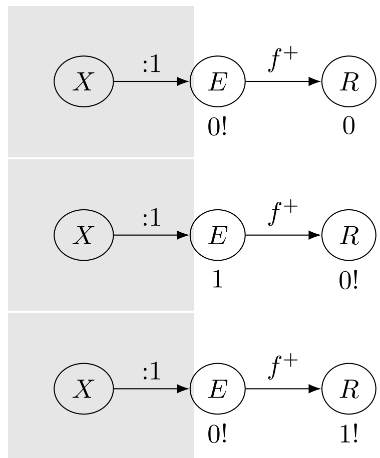


Figure 10: Models for NEG(PERF(*Anna kill Ivan*) contexts: failed event, failed result, and failed event but result anyway ("febra")

5.4 Putting the contributions together

To see how the contributions of the verb stem, the perfective, and negation yield the desired construals, consider again (8b), reproduced here as (17):

(17) Anna ne ubila Ivana.
Anna NEG killed.PERF Ivan
'Anna didn't kill Ivan.' (= (8b))

Consider now the failed event context, where the event does not even start, repeated here as Figure 11. Let's suppose that Ⓧ = whether Anna intends to kill Ivan. Ⓧ → Ⓔ is part of the model, but Ⓔ has the inexplicable value 0. Since Ⓔ = whether there is an Anna-initiated event that feasibly results in Ivan's death, this means that there is no such event. Because of the inexplicable value, the model needs to be altered.

The hearer infers an influence on Ⓔ that explains the otherwise inexplicable value, represented in Figure 12 as Ⓨ. That is, whether there is an intention or circumstance Ⓨ affects whether any (even "partial") event of the description occurs, i.e., whether Ⓔ = 1. In particular, whatever the value of Ⓨ may be, and whatever the Ⓨ-Ⓔ arrow function f may be, f(Ⓨ) = Ⓔ = 0. Figure 12 is an "event collider", since there is a "collision" between influences centered on

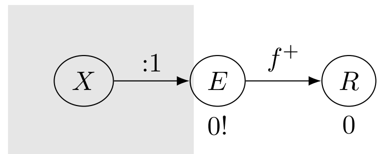


Figure 11: Failed event; inexplicable event value

the event node Ⓔ.

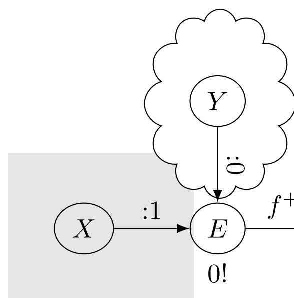


Figure 12: Failed event: An "event collider", with influence Ⓨ inferred to explain the otherwise inexplicable value of Ⓔ

Likewise, in the failed result context, recall that there is an inexplicable value for Ⓡ (shown again in Figure 13). In a similar fashion to the failed event context, the inexplicable value of Ⓡ is inferred to result from an evoked influence Ⓨ, as shown in the "result collider" in Figure 14.

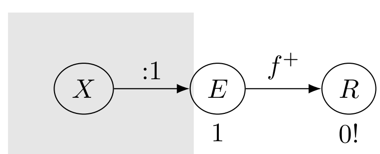


Figure 13: Failed result; inexplicable result value

Third, the "febra" context, shown in Figure 15, has inexplicable values for both Ⓔ and Ⓡ. The febra context includes the case where Anna wants to kill Ivan, but something prevents her from doing so; but someone else kills him.
Since there are two inexplicable values, we suppose the existence of two inferred influences (it also could be a single influence influencing both Ⓔ and Ⓡ), and we have an "event-result double collider".

In all of these contexts, the easiest scenario to think about is the one where Ⓧ = whether Anna has the intention to kill Ivan. But remember that Ⓧ need not always be about her intention; it could be instead about circumstance(s) that would tend to make feasible an action she performs accidentally. For instance,

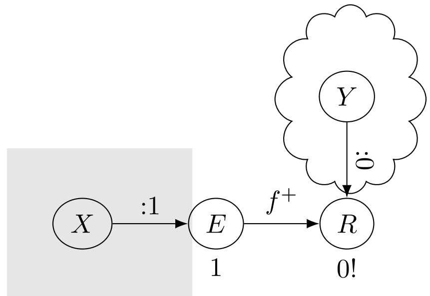


Figure 14: Failed result: A "result collider", with influence Ⓨ inferred to explain the otherwise inexplicable value of Ⓡ

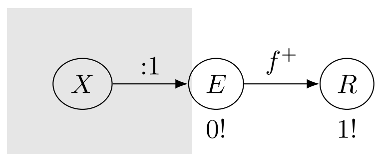


Figure 15: Failed event but result anyway ("febra"); inexplicable event and result values

she could have been holding a gun, aiming at another target, but been pushed in such a way as to make feasible her shooting Ivan. In that case, Ⓧ would be whether she is pushed in such a way.

5.5 Compositionality

To compose the denotation of Russian negative perfectives, we first claim that the causal model under discussion is referred to via an indexical contextual variable $\mathcal{M}$, as a kind of common ground Stalnaker (2002). The perfective operator operates on the model, adding the node Ⓧ to it as a causal influence. Thus we are dealing with a "monster" in the sense of Schlenker's 2003 corollary of Kaplan's definition, namely, an operator that manipulates a contextual variable.

Our proposal is that perfective in Russian is just a monster that adds a node Ⓧ influencing the first node in $\mathcal{M}$ (namely, Ⓔ) and "wanting" that first node in $\mathcal{M}$ to have the value 1.

(18) $\textsc{Perf}(\llbracket \alpha \rrbracket^{\mathcal{M}}) = \llbracket \alpha \rrbracket^{\textcircled{x} \xrightarrow{:1} \mathcal{M}}$

The lexical entry of a verbal stem such as Kill[11] includes not only the denotation returned by the evaluation function when applied to Kill. To do this, we will use a function notated with a box around "Kill", such that $\boxed{\text{Kill}}$ gives us

[11] The stem consists of the prefix *u-* and the root *bi-* 'beat'; the attachment of the prefix shifts the lexical meaning from 'beat' to 'kill'.

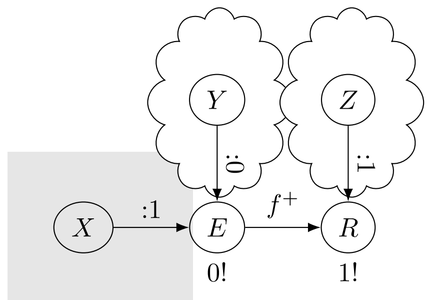


Figure 16: Febra context: An “event-result double collider”, with influence Ⓨ inferred to explain the value of Ⓔ and influence Ⓩ inferred to explain the value of Ⓡ

the causal model contributed by KILL. We assume that the evaluation function requires that the causal model of KILL is the causal model of evaluation.[12]

Let $\mathcal{V}$ be a valuation function that takes a causal model $\mathcal{M}$ and a situation s and returns the truth values that correspond with the nodes of the structure on a single line of $\mathcal{M}$'s valuation table. Here we represent the output of the valuation function with an ordered pair.[13]

(19) Lexical entry of KILL

a. The causal model given by KILL, written $\boxed{\text{KILL}}$, is Ⓔ → Ⓡ, where Ⓔ $= \lambda y \lambda s \,.\, \exists e : \theta_{a/c}(y, e)$ in s and Ⓡ $= \lambda x \lambda s \,.\, x$ is dead in s

b. $[\![\text{KILL}\ ]\!]^{\boxed{\text{KILL}}} = [\![\text{KILL}\ ]\!]^{Ⓔ \to Ⓡ} = \lambda y \lambda x \lambda s \,.\, 1$ iff $\mathcal{V}(Ⓔ(y) \to Ⓡ(x))(s) = (1, 1)$

In (19) we have notated an “agent/causer” thematic role rather than just “agent” because inanimate causers are also possible with KILL. We have not explicitly notated a patient role for Ivan because the patient role is effectively given by the causal model: Ivan is dead because of the event referred to in the denotation of Ⓔ.

Saturating the entity arguments of the verbal stem, and combining the result with the perfective operator in (18) above, yields (20), where Ⓔ $= \lambda s \,.\, \exists e : \theta_{a/c}(\text{ANNA}, e)$ in s and Ⓡ $= \lambda s \,.\,$ IVAN is dead in s :

(20) PERF($[\![\text{ANNA KILL IVAN}]\!]^{Ⓔ \to Ⓡ}$) $= [\![\text{ANNA KILL IVAN}]\!]^{Ⓧ \xrightarrow{:1} Ⓔ \to Ⓡ} = 1$ iff $\mathcal{V}(Ⓔ \to Ⓡ)(s) = (1, 1)$

We consider negation to also be a monster. And as expected, negation takes

[12] This requirement is too strong; KILL should also be able to be evaluated in a model that has Ⓔ → Ⓡ as a part. We leave this point, as well as the question of how lexically-contributed models are integrated into a causal model under discussion, for future work.

[13] These valuations are more than just window dressing; a verb such as *prevent*, for instance, would have a valuation (1,0).

its argument and returns 1 just in case the denotation of the argument in the model of evaluation is 0.

(21) $\text{NEG}([\![\alpha]\!]^{\mathcal{M}}) = 1$ IFF $[\![\alpha]\!]^{\mathcal{M}} = 0$

Adding (21) to (20), we get (22):

(22) $\text{NEG}(\text{PERF}([\![\text{ANNA KILL IVAN}]\!]^{Ⓔ \to Ⓡ})) = \text{NEG}([\![\text{ANNA KILL IVAN}]\!]^{Ⓧ \xrightarrow{:1} Ⓔ \to Ⓡ})$
$= 1$ iff $\mathcal{V}(Ⓔ \to Ⓡ)(s) \neq (1,1)$

As desired, the denotation in (22) has Ⓧ escaping negation. The reason it escapes negation is that negation effectively acts on the valuation statement $\mathcal{V}(Ⓔ(y) \to Ⓡ(x))(s) = (1,1)$ which ultimately comes from (19b). Ⓧ is never valued, so there is no truth value of Ⓧ for negation to act on.

Verb phrases that are not typically thought of as encoding result states may in fact encode results, i.e. they have a causal structure Ⓔ → Ⓡ, but the result is not a state (and see also Copley and Harley (2015) for the claim that results need not be states). Consider, for instance, the sentence in (3a) above, repeated here as (23).

(23) Lena ne pozvonila.
Lena NEG phoned.PERF
'Lena hasn't phoned.'

Here we want to argue that there is an action (event) of phoning that normally results in the other phone ringing. Russian perfective predicates normally contain a result in their denotation. An exception may be constituted by perfective activities with the prefixes *po-* and *pro-*, but these are highly exceptional in many ways; therefore, we leave their investigation within the present framework to future research. For discussion of these prefixes, see Borik (2002); Kagan (2016); Filip (2000); Zinova (2021) and references therein.

6 Discussion

Our account explains the presupposed and evoked influences in Russian negative perfectives through the presupposition of Ⓧ and the need to avoid inexplicable values in the model, which gives rise the to the influence Ⓨ and (sometimes) Ⓩ.

Negative perfectives vs. affirmative perfectives: Recall that affirmative perfectives do not evoke influences (represented by Ⓨ and Ⓩ) at all. This is as expected, since in the absence of negation, all values would be explained or explicable according to the lexical entry of the verbal stem. However, even in affirmative perfectives, we can still detect the presence of presupposed Ⓧ, because of the specificity facts.

Specificity: The specificity effect associated with perfective clauses ((3a), (4a) above) is due to the presence, but lack of explicit valuation of Ⓧ. We therefore are dealing not with just any situation that $\textcircled{E}$ holds of, but rather specifically with the situation that is caused by the specific intentional state or circumstance that characterizes $\textcircled{X}$. Such an intention or circumstance constitutes the identifying property which creates the specificity intuition. Specificity is maintained in (4a), despite the absence of temporal definiteness, as expected in our theory: (4a) asserts that a phoning event caused specifically by the $\textcircled{X}$ eventuality did not take place at any temporal interval.

Feasibility and defeasibility: The feasibility intuition is due to the presupposition that $\textcircled{X}$ influences $\textcircled{E}$, which, again, is not at issue precisely because $\textcircled{X}$ does not get a valuation. Our interpretation of the arrows as conveying defeasible efficacy means that there is an influence that is defeasibly efficacious in a closed model for whether an $\textcircled{E}$ event occurs that would itself be defeasibly efficacious in a closed model for $\textcircled{R}$. This is the case independently of whether the sentence is affirmative or negative. In the negative, however, either $\textcircled{E}$, $\textcircled{R}$, or both must be equal to zero to ensure that $\mathcal{V}(\textcircled{E} \rightarrow \textcircled{R}) \neq (1,1)$. The effect is that we retain the feasibility of $\textcircled{X}$ to cause $\textcircled{E}$, even if at least one of $\textcircled{E}$ and $\textcircled{R}$ is equal to zero.

Maximality: Recalling that the notion of "maximality" of an event has been proposed as the core meaning of perfectives (Filip and Rothstein, 2006; Filip, 2008), we notice that in our system, for affirmative perfectives, $\textcircled{X}$ is presupposed as the *only* explanation of the values in $\textcircled{E}$ and $\textcircled{R}$, because a closed model is supposed. This corresponds to maximality (or exhaustivity), not of the event, but of causal influences on the event and result. In the case of the negative perfective, however, the speaker ends up committing themselves to having at least one more node in their causal model than what is provided by the sentence, namely the node $\textcircled{Y}$, so the influences contributed by the sentence are not the maximal set of influences. We therefore derive maximality of the event in the sense of Filip and Rothstein (2006); Filip (2008) with a closed world and non-interference, i.e., exhaustivity of the causal influences given by the sentence itself. Likewise, non-maximality of the event in the case of the negative perfective is derived from the fact of interference or interruption by additional causal influences, where the causal influences given by the sentence itself (namely, $\textcircled{E}$ and $\textcircled{R}$) are not exhaustive.

7 Explicit causal relations versus quantification over possible worlds

Earlier, we promised to return to the idea of quantification over possible worlds. This idea came up not because possible worlds are typically used in the denotations of perfectives, or even negative perfectives—this is not the prototypical analysis.[14] In fact, we are very much in the right part of phrase structure to use

[14]We acknowledge the existence of habitual perfectives, which may possibly involve quantification over possible worlds, but such uses are highly restricted in Russian; the imperfective is preferred. We believe that in habitual perfective cases, a higher modal operator, which is not contributed by the perfective, takes scope over the AspP.

causal relations, given that verbal denotations are often analyzed with causal relations (as we do here, with Ⓔ → Ⓡ, and the perfective morphology we have been dealing with here is probably part of or adjacent to VP.

The real reason quantification over possible worlds comes up is that it is the standard way to model alternatives to actual events, and the data show that in Russian negative perfectives, alternatives to actual events are considered, since we compare what happened to what was expected to happen. So the question is raised, why do we not use quantification over possible worlds? Our answer is that quantification over possible worlds is not necessary to account for these data, nor is it explanatory.

To see that quantification over possible worlds is not necessary, consider the origin story of the familiar Lewis-Dowty-Stalnaker-Kratzer possible worlds framework. Although quantification over possible worlds may seem like the null hypothesis because of its familiarity, we should remember that one of the major motivations David Lewis had for developing possible world semantics was that he wanted to *explain* causal relations (Lewis, 1973), which for him were too spooky to be primitive in the ontology. In Lewis's view, picking out the set of closest possible worlds *avoided* reference to causal relations; for him, everything we need to do in denotations can and should be done only using the primitives of possible worlds, times, and so forth.

Lewis was taking one side in a long-standing (and still live) philosophical debate over the nature of causation. Theories of causation can be divided (Copley and Wolff, 2014) into dependency (or 'make a difference') theories and production (or 'process', 'generative', or 'mechanistic') theories, the latter including force-dynamic theories. Dependency theories define causation as being fundamentally built on a dependency between occurrences of events. The statement *A causes B* is consequently defined in terms of a dependency of the occurrence of a *B*-event on the occurrence of an *A*-event. The particular kind of dependency can be one of logical dependency (e.g., Mackie 1974), counterfactual dependency (e.g., Lewis 1977), probability raising (e.g., Suppes 1970), or intervention (e.g., Pearl 2000).

Production theories, on the other hand, understand causation as involving a mechanistic relationship between participants, either as a configuration of forces (e.g., Wolff 2007), a transmission of energy (e.g., Dowe 2000), or a transference of some other quantity (e.g., Mumford and Anjum 2011).

Dependency theories, in short, view causation in terms of possibilities, in the sense that possible worlds or situations are primitive, and causation is defined in terms of propositions which are predicates of these possible worlds. Production theories, on the other hand, take causal concepts such as force and transmission of energy to be primitive, with the forces themselves defined as being directed toward possibilia. The relevance of the debate on causation to whether quantification over possible worlds is necessary for our data is therefore that causation and modality are 'two sides of the same coin,' as Ilić (2014) puts it. Either causation can be derived from (sets of) possibilities, as in the dependency perspective, or possibilities (in the sense of possible courses of events) can be derived from causation, as in the production perspective.

Both dependency and production theories are, or can be made, powerful enough to describe anything the other kind of theory can, even if some causal phenomena are better explained with one kind of theory than the other. The difference between a denotation that has quantification over possible worlds and one that has causation is analogous to the difference in phonology between Optimality Theory (generate all the possible forms and choose a particular set of them) and traditional phonological derivations (transform one form into another). In both approaches, there are possible forms that are unattested, but in only the first approach are such possibilia collected into sets by the grammatical computational system. Likewise, one can either treat possible courses of events as atomic possible worlds that are explicitly quantified over, or one can construct such courses of events from one event to the next using causal relations, as a causal chain. Just as in the phonology example, while both approaches are powerful and can handle a great deal of the data, certain data is more easily treated in one than the other. Here we argue that, at least for the phenomena in question, it is better to utilize a notion of causation than to quantify in the denotation over atomic possible worlds, even though both approaches may yield appropriate truth conditions.

We should underline that the production perspective does require primitive possibilia, but organizes them into a course of events (a 'world', or a line in the causal model's truth table, which we represent with the variable s) only through the intermediary of causal relations; crucially, we have no need to quantify over our courses of events.

The reason we don't need to quantify over our courses of events has to do with how we model normalcy. In a possible worlds framework, the normalcy of an outcome P is modeled by identifying the set of most normal, "best", "inertial" or most "stereotypical" possible outcomes, and then saying that all worlds in this set are P-worlds. In our framework, however, normalcy is modeled with respect to a certain causal model, call it $\mathcal{M}$. Any values that are explicable in terms of the values, functions, and structure of $\mathcal{M}$ are normal with respect to that model. Values of nodes that are inexplicable in terms of the values, functions, and structure of $\mathcal{M}$ are not normal, and their presence means that $\mathcal{M}$ is not an appropriate model of reality. Thus, another model $\mathcal{M}'$ which explains that value must be constructed; one that includes the evoked influence(s) Ⓨ and, if needed, Ⓩ.

If one likes our proposal but believes that there must be quantification over possible worlds in the denotation, there is a third option: to use our framework but to interpret the relations represented by arrows as involving quantification over possible worlds. For the reasons we have discussed above, we see no need to do this but it should be possible. If one chooses this direction, the major challenge would be to define the right set of worlds in all of which the event denoted by the VP is predicted to be realized and reach completion.

The most natural direction to take to do this would be in the spirit of modal analyses of the progressive aspect (e.g. Dowty (1979), Portner (1998)). According to this approach, the VP-event should reach completion in the "best" circumstantial or "inertia" worlds in which every element in this eventuality develops normally, in accordance with its inherent properties and the laws of physics and human behavior.

For instance, Anna could try to kill Ivan, plan to kill Ivan, or merely consider the possibility of killing Ivan. In the latter case, Anna may be willing to kill Ivan and consider arguments in favor of doing so but, within the same situation, she may also consider arguments against killing him. Thus, if we consider the situation as a whole, we cannot say that Anna kills Ivan in *all* the inertia worlds relative to that situation. This is so because the situation contains both elements that may cause the killing and elements that may prevent the killing. Then Anna will kill Ivan in some inertia worlds but not in others. In turn, if we want to make sure we eliminate the latter type of worlds, we should take causal relations into account while defining the right set of inertia worlds.

Even if the problem of finding the right set of worlds can be solved, we think that the use of possible worlds here is not as explanatory as the use of explicit structures of causal relations. One reason we say this is that, in our view, Lewis was mistaken in thinking that one can avoid using knowledge of causality to pick out sets of worlds. We do not see how. For, given that correlation is not causation, how would we know which set of worlds to pick out as the best or inertia worlds, if not by using causal information?

Inertia, after all, is a property that springs from the laws of physics. For us, what is at stake here is just a matter of whether one is referring to causal relations explicitly, by writing something like the causal arrow in the denotation, or merely using them implicitly, behind the denotational scenes as it were, in choosing the closest set of worlds. Since we also happen to disagree with Lewis that causal relations need to be explained in terms of possible worlds, we feel that it is not necessary to avoid causal relations. So we choose here to represent causal relations directly, in such a way as to model conditions and influences interacting with each other. That is, our causal models have defeasible causal relations, where there can be multiple possible influences on an outcome, and it can turn out that only one of the influences is efficacious.

A second reason we prefer explicit causal relations is that it is simpler than quantification over possible worlds, in the sense that the latter uses up a lot more ink when all we want to do is to identify multiple conditions or influences interacting. This is not merely an Occam's Razor argument, since what is at stake is the relationship between heads and their denotations—it is also a question of complexity of the syntax-semantics interface, which we would like to be as simple as possible.

8 Conclusion

Russian perfective sentences have curious properties such as as event specificity and feasibility that resist explanation by current analyses. These properties appear most prominently in negative perfective sentences. We claim that these properties arise because the Russian perfective is a causal monster: it adds to the causal model in question an additional, not-at-issue influence which affects the event described by the verb stem. Negative perfective sentences highlight the existence of this additional influence because they do not negate it, while still negating the at-issue content. This configuration leads to inexplicable outcomes in the model, which can only be explained by further altering the causal model to include at least one more influence.

Authors' addresses

Bridget Copley
SFL (CNRS/Paris 8)
59-61 rue Pouchet
75017 Paris, France

Olga Kagan
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva 84105, Israel

References

Altshuler, D. (2010). A birelational analysis of the Russian imperfective. In *Proceedings of Sinn und Bedeutung*, Volume 14, pp. 1–18.

Amaral, P. and F. Del Prete (2020). Predicting the end: Epistemic change in romance. *Semantics and Pragmatics 13*, 11–1.

Baglini, R. and E. A. Bar-Asher Siegal (2020). Modeling linguistic causation. Aarhaus University and Hebrew University of Jerusalem ms.

Borik, O. (2002). *Aspect and Reference Time*. Ph. D. thesis, Utrecht.

Copley, B. (2021). Reconciling causal and modal representations for two salish out of control forms. In *Proceedings of WCCFL 39*. Cascadilla Press.

Copley, B. and H. Harley (2015). A force-theoretic framework for event structure. *Linguistics and Philosophy 38*, 103–158.

Copley, B. and A. Mari (2022). *Forbid* is not *order not*. In *Proceedings of NELS 52*. UMass Amherst.

Copley, B. and P. Wolff (2014). Theories of causation should inform linguistic theory and vice versa. In B. Copley and F. Martin (Eds.), *Causation in Grammatical Structures*, pp. 11–57. Oxford University Press.

Dickey, S. M. (2000). *Parameters of Slavic aspect*. Indiana University.

Dickey, S. M. (2018). Thoughts on the 'typology of slavic aspect'. *Russian Linguistics 42*(1), 69–103.

Dowty, D. (1977). Towards a semantic analysis of verb aspect and the English 'imperfective' progressive. *Linguistics and Philosophy 1*, 45–77.

Dowty, D. (1979). *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Filip, H. (2000). The quantization puzzle. *Events as grammatical objects 39*, 96.

Filip, H. (2008). Events and maximalization: The case of telicity and perfectivity. In S. Rothstein (Ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, pp. 217–256. John Benjamins.

Filip, H. and S. Rothstein (2006). Telicity as a semantic parameter. In *Formal approaches to Slavic linguistics*, Volume 14, pp. 139–156. Michigan Slavic Publications Ann Arbor.

Forsyth, J. (1970). *A grammar of aspect: Usage and meaning in the Russian verb*. Cambridge University Press.

Grønn, A. (2004). The semantics and pragmatics of the Russian factual imperfective.

Ilić, T. (2014). Modality and causation: two sides of the same coin. In B. Copley and F. Martin (Eds.), *Causation in Grammatical Structures*, Oxford Studies in Linguistics. Oxford University Press.

Jakobson, R. (1984). Structure of the Russian verb. In M. Malle and L. R. Waugh (Eds.), *Russian and Slavic Grammar: Studies 1931-1981*, pp. 1–14. Berlin, New York: De Gruyter Mouton.

Kagan, O. (2008). On the semantics of aspect and number. In *Formal Approaches to Slavic Linguistics*, Volume 16, pp. 185–198.

Kagan, O. (2010). Russian aspect as number in the verbal domain. In *Layers of aspect*, pp. 125–146. CSLI Publications.

Kagan, O. (2011). The actual world is abnormal: on the semantics of the bylo construction in Russian. *Linguistics and Philosophy*, 57–84.

Kagan, O. (2016). *Scalarity in the verbal domain: The case of verbal prefixation in Russian*. Cambridge University Press.

Kagan, O. A. (2020). A force-dynamic approach to perfectivity. In *Vzaimodejstvie aspekta so smezhnymi kategorijami*, pp. 149–155.

Kaufmann, S. (2013). Causal premise semantics. *Cognitive science 37*(6), 1136–1170.

Krifka, M. (1992). Nominal reference and temporal. *Lexical matters* (24), 29.

Leinonen, M. (1982). *Russian Aspect, "Temporal'naja Lokalizacija" and Definiteness/Indefiniteness*. Neuvostoliittionsttituutin.

Lewis, D. (1973). Causation. *The Journal of Philosophy*, 556–567.

Martin, F. (2020). Aspectual differences between agentive and non-agentive uses of causative predicates. In E. Bar-Asher Siegal and N. Boneh (Eds.), *Perspectives on causation*, pp. 257–294. Springer.

Nadathur, P. (2019). *Causality, aspect, and modality in actuality inferences.* Ph. D. thesis.

Nadathur, P. (2021). Causality and aspect in ability, actuality, and implicativity. In *Semantics and Linguistic Theory*, Volume 30, pp. 409–429.

Nadathur, P. and H. Filip (2021). Telicity, teleological modality, and (non-) culmination. In *Proceedings of the West Coast Conference on Formal Linguistics*, Volume 39.

Nadathur, P. and S. Lauer (2020). Causal necessity, causal sufficiency, and the implications of causative verbs. *Glossa: a journal of general linguistics 5*(1).

Pearl, J. (2000). *Causality: Models, reasoning and inference.* Cambridge University Press.

Pearl, J. and D. Mackenzie (2018). *The book of why: The new science of cause and effect.* Basic Books.

Portner, P. (1998). The progressive in modal semantics. *Language 74*(4), 760–87.

Schlenker, P. (2003). A plea for monsters. *Linguistics and philosophy 26*(1), 29–120.

Setiya, K. (2022). Intention. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2022 ed.). Metaphysics Research Lab, Stanford University.

Smith, C. (1991). *The parameter of aspect.* Dordrecht: Kluwer.

Stalnaker, R. (2002). Common ground. *Linguistics and philosophy 25*(5/6), 701–721.

Talmy, L. (2000). *Towards a Cognitive Semantics.* Cambridge, Mass.: MIT Press.

Thelin, N. B. (1990). Verbal aspect in discourse: On the state of the art. *Verbal aspect in discourse*, 3–88.

Thelin, N. B. (1991). On the concept of time: Prolegomena to a theory of aspect and tense in narrative discourse. *New Vistas in Grammar: Invariance and Variation.–Amsterdam*, 261–285.

Timberlake, A. (2004). *A reference grammar of Russian.* Cambridge University Press.

Von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of semantics 19*(3), 245–274.

Zinova, Y. (2021). *Russian verbal prefixation: A frame semantic analysis (Volume 8).* Language Science Press.

Zinova, Y. and H. Filip (2014). Meaning components in the constitution of Russian verbs: Presuppositions or implicatures? In *Semantics and linguistic theory*, Volume 24, pp. 353–372.