Modeling explanatory influences for negative perfective declaratives and imperatives in Russian

May 26, 2025

Abstract

Russian negative perfectives have unexpected construals not predicted by existing analyses. These construals involve the existence of additional *influences*: potential causes or causal conditions whose presence may or may not result in a particular outcome. Such influences can explain why a particular outcome was expected or feasible, even if it did not occur. We argue that the perfective itself presupposes such an influence on the event described by the verb phrase. This influence can correspond either to the agent's intentional state or to a particular circumstance. Further, in the case of negative perfective declaratives, we argue that an additional influence must be inferred to explain the otherwise inexplicable outcome that an expected result does not occur. We extend this theory to negative perfective imperatives, where the addressee is warned to prevent an undesirable event that would otherwise be expected to happen.

1 Introduction

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) perfective sentences contribute a meaning component according to which the event, even though ultimately not realized, was at some point expected or feasible. We capture this component by arguing that the perfective contributes the presupposed existence of an influence (an eventuality which, if not overcome by some other influence, would cause the event to happen). Further, in negative

perfective sentences, an overcoming influence is inferred to explain the fact that the event failed to be realized or to reach completion in the end. This overall picture results from 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; Schulz, 2011), which easily represent such inexplicable outcomes and multiple influences.

The paper is structured as follows. Section 2 introduces basic and well-documented facts about the perfective/imperfective distinction in Russian. In Section 3, we turn to the discussion of additional meaning components associated with the perfective aspect, which are either not observed in the literature at all or have received much less attention. These include event specificity ("temporal definiteness" in the terminology employed by Leinonen (1982); Thelin (1990, 1991) as well as Dickey (2000, 2018, 2020)) and feasibility. Having established the data, in Section 4, we propose a formal analysis of these sentences within a causal model framework, along the lines of Schulz (2011). Section 5 continues by showing how denotations of sentences with perfective verbs and negation are built. The compositional analysis based on a formal treatment of causal models is provided in Section 6. In Section 7, we address the ways in which a range of meaning components, such as event specificity, feasibility, maximality, etc. follow from the proposed analysis. Section 8 is dedicated to negative perfective imperatives, and Section 9 compares our analysis to possible worlds approaches for unrealized events. Section 10 concludes the discussion.

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.PFVE car
'Vasja washed a / the car.'

This sentence entails that the event of washing the car reached its normal completion, *i.e.* all the (relevant) parts of the car have been washed. 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.IMPF car
'Vasja washed / was washing / used to wash a / the car.'

[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.

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 meanings of the resulting sentences differ. 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.

3 Perfective verbs: Event specificity and feasibility

3.1 Event Specificity

There exists a strong sense 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.PFVE
'Lena hasn't phoned.'
b. Lena ne zvonila.
Lena NEG phoned.IMPF
'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]

[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 same kind of contrast can be observed in interrogatives:

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

(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.IMPF Lena

---

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.IMPF *War and Peace*
'I haven't read *War and Peace*.'
b. Ja ne pročitala *Vojnu i mir*.
I NEG read.PFVE *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.

'Ivan has met with Lena (at least once)'.

b. Ivan vstretil Lenu
Ivan met.PFVE 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.PFVE Masha
'Dima has never met Masha.'

b. *Esli Dima kodga-libo/kogda-nibud' vstretil Mašu, on ejo ne
if Dima ever/ever met.PFVE Masha he her NEG
zabudet.
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, 2000, 2018). 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.PFVE
'Ultimately/in the end, he never called her.'

[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.

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, 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.IMPF Ivan
'I didn't kill Ivan.'

b. Ja ne ubila Ivana.
I NEG killed.PFVE 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 stem 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. For example, (9) may be uttered if Misha wrote a part of the letter but didn't complete the letter (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 stem did not even begin. The fact that sentences with

[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).

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, feasible, or 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.PFVE 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, she was thwarted. (N.B.: *defeasibility* is 'defeat-ability', not 'de-feasability').

4 Modeling event specificity, feasibility, and defeasibility

To model Russian negative perfective sentences, we start from the components of meaning 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 interferred.

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".[6] The

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

[6] An anonymous reviewer rightly points out that it's possible to have true progressive sentences about the past where it was not feasible or likely, from the perspective of the past, that the event would be completed. For example, suppose that, against all odds, Henriesa crossed a dangerous minefield yesterday. In the present, we can truly say *At 2pm yesterday, Henriesa was crossing the minefield.* This sentence is true even though it was not feasible at 2pm yesterday that she would cross the minefield. This fact seems to indicate on its face that feasibility is not a part of imperfective/progressive semantics. However, we could still say that

idea is that an outcome is feasible if in these worlds, where things proceed in accordance with their properties and circumstances in the actual world—that is, without interruption—the outcome occurs. But the actual world could either be in this 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 here; instead we will define explicitly defeasible causal relations between (occurrences of) eventualities, which will allow us to view them as *influences*: potential causes or causing conditions whose presence may or may not result in a particular outcome. Influences tend to make a particular outcome occur given certain background conditions, but another influence can make the outcome different, by overcoming the first influence's effect. We discuss some issues around the choice of explicit causal relations versus quantification over possible worlds below in section 9.

The kind of framework we need is one that has the ability to represent multiple influences on whether an eventuality occurs, such that the effect of an influence can be defeated or overcome by another influence. The causal relations we use must be fit for this purpose. In the most prominent linguistic treatment of causation at the moment, causal relations are written as "e_1 CAUSE e_2" statements; causal relations are taken to relate Davidsonian events, and multiple influences are not represented. This is not because multiple influences are incompatible with Davidsonian event arguments; rather, it is only because CAUSE has not been defined to admit multiple influences. Typically, something identified as "the cause" (e_1) is given pride of place (Bar-Asher Siegal and Boneh, 2020), and any other influencing conditions are relegated to expression through quantification over the preferred set of worlds. An "e_1 CAUSE e_2" statement is taken to entail that the outcome e_2 occurs.

Exactly because e_1 CAUSE e_2 is taken to entail that the outcome e_2 occurs, something like possible worlds is needed if the outcome does not occur and if we want to preserve the causal relation. We say that 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).

But there exist other approaches to causation that do allow multiple, defeasible influences. Using one of these theories means that we don't have to appeal to quantification over possible worlds in order to say that there is a causal relation when the outcome is defeated. For example, one can represent causation via force dynamics (Talmy, 2000), which allows for multiple influences by representing multiple forces that interact on an object. Here we will use causal modeling, a related but more general and powerful approach, which allows us

it is, and the feasibility seems to be relativized to time. We could say that *Henriesa is crossing the minefield*, uttered at 2pm yesterday, is either false or has an unclear truth value, and that this is precisely because her crossing is not viewed as feasible. From the present perspective, however, she clearly crossed, so somehow it was feasible. We think this point is well-aligned with the use of sets of worlds relativized to times to explain progressive semantics. Likewise, a feasibility requirement on Russian perfectives is satisfied in all past *affirmative* perfectives; the eventuality occurred, so somehow there must have been a way for it to occur.

to represent the influences of multiple potential causes and causal conditions, not just on objects but on outcomes in general.

A causal model is a formal representation of the structure that causal relations give to a conceptual model of the world.[7] Causal models are formally constructed to include a causal structure and a valuation function.[8] The causal structure is represented by a directed acyclic graph (DAGs). In a DAG, there is a set of variables that are the vertices (or nodes) of the graph. These variables can have values. In our case, the variables are *whether*-propositions (Kaufmann, 2013; Nadathur and Lauer, 2020, *a.o.*) and the values they can have are truth values.

The variables are connected by a set of edges (or "arrows"), hence they are "directed".[9] The causal model conveys that the (truth) values of some variables influence the (truth) values of other variables, according to the arrows of the graph, which show the *direction* of causation. For instance, $A? \rightarrow B?$ represents that the value of the *whether*-proposition $B?$ is dependent in some way on the value of the *whether*-proposition $A?$, without further specifying which values of $A?$ and $B?$ go together.

The valuation table that defines the function gives more information about which values go together, according to 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 variables, or nodes, correspond to *whether*-propositions: In this model, we see that whether there is fire depends on both whether there is lightning and also whether there is a lit match.

Variables that have arrows pointing at them are called *endogenous*: their value depends on the values of the variables pointing at them. Variables with no arrows pointing at them are called *exogenous* or *background* variables; their values do not depend on the values of other variables. Any variable with an arrow coming out of it can be understood as an *influence* in the sense we have been discussing.

The move to explictly treat the nodes as *whether*-propositions is just a notational choice. Some authors write the labels for the variables as propositions instead (*e.g.*, "there is lightning" instead of "whether there is lightning"), but the meaning is the same. The explicit use of *whether* in the label just reminds us more explicitly that the node, which is the relatum of the causal relation represented by the arrow, corresponds to a proposition that does not have a truth value. The reason why the nodes themselves are valueless propositions

[7] For a gentle introduction to causal models, see Chapter 1 of Pearl and Mackenzie (2018). Causal models have been perhaps most vigorously explicated by Pearl (2000) and see also work within linguistics such as Schulz (2011); 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); Nadathur (2023).

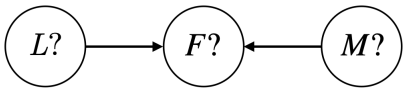
[8] Or equivalently, an equation, which gives another name for these models: "structural equation models."

[9] And they are "acyclic" because there is no way to follow the arrows in a circle; we are not tempted to have such a configuration here so this requirement is not an important one for our purposes.

is that, in order to tell if the value of one node really depends on another, we have to inspect different values in different courses of events, where the courses of events are represented by lines in the valuation table.

We explicitly treat the *whether*-proposition variables as denoting type $\langle s, t\rangle$ functions of the form $\lambda s\ .\ \ldots s \ldots$, which will be useful to the compositional analysis. We name these *whether*-proposition variables using a letter with a question mark, *e.g.* $A?$ instead of A, to remind us of this choice. This naming convention also helps us to avoid confusing the causal model arrow with the material conditional arrow. When we see $A? \rightarrow B?$, we can be sure we are dealing with the causal model arrow and not the material conditional arrow as in the expression $A \rightarrow B$.

When we get more formal about these variables in Section 6, we will write $X?(s) = 1$, but for now we will elide the situation argument and write $X? = 1$.



$L?$: whether there is lightning
$M?$: whether there is a lit match
$F?$: whether there is fire

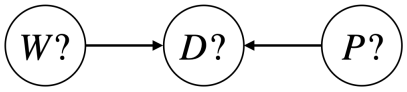
$L?$	$M?$	$F?$
1	1	1
1	0	1
0	1	1
0	0	0

Figure 1: A causal model whose causal structure is a collider

The causal structure in Figure 1 is called a "collider" structure, since two influences in a sense collide, both influencing the same node. The possible values of each node are the truth values of the associated proposition. We can represent the value of $F?$ given the values of $L?$ and $M?$: so the value of $F?$ is given by a function over the possible values of $L?$ and $M?$. In this case, that function happens to be the *OR* function.

Recall that a causal model is comprised of a causal structure along with a valuation function. The same causal structure can be used in combination with a different valuation function, as shown in Figure 2. The function need not even be a familiar Boolean function from propositional logic. In Figure 2, for instance, the plug prevents the water from draining, and the function is not familiar from propositional logic, but represents a prevention or an overcoming: The presence of the plug prevents the water from draining, or equivalently, overcomes the influence of the existence of water in the tub on whether water drains from the tub.

As mentioned above, an arrow belongs in the model exactly when one node's value depends causally on the other node's value. Looking at all the lines of the valuation functions in Figure 2, we can verify that whether the value of the endogenous variable $D?$ is 1 or 0 depends on the values of both exogenous variables $W?$ and $P?$ together. Given this particular valuation table in Figure



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

$W?$	$P?$	$D?$
1	1	0
1	0	1
0	1	0
0	0	0

Figure 2: Same structure, different function

2, we can also speak of $P?$ as an influence that overcomes the influence of $W?$. This is because when $P? = 0$, the value of $D?$ depends on the value of $W?$, such that $W? = 1$ makes $D? = 1$. But when $P? = 1$, and $W? = 1$, $D? = 0$.

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. The flexibility of causal models in this regard means that a node's presence in a model does not mean that the state of affairs it describes actually exists. When we do want to consider a particular state of affairs, we can consider the line in the valuation function that corresponds to that state of affairs.

Cases of prevention or overcoming as in Figure 2 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 one influence defeating or overcoming another. Recall that in examples like (10), the intuition is that there was an influence that feasibly would have caused the event of Anna killing Ivan, but something prevented them from doing so (= defeated the outcome).

(10) Anna ne ubila Ivana.
Anna NEG killed.PFVE Ivan
'(Ultimately), Anna didn't kill Ivan.'

To be more specific, we can identify two kinds of contexts that would make (10) true (and see Martin (2020) and references therein for antecedents to these in the literature):

- a "failed-event" context where Anna either wanted to initiate an event that would feasibly kill Ivan, or was in a circumstance that would have feasibly led to her initiating such an event; either way, the event did not occur, and Ivan remains alive.
- a "failed-result" context where Anna either wanted to initiate an event that would feasibly kill Ivan, or was in a circumstance that would have feasibly led to her initiating such an event; either way, the event occurred but did not result in Ivan's death, and Ivan remains alive.

To preview our analysis of (10), we will propose that the use of the perfective

requires the speaker to presuppose the existence of an influence on the event described by the verb stem. 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. Negation is what ensures the non-occurrence of the event. This means that we have a model as in Figure 3 below, where $X?$ represents whether Anna has an intention to kill Ivan.[10]

The causal model itself is presupposed: that is, regardless of whether the sentence is affirmative or negative, it is presupposed that whether Anna had the intention influences whether she did the action, which in turn influences whether Ivan died. In other words, the intention was of the kind that might feasibly be expected to cause the action, and the action was of the kind that might feasibly be expected to cause the result, in the absence of any preventing influence.

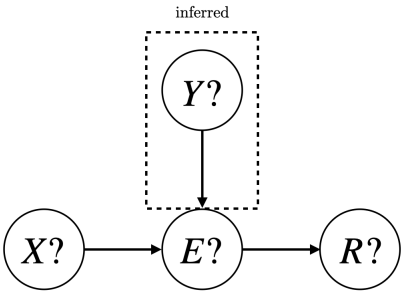
Additionally, a truth value 1 of the node $X?$ is apparently presupposed, since it is not cancelled under negation or in interrogatives. We will represent the presupposed truth value here by putting the values of $X?$ in the valuation table in a gray background, noting that the truth value in that column is always 1—this is what it means for a truth value to not be at issue, there is no alternative presented. $E?$ represents whether there occurs an event, with Anna as the agent, and which is feasibly expected to kill Ivan; and $R?$ represents whether Ivan is dead.

However, we also know, because of negation, that at least one of $E?$ and $R?$ has the value 0. Consider the failed-event context, where they both have the value 0. The combination of truth values where $X? = 1$, $E? = 0$, and $R? = 0$ is not a line in the valuation function. Thus, given $I? = 1$, the values of $E?$ is *inexplicable* in the model, though the value of $R? = 1$ given that $E? = 1$ is explicable. In effect, the fact that negation only operates on some of the truth values introduces a kind of clash, meaning that the model as it stands cannot be used to represent reality, as the combination of truth values we are interested in do not correspond to any line of the valuation table. The hearer must then adduce a different model, one where the combination $I? = 1$, $E? = 0$, and $R? = 0$ is in fact a line in the valuation function of the model, still keeping in line with world knowledge.

The way to do this is to add an additional node $Y?$ to the model with an arrow pointing from it to $E?$,[11] and with the understanding that $Y?$'s influence on $E?$'s value is sufficient to overcome the influence of $X?$ on $E?$'s value. $Y?$ could be, for instance, "whether Ivan is not in the right place for Anna to be able to kill him". In order for the new model containing $Y?$ to account for the fact that $X? = 1$ but $E? = 0$, the $Y?$ condition must influence the $E?$ event

[10] Some philosophically-minded readers may at this point be skeptical, as for many philosophical views, intentions are not straightforwardly 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.

[11] The additional influence in the failed-event case needs to point at $E?$. As we'll see, in the failed-result case, it has to point at $R?$.



not-at-issue	inferred	at-issue	
$X?$	$Y?$	$E?$	$R?$
1	1	0	0
1	0	1	1

$X?$: whether Anna intends to kill Ivan
$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$Y?$: whether there is a condition sufficient to prevent $E? = 1$ given $X? = 1$
$R?$: whether Ivan is dead

Failed event: $X? = 1$, $E? = 0$, $R? = 0$ exists in a line in the table, so this model is ok

Figure 3: Model for failed-event context

sufficiently to ensure that $E? = 0$, even though $X? = 1$.

We will explain our proposal in further detail below. As we do, keep in mind that these causal models are not to be thought of as ordinary semantic models, where a sentence is interpreted on the model and thereby we get a truth value. Instead, in this approach, sentences may contribute a causal model to a Common Ground, or as a kind of model under discussion, intended to represent reality (Baglini and Francez, 2015; Nadathur and Lauer, 2020; Copley and Mari, 2022). To reflect this, we will have a contextual variable that represents the model so that we can build the model compositionally, with different parts of the sentence specifying different parts of the model, as well as different truth values for nodes.

These contributions from parts of the sentence can be either at-issue or not-at-issue. For instance, the value 1 of the node $X?$ contributed by the perfective model is treated as not-at-issue (and accordingly, accommodated, in case the hearer does not already know that Anna intended to kill Ivan).

Finally, there is a part of the causal structure—namely $Y?$ and the arrow coming out of it—which is not contributed by any part of the sentence. Instead, this part of the structure has to be inferred, given the interaction between negation and perfective that leads to an inexplicable truth value.

5 Building the meaning

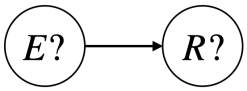
To build the meaning of Russian negative perfectives, we will address the contributions to the causal model made by the verb stem, by the perfective operator,

and by negation, and then show how this picture derives the two construals discussed above, namely failed-event and failed-result. Afterwards, in Section 6, we will relate these models to a compositional semantics.

5.1 The VP: Events and results

We consider telic verb phrases (VPs) to introduce a model that reflects that whether a certain kind of event occurs, influences whether a certain kind of 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).

For the case we are interested in, in (10) above, Anna kill Ivan contributes an event node $E?$ and a result node $R?$, as well as a causal relation between them and a valuation function that is the identity function. The model contributed by Anna kill Ivan is given in Figure 4 below:



at-issue

$E?$	$R?$
1	1
0	0

$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$R?$: whether Ivan is dead

Figure 4: Model for Anna kill Ivan

Anna kill Ivan contributes the causal model in Figure 4 as not-at-issue material. The idea here is that the model in Figure 4 expresses the complicated claim that this kind of action by Anna, all else being equal (i.e., in the absence of other influences) would feasibly cause Ivan's death, and if she didn't do it the death wouldn't happen. This claim is not-at-issue because as we have seen, it survives under negation.

It may seem incredibly counterintuitive to say that any part of the causal contribution of the VP is not-at-issue, but the claim is merely that the causal model (consisting of the causal structure, the labels of the nodes, and the valuation table) is not-at-issue. Figure 4 does not itself contribute the actual values of either $E?$ or $R?$. These values *are* at-issue, and in declarative sentences, the values are decided, as usual, by whether the sentence is affirmative or negative (see section 5.3 below). In the case of affirmation, both nodes get a value of 1; negation is discussed below in section 5.3.

5.2 The contribution of the perfective in Russian

We propose that the perfective in Russian contributes a specific influence whose existence is not at-issue, whose value is presupposed to be 1, and which has an arrow pointing toward the event node $E?$, influencing $E?$ to have a value of 1. This new influence we notate as $X?$.[12]

$X?$ is present wherever perfective morphology is present, that is, even where there is no negation. The addition of $X?$, we claim, is what adds the idea that $E?$ is specific and feasible. The existence of an $X?$ influence is not-at-issue because specificity and feasibility are maintained under negation (3a), in interrogatives (4a), and in the antecedent of conditional sentences as in (11):

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

Unlike the imperfective example in (11a), the perfective example in (11b) requires a context in which Dima was expected or supposed to read War and Peace.

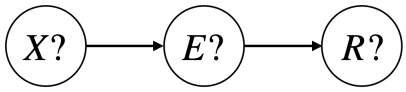
The *whether*-proposition corresponding to $X?$ can be the subject's intention, which makes the model in Figure 5 a representation for intentional action. Or alternatively, $X?$ could represent a circumstance such as her being bumped while intending to shoot someone else, making her gun point at Ivan. This option would make Figure 5 a representation for mere (accidental) actions and non-agentive events (and see Martin (2020) for an argument for treating the latter two similarly).

Thus, there are two possibilities for $X?$, the node that makes $E? = 1$ feasible: either $X?$ is "whether the subject had an intention for $R? = 1$, or it is a circumstance, "whether there was a circumstance that made it feasible that $R? = 1$".

Recall that we represent not-at-issue values on a gray background, with at-issue meaning on a white background. The truth value of $X?$ is not-at-issue, presupposed always to be 1, so it is on the gray background in Figure 5. Remember that the causal structure, the *whether*-proposition labels of the nodes, and the character of the dependency between the values of $E?$ and $R?$

[12] In principle $X?$ could also have a value of 0 as long as its influence was preventive, *i.e.*, $X? = 0$ would tend to make it that $E? = 0$. We elide this possibility throughout.

(represented by the valuation table) are also not-at-issue.



not-at-issue	at-issue	
$X?$	$E?$	$R?$
1	0	0
1	1	1

$X?$: whether Anna intends to kill Ivan (or: whether some circumstance makes it feasible that Anna would kill Ivan)
$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$R?$: whether Ivan is dead

Figure 5: Model contributed by Pfve(Anna kill Ivan)

5.3 Negation

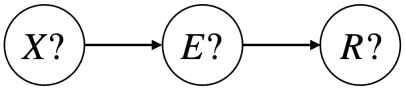
Negation on this account acts on the at-issue meaning, namely the at-issue truth values of $E?$ and $R?$. We get two construals, shown in Figure 6: "failed-event" and "failed-result".

On the failed-event construal, where $E?$ is set to 0, all else being equal, $R?$ is also set to 0 because the model says that when $E? = 0$, $R?$ is also 0. To see this, note that while there is no line of the valuation table of this model which has $E? = 1$ and $R? = 0$, there is a line where $E? = 0$ and $R? = 0$. So $E? = 0$ is the inexplicable value; we mark it with an exclamation point. On the failed-result construal, where $E? = 1$ but $R? = 0$, the inexplicable value is that of $R?$, so on that construal, it is the value of $R?$ that we mark with an exclamation point. Here too, the actual valuation is not a line in the valuation table.[13]

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 (10), reproduced here as (12):

[13] A formally possible alternative structure to the $X? \rightarrow E? \rightarrow R?$ structure we propose here is $X? \rightarrow R?$, where the event is not represented by the sentence itself. On this structure, even a failed-event construal would be represented with a result collider, there being no $E?$ node. To get the failed-event construal, $Y?$ would have to have the meaning that we here attribute to $E?$ "whether an Anna-initiated event that would feasibly kill Ivan takes place", but the inferred truth value of $Y?$ would have to be 0 so as to prevent $R? = 1$. In that case negation would act only on the at-issue truth value of $R?$, being now the only at-issue truth value. While this move is formally possible, we think we should represent the event in the meaning of the verb stem, especially for accomplishments, so we wouldn't want to delete $E?$ from the model.



not-at-issue	at-issue	
$X?$	$E?$	$R?$
1	0	0
1	1	1

$X?$: whether Anna intends to kill Ivan (or: whether some circumstance makes it feasible that Anna would kill Ivan)
$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$R?$: whether Ivan is dead

Failed event: $X? = 1$, $E? = 0!$, $R? = 0$ does not exist in a line in the table
$\Rightarrow$ model must be changed
Failed result: $X? = 1$, $E? = 1$, $R? = 0!$ does not exist in a line in the table
$\Rightarrow$ model must be changed

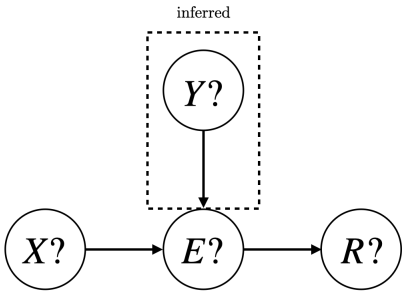
Figure 6: Model contributed by NEG(PFVE(*Anna kill Ivan*), with failed event and failed result valuations that are not in the model's valuation table

(12) Anna ne ubila Ivana.
Anna NEG killed.PFVE Ivan
'(Ultimately), Anna didn't kill Ivan.' (= (10))

Consider now the failed-event context, where the event does not even start. Let's suppose that $X?$ is whether Anna intends to kill Ivan. $X? \rightarrow E?$ is part of the model, but $E?$ has the value 0. Since $E?$ is whether there is an Anna-initiated event that feasibly results in Ivan's death, this means that there is no such event even though she had an intention to kill Ivan. Note again that the valuation $X? = 1$, $E? = 0$, $R? = 0$ is not a line in the valuation table of the model. This means that the model in Figure 6 is not an appropriate model to use to represent reality in this case. What should be done? The model must be altered in such a way as to make a valuation function that includes the actual valuation. This has the effect of explaining the actual valuation, and in particular, the truth value that was inexplicable in the other model. Speakers, then, use negative perfectives exactly when they wish to convey that the altered model is a good representation of reality.

In the failed-event case we are interested in, the actual valuation is $X? = 1$, $E? = 0$, and $R? = 0$. Again, $E?$'s value is the one that needs explaining. In the case of (12), the hearer infers an influence on $E?$ that explains $E?$'s otherwise inexplicable value. We represent this new inferred influence in Figure 7 as $Y?$. Like $X?$, $Y?$ can presumably be either an intention or a circumstance. $Y?$ affects whether any event of the description occurs. We'll assume that $Y?$

$= 1$ tends to ensure that $E? = 0$, overcoming the influence of $X?$ which on its own would tend to make $E? = 1$. This new model includes the actual valuation as a line in its valuation table, so this model is a good representation of reality.



not-at-issue	inferred	at-issue	
$X?$	$Y?$	$E?$	$R?$
1	1	0	0
1	0	1	1

$X?$: whether Anna intends to kill Ivan
$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$Y?$: whether there is a condition sufficient to prevent $E? = 1$ given $X? = 1$
$R?$: whether Ivan is dead

Failed event: $X? = 1$, $E? = 0$, $R? = 0$ exists in a line in the table, so this model is ok

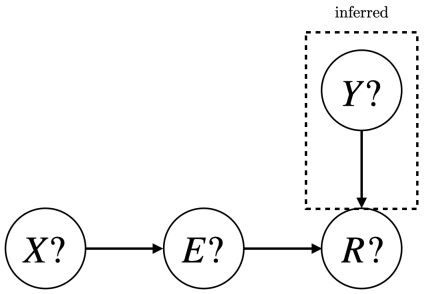
Figure 7: Failed event: An "event collider", with influence $Y?$ inferred to explain the otherwise inexplicable value of $E?$

Similarly, in the failed-result context, a truth value is inexplicable; this time it is the truth value of $R?$. In a similar fashion to the failed-event context, the inexplicable value of $R?$ is inferred to result from an inferred influence $Y?$, as shown in the "result collider" in Figure 8. Again, this new model includes the actual valuation as a line in its valuation table, so this model is a good representation of reality.

In both of these contexts, the easiest scenario to think about is the one where $X?$ = whether Anna has the intention to kill Ivan. But remember that $X?$ need not always be about her intention; it could be instead about one or more circumstances that would tend to make feasible an action she performs accidentally. Recall, 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, $X?$ would be "whether she is pushed in such a way as to make it feasible that she would do an action that would feasibly kill Ivan".

6 Formal analysis

To build our formal analysis in this section, we will define causal models, following Schulz (2011) in order to use her notion of causal entailment to define



not-at-issue		at-issue inferred	at-issue
$X?$	$E?$	$Y?$	$R?$
1	1	1	0
1	1	0	1
1	0	1	0
1	0	0	0

$X?$: whether Anna intends to kill Ivan
$E?$: whether $\exists e$ such that *agent*(Anna, e), and e would feasibly kill Ivan
$Y?$: whether there is a condition sufficient to prevent $R? = 1$ given $E? = 1$
$R?$: whether Ivan is dead

Failed result: $X? = 1$, $E?= 1$, $R? = 0$ exists in a line in the table, so this model is ok

Figure 8: Failed result: A "result collider" with influence $Y?$ inferred to explain the otherwise inexplicable value of $R?$

inexplicable truth values. This move will allow us to present a compositional semantics for declarative Russian negative perfectives.[14]

6.1 Definition of causal models

First, we define causal models, borrowing from Schulz (2011). We slightly alter this framework to align it more closely with the lambda calculus for compositionality. Schulz's framework treats situations as functions from variables (propositions) to truth values, with variables treated as atomic (so it has us write, *e.g.*, $s(P) = 1$). However, this is contrary practice to lambda calculus, which treats situations as atomic, and treats propositions as functions (so we write instead, *e.g.*, $P(s) = 1$). Another reason to treat situations as atomic is that we need s to refer to the world somehow, so that we can write things like "$agent(y, e)$ in s", which can be true of a situation in reality but not (at least not straightforwardly) of a function from propositions to truth values. But we like Schulz's approach otherwise, because, as we'll explain below, it allows us to define inexplicable truth values from the causal model itself. So, we will use her framework, but in order to align her approach with the lambda calculus, we will treat situations as atomic, referring to parts of reality.

[14] Another framework that we could have used for our purposes is that of Alonso-Ovalle and Hsieh (2017, 2021). We have chosen to use the framework of Schulz (2011) instead because it makes explicit things that Alonso-Ovalle and Hsieh (2017, 2021) leave implicit. This choice is discussed more in Section 7 below.

We differ from Schulz also in making the meaning of variables explicit in lambda notation, as discussed above in Section 4. The variables represent *whether* propositions, which can be represented as $\lambda s \ . \ P(s)$. That is, they are functions that take a situation argument and return a truth value. Again, to be perfectly clear about this move in our notation, we notate the variables with question marks.

Another explicit difference between us and Schulz (2011) is that we say explicitly in the definition of causal models (in part 3d. below) that the model has a causal interpretation. The reason we do this is because in this analysis, the causal meaning of the sentences we are interested in can come only from the causal model. Without the condition in 3d., the model is a strictly formal object that gives dependencies of truth values of variables with respect to each other; we need these dependencies to specifically be causal in nature. It may be that, due to world knowledge, the only appropriate way to interpret such dependencies would be to say that they are causal dependencies, but to really make sure our denotations have causal meaning, we add the condition in 3d.

Definition 1 (Causal models): A causal model, also known as a *dynamics*, is a tuple $\mathcal{D} = \langle \Sigma, \beta, F \rangle$ where:

1. Σ is a finite set of variables denoting *whether*-propositions of type $\langle s, t \rangle$
   e.g. $\{X?,$
   $Y?, Z? \ldots\}$
2. $\beta \subseteq \Sigma$ is the set of background (exogenous) variables, and
3. F is a function that maps elements $X? \in \xi = \Sigma - \beta$ to tuples $\langle Z_{X?}, f_{X?} \rangle$, where
   (a) $Z_{X?}$ is an n-tuple of elements of Σ
   (b) $f_{X?} : \{0,1\}^n \rightarrow \{0,1\}$ is a two-valued truth function from n-tuples on $\{0,1\}$ to $\{0,1\}$
   (c) F is rooted in β (explained below in Definition 2)
   (d) F has a causal interpretation.

How does this definition relate to the diagrams? The information represented by the arrows in the diagrams, along with the information represented by the valuation table, is here characterized by the function $f_X?$. This function relates the truth values of the immediate causal ancestors, or "parents" of $X?$, namely, the variables "pointing" at $X?$, to the value of $X?$. The set of background (or exogenous) variables β is the set of variables with no arrows pointing at them; their values do not causally depend on other variables in the set Σ of variables. The set of endogenous variables ξ is the set of variables that have at least one arrow pointing at them; their values depend on the value(s) of at least one other variable in the set Σ of variables.

Rootedness ensures that "walking backwards" through the arrows gets us only to variables in β, ensuring that $\mathcal{D}$ is acyclic (which is necessary for it to be a causal model):

Definition 2 (Rootedness): Let $\beta \subseteq \Sigma$ be a set of variables, and F a function mapping elements $X?$ of $\chi = \Sigma - \beta$ to tuples $\langle Z_{X?}, f_{X?} \rangle$ as above. Let R_F be the relation that holds between the variables $X?, Y?$ if $Y? \in Z_{X?}$. Let R_F^T be the transitive closure of R_F. F is rooted in β if $\langle \Sigma, R_F^T \rangle$ is a partially-ordered set and β is the set of its minimal elements.

6.2 Causal entailment

Schulz presents a notion of causal entailment that will be useful for defining inexplicable truth values. The idea is that facts about influences can, in a certain sense, entail facts about what they influence. This allows us to determine the causal consequences of influences according to the model.

To begin with, we need a way to express that a truth value could start out undetermined, and then be influenced to have a certain value. For this, Schulz (2011) proposes a trivalent (= three valued) logic, such that the truth values are 1, 0, and u (for "undetermined"). A "fixed-point" operator $\mathcal{T}_\mathcal{D}$ maps situations s to new situations $\mathcal{T}_\mathcal{D}(s)$ in order to calculate the causal consequences of the collection of truth values arising when the variables take s as their argument. Effectively, $\mathcal{T}_D$ extends a situation s about which less is known or determined to a situation $\mathcal{T}_D(s)$ about which more is known or determined, by running the dynamics on s (once).

Definition 3 (Causal update with the fixed-point operator $\mathcal{T}_\mathcal{D}$): Let $\mathcal{D}$ be a dynamics and s a situation. Let $X?$ be an endogenous variable whose causal parents have determined truth values. We define the situation $\mathcal{T}_\mathcal{D}(s)$ by:

1. if $X? \in \beta$, then $\mathcal{T}_\mathcal{D}(s) = X?(s)$
2. if $X? \in \xi$, with $Z_{X?} = \{X?_1, \ldots, X?_n\}$, then
   (a) if $X?(s) = u$ and $f_{X?}(X?_1(s), \ldots X?_n(s))$ is defined, $X?(\mathcal{T}_\mathcal{D}(s)) = f_{X?}(X?_1(s), \ldots X?_n(s))$
   (b) if $X?(s) \neq u$ or $f_{X?}(X?_1(s), \ldots X?_n(s))$ is undefined, $X?(\mathcal{T}_\mathcal{D}(s)) = X?(s)$.

$\mathcal{T}_\mathcal{D}$ is called a "fixed-point" operator because after finitely many iterations of applying $\mathcal{T}_\mathcal{D}$, the resulting situation will eventually show no further development. This fixed-point situation, which is called s^*, is used in Schulz's definition of causal entailment:

Definition 4 (Causal entailment): Let $\mathcal{D}$ be a dynamics such that $\mathcal{D} = \{\Sigma, \beta, F\}$. For all $X? \in \Sigma$, and for any truth value t, a situation s causally entails that $X?(s) = t$ iff the fixed point s^* of $\mathcal{T}_\mathcal{D}$ is such that $X?(s^*) = t$.

6.3 Inexplicable truth values

Now that we have defined our causal models, we need to add to this picture the idea that inexplicable truth values can arise, as well as what happens if

they do arise. For us, they arise from the composition of perfective verb forms with negation, and they are inexplicable precisely because they are not causally entailed by the truth values of their causal parents. Inexplicable truth values are defined in Definition 5 below; they must be determined values (*i.e,* either 1 or 0).

Definition 5 (Inexplicable truth values): A determined truth value t is inexplicable for a variable $X?$ in $\mathcal{D}$ iff $X? \neq t$ is causally entailed by the truth values of its causal parent(s) in $\mathcal{D}$.

For example, consider a dynamics $\mathcal{D} = \{\Sigma, \beta, F\}$ as in Figure 5 above, where $\Sigma = \{X?, E?, R?\}$, $\beta = \{R?\}$, and $F : E? \mapsto \langle\langle X?\rangle, =\rangle$ & $R? \mapsto \langle\langle E?\rangle, =\rangle$. Now consider a situation s where $X? = 1$ and $E? = 0$. The value 0 is inexplicable for $E?$ in $\mathcal{D}$ because the function F says that the only parent of $E?$ is $X?$, and the function $f_{E?}$ that gives us the truth value of $E?$ given the truth value of $X?$ is the identity function. Thus, given that its unique causal parent $X? = 1$, it is expected that $E? = 1$. If in fact $E? = 0$, the dynamics $\mathcal{D}$ gives us no way to explain that fact.

When an inexplicable truth value arises, the upshot is that the model, as it stands, can't represent the situation being described by the sentence. In other words, the actual situation being described is impossible according to the model. Thus, the model needs to be altered so that we can understand how the actual situation being described is a possible situation. We take the contributions of the verb stem and the perfective to the causal model under discussion to be non-exhaustive. That is, they contribute nodes and arrows but they do not make the claim that the contributed nodes and arrows are the only nodes and arrows in the model. Therefore an additional influence on the appropriate node can be inferred to explain the otherwise inexplicable truth value.

6.4 Compositionality and at-issue/not-at-issue meaning

As we argued above in Section 4, the causal model itself needs to be presupposed. We express this by presupposing conditions on a causal model $\mathcal{D}$ treated as a discourse parameter, a kind of "model under discussion", following Baglini and Francez (2015), Copley and Mari (2022), and Nadathur (2023). Lexical and functional items can put constraints on $\mathcal{D}$, via the definition of functional application in Definition 6 below:

Definition 6 (Functional application): For $[\![\alpha]\!]^{\mathcal{D}}, [\![\beta]\!]^{\mathcal{D}}, [\![\gamma]\!]^{\mathcal{D}}$ such that $[\![\beta]\!]^{\mathcal{D}}$ and $[\![\gamma]\!]^{\mathcal{D}}$ are the daughters of $[\![\alpha]\!]^{\mathcal{D}}$ and $[\![\gamma]\!]^{\mathcal{D}}$ is in the domain of $[\![\beta]\!]^{\mathcal{D}}$: $[\![\alpha]\!]^{\mathcal{D}}$ is defined iff $[\![\beta]\!]^{\mathcal{D}}$ and $[\![\gamma]\!]^{\mathcal{D}}$ are defined, and if defined, $[\![\alpha]\!]^{\mathcal{D}} = [\![\beta]\!]^{\mathcal{D}}([\![\gamma]\!]^{\mathcal{D}})$.

Both the verb stem and the perfective in fact do put constraints on $\mathcal{D}$, as shown in their denotations below. The verb stem denotation is shown in (13):

(13) $[\![\textsc{Kill}]\!]^{\mathcal{D}}(s)$ is defined iff
$\mathcal{D}$ is a dynamics such that $\{\Sigma, \beta, F\}$ and

$E?$, $R? \in \Sigma$ and
$E? = \lambda s \lambda y$. $\exists e$: $agent(y, e)$ in s and
$R? = \lambda s \lambda x$. x is dead in s and
$F(R?) = \langle \langle E? \rangle, = \rangle$.
If defined,
$[\![\textsc{Kill}]\!]^{\mathcal{D}}(s) = 1$ iff $(E?(y)(s) \& R?(x)(s)) = 1$ in s.

Although when the verb stem is considered on its own, $E?$ is a background variable, and therefore is in β, we do not put this as a condition in the verb stem denotation. The reason is that when perfective is added, $E?$ is no longer a background variable.

When perfective aspect is added, it puts a further condition on the causal model under discussion.[15] In addition to the two nodes from the verb stem, $\mathcal{D}$ must include an influence $X?$ on the first of the two nodes in that model. This means that perfective must somehow be able to identify the first of those two nodes. This could probably be done in various ways, but given the compositional framework we are using, we suggest that perfective is looking for a node that makes reference to an event, and that this node is most naturally and economically understood to be the $E?$ node introduced by the verb stem. If eventive verb stems are verb stems whose first or only node is an event node, this proposal has the advantage of predicting that perfective always co-occurs with eventivity, which seems to be the case.

Definition 7 (Definition of event node): A node $E?$ is an event node just in case it is of the form $\lambda s \ldots \exists e \ldots$.

(14) $[\![\textsc{Pfve}]\!]^{\mathcal{D}}([\![\alpha]\!]^{\mathcal{D}})(s)$ is defined iff
p is of type $\langle s, t \rangle$ and
$\mathcal{D}$ is a dynamics such that $\{\Sigma, \beta, F\}$ and
$X?$, $E? \in \Sigma$ and
$X? = \lambda s$. $X(s)$ and
$E?$ is an event node and
$F(E?) = \langle \langle X? \rangle, = \rangle$.
$[\![\textsc{Pfve}]\!]^{\mathcal{D}}([\![\alpha]\!]^{\mathcal{D}})(s) = 1$ iff $[\![\alpha]\!]^{\mathcal{D}} = 1$ in s.

Unlike the verb stem and perfective aspect, negation has nothing to say about the causal model under discussion, behaving in the usual fashion.

(15) $[\![\textsc{Neg}]\!]^{\mathcal{D}}([\![\alpha]\!]^{\mathcal{D}})(s) = 1$ iff $[\![\alpha]\!]^{\mathcal{D}} = 0$ in s.

Putting these pieces together, the following compositional derivation results:

(16) $[\![\textsc{Neg}]\!]^{\mathcal{D}}([\![\textsc{Pfve}]\!]^{\mathcal{D}}([\![\textsc{Kill}]\!]^{\mathcal{D}}))$(s) is defined iff
$\mathcal{D}$ is a dynamics such that $\mathcal{D} = \{\Sigma, \beta, F\}$ and
$X?$, $E?$, $R? \in \Sigma$ and
$E? = \lambda s \lambda y$. $\exists e$: $agent(y, e)$ in s and

[15] We assume that perfective is composed with the verb stem rather than a perfective form being stored in the lexicon, but for our current purposes, nothing hangs on this assumption.

$R? = \lambda s \lambda x$. x is dead in s and
$F(R?) = \langle\langle E?\rangle, =\rangle$ and
$F(E?) = \langle\langle X?\rangle, =\rangle$.
If defined,
$[\![\textsc{Neg}]\!]^{\mathcal{D}}([\![\textsc{Pfve}]\!]^{\mathcal{D}}([\![\textsc{Kill}]\!]^{\mathcal{D}}))(s) = 1$ iff $(E?(y)(s)\ \ \&\ R?(x)(s)) = 0$ in s.

This denotation requires that either $X?(s) = 1$ and $E?(s) = 0!$ (failed-event), or $X?(s) = 1$ and $R?(s) = 0!$ (failed-result). (The combination $E?(s) = 0!$ and $R?(s) = 0!$ is also possible, but this combination has two inexplicable truth values and would require two inferred influences to explain them.) Both failed-result and failed-event are impossible valuations given $\mathcal{D}$, and so represent impossible situations according to $\mathcal{D}$. But another node (and arrow) can be inferred, changing the model such that the actual situation described is a possible situation, given the changed model that includes $Y?$.

The way the denotations are written, adding a node is not a problem; we can just add it to Σ. Adding the arrow is trickier. For instance, in the failed result (result collider) case, consider that the denotation of the verb stem in (13) above includes the condition that $F(R?) = \langle\langle E?\rangle, =\rangle$. But if another arrow is added pointing from the new node $Y?$ to $E?$, then this condition is no longer true; it must be overwritten somehow by an Overcome function as below in (17) and in Figure 9 (and see the valuation table in Figures 7 and 8), and rewrite $F(R?)$ as in (17):

$E?$	$Y?$	$R?$
1	1	0
1	0	1
0	1	0
0	0	0

Figure 9: Overcome function: The influence of $Y?$ overcomes the influence of $E?$ on $R?$

(17) $\quad F(R?) = \langle\langle Y?, E?\rangle,$ Overcome$\rangle$

The function itself can probably be gotten from world knowledge, since we know what the function needs to express, namely that the influence of $Y?$ overcomes the influence of $E?$ to explain the fact that $R?(s) = 0$. However, since we want to tell a pragmatic reasoning story about inferring a new node in the model, it is inconvenient if this reasoning seems to require falsifying something that is already presupposed.

We can make this analysis work, however, if we can ensure that presupposed conditions defining F do not need to be falsified. That is, these conditions need

to be weaker. In the failed result case, for instance, we need to be able to say something about the influence of $E?$ on $R?$ that is maintained even though an influence $Y?$ on $R?$ is added, where the influence of $Y?$ overcomes the influence of $E?$. Following Kagan (2020), we need to treat influences more like forces in force-dynamic frameworks, such that they can be characterized both in the presence and absence of other influences.

To do this, we borrow Copley and Harley's (2015) idea that forces are defined by their *ceteris paribus* result: the result that would occur in the absence of other forces. In our framework here, the analogue of this idea is that presupposed conditions on F are to be understood as holding only in the absence of other influences. This way, adding another influence does not falsify the presupposed condition on F.

7 Discussion

Our account explains the presupposed and inferred influences in Russian negative perfectives through an influence $X?$ contributed by the perfective, and the need to avoid inexplicable values in the model, which gives rise the to the inferred influence $Y?$. The inexplicable values arise due to the interaction of negation with the contribution of the perfective. Essentially, the perfective presupposes that something was feasible, but the negated sentence asserts that even so, it didn't happen.

Negative perfectives vs. affirmative perfectives: Even in affirmative perfectives, we can still detect the presence of presupposed $X? = 1$, because of the specificity facts and their survival in questions and antecedents of conditionals (and we will say just below how the presupposition that $X? = 1$ yields the specificity facts). But affirmative perfectives do not trigger another influence (represented by $Y?$) at all. This is as expected, since in the absence of negation, the $E?$ and $R?$ values are 1, and these are explicable when $X? = 1$, even in a model that doesn't include $Y?$.

Specificity: The specificity effect associated with perfective clauses ((ia), (4a) above) is due to the presence of the additional influence $X?$. We are therefore dealing not with just any situation that $E?$ holds of, but rather with the specific situation that is caused by the specific intentional state or circumstance that characterizes $X?$. The *whether*-proposition of $X?$, along with the presupposed truth value 1 for $X?$, together give the identifying property which creates the specificity intuition. In this analysis, specificity is maintained, as desired, in *e.g.* (ia), despite the absence of temporal definiteness: (ia) asserts that a reading *War and Peace* event that would have been caused specifically by the $X?$ eventuality did not take place at any temporal interval.

Feasibility and defeasibility: The feasibility intuition is due both to the presuppositions that $X? = 1$, and the presupposition that $X?$ influences $E?$ in such a way that $X = 1$ tends to make it so that $E? = 1$ (encoded in the contextual variable $\mathcal{D}$). Feasibility is essentially having *ceteris paribus* causal conditions which are met in the situation at hand. So, when there is a structure

$A? \rightarrow B?$ and a valuation where $A? = 1$ and $B? = 0$, despite a valuation table that does not include this valuation, this says exactly that the occurrence of an event described by $B?$ was feasible but such an event did not occur; it was defeated.

Feasibility shows up independently of whether the sentence is affirmative or negative. In the negative, however, either $R?$ or both $E?$ and $R?$, must be equal to 0. Yet the truth value 1 of $X?$ is not-at-issue, so is not affected by negation. The effect is that in the negative, we retain the feasibility of $X?$ to influence $E?$.

Maximality: Recall that the notion of "maximality" of an event has been proposed as the core meaning of perfectives (Filip and Rothstein, 2006; Filip, 2008). Negative perfective sentences, on such an account, express non-maximality (or frank non-occurrence) of the event. We notice that in our system, for affirmative perfectives, $X? = 1$ is presupposed as the *only* explanation of the values in $E?$ and $R?$. In other words, the model seems to be "closed"—the speaker is committed to the idea that there are no other nodes. 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 those that are provided by the sentence, namely, the node $Y?$. Thus, the influences contributed by the sentence are not the maximal set of influences. We therefore can derive maximality of the event for affirmative perfectives in the sense of Filip and Rothstein (2006); Filip (2008) with a closed model (*i.e.,* exhaustivity of the causal influences given by the sentence itself) and the consequent lack of interruption of the event by an outside influence. 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 (namely, $X?$, $E?$ and $R?$) are not an exhaustive list of relevant influences.

8 Negative perfective imperatives

So far we have looked at Russian negative perfective verbs in declarative sentences. Our framework also provides an understanding of how Russian negative perfective verbs behave in imperative sentences.[16]

In general, in affirmative imperatives, both imperfective and perfective verbs are acceptable, as shown in (18):

(18) a. Otkryvaj okno!
open.IMP.IMPF window
'Open the window!'

b. Otkroj okno!
open.IMP.PFVE window
'Open the window!' Goncharov (2018) p. 106

[16]We thank an anonymous reviewer for suggesting we look at imperatives.

However, in the case of negative imperatives, while imperfective verbs are always possible, perfective verbs tend to be unacceptable, even when they refer to single bounded events, as shown in (19):

(19) a. Ne otkryvaj okno!
Neg open.Imp.Impf window
'Don't open the window!'
b. #Ne orkroj okno!
Neg open.Imp.Perf window
Intended:'Don't open the window!' Goncharov (2018) p. 106

There are, however, highly limited contexts in which negative perfective verbs are possible in imperative sentences, as in (20), where, for example, the sidewalk is icy, and the speaker suspects that the addressee is not paying enough attention or being careful enough to avoid falling down.

(20) Ostorožno! Ne upadi!
careful Neg fall.Imp.Perf
'Be careful! Don't fall down!' Goncharov (2018) p. 107

An approach in the literature (Bogusławski, 1985; Goncharov, 2018) characterizes Russian negative perfective imperatives as only being acceptable when the predicate describes a non-intentional action, on the basis of examples such as the one in (20). Dickey (2020) gives a more nuanced characterization, saying that the verb itself can be agentive, *i.e.* it can describe a potentially intentional action, as long as the action may result in undesirable consequences which the addressee may not be aware of. Dickey (2020, 571) refers to negative perfective imperative sentences as "preventives", which "serve to warn someone from doing something inadvertently." The idea is that something bad is going to happen if the addressee does not do something about it, and that the addressee is not aware of the danger, or at least not sufficiently attentive to it.

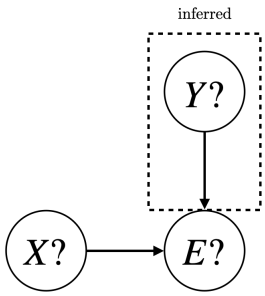
We find that the most acceptable examples of agentive-action preventives are ones where the speaker knows that the addressee has a disposition or urge to engage in the action. Consider, for instance, the imperative in (21a); the context is that the speaker knows the addressee is very angry and, because of that, likely has an urge to hit the other person. Similarly, in (21b), the speaker is (hyperbolically) warning the addressee not to buy everything in the store, conveying their belief that the addressee has a disposition to do so. We also find that the addition of words like *tol'ko* 'just', *požalujsta* 'please', and *ostorožno* 'careful' seems to be doing some work here to make the sentences in (20) acceptable.

(21) a. Tol'ko, požalujsta, ne udar' ego!
only please Neg hit.Imp.Perf him
'Just please don't hit him!'
b. Ostorožnee, ne skupi ves' magazin!
more.careful Neg buy.Imp.Perf whole store

'Be careful, don't buy the whole store!'

To recap: Negative perfective imperatives are generally infelicitous, except for "preventive" cases as in (20), (21a), and (21b), in which the speaker warns the addressee about an undesired event that is feasibly going to happen, unless the addressee does something to prevent it. This undesired event can be either non-agentive ((20)) or agentive ((21a) and (21b)).

This characterization of preventive cases is already a bit reminiscent of the causal models we have proposed for declarative negative perfective sentences: There is something that makes an event feasible even though the event either was not realized (in the case of past negative perfective declaratives), or may not be realized (in the case of negative perfective imperatives, if the addressee heeds the warning). But let's make the connection more evident, using the causal model in Figure 10.



not-at-issue	inferred	at-issue
X?	Y?	E?
1	1	0
1	0	1

X?: whether a condition occurs that makes the undesired event's occurrence feasible
E?: whether the undesired event occurs
Y?: whether the addressee performs an action which is sufficient to prevent the undesired event from occurring given $X? = 1$

Figure 10: Negative perfective imperatives: An "event collider" with influence Y? inferred

The first thing to notice is that there is no R? node. The reason we see no argument for representation of R? in the causal structure associated with these imperatives is that the addressee is being told to perform an action that keeps an event from occurring (an event collider)—rather than being told to prevent the feasible result from occurring while not preventing the event that leads to it (a result collider). So there is no evidence for a result collider, which seems to indicate that there is no R? node. It may be that imperatives only introduce the E? node. But this seems unlikely, as it looks like the same verb stem as in declaratives. So it could be that both E? and R? *are* represented, but for some reason imperatives do not allow a result collider, so we can't "see" R? in

imperatives. For now, because we can't detect a result collider, we won't put an $R?$ node in these models for imperatives.

As in declaratives, we propose for imperatives as well that the perfective contributes the node $X?$ and gives it a not-at-issue truth value of 1. $X? = 1$ tends to influence the undesirable event node $E?$ denoted by the verb stem to have the value 1; and $Y?$ represents an effort of the addressee to overcome the influence of $X?$, which may, if sufficient effort is applied, prevent the undesirable event from taking place. For example, the speaker of (20) believes there is a reason to expect that the addressee will fall down, the speaker of (21a) thinks there is a reason to expect that the addressee is going to hit the other person, and the speaker of (21b) thinks that there is a reason to expect that the addressee is going to buy everything in the store; and in all of these, the undesirable event is expected to occur unless the addressee makes some special effort to prevent it from occurring. This is represented by the model in Figure 10. Now, consider that negation contributes a truth value of 0 for $E?$. This is reminiscent of the "failed-event" case of declaratives.

The question that arises now is how the model in (21) interacts with what it is to be an imperative sentence. We will make a few simple assumptions about imperatives to answer this question. We will not attempt to defend these assumptions for all imperatives here, but we hope that they provide a starting place from which to think about how causal models might interact with imperative meaning in general.

Let's assume to begin with that requests and commands, and indeed preventive warnings, all involve placing something on the addressee's to-do list (*e.g.*, in the sense of Portner, 2007). Let's also assume that what is placed on the addressee's to-do list is the causal model contributed by the imperative sentence as in (21). By placing the causal model on the addressee's to-do list, the speaker conveys that they want the addressee to make reality match any at-issue truth values (in this case, just $E?$'s truth value 1), and that any not-at-issue truth values (here, $X?$'s truth value 1), as well as the model itself, accord with reality.

Another assumption we will make is that in order to felicitously request or command or warn someone to bring about an eventuality or prevent an eventuality from occurring, the speaker must believe that the addressee can control whether that eventuality happens. In our framework, where sentences contribute causal models and uttering an imperative sentence adds its causal model to the addressee's to-do list, let's make it a felicity condition on causal models appearing on to-do lists that they contain an exogenous node whose truth value is controlled by the addressee, and which influences whether the eventuality occurs.

(22) Felicity conditions on uttering imperatives about $E?$
The causal model added to the addressee's to-do list contains a node $A?$ (the *addressee control node*) such that:
a. $A?$ is exogenous in the model

b. The truth value of $A?$ is set by the addressee[17]
c. There is an arrow from $A?$ to $E?$

The condition in (22a) tells us that the truth value of$A?$ is decided by something outside the model. The condition in (22b) tells us who exactly decides it; namely, the addressee. The condition in (22c) ensures that the truth value of $A?$ influences the truth value of $E?$.

Returning to the model in (22), we see that $Y?$ can be this "addressee control node" $A?$: it is exogenous, it seems to be about the addressee exerting either physical or mental effort, which we can suppose is under the control of the addressee; and the speaker seems to assume that such effort could prevent the undesired event from occurring.

And in fact, not only can $Y?$ be the addressee control node in such a model, but it is apparently needed in negative perfective imperatives, in that the preventive contexts (represented by models with $Y?$) are the only ones that support felicitous utterances of negative perfective imperatives.

In the structure in Figure 10, if $X?$ is a circumstance (so, not within the control of the addressee), there is no addressee control node. If on the other hand $X?$ can represent the addressee's intention, similarly to what we saw above in declarative sentences, could $X?$ then be the addressee control node?

It seems that it cannot. In general, the contexts that support negative perfective *imperatives* seem to be a subset of the contexts that support negative perfective *declaratives*, exactly along these lines. All felicitous negative perfective imperatives, as in the (a) cases below, have corresponding felicitous negative perfective declarative, as in the (b) cases below. (These negative perfective declaratives could be uttered later in the discourse, after the event has been successfully avoided.)

(23) a. Ostorožno! Ne upadi!
careful Neg fall.Imp.Pfve
'Be careful! Don't fall down!' = (20)
b. One ne upal.
he Neg fellPfve
'He didn't fall down.'

(24) a. Tol'ko, požalujsta, ne udar' ego!
only please Neg hit.Imp.Pfve him
'Just please don't hit him!' = (21a)
b. Dima ego ne udaril
Dima him Neg hit.Pfve
'Dima didn't hit him.'

(25) a. Ostorožnee, ne skupi ves' magazin!
more.careful Neg buy.Imp.Pfve whole store
'Be careful, don't buy the whole store!' = (25)

[17]We can think of this setting of the truth value of $A?$ along the lines of Pearl (2000)'s *do*-operator, *e.g.* $do(\text{Maria})(A? = 1)$.

b. Lena ne skupila ves' magazin.
Lena Neg bought.Pfve whole store
'Lena didn't buy the whole store.'

However, the converse is not true; there are negative perfective imperatives that are infelicitous (as in the a cases below) but their corresponding negative perfective declaratives are felicitous (as in the b cases below):

(26) a. #Ne napiši pis'mo!
Neg write.Imp.Pfve letter
Intended: 'Don't write a letter!'
b. Miša ne napisal pis'mo.
Misha Neg wrote.Pfve letter
'Misha didn't write / hasn't written the letter.' = (9)

(27) a. #Ne otkroj okno!
Neg open.Imp.Pfve window
Intended: 'Don't open the window!'
b. Anna ne otkryla okno.
Anna Neg opened.Pfve window
'Anna didn't open the window.'

The difference between (20)-(25) on the one hand, and (26) and (27) on the other, is that the events described in (20)-(25) are in some sense held to be out of the control of the addressee—whether the addressee is an agent ((21a) and (25)) or not ((20))—while the events described in (26) and (27) are held to be under the addressee's control. Being in control of an event means, perhaps, being in control of whether it happens or not. In our framework, this would line up with the agent's intention for such an event to happen being the cause of the event occurring—if the agent wants it to happen, it happens; and if the agent doesn't want it to happen, it doesn't happen.

We propose, therefore, that the imperative sentences in (26), (26a), and (27a) are exactly cases where the *whether*-proposition of X? is *whether the addressee-agent has the intention for the event to occur*. The contexts that could just about improve them require that the normally intentional, controlled action be seen as somehow out of control. For (26), it is very difficult to imagine writing a letter without having control over it. We can get (26a) to improve a little bit if we imagine Misha is in love with someone and this love is what is driving them to write the letter against their will (and so the *whether*-proposition of X? would instead be *whether Misha is in love*). It's somewhat easier to imagine how opening the window could be out of control, if, perhaps, there is a button to open the window that one might accidentally press; in that case the *whether*-proposition of X? would be *whether Anna presses the button*. In either case, it would improve the sentence to add a word like *smotri* 'look out' at its beginning, exactly to signal that something is going to happen that the addressee is not completely aware of.

Our question was: Why are the negative perfective verb phrases which are

felicitous as imperatives, apparently a subset of the negative perfective verb phrases which are felicitous as declaratives? The answer seems to be that apparently imperatives prohibit $X?$ from being the intention of the addressee, while declaratives have no such constraint on $X?$. This answer accounts for the fact that negative perfective imperatives require preventive models, and don't allow non-preventive models (*i.e.,* models without $Y?$).

But we still don't have an answer as to why an intentional $X?$ can't count as the addressee control node. After all, $X?$ is exogenous, and has an arrow to $E?$. The answer here may be that it violates the requirement in (22a): namely, it is not actually under the control of the intender whether the intention is held or not. In other words, "the heart wants what it wants," and this cannot be overcome by effort. We think this is a possible reason, but there are other possible reasons too: perhaps the addressee control node is not allowed to be given by the semantics of the sentence itself, as, on our story, $X?$ is but $Y?$ is not, or perhaps nodes with not-at-issue truth values such as $X?$ are not allowed to be addressee control nodes. Deciding this question would take us too far afield, but we note that the framework has allowed us to get more precise about negative perfective imperatives than would have otherwise been possible, and shows potential for further investigation.

9 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.[18] In fact, we are very much in the right part of phrase structure to use causal relations, given that verbal denotations are often analyzed with causal relations (as we do here, with $E? \rightarrow R?$).

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.

There are two reasons we say that quantification over possible worlds is not necessary. First, recall our earlier discussion in Section 3.2 about how a theory of causation which entails its result almost forces one to introduce possible worlds, so that the existence of a causal relation can be maintained even when

[18] Responding to an anonymous reviewer's comment, 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 perfective.

its outcome is defeated. If we simply adopt a theory of causation that allows for causal relations even when the outcome is defeated, we don't need to use quantification over possible worlds.

Second, the familiar possible worlds framework made most powerful and flexible by Angelika Kratzer has an origin story with David Lewis. Although quantification over possible worlds may seem like the null hypothesis because of its familiarity, we should remember that one of the motivations David Lewis had for developing possible world semantics was that he wanted to explain—or explain away—causal relations (Lewis, 1973). He wanted this because of his commitment to explaining everything in terms of a "vast mosaic of local matters of particular fact" (Lewis, 1986). In Lewis's view, picking out the set of closest possible worlds crucially 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. Ultimately his 1973 theory of causation is based on counterfactual statements that use these primitives.

But we don't need to explain causation in terms of counterfactual statements. Instead, we can account for counterfactuality by using two methods: First, we build into our causal relations the idea that, all else being equal (or normal, or expected), if A hadn't occurred, B would not have occurred. This is expressed in the valuation table of the causal model; no complex counterfactual statement is necessary. Our second method to deal with counterfactuality comes into play in cases where all else is not equal, as in the negative perfective cases we have been looking at. These are cases where the presuppositions of the model predict one truth value, but negation gives us a different truth value. Such inexplicable truth values mean that the set of all actual truth values are not a line in the model's valuation table, and their presence means that the model is not an appropriate model of reality. Thus, another model which explains that value must be constructed. Together, these mechanisms do the same work as the identification of the "best", "inertial" or most "stereotypical" possible worlds in possible worlds frameworks. So, it is not necessary to use a quantification over possible worlds framework.

An alternative option, for those who believe that there must be quantification over possible worlds in the denotation, is 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. We think that an appeal to causal relations must be made anyway, in order to make such a move work.

The most natural direction to take, in order to try to create such an account, would be in the spirit of modal analyses of the progressive aspect (*e.g.* Dowty (1979), Portner (1998)) and the use of eventualities to anchor modal bases (Hacquard, 2006, 2010). According to this kind of approach, the VP-event should reach completion in the "best" circumstantial, "stereotypical", or "inertia" worlds in which every element in the eventuality develops normally, in accordance with its inherent properties and the laws of physics and human

behavior. In the case of negative perfectives, the anchor eventuality would probably have to be Anna's thought or the circumstance that made her killing action feasible. We think that such an approach could be made to get the right truth conditions.

Another alternative here is to use Kratzer's premise semantics (Kratzer, 1977, 1981b,a, 2012, a.o.) along with causal modeling. This is the route taken most prominently by Kaufmann (2013) for counterfactuals, as well as Alonso-Ovalle and Hsieh (2017, 2021) for abilitative cases in Tagalog where the assertion of the sentence is contrary to what one would expect to occur given background expectations.[19] But given the complementarity of causation and counterfactual possibilities, we're skeptical that both powerful mechanisms are needed. Moreover, we note that causal relations make explicit what stereotypical conversational backgrounds leave implicit—namely, what influences what, allowing us to add influences compositionally, as we do here.

In short, we think that we think that the use of quantification over possible worlds here is not necessary. We also think it is not as explanatory as the use of explicit structures of causal relations. Effectively, our use of explicit causal structures piggybacks on a consensus that there is already a causal relation in (telic) verb phrases, between an event and a result. The perfective is compositionally near to the verb stem, so it makes sense that it too would have explicitly causal semantics. This is not merely an Occam's Razor argument, since what is at stake is the relationship between heads and their meanings—it is also a question of complexity of the syntax-semantics interface, which we would like to be as simple as possible.

10 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 presupposes an influence that would tend to make the described event feasible (and, through it, would tend to make the result feasible as well). A negated perfective sentence thus presupposes that an event and result were feasible, even while it asserts that either the event or the result did not occur. We propose to model this account using causal models, where the perfective adds an additional, not-at-issue influence which affects the event described by the verb phrase. Negative perfective sentences highlight the existence of this additional influence because negation does not affect the presupposed existence of the added influence, while still negating the at-issue content from the verb phrase. This configuration leads to inexplicable

[19] They treat the unexpectedness of a described eventuality by using a stereotypical conversational background to collect propositions that are expected in the world of evaluation w. Then they look at a set containing these propositions, along with the facts in w that the occurrence of the eventuality causally depends on and the proposition describing the eventuality. A requirement that the occurrence of the eventuality be unexpected is modeled by a requirement that that set be inconsistent.

truth values in the model, which can only be explained by further altering the causal model to include at least one more influence. We extended this theory to account for negative perfective imperatives, which are only used in preventive contexts, where the addressee is warned to prevent an undesired outcome.

References

Alonso-Ovalle, L. and H. Hsieh (2017). Causes and expectations: On the interpretation of the tagalog ability/involuntary action form. In *Semantics and Linguistic Theory*, pp. 75–94.

Alonso-Ovalle, L. and H. Hsieh (2021). Causes and expectations: On the interpretation of the tagalog ability/involuntary action form. *Journal of Semantics 38*(3), 441–472.

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.

Baglini, R. and I. Francez (2015). The implications of managing. *Journal of Semantics 33*(3), 541–560.

Bar-Asher Siegal, E. A. and N. Boneh (2020). Causation: From metaphysics to semantics and back. In *Perspectives on causation: Selected papers from the jerusalem 2017 workshop*, pp. 3–51. Springer.

Bogusławski, A. (1985). The problem of the negated imperative in perfective verbs revisited. *Russian Linguistics 9*(2/3), 225—239.

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.

Dickey, S. M. (2020). Time out of tense: Russian aspect in the imperative. *Journal of Linguistics 56*(3), 541–576.

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.

Goncharov, J. (2018). Intentionality effect in imperatives. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 27*, 105–127.

Grønn, A. (2004). *The semantics and pragmatics of the Russian factual imperfective*. Ph. D. thesis, University of Oslo.

Hacquard, V. (2006). *Aspects of Modality*. Ph. D. thesis, MIT.

Hacquard, V. (2010). On the event relativity of modal auxiliaries. *Natural language semantics 18*, 79–114.

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 34*, 57–84.

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.

Kratzer, A. (1977). What 'must'and 'can'must and can mean. *Linguistics and philosophy 1*(3), 337–355.

Kratzer, A. (1981a). The notional category of modality. In H.-J. Eikmeyer and H. Rieser (Eds.), *Words, Worlds, and Contexts*, pp. 38–74. Berlin: de Gruyter.

Kratzer, A. (1981b). Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic 10*, 201–216.

Kratzer, A. (2012). *Modals and conditionals: New and revised perspectives*, Volume 36. Oxford University Press.

Krifka, M. (1992). *Thematic Relations as Links between Nominal Reference and Temporal Constitution*, pp. 29. Center for the Study of Language (CSLI).

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

Lewis, D. (1973). Causation. *The Journal of Philosophy 70*(17), 556–567.

Lewis, D. (1986). *On the plurality of worlds*. Basil Blackwell.

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, Stanford University.

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

Nadathur, P. (2023). *Actuality inferences: causality, aspect, and modality*, Volume 15. Oxford University Press.

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.

Portner, P. (2007). Imperatives and modals. *Natural language semantics 15*, 351–383.

Schulz, K. (2011). If you'd wiggled a, then b would've changed. *Synthese 179*(2), 239–251.

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.

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*. John Benjamins.

Thelin, N. B. (1991). On the concept of time: Prolegomena to a theory of aspect and tense in narrative discourse. In L. R. Waugh and S. Rudy (Eds.), *New Vistas in Grammar: Invariance and Variation.–Amsterdam*, pp. 261–285. John Benjamins.

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. 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.