A star is drawn: Testing the culmination inferences of Russian perfective accomplishments

Natasha Kasher, Ben-Gurion University of the Negev, IL, kashern@bgu.ac.il
Aviya Hacohen, Ben-Gurion University of the Negev, IL, aviya@bgu.ac.il

It is a widely established view in the event-semantics literature that perfective (PFV) accomplishments denote culmination. It has also been increasingly recognized over the past decades that non-culminating interpretations of PFV accomplishments are more widely available than previously assumed. Hence, in many languages, the culmination inference carried by PFV accomplishments is defeasible.

The Slavic PFV, on the other hand, has been consistently argued throughout the literature to enforce strict culmination requirements on telic accomplishments within its scope, such that non-culminating interpretations are entirely disallowed for such forms. Furthermore, in Slavic languages such as Russian, culmination is taken to be semantically entailed by PFV accomplishments, rather than defeasibly implicated. This is then used to explain the supposed contradiction created in (1), between the entailed culmination of the first clause, and its cancellation in the second clause.

(1) Masha s’ela buterbrod (#no kusoeck ostavila).
Masha PFV-ate-SG-F sandwich-ACC but piece.small left.
‘Masha ate the/a sandwich (#but left a small piece).’

To our knowledge, there exists no prior attempt to systematically probe Russian speakers’ actual judgments regarding this assumed contradiction. The current study aims to fill this gap.

Results from a gradable acceptability judgment task reveal that Russian speakers do in fact exhibit some degree of tolerance when presented with PFV accomplishments as descriptions of partially completed events. Moreover, these speakers provide high acceptability ratings for sentences in which the PFV accomplishment is followed by a cancellation phrase.
1 Introduction

The defining property of accomplishment verbs is that they “proceed toward a terminus which is logically necessary to their being what they are” (Vendler 1967: 101). Thus, a base accomplishment VP, comprised of a dynamic verb and a quantized incremental Theme argument (e.g., Krifka 1989), is integrally associated with a telos (Garey 1957), or a point of culmination (Parsons 1990). Combined with the culmination requirement carried by the perfective (pfV) operator, such telic predicates may only denote culminated events; namely, events in which the referent of the Theme argument has fully undergone the change specified by its governing verb (e.g., Filip 2017: 169). So, for example, to eat an apple entails to eat an apple completely (Tenny 1994: 22–23), and the sentence in (1) would be false in the event that the referent of an apple is only partially eaten.

(1) Adam ate an apple.

Over the past decades, however, it has been increasingly recognized that pfV telic accomplishments do in fact allow for varying degrees of culmination requirements crosslinguistically, and specifically, that non-culminated interpretations of pfV telic accomplishments are more widely available than previously assumed (for a detailed list, see Martin 2019; Persohn 2022). Hence, in many languages, the culmination inference carried by pfV telic accomplishments is defeasible, as demonstrated for Mandarin (2a), Hindi (2b), and English (2c), where the second clause felicitously cancels the culmination inference of the first clause.

(2) a. Ta hua-le yi-fu hua, keshi mei hua-wan.
    he draw-LE one-CL picture but not draw-finish.
    ‘He drew a picture, but he didn’t finish drawing it.’ (Mandarin: Soh & Kuo 2005)

b. Maya-ne biskuT-ko khaa-yaa par us-e puuraa nahiin khaa-yaa.
    Maya-ERG cookie-ACC eat-PFV but it-ACC full not eat-PERF
    ‘Maya ate a cookie, but not completely.’ (Hindi: Arunachalam & Kothari 2011)

c. She ate the sandwich but as usual she left a few bites. (Hay et al. 1999)

The willingness of English speakers to accept pfV telic accomplishments as descriptions of partially completed events is also demonstrated experimentally (e.g., Arunachalam & Kothari 2011).1

1 See also Foppolo et al. (2021) for an interesting recent study investigating the time course of telicity interpretations in Italian. The study uses eye-tracking to test when and how the culmination inference of telic predicates arises during processing. Most relevant for the current study, though, are the results from an offline acceptability judgment task that was used as a pre-test for the main experiments. The results show that Italian speakers \( N = 37 \) accept perfective telic sentences as descriptions of incomplete events at around 30%.
The Slavic PFV, on the other hand, has been consistently argued throughout the theoretical literature to enforce strict culmination requirements on telic accomplishments within its scope, such that non-culminating interpretations are entirely disallowed for such forms (e.g., Filip 2000; 2008; 2017; Grønn 2003; Kazanina & Philips 2007; Martin 2019; Smith 1991/1997).

Furthermore, in Slavic languages such as Russian, culmination is taken to be semantically entailed by PFV telic accomplishments, rather than defeasibly implicated (e.g., Comrie 1976; Filip 2000; 2008; 2017). This is then used to explain the supposed contradiction created in (3), between the putatively entailed culmination of the first clause, and its cancellation in the second clause.

(3) Masha s'ela buterbrod (#no kusοčeck ostavila).
Masha. PFV-ate-SG-F sandwich-ACC but piece.small left.
‘Masha ate the/a sandwich (#but left a small piece).’

Experimental work on adult native speakers’ interpretations of PFV telic accomplishments is rather limited. To our knowledge, the only available source for relevant data is a handful of acquisition and L2 research (acquisition of Russian: Kazanina & Phillips 2007; acquisition of Polish: van Hout 2008; Russian L2: Slabakova 2005; Nossalik 2009; see also van Hout et al. 2010 and van Hout 2018 for a survey of crosslinguistic acquisition data). The primary finding of these studies is that Russian and Polish native adult speakers systematically link the perfective form of the verb with completed events and the imperfective with ongoing processes. Crucially, though, none of these studies directly test the availability of non-culminating interpretations of PFV telic accomplishments. The current study is the first to directly test this in adult native speakers of Russian. Moreover, there exists no prior attempt to systematically probe speakers’ judgments regarding the assumed contradiction in (3) for Russian. The current study aims to fill this gap.

Contrary to the widely assumed view presented above, that Slavic languages such as Russian are uniquely strict in terms of the culmination inference encoded by their PFV telic accomplishments, we want to show that under certain conditions, PFV telic accomplishments may be licensed as descriptions of partially completed events, even in Russian. And further, that Russian speakers may in fact judge sentences such as (3) – where a PFV telic accomplishment is followed by a cancellation phrase – as acceptable.

To this end, we designed a gradable acceptability judgment task, an experimental paradigm which, to the best of our knowledge, has not yet been used to evaluate the availability of Russian non-culminating interpretations of PFV accomplishments. Moreover, and more broadly, gradable acceptability tasks have rarely been used to probe speakers’ interpretations of PFV telic accomplishments in other languages. Indeed, while there is quite a robust crosslinguistic body of experimental research on the topic (e.g., Arunchalan & Kothari 2011; Hacohen & Wolf 2017; van Hout 1998; Liu 2018; Ogiela & Schmitt & Casby 2014; Schulz & Penner 2002; Wright
2014), it almost exclusively comprises studies employing binary judgment tasks. An interesting exception is Aoki & Nakatani (2013), who used a gradable acceptability task to test the claim that culmination in Japanese accomplishments is an implicature (and therefore cancellable) (cf. Tsujimura’s 2003). More recently, Chen (2018) used a Likert-type scale to test interpretations of individual verbs in Mandarin Chinese. There is also some existing work from L2/bilingualism research that uses acceptability scales (see, e.g., Slabakova 2000; 2001; Yin & O’Brien 2019), but it is concerned primarily with issues such as crosslinguistic influence rather than telicity per se. Hence, although some experimental work on telicity that elicits non-binary data does exist, it is either restricted to language-specific lexical features of individual verbs or provides little information as to the intricate complexities of the linguistic phenomenon as such.

The rationale for using a gradable paradigm, rather than the more commonly used binary judgment task, is that the former better reflects the nature of acceptability judgments. Unlike grammaticality judgments, which are typically categorical (a sentence is either grammatical or not), acceptability judgments are considered inherently gradable (Chomsky 1965; Schütze 2016: 62, Sprouse 2007 and references therein). Moreover, since “[i]n natural conversation, there are many moves available to an interlocutor who is asked to judge the validity of a statement” (Sikos & Kim & Grodner 2019: 2), reducing speakers’ available choices to only two possible response options artificially constrains behavior (ibid. and cf. Sorace & Keller 2005). Hence, the use of a gradable judgment paradigm is paramount to exposing speakers’ true judgments. As our findings demonstrate, this methodological choice is particularly crucial when investigating the availability of non-culmination crosslinguistically, and in Russian, in particular.

2 Methods
2.1 Participants
A group of 33 native Russian speakers were recruited online through Russian social media (12 = F, 16 = M, 5 = no gender reported). Participation was on a voluntary basis and no compensation was offered. Participants came from Russia, Ukraine and Belarus, and all declared that Russian was their main language from birth. Age of participants ranged from 18–64, with a median age of 37. The results of three additional participants were excluded from the analysis after a post hoc examination revealed that their responses consisted almost entirely of the highest ratings (4-scores).

2.2 Design and Material
In our design, we manipulated the following variables: aspectual frame (Perfective, Perfective followed by Cancellation phrase, Imperfective); verb type (Creation, Consumption, Destruction);

---

2 Interestingly, much of the available experimental data, particularly those obtained in the early years, come from acquisition studies. Adult data – though for the most part not the focus of the research – are nonetheless almost always collected in these studies.
event type (Partial Completion, Full Completion, No event). The verbal stimuli consisted of eight accomplishments composed of an incremental transitive verb followed by a singular count noun object DP. The choice of VPs was mainly motivated by such factors as frequency, familiarity, and ease of depiction. In light of previous studies discussing potential effects of verb-type on the availability of (non-)culminating interpretations (see e.g., Martin 2019; Ogiela et al. 2014; van Hout 1998; Wright 2014), we also wanted to control for verb type. A list of the base accomplishments used is provided in Table 1.

<table>
<thead>
<tr>
<th>Creation</th>
<th>Consumption</th>
<th>Destruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>narisovat’ zvezdu</td>
<td>s’est’ šokoladku</td>
<td>steret’ nadpis’</td>
</tr>
<tr>
<td>draw star-ACC</td>
<td>eat chocolate.bar-ACC</td>
<td>erase writing-ACC</td>
</tr>
<tr>
<td>složit’ pазl</td>
<td>vypit’ stakan sока</td>
<td>rastvorit’ kубik saxara</td>
</tr>
<tr>
<td>assemble puzzle-ACC</td>
<td>drink glass-ACC juice-GEN</td>
<td>dissolve cube-ACC sugar-GEN</td>
</tr>
<tr>
<td>postroit’ piramidku</td>
<td></td>
<td></td>
</tr>
<tr>
<td>build small.pyramid-ACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sdelat’ taburetku</td>
<td></td>
<td></td>
</tr>
<tr>
<td>make bench-ACC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Base accomplishments tested.

Each of the eight accomplishments appeared in three aspectual frames, as shown in Table 2 for ‘draw star’.

<table>
<thead>
<tr>
<th>Condition 1: Perfective (PFV)</th>
<th>Malčik narisoval zvezdu.</th>
</tr>
</thead>
<tbody>
<tr>
<td>boy PFV-drew star-ACC</td>
<td>‘The boy drew a/the star.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition 2: Perfective + cancellation (PFV + CNCL)</th>
<th>Malčik narisoval zvezdu, no odnovo lučika ne xvataet.</th>
</tr>
</thead>
<tbody>
<tr>
<td>boy PFV-drew star-ACC but one ray not sufficient</td>
<td>‘The boy drew a/the star, but one point is missing.’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition 3: Imperfective (IMP)</th>
<th>Malčik risoval zvezdu.</th>
</tr>
</thead>
<tbody>
<tr>
<td>boy IMP-drew star-ACC</td>
<td>‘The boy was drawing a/the star.’</td>
</tr>
</tbody>
</table>

Table 2: Experimental manipulation of aspectual frames.

The visual stimuli of the critical items comprised eight short animated clips, depicting a human character performing the action denoted by the eight base accomplishments. The scenarios depicted in the clips unfolded naturally towards their intended endpoint, but they all terminated
short before reaching their goal. Note that the events were depicted as approximately 80% completed, similarly to many previous works on the topic (cf. e.g., Kaplan & Raju & Arunachalam 2021). This is illustrated in Figure 1 for ‘draw star’.

In addition to the eight critical items, in which all the events were depicted as partially complete, we also included eight control items. The verbal stimuli in the control items were the same as for the critical items, namely, the eight base accomplishments appearing in three aspectual frames, but the visual stimuli were manipulated, such that five of the clips depicted completed events, and the three remaining clips depicted scenarios in which the intended action did not even begin. The rationale for this unequal division is that in addition to providing the visual control to the partially completed events depicted in the critical items, the control items also served to counterbalance expected responses. Specifically, we wanted to increase the number of low acceptability responses so as to not create a bias towards high acceptability.

In total, there were 48 trials and we used a within-subject design, i.e., all participants were asked to judge all the trials.

2.3 Procedure

The experiment was conducted online using the Qualtrics\textsuperscript{\textregistered} platform. The visual stimuli were presented to all participants in one fixed pseudo-randomized order, while the verbal stimuli were fully randomized for each participant and for each item. Randomization was done automatically using the appropriate feature available through Qualtrics.

At the beginning of each experiment, the participant was told that they would be presented with video clips accompanied by sentences in Russian, and that their task was to evaluate how likely it would be for a native speaker to produce such a sentence. Response options were noted on a 4-point scale, each explicitly labeled: 1 (\textit{ni maleiśego šansa} ‘not a chance’), 2 (\textit{vrjad li} ‘not likely’), 3 (\textit{vozmožno, xotja čto-to ne tak} ‘possible, but slightly off’), 4 (\textit{vpolne verojatno} ‘highly likely’). Importantly, this was not a preference task; the participants were specifically instructed
to rate each of the sentences independently, and not relative to one another. This type of joint presentation, where the various sentence types are presented side by side, has been found to significantly enhance the sensitivity of acceptability measures (see Marty & Chemla & Sprouse 2020). Note also that no midpoint was used in the scale. This was done to control for a potential central tendency bias (cf. e.g., Douven 2018; Stevens 1971) and in order to create a forced-choice scale (cf. Chyung et al. 2018).

3. Result and analysis

The purpose of our study was to experimentally test whether non-culminated readings of perfective telic accomplishments are available for Russian speakers. Equally important was our aim to test whether Russian speakers accept sentences in which such PFV telic accomplishments are followed by a cancellation phrase. Recall that participants were asked to judge how likely it would be for a native speaker of Russian to use the sentences in the verbal stimuli as descriptions of the events presented in the videos. The distribution of responses across aspectual frames is plotted in Figure 2.

First, the graph reveals that in the PFV condition, a total of 65% of data points received low acceptability scores (25% and 40% 1- and 2-ratings, respectively). This confirms that Russian PFV telic accomplishments, just like PFV telic accomplishments crosslinguistically, generally denote culmination, and are therefore deemed by speakers to be largely ill-suited as descriptions of events that do not reach their intended endpoint. Still, we also find that as much as 34% of the time, participants judged such forms as “possible” (25%) or even “highly likely” (9%) descriptions of partially completed events. This is unexpected, given the widely accepted view in the literature that the Russian PFV strongly entails culmination, and should therefore be entirely unacceptable in such contexts.

Figure 2: Acceptability ratings across aspectual frames (partially completed events).

For considerations of space, only data from the partially completed scenarios are discussed here. See Appendix for results of the control items.
Even more surprising, though, are the findings from the PFV-CNCL condition. Here we see that participants freely accept sentences in which PFV telic accomplishments are followed by a cancellation phrase, with a total 81% of judgments being either “highly likely” (49%) or “possible” (32%). Low acceptability scores comprise merely 18% of the data points in this condition. Interestingly, this distribution is essentially identical to the IMP data, which is not predicted to carry any inference of culmination.

A statistical analysis of the three conditions in the partially completed scenarios revealed a main effect of aspectual frame (Friedman Chi-Square = 41.54, p < 0.001, df = 2). However, this effect was exclusively due to the distribution in the PFV. Indeed, a post hoc Wilcoxon Signed-Rank test confirmed that the distributions of the PFV + CNCL and the IMP were not significantly different (p = 0.470).

It is also interesting to examine the response distribution across verb types in the two critical conditions, the PFV and the PFV + CNCL. These data are presented in Figure 3.

![Figure 3: Acceptability ratings for PFV and PFV + CNCL across verb types (partially completed events).](image)

As is clearly visible from the graphs, response distributions were essentially identical, regardless of verb type. This was confirmed by a Friedmann Chi-Square, which did not yield a significant effect of verb type in either of the aspectual frames (p = 0.1975 for PFV, p = 0.0509 for PFV + CNCL).

4. General discussion

It is a long established and widely accepted view in the event semantics literature that Slavic languages constitute a language family which is particularly strict in terms of the culmination
inference encoded by PFV telic accomplishments (e.g., Filip 2017). This inference, it is argued, amounts to an entailment in Slavic languages, and is therefore non-defeasible (e.g., Comrie 1976; Filip 2000; 2008; 2017). This, in turn, is claimed to be the reason why in Slavic languages, such as Russian, PFV telic accomplishments are only allowed as descriptions of fully completed events, and why sentences in which a PFV telic accomplishment followed by a cancellation phrase purportedly result in a contradiction (e.g., Martin 2019).

Our study suggests that this commonly held typology may need to be revised. Our primary finding is that while PFV telic accomplishments in Russian do generally denote culmination, Russian speakers, nevertheless, also allow for a considerable degree of non-culminating interpretations of such constructions. Furthermore, and even more surprisingly, our data reveal high acceptability judgments for sentences in which PFV telic accomplishments are followed by a cancellation phrase. It is particularly striking that the response pattern in the PFV + CNCL condition was not significantly different from the judgments yielded in the imperfective condition. This finding is interesting first, because to our knowledge, it is the first and only time that acceptability judgments of such constructions in Russian have been directly elicited in a systematic, experimental setting. Perhaps more importantly, though, such a finding is striking because it indicates that contra to what has been consistently argued in the theoretical literature, Russian PFV telic accomplishments followed by a cancellation phrase do not necessarily result in a contradiction.

The interesting question to ask here is what exactly is being canceled. In our view, what is being canceled is not the culmination per se, but rather the maximal interpretation of culmination (cf. Martin 2019; Martin & Demirdache 2020), namely, the point “beyond which the event cannot proceed” (Declerck 1991: 121; Depraetere 1995: 3). Hence, the Russian perfective does encode event completion (unlike, e.g., the perfective in Mandarin or Hindi), but it also allows for non-maximal interpretations of PFV telic accomplishments. In other words, and following Martin’s (2019) terminology, while Russian has no non-culminating accomplishments, our study shows that it does have non-maximal accomplishments on par with what has been reported for languages such as English and Spanish (e.g., Martin & Arunachalam 2022). Further research is needed to fully develop this preliminary proposal.

In addition to the interesting theoretical implications of the current study, it also highlights the importance of methodology. In particular, it demonstrates the unique insight gained through the use of gradable acceptability judgments. It has been argued from a theoretical perspective, that the cancellation of the culmination inference in sentences like (3) results not in contradiction, but rather in “degraded acceptability” (Kennedy & Levin 2008: 160). But what does “degraded acceptability” mean from an experimental psycholinguistic perspective? If the measure used

---

4 The original claim is made for degree achievements in English.
to collect such acceptability judgments exclusively employs binary judgment tasks, it can only mean that more speakers find it unacceptable. But what if that is not all it means? What if some of the “yes/acceptable” responders would actually prefer to say something more like “yes, to some extent/sort of”? Conversely, what if the “no/unacceptable” actually masks a “not so much” answer? Clearly, this is very different from simply averaging across speakers and arriving at some average response per predicate/sentence-type. Using a gradable acceptability task, as we have done in the current study, allows us to uncover such subtle meaning intricacies that may be glossed over with a binary response paradigm.

In sum, what our findings suggest is that in contrast to the widely accepted view, Russian in particular, and arguably Slavic languages more broadly, are not unique with respect to the culmination inference encoded by their PFV telic accomplishments. This, in turn, provides some preliminary support for a more universalist view of event semantics, whereby the range of crosslinguistic differences is more restricted than previously assumed. However, we emphasize that such a suggestion is merely speculative at this point. While it may still be the case that the culmination inference of PFV accomplishments is indeed stricter in Russian than in other languages (cf. Minor et al. 2022), the contribution of the current study is in showing that Russian has some degree of tolerance in this respect. As such, the current study provides a necessary first step that we hope will inspire and initiate extensive research in the near future.
Data availability

The Appendix with additional results, as well as all the raw data and analysis scripts, can be found at this link: https://osf.io/g4637/?view_only=4a355520602d44968e72eff6b77e512f

Abbreviations

PFV = Perfective, IMP = Imperfective, ACC = Accusative, VP = Verb Phrase, CL = Clitic, ERG = Ergative, GEN = Genitive, CNCL = Cancellation, 1 = 1st person, 2 = 2nd person, 3rd = third person, SG = singular, PL = plural, F = feminine, M = masculine

Ethics and consent

The Human Subjects Research Committee of Ben-Gurion University has reviewed and approved this study.

Acknowledgements

We thank the anonymous reviewers for their valuable feedback.

Competing interests

The authors have no competing interests to declare.

References


Depraetere, Ilse. 1995. On the necessity of distinguishing between (un)boundedness and (a)telicity. Linguistics and philosophy 18(1). 1–19. DOI: https://doi.org/10.1007/BF00984959


