Temporal information and event bounding across languages: Evidence from visual world eyetracking*

Serge Minor
UiT The Arctic University of Norway

Natalia Mitrofanova
UiT The Arctic University of Norway

Gustavo Guajardo
UiT The Arctic University of Norway

Myrte Vos
UiT The Arctic University of Norway

Gillian Ramchand
UiT The Arctic University of Norway

Abstract We explore the typological question of what the interpretation of grammatical perfectivity is, and how it connects to the related aktionsartal notion of BOUNDEDNESS/TELICITY on the one hand, and the tense category PAST on the other. We report on a comparative experimental paradigm of past tense accomplishment sentences in Russian, Spanish and English respectively, in which we use an online visual world paradigm – comparing looks to an ongoing representation (OE) with a result state representation (CE) – to track the triggering of entailments of culmination during auditory processing. In all three languages, the results revealed at-ceiling preference for OE in the imperfective condition both in the offline task and the online gaze patterns. In the perfective condition, we found robust differences. In Russian, the choice of the result state (CE) picture in the offline task was at ceiling (95%); for Spanish it was high, but not quite at ceiling (83%); in English there was no statistical preference for the CE picture in the Simple Past condition (54%, not significantly different from chance, $p = 0.39$). Analysis of the participants' online gaze patterns yielded parallel results. Our results for English suggest that even on telic predicates, the simple past form does not obligatorily enforce a completed-event interpretation, contrary to previous assumptions in the literature (Smith 1995).

Keywords: aspect, perfectivity, Russian, English simple past, Spanish, eyetracking, visual world

1   Introduction

Aktionsart categories are rooted in the semantic classification of verbal predicates within and across languages; they classify the event represented in terms of dynamicity, duration and boundedness (telicity) (Dowty 1979, Vendler 1967). Tense marking is a way of specifying the way in which the events so described are anchored to temporal points in our situational reality (Comrie 1985). But what does aspect marking do, and why do so many languages find it convenient to mark it? In our study, we explore the typological question of what the interpretation of grammatical perfectivity is, and how it connects to the related aktionsartal notion of BOUND-EDNESS/TELICITY on the one hand, and the tense category PAST on the other. We report on a comparative experimental paradigm of past tense sentences in Russian, Spanish and English respectively, in which we use an online visual world paradigm to track the triggering of entailments of culmination during auditory processing. Our results are surprising, and force us to reexamine basic premises concerning the formal modeling of perfectivity and crosslinguistic variation.

1.1   Background: Perfectivity

First of all, it is important to acknowledge that we cannot take for granted that the linguistic category named 'perfectivity´ corresponds to the same thing across languages either truth conditionally or operationally. Nevertheless, we want to be able to: (i) compare different morphological packaging which results in similar truth conditions across languages; (ii) say something about how those different morphological 'moving parts' work compositionally and interact with the rest of the language system in question.

Given clear definitional diagnostics for the perfective vs. imperfective distinction, one can meaningfully ask the question whether a particular form or construction in a language allows, or forces, perfective or imperfective interpretations.

The following tests are the ones that are classically thought to be definitional of perfectivity:

- Entailment tests: Infelicity of denying the completion entailment (Giorgi & Pianesi 1997)

- Narrative Progression: Ability to advance the reference time in narrative progression (Kamp & Rohrer 1983)

Intuitively, aspectual distinctions correspond to different *ways* of viewing an eventuality and linking it to the anchoring point/interval (Comrie 1976, Smith 1995). Most modern aspectual semanticists implement this using a classical Kleinian

analysis building on basic intuitions of Reichenbach 1947 (see Klein 1991, Klein 1994). In this view, given an eventuality E, a situational perspective taken on it R, and an anchoring interval S, perfective and imperfective viewpoints differ in the following way:

- A **perfective** viewpoint corresponds to R holistically including E: $[_R [_E]]$;

- An **imperfective** viewpoint corresponds to R being contained in E: $[_E [_R]]$

This definition of viewpoint aspect is distinguished from *lexical aspect* or *aktionsart* which is traditionally cashed out in terms of different event internal topologies of dynamicity, duration, transition, and sometimes result or telos (Vendler 1967, Dowty 1979, Smith 1995, Pustejovsky 1995).

The attractiveness of the Kleinian analysis is that it seeks to derive both core diagnostics of the perfective from the same representation. It handles performance on the entailment tests because in the case of accomplishment, any event that contains a telos as a subpart will always have that telos instantiated in the perfective condition where the reference interval completely includes the event. In the case of narrative progression, the encapsulation of event time inside the reference time enforces discreteness when reference times are ordered.

Our three languages (Russian, Spanish and English) each possess verbal forms that have been analysed as exemplifying the perfective/imperfective distinction. In the next two subsections, we look at the judgement data from our three languages with respect to the core diagnostics of narrative progression and completion, to establish the plausibility of the parallelism claimed in the literature and to highlight any areas where the judgements are less clear.

1.2 Narrative Progression

In all our three languages what has been called the **perfective**, or preterite past, produces narrative discreteness, and the **imperfective/progressive** past produces overlap. The following examples illustrate this for Spanish (1-2) and Russian (3-4):

(1) Cuando leí   el libro, sonó   el teléfono.
     when   read.1S.PRET the book rang.3S.PRET the phone

     'When I read the book, the phone rang.'

(2) Cuando leía   el libro, sonó   el teléfono.
     when   read.1S.IMP the book rang.3S.PRET the phone

     'When I was reading the book, the phone rang.'

(3)   Kogda ja pro-čita-l   knigu, po-zvoni-l   telefon.
      when I PFV-read-PST.M.SG book   PFV-ring-PST.M.SG phone

      'When I read the book, the phone rang.'

(4)   Kogda ja čita-l   knigu, po-zvoni-l   telefon.
      when I read.IMP-PST.M.SG book   PFV-ring-PST.M.SG phone

      'When I was reading the book, the phone rang.'

In (1) and (3) (and the English translation) the reading event is represented by a perfective verb form, and the sentence is understood as describing a sequence of events where the ringing of the phone follows the reading the book. In (2) and (4) the first verb is imperfective (=progressive in the English translation), and the two events are understood as overlapping (i.e. the phone rang within the time period that the reading event was taking place).

While the data seem unproblematic at first blush, it should be noted that an account of the English simple past form as a perfective is only possible by systematically excluding its use on states. When a stative verb is in the simple past, it produces narrative overlap with any preceding eventuality, as the following example shows.

(5)   When I owned those shares, the stock prices went up and down many times.

To maintain the perfectivity of the English past with respect to narrative progression, we are forced into having two semantically different (homophonous) past tense inflections, or at least we have to stipulate that past tense statives are 'non-canonical' perfectives in some way (cf. Smith 1991).

1.3   Entailment Tests for Completion

As pointed out above, telic predicates marked with perfective aspect have traditionally been taken to trigger entailments of event completion (Smith 1991; Klein 1994; see e.g. Dowty 1979; Parsons 1990 on the English simple past). In recent formal semantic work, however, it has been proposed that this is subject to cross-linguistic variation. Specifically, it has been noted that in some languages perfective forms of accomplishment verbs are consistent with non-culmination (see e.g. Singh 1991 for Hindi, Koenig & Muansuwan 2000 for Thai, Bar-el, Davis & Matthewson 2005 for Squamish). Importantly, however, with respect to this dimension of variation, all three of our languages, Russian, Spanish and English, have been classified typologically as having 'strong' perfective forms, i.e. *by default* triggering culminating inferences with accomplishments (**?**, Martin & Gyarmathy 2019). In the following section we consider this claim more closely as applied to the three languages.

English

Multiple cases where expected completion entailment fail to arise in English have been discussed in the literature. First of all, when it comes to judgements of culmination, we must be careful to distinguish from the outset verbs that have a linguistically specified telos (achievements and accomplishments), and those that do not (activities and states) – because it is only for the former that the question is meaningful. To confuse the picture, some predicates in English have been claimed to be lexically ambiguous between telic and atelic readings. For instance, Martin 2019 argues, following Pancheva 2003, that the predicate *read the book* in the following example is genuinely ambiguous between an accomplishment and an activity reading.

(6)   John read the book for two hours/John read the book in two hours.

Furthermore, there are predicates that we may still want to call accomplishments that get a default inference of culmination, but where the culmination entailment can 'disappear' in a certain contexts. For example, it has been argued in a class of lexically modalized predicates the internal modal operator is responsible for non-culmination in examples such as (7) (Martin & Schäfer 2017).

(7)   Mary taught the song to John but he could not get it into his head.

With another class of more run of the mill accomplishments, it turns out that modification and context can give rise to non-culmination. Specifically, the subclass of *gradable* accomplishments when modified by durative adverbials can assert a maximal stage of an event that is nevertheless not 'complete' (Smollett 2005, Rappaport Hovav 2008, Piñon 2008, Kennedy 2012, Deo & Piñango 2011).

(8)   John ate a sandwich for several minutes until he found a bug in it. (example from Wright 2014)

(9)   John built a sand castle/ baked a cake for an hour. (example from Deo & Piñango 2011)

We could think of these cases as simple cases of coercion (Moens & Steedman 1988, de Swart 1998), or alternatively as triggered by the existence of covert degree operators that are either simple existential, or maximal as in analyses for English by Piñon 2008 and Kennedy 2012.

Yet another sub-case of non-culminating accomplishments in English has been argued by Martin 2019 to rely on indeterminacies/vagueness in the interpretation of (in)definite descriptions, so that the object is not completely affected at one level of granularity, but *is* fully affected at a looser level of granularity sufficient to license the use of the 'perfective' past tense.

(10)   I mowed the lawn (but not all of it). (from Rappaport Hovav 2008)

(11)   Mary ate the sandwich, but as usual she left a few bites. (from Hay, Kennedy & Levin 1999)

We believe that these escape clauses (as well as the persistent problem of past tense on states) raise doubts about the utility and force of the Kleinian proposal for English.

Furthermore, a number of previous small-scale experimental judgement studies have produced results which question the validity of analyzing the English Simple Past form of telic predicates as a 'strong perfective'. In a video judgement task (8 items, $n = 24$), Arunachalam & Kothari 2011 found that a full 47% of the time English speakers accepted sentences with simple past telic verbs as descriptions of incomplete events. This was comparable to the weak (simple) perfective in Hindi (53%), but contrasted with the strong (complex) perfective (29%). Similarly, in a sentence-video matching task (4 items, $n = 22$ participants), Jeschull 2007 found that adult English speakers selected videos of completed events only 49% of the time when presented with sentences containing non-particle simple past telic verbs (compared to 91% for particle verbs). Finally, in a picture judgement task (2 items, $n = 16$), van Hout 2008 found that 75% of the time adult English speakers accepted sentences with non-particle simple past telic verbs (*eat/drink*) as descriptions of incomplete events (as compared to 29% for particle verbs *eat up/drink up*).

Russian vs Spanish

Turning now to our other two languages, in Russian, perfective verbs such as in (12) strongly entail that the telos or final boundary was reached:

(12)   Maša na-pisa-la   pis'mo. #No ne   zakončila.
        Masha PFV-wrote-PST.F.SG letter   but NEG finish.PFV - PST.F.SG

        'Masha wrote a letter. #But she didn't finish it.'

Similarly, verbs in the Preterite form in Spanish indicate that the corresponding event reached a final boundary, i.e. ceased to occur (Fábregas 2015). In the case of accomplishment predicates, and in the absence of indications to contrary, this can be taken to mean that the event was completed, i.e. reached it's maximal/result stage. However, in contrast to the Russian Perfective, in the case of the Spanish Preterite this is not a necessary entailment.

For instance, Arche 2014 reports that most Spanish speakers find the following sentence contradictory. This indicates that the verb in the Preterite is interpreted as conveying completion of the colouring-event:

(13)   Marta coloreó   un castillo, pero no lo terminó.
         Marta colour.3S.PRET a   castle   but   not it finish.3S.PRET

'Marta coloured a castle, but she did not finish it.'

However, in the presence of an adverbial that imposes a temporal boundary on the duration of the event (e.g. *for ten minutes*), the completion interpretation is easily cancelled. For instance, the following sentence no longer produces a contradictory effect:

(14)   Marta coloreó   un castillo durante diez minutos, pero no lo
         Marta colour.3S.PRET a   castle   for   ten minutes but not it
         terminó.
         finish.3S.PRET

'Marta coloured a castle for ten minutes, but she did not finish it.'

In stark contrast to Spanish, Russian necessitates the use of Imperfective aspect in contexts such as (14) (Janda & Fábregas 2019).

This view is supported by corpus analysis. In an analysis of the Spanish-Russian portion of the Russian National Corpus, Janda & Fábregas 2019 found that 22% of Preterite forms in Spanish were translated by Imperfective verbs in Russian (by contrast, only 8% of Spanish Imperfect form were translated using some kind of Perfective verb in Russian). Interestingly, most of the discrepancy in the translation of Spanish Preterite forms is accounted by one class of cases, namely constructions where a timespan adverbial imposes a temporal boundary on the duration of the event, as in (14).

The cancellability of the completion entailment, as well as the apparent existence of speaker- and verb-conditioned variation in the judgements of examples like (13) (Arche 2014), suggests that verbs in the Preterite form in Spanish do not semantically entail that the event reached its result state/maximal stage. Instead, they entail the existence of a final boundary, which in the case of accomplishment verbs may by default *pragmatically implicate* event completion. If this is the case, we can expect that the emergence of the completion implicature may be influenced by the lexical properties of particular verbs, as well as contextual factors and world knowledge.

The upshot of all of this recent work in formal semantics is that it is not so straightforward to classify a tense/aspect form as perfective or imperfective based on entailments of culmination. As we have seen, there are many cases in which the tense form in question is analysed as 'perfective' but where the sentences are nevertheless judged as non-culminating. Comparing the experimental results for our three languages, we are impeded by the fact that tasks and materials vary across studies. It remains unclear whether the effects documented for English are really a sign that the English past tense is different from other perfectives, or just due to

the fact that English has been scrutinized more carefully with respect to aktionsart categories, with a few more 'unusual verbs'. It would thus be desirable to be able to compare completion entailments directly across languages, keeping task and materials as similar as possible.

1.4 Interim Summary

To summarize this section, we think that while the theoretical literature seems to form a consensus around the Kleinian representation of perfectivity, and uniformly ascribes it to the English simple past, the Spanish pretérito and the Russian perfective, we have seen some indications that the three forms are different with respect to how they behave on the criterial diagnostics.

The goals of our particular study were therefore the following:

- To probe the interpretation of perfective forms using an online measure (synchronous with online processing), less subject to metalinguistic awareness and the difficulty of interpreting offline entailment judgements.

- To test a broader selection of items and larger samples of participants than previous studies.

- To compare the interpretation of aspectual forms in multiple languages by applying a unified experimental paradigm.

The question is whether such a paradigm, keeping context and task effects the same, would show cross-linguistic differences in levels of perfectivity, here operationalized as a picture matching task probing the existence of a completion entailment for accomplishment predicates.

2 The Visual World Paradigm

Our study uses a combination of an offline picture-selection task and an online visual world (VW) task, where we record the participants' eye movements while they listen to the test sentences. In order to motivate the use of this paradigm to understand the competence of our speakers, we need to discuss briefly what the visual world paradigm is and why we think it is an interesting window on the linguistic systems of our speakers.

The visual world paradigm (VWP) has been a richly productive methodology in the area of linguistic processing, ever since the seminal study of Tanenhaus, Soivey-Knowlton, Eberhard & Sedivy 1995. The usefulness of the paradigm stems from the general fact that human eye movements track the focus of linguistic attention, if that

attention is given a visual manifestation. The specific linking hypothesis between the eye movement behaviour and processing activity in the brain is still controversial, with a number of subtly different proposals contrasting mostly with respect to the relative timing and interaction between top-down and bottom-up cues to attention, and the relationship between the visual and linguistic information inputs (see Huettig 2015 for an overview). Overall, however, the paradigm has proved to be highly sensitive to aspects of linguistic input, with eye movements tracking attention in a temporally fine grained way, beyond the conscious control of the language user. The visual world paradigm has contributed to our understanding of not only spoken word recognition (Allopenna, Magnuson & Tanenhaus 1998) but also of the time course of sentence processing and ambiguity resolution (Tanenhaus et al. 1995). In one type of study, the visual world paradigm has been used to show that humans engage in predictive processing, and that bottom-up information about the word being heard is used incrementally to predict what is going to be said (cf. Kamide 2008, Huettig 2015; Kuperberg & Jaeger 2016 for recent overviews). The evidence shows that this applies not just to lexical recognition, but that grammatical information, such as gender and case marking, is also processed incrementally during sentence processing. For instance, grammatical gender information on agreeing elements (such as determiners and adjectives) may speed up the processing of a subsequent noun (Guillelmon & Grosjean 2001; Lew-Williams & Fernald 2007; Lew-Williams & Fernald 2010; Hopp 2013; Hopp 2016), and information on case morphemes can be integrated rapidly to predict the continuation of a sentence (Kamide, Scheepers & Altmann 2003).

2.1 Previous Studies on the Processing of Grammatical Aspect

As we have discussed, grammatical aspect has been analyzed as a category that controls which perspective is being taken on the event described by the verb in combination with its arguments (Smith 1995). Events or situations in the world are arguably more complex than objects in that they involve a temporal dimension and express relationships among objects/participants. In language, events are typically evoked by means of verbs, but the identity criteria for (even concrete) events are more difficult to characterize than for concrete objects (Davidson 1969). Even visually the problem is more complex, the event of 'eating' contains many temporal phases, each of which can look quite different qualitatively from the other. Using a static visual representation to image an event faces choices not found with object imageability, in that the event *looks different* at different points during its temporal unfolding. Aspect marking on lexical verbs is a linguistic device that is thought to be relevant to precisely the question of which portion of the event is being singled out in the description.

To represent a verbal concept in its entirety, a dynamic representation would theoretically be more satisfactory. The present study (and others that have tested aspectual distinctions) deliberately uses two different types of *static* visual representation which focus on different temporal portions of the depicted event — a snapshot of the ongoing event (OE) and a snapshot of the immediate aftermath of the completed event (CE). If aspectual marking makes salient a particular dimension of an event, we would expect eye tracking in a picture matching paradigm to track (i) the nature of that focus, and (ii) the time course of when that choice is made. In the terms of our previous discussion, we can assume that *if a verb form entails a final state, gaze will be directed to an image of that final state in processing*.

Indeed, aspectual information coded by grammatical morphemes have been shown recently in a number of eye tracking VWP studies to facilitate event recognition (Zhou, Crain & Zhan 2014 for Mandarin and Foppolo, Bosch, Greco, Carminati & Panzeri 2021 for Italian). In each of these two latter cases, the perfective morpheme or functional element triggers preferential looks towards the completed event picture, thus corroborating the general semantic judgements that in the context of telic verbs, perfectivity generates a culmination entailment by default.

Our own previous work using a similar OE/CE-paradigm demonstrated an analogous effect in the processing of grammatical aspect by Russian speakers (Minor, Mitrofanova & Ramchand 2022). Even more interestingly, we found that the pattern of looks was finely tuned to the unfolding of aspectual information in the linguistic stimulus, sensitive to the different points of disambiguation *within* the verb (corresponding to prefixal vs. suffixal encoding of the aspectual contrast).

3 The Experiment

In this section, we report on three VW eye-tracking experiments involving a sentence-picture matching task: Russian (*n*=124), Argentinian Spanish (*n*=32) and English (*n*=66). The experiments had the same overall design, with some variation in the choice of items.

3.1 Procedure

While looking at a blank screen, the participants heard a short context sentence that located the narrative in the past:

> *It was a bright and sunny day...*
> *It was the first period at school...*
> *It was early in the morning...*

At the onset of the target sentence, the participants were given a visual display:

Two pictures side by side representing an **ongoing event** (OE) and the corresponding **completed event** (= the result state that obtained once the action was complete) (CE).

| (a) Ongoing Event (OE) | (b) Completed Event (CE) |
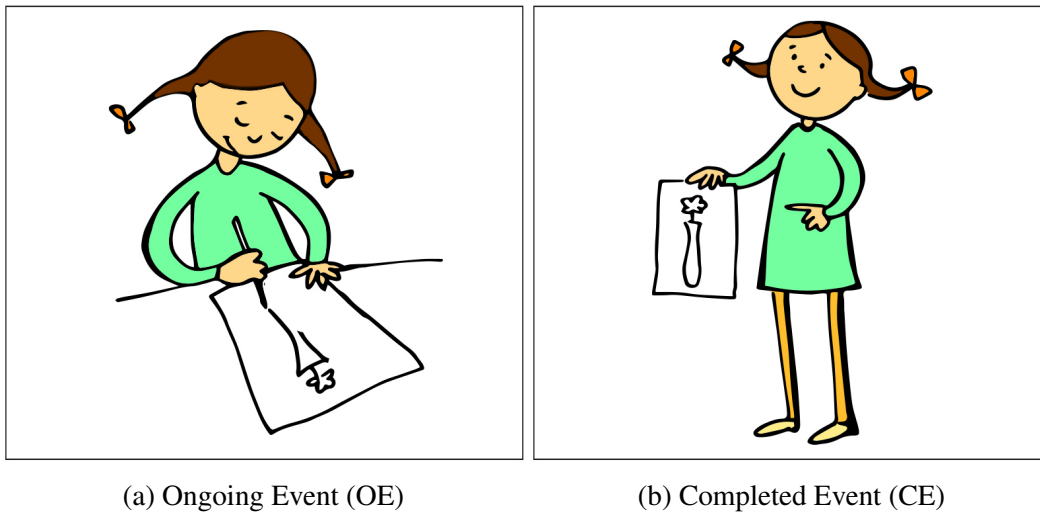
While looking at the pictures, the participants heard a test sentence and were asked to choose which picture best matched the sentence. We manipulated the aspect of the verb in the test sentence.

Russian:

(15) Devočka na-risova-la/risova-la   tonkuju vazu.
      girl   PFV-draw-PST.F.SG/draw.IMP-PST.F.SG thin   vase
      'The girl drew a thin vase.'

Spanish:

(16) Una nena dibuj-o/dibuj-aba   un florero delgado.
      a   girl draw-3S.PRET/draw-3S.IMP a   vase   thin
      'The girl drew a thin vase.'

English:

(17)   The girl drew/was drawing a thin vase.

In each experiment there were 24 test items involving accomplishment predicates. There were also 24 filler items, where the two pictures represented different event kinds. We recorded the participants' eye-movements (SMI RED500/EyeLink Duo) and offline responses.

Based on the theoretical literature, we made the following predictions:

•The Russian Imperfective, Spanish Imperfective, English Progressive would show a preference for the **Ongoing event** picture.

•The Russian Perfective, Spanish Perfective, English Simple Past would show a preference for the **Completed event** picture.

The leading research question was whether we would find crosslinguistic differences in the tendency of the 'perfective' to focus on the result of an accomplishment event (Russian > Spanish > English).

3.2   Results

Below we present the offline (picture selection) and online (gaze) results for the three languages. Details of the statistical analysis procedure can be found in the Appendix.

3.2.1   Russian

In their offline responses, the participants exhibited an almost universal preference for the Ongoing Event picture when they heard target sentences containing an Imperfective verb (98% of the Imperfective trials). Conversely, they exhibited an at-ceiling preference for the Completed Event picture when presented with target sentences containing a Perfective verb (95% of the Perfective trials). For the statistical analysis, the Ongoing Picture was coded as the accurate response in the Imperfective trials and, conversely, the Completed Picture was coded as the accurate response in the Perfective trials. The analysis confirmed that the accuracy was significantly above chance in both the Imperfective ($B = 5.85$, $SE = 0.78$, $Z = 7.47$, $p < 0.001$) and Perfective conditions ($B = 4.79$, $SE = 0.51$, $Z = 9.44$, $p < 0.001$). Furthermore, the analysis did not reveal a significant difference between the accuracy scores in the two aspectual conditions ($B = -1.06$, $SE = 0.94$, $Z = -1.14$, $p = 0.26$). In the next figure, we show the proportion of looks to the Target picture separately for the two aspectual conditions: Perfective and Imperfective. For the Imperfective

	Imperfective	Perfective
ONGOING EVENT	98%	5%
COMPLETED EVENT	2%	95%

Figure 2: Offline Results: Russian

trials looks to the Ongoing Event picture were coded as Target, while looks to the Completed event picture were coded as Competitor. The opposite coding was employed for the Perfective trials. All looks outside of the pictures (i.e. looks to White Space) were removed (7.77% of the data). We calculated the proportion of looks to the target picture in 60 consecutive 50 ms time bins starting from the verb onset (cf. Fig. 3). A cluster-based permutation analysis (Maris & Oostenveld 2007) revealed that in the Imperfective condition the probability of looks to the Target picture was significantly above chance in the time window from 350 to 3000 ms after the verb onset ($z = 463.09, p < 0.001$, represented by shading in Fig. 3). In the Perfective condition, the probability of looks to the Target picture was significantly above chance in the time window from 550 to 3000 ms after the verb onset ($z = 418.95, p < 0.001$).
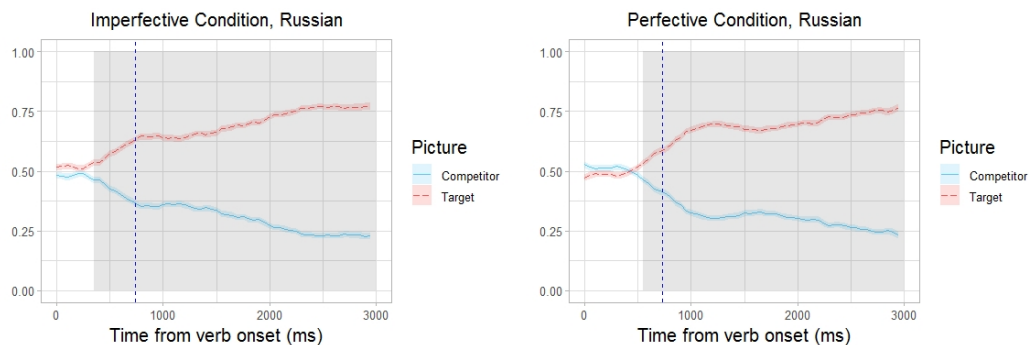

Figure 3: Russian experiment: Proportion of looks to the Target and Competitor pictures in the Imperfective and Perfective conditions. Shading represents the time windows where the probability of looks to the Target picture was significantly above chance. The dashed vertical blue lines mark the average verb offsets in the two conditions.

3.2.2 Spanish

The Spanish-speaking participants exhibited an at-ceiling preference for the Ongoing Event picture in the imperfective (Imperfect) condition (97% of the imperfective trials). In trials involving verb in the perfective (Preterite) form participants chose the Completed Event picture 83% of the times.

	Imperfective	**Perfective**
ONGOING EVENT	97%	17%
COMPLETED EVENT	3%	83%

Figure 4: Offline Results: Spanish

Statistical analysis revealed that the preference for the target picture was significant in both the imperfective (intercept $B = 27.43$, $SE = 7.77$, $Z = 3.53$, $p < 0.001$) and the perfective (intercept $B = 2.56$, $SE = 0.42$, $Z = 6.02$, $p < 0.001$) conditions. However, we found a significant effect of aspectual condition ($B = -3.46$, $SE = 0.91$, $Z = -3.82$, $p < 0.001$), indicating that the probability of an accurate response was significantly higher in the imperfective condition as compared to the perfective condition.

Fig. 5 shows the proportion looks to the Target and Competitor pictures by aspectual condition, with all looks outside of the pictures removed (3.45% of the data) A cluster-based permutation analysis revealed that in the imperfective condition the probability of looks to the Target (i.e. the Ongoing Event picture) was significantly above chance in the time window from 900 to 3000 ms following the verb onset ($z = 165.6$, $p < 0.001$). In the perfective condition the analysis identified two time windows where the probability of looks to the Target (i.e. the Completed Event picture) was significantly above chance: from 800 to 1850 ms after the verb onset ($z = 99.99$, $p < 0.001$) and from 1950 to 3000 ms after the verb onset ($z = 71.42$, $p < 0.001$).

3.2.3 English

Participants exhibited an at-ceiling preference for the Ongoing Event picture in the Past Progressive condition (95% of the Progressive trials). However, in the Simple Past condition the participants did not show a strong preference for either picture, choosing the Completed Picture in only 54% of the trials. Once again, we coded the selection of the Ongoing Event picture in the imperfective (Progressive)
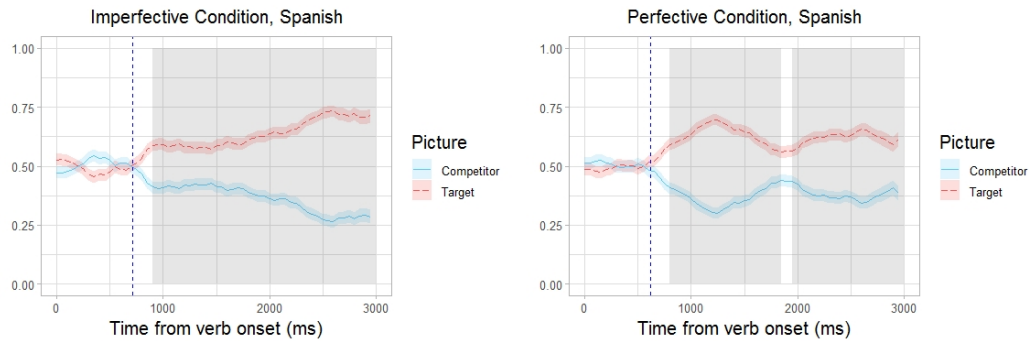
Figure 5: Spanish experiment: Proportion of looks to the Target and Competitor pictures in the imperfective (Imperfect) and perfective (Preterite) conditions. Shading represents the time window where the probability of looks to the Target picture was significantly above chance. The dashed vertical blue lines mark the average verb offsets in the two conditions.

condition and the Completed Event picture in the perfective (Simple Past) condition as 'accurate', and the opposite choice as 'inaccurate'. Statistical analysis revealed that in the Progressive condition the participants' performance was significantly above-chance (intercept $B = 6.24$, $SE = 0.77$, $Z = 8.09$, $p < 0.001$). On the other hand, in the Simple Past condition the participants' performance was not significantly different from chance (intercept $B = 0.26$, $SE = 0.3$, $Z = 0.86$, $p = 0.39$). Finally, the analysis revealed a significant effect of aspectual condition ($B = -5.8$, $SE = 0.79$, $Z = -7.37$, $p < 0.001$), indicating that the proportion of accurate responses was significantly higher in the Past Progressive condition as compared to the Simple Past condition.

	Progressive	Simple Past
ONGOING EVENT	95%	46%
COMPLETED EVENT	5%	54%

Figure 6: Offline Results: English

Fig. 7 shows the proportion of looks to the Target and Competitor pictures separately for the two aspectual conditions: Simple Past and Past Progressive. All looks outside of the pictures were removed (6.5% of the data). A cluster based permutation analysis revealed that in the Progressive condition the probability of

looks to the Target picture was significantly above chance in the time window from 450 to 3000 ms after the verb onset ($z = 283.62, p < 0.001$, represented by shading in Fig. 7). The analysis of the Simple Past condition did not reveal any significant clusters of above-chance looks to the Target picture.
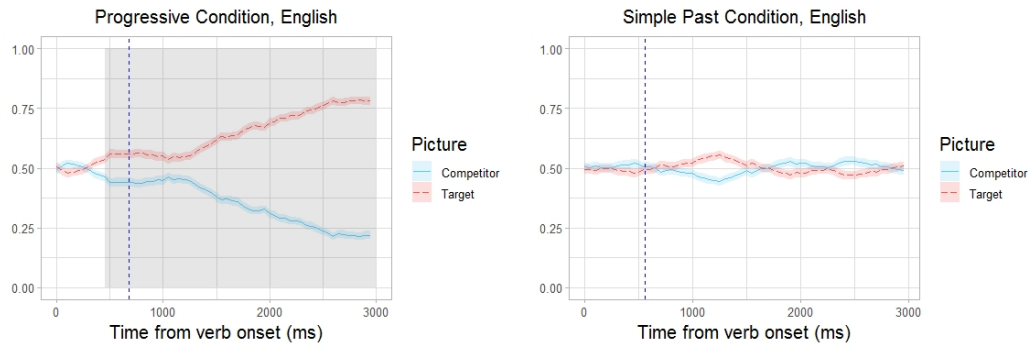

Figure 7: English experiment: Proportion of looks to the Target and Competitor pictures in the Past Progressive and Simple Past conditions. Shading represents the time window where the probability of looks to the Target picture was significantly above chance. The dashed vertical blue lines mark the average lexical verb offsets in the two conditions.

4   Discussion

These results confirm our prediction that the imperfective forms in all the three languages draw attention to the in-progress representation of the event.

   With respect to the perfective forms, our results suggest that perfective accomplishment verbs in Russian strongly highlight the result state of the event. In Spanish the perfective also highlights event completion, but to a lesser extent than in Russian. Our results for the English Simple Past condition are striking. They suggest that even on telic predicates, the Simple Past form does not encode a preferential cognitive salience for either the activity portion of an event *or* its result state.

   We think that this pattern of data indicates that we should separate specification of event completion from narrative sequencing in our classification of inflectional marking on verbs. In this context, we will define PERFECTIVE as a marking that imposes a *boundedness* requirement on the event description it combines with. We leave it open for the purposes of this paper whether one cashes this out in terms of non-cumulativity/non-divisiveness (Krifka 1989), or in terms of (modalized) maximality (Filip & Rothstein 2005, Filip 2008, Tatevosov 2014).

For Russian, we assume that what has been called perfective marking is truly perfective in the above sense. In addition the perfective specification in question is 'low', and is specified at the level of vP/VP. The choice of prefix in Russian encodes the specific kind of boundedness in play (resultative, delimitative, inchoative, etc.). All the items in our particular experiment were in fact resultative in addition to being bounded. In Spanish, the pretérito is also perfective in the above sense. However, in this language the boundedness feature is 'high' in the structure (outside the minimal vP), and is therefore licensed by bounding adverbials. In some cases boundedness that is pragmatically inferred can also license the use of the perfective in Spanish. This difference is reflected in a reduction of certainty on the part of the Spanish speakers when associating the perfective verbal form with the completed event picture.

In English, we must concur with de Swart 1998 that the simple past in English is non-aspectually specific. In particular it does *not* entail boundedness, unlike the inflectional marking in the other two languages. In the absence of a grammatical boundedness requirement, predicates which *can* be 'telic' in fact exhibit broad flexibility, with different degrees of preference for telic/atelic interpretations.

We thus expect that in English, judgements of completion will vary radically from predicate to predicate, even within the accomplishment category. The following diagram shows the inter-item variation that we found for English verbs in our experiment. Note that even the verb with the highest tendency to be interpreted as matching the completed picture does not reach 90% target behaviour, unlike nearly all of the Russian verbs.

(a) Inter Item Variation in English   (b) Inter Item Variation in Russian
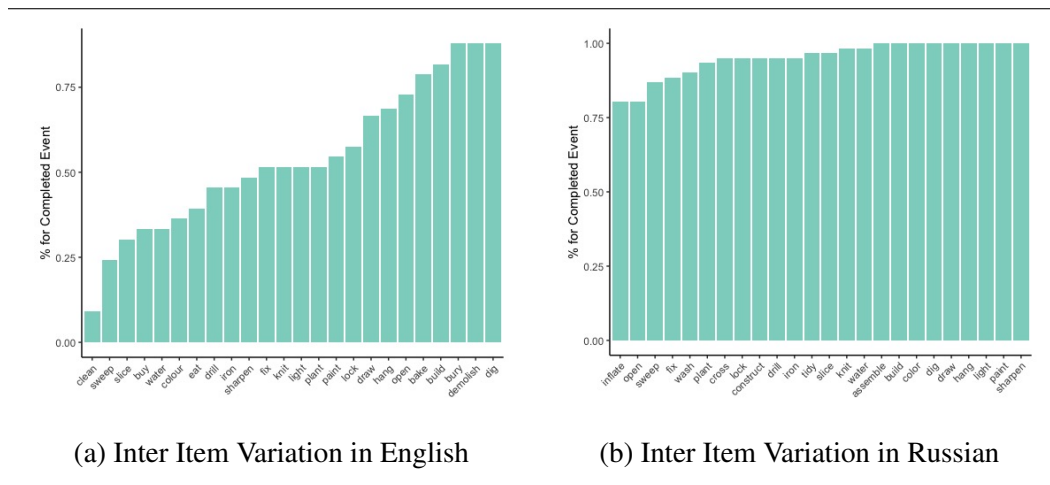
We can see that the Russian grammatical marking leads to speakers giving largely uniform results across the board, even though pragmatic effects and the actual pictures are mostly the same between the English and Russian experiments.

How then do we account for the role of the English Simple Past in advancing the reference time in a narrative? Based on our results, we would like to propose that the rules of narrative progression are in principle *dissociable* from grammatical aspect. At least in English, the rules of narrative progression are tied to the sortal distinction between **stative** and **non-stative** eventualities— states in English require overlap while non states favour sequencing. Connected to this we would therefore analyse the English progressive in *-ing* as a *stativizer* (Vlach 1981; Hallman 2009; Ramchand 2018). In the (non-stativizing) Simple Past form, activities, accomplishments and achievements favor a sequencing interpretation, even when they are not bounded. Bohnemeyer & Swift 2004 has independently argued for a distinction between languages like English which are fundamentally organized around a stative/dynamic distinction and others like Russian and Dutch which are grammatically organized around a bounded/unbounded distinction.

If we are right, and the English simple past is in fact underspecified for aspect, a question naturally arises as to why completion entailments for this form have been overestimated in the literature. We think that this is possible due to a focus on examples with the clearest judgements— since the English effects are highly contextual, these are fairly easy to find. Another problem has been the implicit reliance on the diagnostic of narrative sequencing as correlating with the completion profile (as it does in the Slavic languages). We have shown in this series of experiments that that assumption is probably false.

One of the important theoretical implications of our study is that we need to be able to dissociate completion entailments and narrative sequencing in our formal analysis of perfectivity. We need to move away from the Kleinien topological characterization, and towards accounts which state boundedness or quantizedness more directly. If we choose to reserve the label 'perfective' for boundedness in this sense, we can keep it separate from aktionsart on the one hand, and rules of narrative sequencing on the other. These three kinds of classification clearly interact with each other, but may do so in slightly different ways across languages. Languages also seem to vary with respect to which of these kinds of distinctions are lexically encoded and which are encoded via higher functional elements. In Russian, perfectivity seems to be lexically encoded on the verb, while in Spanish it is a functional element bundled with tense. This analysis predicts the scopal interaction with bounding adverbials that we have seen. Yet other languages, like English, do not obligatorily encode boundedness, but have it fall out from an interaction between aktionsart type, pragmatics, and discourse context.

References

Allopenna, P.D., J. S. Magnuson & M.K. Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements:evidence for continuous mapping models. *Journal of Memory and Language* 4(38). 419–439.

Arche, Maria. 2014. About the primitives of aspect across languages. *Natural Language and Linguistic Theory* 32. 711–733. doi:http://www.jstor.org/stable/43697737.

Arunachalam, Sudha & Anubha Kothari. 2011. An experimental study of Hindi and English perfective interpretation. *Journal of South Asian Linguistics,* 4. 27–42.

Bar-el, Leora, Henry Davis & Lisa Matthewson. 2005. On non-culminating accomplishments. In Leah Bateman & Cherlon Ussery (eds.), *North East Linguistics Society (NELS 35)*, 87–102.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. doi:10.18637/jss.v067.i01.

Bohnemeyer, Juergen & Mary Swift. 2004. Event realization and aspectual interpretation. *Linguistics and Philosophy* 27(3). 263–296. doi:https://doi.org/10.1023/B:LING.0000023371.15460.43.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems* Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.

Comrie, Bernard. 1985. *Tense* Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.

Davidson, Donald. 1969. *The Individuation of Events* 216–234. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-1466-2_11.

Deo, Ashwini & Maria Piñango. 2011. Quantification and context in measure adverbs. In *Semantics and Linguistic Theory (SALT) 21*, 295–312. doi:https://doi.org/10.3765/salt.v21i0.2614.

Dowty, David R. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.

Filip, Hana. 2008. Events and maximalization: The case of telicity and perfectivity. In Susan Rothstein (ed.), *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, 217–256. Amsterdam: John Benjamins. doi:10.1075/la.110.10fil.

Filip, Hana & Susan Rothstein. 2005. Telicity as a semantic parameter. In James Lavine, Steven Franks, Hana Filip & Mila Tasseva-Kurktchieva (eds.), *Formal Approaches to Slavic Linguistics (FASL) XIV - The Princeton University Meeting*, 139–156. Ann Arbor, MI: University of Michigan Slavic Publications.

Foppolo, Francesca, Jasmijn E. Bosch, Ciro Greco, Maria N. Carminati & Francesca Panzeri. 2021. Draw a star and make it perfect: Incremental processing of telicity. *Cognitive Science* 45. doi:10.1111/cogs.13052.

Fábregas, Antonio. 2015. 'imperfecto' and 'indefinido' in spanish: what, where and how. *Borealis – An International Journal of Hispanic Linguistics* 4. 1–70. doi:10.7557/1.4.2.3534.

Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense and Aspect: From Semantics to Morphosyntax*. New York: Oxford University Press.

Guillelmon, D. & F. Grosjean. 2001. The gender marking effect in spoken word recognition: the case of bilinguals. *Memory and Cognition* 29. 503–511. doi:https://doi.org/10.3758/BF03196401.

Hahn, N, J Snedeker & H Rabagliati. 2015. Rapid Linguistic Ambiguity Resolution in Young Children with Autism Spectrum Disorder: Eye Tracking Evidence for the Limits of Weak Central Coherence: Rapid lexical ambiguity resolution in ASD. *Autism Research* 8(6). 717–726. doi:10.1002/aur.1487.

Hallman, Peter. 2009. Instants and intervals in the event/state distinction. *Manuscript, UCLA* .

Hay, Jennifer, Christopher Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in "Degree Achievements". In Tanya Matthews & Devon Strolovitch (eds.), *Semantics and Linguistic Theory SALT IX*, 127–144. Ithaca, NY: CLC Publications.

Hopp, H. 2013. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29. 33–56. doi:https://doi.org/10.1177/0267658312461803.

Hopp, H. 2016. Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research* 32. 277–307. doi:https://doi.org/10.1177/0267658315624960.

van Hout, Angeliek. 2008. Acquiring telicity cross-linguistically: On the acquisition of telicity entailments associated with transitivity. In Melissa Bowerman & Penelope Brown (eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*, 255–278. New York/London: Erlbaum.

Huang, Y & J Snedeker. 2020. Evidence from the Visual World Paradigm Raises Questions about Unaccusativity and Growth Curve Analyses. *Cognition* 200. 104251. doi:10.1016/j.cognition.2020.104251.

Huettig, Falk. 2015. Four central questions about prediction in language processing. *Brain Research* 1626. 118–135.

Janda, L. & A. Fábregas. 2019. Seeing from without, seeing from within: Aspectual differences between spanish and russian. *Cognitive Linguistics* 30(4). 687–719. doi:10.1515/cog-2018-0054.

Jeschull, Liane. 2007. The pragmatics of telicity and what children make of it. In

Alyona Belikova, Luisa Meroni & Mari Umeda (eds.), *Generative Approaches to Language Acquisition, North America (GALANA) 2*, 180–187. Somerville, MA: Cascadilla Proceedings Project.

Kamide, Y. 2008. Anticipatory processes in sentence processing. *Language and Linguistics Compass* 2(4). 647–670. doi:https://doi.org/10.1111/j.1749-818X.2008.00072.x.

Kamide, Y., C. Scheepers & G. T. Altmann. 2003. Integration of syntactic and semantic information in predictive processing. *Journal of Psycholinguistic Research* 32(1). 37–55. doi:10.1023/A:1021933015362.

Kamp, Hans & Christian Rohrer. 1983. Temporal reference in French. Ms. University of Stuttgart.

Kennedy, Chris. 2012. The composition of incremental change. In Violeta Demonte & Louise McNally (eds.), *Telicity, Change, State: a cross-categeorial view of event structure*, 103–121. Oxford University Press.

Klein, Ewan. 1991. Comparatives. In A. von Stechow & D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössichen forschung*, Berlin: Walter de Gruyter.

Klein, Wolfgang. 1994. *Time in Language*. London and New York: Routledge.

Koenig, Jean-Pierre & Nuttanart Muansuwan. 2000. How to end without ever finishing: Thai semi-perfectivity downloaded from. *Journal of Semantics* 17. 147–184.

Krifka, Manfred. 1989. Nominal reference and temporal constitution and quantification in event semantics. In P. van Emde Boas R. Bartsch, J. van Bentham (ed.), *Semantics and Contextual Expression*, 75–115. Dordrecht: Forsi.

Kuperberg, G. & T. F. Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31. 32–59. doi:10.1080/23273798.2015.1102299.

Lew-Williams, C. & A. Fernald. 2007. Young children learning spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science* 18(3). 193–198. doi:10.1111/j.1467-9280.2007.01871.x.

Lew-Williams, C. & A. Fernald. 2010. Real-time processing of gender-marked articles by native and non-native spanish speakers. *Journal of Memory and Language* 63(4). 447–464.

Maris, E & R Oostenveld. 2007. Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods* 164. 177–190. doi:10.1016/j.jneumeth.2007.03.024.

Martin, F. & Z. Gyarmathy. 2019. A finer-grained typology of perfective operators. In C. Piñon (ed.), *Empirical Issues in Syntax and Semantics 12*, Paris:CSSP.

Martin, F. & F. Schäfer. 2017. Sublexical modality in defeasible causative verbs. In Maria Luisa Rivero Ana Arregui & Andrés Salanova (eds.), *Modality Across*

Syntactic Categories, 87–108. Oxford University Press.

Martin, Fabienne. 2019. Non culminating accomplishments. *Language and Linguistics Compass* doi:https://doi.org/10.1111/lnc3.12346.

Minor, Serge, Natalia Mitrofanova & Gillian Ramchand. 2022. Fine-grained time course of verb aspect processing. *PLOS ONE* 17. e0264132. doi:10.1371/journal.pone.0264132.

Moens, Marc & Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics* 14. 15–28.

Oakes, L., H. Baumgartner, F. Barrett, I. Messenger & S. Luck. 2013. Developmental changes in visual short-term memory in infancy: evidence from eye-tracking. *Frontiers in Psychology* 4. 697. doi:10.3389/fpsyg.2013.00697.

Pancheva, Roumyana. 2003. The aspectual makeup of perfect participles and the interpretations of the perfect. In Monika Rathert Artemis Alexiadou & Arnim von Stechow (eds.), *Perfect Explorations*, 277–306. Mouton de Gruyter.

Parsons, Terence. 1990. *Events in the Semantics of English: A Study in Subatomic Semantics*. Cambridge, MA.: MIT Press.

Piñon, Christopher. 2008. Aspectual composition with degrees. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and Adverbs: Syntax, Semantics and Discourse*, 183–219. Oxford University Press.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, Ma.: MIT Press.

Ramchand, Gillian. 2018. *Situations and Syntactic Structures: Rethinking Auxiliaries and Order in English* Linguistic Inquiry Monograph 77. MIT Press.

Rappaport Hovav, Malka. 2008. Lexicalized meaning and the internal temporal structure of events. In Susan Rothstein (ed.), *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, 3–42. John Benjamims.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: Macmillan.

Singh, Mona. 1991. The perfective paradox: Or how to eat your cake and have it too. In Christopher Johnson Laurel Sutton & Ruth Shields (eds.), *Berkeley Linguistics Society (BLS)*, 469–479.

Smith, Carlota S. 1991. *The Parameter of Aspect*, vol. 43 Studies in Linguistics and Philosophy. Dordrecht: Kluwer.

Smith, Carlota S. 1995. *The Parameter of Aspect* Studies in linguistics and philosophy; 43. Dordrecht: Kluwer revised edn.

Smollett, Rebecca. 2005. Quantized direct objects don't delimit after all. In Henk J. Verkuyl, Henriette de Swart & Angeliek van Hout (eds.), *Perspectives on Aspect*, 41–59. Dordrecht: Springer Netherlands. doi:10.1007/1-4020-3232-3_3.

de Swart, Henriette. 1998. Aspect shift and coercion. *Natural Language and Linguistic Theory* 16. 347–385.

Tanenhaus, M. K., M. J. Soivey-Knowlton, K. M. Eberhard & J. C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension.

Science 268(5217). 1632–1634.

Tatevosov, Sergei. 2014. Perfectivity in Russian: A modal analysis. In Jyoti Iyer & Leland Kusmer (eds.), *North East Linguistic Society (NELS) 44*, vol. 2, 196–210.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, N.Y.: Cornell University Press.

Vlach, Frank. 1981. The semantics of the progressive. In Philip Tedeschi & Annie Zaenen (eds.), *Syntax and semantics, vol. 14: tense and aspect*, 271–292. New York: Academic Press.

Wright, Tony Allen. 2014. *Strict vs. Flexible Accomplishment Predicates*: University of Texas at Austin PhD dissertation.

Yang, W, A Chan, F Chang & E Kidd. 2020. Four-year-old Mandarin-speaking children's online comprehension of relative clauses. *Cognition* 196.

Zhou, P., S. Crain & L. Zhan. 2014. Grammatical aspect and event recognition in children's online sentence comprehension. *Cognition* 133(1). 262–276.

Appendix: Data analysis

Offline judgements

To analyze the results of the offline picture selection task, we coded the selection of the Ongoing Event picture in the Imperfective condition and the Completed Event picture in the perfective condition as 'accurate' (1). The opposite choice in the two conditions was coded as 'inaccurate' (0).

To test whether the selection of the Target picture was significantly above chance in each of the aspectual conditions, we used the package *lme4* (Bates, Mächler, Bolker & Walker 2015) of the statistics software R version 3.6.0 to fit two mixed effects logistic regressions estimating the log-odds of an accurate response in the Past Progressive and Simple Past trials. The models included an intercept term and random intercepts for participants and items.

To compare the offline responses in the two conditions, we further fit a mixed effects logistic regression estimating the probability of an accurate response based on the aspectual condition (Perfective vs Imperfective) as the predictor variable. The model included a random slope for condition by participant and by item.

Gaze data

For the Imperfective trials looks to the Ongoing Event picture were coded as Target (1), while looks to the Completed event picture were coded as Competitor (0). The opposite coding was employed for the Perfective trials. All looks outside

of the pictures (i.e. looks to White Space) were removed prior to the analysis. We calculated the proportion of looks to the target picture in 60 consecutive 50 ms time bins starting from the verb onset. The data was made fully binary by replacing all proportions ≥ 0.5 with 1, and all proportions < 0.5 with 0. We then performed a cluster-based permutation analysis (Maris & Oostenveld 2007; see Hahn, Snedeker & Rabagliati 2015; Oakes, Baumgartner, Barrett, Messenger & Luck 2013; Yang, Chan, Chang & Kidd 2020; Minor et al. 2022 for applications of this method in visual world eye-tracking) to assess whether the probability of looks to the Target picture was significantly above chance (0.5). The analysis was performed separately in each aspectual condition (Perfective and Imperfective). For each time bin, we fit a mixed effects logistic model evaluating the log-odds of looks to the Target picture. The model included an intercept term, and items and participants as random intercepts. The intercept represented the difference between the log-odd of looks to the Target picture and 0, which corresponds to chance (0.5) probability. The cluster-based permutation analysis followed the same overall procedure as described above. We identified clusters of adjacent time bins where the Intercept p value was below 0.1 and calculated their sum statistics. The null hypothesis in this case was that the probability of looks to the Target picture was at chance level, i.e. that the probability of looks to the Target and Competitor pictures was not different. We obtained the distribution of the sum statistic under the null hypothesis by randomly permuting the picture labels (Target vs Competitor) by participant, and determining the maximal cluster sum statistic. This was performed 1000 times. We then derived the p values of the clusters in the original dataset by calculating the proportion of permutations with an equal or larger value of the sum statistic.

Cluster based-permutation analysis is well-suited for the analysis of VW eye-tracking data because: a) it provides correction for multiple comparisons without sacrificing statistical power, b) does not lead to an inflated Type 1 error rate due to autocorrelation, and c) allows localization of the effect in time without relying on an arbitrary selection of time windows Yang et al. 2020; Huang & Snedeker 2020.

Serge Minor
Institute for Language and Culture
HSL Faculty, UiT
Breivika Campus, N 9037 Norway
sergey.minor@uit.no

...

Gillian Ramchand
Institute for Language and Culture
HSL Faculty, UiT
Breivika Campus, N 9037 Norway
gillian.ramchand@uit.no