The Processing and Acquisition of Russian Aspect: Evidence from VW Eye-Tracking

Serge Minor, Natalia Mitrofanova, Gillian Ramchand

UiT – The Arctic University of Norway

Aspect marking (the marking of a verb for whether the situation it describes includes boundaries, or is dynamically homogenous) is found in almost half of the world's languages (cf. Dahl & Velupillai 2013). Previous research on the acquisition of grammatical aspect in Slavic yielded diverse results (Weist et al. 1991, Stoll 1998, Vinnitskaya & Wexler 2001, van Hout 2005, Kazanina & Phillips 2007, Gagarina 2009). However very few online processing studies have been conducted. We present the results of a Visual World eye-tracking study which was aimed at investigating the processing of verbal aspect by Russian-speaking children and adults. As far as we know, this is the first large online processing study on the acquisition of grammatical aspect in Slavic. The **goals of the study** were to investigate a) whether children and adults show sensitivity to grammatical aspect both in their offline and online responses, and b) how quickly the children and adults react to aspectual information in the verb. For most verbs in Russian aspectual information is available early (in the prefix or the root), and always before the tense inflection. This means that if aspect is processed incrementally, we should see evidence of this before the verb offset. **The study**. Each trial included an audio preamble which located the narrative in the past (e.g. *It was a sunny day*), followed by a sentence-picture matching task where the participants were presented with a pair of pictures on a screen: one representing an Ongoing Event, i.e. action in progress (Fig. 2) and one representing a Completed Event, i.e. the result that obtained after the action was completed (Fig. 1). While looking at the pictures, the participants were presented with an audio stimulus of a sentence in the past tense with either a perfective or an imperfective verb (cf. ex. 1-2), and were asked to choose the picture that best corresponded to the sentence. Each participant was presented with 24 test trials involving different verbs/event types, as well as filler trials (24 for the Adults, 12 for the Children) which involved pictures representing different event types (i.e. described by different lexical verbs). The participants' eye-movements and offline responses were recorded. **Participants.** We tested 124 adult speakers (mean age=22) and 174 children in three age groups: 3-4 y.o. (n=67), 5-6 y.o. (n=69) and 7-8 y.o. (n=38). **Offline results**. Offline scores revealed that the children of all age groups and the adults pointed to pictures representing Completed Events significantly more often upon hearing sentences with a Perfective verb, and to pictures of Ongoing Events when presented with sentences containing an Imperfective verb (cf. Table 1). Within the children, the 5-6-year-olds performed significantly better than the 3-4-year-olds (p<0.001), and the 7-8-year-olds were marginally better than the 5-6-year-olds (p=0.053). The 5-6 and 7-8-year-olds performed significantly better in the imperfective than the perfective condition (p < 0.01). **Online results**. To investigate the effect of aspect on the gaze patterns, we conducted a cluster-based permutation analysis (Maris & Oostenveld 2007) comparing looks to the Ongoing Event picture in two conditions: Imperfective and Perfective, cf. Fig. 3. The analysis revealed significant clusters of difference between the conditions for all groups of participants. This indicates that all groups of participants showed online sensitivity to grammatical aspect. To evaluate the latency of the effect, we ran a series of paired-sample *t*- tests comparing the proportions of looks to Target and Competitor pictures in 50 ms time bins from the verb onset, for the Adults and the Children who achieved high offline accuracy (>0.85, n=73, mean age=5;10), see Fig. 4 (upper panels). We applied the Holm-Bonferroni method to correct for multiple comparisons. The onset latency of the effect for the adults was almost 300 ms before the average verb offset (at 734 ms, represented by the vertical blue line). For the children, the effect started at about 100 ms after the average verb offset. Given that 4-8 y.o. children generally need between 300-450 ms to plan and execute a saccade (Yang et al., 2002; Trueswell 2008), we can conclude that both the adults and the children reacted to grammatical aspect already before the offset of the verb. Finally, we compared the gaze patterns in the test trials to those in the filler trials where it was sufficient to process the lexical meaning of the verb to identify the target picture. As evident from Fig. 4 (lower panels), the estimates of effect latency are only slightly earlier in the filler trials for both the adults and the children. We conclude that the processing of grammatical aspect in both adults and children happens incrementally and rapidly – already on a *sub-word* level – comparable to the speed of lexical processing.

Examples:

1. Deduška sverlil bol'šuju dyrku
   grandpa drill.IMP.PAST big hole
   'Grandpa was drilling a big hole.'

2. Deduška prosverlil bol'šuju dyrku
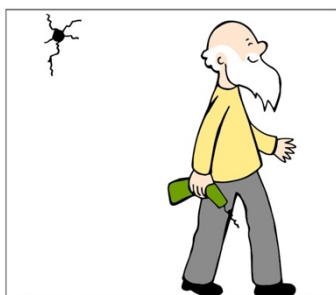   grandpa drill.PF.PAST big hole
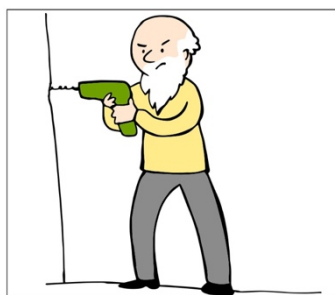   'Grandpa drilled a big hole.'

Fig. 1. Completed event

Fig. 2. Ongoing event

Group	Overall	Imp	Perf
3-4 y.o.	0.66	0.70	0.62
5-6 y.o.	0.82	0.90	0.73
7-8 y.o.	0.87	0.96	0.78
All children	0.77	0.84	0.70
Adults	0.97	0.98	0.95

Table 1. Offline accuracy by participant group and condition (Imperfective vs Perfective).
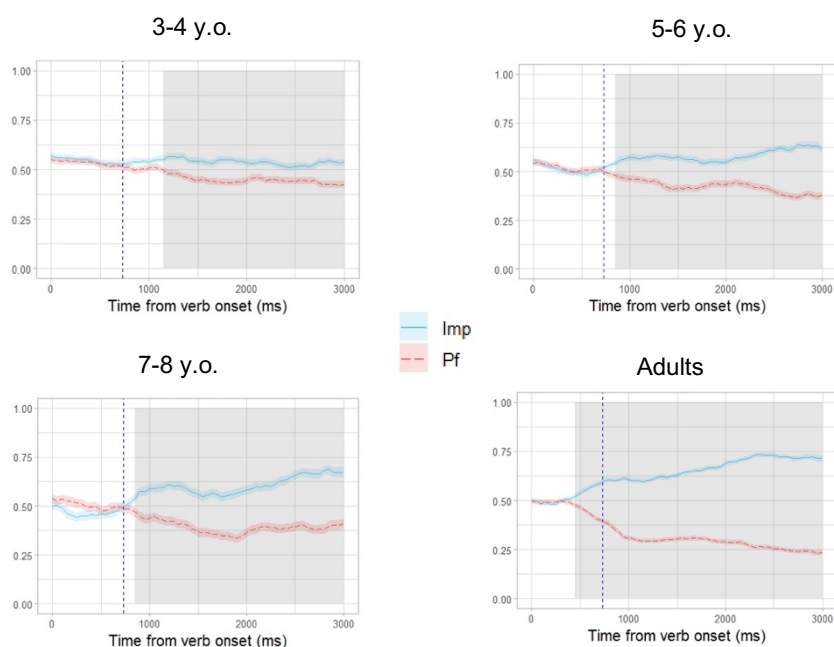

Fig. 3. Prop. of looks to the Ongoing Event picture in 50 ms time bins starting from verb onset by condition (Imperfective vs Perfective) for the four groups of participants. Vertical dashed blue lines represent the average verb offset. Shaded areas represent significant clusters identified via a cluster-based permutation analysis.
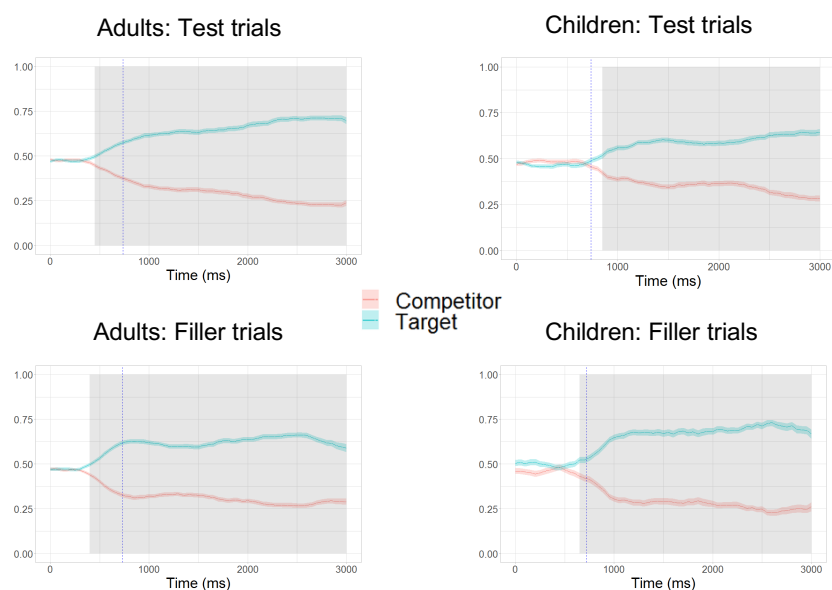

Fig. 4. Prop. of looks to Target vs Competitor pictures from verb onset in the test and filler trials, for the Adults and the Children with >85% offline accuracy. Vertical dashed blue lines represent the average verb offset. Shaded areas represent time bins with significant differences corrected for multiple comparisons (Holm-Bonferroni method).