1. Introduction

Both late and early bilingualism can result in incomplete acquisition of the non-dominant language of the acquirer (see Montrul, 2008, for a thorough discussion). That is, due to lack of abundant and variable linguistic input and opportunities for language use, both the acquisition of a second language (L2) after puberty and the acquisition of the first and non-dominant language in the context of bilingualism, i.e., heritage language (HL) acquisition, may lead to a non-convergent grammar that is different from the monolingual baseline. Despite onset of acquisition in a naturalistic environment before puberty, heritage language acquirers have been reported to have deficits so often observed among L2 learners: problems with discourse pragmatics (Kagan, 2006; Polinsky, 2007), lexicon (Polinsky, 2007; Montrul, 2009); and morpho-syntax, e.g., gender agreement (Montrul, Foote, & Perpiñán, 2008; Polinsky, 2006); differential object marking (Montrul & Bowles, 2008); and tense/aspect/mood (Montrul, 2002, 2007; Pereltsvaig, 2004, 2005; Polinsky, 1997, 2008; Laleko, 2008, 2010). Some of these difficulties, for both L2 and HL acquirers, have been attributed to reanalysis and/or incomplete acquisition of grammatical categories, to a lack of automaticity, as well as to high processing costs involved in the manipulation of the two languages.

Documenting an elaborate linguistic profile of low proficiency American heritage speakers of Russian, Polinsky (1997, 2000) has found that in the worst-case scenario, regardless of the specific path (simultaneous or sequential bilingualism), interrupted acquisition of Russian and switch to English as the dominant language before puberty may result in a limited bilingual, a semi-speaker of Russian, whose grammar is constrained by universal principles, yet systematically different from baseline monolingually acquired (L1) grammars. Polinsky’s further research (2006, 2007, 2008, among other studies) has shown that systematic divergences that low proficiency heritage languages exhibit in grammar are correlated with poor lexical knowledge and low speech rate, and, mainly, that impoverished verbal morphology leads to a restructured and reduced grammatical system. Montrul (2004, 2005) has shown that, despite some non-convergence in the linguistic behavior of heritage language speakers of Spanish, especially in the area of morpho-syntax, the degree of native-like performance and competence in syntax-semantics and syntax-discourse interfaces increases with higher proficiency heritage speakers. Observing the same pattern in L2 acquisition, Montrul (2005) argues that the effects of incomplete acquisition may be limited to syntax-related interfaces and suggests that language acquisition research should focus on finding differences and similarities between HL acquisition and L2 acquisition.

In this paper, I present the results of a portion of a larger study which focuses on the way high-proficiency literate HL and L2 language learners of Russian comprehend aspectual distinctions, which pose an observable difficulty for both early and late acquirers. The Stop-Making-Sense Task discussed here taps into the participants’ sensitivity to a possible mismatch between the aspectual forms they encounter and the disambiguating adverbials in the same sentence. The larger study, and this task in particular, seeks answers to the following research questions:

* I am thankful to Roumyana Slabakova for sharing her proficiency measure tool; to Mila Tasseva-Kurktchieva for thoughtful discussions of the study during preparation and analysis; and to the SLRF 2010 audience and two anonymous reviewers for insightful comments to the presentation and the manuscript. All errors are mine only.

1. Which aspectual contrasts, those based on telicity or on boundedness, present a greater difficulty for incomplete acquirers? Are perfectives and imperfectives acquired to the same degree of success?

2. Which morphological mechanisms for marking aspectual information – prefixation, suffixation, or a combination of both – pose a greater difficulty for incomplete acquirers?

3. Do heritage speakers have advantage over proficiency-matched foreign language learners in their aspectual knowledge?

4. Do heritage learners pattern more with monolingual L1 speakers of Russian or L2 learners of Russian in their aspectual judgments?

2. Russian aspect

Rich verbal morphology makes Russian aspect notoriously complex for linguists, language teachers, and language learners. Morphologically, all Russian finite and non-finite verb forms are either Perfective or Imperfective, with no apparent consistency of marking one or the other. Risking an overgeneralization, we can say that, as a rule, prefixation, suffixation, and suppletion are the main active mechanisms for the formation of aspectual pairs, with prefixation usually generating perfective verbs from primary imperfectives and suffixation generating secondary imperfectives (SI) from perfectives, as shown in (1).

\[(1)\]

<table>
<thead>
<tr>
<th></th>
<th>Imperfective</th>
<th>Perfective</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. prefixation</td>
<td><em>pisat</em></td>
<td><em>DO-pisat</em></td>
</tr>
<tr>
<td>b. SI suffixation</td>
<td><em>da-VA-tj</em></td>
<td><em>atr</em></td>
</tr>
<tr>
<td>c. suppletion</td>
<td><em>saditšja</em></td>
<td><em>sesh</em></td>
</tr>
</tbody>
</table>

In (1a), a prefix attaches to a simplex imperfective verb to form a perfective counterpart of the pair; in (1b), a SI suffix attaches to a perfective verb to create its imperfective counterpart by deperfectivizing it; and in (1c) the two members of the aspectual pair are stored in the lexicon. While quite a few aspectual pairs are based on suppletion and there is a considerable number of biaspectral verbs (mainly borrowings), in this study, I am particularly interested in the first two mechanisms of aspect marking. I look into perfectivizing prefixes and SI suffixes because their acquisition involves learning the rule-based mechanisms of overt aspectual morphology rather than lexical knowledge.

The prefixation mechanism involves a number of polysemantic derivational prefixes (19-21 according to Slabakova, 2005a), some of which are believed to yield aspectual changes only (*pisat* – NA-*pisat* ‘write – write something down’) while others seem to add a change of lexical meaning to the derived word as well (*pisat* – DO-*pisat* ‘write – finish writing something’). In contrast to the rich prefixation mechanism, aside from the non-productive -*a-*, there is only one productive SI imperfectivizing suffix -*(y)va-, which has a few phonologically conditioned allomorphs. For a learner, the emerging pattern of the way overt aspectual morphology works may be as follows: while SI suffixation consistently leads to imperfectivity, prefixation results in perfectivity, only if the verb does not undergo further SI suffixation. Thus on the surface, in the overall morphological realization of aspectual meaning, the mechanism of prefixation is less consistent and more complex than the more regular SI suffixation. This complexity could be one of several potential difficulties in the acquisition of Russian aspect; however, this is only part of the difficulty because these morphological patterns are only reflexes of the underlying syntactic and semantic contrasts.

In this study, I follow the assumptions in Slabakova (2001, 2005a,b) for English and Slavic aspect, which are compatible with the analyses in other recent proposals (Nossalik, 2009; Laleko, 2010) that distinguish lexical (*inner*) aspect from grammatical (*outer*) aspect. The two types of aspect occupy two separate syntactic positions (i.e., one inside and the other above the little vP) and semantically are associated with two distinct features, i.e., telicity and boundedness, respectively. Languages may converge and/or differ in how they mark aspectual information morphologically – for example, while English marks boundedness on the verb and telicity on the direct object, Russian marks both semantic features via verbal morphology, but by two different morphemes. In what follows, I briefly describe the interaction of the two semantic features in English and then Russian, and follow by discussing
relevant parametric differences in aspect marking between English and Russian. I conclude by outlining several implications for language acquisition.

From the point of view of lexical aspect, predicates can be grouped based on telicity, i.e., on whether they denote events with an inherent limit/endpoint (along the lines of Vendler, 1957). Telic predicates contain inherent endpoints, as in *recognize a name* (achievement) and *write a book* (accomplishment). Atelic predicates do not have an inherent endpoint, as in *love music* (state) and *read books* (activity). As argued by Slabakova (2001,2005a,b), the value of the universal semantic feature telicity is not specified in the lexicon for all verb types.

Vendler’s classes can also be grouped based on their dynamicity, i.e., whether they denote events that can extend over time and contain a process or are devoid of process and hold at instants. While non-dynamic states and achievements are stored in the lexicon as [-telic] and [+telic], respectively, dynamic activities and accomplishments, which constitute the majority of verbal stems (Slabakova, 2001; Travis, 1994), are lexically underspecified as [α telic] and get their telicity value set compositionally. Importantly, languages may differ parametrically in the ways the value-setting can be achieved: lexicalized in some verbs; encoded by derivational morphology, (e.g., by prefixes in Russian); or, as in English, encoded by inflectional markers on the direct object. If a lexically underspecified English verb is followed by a quantized object (2a), the predicate is telic, and if the object is non-quantized (2b), the predicate is atelic.

(2) a. *He wrote the/three/those letters.* [+telic], accomplishment
    b. *He wrote Ø letters/fiction.* [-telic], activity

In Russian, the telicity value of underspecified verbs is achieved via prefixation: regardless of the cardinality of the object, non-prefixed verbs of this type constitute atelic predicates (3a) and prefixed verbs constitute telic predicates (3b).

(3) a. *Kolja čítal (eti) pisjmа.* [-telic], activity
    Kolja Ø.read.PAST (these) letters
    ‘Kolja would read/was reading (these) letters.’

    b. *Kolja PROčítal (eti) pisjmа.* [+telic], accomplishment
    Kolja PREFIX.read.PAST (these) letters
    ‘Kolja read (these) letters.’

In contrast to lexical aspect, which is a property of predicates, grammatical aspect applies to events described by whole sentences and reflects different ways of viewing the internal temporal constituency of the whole situation. This is closely related to the checking of the feature boundedness, which, unlike telicity, refers to whether an event has reached its actual endpoint. Together with telicity, boundedness helps to build the full compositional aspectual interpretation of the clause as either Perfective or Imperfective. For example, two events can be both telic, but different in grammatical aspect (4).

(4) a. *I ate a piece of cake last night.* [+telic], [+bound] → Perfective
    b. *I was eating a piece of cake when she called.* [+telic], [-bound] → Imperfective

The Perfective viewpoint looks at the situation from outside, disregards the internal structure of the situation (5a), and renders the event that has a potential endpoint as completed (bounded); the predicate not only has a potential endpoint, but has actually reached it. The Imperfective viewpoint looks at the situation from inside focusing solely on the internal structure of the situation regardless of

---

1 As pointed out by Slabakova (2005a), with the highly polysematic nature of Slavic prefixes, it is best to speak of *uses* or *senses* of a particular prefix. E.g. the same prefix na- may be used in its purely telicizing sense *napisal pis’mo*– ‘to write a letter’, or also have an additional lexical meaning of ‘in excess’ in *nagotovil edy* – ‘to cook too much food’ or an additional meaning of ‘onto/over’ in *nabrostit jel* (na pleči) – ‘throw a coat on (one’s shoulders)’.
its beginning or end (5b). The Imperfective viewpoint subsumes the habitual (5c) and the ongoing (5b) viewpoints, since both these meanings are unbounded.

(5)  
   a. He has read this book.  [+telic], [+bounded] → Perfective  
   b. He was reading this book.  [+telic], [-bounded] → Imperf., ongoing  
   c. He reads this book every night.  [+telic], [-bounded] → Imperf., habitual

Grammatical aspect has scope over clauses and is most often expressed by inflectional morphology combining tense and aspectual information (e.g., by verbal suffixes, both in Russian and English). Both English and Russian use imperfectivizing suffixation to signal an ongoing event, but in Russian the suffix is selectionally restricted only to telic predicates. Another difference from English is that the Russian SI suffix -(y)va denotes both ongoing and iterative/habitual telic events (6a -d). The SI suffixes can attach to both types of telic predicates: those specified for telicity in the lexicon (6b) and those that derive a telic interpretation via prefixation (6d):

(6)  
   a. Kolja zakazal bilet na poezd.  [+telic], [+bounded] → Perfective  
      Kolja order.PAST ticket for train  
      ‘Kolja ordered a train ticket.’
   b. Kolja zakazYVAL bilet na poezd.  [+telic], [-bounded] → Imperfective  
      Kolja order.SI.PAST ticket for train  
      ‘Kolja was ordering/would order a train ticket.’
   c. Kolja PEREčital eti pisjma.  [+telic], [+bounded] → Perfective  
      Kolja PREFIX.read.PAST these letters  
      ‘Kolja reread these letters.’
   d. Kolja PEREčitYVAL eti pisjma.  [+telic], [-bounded] → Imperfective  
      Kolja PREFIX.read.SI.PAST these letters  
      ‘Kolja was rereading/would reread these letters.’

To sum up, from the semantic point of view, telicity encodes presence/absence of an inherent limit of the event in predicates. Boundedness indicates that the event described by the whole sentence has reached its actual limit. Calculation of telicity is crucial to and precedes the calculation of boundedness. Language-specific ways of marking these features may differ. Morphologically, in English, telicity is marked on the object while boundedness is marked by verbal suffixes. In Russian,

2 It is worth noting that while most perfectives formed with purely telicizing morphemes do not allow SI imperfectivization, most of those that are formed with telicizing prefixes that add some lexical information to the denotation of the verb do. For example, in the aspectual triplet pisatj–dopisatj – dopisyvati, the prefix do- has the additional meaning of finishing. At the same time pisať can form an aspectual pair, which does not allow SI pisať – napisať with na- as a purely telicizing prefix. However, other verbs allow triplets even with purely telic prefixes. I believe that knowing which stems can combine with which affixes is part of encyclopaedic rather than grammatical knowledge.

3 There are two types of aspectual morphemes that are excluded from discussion since they are not the object of this study. Like prefixes, a semelfactive suffix -nu- can create a perfective member of an aspectual pair (prygať-prygnutj ‘to jump – to jump once’). However, the status of nu- (and semelfacts) and the delimitative po- is not straightforward, since in some accounts po- is considered a telicizing morpheme while in others a marker of grammatical aspect. There are instances where delimitative prefixes po- and pro- may be used on atelic verbs as boundedness markers rather than telicity markers in sentences like the following:

   On porešal zadaču, no tak ee i ne rešil.  
   ‘He spent a little time solving the problem, but did not solve it.’

   Slabakova (2005a) calls such an instance of po- an external prefix, which has adverbial properties and has no effect on the telicity of the verb. This adverbial nature also suggests that these prefixes may be syntactically different from telicizing prefixes which occupy an aspectual projection inside the vP. See Nossalik 2009 for a similar observation.
both aspectual features are overtly marked on the verb, but by different morphemes: telicity by the presence/absence of prefixes and boundedness by the presence/absence of SI suffixes.

From the point of view of syntactic representation, for predicates lexically underspecified for telicity, Slabakova (2001, 2005a,b) assumes two separate aspectual projections/phrases (AspP): the lower AspP for lexical and the upper AspP for grammatical aspect. While both prefixes and SI suffixes participate in the compositional marking of the aspectual meaning of the whole clause, Slabakova’s crucial claim is that Russian prefixes and SI suffixes separate the tasks of checking lexical and grammatical aspect. That is, first the (a)telicity value of the verb is calculated based on the presence or absence of a prefix and then the (un)boundedness feature is checked based on the presence or absence of a SI suffix. In contrast, lexically telic predicates are VPs that have only one aspectual projection – that for grammatical aspect – and are thus structurally simpler than the vPs headed by verbs lexically underspecified for telicity. In both languages, when lexically telic predicates like (6b) appear in the ongoing interpretation (marked with –ing or -(y)va-, respectively), a semantic shift turns the non-dynamic achievement into a dynamic accomplishment by means of coercion, a pragmatically induced process, which happens at the CP level and should not affect the syntactic structure of the predicate.

Successful acquisition of the Russian aspectual system involves acquisition of the semantic features telicity and boundedness. A learner acquiring telicity needs to deduce that, with some exceptions, prefixed verbs are [+telic] and to learn all the polysemantic derivational prefixes with their subsets of lexical meaning, along with each individual verb root and its subset of prefixes – a rather complex lexical task, indeed. Acquisition of boundedness, at least with respect to secondary imperfectivization, seems to require less from the learner. A learner needs to deduce (1) that all verbs suffixed with a SI morpheme are both [-bounded] and [+telic] (that the regular inflectional -(y)va- with its allomorphs can attach only to telic verbs) and (2) that unlike the English –ing, a Russian SI morpheme cannot appear on [-telic] predicates, but can encode both ongoing and iterative/habitual [+telic] events.

3. Acquisition of Russian aspect

Acquisition of Russian aspect has received attention in recent years, both in adult second language acquisition research (Slabakova, 2003, 2005a,b; Nossalik, 2008, 2009) and heritage language acquisition research (Bar-shalom & Zaretsky, 2008; Gupol, 2009; Polinsky, 2008; Laleko, 2008, 2010). While the studies differ in their research questions, acquisition focus and methodology, what seems to be a common observation is that there are asymmetries in the acquisition of Russian aspectual contrasts, i.e., the Imperfective is more difficult to acquire than the Perfective.

Bar-Shalom & Zaretsky (2008) report that, despite various vocabulary errors, problems with case and subject-verb agreement and instances of code-switching in their narrative study, Russian-English bilingual children (ages 4-10) had successfully acquired both the semantics of Perfective and Imperfective and the morphological mechanisms for marking aspectual contrasts, including the aspectual pairs. In contrast, Gupol (2009) reports incomplete acquisition of Russian aspect by Russian-Hebrew bilinguals (ages 4-8) since most of the elicited production errors were tense and aspect errors, including contextually inappropriate use of grammatical aspect. Gupol claims that having acquired the inventory of Russian inflectional and derivational morphology, bilingual children do not have full command of their function: i.e., they do not accept imperfective morphology for accomplishment verbs in ongoing situations and use perfective morphology in both incomplete and on-going situations.

Looking into aspectual systems of adult heritage speakers of lowest proficiency, Polinsky (2008) found non-convergence with the monolingual baseline in several areas during spontaneous production and forced choice tasks, which led her to argue for restructuring of aspect in heritage grammars due to impoverished functional morphology and a compressed lexicon. Polinsky found a smaller set of prefixed perfective forms, with variation across speakers, overgeneralized use of the imperfectivizing suffix -(y)va- in production and poor judgments of aspectual distinctions in comprehension. There was also overall loss of Perfective-Imperfective aspectual pairs, with retention of only one member of the pair for both meanings – that is, use of imperfective forms in perfective contexts and vise versa.

Laleko (2010) compared high proficiency adult heritage Russian speakers and monolingual controls in patterns of production, interpretation, and acceptability judgments of aspectual contrasts. Despite error-free production, the HL group showed non-convergence with monolingual baseline in
the use of the Imperfective in comprehension tasks. Laleko reports that heritage speakers in her study were less likely than the controls to accept imperfective verb forms for completed events (despite the presence of contextual discourse-pragmatic triggers of imperfectivity) and show lower accuracy interpreting annulled result implicature. Laleko (2010) explains her findings by arguing that advanced heritage language speakers show selective (covert) aspectual restructuring, mainly at the highest level of sentential structure, C-domain, from which syntax is mapped onto discourse-pragmatics. She maintains that high proficiency heritage speakers are perceived as indistinguishable from the controls in the production because the non-convergence of their heritage aspectual systems manifests itself in infelicity rather than ungrammaticality and in a narrower range of contextual use of the Imperfective.

Slabakova’s (2003, 2005b) study was devoted to the L2 acquisition of telicity, a feature of lexical aspect. She found that advanced and high intermediate adult learners of Russian knew that it is the presence or absence of the perfective prefix on the verb, rather than the form of objects, that contribute to the calculation of telicity in Russian, paid no attention to the form of the object, and were able to arrive at correct entailments in the interpretation task. As she predicted, low-intermediate learners were not yet able to overcome the effects of L1 transfer and made errors, which significantly depended on the cardinality of the object. Slabakova argues that the perceived difficulty in acquiring Russian aspect does not consist in acquiring the grammatical mechanisms of telicity marking, but rather in learning the lexical component of telicity marking. Slabakova also reports that non-convergence in interpretations of imperfective sentences (with primary imperfectives) between low-intermediate L2 and the monolingual control group can be explained a degree of salience of interpretations between L2 learners and native speakers, with the latter commanding a greater repertoire of possible contexts for the use of the Imperfective.

Nossalik (2008, 2009) tested adult L2 Russian acquisition of boundedness, i.e., grammatical aspect, in predicates with verbs that are lexically underspecified for telicity. Nossalik reports successfully acquired complex syntactic structure of such predicates by advanced and high-intermediate learners, who were able to “block” the English mechanisms of aspectual marking by disallowing the atelic verbs to be inflected with -(y)var- and to receive future tense interpretation with present tense inflections and, thus, arrive at correct semantic entailments, including aspectual shifts. At the same time some low-intermediate learners showed effects of transfer in the interpretation of the Imperfective. Nossalik (2009) conducted a truth value judgment task and a grammaticality judgment task, which showed that near-native speakers behaved indistinguishably from monolingual controls. The convergence of the L2 learners with the L1 baseline in their syntactic knowledge led Nossalik to argue that purely morpho-syntactic properties of Russian aspect are acquirable in second language context. However, in this study, even advanced learners did not converge with the monolingual controls at the interfaces between syntax and the lexicon and syntax and pragmatics. Nossalik concluded that restructuring of L2 grammar and native-like attainment of tense/aspect knowledge is possible, but not ubiquitous.

Based on the studies outlined above, in the worst case scenario, both child and adult heritage speakers suffer from morphological and representational deficits in both production and comprehension (Gupol, 2009; Polinsky, 2008), and in the best case scenario they maintain practically error-free production, but may retain problems mapping aspectual information onto discourse-pragmatics level (Bar-Shalom & Zaretsky, 2008; Laleko, 2010). Importantly, while some of the HL studies tested the interpretation of Russian aspect, none of them focused on teasing apart what kind of aspectual contrasts may be restructured or incompletely acquired: those based on telicity or those based on boundedness. Several L2 studies (Slabakova, 2003, 2005; Nossalik, 2008, 2009) tested empirically whether English-speaking adult L2 learners are capable of acquiring the two aspectual contrasts and interpreting aspectual morphology in a native-like way. Their empirical findings support the Full Transfer/Full Access (Schwartz & Sprouse, 1996) assumption that, despite starting out constrained by the representations of their L1, adult L2 learners can overcome the parametric differences with their native grammar and restructure their grammar to that of L2 parameters, and, thus, fully acquire a functional category.

A common thread between the L2 studies and those on heritage language acquisition is that the Imperfective, which is unmarked in Russian and allows a wider range of interpretations than the perfective, may present a greater difficulty for incomplete acquirers. The unmarked and structurally and semantically complex Imperfective rather than marked and lexically complex Perfective seems to present a greater difficulty for children and adults acquiring Russian.
4. The experiment

In Mikhaylova (Forthcoming), I compared proficiency-matched HL and L2 learners in their sensitivity to telicity and boundedness and the associated morphological mechanisms using a Semantic Entailments Task. The participants were asked to choose whether one particular continuation or both continuations provided for each sentence like (3) or (6) were logically possible (7).

(7) Valja PROčitala detektiv…
‘Valya read the detective story…’

a) ...i ej ne ponravilsja konec
‘...and she didn’t like the ending’

b) ...i ona hotela uznat j konec
‘...and she really wanted to find out the ending’

c) oba varianta vozmožny
‘both variants are possible’

The learners had to interpret the event in each sentence as either completed or incomplete, and the relevant verbal morphology was the only clue available for calculating correct interpretations, with no other contextual clues to rely on. Overall, the L2 group seemed to be going through what Slabakova (2008) calls a morphological bottleneck: they were significantly more accurate on sentences with no overt aspectual morphology, performing almost at chance in their interpretation of all the aspectually affixed types of predicates. At the same time, the heritage group was statistically similar to both the L1 and the L2 groups. Interestingly, the HL speakers were statistically more accurate than L2 learners in boundedness contrasts in the lexically underspecified perfective and SI accomplishment predicates, i.e., the HL group had a morphological advantage over the L2 group.

The task presented in this paper tested the same groups of participants and targeted the same aspectual contrast conditions using a different method: adding a disambiguating adverbial to each sentence and shifting the participants’ attention slightly from the logic of the sentences to the grammaticality of the sentences.

4.1. Participants

The participants whose results are discussed in this paper are those heritage speakers and foreign language learners from a larger pool who scored within the native speaker range on the independent proficiency measure (see Section 4.2) and, thus, were deemed of high proficiency. All the participants in the two test groups were dominant in English. All were literate in Russian and, by the time of the study, had been exposed to formal instruction (were or had been enrolled in courses of Russian at the college level). The group of heritage language speakers (HL) consisted of 22 high proficiency heritage speakers of Russian (see Table 1 for age range). The group of learners of Russian as a foreign language (L2) consisted of 11 high proficiency foreign language learners of Russian. The control group (L1) consisted of 30 monolingual native speakers of Russian, tested in Russia. All controls were college students without professional (meta-)linguistic training (i.e., not majoring in Linguistics or language sciences).

Table 1. Age of the participants at the time of study

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean (range)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>30</td>
<td>21 (16 – 40)</td>
<td>5.47</td>
</tr>
<tr>
<td>HL</td>
<td>22</td>
<td>21 (19 – 28)</td>
<td>1.94</td>
</tr>
<tr>
<td>L2</td>
<td>11</td>
<td>28.5 (20 – 64)</td>
<td>14.20</td>
</tr>
</tbody>
</table>

4.2. Cloze test (proficiency measure)

To measure the proficiency level of the test groups, I replicated Slabakova’s (2005b) proficiency measure (cloze test), which was originally used in her study of L2 acquisition of Russian telicity. The cloze test consisted of a continuous text (a story about seasons) with 30 blank spaces substituting
single words. Participants were asked to fill in the blanks choosing the only correct option of the three options provided in the drop-down menu. All participants, including the control group, completed the cloze test online after finishing the Semantic Entailment Task and before filling out the linguistic background questionnaires. Table 2 shows mean accuracy of the control group and the two test groups.

Table 2. Accuracy on the proficiency measure (% correct choices)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean (range)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>30</td>
<td>96.4 (80 – 100)</td>
<td>4.9</td>
</tr>
<tr>
<td>HL</td>
<td>22</td>
<td>95.6 (83.3 – 100)</td>
<td>4.0</td>
</tr>
<tr>
<td>L2</td>
<td>11</td>
<td>87 (80 – 96.7)</td>
<td>6.2</td>
</tr>
</tbody>
</table>

It is worthwhile to say a few words about the proficiency level of the two test groups reported here. Following Slabakova’s methodology, those heritage language learners and foreign language learners of Russian whose scores on the cloze test were within the range of scores in the control group of monolingual native speakers, were considered to be of high proficiency and were selected for analysis in this paper. However, when comparing the two test groups in this study, it is important to remember that based on a one-way ANOVA and Bonferroni Post Hoc test, statistically the three groups were not balanced in their proficiency. More specifically, the L1 group was statistically more accurate on the proficiency measure than the L2 group (F(2,62) = 16.17; p < 0.0001; d = 1.93). And the HL learners, whose average proficiency score was almost identical to that of L1 controls (no statistical difference), were more accurate on the proficiency measure than the L2 group (F(2,62) = 16.17; p < 0.0001; d = 1.77).

Given such distribution of accuracy on the proficiency measure, it would be logical to expect that the HL group would behave similarly to the L1 group and differently from the L2 group on the target task as well. However, based on the results of the Semantic Entailments Task in Mikhaylova (Forthcoming), the HL participants could behave similarly to the L2 group as well.

4.3. The task

The Stop-Making-Sense Task tested the participants’ sensitivity to telicity and boundedness contrasts and their morphological realization in different predicate types. This was a self-paced reading task (created via Linger software package and administered on the researcher’s computer) in which the participants were asked to read Russian sentences until the sentence stopped making sense. The participants saw only one word of the sentence at a time, with the other words covered by dashes and no opportunity to back-track. The participants tapped “J” on the keyboard to move to the next word, until they felt the sentence stopped making sense. As soon as they noticed an error, they were asked to press “F” on the keyboard and they were taken to the next sentence. If the sentence made sense all the way through, the participants were automatically taken to the next item. Before starting the experiment, each participant completed a training session with the researcher; the main experiment was done without the researcher’s immediate presence. Most participants completed this task during a separate session, either before or after the Proficiency Measure and the Semantic Entailment Task described in Mikhaylova (Forthcoming).

The task consisted of 42 target items (3 conditions, 14 items per condition) and 66 filler items. The order of the sentences was automatically randomized for each participant. All the items were based on vocabulary from the textbook *Golosa*, Volumes 1-2, a widely used college textbook for beginner level Russian courses. For predicates with verbs not specified for telicity in the lexicon (activity-accomplishment type), I used only those stems that could appear both in prefixed and prefixed-SI forms. To control for effects of phonological transparency, the SI test items only contained -(y)va-. Each target sentence started with a disambiguating adverbial, which was either compatible or incompatible with the form of the predicate. All fillers also started with adverbials – for consistency.

---

4 An anonymous reviewer raised a concern about the comparison of two groups that are statistically unbalanced for proficiency. I find it is very informative to compare the two types of literate language learners at the highest level of proficiency available in college level classrooms and scoring within the native speaker range on the proficiency measure, even if one population performs statistically worse than the other.
and as a distraction from the target items. The task was based on the same three conditions and sets of verb stems that were used in the Semantic Entailment Task in Mikhaylova (Forthcoming). In each condition half of the sentences were perfective and half imperfective, contrasting only in one aspectual morpheme:

CONDITION 1: activity-accomplishment predicates contrasting in telicity
CONDITION 2: accomplishment predicates contrasting in boundedness
CONDITION 3: achievement predicates contrasting in boundedness

The first condition tested telicity contrasts in dynamic predicates which are [α telic] in the lexicon and which encode both telicity and boundedness via presence or absence of aspectual morphology. The condition included sentences with [-telic; -bounded] activities (1A), which carry no overt aspectual markers, and [+telic; +bounded] accomplishments (1B), which carry a telicizing prefix (either used in a purely telicizing sense or adding additional lexical meaning to the verb), but no SI suffix. This condition manipulated the compatibility of the predicate with adverbials equivalent to the English in X time/for X time. Only [+telic] predicates would be semantically incompatible with the adverbial for X time while [-telic] predicates would be impossible with in X time. For each participant Linger automatically selected which seven of the fourteen sentences would be presented in a plausible adverbial/predicate combination and which would appear in illicit combinations. Types (1A) and (1B) are structurally the same, but differ in morphological complexity.

CONDITION 1: pairs based on TELICITY for lexically underspecified predicates

A. \( V^5 \) (no aspectual affixes) [-telic; -bounded] activity = Imperfective
   \( Dva \ \text{časa}/*/Za \ dva \ \text{časa} \ Vladimir \ pisal \ pisjmo. \)
   For 2 hours/*In 2 hours Vladimir was writing a letter.

B. PREFIX +V [+telic; +bounded] accomplishment = Perfective
   \( *Dva \ \text{časa}/Za \ dva \ \text{časa} \ Vladimir \ DOpisal \ pisjmo. \)
   *For 2 hours/In 2 hours Vladimir wrote/finished a letter.

Conditions 2 and 3 manipulated the compatibility of the boundedness value of the predicate with an all X time period type of modifier. Only [+bounded] events would be semantically incompatible with the adverbial all X time period. Conditions 2 and 3 tested the same semantic feature, but in different types of predicates. All [+bounded] sentences were implausible, and all [-bounded] sentences were plausible. As in Condition 1, Condition 2 included verbs lexically underspecified for telicity, in which both aspectual features are encoded with presence/absence of aspectual morphology: i.e., prefixed [+telic; +bounded] accomplishments (2C) and prefixed [+telic; -bounded] accomplishments also marked by an SI suffix (2D). The former (2C) is essentially the same type of predicate as (1B), since dynamic verbs are capable of forming aspectual triplets differing in both telicity and boundedness. In the two conditions, (1B) and (2C) contained the same verbal stem in different sentences and against different adverbial modifiers. Types (2C) and (2D) are structurally the same, but (2D) is morphologically more complex.

CONDITION 2: pairs based on BOUNDEDNESS for lexically underspecified predicates

C. PREFIX+V [+telic; +bounded] = Perfective
   \( \text{*Ves\' den\' Vladimir \ DO\čital \ detektiv.} \)
   ‘*All day Vladimir finished reading the detective story.’

5 The terms \( V, \) PREFIX and SI are used here only as a reference to the presence/absence of overt aspectual morphology on the verb, rather than to provide its full morphological decomposition. While pisal actually has two affixes (thematic vowel -a- and Past Tense suffix –l), neither of them are aspectual morphemes that affect the aspectual status of the predicate (or fill an aspectual projection) as telicizing prefixes and SI suffixes do; hence the predicate is labeled \( V. \)
Finally, Condition 3 included non-dynamic achievements, which are lexically specified as [+telic] and only encode the predicate’s boundedness via presence/absence of a SI suffix. Achievement verbs without aspectual morphology head [+telic; +bounded] predicates (3E) and those carrying an SI suffix head [+telic; -bounded] predicates (3F). The non-dynamic predicates in this condition are structurally simpler than those in Conditions 1-2. Type (3F) is structurally and morphologically the simplest among all the predicates manipulated in the task (A-F) and type (2D) is the most structurally and morphologically complex.

**CONDITION 3: pairs based on BOUNDEDNESS for lexically specified telic achievement predicates**

**E. V (no aspectual affixes) [+telic; +bounded] = Perfective**

*Vesj večer Vladimir zakazal bilet v Moskvu.*

‘All evening Vladimir ordered a ticket to Moscow.’

**F. V+SI [+telic; -bounded] = Imperfective**

*Vesj večer Vladimir zakazYVAL bilet v Moskvu.*

‘All evening Vladimir was ordering a ticket to Moscow.’

To reiterate, in Conditions 1 and 2, the described events are encoded by lexically underspecified verbs differing in their telicity and boundedness values, respectively, while Condition 3 included lexically [+telic] predicates, differing in their boundedness. That is, events denoted by non-prefixed (and, thus, [-telic]) predicates as (A) in Condition 1 are usually interpreted as Imperfective – they have no inherent limit, and, without semantic coercion via adverbial modification or context, they are [-bounded] by default. There is also a selectional restriction that [-telic] predicates never undergo further SI suffixation. Events denoted by prefixed (and, thus, [+telic]) predicates of the lexically underspecified type (B in Condition 1, and C in Condition 2) get a Perfective interpretation – they contain an inherent limit marked by the prefix, and, without further SI suffixation, get interpreted as completed events. Overt morphology on the prefixed-SI suffixed verbs (D in Condition 2) signals that predicates with these lexically underspecified verbs are [+telic], yet the event is unbounded – at the moment of speech, the inherent limit has not been reached.

Conditions 1 and 2 include the semantically and structurally complex predicates with lexically underspecified verbs; so, both telicity and boundedness need to be calculated for successful interpretation of the sentence. Condition 3, in contrast, includes structurally simpler predicates with lexically marked [+telic] non-dynamic verbs; so the processing task only involves calculation of the boundedness feature. Condition 2 is both structurally, semantically, and morphologically more complex, and, therefore, should be the most difficult from the point of view of processing.

### 5. Results and discussion

In order to answer the research questions posed at the beginning of this paper, I examined each group’s performance on the task. Based on the total means for the whole task, the control group performed with an overall higher accuracy than the HL and the L2 group on the Stop-Making-Sense (SMS) Task, with the HL group outperforming the L2 group (see Table 3).

**Table 3. Average accuracy on the SMS task (% correct choices)**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean (range)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>30</td>
<td>91.0 (88.4 – 93.6)</td>
<td>6.9</td>
</tr>
<tr>
<td>HL</td>
<td>22</td>
<td>85.2 (81.9 – 88.6)</td>
<td>12.1</td>
</tr>
<tr>
<td>L2</td>
<td>11</td>
<td>70.8 (58.8 – 82.8)</td>
<td>17.8</td>
</tr>
</tbody>
</table>
A one-way ANOVA and Bonferroni Post Hoc test revealed that while the mean of total accuracy scores for foreign language learners was statistically different from that of the monolingual controls (F(2,62)=12.96; p<0.0001; d=2.93) and that of heritage learners (F(2,62)=12.96; p=0.005; d=2.00), as in the proficiency measure, there was no statistical difference between the heritage learners and the controls (F(2,62)=12.96; p=0.142).

Figures 1-2 show that the mean scores of monolingual controls were equally high on all three conditions and on specific predicate types within each condition (A-F) defined in the previous section. It also appears that the two test groups show some asymmetries in their sensitivity to aspectual contrasts and that the two test groups did not converge in the patterns of their sensitivity to the three conditions and the types of predicates. The following sections discuss the results of each group individually, and Section 5.4 discusses similarities and differences between the groups.
5.1. The control group

A series of paired-samples T-Tests showed that there was no statistical difference between the accuracy of native speakers on the three conditions (see means for each condition in Figure 1), i.e., the L1 group was equally sensitive to aspectual contrasts based on telicity and boundedness. There was also no statistical difference between the scores of the control group on the members of each pair, i.e., between 1A and 1B, between 2C and 2D, and between 3E and 3F. The battery of tests did not show any difference between the accuracy of the control group on judgments of perfective predicates (between 1B and 2C; 1B and 3E; or 2C and 3E) or of imperfective predicates (between 1A and 2D; 1A and 3F; or 2D and 3F). This suggests that both types of aspectual contrasts and the associated morphological mechanisms are stable in adult L1 grammars. The lower scores on primary imperfectives (1A) and SI achievements (3F) are not statistical.

5.2. The HL group

The mean scores in Figures 1-2 show that heritage speakers achieved higher accuracy on both boundedness conditions (Conditions 2-3) than on telicity contrasts (Condition 1). A series of paired-samples T-Tests revealed that in predicates with lexically underspecified activity-accomplishment verbs the HL group was statistically less accurate on pairs differing in telicity (Condition 1) than on those contrasting in boundedness (Condition 2): \( t(21) = -2.222; p = 0.037; \text{Cohen's } d = 0.51 \). The HL group also performed statistically worse on Condition 1 than on boundedness in lexically telic achievements in Condition 3 (\( t(21) = -2.642; p = 0.015; \text{Cohen's } d = -0.61 \)), with no difference between the two boundedness conditions. Just as with the native controls, the accuracy of the HL group within each condition was balanced, i.e., there was no statistical difference between their scores on predicate pairs 1A and 1B, 2C and 2D, and 3E and 3F (see Figure 2 for the means). So, in answer to the first research question we can say that telicity contrasts pose a greater difficulty to heritage speakers than boundedness contrasts.

Also, just as with the control group, the battery of tests did not show any difference in the way heritage speakers judged Perfective predicates (i.e. they achieved statistically indistinguishable scores on the two sets of accomplishments in 1B and 2C; accomplishments in 1B and achievements in 3E; and on accomplishments in 2C and achievements in 3E). However, unlike the control group, the HL group was statistically less accurate on primary imperfectives ([+telic] activities) in 1A than on the two types of [+telic] secondary imperfectives. More specifically, their scores on morphologically unmarked activities (1A) were statistically lower than their scores on structurally simpler but morphologically more complex SI-suffixed achievements (3F): \( t(21) = -2.642; p = 0.015; \text{Cohen's } d = -0.61 \). Activities, which have no overt aspectually marked, presented a greater challenge than the morphologically more complex prefixed and SI-suffixed accomplishments (2D): \( t(21) = -2.300; p = 0.032; \text{Cohen's } d = -0.52 \). Finally, there was no statistical difference between the HL group’s mean scores on the two types of SI predicates (3D and 3F).

Given the results, not only boundedness contrasts instantiated by SI morphology seem rather stable in the HL grammar regardless of predicate type, but also judgments of the Perfective seem more stable than the judgments of the Imperfective. This suggests that primary imperfectives rather than secondary imperfectives or perfectives are the vulnerability point in this task. At the same time the group’s lower performance on primary imperfectives seems related to their difficulty with telicity contrasts. To be fair, native speakers also scored lower on the primary imperfectives than on SI predicates, but that difference was not statistical. It is difficult to provide a definite answer to the second research question judging only by this task. In this task, the regular SI suffixation mechanism does not seem to pose a difficulty for heritage speakers in on-line processing, even in the SI accomplishments; however, the irregular prefixation mechanism may be mistaken for lack of aspectual morphology and, thus, may be more challenging to parse on-line.

5.3. The L2 group

If we analyze the means of the L2 group’s results on the three conditions (see Figure 1), the pattern is different from that of native controls, and even that of heritage speakers. The L2 group was
less accurate in Condition 2 than in the other two conditions: boundedness contrasts in accomplishment predicates posed a greater difficulty to the L2 learners than telicity contrasts in the activity-accomplishment predicates in Condition 1 ($t(10)=2.241; p=0.049$; Cohen’s $d=0.50$) and even more so than boundedness contrasts in achievements in Condition 3 ($t(10)=3.218; p=0.009$; Cohen’s $d=0.58$). This suggests that while the foreign language learners know the mechanism of boundedness marking, the structural and morphological differences between types of predicates may affect their ability to process functional morphology correctly on such a demanding on-line task.

Looking closer at predicate types within each condition, we see that the L2 group was significantly higher in accuracy on [+bounded] suffixed predicates 2D and 3F types in Condition 2 and Condition 3 than on the [+telic] predicates 2C and 3F ($t(10)=3.185; p=0.01$; Cohen’s $d=-1.39$ and $t(10)=-2.613; p=0.026$; Cohen’s $d=-0.93$, respectively). An interesting pattern emerges when we compare the performance of the L2 group on perfective and imperfective predicates: unlike the HL group, they perform statistically similarly on both types of the imperfective (no pair-wise difference between 1A, 2D, and 3F), but their scores on the [+telic; +bounded] accomplishments 2C in Condition 2 were significantly different from 1B in Condition 1 ($t(10)=3.325; p=0.008$; Cohen’s $d=0.88$) and also from [+telic; +bounded] achievements 3E in Condition 3 ($t(10)=2.947; p=0.015$; Cohen’s $d=0.48$). Based on the pair-wise comparisons of predicate types, boundedness contrasts in accomplishment seem to be less stable in the L2 group, as is the prefixation mechanism, which seems to be acquired in the telicity condition, but breaks down in the boundedness condition. Under the strenuous working memory load in the task, i.e., seeing only one word of the sentence at a time without the possibility of backtracking, the L2 participants may fail to parse the prefixed predicates as such due to working memory limitations and may fail to compute the correct aspectual interpretation of the sentence. There are several other possible explanations to this difference in patterns between the HL and the L2 groups, including small sample size of the L2 group in comparison with the HL group, task effects, and a bias towards accepting rather than rejecting sentences.

### 5.4. The three groups compared

If we compare the mean scores of the two test groups on each condition and then within each condition, we can say that both telicity and boundedness features have been acquired, although to a different degree of stability and that there is a consistent advantage of the HL participants over the L2 participants. After careful examination of Figures 1-2 and the results presented above, we can deduce the answers to the last two research questions posed: i.e., that heritage learners have some advantage over foreign language learners and that heritage learners sometimes but not always show patterns of sensitivity to aspectual contrasts that are similar to those of monolingual controls. What is interesting is that the HL and the L2 groups seem to diverge in the patterns of their accuracy: the HL group struggles most with telicity in activity-accomplishments (Condition 1), while the L2 group with the boundedness contrasts (Conditions 2 and 3). At the same time, it is in its most challenging condition that the HL group converges with the L2 group. A one-way ANOVA supports these conclusions: HL and L2 learners are statistically indistinguishable in Condition 1, but both groups differ significantly from the controls: $F(2,60)=4.6; p=0.014$. Bonferroni Post Hoc showed L1 scores are statistically higher than those of the L2 group ($p<0.05$; Cohen’s $d=1.10$) and than those of the HL group ($p=0.046$; Cohen’s $d=0.89$). And it is in the boundedness conditions that the HL group converges with the monolingual controls, both differing significantly from the L2 group in Condition 3 ($F(2,60)=7.46; p=0.001$; Cohen’s $d=1.68$ for the L1-L2 pair and Cohen’s $d=1.54$ for the HL-L2 pair). The L2 group lagged particularly behind on the morphologically complex Condition 2 ($F(2,60)=21.77; p<0.0001$; Cohen’s $d=4.25$ for the L1-L2 pair and Cohen’s $d=3.50$ for the HL-L2 pair). Based on the results of this task, it appears that in on-line processing, the regular SI morphology used to mark the grammatical aspect is easier for the two test groups than the derivational prefixation mechanism used to mark lexical aspect.

Before moving to the conclusion, it is worth addressing why there was such a difference in the scores of foreign language learners on the predicates 2C and 1B, considering they contained the same type of predicate. One explanation could be the set-up of the task, which was based on the manipulation of telicity and boundedness contrasts. Some of the 1B predicates appeared with plausible adverbials and others with implausible adverbials, so the participants had to both reject and accept those sentences. In contrast, all 2C and 3E predicates appeared with an implausible adverbial and their
imperfective counterparts with a plausible adverbial – so, the participants needed to reject all the 2C and 3E items. However, this set-up affected neither the control group, nor the HL group (Table 4). Also, there is a possibility that in a task that involves accepting/rejecting an item, a bias towards either accepting or rejecting sentences could be a potential confounding factor. That is, the L2 group may have failed to reject implausible sentences in Conditions 2 and 3 because this type of learner is known to be more likely to accept sentences they are not sure about rather than reject them. In a set-up like this one, 50% of the items were plausible and the other 50% implausible, and a 50%/50% acceptance/rejection rate would indicate that the participants were not biased one way or the other, regardless of the accuracy of judgment. Table 4 illustrates that the foreign language learners departed from the ideal, especially in aspect related sentences.

Table 4. Acceptance ratio in the SMS task (% accepted sentences, regardless of accuracy)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>The whole task, including fillers</th>
<th></th>
<th>Target items only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>L1</td>
<td>30</td>
<td>49.85 %</td>
<td>2.91</td>
<td>50.79 %</td>
</tr>
<tr>
<td>HL</td>
<td>22</td>
<td>50.25 %</td>
<td>3.88</td>
<td>50.65 %</td>
</tr>
<tr>
<td>L2</td>
<td>11</td>
<td>57.58 %</td>
<td>9.25</td>
<td>64.29 %</td>
</tr>
</tbody>
</table>

According to the results of a one-way ANOVA and Bonferroni Post Hoc Test, the control group and the HL group were indistinguishable and showed no bias either way. In contrast, the L2 group, in general, accepted significantly more sentences ($F(2,60)=11.033; \ p<0.0001$) than the L1 group ($p<0.0001; \text{Cohen’s } d=2.66$) and than the HL group ($p<0.0001; \text{Cohen’s } d=2.52$). In addition, the L2 group was especially biased toward accepting the target sentences, i.e., those manipulating aspectual knowledge, which made this group stand apart from the two other groups ($F(2,60)=9.545; \ p<0.0001$). The difference was statistical between the L2 and L1 groups ($p<0.0001; \text{Cohen’s } d=2.66$) and the L2 and HL groups ($p<0.001; \text{Cohen’s } d=2.68$). These results could reflect the fact that the L2 group was also lower in general proficiency than the other two groups and, thus, less confident in judgments. However, the L2 group is also less stable in their aspectual judgments than the other two groups, which accepted and rejected target items at the same rate as the sentences in the whole task.

6. Conclusion

The results of the experiment support the existing assumption that aspect is difficult for HL and L2 learners; high accuracy in the proficiency measure was no guarantee for equally high proficiency in aspectual interpretations. Aspect is a complex category and some aspectual contrasts pose a greater difficulty than others for both test groups. Overall, morphology may be the bottleneck not only of L2 acquisition (Slabakova, 2008), but also, to some degree, for HL acquisition as well. Nevertheless, there seems to be some advantage in early bilingualism versus late bilingualism, which is selective rather than categorical (Montrul, 2008). What is also important is that despite patterning with the native speakers in so many ways, high proficiency heritage speakers also converged with the foreign language learners in their sensitivity to telicity contrasts, a feature of lexical aspect, and the associated morphological realization. This finding is consistent with Slabakova’s (2005) conclusion that L2 learners struggle with the lexical component of aspectual knowledge more so than with the syntax component and with Polinsky’s (2008) claim that a reduced lexicon and morphological repertoire may affect HL grammars.

References


