

ARTICLE

The Acquisition of Verbal Morphology by Child Classroom EFL Learners in Russia and China: The Effect of Age and L1

Athina Ntalli , Theodora Alexopoulou, Henriëtte Hendriks and Ianthi Maria Tsimpli

Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK

Corresponding author: Theodora Alexopoulou; Email: ta259@cam.ac.uk

(Received 05 January 2023; revised 24 January 2025; accepted 11 February 2025)

Abstract

We investigate the effects of age and first language (L1) on the acquisition of verb morphology in L2 English by Chinese and Russian children learning English as a foreign language in EFL schools in Shanghai and Moscow. We tested children 5 years after they started their EFL classes and considered two groups in each country: one group started EFL classes at the age of 4 and was tested at the age of 9, while the other group started at 7 and was tested at 12. We assessed the production of 3SG-agreement and past tense using two elicited production tasks (TEGI). Our results show that later starters consistently outperform earlier starters. Unexpectedly, Chinese children showed higher accuracy with 3SG-agreement than their Russian counterparts. Finally, learners were more accurate with regular past tense than 3SG-agreement.

Keywords: child L2 acquisition; morphosyntax; EFL; age effects; L1 effects

1. Introduction

Children all over the world start learning English as a foreign language (EFL) at increasingly younger ages with attendance at afternoon schools outside their regular/mainstream day schools being common in many countries. There is limited research into classroom EFL learning in these younger ages, in particular outside Western Europe, and, specifically, on the impact of an earlier start on young L2 classroom learners. Our goal in this study is to investigate how the age at which children start their classroom EFL lessons impacts their acquisition of verbal morphology and what is the role of their L1. Specifically, we focus on 3SG-agreement and the past tense and present an empirical study of classroom learners in Moscow and Shanghai.

Unlike previous work, this field study is conducted in two countries outside Europe and North America, namely China and Russia, aiming to contribute to our understanding of young learners' EFL learning in these contexts. Additionally, research on young EFL learners has primarily been conducted in mainstream day schools with designs focused on the national context of individual countries (e.g., García Mayo & Garcia Lecumberri, 2003;

Muñoz, 2006). In the current study, we present a cross-national design, exploiting the opportunities presented by international EFL schools, which run programs in various countries following a standardised curriculum and teaching approach. Such schools allow us to investigate young learners' EFL development across countries and abstract from the specific educational context of individual countries. In this study, we examine the acquisition of verbal morphology by children attending after-school EFL classes at the same school, following the same curriculum and pedagogical approach. Our study, therefore, complements previous research investigating EFL acquisition in a single national context and within mainstream education.

Morphology is particularly challenging for both adult second language (aL2) and child second language learners (cL2) in naturalistic settings (e.g., DeKeyser, 2005; Hawkins & Chan, 1997; Lardiere, 2007; Paradis, Tulpar, & Arppe, 2016) as well as instructed settings (e.g., Housen, 2002; Yang & Huang, 2004). Verbal morphology is a well-researched topic; yet, investigating it in young learners through a cross-national design can provide new insights into how age and L1 might impact child L2 acquisition in an EFL-instructed setting. Our cross-national design involves two typologically distinct L1s: Russian, a language with a rich paradigm of verbal agreement and tense morphology, and Chinese, an isolating language with no agreement or tense morphology. Children may thus face a different learning challenge since Russian learners are already familiar with marking agreement and tense features morphologically on the verb, while Chinese learners are not.

In addition to L1, we investigate the effect of age (of onset) on the acquisition of verbal morphology by comparing children starting at different ages but with matched years of instruction.

1.1. The age factor in child second/foreign language acquisition

Research on the role of age in the acquisition of L2 grammar in naturalistic settings shows that children who start their L2 acquisition early have an advantage in ultimate attainment compared to those who start later in childhood or after it (e.g., Hyltenstam & Abrahamsson, 2003; Johnson & Newport, 1989). For example, young children who immigrate to a new country and are immersed in the L2 at a very young age usually achieve native-like acquisition of the L2 grammar, whereas those who immigrate as adults often do not.

The advantage of an earlier start with L2 for ultimate attainment is well established for immersed learners who complete their primary and secondary education in their L2. However, the picture is more nuanced when we turn to classroom/FL learning studies, where learners are typically educated in their L1 and receive limited hours of FL learning classes as part of their mainstream curriculum or in FL schools. For instance, studies examining grammatical abilities in classroom settings and comparing earlier and later starters show that older starters are faster learners than younger ones (e.g., Cadierno *et al.*, 2020; Cadierno *et al.*, 2022: comparing children starting EFL classes when they were 7–8 and 9–10, respectively, testing them over three consecutive years from onset; Cenoz, 2003: testing children with different ages of onset at 4, 8, and 11 after 600 hours of instruction; Jaekel, Schurig, Florian, & Ritter, 2017: testing children with ages of onset at 6–7 and 8–9 firstly in Year 5, that is, when children were 10–11, and subsequently in Year 7, when children were 12–13; Muñoz, 2006: testing learners with different ages of onset at 8, 11, and 14, and in adulthood at different time points controlling for instruction time in English; Myles & Mitchell, 2012: comparing children starting French at the ages of 5, 8,

and 11, testing them three times: 9 weeks after they had been attending a course, 19 weeks, and two months later).

Older starters are thought to perform better due to cognitive maturity, the benefit of explicit instruction, and the ability to employ a wider range of learning strategies (e.g., Muñoz, 2006). Not only do older learners learn faster, but they also do not seem to be disadvantaged for ultimate attainment since the majority of FL classroom studies have failed to establish an ultimate attainment advantage for younger starters (e.g., Muñoz, 2006; Pfenninger & Singleton, 2016: children tested at the beginning and end of secondary school, with early starters beginning English classes in Grade 3 of primary education and later starters upon entering secondary school).

Some studies, however, have indicated an advantage for younger starters in other language domains. Jaekel, Schurig, and Ritter (2022), assessing proficiency through reading and writing, found earlier learners who started at Grade 1 outperforming later learners starting at Grade 3 when tested in Grade 5. An early start advantage has also been found in phonemic discrimination for earlier starters (age of onset at 9 or later) compared to later starters (age of onset at 12–13) when all were tested as adults with an average age being 19.4 (Larson-Hall, 2008). Finally, the same advantage has been shown for listening comprehension between early starters (who began English classes in elementary school) and late starters (who began English classes in secondary school) when tested as 18-year-olds (Pfenninger, 2014).

Most of the studies above have focused on children with an early age of onset, but typically after school entry age (6–7 years old). Few studies have considered children who start learning younger and before starting school. There are hypotheses suggesting that for morphosyntax, age effects begin at an even younger age (e.g., Meisel, 2009). Hence, more research is needed to determine whether an earlier age of onset in EFL learning provides any advantage for grammar acquisition.

1.2. *The L1 factor in child second language acquisition*

The age factor is intertwined with the L1; L2 children have already established their L1 when they start acquiring their L2, and hence the L1 may influence the acquisition of the L2. L1 effects are well-attested in various domains of child L2 (cL2). For example, Haznedar (1997) as well as Mobaraki, Vainikka, and Young-Scholten (2008) found word order transfer from the children's L1 to their L2 at the initial stages of acquisition. Paradis (2011), testing children on the acquisition of verb morphology in English and with a mean exposure of 20 months, reports higher accuracy with past tense and agreement by children whose L1 marks tense morpho-phonologically compared to those whose L1 does not. In another study, Paradis et al. (2016) examining Chinese children's accuracy in verb morphology after 6 years of immersion, found that not all children reached native-like levels of accuracy for one or more verb morphemes and pointed to the possibility of persistent L1 effects even for children with an early age of onset. Yang and Huang (2004), focusing on instructed learners of English with L1 Cantonese, attributed the learners' rather low performance on verb morphology to L1 influence due to the absence of tense and agreement morphology in Cantonese. Finally, Rocca (2007), testing Italian children learning English, found that they frequently used stative progressives that she interpreted as transfer of a prototypical feature of the Italian imperfective.

In sum, evidence for L1 effects on verb morphology, especially after the initial stages of learning (for a study on the initial stages of development, see Paradis, Rice, Crago, &

Marquis, 2008), is robust, as shown above. Still there is some unclarity on the effects of L1 in an EFL context of limited input.

1.3. Other factors that play a role in child L2 acquisition

Various other child-internal factors have been found to influence the L2 acquisition of grammar. Roehr-Brackin and Tellier (2019) showed that analytic-language ability predicts achievement in L2 French grammar for English-speaking children learning French in classrooms. They further highlighted that analytic language ability and, more broadly, aptitude are dynamic in children under 12 years of age. Motivation (e.g., Djigunović & Nikolov, 2019; Skehan, 2002), attitudes towards English L2 learning (Myles, 2022), personality characteristics such as anxiety, extroversion/introversion, risk-taking, and self-esteem (e.g., Fenyvesi, Hansen, & Cadierno, 2020), as well as gender (e.g., Cadierno *et al.*, 2022), have also been shown to impact L2 achievement.

Turning to contextual factors influencing L2 achievement in an FL setting, use of and exposure to English have been shown to be consistent predictors (Tragant & Muñoz, 2023). One of the strongest predictors of L2 achievement is the amount of out-of-school exposure to the language through media (e.g., Lindgren & Muñoz, 2012). Other factors playing a role include the L1 literacy (e.g., Pfenninger & Singleton, 2016), the educational background of learners' parents (e.g., Butler & Le, 2018), the L2 proficiency of parents (Muñoz & Lindgren, 2011), the cultural capital learners live in, and the status of the L2 language in society.

1.4. Verb morphology in English, Russian and Chinese and L2 learning

We chose Chinese and Russian as two languages that vary in different ways from English in relation to the features we are examining: Russian is broadly similar to English while Chinese is typologically quite distant from English. We capture cross-linguistic variation through generative parametric variation, meaning variation in the features associated with the functional category of Tense. Tense is a functional category encoding features of tense and agreement. Simplifying somewhat, tense features in English are realised through the auxiliaries *be* and *have* and bound morphology *-ed* in the past.¹ Agreement is morphologically expressed through the third person singular in the present through an affixed *-s* (and on the auxiliaries *be* (*is/are/am*) and *have* (*have/has*)).

Features may be interpretable or uninterpretable (Chomsky, 1995: 277). The interpretable features are associated with semantic features at the Logical Form (LF)² (e.g. past tense interpretation in “she studied”) while uninterpretable features are formal features void of semantic meaning (e.g., 3SG agreement in “she studies”). Thus, tense is an interpretable feature while agreement *-s* is an uninterpretable one. This distinction is central to the Interpretability Hypothesis (Tsimplici & Dimitrakopoulou, 2007), which suggests that L2 learners find uninterpretable features more challenging than interpretable ones. According to this hypothesis, only semantic/LF-related features are accessible to L2 learners while formal/uninterpretable features not realised in the L1 are not

¹These items also encode aspectual features like perfect (*-ed, have*), progressive (*be*), and so forth, potentially involving additional functional projections.

²Logical form as a distinct level of representation containing the syntactic information that is relevant for semantic interpretation.

available to L2 learners. The Interpretability Hypothesis predicts that, other things being equal, tense features will be more easily acquired than 3SG-agreement *-s*. Adding the effect of the L1–L2 typological similarity, we can predict that Russian learners will find tense and agreement morphology less challenging than Chinese learners, since the corresponding (abstract) tense and agreement features are part of the L1 Russian grammar, but not part of the L1 Chinese grammar.

Let us consider the respective phenomena under examination in the three languages in more detail. English marks person agreement only on third person singular in present (e.g., *She play-s basketball.*) and marks verbs for past tense (e.g., *She climb-ed the ladder. / He made a house.*).

We here consider Mandarin, Cantonese, and Shanghainese and refer to all three as Chinese, as they are all typologically similar with respect to the phenomena under focus; specifically, they are all isolating, they do not encode tense and agreement morphologically, and they may mark aspect through (mainly) unbound particles (Li & Thompson, 1981).

Example in Chinese with generic/habitual interpretation:

1. *Laoshi jiao xuesheng.*
Teacher teach student.
“A teacher teaches students.”

Example in Chinese with past interpretation:

2. *Ta shua le liba.*
PR3SG. paint LE-PERF ASP fence.
“He painted the fence.”

Example (1) has a plain verb with no marking. Such examples usually have a habitual interpretation corresponding to the simple present.

Example (2) involves a personal pronoun and the aspectual marker *le*. “Ta” has a third-person interpretation, thus encoding interpretable person features, but crucially, there is no third-person agreement morphology on the verb.

“Le” is an unbound perfective marker usually found with past events but also used for future events. To indicate time reference, Chinese uses temporal adverbs, while discourse, context, and world knowledge can also situate an event in time. Thus, Chinese differs from English in that it lacks grammatical tense manifested in affixal morphology (or an unbound morpheme); further, as an isolating language, it lacks agreement.

Russian marks verbs with morphological affixes for both tense and agreement (Mezheвич, 2008); both are instantiated in the present and simple past as examples 3 and 4 show.

Example in Russian with generic interpretation:

3. *Stomatolog lech-it zuby.*
Dentist cure-PRESENT.3SG teeth.
“A dentist cures teeth.”

Example in Russian with past and perfective interpretation:

4. *On po-krasi-l zabor.*
He PERF-paint-PAST.MASC.SG fence.
“He painted the fence.”

In example (3), the present form of the verb is inflected for third person singular.

In Example (4), the prefix *po-* encodes perfective aspect while the affix *-l* conflates tense, gender, and number features. So, in the past (perfective), the verbal form agrees in gender and number with the subject.

To summarise, English and Russian realise both grammatical tense and agreement as morphological affixes on the verb while Chinese does not. Thus, Chinese learners need to acquire tense and agreement features which are absent from their L1; Russian learners can rely on the tense and agreement features in their L1 to acquire tense and agreement verbal morphology in English.

2. The present study

Our study investigates the role of age (of onset) and L1 on the acquisition of verbal morphology by L2 child classroom learners in an instructional FL classroom environment. Specifically, we consider the following research questions:

Research Question 1: Will the starting age of learning EFL play a role in the acquisition of verbal morphology? Will younger starters have an advantage in verbal morphology? Or will older learners outperform younger ones as much as previous research has shown, albeit in different contexts and with overall older child learners?

Research Question 2: Can we detect L1 influence? Will the presence of tense and agreement features in L1 Russian facilitate the acquisition of verbal morphology for Russian-speaking children in contrast to Chinese-speaking children who lack tense and agreement morphology in their L1? Will Russian learners achieve higher accuracy as a result of the L1 influence?

Research Question 3: Are uninterpretable features (i.e., 3SG-agreement) more challenging than interpretable (i.e., past tense) ones?

Previous studies suggest an advantage for older children regarding the first research question. Yet, no study has considered classroom learners in EFL schools complementing mainstream education. Further, unlike most of the previous studies, we consider learners with a very early age of onset (AoO: 4 and 7) to establish whether a potential advantage would manifest with an earlier AoO, at 4 in our study in comparison to learners with older ages of onset in previous studies. We further anticipate that Russian learners, whose language instantiates tense and agreement features, will outperform Chinese learners, whose language lacks these features. Finally, following the Interpretability Hypothesis, we expect that all learners will find past tense features easier to acquire than 3SG-agreement.

3. Method

3.1. Study design and rationale

To investigate the role of age, we tested 9- and 12-year-old children starting their EFL learning at 4 and 7, respectively, thus having had 5 years of instruction in the afternoon EFL schools we recruited them from. We compared Chinese and Russian learners and the acquisition of tense marking with 3SG-agreement. We present our method in detail in the following sections.

3.2. Instructional setting

The data collection was carried out in EF (Education First/English First) schools in Shanghai and Moscow which are private afternoon schools teaching EFL. These schools are complementary to mainstream day schools that children attended.

A standardised curriculum was followed in both countries, supported by books and online resources produced by EF. There was some variation in the number of teaching hours offered in each country. Thus, EF in Shanghai offered 2 academic hours (1 hour = 40'-45') per week for all ages, while EF schools in Moscow gradually increased weekly teaching hours with age, offering 3 academic hours (1 hour = 40'-45') to children 3–6, raising to 4 hours to children 7–9, and to 5 hours to children older than 9.

EF employed native speakers of English or qualified non-native speakers and provided in-house training. In Shanghai, EF teachers were (mostly) native speakers of English while in Russia, most teachers were Russian native speakers. All teachers followed the same EF internal curriculum and used the same materials (e.g., book series, online/interactive activities).

Teaching followed mostly a communicative approach blended with some behaviourist elements (e.g., drilling, rewards), especially for younger learners. There was more focus on form for older learners, although the approach remained communicative. At more advanced ages, drilling was abandoned. It is important to highlight that different book series were used for different ages so that children starting at 4 and children starting at 7 would not use the same books. The book series is different in content and teaching practice, as we will see below.

In our classroom observations, we saw that the lessons in both countries were highly interactive, and learners were engaged. The teachers would encourage participation, and learners generally seemed motivated (which was also confirmed through our parental questionnaires). The learning atmosphere was positive and relaxing, and students seemed comfortable.

Overall, the instructional settings in EF in the two countries were fairly similar, following a standardised curriculum, teaching resources, and similar approaches to teacher training and classroom pedagogy.

3.3. Teaching of the grammatical phenomena under focus in EF language schools

A child starting to learn English at EF at 3 or 4 would have a prolonged A0 CEFR level which lasted up to the age of 6. At this level only very basic structures were introduced. Simple present and present continuous were introduced with the former expected to be mastered at the ages of 7–8. The simple past was introduced later than the simple present at the ages of 7–8 to be mastered at the ages of 8–9. This means that children starting English at 4 would be exposed to the simple past at around the same age with children starting at 6. The key difference was that the younger groups would have earlier exposure to the language, including the present tense with the past tense introduced later.

3.4. Educational context in China and Russia

At the time of the data collection, in Shanghai, literacy in Mandarin officially started at age 6 (primary school one) in mainstream day schools. Preschool education started at age 3 up to 6 and was not obligatory, although the enrolment rate in urban areas was very high (98–99%) (Zhou, 2011). In preschool education, children were not explicitly taught academic skills; teachers followed informal literacy practices (China Preschool Education Research Society, 1999 as referenced in Li, 2013; Li, Corrie, & Wong, 2008).

In Moscow, L1 literacy also started when children entered state day schools around the ages of 6.5–8 (Bodrova & Yudina, 2018). Preschool education was not compulsory, but

similarly to Shanghai, the vast majority of parents enrolled their children in preschool education which combined childcare and education (Bodrova & Yudina, 2018). Thus, in Shanghai, children were introduced to literacy slightly earlier than in Moscow, something that was also confirmed through responses to our parental questionnaires. According to Chinese parents, on average, their child started to learn to read and write in Chinese at age 4, while according to Russian parents, the respective average age reported was 5.

Turning to cultural attitudes towards schooling, in Shanghai, there was a significant tendency towards early learning and achievement. As Sun, de Bot, and Steinkrauss (2014) argue, “Chinese parents’ mindset is along the lines of the earlier, the better” and “we must not lose at the starting line” (p. 2). Education was of the outmost importance and a top priority for Chinese parents who had very high demands and expectations for their children (Hu & Szente, 2009). The “one-child policy” held until recently in China had meant that parents paid even greater attention to their child’s academic achievements (Hu & Szente, 2009) which were viewed as the determinant for their social upgrade in a very competitive market. It is not surprising then that in China, private English language schools, but also early learning centres aiming to boost children’s cognitive and general learning development, have been proliferating. By the time they reach 15 years of age, learners quit private language schools to focus on preparations for the “Gaokao” exam, the national exam determining their entrance to university.

In Moscow, children generally started English around one year later (not only privately but in day schools as well, as we shall see below) and continued attending classes at EF up to the age of 17 or so. Unfortunately, there is not much literature on parents’ attitudes towards education (at least in the English language). However, based on personal discussions with the EF teachers in Moscow and the children’s parents during the data collection, it seemed that Russian parents also value education and an early start.

3.5. Participants

L1-Chinese-speaking children

A total of 73 child participants were tested during the data collection in Shanghai. All children were native speakers of Chinese. They formed two groups: the first group involved 39 nine-year-olds (20 girls and 19 boys), and the second group involved 34 twelve-year-olds (16 girls and 18 boys). The younger group had a mean age of 9;7,³ and the older group had a mean age of 12;6. Children were not exposed to English at home, as the home language as well as the parents’ L1 was mostly Chinese, including Mandarin and Cantonese; a few parents also reported Shanghaiese as the home language which like Mandarin and Cantonese, does not instantiate tense and agreement. Children came from families of a high socioeconomic status and highly educated parents. The vast majority of the parents (124/144) had a university degree (>85%); the remaining few (20/144) had professional training or low/upper secondary education.

L1-Russian-speaking children

A total of 74 child learners were recruited at EF schools in Moscow. As with the Chinese children, two groups were formed, 32 nine-year-olds (14 girls and 18 boys) and 42 twelve-year-olds (19 girls and 23 boys). The younger group had an average age of 9;8 and the older

³Participants were pooled through filtering out their year of birth on a database of EF, so any age range is within 12 months.

ones 12;6. The children's L1 and home language was always Russian. Russian was the mother tongue of all parents except for one father whose first language was Armenian. Children came from families of high socio-economic status. The vast majority of the parents (> 90%) had postgraduate education while only a few had professional training (7/147), and one reported elementary school while his/her partner had a postgraduate qualification.

Table 1 shows an overview of the groups of participants.

3.6. Length of exposure and input

We operationalised length of exposure as the number of years of instruction at EF schools. In addition, we considered the academic hours of attendance for each child in EF since, as we saw earlier, for Russian children teaching hours increased with proficiency while in China the hours of classes remained constant for all levels. Further, Shanghai schools offered extra mini-intensive courses over holidays which were not available in Moscow. Table 2 summarises the mean academic hours for Russian and Chinese age groups. As can be seen, the younger group in Moscow had fewer hours of attendance than their counterparts in Shanghai. The gap closes for the older groups.

We also need to consider English classes in day schools. In Shanghai, children started English at the age of 6 (primary 1) and received around 4 hours of instruction per week on average (Tan, 2012). In Moscow, children started learning English at day schools at 8 years of age and received 3 academic hours per week. Hence, our Russian participants had received less overall teaching input in English than their Chinese counterparts at the time of data collection. The Russian nine-year-olds would have attended one year of English classes at their day school in comparison to 3 years attended by the Chinese nine-year-olds; in addition, Chinese nine-year-olds would have had more teaching hours per week overall in their day schools. These discrepancies are less pronounced for the twelve-year-olds; the Russian twelve-year-olds would have attended 4 years of English day school instruction in comparison to 6 years attended by the Chinese twelve-year-olds.

3.7. Proficiency

To assess the learners' proficiency, we utilised the CEFR (Common European Framework for Reference for Languages, Council of Europe, 2001) levels of EF classes attended at the

Table 1. Overview of participants

	Chinese		Russian	
	9-year-olds	12-year-olds	9-year-olds	12-year-olds
N	39	34	32	42
Mean age	9;7	12;6	9;8	12;6

Table 2. Mean academic hours and ranges of attendance in EF by both Chinese and Russian children

Academic hour = 40 mins	CH_9y.o.	CH_12y.o.	RU_9y.o.	RU_12y.o.
Academic hours in EF	880 [384–1455.5]	821 [546–1065]	592 [242–697.5]	804 [405–921.5]

time of testing, based on teachers' assessments. Additionally, we administered the Word Finding Vocabulary task (Renfrew, 1995).

3.8. Class CEFR level

Table 3 shows the proficiency levels of the class children attended at the time of testing. The majority of the twelve-year-olds have reached B1 in both nationality groups which are comparable in proficiency. However, there is variation in the proficiency of the younger groups. In both nationality groups, a small number of nine-year-olds (5) have reached B1 with the rest in A. However, in the Russian group, the remaining children are split into A1 and A2 (14 and 13, respectively), while in the Chinese group, the vast majority (29) are A2; thus, more than twice Chinese year-olds are in A2 than Russian ones, which means that the Russian group is of lower overall proficiency.

3.9. Word finding vocabulary task

The Word Finding Vocabulary Task (Renfrew, 1995) captures children's production of vocabulary. Children were asked to name the picture they saw on a computer screen (e.g., a cup). Each picture was presented individually on a PowerPoint presentation. There were in total 50 test items. Target words as specified in the task were scored with 1, and any other answer was scored with 0.

Table 4 summarises the mean vocabulary scores and ranges across groups. As can be seen, the group of Chinese nine-year-olds had a mean score of 17/50, while the group of Chinese twelve-year-olds had a 19.5/50 average score. As for the Russian children, the group of the nine-year-olds produced, on average, 18 correct words out of the 50, with the group of twelve-year-olds scoring 24.5/50.

Table 3. Proficiency levels corresponding to the class children attended at time of testing

	Chinese-speaking children	Russian-speaking children
9-year-olds	5: A1 29: A2 5: B1	14: A1 13: A2 5: B1
12-year-olds	31: B1 3: B2	39: B1 3: B2

Table 4. Percentages of correct responses, average scores (standard deviations) and their ranges in brackets concerning children's scores on the renfrew word finding vocabulary task

	Chinese-speaking children	Russian-speaking children
9-year-olds	34% 17/50 (4.5) [9–30]	36% 18/50 (5.3) [5–27]
12-year-olds	39% 19.5/50 (5.3) [11–33]	49% 24.5/50 (4.8) [17–33]

Table 5. Percentages of correct responses, average scores (standard deviations) and their ranges in brackets concerning children's scores on the renfrew word finding vocabulary task

	Chinese-speaking children	Russian-speaking children
9-year-olds	39% 14.1/36 (3.4) [8–22]	32% 11.5/36 (4.1) [3–17]
12-year-olds	43% 15.4/36 (3.8) [10–26]	42% 15.1/36 (3.0) [11–21]

Despite their lower proficiency, Russian nine-year-olds had vocabulary scores as high as the Chinese nine-year-olds. This discrepancy between CEFR level and vocabulary scores could be explained by the relatively high proportion of Russian–English cognates in the test. There were 14 cognate words out of a total sample of 50 (9 kloun/clown, 10 krokodil/crocodile, 12 kenguru/kangaroo, 19 gitara/guitar, 25 drel/drill, 32 binokl/binoculars, 36 parashut/parachute, 37 magnit/magnet, 40 iglu/igloo, 42 mikrofon/microphone, 43 sedlo/saddle, 46 raketka/raquet, 48 kompas/compass, 49 termometr/thermometer). Note that microphone (mai ke feng) and guitar (ji ta) may also be considered cognates in the Chinese language.

We removed the cognates from the word list and recalculated the vocabulary scores without them. Table 5 presents how the scores change. We can now see that the vocabulary scores correlate with the CEFR class levels. Younger Russians had lower vocabulary scores than their Chinese counterparts which corresponds to their lower CEFR proficiency. Older groups seem to have the same vocabulary size which corresponds to their comparable CEFR levels. It is noticeable, though, that there is only one point of difference in the average vocabulary score between Chinese nine-year-olds and twelve-year-olds, perhaps meaning that the task may not capture that well progression from A2 to B1.

ANOVA analysis showed that there was a significant difference between the group means, $F(3, 143) = 8.37, p < 0.001$. Post-hoc tests using the Bonferroni method showed that only the Russian nine-year-olds had significantly lower scores on the Renfrew task than all the other groups (RU9 – CH9: $p = 0.014$, RU9 – CH12: $p < 0.001$, RU9–RU12: $p < 0.001$). There were no significant differences observed among the other groups (CH9, CH12, RU12) in their Renfrew task scores.

For statistical analyses in subsequent sections, we consider only the vocabulary scores without the cognates.

3.10. Experimental materials and questionnaires

Test of early grammatical impairment (TEGI)

To evaluate the influence of age and L1 on acquisition, we investigated the accuracy in the use of L2 English verbal morphology. Specifically, we used the TEGI (Rice & Wexler, 2001), a battery of elicited production tasks which includes production probes for third person singular (3SG) and past tense (regular and irregular). (See for TEGI studies in cL2: Chondrogianni & Marinis, 2011; Paradis et al., 2016.)

The TEGI part capturing the 3SG-agreement consists of 10 test items following a practice item. The child would see a picture on a laptop screen depicting a person (e.g., a

teacher, a dad), and the experimenter asked the child what each person does. For example, in the case of a picture depicting a teacher, the experimenter would say, “Here is a teacher. Tell me what a teacher does.” The use of the indefinite article seeks to elicit a generic description of what teachers do, with the target answer being “A teacher teaches.” Although the children were shown pictures, they were explicitly asked not to describe the people in the pictures.⁴

We followed the TEGI guidelines with one exception: when the child gave a wrong answer, the experimenter did not provide the correct answer. This was to avoid the possibility of favouring older children who are better test-takers and may have had more practice than younger children with these types of tasks. However, the experimenter did correct if the child used the present continuous. For example, if the child said, “A teacher is teaching.” the experimenter would say, “Do not tell me what this teacher (pointing to the picture) is doing. Tell me what a teacher does. Any teacher.” or “Do not describe the picture. Tell me what any teacher does.”

The TEGI probe for past tense included 2 trial items and 18 test items: 10 targeting regular verbs and 8 targeting irregular verbs. Children were shown two pictures on a computer screen. The experimenter described the first picture and asked the child to describe the second one. For example, the experimenter said, “Here the boy is painting the fence (pointing at the first picture). Now he is done (pointing at the second picture). Tell me what he did. He ...” Target answer: *He painted (the fence).*

As with the previous test, we followed the procedure indicated by the TEGI manual except for providing the correct answer in the example item when children did not provide the target.

3.11. Language background questionnaire

To collect information about children’s exposure to English and use outside their EF schools, we designed and administered a language background questionnaire. Questions concerned the AoO of learning English, for example at another school, parental L1, English language use at home, motivation, use of media in English, etc. (See [Supplementary Material](#) for details on the language background questionnaire).

3.12. Procedure

Our study was given approval from the Ethics Committee of our university’s School of the Humanities and Social Sciences. We administered parent information sheets through EF schools and obtained parental consent before testing the children. We tested children in a dedicated classroom in their EF schools. Each child had a one-to-one session with the experimenter, who always asked each child for oral consent before the testing. The whole process was recorded using a professional recording device. The testing session lasted on average 45 minutes with TEGI tasks lasting approximately 10’-15’ (we administered more tests not discussed in the present paper). Data were later transcribed orthographically by the experimenter.

⁴It is unclear if the children are able to interpret the indefinite article correctly since neither Russian, nor Chinese has a definite article.

3.13. Data coding and scoring of the probes

3.13.1. Coding and scoring of the TEGI 3SG

Correct answers were considered those which had a third-person singular subject and a verb marked with 3SG *-s* (e.g., “A father/He plays with his children.”) and were scored with 1. We scored with 0: *i.* the cases when the child used a third person singular subject and did not mark the verb with 3SG *-s* (e.g., **she play*), *ii.* the instances of periphrastic marking (e.g., **He is play, she is playing, she can play; he does play*), *iii.* Any other response using another tense form (e.g., simple past), plural subject, progressive participle *-ing*, *does + verb-ing* or noun, copula “be,” and *iv.* “No responses.”

3.14. Coding and scoring of the TEGI PAST

Correct answers were those inflected with *-ed* (e.g., *She cleaned*) or the correct suppliance of the irregular form (e.g., *He made*) and were scored with 1. We also scored with 1 answer using a lexical verb other than the targeted one but correctly inflected (e.g., *tidied* instead of target *cleaned*). Similarly, for irregular verbs (e.g., *built* instead of *made*). Instances of overregularization, e.g., **rided* instead of *rode*, were coded separately and were scored with 1. All other answers were scored with 0: omission of inflection (e.g., *She clean-* or *He make*), instances of periphrastic marking (e.g., “*The girl is/was clean her room. The boy is jumping.”), use of another tense, use of progressive participles as the main verb, instances of “did” + *verb-ing* or noun, “be” as a main verb, use of verbs such as *hit, put, run* for which we cannot be sure whether they were marked for past tense, use of another type of verb (e.g., regular instead of irregular), and “no responses.”

3.15. Statistical modelling

We conducted mixed-effects logistic regression (GMELR) analyses in R (R Core Team, 2014) to answer statistically our research questions repeated below:

- a. whether there is an early age of onset advantage or an older age advantage as evidenced through accuracy in suppliance of agreement and past tense morphology,
- b. whether Russians use 3SG-agreement and past tense markings more accurately than Chinese, and
- c. whether 3SG-agreement is more challenging than past tense marking.

All categorical variables included as fixed effects were effect-coded (e.g., $-0.5, 0.5$) (Brehm & Alday, 2022). Continuous independent variables were z-score transformed. In all analyses, the level of significance was taken to be $p < .05$.

For the first two questions, we conducted three separate analyses, each considering accuracy in suppliance of 3SG-agreement, past regular tense or past irregular tense as the dependent variable in their respective models. The same procedure was followed for all three models. Each model would initially only include vocabulary scores and two variables serving as proxies for children’s input: hours of attendance at EF (formal instruction input) and use of media in English (out-of-school exposure). Subsequently, we introduced age (9: $-.5$, 12: $.5$) and first language (CH: $-.5$, RU: $.5$) variables into the model to assess their effects beyond differences in vocabulary and input. Finally, we included the interaction between age and first language. Model comparisons were performed using the *anova* function with criteria being the *p*-values and the model fit

indices (e.g., AIC). Significant improvements in model fit due to the addition of fixed effects or interactions guided the retention of these variables in the final models.

To decide on the variables to include in our models, we considered the variation within them. Hence, we excluded SES measured through parental education as the vast majority of the children came from parents with high SES. Motivation and classroom enjoyment were also not controlled for because, again, parental questionnaires showed little variation. Further, a few children reported they started English outside EF at an earlier age than 4, but these were only 2 Chinese nine-year-olds and 9 Chinese twelve-year-olds. By contrast, no Russian nine-year-old had any previous exposure while there were 5 Russian twelve-year-olds reporting an earlier start outside EF. Due to the small variation and the fact that the number of children who had started learning English outside EF do not differ too much across populations, we did not include this aspect in our models.

To make sure there were no collinearity issues, we carried out a Variance Inflation Factor (VIF) analysis in R and found that there was no significant multicollinearity among all variables. Hence, including all variables described above in our models is justified.

Turning to the random structure, this included a random intercept for participants and a random intercept for items in all three models.

Finally, to examine statistically whether 3SG-agreement is more difficult than past tense, we again carried out a mixed effects logistic regression to compare children's performance on 3SG-agreement and regular past tense. We excluded irregular past from this analysis since the irregular forms do not involve the application of a grammatical rule. Thus, the 3SG *-s* morpheme is compared with the *-ed* past morpheme. Considering again accuracy as the dependent variable, we included inflection (i.e., 3SG-agreement (coded as $-.5$) or past tense (coded as $.5$) as a fixed effect, and we specified a random intercept for participants and items, using $(1|part) + (1|item)$ to define the random structure.

4. Results

4.1. Binomial logistic regression for TEGI 3SG-agreement

Figure 1 presents the means and standard deviations (SDs) of raw scores for 3SG agreement accuracy by group. Three children of the Chinese 9-year-olds and two of the Chinese 12-year-olds failed to do the task because of lack of time or lack of understanding of the instructions (likely a result of low proficiency). Similarly, one Russian 9-year-old was excluded due to unintelligible responses. All 42 Russian 12-year-old participants did the task.

Results of the final model showed three significant main effects (cf. Table 6). The first significant main effect concerned age ($\beta = 2.920$ (SE = 0.571), $z = 5.113$, $p < 0.001$); older children were more accurate in their use of 3SG-agreement morphology than younger ones. The second significant main effect was the participants' first language ($\beta = -1.428$ (SE = 0.557), $z = -2.565$, $p = 0.010$) with Chinese children outperforming the Russian children. Finally, there was a significant main effect of vocabulary ($\beta = 0.997$ (SE = 0.244), $z = 4.095$, $p < 0.001$) with higher scores related to higher 3SG-agreement accuracy. The interaction between age and first language did not improve the model fit. Figure 2 illustrates the predicted probabilities of accuracy in the production of the 3SG-agreement morpheme across different age groups and language groups.

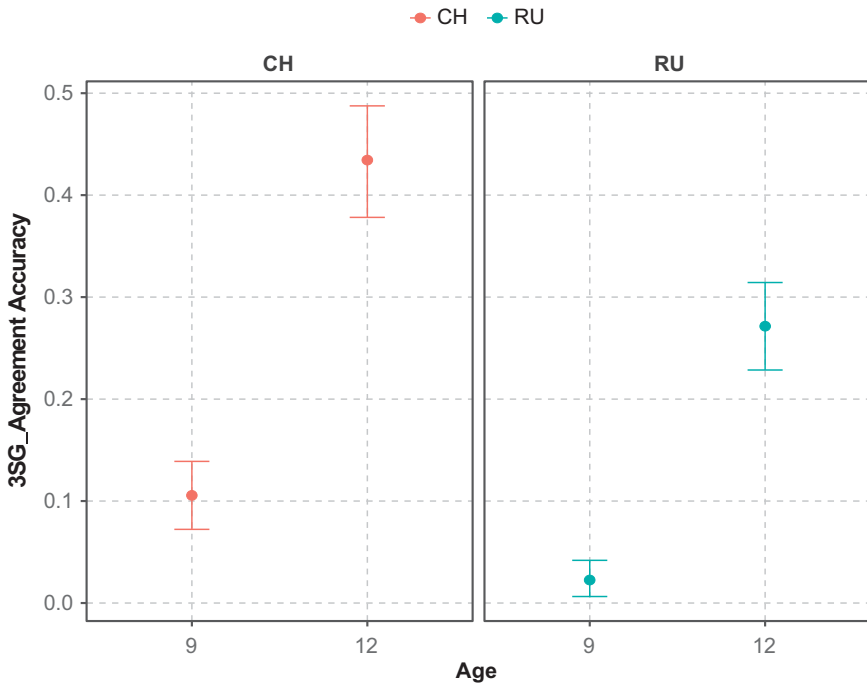


Figure 1. Mean accuracy with 3SG-agreement and SDs per L1 and age group.

Table 6. The optimal model predicts accuracy based on the vocabulary score, the hours of instruction, the use of media in English per week, the age, and the L1. The model also includes a random intercept for participant and a random intercept for item. (Optimal Model: $acc \sim zvoc + zhours + zmedia_use_min + age + first_lang + (1 | part) + (1 | item)$)

Fixed effects				
	Estimate	SE	z-value	<i>p</i>
Intercept	-3.096	0.344	-9.008	<0.001
Vocabulary	0.997	0.244	4.095	<0.001
Hours of instruction in EF	-0.271	0.255	-1.066	0.286
Use of media in English per week	0.133	0.229	0.580	0.562
Age	2.920	0.571	5.113	<0.001
L1	-1.428	0.557	-2.565	0.010
Random effects				
Groups	Name	Variance	Std. Dev.	
Participant	(Intercept)	4.656	2.158	
Item	(Intercept)	0.138	0.372	

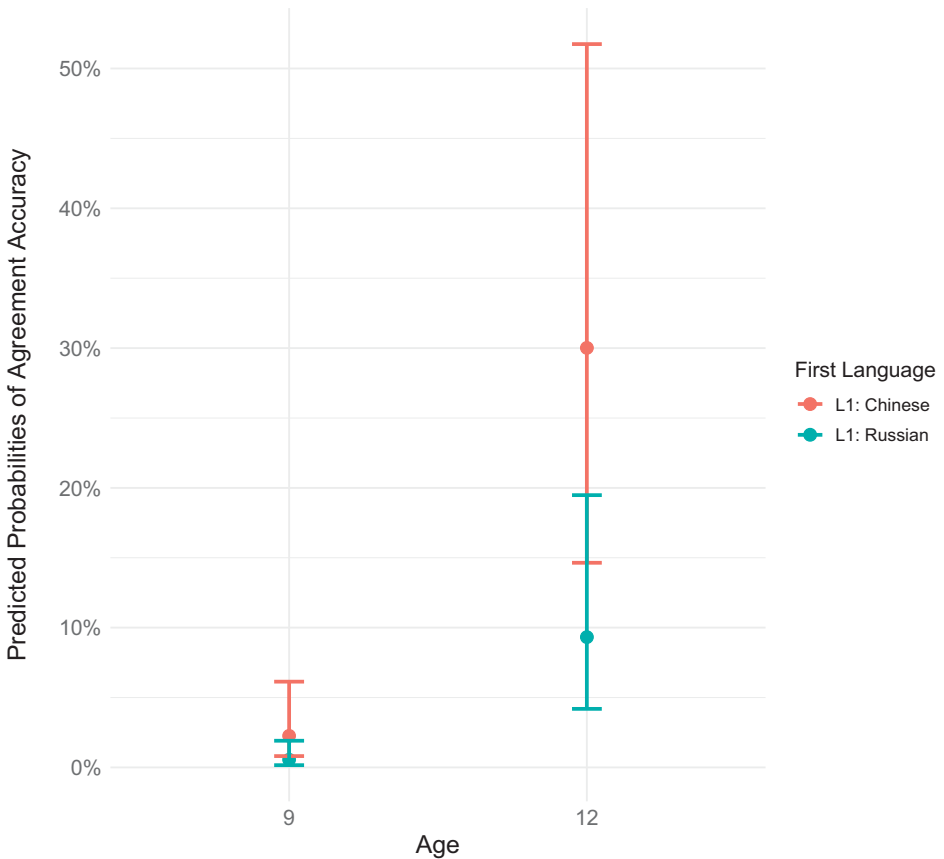


Figure 2. Predicted accuracy in 3SG-agreement after controlling for vocabulary, hours of instruction, and use of media. The point indicates the predicted average probability of a 3SG-agreement instance being produced accurately according to L1 and age. The whiskers indicate the 95% confidence interval.

4.2. Binomial logistic regression for TEGl past regular

Figure 3 presents means and SDs of raw scores for accuracy with regular past forms by group. Only 32 out of 39 Chinese 9-year-olds and 29 out of 32 Russian 9-year-olds completed the task. Additionally, two Chinese 12-year-olds did not complete the task, while all Russian 12-year-olds did. In cases where the task was not completed, this was due to time restrictions and/or difficulties with proficiency or comprehension of the instructions.

Results of the optimal model revealed two significant main effects (cf. Table 7). Specifically, there was a significant main effect of age ($\beta = 3.607$ (SE = 0.668), $z = 5.402$, $p < 0.001$), with older children being more accurate in their use of past tense morphology than younger ones, and a significant effect of vocabulary ($\beta = 1.213$ (SE = 0.277), $z = 4.380$, $p < 0.001$), with higher vocabulary scores related to higher accuracy with regular past tense. The interaction between age and first language did not improve the model fit. Figure 4 illustrates the predicted accuracy in the production of the regular past tense morpheme across different age groups and language groups.

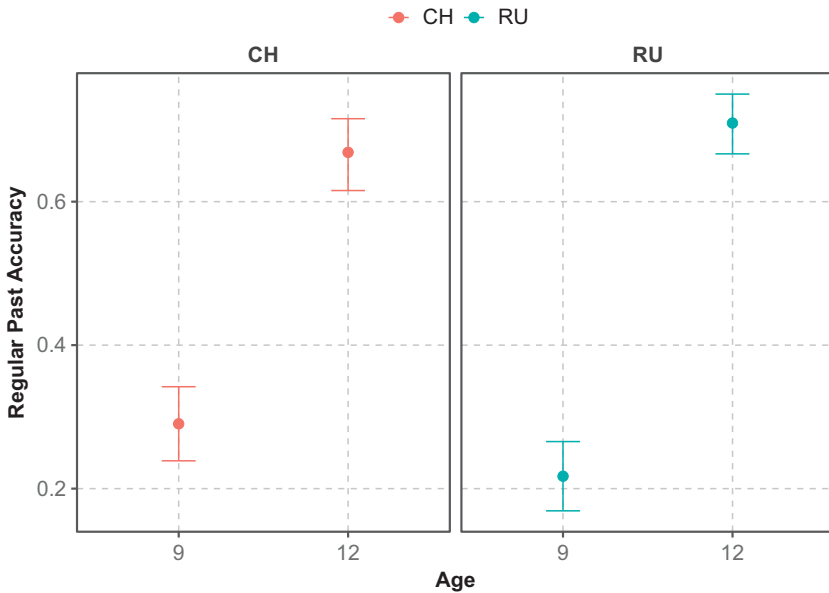


Figure 3. Mean accuracy with regular past and SDs per L1 and age group.

Table 7. The optimal model predicts accuracy based on the vocabulary score, the hours of instruction, the use of media in English per week, the age, and the L1. The model also includes a random intercept for participant and a random intercept for item. (Optimal Model: $\text{acc} \sim \text{zvoc} + \text{zhours} + \text{zmedia_use_min} + \text{age} + \text{first_lang} + (1 | \text{part}) + (1 | \text{item})$)

Fixed effects				
	Estimate	SE	z-value	p
Intercept	-0.559	0.338	-1.655	0.098
Vocabulary	1.213	0.277	4.380	<0.001
Hours of instruction in EF	0.003	0.290	0.010	0.992
Use of media in English per week	0.253	0.291	0.870	0.384
Age	3.607	0.668	5.402	<0.001
L1	0.003	0.651	0.005	0.996
Random effects				
Groups	Name	Variance	Std. Dev.	
Participant	(Intercept)	8.089	2.844	
Item	(Intercept)	0.230	0.480	

4.3. Binomial logistic regression for TEGI irregular past

Figure 5 displays the means and SDs of raw scores for irregular past accuracy by group. The number of participants is the same as given for the regular past above.

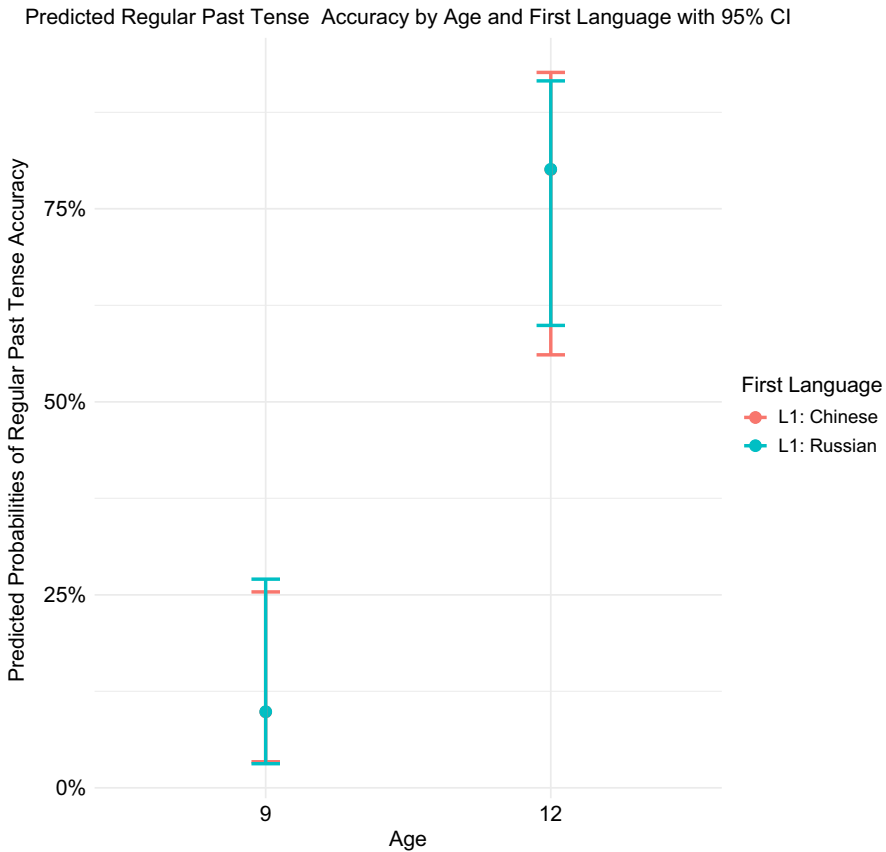


Figure 4. Predicted accuracy in regular past tense after controlling for vocabulary, hours of instruction, and use of media. The point indicates the predicted average probability of a past tense -ed morpheme being produced accurately according to L1 and age. The whiskers indicate the 95% confidence interval.

The final model demonstrated two significant main effects (cf. Table 8); a significant main effect of age ($\beta = 3.420$ (SE = 0.834), $z = 4.102$, $p < 0.001$) with older children outperforming younger ones, and a significant effect of vocabulary ($\beta = 1.792$ (SE = 0.447), $z = 4.012$, $p < 0.001$) with higher vocabulary scores related to higher accuracy with regular past tense. The interaction between age and first language did not improve the model fit. Figure 6 illustrates the predicted probabilities of accuracy of the production of the irregular past tense form across different age groups and first language groups.

4.4. Binomial logistic regression for the asymmetry in the acquisition of features

Figure 7 shows the mean accuracy for agreement and past tense across each age and L1 group.

The model showed a significant main effect of inflection ($\beta = 2.434$ (SE = 0.154), $z = 15.821$, $p < 0.001$) with accuracy being higher on past tense compared to 3SG-agreement (cf. Table 9). This holds across ages and L1s.

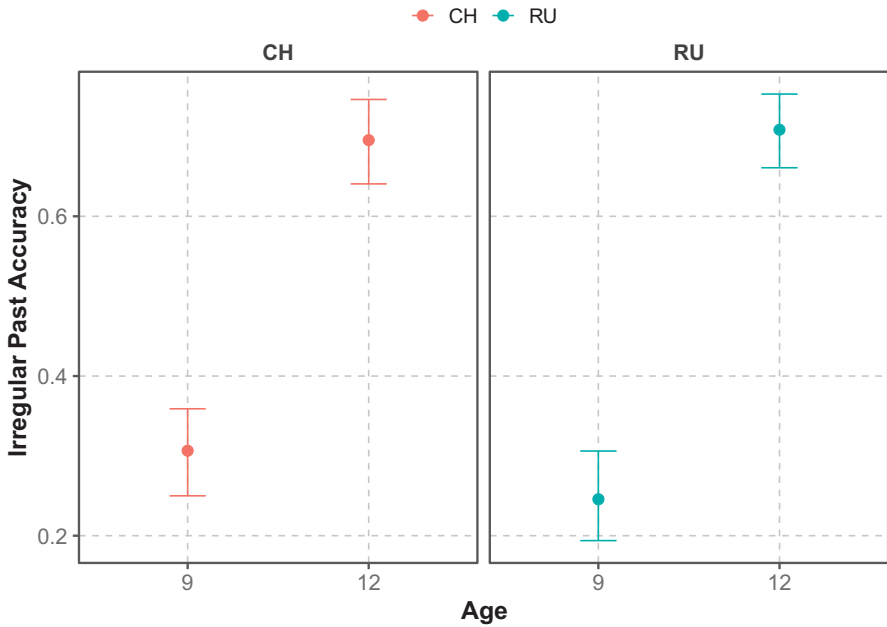


Figure 5. Mean accuracy with irregular past and SDs per L1 and age group.

Table 8. The optimal model predicts accuracy based on the vocabulary score, the hours of instruction, the use of media in English per week, the age, and the L1. The model also includes a random intercept for participant and a random intercept for item. (Optimal Model: $acc \sim zvoc + zhours + zmedia_use_min + age + first_lang + (1 | part) + (1 | item)$)

Fixed effects				
	Estimate	SE	z-value	p
Intercept	-0.406	0.380	-1.070	0.284
Vocabulary	1.792	0.447	4.012	<0.001
Hours of instruction in EF	0.093	0.387	0.240	0.810
Use of media in English per week	0.311	0.385	0.809	0.418
Age	3.420	0.834	4.102	<0.001
L1	0.316	0.831	0.381	0.703
Random effects				
Groups	Name	Variance	Std. Dev.	
Participant	(Intercept)	10.417	3.228	
	(Intercept)	0.111	0.333	

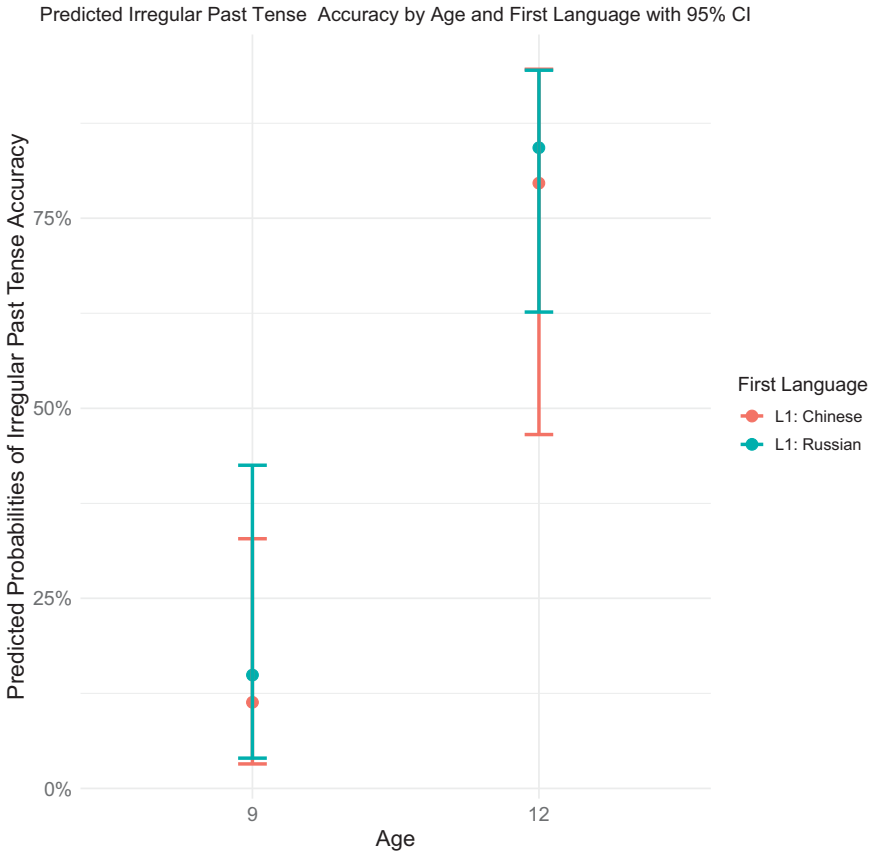


Figure 6. Predicted accuracy in irregular past tense after controlling for vocabulary, hours of instruction, and use of media. The point indicates the predicted average probability of a past tense -ed morpheme being produced accurately according to L1 and age. The whiskers indicate the 95% confidence interval.

5. Discussion

We begin our discussion considering the proficiency of learners before turning to our main questions.

5.1. Proficiency

Based on our research design, according to which all children would be 9 or 12 years old and would have had 5 years of instruction at EF, we expected that the two age groups across countries would be in the same proficiency levels. As we saw, Russian nine-year-olds were of lower proficiency than Chinese nine-year-olds, as almost half of the Russian nine-year-olds were in an A1 CEFR class level with only an eighth of the Chinese in that same level. Renfrew’s results excluding the cognate words, further confirmed the lower proficiency of Russian nine-year-olds. Older groups were of comparable proficiency with the vast majority in B1.

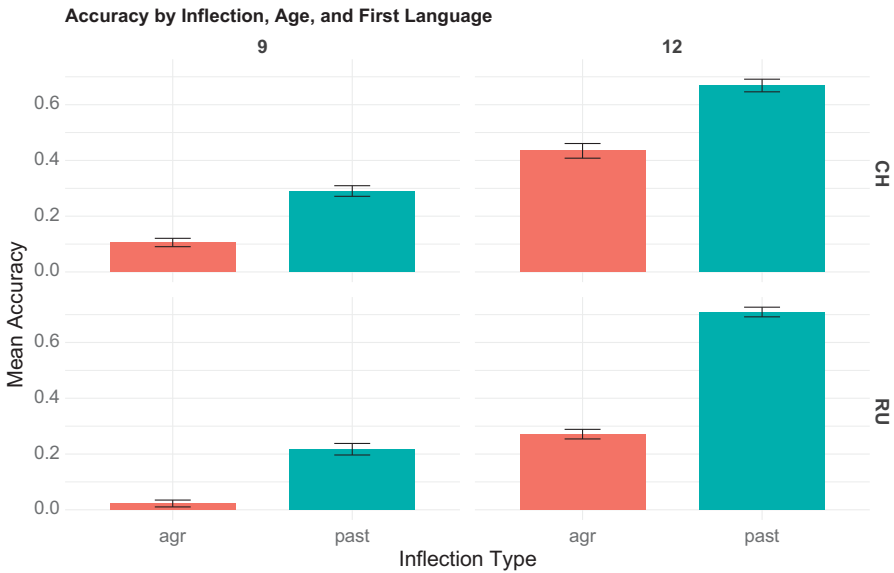


Figure 7. Mean accuracy for agreement vs. past tense for each age and L1 group.

Table 9. The model predicts accuracy based on the type of inflection and includes a random intercept for the participant and a random intercept for the item. (Model: $acc \sim inflection + (1 | part) + (1 | item)$)

Fixed effects				
	Estimate	SE	z-value	p
(Intercept)	-1.568	0.274	-5.731	<.001
Inflection	2.434	0.154	15.821	<.001
Random effects				
Groups	Name	Variance	Std. Dev.	
Participant	(Intercept)	7.562	2.750	
Item	(Intercept)	0.197	0.444	

The question is why the younger groups differ in proficiency despite the fact that they have been attending the same language schools for the same number of years and how it is explained that older groups did not differ.

Despite the similar settings in EF schools in Shanghai and Russia, there are still some important differences in the English input children receive in each country which might help explain the differences in proficiency for the younger groups. In Shanghai, children start learning English in their day schools earlier (at 6 as opposed to 8) and for more hours (4 on average vs. 3) compared to Russians. Another reason may be the age literacy in L1 starts; parental questionnaires show that on average, Chinese children started literacy one year earlier than Russian children, at around 4 years of age. Taken together, these

differences mean lower amounts of English instruction and later schooling for Russian children which may explain their lower proficiency. Additionally, parental attitudes and expectations may further compound these differences and indirectly lead to the observed proficiency difference between Russian and Chinese nine-year-olds.

It is unclear why these differences do not impact the older groups which show matching proficiency, as discussed below.

Older groups received comparable averages of EF academic hours (CH:821 vs. RU: 804) though with a higher average of hours in Chinese schools. Chinese children would also have had more input at their day school as they started English earlier and for more hours. We would expect that these differences in teaching input would perhaps lead to higher proficiency for the Chinese learners. It is striking that older groups are at the same EF CEFR class level, but Chinese seem to have reached that with many more hours of instruction coming from their day schools.

This cross-national study highlights the significance of wider educational factors in children's learning and progress with the L2. Specifically, schooling is key as countries differ in their policies of when English is introduced, the intensity of schooling, and so forth. Other factors might play a role as well, such as parental attitudes to early years learning and potentially the similarity between L1 and L2 which we cannot explore further in the present study. However, it is clear that while the emphasis of the Chinese education system and parents on a strong and intense early start does seem to confer an advantage in L2 English proficiency to younger Chinese learners, this advantage does not seem to extend to later childhood, by which time Russian children seem to achieve similar proficiency with Chinese children.

5.2. The impact of age (of onset) on verbal morphology in EFL learning

Older children were consistently more accurate in their use of the target verbal morphology than younger children, in line with previous research. Through our statistical analyses, we controlled for input differences, adding such variables as covariates in our models (formal instruction hours at EF and informal learning through use of multimedia in English). After controlling for input, we found a significant effect of age.

We need to note that there were both quantitative and qualitative aspects of input we could not control for. Older starters receive more and richer input. Younger starters have fewer teaching hours than older ones because older children have consistently attended English in their day school for the whole 5-year period, while nine-year-olds will have had English at their day school for 1 or 3 years depending on the country. Further, there were qualitative differences in the EF curriculum across different ages related to the content covered, the timing in the introduction of the phenomena in question, and the teaching approach (see the section "Teaching of the grammatical phenomena under focus in EF language schools"). Yet, it is clear that in this real-life context and given the specific educational backgrounds, children attending afternoon EFL schools outside their regular/mainstream day schools show an older age advantage in grammar.

The higher accuracy of the older learners might be attributed to their overall higher proficiency which as discussed in the previous section, may be linked to a range of broader educational factors. For example, L1 literacy interacts with L2 proficiency (Cadierno *et al.*, 2022). Younger children in both L1 cohorts would have less well-developed literacy skills than older ones. The more well-developed L1 skills can assist and accelerate L2 acquisition (Tribushinina, Dubinkina-Elgart, & Rabkina, 2020). Through schooling, older ones

should have developed learning strategies that help their learning. They have more mature cognitive skills which are known to help general learning but also explicit language learning (Jaekel et al., 2017; Muñoz, 2006, 2010; Pfenninger, 2014). Finally, older learners have more advanced language aptitude skills (e.g., language analytic ability); recent research has shown that language aptitude is dynamic in childhood, being developed up to the age of 12 (Roehr-Brackin & Tellier, 2019). We would thus expect older learners to have more developed language analytic ability compared to younger learners which would directly facilitate the acquisition of verbal morphology.

In sum, the age advantage in grammar observed in immersed or naturalistic contexts for younger children is not attested in instructed L2 learners with limited input even if children start their classes at a very young age. Older learners benefit from cognitive maturity, higher literacy skills, and more effective learning strategies, as previous research has consistently demonstrated, while they also had more input.

5.3. First language effects on verbal morphology in EFL learning

Turning to our question of L1 effects in the acquisition of 3SG-agreement and past tense comparing Russian and Chinese learners, we found that Chinese learners did significantly better in 3SG-agreement while the two L1 cohorts did not differ in past tense (regular and irregular). Further, there was no significant interaction between age and L1, suggesting that the effect of age does not vary depending on the learners' first language. In other words, L1 effects are not influenced by age.

The fact that Chinese do better in 3SG-agreement is against our hypothesis. We argue that this is due to factors such as the quantity of input that we have already discussed, as well as the parental attitudes and the general cultural attitude towards schooling and English education (Bolton & Graddol, 2012) which may influence Chinese children's learning.

Another potential explanation for the observed "L1 effect" could be the higher number of native English teachers in Shanghai compared to non-native teachers in Moscow. This difference might influence the quality of language input, potentially explaining why Chinese learners demonstrate higher accuracy with 3SG-agreement. However, it is crucial to recognise that non-native teachers can possess near-native proficiency and employ highly effective pedagogical methods. On the other hand, not all native speakers may excel in teaching grammar. In any case, in contexts of limited input, where learning tends to rely more on explicit teaching methods, the influence of a teacher's nativeness may be minimal compared to the broader impact of teaching effectiveness.

We also considered the possibility of whether the teaching method followed at their day mainstream schools may be important in interaction with the type of task we used to assess children's production of morphemes. In a study in China on teachers' and learners' beliefs, questionnaire results revealed that the vast majority of the participants believed that the English examinations in schools and universities mostly assess a "command of English grammar" (Pan & Block, 2011). In classrooms in Shanghai mainstream schools, teachers follow a traditional approach to teaching grammar (e.g., Rao, 2013) with a focus on form and exam oriented. Unfortunately, there is a dearth of research in EFL teaching in Russia. To the best of our knowledge, there is only one study (Tribushinina, et al., 2020) conducted in West Siberia, according to which EFL books are also form-focused and the teaching of grammar is explicit. Hence, it seems that the teaching methodology in interaction with the type of task we used may not explain these results.

This study appears to contradict findings from ESL studies in immersed contexts where children learn English naturalistically. In ESL contexts, Chinese learners have been found to be slower in acquiring verb morphology compared to learners with L1s with rich verbal morphology (e.g., Paradis, 2011; Paradis *et al.*, 2016). While this contrast deserves further empirical investigation, we may speculate that immersed learners might not have benefited from the form-focused instruction typical of EFL teaching which also appears to be a salient feature of English instruction in mainstream education in China.

5.4. Morpheme effects in EFL learning

The final question we aimed to address in this study was whether 3SG-agreement as a purely syntactic feature will be more challenging than the past tense which is interpretable.

Statistical analyses showed a significant effect of inflection with better performance on past tense than 3SG-agreement across age and L1 groups. This is in line with our hypothesis.

It is possible that the past tense is taught for a longer period of time because of the fact that regular and irregular verb forms exist, where the irregular ones may have to be rote-learned and thus may be more frequently presented in the input children receive. On the other hand, the past tense is introduced later than 3SG-agreement in the curriculum, and children get exposed to agreeing forms such as *is/are* or *have/has* as soon as they start English classes, and so exposure alone may not sufficiently explain this difference in accuracy across age and L1 groups.

We attribute this asymmetry in the acquisition of 3SG-agreement and past tense to the nature of the features involved, with interpretable past tense being earlier/more easily acquired. One might argue that agreement morphology is the exception in an unmarked paradigm and, thus, less salient for the learner and less noticeable. If that was the case, we would expect irregular past forms to be more accurate than regular past *-ed* forms, which was not attested. This would be expected because irregular forms do not depend on a morphological mapping in the same way as regular forms which we did not see. They could be considered salient as they are rote-learned as whole words.

6. Conclusion

The cross-national comparison adopted in this study brought up many aspects which are often overlooked when considering a single country's context; from educational differences to parental attitudes and to even societal values, and emphasised that a range of factors may contribute to L2 learning. In this context, we found an older age advantage for grammar learning which was attributed to the higher cognitive maturity and analytic abilities of older children, while other underlying factors may include the higher quantity of input and the broader facilitative effects of schooling and literacy. We also found that at younger ages, Chinese children reach higher proficiency, again reflecting more input and potentially a stronger emphasis of parental and societal attitudes in China to early achievement. The same factors may also explain why Chinese learners outperform Russians in accuracy in agreement morphology. This study's results reflect the reality for many learners around the world and how English is learned in afternoon EFL schools given the limited hours offered in such contexts. Vocabulary consistently predicted outcomes in all analyses, possibly reflecting the overall amount of input each child received. If vocabulary knowledge

serves as an indirect measure or proxy for input, this suggests that the quantity of input is significant in these EFL contexts. Finally, learnability of the morphemes varied with past tense shown to be acquired more easily than 3SG-agreement. From a pedagogical perspective, it might be useful to consider the nature of the morphemes, as some may need more practice than others or different teaching strategies need to be employed.

We believe that cross-national comparisons can be very informative, and the current study is opening the ground for such research. Further research should examine the ultimate attainment of different starters in learning English, focusing on grammar achievement controlling for curriculum effects and controlling for the timing in the introduction of the phenomena. Administering detailed language background questionnaires to better control for exposure to English and perhaps match the groups in terms of hours, considering both school and out-of-school exposure, is essential.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S030500092500008X>.

Acknowledgments. We thank the EF schools in Shanghai and Moscow and all the teaching staff who welcomed us at their premises. We are grateful to all the children who participated in this study and their parents. This research was conducted within the EF Education First Research Lab for Applied Language Learning at the Department of Theoretical and Applied Linguistics, University of Cambridge. The first author gratefully acknowledges a studentship from EF Education First.

Competing interest. The authors have no conflicts of interest to disclose.

References

- Bodrova, E., & Yudina, E.** (2018). Early childhood education in the Russian Federation. In J. L. Roopnarine, J. E. Johnson, S. F. Quinn, & M. M. Patte (Eds.), *Handbook of international perspectives on early childhood education* (pp. 59–70). New York: Routledge. <https://doi.org/10.4324/9781315562193>
- Bolton, K., & Graddol, D.** (2012). English in China today: The current popularity of English in China is unprecedented, and has been fuelled by the recent political and social development of Chinese society. *English Today*, 28(3), 3–9. <http://doi.org/10.1017/S0266078412000223>
- Brehm, L., & Alday, P. M.** (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, 125, 104334. <https://doi.org/10.1016/j.jml.2022.104334>
- Butler, Y. G., & Le, V.-N.** (2018). A longitudinal investigation of parental social-economic status (SES) and young students' learning of English as a foreign language. *System*, 73, 4–15. <https://doi.org/10.1016/j.system.2017.07.005>
- Cadierno, T., Hansen, M., Lauridsen, J., Eskildsen, S., Fenyvesi, K., & Hannibal Jensen, S.** (2022). A multi-factor study of the development of English receptive skills by young Danish children. *Bilingualism: Language and Cognition*, 25(4), 566–578. <https://doi.org/10.1017/S1366728921001085>
- Cadierno, T., Hansen, M., Lauridsen, J. T., Eskildsen, S. W., Fenyvesi, K., Hannibal Jensen, S., & aus der Wieschen, M. V.** (2020). Does younger mean better? Age of onset, learning rate and short-term L2 proficiency in young Danish learners of English. *Vigo International Journal of Applied Linguistics*, 17, 57–86. <https://doi.org/10.35869/vial.v0i17.1465>
- Cenoz, J.** (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code-mixing. In M. García Mayo, & M. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77–93). Clevedon: Multilingual Matters. <https://doi.org/10.21832/9781853596407-005>
- Chomsky, N.** (1995). Bare phrase structure. In G. Webelhuth (Ed.), *Government and binding theory and the minimalist program* (pp. 383–439). Oxford: Blackwell.
- Chondrogianni, V. & Marinis, T.** (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, 1(3), 318–345. <https://doi.org/10.1075/lab.1.3.05cho>

- Council of Europe** (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- DeKeyser, R. M.** (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1–25. <https://doi.org/10.1111/j.0023-8333.2005.00294.x>
- Djigunović, J. M., & Nikolov, M.** (2019). Motivation of young learners of foreign languages. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 515–533). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-28380-3_25
- Fenyvesi, K., Hansen, M. B., & Cadierno, T.** (2020). The role of individual differences in younger vs. older primary school learners of English in Denmark. *International Review of Applied Linguistics in Language Teaching*, 58(3), 289–322. <https://doi.org/10.1515/iral-2017-0053>
- García Mayo, M., & García Lecumberri, M. L.** (Eds.) (2003). *Age and the acquisition of English as a foreign language*. Multilingual Matters. <https://doi.org/10.21832/9781853596407>
- Hawkins, R., & Chan, C.** (1997). The partial availability of universal grammar in second language acquisition: The 'Failed Functional Features Hypothesis'. *Second Language Research*, 13(3), 187–226. <https://doi.org/10.1191/026765897671476153>
- Haznedar, B.** (1997). L2 acquisition by a Turkish-speaking child: Evidence for L1 influence. In E. Hughes & A. Greenhill (Eds.), *Proceedings of the 21st Boston university conference on language development* (pp. 257–268). Somerville, MA: Cascadilla Press.
- Housen, A.** (2002). A corpus-based study of the L2-acquisition of the English verb system. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Language learning & language teaching* 6 (pp. 77–116). John Benjamins Publishing Company. <https://doi.org/10.1075/llt.6.08hou>
- Hu, B. Y., & Szente, J.** (2009). Exploring the quality of early childhood education in China: Implications for early childhood teacher education. *Journal of Early Childhood Teacher Education*, 30(3), 247–262. <https://doi.org/10.1080/10901020903084330>
- Hyldenstam, K., & Abrahamsson, N.** (2003). Maturation constraints in SLA. In C. J. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–588). Oxford: Blackwell. <https://doi.org/10.1002/9780470756492.ch17>
- Jaekel, N., Schurig, M., Florian, M., & Ritter, M.** (2017). From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning*, 67(3), 631–664. <https://doi.org/10.1111/lang.12242>
- Jaekel, N., Schurig, M., & Ritter, M.** (2022). Investigating cohort effects of early foreign language learning. *The Language Learning Journal*, 52(1), 104–116. <https://doi.org/10.1080/09571736.2022.2108124>
- Johnson, J., & Newport, L. E.** (1989). Critical period effects in second language learning. The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Lardiere, D.** (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum. <https://doi.org/10.2307/416851>
- Larson-Hall, J.** (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research*, 24, 35–63. <https://doi.org/10.1177/0267658307082981>
- Li, C. N., & Thompson, S. A.** (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Li, H.** (2013). Teaching Chinese literacy in the early years: A comparison of L1 and L2 preschool classrooms in Shenzhen and Singapore. *Asia-Pacific Journal of Research in Early Childhood Education*, 7(3), 19–43.
- Li, H., Corrie, L. F., & Wong, B. K. M.** (2008). Early teaching of Chinese literacy skills and later literacy outcomes. *Early Child Development and Care*, 178(5), 441–459. <https://psycnet.apa.org/doi/10.1080/03004430600789365>
- Lindgren, E., & Muñoz, C.** (2012). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10(1), 105–129. <https://doi.org/10.1080/14790718.2012.679275>
- Meisel, J. M.** (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft*, 28(1), 5–34. <https://doi.org/10.1515/ZFSW.2009.002>
- Mezhevich, I.** (2008). A feature-theoretic account of tense and aspect in Russian. *Natural Language & Linguistic Theory*, 26(2), 359–401. <https://www.jstor.org/stable/20533008>
- Mobaraki, M., Vainikka, A., & Young-Scholten, M.** (2008). The status of subjects in early child L2 English. In B. Haznedar & E. Gavrusova (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 209–236). Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.46.11mob>

- Muñoz, C. (Ed.) (2006). *Age and the rate of foreign language learning: Second language acquisition*. Clevedon: Multilingual Matters. <https://doi.org/10.21832/9781853598937>
- Muñoz, C. (2010). On how age affects foreign language learning. In *Advances in research on language acquisition and teaching: Selected papers*. GALA.
- Muñoz, C., & Lindgren, E. (2011). Out-of-school factors – The home. In J. Enever (Ed.), *ELLiE: Early language learning in Europe* (pp. 103–124). London: British Council.
- Myles, F. (2022). Attitudes and motivation towards learning French in primary school the role of developmental changes between the ages of 5 and 7. *Language Teaching for Young Learners*, 4(2), 240–263. <https://doi.org/10.1075/ltyl.22005.myl>
- Myles, F., & Mitchell, R. (2012). *Learning French from ages 5, 7, and 11: An investigation into starting ages, rates and routes of learning amongst early foreign language learners*. (ESRC End of Award Report, RES-062-23-1545). Swindon, England.
- Pan, L., & Block, D. (2011). English as a “global language” in China: An investigation into learners’ and teachers’ language beliefs. *System*, 39(3), 391–402. <https://doi.org/10.1016/j.system.2011.07.011>
- Paradis, J. (2011). Individual differences in child second language acquisition. Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213–237. <https://doi.org/10.1075/lab.1.3.01par>
- Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, 29(4), 689–722. <https://doi.org/10.1017/S0142716408080296>
- Paradis, J., Tulpar, Y., & Arppe, A. (2016). Chinese L1 children’s English L2 verb morphology over time: Individual variation in long-term outcomes. *Journal of Child Language*, 43(3), 553–580. <https://doi.org/10.1017/S0305000915000562>
- Pfenninger, S. E. (2014). The misunderstood variable: Age effects as a function of type of instruction. *Studies in Second Language Learning and Teaching*, 3, 529–556. <https://doi.org/10.14746/sslt.2014.4.3.8>
- Pfenninger, S. E., & Singleton, D. (2016). Age of onset, socio-affect and cross-linguistic influence: A long-term classroom study. *Vigo International Journal of Applied Linguistics*, 13, 147–179. <https://revistas.uvi.go.es/index.php/vial/article/view/78>
- R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Rao, Z. (2013). Teaching English as a foreign language in China: Looking back and forward: Reconciling modern methodologies with traditional ways of language teaching. *English Today*, 29(3), 34–39. <http://doi.org/10.1017/S0266078413000291>
- Renfrew, C. (1995). *Word Finding Vocabulary Test*. 4th Edition. Oxon: Winslow.
- Rice, M., & Wexler, K. (2001). *Test of early grammatical impairment*. New York: The Psychological Corporation. <http://hdl.handle.net/1808/32569>
- Rocca, S. (2007). *Child second language acquisition: A bi-directional study of English and Italian tense-aspect morphology*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.35>
- Roehr-Brackin, K., & Tellier, A. (2019). The role of language-analytic ability in children’s instructed second language learning. *Studies in Second Language Acquisition*, 41, 1111–1131. <https://doi.org/10.1017/S0272263119000214>
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–94). John Benjamins. <https://doi.org/10.1075/llt.2.06ske>
- Sun, H., de Bot, K., & Steinkrauss, R. (2014). A multiple case study on the effects of temperamental traits in Chinese preschoolers learning English. *International Journal of Bilingualism*, 19(6), 703–725.
- Tan, C. (2012). The culture of education policy making: Curriculum reform in Shanghai. *Critical Studies in Education*, 53(2), 153–167. <https://doi.org/10.1080/17508487.2012.672333>
- Tragant, E., & Muñoz, C. (2023). *Ten years of English learning at school*. Palgrave Macmillan.
- Tribushinina, E., Dubinkina-Elgart, E., & Rabkina, N. (2020). Can children with DLD acquire a second language in a foreign-language classroom? Effects of age and cross-language relationships. *Journal of Communication Disorders*, 88, 106049. <https://doi.org/10.1016/j.jcomdis.2020.106049>
- Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from Wh-interrogatives in second language acquisition. *Second Language Research*, 23(2): 215–242. <https://doi.org/10.1177/0267658307076546>

- Yang, S., & Huang, Y. Y.** (2004). The impact of the absence of grammatical tense in L1 on the acquisition of the tense-aspect system in L2. *International Review of Applied Linguistics in Language Teaching*, **42**(1), 49–70. <https://doi.org/10.1515/iral.2004.002>
- Zhou, X.** (2011). Early childhood education policy development in China. *International Journal of Child Care and Education Policy*, **5**(1), 29–33 <https://doi.org/10.1007/2288-6729-5-1-29>

Cite this article: Ntalli, A., Alexopoulou, T., Hendriks, H., & Tsimpli, I.M. (2025). The Acquisition of Verbal Morphology by Child Classroom EFL Learners in Russia and China: The Effect of Age and L1. *Journal of Child Language* 1–28, <https://doi.org/10.1017/S030500092500008X>