Factors Affecting Language Proficiency in Heritage Language: The Case of Young Russian Heritage Speakers in Spain

Tamara Vorobyeva
PhD student of Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain
vorobyeva.tamara@gmail.com

Aurora Bel
Associate professor of Language Acquisition and Bilingualism, Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain
aurora.bel@upf.edu

Abstract

This study focuses on the issue of language proficiency attainment among young heritage speakers of Russian living in Spain and examines factors that have been claimed to promote heritage language proficiency, namely, age, gender, age of onset to L2, quantity of exposure and family language use. A group of 30 Russian-Spanish-Catalan trilingual children aged 7–11 participated in the study. In order to measure heritage language proficiency (L1 Russian), oral narratives were elicited.

The results demonstrated a significant relationship between L1 proficiency and three sociolinguistic variables (age of onset to L2, quantity of exposure and family language use). Additionally, the multiply regression model demonstrated that the only significant variable affecting language proficiency was family language use and it accounted only for 33% of the variation of children’s language proficiency. The study raises the question about what are the other, yet unknown factors, which can affect heritage language proficiency.
Keywords

heritage speakers – child bilingualism – language proficiency – language dominance – Russian – input exposure

1 Introduction

A heritage speaker is typically a bilingual speaker who comes from immigrant and/or ethnic minority groups and usually starts acquiring the language of her/his parents (L1) but then shifts to a more socially dominant language (L2). Such an abrupt or gradual shift often limits typical L1 development and can result in an incomplete acquisition of different L1 features (Anderson, 2001; Montrul, 2008) or L1 attrition in adult heritage speakers (Polinsky, 2008, 2011, 2018).

During language acquisition, exposure to the L1 is often limited and occurs mainly in a home environment, whereas exposure to the L2 is overwhelming: it is the dominant language of schooling and communication in the society. As a result, heritage speakers are usually more proficient in L2 than in L1 (Benmamoun, Montrul and Polinsky, 2013; Polinsky, 2018). As previous research has revealed, one of the important characteristics of heritage speakers is that their L1 proficiency can significantly vary: some of them have limited productive abilities and/or even no literacy skills; meanwhile, others possess a native-like mastery of written and spoken modalities (Montrul, 2005, 2010; Valdés, 2005). The range of variability is supposed to be a result of differential input conditions, including such aspects as quantity of exposure to L1 (Gathercole and Thomas, 2005; Sorace, 2005; Unsworth, 2014, 2016a), age of onset to bilingualism (Montrul, 2008, 2016), chronological age (Gagarina and Klassert, 2018; Gathercole, Thomas, Hughes, Roberts and Hughes, 2013; Mancilla-Martinez and Lesaux, 2011), and socioeconomic status of the heritage speakers’ families, among others (Gathercole and Thomas, 2005; Hart and Risley, 1995; Jia, 2008).

One of the important factors that affect heritage language proficiency is exposure to L1 (Gathercole and Thomas, 2005; Sorace, 2005; Unsworth, 2008, 2014, 2016a). This factor includes two aspects: the quantity and quality of exposure to the L1. First, the quantity of exposure can vary depending on factors such as what languages are spoken at home (L1 or/and L2), the frequency they are being used, and an overall amount of L1 exposure (at home and outside the home). Another aspect is the quality of exposure that can be reduced to certain contexts of acquiring the L1 (naturalistic or formal), variety (e.g., peninsular Spanish vs. Puerto Rican Spanish), and register (written or oral). All these
aspects are crucial for normal L1 language development of heritage speakers (Jia and Paradis, 2015; Montrul, 2016). Moreover, such factor as age of onset to bilingualism (or age of onset to L2) was also reported to play a critical role in child language development. Various studies have demonstrated that the later a child comes into contact with L2 the better his/her L1 competencies will be (Montrul, 2008, 2016; Polinsky and Kagan, 2007). On the contrary, when the shift to L2 occurs early in childhood, L1 skills can be dramatically affected, which would result in incomplete L1 acquisition in the case of young heritage speakers (Anderson, 2001) or L1 attrition in adult heritage speakers (Polinsky, 2008, 2011).

The current study focuses on the role of external factors that have been claimed to promote heritage language proficiency, namely, the overall quantity of L1 exposure (including both exposure at home and outside the home), quantity of L1 exposure at home (referred as family language use in our study), and age of onset to L2 (Albirini, 2014; Gathercole and Thomas, 2005; Montrul, 2016; Unsworth, 2016b). We also examine the influence of demographic factors (age and gender) on heritage language proficiency. The study deals with a different sociolinguistic environment still not explored enough in the previous research and, namely, the context of Catalonia, a bilingual community where Catalan and Spanish are both present in social and educational contexts. Participants of the study are young Russian heritage speakers from Barcelona between ages 7 and 11. The growth of a Russian-speaking population in Spain started in the 90s and according to Spanish National Statistics Institute (2017), it jumped from 40m in 1999 to more than 47m in 2016. This pronounced demographic shift was almost entirely the result of immigrants from the former USSR countries such as Russia, Ukraine, Belorussia, and Kazakhstan.

The surge of the Russian-speaking population has provoked an increasing demand for Russian language classes where the second-generation Russian immigrants (mainly children from 3 years old) can learn the Russian language and get familiar with Russian culture. The children of immigrants who were born in Spain or migrated with their parents before 6 years old comprise the participant group in the present study. They are young multilingual children who speak Russian (L1 heritage language), Spanish and Catalan (L2s, dominant languages of communication and schooling in a bilingual environment of Barcelona).

Language proficiency levels have been demonstrated to be crucial for heritage language knowledge and the maintenance of language target features (Albirini, 2014; Montrul, 2004; Polinsky, 2000, 2006, 2008, among others). Since the proficiency of heritage languages can vary tremendously, assessing it is not a simple task (Benmamoun et al., 2013; Montrul, 2005, 2010, 2016).
Previous research on heritage language has implemented various language proficiency assessment techniques, some of them being vocabulary measures (Fairclough and Ramírez, 2009; Fairclough, 2011; Polinsky, 2000, 2006) or grammatical knowledge assessment (Alarcón, 2011; Gagarina, Klassert and Topaj, 2010). The evaluation of language skills based on oral narratives presented in this study has been widely used in the previous research measuring both L1 and L2 proficiency. It provides rich information about the linguistic development of children and allows for the evaluation of productive language abilities in monolingual and bilingual learning contexts (Gutiérrez-Clellen, 2004; Pearson, 2002). Previous studies on Russian as a heritage language have implemented different proficiency assessment measures based on oral narrative tasks. Some of them include (but are not limited to) rate of speech calculation (Polinsky, 2008), oral proficiency interviews (OP1) based on the American Council on the Teaching of Foreign Languages (ACTFL) guidelines (Kagan and Friedman, 2003), and the multilingual assessment instrument (MAIN) proposed by Gagarina et al. (2012, 2015).

All in all, the variability in heritage language proficiency has been explained by different external and internal factors, though it is not clear which of the factors, or combination of them, determine the native-like mastery of heritage language. Thus, the goal of this study is to examine the relationship between heritage Russian proficiency and such explanatory socio-linguistic and demographic factors as the quantity of L1 exposure, family language use, age of onset to L2, age, and gender in a societal context that has not been covered by previous research.

1.1 External Factors and Language Development

The context of heritage language acquisition is characterized by the limited quantity of exposure to L1 (Montrul, 2016). Therefore, it can result in the incomplete acquisition of heritage language in childhood or attrition in adulthood in different language domains (Montrul, Foote and Perpiñán, 2008; Polinsky, 2006, 2008, 2011 for morphology; Anderson, 1999; Pham and Kohnert, 2014 for lexicon). Additionally, the quantity of exposure to L1 in majority language dominant environment can vary depending on such factors as what languages are spoken at home, outside home (only L1 or both L1 and L2), and the frequency of their use. Unlike monolingual language environments when the exposure to L1 prevails, heritage speakers are exposed to different languages at the same time and the proportion of this exposure is often not in favour of L1.

If exposure to L1 input is not sufficient, some aspects of grammar are not completely acquired. Hoff et al. (2012) demonstrated that children with heritage language exposure under 40% fell below the proficiency levels of
monolingual peers when their grammatical knowledge was measured. Similar results were demonstrated by Thordardottir (2011) but for lexical development. The author reported that bilinguals can reach the monolingual language levels of receptive and productive vocabulary in both languages if the exposure will be from 40% to 50% in each language respectively. In addition, analysing language development (use of verbs, articles, and subjects in Italian) of four English-Italian bilinguals, La Morgia (2015) concluded that 20% of weekly language input may not be enough to reach the monolingual level of language performance.

According to Montrul (2012), the quantity of L1 exposure also includes such aspects as the modality of input (only oral or both oral and written) and the setting of language acquisition (at home versus formal instruction). Heritage speakers mainly acquire L1 in a naturalistic setting from their caregivers and/or other family members such as siblings (Benmamoun et al., 2013; Montrul, 2012, 2016). Indeed, a sufficient amount of L1 exposure at home has proved to have a positive contribution to heritage language development (De Houwer, 2007; Gathercole and Thomas, 2009; Lü and Koda, 2011). In their study, Paradis, Nicoladis, Crago, and Genesee (2010) found that the language of greater exposure at home (based on a parental questionnaire) was directly correlated with children’s rate of acquisition of morpho-syntax (measured via elicitation task). The children who had English as the language of greater exposure at home were more proficient with English verb forms, whereas the children with more French input were more proficient with French verb morphology. Similar findings for English-Chinese bilingual children were demonstrated by Lu and Koda (2011). For instance, the use of Chinese at home supported the development of literacy skills and oral vocabulary acquisition in Chinese. Also, in her study De Houwer (2007) showed that having at least one parent who speaks the minority language at home positively affects the L1 development of a child. In contrast, if both parents use the L2 at home, the probability that children will speak minority language declines significantly. To conclude, all the findings listed above indicate the importance of the minority language use at home and outside the home for successful heritage language development and maintenance.

The quantity of L1 exposure is closely connected with the age of onset to L2 (Gathercole and Thomas, 2005; Montrul, 2016; Sorace, 2005). The age of the onset to bilingualism is the age when a child starts hearing and thereby acquiring L2. If children start acquiring both languages from birth (i.e., parents are native speakers of different languages), then they are simultaneous bilinguals (De Houwer, 2009; Genesee and Nicoladis, 2007). When children start acquiring L1 from birth and then L2 after the age of 3–4 (Meisel, 2011) they are
sequential bilinguals. The early age of onset to L2 has proved to be a predictor of successful L2 acquisition (Montrul, 2008). On the other hand, it’s also a crucial factor in the incomplete acquisition of a child’s L1 or the attrition of vulnerable L1 language domains in adulthood (Anderson, 2001; Montrul, 2008; Polinsky, 2008). Several studies have demonstrated that sequential bilinguals tend to be more proficient in their L1 when compared to simultaneous bilinguals. In their longitudinal study, Jia and Aaronson (2003) found that within a year of arrival to the U.S., young Chinese-speaking immigrants (between ages 5 and 9) had switched from L1 to L2, whereas older children continued to use the L1 with their Chinese-speaking parents and siblings. In other research on heritage languages in the U.S. Carreira and Kagan (2011) found that the earlier heritage speakers start learning English, the less likely they will use their heritage language in adulthood. Furthermore, the respondents who were U.S.-born tended to rate their heritage language skills lower than those who arrived later. Caldwell-Harris, Staroselsky, Smashnaya and Vasilyeva (2012) demonstrated that proficiency of L1 Russian increased with later age of onset to English. The participants were divided into three groups depending on their age of arrival to the U.S. Late arrivals (those who arrived at 10 years and later) tended to have a similar proficiency level in L1 and L2 (based on a self-report evaluation), middle arrivals (from 6 to 9) became more dominant in L2 compared to L1, and finally, early arrivals (from birth to 5) had limited L1 skills even if their home language was exclusively Russian. As Montrul (2016) notices, the earlier L2 acquisition starts in childhood, the more vulnerable are the L1 linguistic skills in a minority language. Thus, the general rule of how the age of onset influences proficiency of L1 can be formulated as follows: the later the acquisition of L2 starts, the higher is L1 proficiency.

Not only quantity of exposure and age of onset, but also age has been shown to influence language development. According to the previous studies, aging populations (both monolingual and bilingual) are known to outperform younger populations (Gathercole et al., 2013; Mancilla-Martinez and Lesaux, 2011; Stine-Morrow and Shake, 2009). In their study, Gathercole et al. (2013) tested vocabulary and grammar performance of four age groups (age range 2–15) of English-Welsh bilinguals. The results demonstrated that older bilinguals had better vocabulary and grammar skills in both languages than younger ones. Similar results were shown by Mancilla-Martinez and Lesaux (2011), but for Spanish heritage speakers (aged 4 to 11). They found a positive correlation between the child’s age and vocabulary growth when measured using standardized tests for Spanish and English language.

Additionally, several studies have demonstrated that gender is a predictor of successful heritage language acquisition (Portes and Hao, 1998; Zentella, 1997)
and L2 learning (Ellis, 1994; Larsen-Freeman and Long, 1991; Slik et al., 2015). Portes and Hao (1998) showed that females performed better than males on heritage language proficiency measures. The authors reasoned that females tend to spend more time with their parents and, hence, were more exposed to the heritage language. In their study, Slik et al. (2015) collected data from more than 25,000 adult learners of L2 Dutch with 88 different L1s. The results proved that female learners independent of L1 outperformed males in speaking, writing, reading, and listening. To better understand the reason for the gender gap, the researchers proposed that additional studies on cognitive abilities in males and females were needed in L1 and L2 acquisition research.

1.2 Proficiency Assessment of Heritage Speakers

One of the distinctive aspects of heritage languages is the variability in language skills that heritage speakers exhibit; some are proficient as native speakers, some speak fluently but have no literacy skills, some understand the language but have limited productive skills (Benmamoun et al., 2013; Montrul, 2005). This diversity makes the heritage language assessment and, as a result, the placement of heritage speakers in the classroom particularly challenging.

In previous studies, different diagnostics of heritage language proficiency in adults were implemented. For instance, Polinsky (2008) used speech rate measures in spontaneous oral productions for evaluating heritage speakers of Russian. Their proficiency proved to be correlated with the level of grammatical knowledge: speakers with poor speech production demonstrated difficulties in grammatical gender control. Also, some formal standardized methods were elaborated to assess the proficiency of heritage speakers. Kagan and Friedman (2003) suggested using ACTFL oral proficiency interviews (OP1) as part of a placement procedure for heritage speakers with different L1s.

Several methods have been used in an attempt to properly assess language competencies of heritage children. Gagarina et al. (2010) created a proficiency test for typically developing bilingual children of pre-school and elementary school age. Originally, the test was designed for Russian-German bilinguals but according to the authors, it is suitable for any bilingual child with L1 Russian. The test evaluates both grammar and lexical knowledge in production and comprehension. The grammar part consists of elicited oral production tasks (assessing knowledge of the case and verbal inflections) and a picture selection task with auditory sentence presentation (assessing knowledge of different grammatical constructions: negation, passive voice, relative clauses, etc.). The lexical part includes the production of nouns and verbs through a picture-naming task and a picture selection comprehension task with an auditory presentation. Additionally, Gagarina et al. (2012) proposed a proficiency
assessment instrument based on a narrative task. The tool is used for a proficiency assessment, speech diagnosis, and an assessment of narrative abilities of multilingual children. It consists of telling and retelling stories based on visual stimuli that are suitable for children from various socio-cultural settings and different L1 and L2 pairs. The authors also provide an evaluation scheme based on analyses of the micro- and macrostructure of the narratives.

All in all, a large body of experimental work on proficiency assessment is based on narratives. Additionally, narratives have proven to be the effective way of evaluating linguistic skills because they consider the diversity of speakers (as in case of heritage speakers) and it is appropriate for any proficiency level, even with an illiterate population (Flecken, 2011; Gutiérrez-Clellen et al., 2000). Finally, narratives are relatively easy to elicit, requiring little material for collecting a rather reliable set of linguistic data (Pavlenko, 2008).

2 Research Questions for this Study

As was previously stated, language development can be influenced by many aspects, and among them are socio-linguistic and demographic factors such as quantity of exposure, age of onset to L2, age, and gender. In addition, the previous studies analysed the relationship between these factors and heritage language proficiency, and a correlation between them. For instance, in the study by Swender, Martin, Rivera Martinez and Kagan, 2014) the researchers found that heritage language use at home and age of onset to L2 proved to be correlated with higher proficiency ratings in heritage Spanish and Russian in the USA. In Albirini’s (2014) study, quantity of exposure and language use was shown to be strongly predictive of higher proficiency ratings in heritage Arabic. The current study examines some of the above-mentioned factors that may influence heritage language proficiency, while incorporating a new sociolinguistic environment and looking at three languages instead of two, which is the most common background in the field of heritage language studies. Namely, we extended the linguistic research to a group (Russian-Spanish-Catalan speaking children) that has not been investigated previously regarding variability in language proficiency.

The specific research questions were as follows:

i. To what extent sociolinguistic and demographic factors such as age, gender, age of onset to L2, quantity of exposure and family language use show a correlation with Russian language proficiency?

ii. Which of these factors are the most important predictors of language proficiency among Russian heritage speakers?
Language proficiency in this study was measured with the help of an oral narrative task. Also, a proficiency index based on accuracy and grammaticality constructs was calculated (see section 3.5 for a detailed description). As previously stated, narratives constitute an appropriate and widely used tool for assessing proficiency in a bilingual environment. The information about age, gender, age of onset to L2, quantity of exposure, and family language use was collected via a parental questionnaire and described below in section 3. To identify associations between Russian language proficiency and sociolinguistic and demographic factors, descriptive statistics and correlation indexes were done. In addition, multiple regression analyses enabled us to investigate which factor was the most important predictor of higher heritage language proficiency.

3 The Study

3.1 Participants

The background information of participants was collected through the parental questionnaire administered in paper and pencil in a community-based school in Barcelona where the participants of the study were recruited. There is a range of Russian Saturday schools operating in all major cities in Spain. The Russian community school in Barcelona provides Russian language and literature classes for the children. Apart from the Russian language tuition, the school’s curriculum includes other disciplines taught in Russian as well as History, Literature, Geography, Arts, and Music.

The participants in the study included 30 multilingual children (13 boys, 17 girls) ranged in age between 7 and 11 (mean=8.8) and divided into two age groups for further statistical analysis. Group 1 consisted of 15 children aged 7–8 years (mean=7.5) and Group 2 included 15 participants aged between 9–11 (mean=10.3). All children started the Russian community school at the age of 3–4 and receive approximately 120 hours of Russian lessons per year. Fourteen children come from monolingual Russian-speaking families and sixteen children are from bilingual families. In the bilingual families all mothers speak Russian, seven fathers speak Spanish and nine fathers speak Catalan with their children. Twenty-two participants were born in Spain and eight were born in Russian-speaking surrounding and moved with their families to Spain at the

---

1 One child was born in Ukraine and one in Belorussia. They moved to Spain before they turned one year old. Their parents are native speakers of Russian and do not speak Ukrainian or Belorussian.
ages between 3 and 8 (average age of arrival=4.6). Nearly 90% of the families are in the mid and mid-upper income group and almost 70% of the parents have higher education. Ninety per cent of all parents speak at least one foreign language (mainly English) as well as Spanish and/or Catalan.

3.2 **Age of Onset to L2 (Spanish and/or Catalan)**

To calculate the age of onset for each participant the earliest age of exposure to L2 was converted into months. For example, if a child has been exposed to Spanish at 6 months age and to Catalan at 24 months, the number 6 was chosen as the age of onset parameter. Those children who were exposed to L2 Spanish and/or Catalan from birth scored 0 (the case of bilingual families where Russian and Spanish and/or Catalan is spoken). The average age of onset to L2 Spanish and/or Catalan of children from monolingual Russian-speaking families is 2.2.

3.3 **Quantity of Exposure to Russian**

Quantity of exposure to Russian was based on children’s current amount of overall exposure to Russian in different contexts (following Gathercole and Thomas, 2009; Unsworth, 2013). The parents were asked to report an average quantity of exposure (in percentages) on a weekly basis to Russian, Spanish, Catalan, and another language(s) (i) at home, (ii) at school, including both ‘local’ and Russian community school, (iii) after-school, extra-curricular activities. On average, the parents from both monolingual and bilingual families reported Russian as being used in 40% of all time in monolingual families and in 20% of the time in bilingual families. Spanish and/or Catalan were reported to be the main languages of communication overall and also the main languages of schooling (70%) and after-school activities (75%).

3.4 **Family Language Use**

Fourteen children come from monolingual Russian-speaking families; sixteen children come from mixed bilingual families, all with Russian-speaking mothers, and L1 Spanish and/or Catalan speaking fathers. To calculate the amount of language use in the family, the parents were asked to report an average percentage of Russian, Spanish, Catalan, and another language(s) use among different family members on a weekly base. This included aspects such as (i) language of communication between parents, use of languages of (ii) father, (iii) mother, (iv) grandparents, (v) caregivers/tutors (if applicable)

---

2 In this study, Catalan and Spanish are equally considered as L2s of the participants.
when speaking to a child, and languages that the children speak with their (vi) siblings (if any).

Finally, the average percentage from 0 to 100 for the Russian language use at home was calculated individually for each participant. The amount of Russian use at home in monolingual families was 90%, and in bilingual families, it was around 60%. The characteristics of the participants presented above are summarized in Table 1 below.

<table>
<thead>
<tr>
<th>Table 1 Sociolinguistic variables of the participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Participants</td>
</tr>
<tr>
<td>monolingual families</td>
</tr>
<tr>
<td>bilingual families</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>male</td>
</tr>
<tr>
<td>female</td>
</tr>
<tr>
<td>Age (in years)</td>
</tr>
<tr>
<td>group 1</td>
</tr>
<tr>
<td>group 2</td>
</tr>
<tr>
<td>Age of onset to L2 (years, months)</td>
</tr>
<tr>
<td>monolingual families</td>
</tr>
<tr>
<td>bilingual families</td>
</tr>
<tr>
<td>Overall quantity of exposure (Russian, in %)</td>
</tr>
<tr>
<td>monolingual families</td>
</tr>
<tr>
<td>bilingual families</td>
</tr>
<tr>
<td>Family language use (Russian, in %)</td>
</tr>
<tr>
<td>monolingual families</td>
</tr>
<tr>
<td>bilingual families</td>
</tr>
</tbody>
</table>

3.5 Proficiency Assessment of Heritage Speakers

The assessment procedure was intended to gauge Russian language proficiency of the young heritage speakers. A picture description task eliciting an oral narrative was employed. The children were presented with the picture book “Frog, where are you?” (Mayer, 1969). The children were given a paper copy of the book and were asked to narrate a story in Russian. They could see two pages of the book at once and turned the pages in their own pace. The data collection guidelines were adapted from Berman and Slobin (1994). Materials included
a paper copy of the book, recording equipment (audio digital recorder with an external microphone and video camera), and a protocol for controlling the date of data collection, order and number of children. The task was conducted one-on-one in a quiet empty classroom with a table and two chairs.

After collecting the recordings, they were transcribed by the researcher using Pearson’s (2002) convention for transcription and were then verified by two native Russian speakers. Disagreements between the transcriber and each verifier were resolved by listening to the audio to determine the final transcriptions to be used for analysis. The verified transcriptions were coded following the clan protocol (MacWhinney, 2000), allowing for semi-automated analysis with the clan programs. Exact repetitions, revisions, and interjections (e.g., uh, um) were excluded. The corpus was divided into utterances based on the intonation, pauses, and syntax of a sentence. The utterance definition was adapted from Loban (1976): each utterance should express a complete idea and grammatically should consist of one main clause and all its subordinate clauses. In this study, language proficiency assessment was based on error calculation in the narratives. An error was understood as belonging to the morpho-syntactic or the lexical-semantic levels. We did not provide a detailed analysis of errors here, e.g., case, number or gender errors were all coded as morpho-syntactic errors. Also, the errors were not subcategorized into omission, commission, or addition. We believe that the detailed overview of errors worthy to be an object of another paper. Phonetical errors were not included in the analyses (only a few errors were made). Discourse-functional aspects were not analysed since the goal of the study was an assessment of grammar and vocabulary skills only. All errors were semi-automatically annotated by the researcher (a native Russian-speaking philologist was consulted) using a coding scheme that captured the morpho-syntactic and lexical errors. Unintelligible forms were disregarded from the analysis. In total, morpho-syntactic errors were the most frequent and comprise 83% of all errors followed by lexical-semantic errors (17%).

The sentences below are some examples of common errors made by the heritage speakers in their oral narratives. For instance, some of the participants had difficulties with lexical access and retrieval. Consider the following example 1, when a child was pointing to the reindeer on the picture in the book but could not remember a word for it in Russian.

(1) *Malčik videl čto eto byl Ø
Boy saw that this was Ø
‘The boy saw that it was a reindeer.’

The possible sentence in native Russian should be like this (1a).
(1a) Malčik videl čto eto byl olen’
Boy saw that this was reindeer
‘The boy saw that it was a reindeer.’

Also, there were several transfer errors, mostly from Spanish or Catalan (the
transferred items were phonologically similar, so it was difficult to determine
the language of the transfer). In example 2, a child instead of Russian word les
(forest) used a Spanish word bosque (or Catalan bosc).

(2) *Malčik pošel v bosk i kričal
Boy went in bosk and yelled
‘The boy went to the forest and he was yelling.’

The corresponding native Russian sentence should be the following (2a).

(2a) Malčik pošel v les i kričal
Boy went in forest and yelled
‘The boy went to the forest and he was yelling.’

Interestingly, the participants made several errors with prepositions. Consider
element 3 in which the child used the preposition na (on) instead of v (in).

(3) *Potom oni uekhali iskat’ na park
Then they went search on park
‘Then they went to search in the park.’

The possible sentence in native Russian involves preposition v ‘in’ (3a).

(3a) Potom oni uekhali iskat’ v park
Then they went search in park
‘Then they went to search in the park.’

Additionally, in example 4 we can observe the absence of an obligatory prepo-
sition na (at) with the verb smotrel (look) when referring to animates.

(4) *I sobačka smotrela Ø ljagušku
And doggie looked Ø frog
‘And doggie looked at the frog.’

The corresponding native Russian sentence should have the preposition na
‘at’ (4a).

(4a) I sobačka smotrela na ljagušku
And doggie looked at frog
‘And doggie looked at the frog.’
In the example 5 the participant used incorrectly the conjunction počemu (why) instead of potomu čto (because). The error may be caused by cross-linguistic influence from L2s; in Spanish and Catalan there is a same word porque meaning both why and because.

(5) *Sobačka hotela poigrat’ počemu ona dumala čto èto mjačik
   Doggie wanted to play why she thought that it ball
   ‘The doggie wanted to play because he thought it was a ball.’

The possible sentence in native Russian is the following (5a).

(5a) Sobačka hotela poigrat’ potomu čto ona dumala čto èto mjačik
   Doggie wanted to play because she thought that it ball
   ‘The doggie wanted to play because he thought it was a ball.’

Most errors were made on morpho-syntactical level, namely, the heritage speakers had problems with the following features: case (examples 6, 7) and number marking (7), gender marking (6), and aspect (8). Consider the following utterances from the narratives below.

Incorrect verbal gender agreement and a noun case error (6).

(6) *Sobačka igral-sja s komarik-i
   Doggie played-M with mosquitoes-NOM
   Doggie played-F with mosquitoes-INS (correct)
   ‘The doggie played with mosquitoes.’

Incorrect number and case of a noun(7).

(7) Našel dv-a lâgušk-a
   Found two-NOM frog-NOM.SG
   Found two-ACC frogs-GEN.PL (correct)
   ‘He found two frogs.’

Wrong use of imperfective aspect instead of perfective (8).

(8) On hotel vylez-at’ iz dyrki čtoby posmotret’
   He wanted get-IPFV out hole to look
   He wanted get-PFV out hole to look
   ‘He wanted to get out of the hole to look.’

The examined errors are common in heritage Russian and were documented in the narratives of the Russian heritage speakers with different L2s (e.g., Mikhaylova and Ravitch, 2018; Protassova, 2014).

To assess the children’s narratives, two measures were developed. The accuracy measure reflected a ratio of error-free utterances over the total number of
utterances (following Larsen-Freeman, 1978, 2006). Error-free utterances were defined as utterances in which no error was found regarding morphology, syntax, word choice, word use, etc.

The grammaticality measure was quantified as a ratio of error-free words divided by the total number of words (following Chastain, 1990). As it was said before, because of the variability in the narratives, it was assumed that a more valid indicator of proficiency would be provided by a combination of the two measures. In order to make the measures equivalent, they were converted into a decimal figure by multiplying by 10 (e.g., 0.943 × 10 = 9.43). Both measures made up a range scale from 0 to 10. Finally, the figure of one proficiency measure was added to another, resulting in a single proficiency score for each child. The final value range was from 9.92 to 20.00. All the proficiency measures are summarized in Table 2.

### Table 2: Group-average proficiency measures

<table>
<thead>
<tr>
<th>Total Utterances</th>
<th>Total Errors</th>
<th>Error-free Utterances</th>
<th>Accuracy Measure</th>
<th>Grammaticality Measure</th>
<th>Proficiency Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>30.13</td>
<td>12.83</td>
<td>21.27</td>
<td>7.12</td>
<td>9.36</td>
</tr>
<tr>
<td>SD</td>
<td>7.12</td>
<td>13.40</td>
<td>9.07</td>
<td>2.60</td>
<td>0.70</td>
</tr>
</tbody>
</table>

4 Results

Different statistical tests were conducted to see which factors affect heritage language proficiency: t-tests, correlations and a multiple regression analysis. As noted above, the demographic (age and gender) and sociolinguistic variables (age of onset to L2, quantity of exposure, and family language use) were obtained with the help of a parental questionnaire. Table 1 in section 3.4 presents the descriptive statistics corresponding to age, gender, age of onset to L2, quantity of exposure, and family language use.

To address the question of whether demographic variables (age and gender) have any effect on heritage language proficiency, t-tests were performed. Age differences were analysed by classifying participants into two age groups. In this case, age was treated as a discrete variable: each age group was coded 1 and 2 (1 = ages 7–8; 2 = ages 9–11). The analyses revealed that children aged between 7 and 8 obtained similar results (M = 13.87, SD = 6.24) to those of the older group
aged 9–11 (M = 14.04, SD = 5.81). Thus, in the subsequent independent-samples T-test, no significant difference was observed between groups (t28=-0.080, p=0.687).

Regarding gender, no significant difference was found in proficiency scores between males and females (t21.428=.946, p=0.355). However, the proficiency scores of females were slightly higher (M = 17.18, SD = 2.62) than the scores in the male group (M = 16.07, SD = 3.52).

To investigate the influence of age of onset to L2, quantity of exposure, and family language use on L1 proficiency, correlations, and a multiple regression analysis was carried out. The analyses were conducted in two steps. First, each of these three variables was treated separately and correlated with L1 proficiency. Second, a multiple regression analysis involving the three independent variables was performed to ascertain which of them can predict the variability in heritage language proficiency and to what degree.

For single correlations, a Spearman’s rank-order was run. Concerning the relationship between age of onset to L2 and proficiency scores in Russian, there was a strong, positive correlation, which was statistically significant (rs = .679, p < .001). In the case of the relationship between quantity of exposure to Russian and L1 proficiency, the Spearman’s correlation calculation also yielded a positive correlation between both variables, which was statistically significant (rs = .651, p < .001). Finally, the strongest and significant positive correlation was between family language use and L1 proficiency scores gave a significant association (rs = .733, p < .001). As a summary, Table 3 shows the correlation coefficients between the factors just reported.

To unveil a more complex and accurate picture of ‘true’ relationships, multiple regression analyses were performed for estimating the relative weights and importance of different factors. The results yielded statistically significant values (F (3, 29) = 158.912; p < .001) and showed that together, age of onset to L2, quantity of exposure and family language use accounted for 33% of the variability in language proficiency (Table 4). Moreover, the coefficients indicated that the only significant variable affecting heritage language proficiency was

<table>
<thead>
<tr>
<th>Language proficiency</th>
<th>Age of onset</th>
<th>Quantity of exposure</th>
<th>Family language use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman correlation</td>
<td>rs = .679</td>
<td>rs = .651</td>
<td>rs = .733</td>
</tr>
<tr>
<td>Significance</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
<td>p &lt; .001</td>
</tr>
</tbody>
</table>
family language use ($\beta=.578$, $t=3.745$, $p < .001$ as depicted in Table 5). Therefore, it can be suggested that family language use was the only significant predictor of Russian heritage proficiency of those addressed here. This means that the proficiency scores in Russian increase together with the use of Russian as the language of communication in the home environment.

5 Discussion and Conclusion

The first study’s aim was to evaluate the existence of effects of factors such as age, gender, quantity of exposure, family language use, and age of onset to L2 on the L1 proficiency of young Russian heritage speakers based on oral narrative analyses. The results of the study demonstrated that proficiency scores were positively associated with quantity of exposure, family language use, and age of onset to L2. However, no interaction was found between age (operationalized as two age groups) and L1 Russian language proficiency. According to previous studies, older monolingual and bilingual children are known to outperform younger ones (Gathercole et al., 2013; Mancilla-Martinez and Lesaux, 2011; Stine-Morrow and Shake, 2009). Thus, one might expect that older heritage speakers are more proficient than younger ones. However, the findings in our study do not support this assumption since our results show that age does not affect heritage language proficiency. One explanation for the differences

<table>
<thead>
<tr>
<th>Table 4: Analysis of variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Squares</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5: Multiple regressions on language proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Family language use</td>
</tr>
<tr>
<td>Quantity of exposure</td>
</tr>
<tr>
<td>Age of onset</td>
</tr>
</tbody>
</table>
observed between our findings and the ones in the literature reported could lie in the differences in children’s ages. For example, in Mancilla-Martinez and Lesaux’s (2011) study children were very young (24, 30, and 36 months), and probably at this age the difference becomes prominent. However, it might be that at older ages, as were the children in our study (7–11), such differences are attenuated and no sharp changes in L1 language knowledge proficiency emerge. As Gathercole et al. (2013) themselves claim, slow rates of language acquisition among children who acquire two languages mainly affect the initial stages of language development. On the other hand, there is evidence that some linguistic domains are acquired at slower rates than others; referring again to Mancilla-Martinez and Lesaux’s (2011) study, the authors evaluate children’s vocabulary, a linguistic level which is known to experience an important growth at the ages evaluated (between 2 and 3 years of age). In our study, we assessed not only vocabulary but also the knowledge of morphology. This can further explain the lack of impact of the age factor in our results. Also, this finding appears to be supported in teaching practice: it is a common procedure among educational institutions to place heritage speakers in groups based on their background and their language proficiency, but not on their age (Kagan, 2005; Kagan and Dillon, 2008).

With regard to the potential role of gender in heritage language knowledge, previous research claimed that females are better than males both in L1 and L2 language acquisition (Larsen-Freeman and Long, 1991; Portes and Hao, 1998; Slik et al., 2015). In this study, the statistics did not demonstrate such an effect on the L1 proficiency of the children. However, measured qualitatively, female heritage speakers performed better than males. Perhaps more information (a larger dataset and more participants) would be needed to determine the gender gap in language proficiency of heritage speakers.

The current study supports other research findings (Jia and Aaronson, 2003; Montrul, 2016, 2008) in that there is an effect of age of onset to L2 on the linguistic development of heritage speakers. In concrete terms, the children who started acquiring L2 later (due to later migration or those who were born in only Russian-speaking families) had a higher proficiency in L1 than those who started acquiring L2 from birth. This means that growing up in an exclusively monolingual family or in a monolingual environment during the first years of life has a positive impact on heritage language proficiency in oral production.

Turning to the experience with the heritage language, the age of onset to the L2 is closely connected with the quantity of exposure to the L1. It has been also shown to correlate with heritage language proficiency in our study. Thus, we can suggest that the later the L2 acquisition starts, the more L1 input the child receives, and linguistic domains have more time to develop. The finding that
quantity of exposure positively affects heritage language proficiency is consistent with previous literature (Gathercole and Thomas, 2009; Thordardottir, 2011; Unsworth, 2008, 2013). In relation to this matter, several studies have demonstrated that language exposure is necessary and beneficial for children’s acquisition of vocabulary and grammar in both the minority and majority languages of heritage speakers (Cobo-Lewis, Pearson, Eilers, and Umbel, 2002; Scheele, Leseman and Mayo, 2010, among others). The results of our study show that regular contact with the Russian language in a variety of contexts and environments (at community school, at leisure, with friends, etc.) improves lexical and morpho-syntactic performance as well as basic narrative skills.

A noteworthy finding of the current study is that family language use is the main significant factor affecting Russian heritage language outcomes, as measured in this piece of research. This finding comes both from the correlation analyses and the regression model used for the identification of the heritage language proficiency predictors. In fact, it was the only significant variable of the model affecting language performance. This result is consistent with the findings from the studies of monolingual and bilingual child populations (De Houwer, 2007; Unsworth, 2015). It is of relevance to mention here that, in her study, Unsworth (2015) found that the frequency of English use at home was significantly correlated with language proficiency measures. She suggested that a potential interpretation for this finding can be the existence of multiple conversational partners who are exclusively English-speaking. It goes in line with our findings, given that the scores of family language use at home increase when both parents of the child are Russian native speakers.

Not only the parental language but also the frequency of heritage language use between siblings can affect L1 proficiency. Apparently, if the siblings address each other in Russian, the amount of family language use increases. Previous studies have found that siblings influenced the language environment in the bilingual homes: older siblings tend to use the language they are more proficient in with their family members (Bridges and Hoff, 2014; Obied, 2009; Stevens and Ishizawa, 2016). The qualitative data from the current study support these findings, showing that the children who have one or more siblings with whom they communicate in the heritage language are more proficient in Russian. Thus, the implementation that one can make is to pass the heritage language to the older sibling in the family to create a linguistic environment for the successful L1 learning for a younger one; if the older child speaks the heritage language at home, the younger one will most probably follow this linguistic practice. Another factor that may increase heritage language use at home is the presence of grandparents as child caregivers (Ishizawa, 2004). Even one grandparent in a household not only has
a positive effect on grandchildren’s heritage language abilities but also favors their cultural awareness: grandparents play a significant role in passing on the cultural and social traditions that differ from that of the dominant culture (Kamo, 1998). In Russian families, grandparents – mostly grandmothers – are traditionally the major providers of child care for preschool children. This fact makes them the main source of the Russian language for their grandchildren. The role of grandparents in child heritage language acquisition and development has not been studied in the present work and more research into that issue is necessary. All in all, it is implied in the present study that the language use at home and its result, language proficiency, are intertwined. Thus, consistent and frequent use of the Russian language is fundamental to successful heritage language maintenance. So, the main implication of this study for the Russian community in Spain is that the family should play a more proactive and systematic role in providing children with language support. The parents (or other involved stakeholders) should encourage their children to use the heritage language as much as possible inside and outside the home providing the children with a stimulating and nourishing environment for heritage language and culture learning. As far as the classroom is concerned, a proper placement procedure should be implemented based on the heritage language skills (and not on age). In turn, the heritage speaker’s language abilities may be assessed with the help of oral narratives (as in our study) or other relevant instruments created specifically for these speakers.

So far, the participants of the study are not only speakers of heritage Russian, but also multilinguals: they were born in bilingual environment with Spanish and Catalan as the dominant languages. Thus, input and output of dominant Spanish and Catalan are stronger than of Russian. The brief overlook on errors of the heritage speakers in the oral narratives demonstrated that some heritage speakers are prone to lexical transfer from Spanish and Catalan to Russian. On the morpho-syntactic level, we can observe simplification of the nominal declension system, probably because Spanish and Catalan do not show rich noun morphology. However, the morpho-syntactic errors could not be attributed solely to transfer from dominant languages (Cenoz, 2003; Butler, 2013) but also to incomplete acquisition or attrition of some linguistic features in Russian heritage language (Polinsky, 2006, 2008, 2011). To attest if the found errors in the narratives were due to incomplete acquisition or attrition, it would be interesting to compare the results of our study with those of adults or older children.

Finally, the results of this study show that age of onset to bilingualism, language exposure and, family language use accounted for 33% of the
variation of children’s language proficiency. That is, 67% of the variance cannot be explained by any of these variables so that it must be accounted for by other, yet unknown factors. Thus, this study raises the question that family language use alone is not in charge of successful heritage language acquisition. Even in households where the heritage language is spoken 100% of the time, some children have limited language abilities in the heritage language. Future research, probably including (but not limited to) attitudes and motivation of parents, language use and practices with siblings and grandparents, socio-economic status of the families, and a fine-grained examination of the at-home and outside-the-home environments could reveal the keys of successful heritage language acquisition and development.

Acknowledgments

This research has benefited from support by the Spanish Ministry of Economy and Competitiveness (Grant FFI2016-75082-P). Special thanks are due to our participants, since without them this work would not have been possible. We would also like to thank the Raduga Association and Istra school for their collaborative support.

References


