Developing reading and listening proficiency with Duolingo: A cross-sectional analysis

Xiangying Jiang∗, Lucy Portnoff∗, Luke Plonsky†, Bozena Pajak∗

Abstract

Online language courses have seen accelerated growth but evidence of their effectiveness based on standardized measures remains limited. The current study adds to this research by investigating the proficiency outcomes of Duolingo learners with ACTFL reading and listening proficiency tests at two points of learning. We tested a sample of learners who had US IP addresses (but might not all reside in the US) when they finished Unit 7 in Duolingo’s Spanish and French courses and compared them to learners who finished Unit 5 in the same courses (Jiang, Rollinson, Plonsky, Gustafson, & Pajak, 2021). The findings shed light on the learner progress between the two units. In Spanish, the average difference in performance between the two groups was one ACTFL sublevel in both reading and listening, going from Intermediate Low to Intermediate Mid in reading and from Novice Mid to Novice High in listening. In French, the difference between the groups was smaller than in Spanish, possibly due to higher performance at the end of Unit 5, but on average learners achieved Intermediate Mid in reading and Intermediate Low in listening at the completion of Unit 7. The proficiency outcomes at the completion of Unit 5 were comparable to those of US-based university students after four semesters of Spanish and French classes. At the completion of Unit 7, proficiency outcomes were comparable to those of US-based university students after five semesters, except for Spanish listening, which remained comparable to four semesters of university classes. Across most measures, the difference between Duolingo Spanish and French learners completing Unit 5 versus Unit 7 was comparable to those of US university students between the 4th and 5th semester.

Keywords

Key words: Duolingo, efficacy, reading and listening, proficiency development, foreign language

1 Introduction

As the prevalence of online language learning grows, so does the need to demonstrate the effectiveness of these systems. Duolingo is one of the language-teaching systems that offer free online courses available on mobile apps and the web. The effectiveness of Duolingo was shown in Jiang, Rollinson, Plonsky, Gustafson, and Pajak (2021), which reported the reading and listening proficiency outcomes of Duolingo learners after completing the beginner-level content of its Spanish and French courses (Units 1-5). The study focused on Duolingo learners who reported having little to no prior knowledge in the target language and using Duolingo as their only learning tool. Results from the ACTFL Reading Proficiency Test (RPT) and Listening Proficiency Test (LPT) demonstrated that on completion of Unit 5, these students reached Intermediate Low in reading comprehension and reached Novice Mid or Novice High in listening comprehension. As a follow-up, the current study sought to evaluate the listening and reading proficiencies of a different group of learners at the completion of Unit 7 in the same courses. The same participant selection criteria and data collection procedures were followed.

In this paper, we first report new Unit 7 data from US-based Duolingo learners, followed by a cross-sectional comparison with Unit 5 data (Jiang, Rollinson, Plonsky, et al., 2021). Unit 5 and Unit 7 are important milestones in the Duolingo courses as they correspond to a meaningful proficiency level aligned with the Common European Framework of Reference (CEFR) (see Figure 1). Specifically, Unit 5 indicates the end of its A2 content, while Unit 7 marks the first half of its B1 content (see Table 1). Please note that Duolingo has recently transitioned to a new home screen design of its course structure (see Figure 2). The data for the current study, however, was collected when the participants were using the original version (see a brief comparison between the two versions in the next section). In addition, this study benchmarked our new Unit 7 results against the proficiency data of US-based university students across semesters, following Jiang, Rollinson, Plonsky, et al. (2021). The university proficiency data used for comparison in Jiang, Rollinson, Plonsky, et al. (2021) were the foreign language proficiency test data of Winke, Gass, Soneson, Rubio,

∗ Duolingo, Inc.
† Northern Arizona University

Corresponding author:
Xiangying Jiang
Duolingo, Inc. 5900 Penn Ave
Pittsburgh, PA 15206, USA
Email: assessment-study@duolingo.com

© 2022 Duolingo, Inc. All rights Reserved.
and Hacking (2014-2017). Jiang, Rollinson, Plonsky, et al. (2021) found that at the completion of Unit 5 in the Duolingo Spanish and French courses, the proficiency scores of the Duolingo participants were comparable to those of US-based university students after taking four semesters of Spanish or French. In the current study, we compared Duolingo Unit 7 data with the scores of US-based university students who had completed five semesters of Spanish or French courses. We used the summary data reported in Tschirner (2016) for comparison because French scores of the fifth-semester students were lacking in Winke et al. ’s (2014-2017) foreign language proficiency test data. Although the proficiency data of university students were used for external benchmarking, the demographics of the university students and Duolingo learners are different in several respects. For example, the Duolingo learners were older, overall, and exhibited a larger range of ages. In addition, many of the Duolingo learners possessed a university degree and brought to the task of learning more varied motivations than the university students. In addition, although the Duolingo group all used US-based IP addresses, we could not confirm their physical location.

In sum, the goal of this paper is threefold: 1) to report the proficiency outcomes of US-based Duolingo learners at the completion of Unit 7, 2) to explore the progress learners may be able to make between Unit 5 and Unit 7 of the Duolingo Spanish and French courses, and 3) to benchmark the Unit 5 and Unit 7 data against semester-based proficiency scores of US-based university students taking foreign language courses. Before we dive into the details about the current study, we first briefly introduce the Duolingo course structure and its alignment with CEFR.

2 Duolingo Course Structure and CEFR Alignment

The Duolingo Spanish and French courses (from the English user interface) are aligned with the Common European Framework of Reference (CEFR), an international standard for describing the abilities of language learners at various levels of proficiency. The CEFR divides language proficiency into three broad levels – A (Basic User), B (Independent User), and C (Proficient User) (Council of Europe, 2001). Each broad level is then further divided into two levels, namely, A1 and A2, B1 and B2, and C1 and C2 (see Figure 1).

As mentioned earlier, the data for the current study was collected when the participants were using the original version of Duolingo’s course structure. To help readers see how research findings based on the original version are still applicable to the new version of the course structure, a brief comparison of the two versions (see Figure 2) follows.

In the original version of the Duolingo course structure, each section of the course concludes with a ‘checkpoint.’ Each circle in the original home screen represents a “skill,” which includes 4-5 lessons at 5 difficulty levels (“crown” levels), where higher difficulty is achieved through exercises requiring progressively more recall and production. Learners are required to complete at least one difficulty level in each row to move on to the next row.

In the new version of the course structure, a circle on the path is roughly equivalent to one “crown” level (difficulty level) in the original version. The content of each original “skill” is now condensed into three levels in the new structure, and levels from different “skills” are interspersed throughout the path together with additional circles for review, as well as reading and listening practice. The content in the new path is organized into small units. See this blog post (Munson, Yu, Rajgarhia & Noh, 2022) for information about the new Duolingo learning path and the differences between the original and new designs of the home screen.
Table 1 shows the correspondence between sections and units in the two versions of the Spanish and French courses. As shown in the table, no changes were made to the scope of the CEFR-aligned course sections. The main difference is in the organization of the content. For example, the content through A2 was taught in five units in the original tree structure but approximately 80 units in the new path structure. Duolingo’s Spanish and French courses (from the English user interface) currently have content through B1. The shaded rows in Table 1 represent the sections that were evaluated in this study.

The lessons in the Spanish and French courses include several activity types targeting learning and practice in vocabulary, grammar, reading, listening, writing, and speaking. In addition to word- and sentence-based activities, Duolingo lessons include activity types that rely on comprehension of short passages or dialogues, which increase the amount of input in the target language. To facilitate listening and speaking development, Duolingo provides learners with many opportunities to listen to the target language and speak it out loud. All course content is accompanied by audio and learners are allowed to play the audio at varied speeds as often as they need. In addition, speech recognition technology is used for all speaking exercises in order to provide learners with feedback.

Lessons are the primary method of teaching new material on Duolingo, but other modes of learning are available outside of the main course structure. For example, learners can complete generalized practice sessions, which review content they have studied throughout the entire course. For “skill”-specific practice, learners can return to any “skill” for which they have completed all difficulty levels in order to refresh their knowledge of a particular functional topic or grammar concept. Another relevant feature is Stories, which provides discourse-level reading and listening comprehension practice, reinforcing and enriching learners’ knowledge by situating the lesson content in everyday contexts. Due to the large degree of user autonomy in navigating the platform, there is considerable variation in the types of sessions that learners choose to complete.

As a result, there can be substantial variation among individual learners on both the percentage of the course material they engage with, as well as on the total amount of time spent learning.

3 The Current Study

This study had three goals. The first goal was to measure the reading and listening proficiencies of Duolingo learners after completing Unit 7 in its Spanish and French courses. The second goal was to use these findings to quantify the progress learners can make between Unit 5 and Unit 7. The final goal was to compare the proficiency levels of Duolingo learners at both units with those of US-based university students. Our interest is not in demonstrating the superiority of one approach or context of learning but, rather, to benchmark Duolingo’s curriculum to language learning that takes place in a more familiar setting. In particular, the current study answered the following research questions:

1. What levels of reading and listening proficiency did US-based Duolingo learners achieve when they completed Unit 7 of the Spanish and French courses?
2. How did the reading and listening levels of US-based Duolingo learners compare cross-sectionally after completion of Unit 5 and Unit 7 of the Spanish and French courses?
3. How did the reading and listening proficiency levels of US-based Duolingo learners at both points compare with proficiency levels of US-based university students at various semesters of Spanish and French courses?

4 Methods

4.1 Participants

The participants of the study were 240 learners who had completed Unit 7 (144 Spanish and 96 French learners). They were learners who: 1) were aged 18 or above; 2) had an IP Address in the United States; 3) had self-reported little or
no prior proficiency in the target language; 4) reported using Duolingo as the only tool to learn the target language; and 5) had the necessary computer equipment for online testing. Some of the data reported here (from 59 Spanish and 47 French learners) was also included in a previous study (Jiang, Chen et al., 2021), which focused on both US-based and non-US-based learners. The current study focuses on US-based learners only.

This study used a combination of app-recorded data and responses to background survey questions to select participants who met all these criteria. With regard to prior proficiency in the target language, we included learners who self-reported prior proficiency of 0-2 on a 0-10 scale, with 0 meaning “I have no knowledge of the language at all,” and 10 indicating “I have perfect knowledge of the language.” Note that Duolingo collects this information from all learners at the beginning of the course for the purposes of learner analytics, and not for course placement.

Demographic and other background information were collected through a survey (see Appendix A). Some general characteristics of our participants include the following: The average age of the participants learning Spanish was about 48 and that of those learning French was about 40. In terms of gender, about 43% of the participants learning Spanish identified themselves as female and about 57% as male, while the percentage of French learners was 38% female and 61% male. About 75% of the participants listed their ethnicity as Caucasian, 13% as Asian, and 3% as African American. Twenty-nine percent of the participants reported having a bachelor’s degree as their highest level of education, 40% having a master’s degree, and 17% having a doctoral degree. Finally, 73% of the participants reported speaking only English before age 6; 8% were early bilingual speakers of English and another language; and 19% of the participants did not speak English before age 6 (their first languages varied widely and only one participant was a heritage speaker of Spanish). For a detailed by-course description of participant background information, see Appendix B.

Information about participants who were tested at the completion of Unit 5 can be found in Jiang, Rollinson, Plonsky, et al. (2021). There were a total of 225 US-based learners (135 Spanish and 90 French learners). The Unit 5 and Unit 7 participants met the same selection criteria described above, used the same instruments, and followed the same data collection procedures described below. Data collection at the completion of Unit 5 took place in 2020 and data at the completion of Unit 7 were collected in 2021. There was no overlap between participant samples across the two studies, except for one participant who happened to be assessed at both points in the course.

5 Instruments

5.1 The Background Survey

The background questionnaire (see Appendix A) included questions related to participants’ language background, reasons for learning the language, level of education, age, gender, whether they took classes or used other programs/apps during the time they learned on Duolingo, and whether they had the appropriate computer equipment for test-taking. The latter questions confirmed eligibility to satisfy criteria #4 and #5 for participant selection; see Participants above.

5.2 ACTFL Listening Proficiency Test (LPT) and Reading Proficiency Test (RPT) in Spanish and French

The ACTFL LPT and RPT are standardized tests for the global assessment of reading and listening ability (ACTFL, 2013, 2014). They measure how well test-takers spontaneously comprehend the texts and discourse they read or listen to as described in the ACTFL Proficiency Guidelines (ACTFL, 2012). The ACTFL proficiency scale includes four broad levels: Novice, Intermediate, Advanced, and Superior. Each of the first three levels includes three sublevels: low, mid, and high. Altogether, there are ten levels in its proficiency rating scale, from low to high in the order of Novice (low, mid, high), Intermediate (low, mid, high), Advanced (low, mid, high), and Superior. We used Form E of the tests in this project, which targeted proficiency levels between Novice Low and Advanced Low. The tests were administered to each participant online by a remote proctor. Participants were asked to read or listen to 15 passages and answer three multiple-choice questions after each passage in the target language. Each test was given an ACTFL rating immediately after the test was submitted. In line with previous research involving ACTFL ratings (e.g., Isbell, Winke, & Gass, 2019; Loewen, Isbell, & Sporn, 2020; Rubio & Hacking, 2019; Tschirmer, 2016), we coded each sublevel numerically following a 1-10 point scale. Please note that the ACTFL proficiency scale is not an interval scale with equal distance between sublevels. See Table 2 for the mapping between the point scale and each sublevel.

5.3 Data Collection Procedures

Shortly after completing Unit 5 or Unit 7 in Duolingo’s Spanish and French courses, a random sample of learners whose IP address was in the US and whose self-reported prior proficiency in the language was 0-2 received an email which invited their participation in the research study. Learners aged 18 and above who were interested in participating completed a background survey to verify eligibility and provide additional demographic information. Learners who responded that they had taken classes or used other programs/apps to learn the language during the time they used Duolingo were disqualified, and so were those who did not have access to the appropriate testing equipment.
Qualified participants were emailed on a rolling basis and invited to take the ACTFL reading and listening proficiency tests one at a time, with the order of tests (reading, listening) randomized across participants. The participants were first contacted by Duolingo researchers and then received an email from Language Testing International (LTI), the organization that administers the LPT and RPT, with their test ID and instructions about how to schedule a time for the test. The participants had two weeks to finish the first test. After the participants finished the first test, we ordered the second test for them and they were again contacted by LTI to take the second test. They went through the same process to schedule and take the second test within two weeks. Each participant was paid $100 after completing both tests.

5.4 Analyses

Descriptive statistics and t-tests were used to answer the first and second research questions on the proficiency outcomes of Duolingo learners after Unit 5 and Unit 7. Note that the data from Unit 5 assessments have been reported in Jiang, Rollinson, Plonsky, et al. (2021).

For the third research question on the comparison of reading and listening proficiency outcomes between US-based university students and US-based Duolingo learners, the university students’ data were from Tschirmer (2016, p. 211). Tschirmer (2016) provided ACTFL listening and reading proficiency scores at the end of the 2nd, 3rd, 4th, 5th, 6th, and 8th semester of their undergraduate study from more than 3,000 participants learning seven languages at 21 institutions across the United States. The main findings were reported based on listening and reading proficiency levels in Spanish and French, which made up 82% of all tests completed. For the purpose of comparison, descriptive statistics and a series of t-tests were conducted.

6 Results

6.1 Reading and Listening Proficiency Outcomes of US-based Duolingo Learners at Unit 7

Table 3 shows the average proficiency scores of US-based Duolingo learners when they reached the end of Unit 7 in its Spanish and French courses. Based on the numerical coding of ACTFL ratings presented in Table 2, a score of 4 corresponds to Intermediate Low in the ACTFL proficiency scale and a score of 5 corresponds to Intermediate Mid. The mean scores in Table 3 indicate that, on average, Duolingo learners achieved Intermediate Mid in reading in both Spanish and French, reached Intermediate Low in French listening, and were at Novice High in Spanish listening.

6.2 Cross-sectional Comparisons of US-based Duolingo Learners at Unit 5 and Unit 7

As mentioned earlier, the Unit 5 data was published in Jiang, Rollinson, Plonsky, et al. (2021). To answer the second research question about differences between the beginner and intermediate sections of the courses, we compiled the Unit 5 and Unit 7 data for comparison purposes.

Table 4 shows the number of learners who scored at each ACTFL sublevel. The levels range from Novice Low to Advanced Low based on the test form (Form E) used in the study, with no Under Rating or Above Rating scores. The score distributions are visually represented in Figure 3.

The mean scores and standard deviations at Unit 5 and Unit 7 for reading and listening of Spanish and French are shown in Table 5. The results indicate that, on average, Duolingo learners of Spanish reached Intermediate Low in reading and were at
Table 4. Number of Participants at Each ACTFL Proficiency Rating (See Table 2 for Rating Abbreviations)

<table>
<thead>
<tr>
<th></th>
<th>NL</th>
<th>NM</th>
<th>NH</th>
<th>IL</th>
<th>IM</th>
<th>IH</th>
<th>AL</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spanish reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 5</td>
<td>4</td>
<td>9</td>
<td>16</td>
<td>44</td>
<td>42</td>
<td>7</td>
<td>10</td>
<td>132</td>
</tr>
<tr>
<td>Unit 7</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>42</td>
<td>16</td>
<td>56</td>
<td>136</td>
</tr>
<tr>
<td><strong>Spanish listening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 5</td>
<td>35</td>
<td>29</td>
<td>24</td>
<td>18</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>131</td>
</tr>
<tr>
<td>Unit 7</td>
<td>14</td>
<td>25</td>
<td>27</td>
<td>15</td>
<td>32</td>
<td>14</td>
<td>14</td>
<td>141</td>
</tr>
<tr>
<td><strong>French reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 5</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>16</td>
<td>25</td>
<td>13</td>
<td>16</td>
<td>88</td>
</tr>
<tr>
<td>Unit 7</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>28</td>
<td>35</td>
<td>13</td>
<td>14</td>
<td>95</td>
</tr>
<tr>
<td><strong>French listening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 5</td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>43</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>89</td>
</tr>
<tr>
<td>Unit 7</td>
<td>6</td>
<td>9</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>9</td>
<td>4</td>
<td>93</td>
</tr>
</tbody>
</table>

Figure 3. Score distributions of Duolingo learners at the completion of Unit 5 and Unit 7.

Novice Mid in listening after completing Unit 5. In comparison, Duolingo learners who completed Unit 7 reached Intermediate Mid in reading and were at Novice High in listening. The difference between Unit 5 and Unit 7 was one ACTFL sublevel in both reading and listening.

On the other hand, Duolingo learners of French reached Intermediate Low in reading and Novice High in listening after completing Unit 5, which is much higher than those of Spanish learners after the same unit. After completing Unit 7, learners’ average scores reached Intermediate Mid in reading and Intermediate Low in listening; however, the difference between Unit 5 and Unit 7 was less substantial compared to learners of Spanish, partially due to higher learner performance after Unit 5, as indicated in the statistical comparisons explained in the following two paragraphs.
Table 5. Comparison between Duolingo learners at Unit 5 with Duolingo learners at Unit 7

<table>
<thead>
<tr>
<th>Language Skill</th>
<th>Study</th>
<th>N</th>
<th>Mean score (SD)</th>
<th>Difference in ACTFL sublevels</th>
<th>Adjusted p-value</th>
<th>Effect size (Hedges’ $g$)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish reading</td>
<td>Unit 5</td>
<td>132</td>
<td>4.30 (1.34)</td>
<td>1.39</td>
<td>$p &lt; 0.01$</td>
<td>1.04</td>
<td>[.78, 1.29]</td>
</tr>
<tr>
<td></td>
<td>Unit 7</td>
<td>136</td>
<td>5.69 (1.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish listening</td>
<td>Unit 5</td>
<td>131</td>
<td>2.80 (1.54)</td>
<td>1.08</td>
<td>$p &lt; 0.01$</td>
<td>.64</td>
<td>[.39, .88]</td>
</tr>
<tr>
<td></td>
<td>Unit 7</td>
<td>141</td>
<td>3.88 (1.82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French reading</td>
<td>Unit 5</td>
<td>88</td>
<td>4.82 (1.55)</td>
<td>.19</td>
<td>$p &gt; 0.05$</td>
<td>.14</td>
<td>[-.15, .43]</td>
</tr>
<tr>
<td></td>
<td>Unit 7</td>
<td>95</td>
<td>5.01 (1.16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French listening</td>
<td>Unit 5</td>
<td>89</td>
<td>3.61 (1.22)</td>
<td>.43</td>
<td>$p &lt; 0.05$</td>
<td>.32</td>
<td>[.02, .61]</td>
</tr>
<tr>
<td></td>
<td>Unit 7</td>
<td>93</td>
<td>4.04 (1.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5 also summarizes the results of statistical comparisons. We compared Duolingo participants after Unit 5 with Duolingo participants after Unit 7 to evaluate the differences in their outcomes. For each comparison, we carried out separate Welch’s two-sample t-tests using the R statistical package (R Core Team, 2020) and adjusted the p-value using the Benjamini-Hochberg method with alpha = 0.05.

We found that participants in Unit 7 scored significantly higher than participants in Unit 5 in Spanish reading ($p < 0.01$, $g = 1.04$), Spanish listening ($p < 0.01$, $g = .64$), and French listening ($p < 0.05$, $g = .32$), with effect sizes ($g$) ranging from small in the case of French listening to moderate and large for Spanish listening and reading, respectively (Plonsky & Oswald, 2014). The difference between Duolingo participants in Unit 5 and Unit 7 for French reading was very small and not statistically significant, but directionally consistent with the other results.

6.3 Benchmarking with the Proficiency Outcomes of US-based University Language Courses

As mentioned earlier, data on the proficiency outcomes of US-based university students were the summary data reported in Tschirner (2016, p. 211). First, we report the comparisons between US-based Duolingo learners and university students on reading and listening proficiency scores of Spanish, and the analysis of French learners follows.

6.3.1 Comparisons on Spanish Reading and Listening Proficiency Scores

Table 6 shows the descriptive statistics of Duolingo learners at the end of Unit 5 and Unit 7 in the Spanish course and of university students at the end of the above-mentioned semesters in Spanish reading and listening. The information in Table 6 is visually represented in Figure 4 below.

The visual representation illustrates how the scores of Duolingo learners at the completion of Unit 5 were similar to those of university students at the end of the 4th semester of classes, while scores of Duolingo learners at the completion of Unit 7 were similar to those of university students at the end of the fifth semester of classes in both reading and listening.

Table 7 summarizes four pairs of comparisons on Spanish reading and Spanish listening. To determine whether there were statistically significant differences, we compared Duolingo learners at the end of Unit 5 with 4th-semester university students, and Duolingo learners at the end of Unit 7 with 5th-semester university students, for both reading and listening. We carried out four separate Welch’s two-sample t-tests using the R statistical package (R Core Team, 2020) and adjusted the p-value using the Benjamini-Hochberg method with alpha = 0.05.

For the comparison between Unit 5 and 4th-semester scores, no significant difference was found between Duolingo study participants and 4th-semester university students in Spanish reading or listening. For the comparison between Unit 7 and 5th-semester scores, Duolingo participants scored significantly higher than university students in Spanish reading ($p < .05$, $g = .19$) but significantly lower than the 5th-semester university students in Spanish listening ($p < .05$, $g = -.22$). In both of these comparisons, however, the magnitude of the differences was quite small as indicated in the effect size indices ($g = .19$ and -.22, respectively).

In sum, the scores of Duolingo learners who completed Unit 5 were comparable to those of university students at the end of four semesters of classes on both Spanish reading and listening based on ACTFL proficiency tests. Duolingo learners who completed Unit 7 scored significantly higher than university students at the end of five semesters of classes in Spanish reading but significantly lower in Spanish listening, although the effect sizes were very small.

6.3.2 Comparisons on French Reading and Listening Proficiency Scores

Table 8 shows the descriptive statistics of Duolingo learners in the French course at the end of Unit 5 and Unit 7, and university students at the end of the above-mentioned semesters in French
Table 6. Descriptive Statistics for ACTFL Sublevels in Spanish Reading and Listening for Duolingo Learners and University Students

<table>
<thead>
<tr>
<th>Study</th>
<th>Spanish reading</th>
<th>Spanish listening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Duolingo Unit 5</td>
<td>132</td>
<td>4.30 (1.34)</td>
</tr>
<tr>
<td>Duolingo Unit 7</td>
<td>136</td>
<td>5.69 (1.33)</td>
</tr>
<tr>
<td>University 2nd semester</td>
<td>242</td>
<td>3.11 (1.35)</td>
</tr>
<tr>
<td>University 3rd semester</td>
<td>222</td>
<td>3.37 (1.43)</td>
</tr>
<tr>
<td>University 4th semester</td>
<td>338</td>
<td>4.09 (1.64)</td>
</tr>
<tr>
<td>University 5th semester</td>
<td>432</td>
<td>5.26 (2.45)</td>
</tr>
<tr>
<td>University 6th semester</td>
<td>208</td>
<td>6.30 (2.16)</td>
</tr>
<tr>
<td>University 8th semester</td>
<td>121</td>
<td>6.94 (1.87)</td>
</tr>
</tbody>
</table>

Figure 4. Visual representation of ACTFL Spanish Reading and Listening proficiency scores for Duolingo learners and university students. The blue vertical bars with 95% confidence intervals (CIs) indicate the average scores at each university semester, while the horizontal green lines with CIs represent the average scores of Duolingo learners at Unit 7 (upper line) and Unit 5 (lower line).

Table 7. Summary of T-tests between Duolingo Learners and University Students on Spanish Reading and Listening

<table>
<thead>
<tr>
<th>Language Skill</th>
<th>Study</th>
<th>N</th>
<th>Mean score (SD)</th>
<th>Adjusted p-value</th>
<th>Effect size (Hedges’ g)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish reading</td>
<td>Duolingo Unit 5</td>
<td>132</td>
<td>4.30 (1.34)</td>
<td>p &gt; 0.05</td>
<td>.13</td>
<td>[-.07, .32]</td>
</tr>
<tr>
<td></td>
<td>University Semester 4</td>
<td>338</td>
<td>4.09 (1.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish reading</td>
<td>University Semester 5</td>
<td>432</td>
<td>5.26 (2.45)</td>
<td>p &lt; 0.05</td>
<td>.19</td>
<td>[.00, .86]</td>
</tr>
<tr>
<td>Spanish listening</td>
<td>Duolingo Unit 5</td>
<td>131</td>
<td>2.80 (1.54)</td>
<td>p &gt; 0.05</td>
<td>-.02</td>
<td>[-.22, .18]</td>
</tr>
<tr>
<td></td>
<td>University Semester 4</td>
<td>317</td>
<td>2.83 (1.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish listening</td>
<td>Duolingo Unit 7</td>
<td>141</td>
<td>3.88 (1.82)</td>
<td>p &lt; 0.05</td>
<td>-.22</td>
<td>[-.42, -.03]</td>
</tr>
<tr>
<td></td>
<td>University Semester 5</td>
<td>392</td>
<td>4.39 (2.44)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2022 Duolingo, Inc
Table 8. Descriptive Statistics for ACTFL Sublevels in French Reading and Listening for Duolingo Learners and University Students

<table>
<thead>
<tr>
<th></th>
<th>French reading</th>
<th>French listening</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Duolingo Unit 5</td>
<td>88</td>
<td>4.82 (1.55)</td>
</tr>
<tr>
<td>Duolingo Unit 7</td>
<td>95</td>
<td>5.01 (1.16)</td>
</tr>
<tr>
<td>University 2nd semester</td>
<td>120</td>
<td>2.94 (1.27)</td>
</tr>
<tr>
<td>University 3rd semester</td>
<td>86</td>
<td>3.63 (1.28)</td>
</tr>
<tr>
<td>University 4th semester</td>
<td>215</td>
<td>4.52 (1.45)</td>
</tr>
<tr>
<td>University 5th semester</td>
<td>166</td>
<td>5.29 (2.20)</td>
</tr>
<tr>
<td>University 6th semester</td>
<td>62</td>
<td>6.65 (1.52)</td>
</tr>
<tr>
<td>University 8th semester</td>
<td>124</td>
<td>6.56 (1.83)</td>
</tr>
</tbody>
</table>

Figure 5. Visual representation of ACTFL French Reading and Listening proficiency scores for Duolingo learners and university students. The blue vertical bars with 95% confidence intervals (CIs) indicate the average scores at each university semester, while the horizontal green lines with CIs represent the average scores of Duolingo learners at Unit 7 (upper line) and Unit 5 (lower line).

Given the proximity of the Duolingo scores to the university scores after the 4th and 5th semesters of study, we also compared Duolingo French learners at the end of Unit 5 with 4th-semester university students, and at the end of Unit 7 with 5th-semester university students (see Table 9). For each comparison, we carried out two separate Welch’s two-sample t-tests (one for reading and one for listening) using the R statistical package (R Core Team, 2020) and adjusted the p-value using the Benjamini-Hochberg method with alpha = 0.05.

For the comparison between Unit 5 and four semesters, no significant difference was found between Duolingo study participants and 4th-semester university students in French reading. However, Duolingo participants scored significantly higher than university students in French listening (p < .05, g = .67), with a moderate effect size. For the comparison between Unit 7 and five semesters, Duolingo participants scored significantly higher than university students in French listening (p < .05, g = .26), although the effect size was small. No significant difference was found between Duolingo study participants and 5th-semester university students in French reading.

7 Discussion

7.1 Summary of Findings

This study sought to answer three research questions. The first research question investigated the reading and listening proficiency levels of US-based Duolingo learners when they reached the end of Unit 7. The findings indicated that at the completion of Unit 7, the first half of the Duolingo B1 course content, Duolingo learners, on average, reached
Intermediate Mid in Spanish reading and French reading, reached Intermediate Low in French listening and Novice High in Spanish listening, based on ACTFL’s proficiency tests and rating scale.

The second research question compared the reading and listening levels of US-based Duolingo learners at the completion of either Unit 5 or Unit 7. The findings indicated that compared to the scores of Duolingo learners at the completion of Unit 5, Duolingo Spanish learners at the end of Unit 7 scored substantially higher – nearly a full ACTFL sublevel with moderate to large effect sizes. Duolingo French learners showed a significant difference in listening across Unit 5 and Unit 7, but the effect size was small. No significant difference was observed in French reading, which was already Intermediate Low at the completion of Unit 5. Although the design that produced these findings was cross-sectional rather than longitudinal, the different patterns observed in this study might reflect distinct stages of acquisition that learners of these two languages progress through.

The by-semester university data itself is a demonstration of the cross-sectional development of reading and listening proficiency of US-based university students taking Spanish and French classes. In the third research question, we examined Duolingo learners’ proficiency development in relation to the proficiency scores of US-based university students. At the completion of Unit 5, Duolingo Spanish learners’ average reading and listening proficiency scores were not significantly different from those of the university students at the end of the fourth semester; at the completion of Unit 7, Duolingo Spanish learners scored significantly higher in reading but significantly lower in listening compared to university students at the end of the fifth semester, both with very small effect sizes. At the completion of Unit 5, Duolingo French learners’ average reading proficiency scores were not significantly different from those of the university students at the end of the fourth semester, and their listening scores were significantly higher. Similarly, at the completion of Unit 7, Duolingo French learners’ average reading proficiency scores were not significantly different from those of the university students at the end of the fifth semester, but their listening scores were significantly higher. Although the comparisons were with the 4th- and 5th-semester university students, the university dataset includes data at six points of learning in the developmental timeline; the Duolingo data, however, only includes two points of learning: Unit 5 and Unit 7. For future research, it will be important to track the developmental pattern of Duolingo learners as they study more advanced content, for example, at the end of the B1 content (Unit 9) and later in the B2 material.

One pattern that can be inferred from the comparisons (Duolingo Unit 5 vs. university semester 4 and Unit 7 vs. semester 5) is that the French listening proficiency of Duolingo learners potentially developed faster than that of university students. While Spanish listening scores increased every semester in the university data, reaching Intermediate High in semester 8, French listening scores among the university students increased until semester 6, but then regressed and remained at Intermediate Low for semester 8. More research is needed to help understand the disparity between Spanish and French in listening development among university students, but one hypothesis is that this might be due to changes in the student population in later semesters in French (e.g., more students placed directly into semester 7).

Another pattern we observed across Duolingo learners and university students is that learners’ listening proficiency was much lower than their reading proficiency. Although both listening comprehension and reading comprehension are receptive skills, the two types of comprehension processes have been found to be mostly different (Wolf, Muijselaar, Boonstra, & de Bree, 2019). At the early stages of language learning, listening comprehension demands a higher level of attention, exerts a heavier load on working memory,
and requires speedy decoding and processing of transient audio input (see, e.g., Bloomfield et al., 2010; Wallace, 2022). In contrast, learners’ decoding process in reading is facilitated by the presence of visually-displayed text (Spoden, Fleischer, & Leucht, 2020; Vandergrift & Baker, 2015). As a result, listening comprehension is often more challenging than reading comprehension for second language learners and takes longer to develop. Tschirner (2016) also observed that listening lagged behind reading and attributed students’ lower listening proficiency to insufficient attention to auditory input in classroom instruction and called for more emphasis on listening development in instructional practices.

For the purpose of comparison with US-based university language programs, this study only reported data from Duolingo learners within the United States, but Duolingo Spanish and French courses (from the English user interface) serve learners all over the world. Jiang, Rollinson, Plonsky, et al. (2021) reported the reading and listening proficiency of global (both US and non-US) Duolingo learners at the completion of Unit 7. While there was no difference in the proficiency of French reading and listening between the US and global samples of Duolingo learners, the Spanish reading and listening proficiency of the global sample was much higher than the US sample reported in the current study. Since Duolingo Spanish and French courses serve global learners, it might not be necessary to restrict future research on Duolingo efficacy to US-based learners, unless the comparisons are with US-based learner groups. However, we also recognize that many non-US-based Duolingo users may live in multilingual contexts which can provide numerous linguistic affordances which limit comparisons to university-based learners.

The availability of the university proficiency data in Tschirner (2016) made the comparison of US-based Duolingo learners and US-based university students possible. The aim of comparing learning outcomes from the two contexts was to benchmark the progress made by Duolingo learners in relation to a more familiar and traditional setting. It is worth noting that there are important differences between the Duolingo study participants and the university student sample. First, the university sample consisted of mostly full-time students from a more homogeneous age range of approximately 18-22, while the Duolingo study participants were mostly post-university adults with an average age of 40 and higher. These differences may confer an advantage for the Duolingo-based sample. Second, as reported in Winke et al. (2014-2017), the university student sample likely included students with multiple years of K-12 language learning experience or heritage status in the target language, while the Duolingo study participants were restricted to those who had a prior self-reported proficiency of 0-2 (on a scale of 0-10, from no knowledge to full proficiency) and used Duolingo as their only learning tool. Finally, the motivations for language learning also differed. The university students took the language courses for credit or degree requirements, while the majority of the Duolingo study participants were learning the language for fun or leisure, for travel, for family and social reasons, and for memory or brain acuity.

7.2 Limitations and Future Directions

This study has several limitations. First and foremost, the inferences about proficiency development across Unit 5 and Unit 7 were based on analyses of cross-sectional instead of longitudinal data. In other words, with the exception of one individual, the participants at Unit 5 and those at Unit 7 were different groups of learners. There may be a test-retest effect in that one participant. However, given the size of the sample, we believe that any such effect would have a minimal impact, if any, on our findings. The differences we observed might have been due to unrelated differences in the specific learner samples we tested. In the future, longitudinal data testing the same learners at both points would allow us to better control for such participant-related factors.

Although the design of the study presents some level of ecological validity due to the fact that the participants reached the end of Unit 5 or Unit 7 independently and had a high degree of autonomy during the learning process, for future research, a pre- and post-test design will allow more control of learning time and participant factors that were self-reported in the present study. These include prior proficiency, exposure to the target language outside of Duolingo, and the exclusion of other learning tools.

Furthermore, the reading and listening proficiencies assessed in the study are both receptive skills. Learners were not assessed in productive skills such as speaking (as in Jiang, Rollinson, Chen, et al., 2021) and writing, or overall proficiency. In future studies, Duolingo’s effectiveness in developing learners’ productive skills or overall proficiency should be evaluated as well. Doing so will provide a better understanding of whether and to what extent Duolingo learners’ success in receptive skills can also be observed in productive skills or overall proficiency.

The findings of the current study should not be overgeneralized because they do not represent the overall effectiveness of Duolingo or university language courses. Participants of the current study were only compared on reading and listening skills, while teaching effectiveness can be reflected in other skills and abilities such as speaking and writing. It is important to consider the comparability—and perhaps complementarity—of Duolingo courses to university classes in developing skills and abilities in addition to reading and listening.

Finally, the comparison between Duolingo learners and university students should not be interpreted as competition between online language learning apps and university language programs. The aim in comparing learning outcomes from the two contexts is, rather, as a means to benchmark the
progress made by Duolingo learners relative to an external, more familiar, and traditional setting. As previously mentioned, the demographics of the Duolingo learners and university students differed to some extent. Although the Duolingo participants had IP addresses within the United States, they might not be 100% US learners. Similarly, the American university students included international students and heritage speakers of Spanish/French. The university students were from a more homogeneous age range, while the Duolingo participants were more varied demographically and included mostly post-university adults. Moreover, the Duolingo participants’ motivations for language learning could also be more varied than university students, who at Semester 4, included both those studying to meet a requirement and some who would later declare majors and minors in the language, and at Semester 5, included mostly majors and minors of the language.

7.3 Pedagogical Implications

The results of the current study suggest that Duolingo can help self-directed learners develop reading and listening proficiency. The proficiency outcomes of Duolingo learners were comparable to those of US-based university students in their fourth semester of lower-division or fifth semester of upper-division language programs. Although Duolingo courses mostly teach vocabulary and grammar at the sentence level (with some longer-form content available in the form of short stories and podcasts), the results of this study also suggest that seemingly discrete vocabulary and grammar knowledge can be applied to integrative tasks such as listening and reading comprehension.

In addition to self-directed learners, classroom teachers have used Duolingo to their advantage and to the benefit of their students (Munday, 2016, 2017), suggesting that the app is also a useful tool to complement other types of language instruction. For instance, if vocabulary and grammar practice can be largely done by students as homework using apps such as Duolingo, more classroom time can be directed toward the teaching of culture and other communicative skills.

7.4 Conclusion

This study evaluated the reading and listening proficiency outcomes of Duolingo learners who had little to no prior knowledge of the target language and used Duolingo as their only learning tool. The findings demonstrated that, on average, learners who finished the first seven units of the Duolingo Spanish or French course reached ACTFL Intermediate levels in both reading and listening proficiency. This finding indicates considerable progress compared to those of the learners who were assessed after finishing the first five units of the same courses. The proficiency scores of Duolingo learners at the completion of Unit 5 were comparable to those of students at the end of their fourth semester. At the completion of Unit 7, Duolingo learners’ scores were comparable with those of university students at the fifth semester in upper-division language programs. In conducting this cross-sectional analysis, we hope to have shed light on the potential effectiveness and comparability of Duolingo, as measured through standardized tests, to more traditional settings. Future studies will continue to build on our findings at other levels of study, in other linguistic domains, and in other target languages.

Notes

1. The two data sources overlap to some degree but they are different. Winke et al. (2014-2017) is a three-year collection of ACTFL reading, listening, and speaking proficiency ratings of students from three American universities, i.e., Michigan State University, University of Minnesota, and University of Utah at Semesters 2, 3, 4, 5, 6, 7, 8 in seven languages, while Tschirner (2016) reported ACTFL listening and reading proficiency scores of undergraduate students at 2nd, 3rd, 4th, 5th, and 6th semester, 4th year, and 2 years abroad from more than 3,000 participants learning seven languages at 21 institutions across the United States. The majority (83% of reading and 88% of listening scores) of the data reported in Tschirner (2016) came from the same three universities in Winke et al. (2014-2017), which indicates the overlap. The main findings of Tschirner (2016) were reported based on the listening and reading proficiency levels in Spanish and French, which made up 82% of all tests completed.

Author Biographies

Xiangying Jiang is a lead learning scientist and works on learning assessment at Duolingo. She has a PhD in Applied Linguistics (Northern Arizona University, 2007) and was Associate Professor of TESOL at West Virginia University before joining Duolingo.

Lucy Portnoff is a data scientist at Duolingo. She graduated from the University of California, Berkeley with an undergraduate degree in Mathematics.

Luke Plonsky is Associate Professor of Applied Linguistics at Northern Arizona University. His work, focusing primarily on second-language acquisition and research methods, has appeared in over 80 articles, book chapters, and books. Luke is Associate Editor of Studies in Second Language Acquisition, Managing Editor of Foreign Language Annals, and Co-Director of the IRIS Database.

Bozena Pajak holds a Ph.D. in Linguistics (University of California, San Diego, 2012). Before joining Duolingo in 2015, she was a Research Associate and a Lecturer in Linguistics at Northwestern University. Her research focused primarily on the acquisition of additional languages in adulthood. She is currently the VP of Learning and Curriculum at Duolingo.
Acknowledgements

We would like to thank Cindy Blanco, Haoyu Chen, Erin Gustafson, Joseph Rollinson, and Lucy Skidmore for their help in collecting data and providing feedback on the report.

References


Appendix

The Background Survey

What language(s) was/were spoken in your home before you were 6 years old?

___________________

What other languages do you speak?

___________________

Why are you learning Spanish? (Check all that apply)

For travel
For school
For job-related purposes
For fun/leisure
For memory/brain acuteness
For social purposes
Other:

What other languages have you studied?

___________________

What is your highest level of education?

Some high school
High school
Associate’s degree
Bachelor’s degree
Master’s degree
Ph.D.
Trade school
Prefer not to answer

What is your age?

___________________

What gender do you identify as?

Male
Female
Other:
Prefer not to answer

Please specify your ethnicity.
Caucasian
African American
Latino or Hispanic
Asian
Other:
Prefer not to answer

How much Spanish do you think you knew before studying on Duolingo?

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nothing   Perfect

How much Spanish do you think you know now?

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nothing   Perfect

In which area or areas do you think Duolingo helped you the most? (Check all that apply)

- Vocabulary
- Grammar
- Pronunciation
- Listening
- Speaking
- Reading
- Writing

How much time (in hours) per week did you use Duolingo to learn Spanish?

___________________

In addition to the Duolingo Spanish lessons, what other Duolingo resources did you use to learn the language? (Check all that apply.)

- Duolingo Spanish Stories
- Duolingo Spanish Podcasts
- Duolingo Tips in Spanish
- Nothing else

What do you like about learning Spanish on Duolingo?

___________________

What do you want to see changed on Duolingo?

___________________

Did you have experience learning Spanish before using Duolingo?

- Yes
- No

—(If yes) How did you learn Spanish before using Duolingo? (check all that apply)
Being around Spanish speakers  
High school Spanish classes  
College Spanish classes  
Language apps  
Internet-based materials such as podcasts and YouTube  
Textbooks and other materials in print  
Other:

Did you take Spanish classes **during** the time you used Duolingo?  
Yes  
No

Did you use other programs or apps to learn Spanish **during** the time you used Duolingo?  
Yes  
No

To participate in the study, you will need to take online tests. Do you have access to a webcam-enabled, internet-connected computer?  
Yes  
No

To sign you up for the tests and get your official certificates, we’ll need your name. Please fill in your name below. Thank you!  
First name  
Last name
### B Appendix

**Background Information of the Participants in the Spanish and French Courses (including Unit 5 participant information from Jiang et al. (2021) for comparison)**

<table>
<thead>
<tr>
<th></th>
<th>Spanish</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unit 5 (n=135)</td>
<td>Unit 7 (n=144)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>46.89 (18.52)</td>
<td>47.92 (15.10)</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td><strong>Language before age 6</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English only</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>English and other–Early Bilingual¹</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Not English–Other²</td>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td><strong>Highest level of education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Trade school</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Associate degree</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>51</td>
<td>29</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>Doctoral degree</td>
<td>18</td>
<td>37</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Asian</td>
<td>18</td>
<td>27</td>
</tr>
<tr>
<td>Caucasian</td>
<td>105</td>
<td>107</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>76</td>
<td>62</td>
</tr>
<tr>
<td>Male</td>
<td>59</td>
<td>82</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Prefer not to answer</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1. The early bilinguals learning Spanish (n=27) spoke English and one of the following languages: Chinese (11), Hindi (3), Arabic (2), Hebrew (2), Italian (2), Russian (2), Bengali, Cantonese, Japanese, Korean, Norwegian, Polish, Punjabi, Swahili, Taiwanese, Tamil, Turkish, Urdu, Uzbek, Vietnamese, and Yiddish. The early bilinguals learning French (n=10) spoke English and one of the following languages: Spanish (4), Tagalog (2), Arabic, Japanese, Slovak, and Tamil.

2. Spanish learners who did not speak English before age 6 (n=56) spoke one of the following languages: Chinese (13), Russian (8), German (4), French (4), Hindi (3), Japanese (3), Dutch (3), Arabic (2), Tamil (2), Farsi, Finnish, Greek, Hebrew, Italian, Kannada, Malayalam, Romanian, Spanish, Turkish, Ukrainian, Urdu, Vietnamese, and Zulu. French learners who did not speak English before age 6 (n=30) spoke one of the following languages: Spanish (8), Chinese (6), Russian (3), Arabic, Bengali, Bulgarian, Cebuano, Filipino, Hebrew, Indonesian, Italian, Japanese, Kazakh, Korean, Punjabi/Hindi, and Russian.