Motivation and Oral Proficiency Outcomes in App-Based Foreign Language Learning

CALICO 2019

Dr. Shawn Loewen, Dr. Daniel R. Isbell & Zachary Sporn
Founded in 2007 in Berlin, Germany

• Premium, subscription-based language learning app and web platform
• Focus on real-life conversation skills and communicative competence: Every lesson aligned with CEFR Can Do statements for functional language ability
• Founded in 2007 in Berlin, Germany

+Babbel
14 learning languages (L2s) accessible via 8 display languages (L1s)

Localized, not translated:
Courses are designed, scripted and recorded by a team
over 150 teachers, linguists and translators

Every lesson features dialogues and audio recordings by
native speakers with a variety of accents
Babbel's Review Manager feature based on spaced repetition.
Vast library of "progressive" and modular content for each
language pair.

Choose your level.

En el hotel.

En el taxi!

El viento a Lima.

Download all 10 lessons.

Language Learning:
Upper Beginner/Pre-Intermediate level or
and the Jungle. The course is based on
also be within the locals, in the hotel, in the
tours and around Lima's nightlife. Tour
John Ecuador on his trip to Peru: making

Intermediate Courses

5 courses

Grammar practice in easy, understandable steps. Drills and exercises.

Beginner's Courses

6 courses

Reading and writing at a pre-intermediate level of learning.
These in-depth courses will practice intensive learning, speaking.
Here are some courses to refresh and improve your basic skills. With

CALICO 2019

6
Ser or Estar?

Family relations can be identified with "ser" or "estar."
Explicit vocabulary and grammar instruction
Babbel's ASR provides feedback on pronunciation and intonation.
1. What linguistic gains are made by English native speaker university students who exclusively use Babbel for autonomous study of Spanish as a second language (L2)?

Both pre- and post-tests consist of:
- an ACTFL Oral Proficiency Interview—Computer version speaking test
- a grammar test
- a vocabulary test (based on LexTale ESP)

2. Is there a relationship between how learners use Babbel, their self-reported motivation and gains, if any, on test scores?
Pretests (Jan. 15 - 30)

Babbel Study (Jan. – April)

Posttests (April 15 – 30)

- Babbel Introduction
- Quasi-individual
- Language lab
- Self-paced
- Weekly progress emails

- Babbel subscription
- $75
- Language lab
- In groups

Week 4 – Progress Survey

Week 8 – Progress Survey

CALICO 2019
Data regarding app participants’ app usage gathered by Babbel, but corresponding objective measures (i.e., pre- and post-test scores) were exclusive to MSU.

Included but not limited to:

- Names and total number of Spanish lessons completed
- Days of study
- Minutes of study
- Use of Review Manager feature
- App usage analytics provided by Babbel
Participants

• MSU undergraduate students
• Recruited in Jan. 2018 by Visiting TESOL and linguistics classes
• Posted Flyers
• Snowball recruitment
• Recruited in Jan. 2018 by MSU undergraduate students

Participants

Want to learn Spanish in just 15 minutes a day?
Participants

- Median age = 22
- Average of 2 previous Spanish classes
- No current Spanish students
- Previous Spanish study okay

BUT

- Previous Spanish study okay
- No current Spanish students
- Average of 2 previous Spanish classes
- Median age = 22

Participants
Participants

- 83 participants recruited and pretested
- 54 post-tested based on
- At least 3 hours of total study
- Weekly average of at least 20 minutes
- No gaps of 4 or more weeks
- Attrition rate of roughly 30%
- Attrition rate is far lower than that of up to 97% encountered in other studies involving mobile apps: Nielsen, 2011; Vesselinov & Greco, 2016; Despain, 2003; Van Deusen-Scholl, 2015; Lord, 2015; Van Deusen-Scholl, 2011; Vesselinov & Greco, 2016; Despain, 2003; Van Deusen-Scholl, 2015; Lord, 2015.
The ACTFL LTI OPIc® is a digital Oral Proficiency Interview (OPI), in which the test-taker answers prompts delivered via a computer, speech samples double-rated by trained assessors according to the ACTFL guidelines in order to assign a rating. Questions are delivered via a virtual avatar through a computer program, and at a time convenient to the candidate and proctor. Test-taker oral proficiency testing on a large scale provides valid and reliable oral proficiency scores and student preferences. Foreign Language Annals (Thompson et al., 2016) highly correlated with OPI ratings for Spanish.

ACTFL Guidelines present levels of proficiency as ranges and describe what an individual can do at each level. ...[The Guidelines] are an instrument for the evaluation of functional language ability.

ACTFL, 2014, p. 3

CALICO 2019
Why measure oral proficiency gains of a cohort of (mostly) novice level language learners?

Babbel focuses on oral communication skills: "Speak the language like you always wanted to."

BUT…

Doubt exists in CALL literature as to whether meaningful development of oral abilities is possible in the context of commercial online language learning platforms (e.g., Lord, 2015). Nonetheless, we believe it is essential for apps like Babbel to provide transparency on learners’ potential proficiency outcomes.


Novice

• Can communicate short messages on highly predictable, everyday topics that affect them directly.

- May be difficult to understand, even by sympathetic interlocutors.
- Use isolated words and phrases that have been encountered, memorized, and recalled.
- Can communicate short messages on highly predictable, everyday topics that affect them directly.

Novice
Intermediate

- Have the ability to create with language when talking about familiar topics related to their daily life.
- Can recombine learned material to express personal meaning.
- Produce sentence-level language, typically in present time.
Grammar Test

- 30 identification and error correction items
- One point for identification
- One point for correction
- 60 points total
- Reliability, \( \alpha = .93 \)

CALICO 2019
Spanish Grammar Test

Error Correction

Each sentence below contains one grammatical error. For each sentence, you will:

- Find and underline the error. Even if you do not know why it is wrong or how to fix it, you can still underline it.
- Write the corrected form on the line. Try your best, but if you have no idea how to correct the error, you may leave it blank.

Example:

0. Adriana está feliz.

__________________________

1. Son las tres con cuarto, ¡ya es tarde!

________________________________________

2. Lamento que María esté enfermo.

________________________________________

CALICO 2019
Spanish Grammar Test

Each sentence below contains one grammatical error. For each sentence, you will:

• Find and underline the error. Even if you do not know why it is wrong or how to fix it, you can still underline it. If you have no idea what the error is, you can skip the question.

• Write the corrected form on the line. Try your best, but if you have no idea how to correct the error, you may leave it blank.

Example:

0. Adriana está feliz.
   _____

1. Son las tres con cuarto, ¡ya es tarde!
   __________________________

2. Lamento que María esté enferma.
   __________________________
Vocabulary Test

• Based on LexTALE-Esp (Izura, Cuetos, & Brysbaert, 2014)

Vocabulary Recognition

• 60 points total
• 30 non-words
• 60 real words

Reliability, α = .87

CALICO 2019
Spanish Vocabulary Test

Word Recognition

In total, the test should only take a few minutes to complete. Please turn the page to start the test.

All you have to do is tick the box next to the words you know. If, for instance, in the example below you recognize “sí”, “sacapuntas”, “bien”, and “casa”, you indicate this as follows:

All you have to do is tick the box next to the words you know. If, for instance, in the example below you recognize “sí”, “sacapuntas”, “bien”, and “casa”, you indicate this as follows:

If you indicate you know a fake word, your score will be lowered. However, errors are penalized. So, if you indicate you know the word, you should be familiar with the word. Be careful!

For this test, you will see a list of 90 words that look “Spanish”. Only some of them are real.

Spanish Vocabulary Test
Motivation Questionnaire:

<table>
<thead>
<tr>
<th>Agree</th>
<th>Agree</th>
<th>Agree not</th>
<th>Disagree</th>
<th>Disagree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly</td>
<td>Agree</td>
<td>Agree</td>
<td>Disagree</td>
<td>Disagree</td>
<td>Disagree</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Strongly</td>
<td>Agree</td>
<td>Agree</td>
</tr>
</tbody>
</table>

4. I am interested in using Babbel (Circle one):

5. I am interested in learning Spanish (Circle one):

Assessed both Babbel and interest in Spanish...
Cumulative Study Hours

- Mean for all participants = 11.6 hours
- Maximum = 27.7 hours
- Median = 9.7 hours
- Minimum = 2.3 hours
<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int High</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Int Mid</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Int Low</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Novice High</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Novice Mid</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Novice Low</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

CALICO 2019
<table>
<thead>
<tr>
<th></th>
<th>Novice Low</th>
<th>Novice Mid</th>
<th>Novice High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int High</td>
<td>17</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Int Mid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int Low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Novice Low</th>
<th>Novice Mid</th>
<th>Novice High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int High</td>
<td>17</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Int Mid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int Low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CALICO 2019</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Pretest</td>
<td>Int High</td>
<td>Int Mid</td>
<td>Int Low</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Posttest</td>
<td>Int High</td>
<td>Int Mid</td>
<td>Int Low</td>
</tr>
</tbody>
</table>

**CALICO 2019**
<table>
<thead>
<tr>
<th>Pretest</th>
<th>Novice Low</th>
<th>Novice Mid</th>
<th>Novice High</th>
<th>Int Low</th>
<th>Int Mid</th>
<th>Int High</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>12</td>
<td>17</td>
</tr>
</tbody>
</table>

Posttest
<table>
<thead>
<tr>
<th>OPCI Change</th>
<th>Change</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>48</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>39</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>%</td>
<td>n</td>
<td>Change</td>
</tr>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Grammar</td>
<td>6.7</td>
<td>19.6</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>9.2</td>
<td>20.2</td>
</tr>
</tbody>
</table>

Average Grammar & Vocabulary Score Increase
Motivation: Spanish Interest
Change in OPlc Score according to hours studied
Change in Grammar Score according to hours studied
Changes in Vocabulary score according to hours studied

![Graph showing changes in Vocabulary score](image)
<table>
<thead>
<tr>
<th>Prior Classroom Learning Hours</th>
<th>Study Babbel Hours</th>
<th>Study Babbel Interest</th>
<th>Prior Babbel Spanish Study Interest</th>
<th>Babbel Vocab Change</th>
<th>Babbel Grammar Change</th>
<th>Babbel OPIC Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>1.16</td>
<td>0.30</td>
<td>-0.11</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.11</td>
</tr>
<tr>
<td>0.36</td>
<td>2.22</td>
<td>0.49</td>
<td>0.49</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>0.56</td>
<td>1.22</td>
<td>0.17</td>
<td>0.22</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.21</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.29</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.65</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
<td>0.54</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Correlations
Limited Mixed Effect Regression Results

- Time was a significant predictor for all three test scores.
- Interest in Spanish was a significant predictor only for OPiC scores.

Increases in OPiC scores over time were dependent on the amount of time a participant spent on Babbel and also on a participant's overall level of interest in learning Spanish.
Limitations

- Sample size
  - Relatively homogeneous population. University students, 75% of initial sample was female

- Intensity of study. Participants used the app for less than 1 hour per week, with considerable individual variation.

- Possible causes for quitting:
  - Frustration with challenge of Spanish learning
  - Frustration with app
  - Lack of time
  - Lack of time
  - Frustration with app
Main Findings

• Virtually all study participants made a measurable gain in their grammar and vocabulary knowledge and/or ability to communicate orally in Spanish.

• Babbel enables learners to transfer receptive, input-based learning and explicit grammar and vocabulary instruction to oral production at the Novice and Intermediate ACTFL levels.

• Babbel enables learners to transfer receptive, input-based learning and transfer receptive, input-based learning and explicit grammar and vocabulary instruction to oral production at the Novice and Intermediate ACTFL levels.

• Babbel enabled learners to transfer receptive, input-based learning and explicit grammar and vocabulary instruction to oral production at the Novice and Intermediate ACTFL levels.
Thanks for your attention.

Questions?
Zach Sporn
Senior Communications Manager, Babbel
zsporn@babbel.com
Change in OPIC Score according to hours studied

Table 6.

<table>
<thead>
<tr>
<th>Time × Spanish Interest</th>
<th>Time × Babble Study Hours</th>
<th>Classroom Experience</th>
<th>Time</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.03</td>
<td>1.17 (0.29)</td>
</tr>
<tr>
<td>0.19 (0.11)</td>
<td>0.03 (0.08)</td>
<td>0.31 (0.08)</td>
<td>0.10 (0.02)</td>
<td></td>
</tr>
<tr>
<td>0.42 (0.26)</td>
<td>0.31 (0.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Random Effects

<table>
<thead>
<tr>
<th>Participant (Intercept)</th>
<th>Participant (slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.23 (1.11)</td>
<td>0.11 (0.33)</td>
</tr>
</tbody>
</table>

Model R² = 29, R² conditional = 9.1, χ² conditional = 43.217, df = 4, p < .001.
Novice

- Can communicate short messages on highly predictable, everyday topics that affect them directly.
- Use isolated words and phrases that have been encountered, memorized, and recalled.
- May be difficult to understand, even by sympathetic interlocutors.
• Have the ability to create with language when talking about familiar topics related to their daily life.
• Can recombine learned material to express personal meaning.
• Produce sentence-level language, typically in present time.

Intermediate
<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPN SA</td>
<td>1.23</td>
<td>1.05</td>
<td>1</td>
</tr>
<tr>
<td>Sex</td>
<td>75% female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year of Birth</td>
<td>1997</td>
<td>5.44</td>
<td>1995.21</td>
</tr>
<tr>
<td>Classroom Experience (years in HS + courses in university)</td>
<td>1.80</td>
<td>1.88</td>
<td>1.80</td>
</tr>
<tr>
<td>Babbel Interest</td>
<td>5.23</td>
<td>0.77</td>
<td>5</td>
</tr>
<tr>
<td>SPN Interest</td>
<td>4.84</td>
<td>0.92</td>
<td>5</td>
</tr>
<tr>
<td>Vocab Score</td>
<td>10.87</td>
<td>9.43</td>
<td>9</td>
</tr>
<tr>
<td>Grammar Score</td>
<td>9.19</td>
<td>10.65</td>
<td>5</td>
</tr>
<tr>
<td>Minutes_Total</td>
<td>489.96</td>
<td>457.78</td>
<td>329</td>
</tr>
<tr>
<td>OPIc Rating</td>
<td>1.63</td>
<td>1.14</td>
<td>1 (NL)</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Descriptives – Whole Group (n = 83)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Did Not Quit (n = 54)</th>
<th>Quit (n = 29)</th>
<th>Mean or Total Difference (Quit – Did Not Quit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPN SA</td>
<td>1.44</td>
<td>0.83</td>
<td>0.61</td>
</tr>
<tr>
<td>Sex</td>
<td>68% Female</td>
<td>68% Female</td>
<td></td>
</tr>
<tr>
<td>Year of Birth</td>
<td>1994.44</td>
<td>1996.66</td>
<td>2.22</td>
</tr>
<tr>
<td>Babbel Interest</td>
<td>4.80</td>
<td>5.24</td>
<td>0.44</td>
</tr>
<tr>
<td>SPN Interest</td>
<td>0.96</td>
<td>0.75</td>
<td>0.21</td>
</tr>
<tr>
<td>Classroom Experience (years in HS +</td>
<td>2.08</td>
<td>1.99</td>
<td>0.09</td>
</tr>
<tr>
<td>OPIc Rating</td>
<td>1.33</td>
<td>1.54</td>
<td>0.22</td>
</tr>
<tr>
<td>Vocab Score</td>
<td>12.89</td>
<td>11.74</td>
<td>1.15</td>
</tr>
<tr>
<td>Grammar Score</td>
<td>11.70</td>
<td>11.74</td>
<td>0.00</td>
</tr>
<tr>
<td>Minutes Total</td>
<td>436.37</td>
<td>436.75</td>
<td></td>
</tr>
</tbody>
</table>

Group Comparisons
Initial Differences: Spanish Self-Assessment

- Difference = -0.61
- \( d = -0.61 \)
- \( p = 0.001 \)
Initial Differences:

- $d = 0.40$
- $p = -0.48$
- Difference = $-0.53$

Initial Differences: OPIC Rating
Initial Differences: Vocabulary Score

- Difference = -5.75
- $d = -0.63$
- $p = 0.007$

$\Delta = 0.007$

$3.75$
Initial Differences: Vocabulary Score

• Difference = -5.34
  • $d = -0.51$
  • $p = 0.029$

Grammar Score

M = 11.06
SD = 11.79

M = 5.72
SD = 7.05