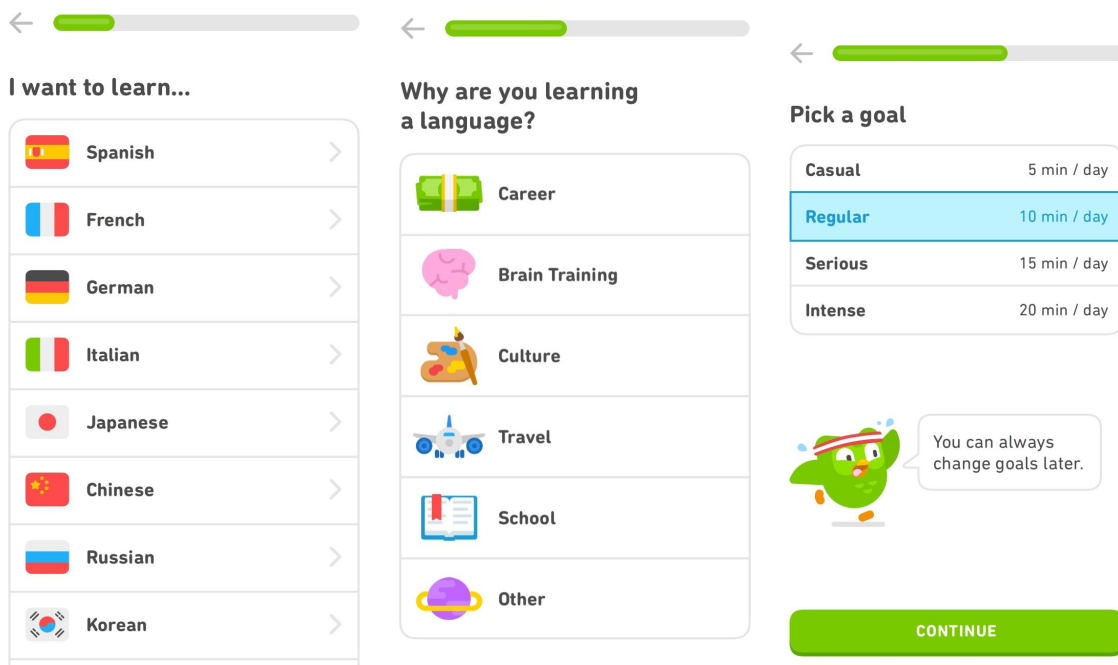


Duolingo UX Project

Sam Ayele

Duolingo is a popular mobile assisted language learning app that aims to teach over 300 million users a new language for free. It's scientific and gamified approach to learning entertains users while teaching them how to speak, write, and understand a foreign language (Teske 2017). Duolingo offers 22 languages, including fake ones like Klingon and High Valerian.



Duolingo got its start in 2009 when professor Luis von Ahn from Carnegie Mellon and his graduate student Severin Hacker wanted to create a program that would educate people for free. Dr. Lusi von Ahn grew up in Guatemala where it was very expensive for people to learn English. He had a vision to bring language to people around the world at no cost. Since then, Duolingo has kept that promise and plans their business model around periodic advertisements that do not take away from the learning.

I have been learning Dutch on Duolingo for about three months. I am interested in the long-term effects of Duolingo and if users believe it has helped them speak a different language.

Users tend to be people who have an intrinsic motivation to learn a new language. They may be travelling to a different country, want to practice outside of a classroom, or are interested in a specific culture.

I love that Duolingo feels more like a game and less like a chore. It gives good feedback and rewards for completing a lesson everyday. Duolingo has daily streaks which encourage users to continue using the app (Weathers & Poehlman 2017). I also enjoy the green owl mascot, Duo. However, after a while, the lessons start to get tedious. Duolingo is also notorious with their aggressive push notifications that have annoyed some users. I think this system is excellent for evaluating language acquisition and decision making. I think Duolingo has the potential to change the world if they have a great user interface and enough support to keep users learning.

Duolingo approaches language learning through a variety of tasks. These include listening and speaking tasks, pair matching, and sentence translations. They also use A/B testing on different users to refine the way they structure a course. Based on their findings, they can track specific subjects learners have difficulty with as well as interfaces that make the app more desirable. In my expert interviews, I saw three, including my own, Duolingo apps that have different layouts and features (see Appendix A).

Experienced Users:

The first person I interviewed was a 32 year old cognitive science major. User A downloaded Duolingo when it first came out in 2012. Although User A has changed his phone a couple of times since then, he has always used the Android app. User A studied Spanish in high school and downloaded Duolingo to help his fluency. Although User A has had the app for 8 years, he's had a hard time committing to it. The last time he used the app was a week before the interview and his longest streak was probably seven days. User A had about three lessons

completed in Spanish. He has also attempted to learn other languages like French, German, and Japanese but did not make any headway with these either. User A feels like he does not have a good reason to learn a language. He feels embarrassed when he uses the app in a public setting because he fears other people might hear his listening and speaking tasks. User A only uses Duolingo in a bored emotional environment. He claims “the gamified approach makes it harder for me to take it more seriously”.

User A’s approach to Duolingo is to “binge, stop, and start again”. He does not use it consistently but when he does open the app, he tends to open it again for the next couple of days. But, once he stops using the app for a day, it is harder to go back to it. He is convinced that the app is not actually working to get him at a conversational level. He claims that the feedback is too lenient and often gives him credit for completing a task he did not think he deserved. He doesn’t think that the app actually trains the brain to speak a new language but to be proficient at a task.

When I asked User A to open up Duolingo, the first thing that popped up was an ad for Duolingo plus, User A was visibly annoyed. When asked about what he thinks of the advertisements on Duolingo, he said they are annoying but understandable to keeping Duolingo free. He did, however, complain about the aggressive daily push notifications. After dismissing the ad, I noticed that User A had only attempted three lessons in Spanish. User A is a perfectionist and wanted to complete an entire lesson before moving on to the next. He complained that the basic lessons were very repetitive but still offered higher levels of completion.

One difference between my interface and User A’s is that he does not have gems as a Duolingo currency for completing a lesson. The android app includes linglots which are less

frequently earned and make the android app less of a game than the IOS version. User A has never bought anything in the store app with these linglots because he “thinks its dumb”.

The next person I interviewed was a 22 year-old literature major who likes reading, writing, and doing puzzles. User B has used the app since junior high school in 2014. They had studied German in high school and college. They wanted to use Duolingo to practice and refine their skills. User B had a 1,000 day streak, meaning they’ve been used the app consistently for more than three years. They use the app daily for about 20 minutes when their push notifications goes off at the end of the day. Unlike User A, User B does not mind using the app in public settings like the dining hall. Although, they do use headphones when completing a listening task.

User B has completed every lesson in German as well as attempted Spanish, Czech, Russian, Klingon, and Valerian. Though, they tend to focus on one language at a time. User B practices Duolingo on an iPhone. Like User A, their interface also uses linglots instead of gems. User B has 5871 linglots saved up. They said “I’ve bought pretty much everything in the store already so they’re pretty much meaningless now”. User B often buys streak freezes in the store so they can continue their streak if they accidentally miss a day.

Both User A and User B commented on the difference between learning on Duolingo vs in a class. User A said that a teacher can provide better feedback and conversational topics than a computer generated sentence. User B agrees saying “a real class makes conversations easier while Duolingo is better for just practice”. Both users told me they have never used the sentences they learn in the app outside of Duolingo.

Based on these two interviews, I was able to extract three tasks to evaluate. First, completing a daily streak. When users first launch the app, they select a language as well as a daily goal to complete. These goals are 5, 10, 15, or 20 minutes per day depending on the level

the user wants to practice at. This is an essential task for Duolingo. After they complete their daily goal, I want users to find the practice mode to brush up on the skills they learned. Finally, I am going to test users on their ability to find the store and buy an outfit for Duo. Since Duolingo uses A/B testing, I am going to use my personal phone for each of the trials so the users have a consistent interface.

Novice Users:

I recruited 5 new Duolingo users for this part of the study. User C is a 21 year old cognitive science major who speaks English and Spanish fluently. User C has wanted to learn French for a long time but never had the time to get into it. He decided to pick the French module for the task evaluations. User C has downloaded Duolingo just to see what it was like but never really got into it.

User D is a 21 year old cognitive science and psychology major. User D has taken four years of Spanish in high school but doesn't consider herself fluent in it. She also understands Tamil because of her Southeast Asian background but does not speak it very well. She has heard of Duolingo but never downloaded or used any other mobile assisted language learning apps. This user picked Japanese for evaluation.

User E is a 29 year old psychology and cognitive science major. User E speaks Hebrew and English fluently, with Hebrew being his native language. User E took Arabic in high school and wants to become more fluent in it. User E has never heard of Duolingo nor used any platforms similar to it.

User F is a 20 year old Astrophysics major who is only fluent in English. She comes from a Vietnamese background and has been going to Vietnamese school since she was in the third grade. She has been unsuccessful in obtaining fluency because she does not have a strong desire to learn a new language. She also took Mandarin for four years in high school. Although

she's never used Duolingo, User F has used Rosetta Stone for Vietnamese on and off for about 6 years. User F wanted to challenge herself with a language she is not familiar in so she chose Greek for the evaluation.

User G is a 24 year old psychology major. She is a native Spanish speaker who is also fluent in English. User G studied French in high school for two years but doesn't remember any of it. She has a desire to learn French and Italian for future travels but is not committed. She chose to evaluate the French module at the casual level of 5 minutes per day.

I interviewed each user using a semi-structured script (see Appendix G). I asked for permission to video record our interview as well as screen record my iPhone. I used my personal device each time and logged out every time a new user was interviewed. I walked each user through the onboarding procedure, asking them to pick a language and a daily goal. From then on, I asked them to complete their daily goal, navigate to practice mode, and buy an outfit from the store. After they completed these tasks, I asked them what they found easy, difficult, confusing, or surprising. To conclude, I asked participants to pick six adjectives from the Microsoft desirability toolkit that best described their experience (see Appendix B and F).

Results:

On average, novice users made ($M= 2.6$) errors on task 1 (see Appendix E). This included selecting a language, choosing a daily goal, and completing the daily goal. On task 2, users made an average of ($M= 16$) errors before locating practice mode and completing a practice session. Finally, users made an average of ($M=2$) errors on task 3, finding the store's app and buying something. Overall, users made ($M= 6.87$) errors across all three tasks. The average time to complete all three tasks was ($M= 15:32$ minutes).

None of the users made errors in picking a language and a daily goal. When picking a daily goal to complete, Users C, D, and F noted that the difference between regular and serious

(10 or 15 minutes a day) does not seem that significant to them. User C claimed “I don’t really see the difference between 5 minutes a day. I don’t know why one is just regular and the other is serious? It seems like the same amount of time everyday”.

Although each user had the same equipment, I received varying feedback based on the language they were practicing. Out of all the languages I saw, the French module seemed to be the most developed and user friendly. Users C and G liked the visual representations of cues in the beginner lessons (see Appendix C). French Dutch, Spanish, and English were the only languages that made use of these visual cues.

User D, E, and F chose Japanese, Arabic, and Greek, languages that do not use English letters. This created different difficulties that Spanish or French learners did not experience. First, users complained it was difficult to follow the instructions. Duolingo prompted them to “translate this sentence” but it would be ambiguous whether they were supposed to translate into English words or the character sound in English characters (see Appendix D). While completing their lessons, Users D and F noted that they were just guessing what Duolingo wanted them to do. They were presented characters from the languages with no meaning provided. Although it was a beginner lesson, they wished they were in the loop about what they were actually practicing.

Another problem that came up with character languages was an insufficient keyboard. When the users were asked to translate an English sentence into their respective language, an English keyboard would pop up. Unless the users had previously installed an international keyboard on their device, these translations were impossible. User E and User F were unable to complete their daily goal because of this problem.

With a few exceptions, most users did not have problems with recognition based translation tasks. Even users who selected a character language improved with use. The most

difficult task users experienced were speaking tasks. They usually needed two or more attempts to pronounce the words correctly. Like User A, User C was not confident about the system feedback. He felt that he did not pronounce the words correctly, therefore did not deserve the credit the system gave to him. This makes Duolingo less credible as a language learning system. In the pairing exercise, User D noted that the fonts were too small to differentiate. She often touched a word she did not intend, making avoidable mistakes in the exercise.

Task 2 was the most difficult out of the three. I chose this specific task because it took me three weeks to find the practice mode feature on my phone. My interface includes five “hearts” or chances to learn a new lesson. If more than five mistakes were made, users can go to practice mode to gain another heart and keep learning. Duolingo uses this feature to discourage binging behaviors like User A. However, when prompted to go to practice mode to gain another heart, users were very confused about where to locate this. They scrolled through the entire app, some even managing to click on the correct icon and exiting, before finally finding the feature. User E had the biggest problem with this task, making 26 errors before finally locating practice mode. However, when asked about any difficulties in the post-interview, User E did not mention this trouble. In fact, he said navigating the app was one of the easiest tasks. User G, however, begs to differ. She said navigating around the app was one of the more difficult parts of Duolingo.

The last task was the easiest of all. Users completed task 3 the fastest and made the least mistakes. This is because of several reasons. First, the store has great visibility. The chest icon makes it easy to locate in menu, it is always visible on the home screen. Second, when completing task 2, users already located the store when scrolling around. Although some users needed a few seconds to relocate the chest, most automatically went to it after my instruction. All of the users were pleasantly surprised with the things they could purchase for Duo. User D

exclaimed “Are these outfits for the owl? How cute! I would literally spend all my money on these”. User G said she was surprised about the fact that they even had outfits available for purchase. She said “It’s just not something I would expect out of a language app. It’s kinda random”.

Overall, new users had a great time with Duolingo. Despite the navigation and keyboard problems, half the users thought the app was easy to use. A third of the users found it motivating, convenient, and at times confusing. Most participants appreciated the great visual and audio feedback. However, Duolingo needs to make practice mode more salient and provide better support for character languages like Greek and Japanese. Users really resonated with the greater mission of providing language education for free. User E wished his grandparents had this technology in their day to help them read and write in Arabic. Four out of five users explicitly mentioned they would download the app on their phone and continue to use it.

Summary:

Duolingo is a gamified language learning system that depends on continuous daily use. It relies on a conversing interface to teach users how to speak a different language. Duolingo uses touch interaction for translating sentences and speech interaction for listening and speaking new phrases. Although it is marketed to new language learners, students like User B use it to practice languages they learn in a classroom.

Duolingo follows design heuristics pretty well. Its minimal green and white aesthetic design makes Duolingo stand out from its competitors. Its cleverly designed owl gives the company a mascot users can remember. Duolingo also has great visibility of system status. The bar at the top of each lesson lets users know how far they are progressing into the lesson (see Appendix D). Duolingo uses a good mix of recognition and recall in their lessons. They ask users to translate sentences using either a word bank or a keyboard. The word bank promotes

recognition while the free typing needs the use of recall. This production of speech helps improve overall comprehension of the language (Hoopman & Macdonald 2018).

What Duolingo does really well, however, is mitigating frustrations online second language learners usually feel (Sun 2013). Duolingo is a low risk app. Since it is free, it is easily accessible to all kinds of learners. When users make mistakes, Duo the owl encourages users that making mistakes is alright. The first time new users made a mistake, Duolingo explained that it was no big deal and encouraged users to keep going. Mistakes can be frustrating and disencouraging. Duolingo has found a way to make mistakes into a learning opportunity without losing their users to feelings of disappointment.

Improvements:

Although Duolingo does a lot of things well, there is much room for improvement. User A mentioned that he could not get past the basic levels of Spanish even though he had Duolingo for almost 8 years. He is a perfectionist and wants to complete all the level of lesson 1 before moving on. Perfectionism often hinders our ability to learn a new language (Ghorbandordinejad 2015). Duolingo can work around users like this by shortening the length of the basic lessons. Instead of encouraging users to learn the same thing over and over again, nudge them to move on and relearn the basics in practice mode. They should also make practice mode more visible in the A/B condition that involves hearts. Users may not be aware that this feature exists and might be discouraged from returning if they lose too many hearts.

One of the biggest problems mentioned by User B, User D, and User F is the lack of an international keyboard for languages that do not use English characters. Most users don't want to go through the trouble of downloading an international keyboard just to learn on Duolingo. Duolingo should develop their own keyboard that pops up when users are in a character language module. This would encourage more users to try a language they normally would not.

Finally, Duolingo needs to be more flexible for experienced users. Although new users can utilize gems or linglots by buying new items from the store, experienced users like User B cannot do anything with their saved up currency. By providing more outfits or features to buy, experienced users can have another incentive to keep coming back to the app.

Since language is an essential part of being human, Duolingo has the potential to change the world. By providing a way to let people learn without economics barriers, Duolingo is fostering connections between people that would not have been possible before.

References

Ghorbandordinejad, F. (2015). Examining the Relationship between Students' Levels of Perfectionism and their Achievements in English Learning. *Research in English Language Pedagogy (RELP)*, 36–45. Retrieved from <http://search.proquest.com/docview/2301548971/>

Hopman, E., & Macdonald, M. (2018). Production Practice During Language Learning Improves Comprehension. *Psychological Science*, 29(6), 961–971.
<https://doi.org/10.1177/0956797618754486>

How Duolingo Built a \$700 Million Company Without Charging Users. (2019, February 24). Retrieved from
<https://producthabits.com/duolingo-built-700-million-company-without-charging-users/>.

Susan Y. H. Sun (2014) Learner perspectives on fully online language learning, *Distance Education*, 35:1, 18-42, DOI: 10.1080/01587919.2014.891428

Teske, K. (2017). Duolingo. *CALICO Journal*, 34(3), 393–401. <https://doi.org/10.1558/cj.32509>

Weathers, D., & Poehlman, T. (2017). 19-C: Keeping the Streak Alive! *Advances in Consumer Research*, 45. Retrieved from <http://search.proquest.com/docview/2126571106/>

Appendix A

UX Expert Interviews

- 1) Introduction
 - a) Greet and thank
- 2) Set expectations
 - a) I'm doing a user experience project for a class called Human Factors. This is going to be a 45 min interview about the learning service Duolingo. We're going to talk a little bit about how you use Duolingo
 - b) You are free to stop this interview at any point if you feel uncomfortable but I would appreciate all the information you can give me
- 3) Any questions?
- 4) Permission to record the conversation
- 5) Start interview
 - a) How old are you?
 - b) What is your major?
 - c) What are some of your hobbies?
 - d) When did you start using Duolingo?
 - e) Why did you want to start using Duolingo?
 - f) Do you usually study on the website or the app?
 - g) How often do you use Duolingo?
 - h) When was the last time you used Duolingo? Is that how you typically use it?
 - i) If you stopped using it, why?
 - j) Did you learn a language in high school? If you did is that the language you are learning now?
 - k) What made you choose this language?
 - l) Can you show me a bit how you use Duolingo?
 - m) What do you think of Duo the owl?
 - n) What is your longest streak?
 - o) If you broke your streak, how did that make you feel?
 - p) Did you ever get back on the streak?
 - q) What goes through your head when you run out of hearts?
 - r) Do you use the discussion section?
 - s) How many gems do you currently have?
 - t) Do you tend to save or store the gems?
 - u) Would you consider getting the paid version of Duo?
 - v) Do you have push notifications on Duolingo?
 - w) When was the last time you spent some gems on something?
 - x) Do you practice this language outside of this app?
- 6) Any other thoughts you'd like to share? Questions? Goodbye

Appendix B

The complete set of 118 product-reaction cards				
Accessible	Creative	Fast	Meaningful	Slow
Advanced	Customizable	Flexible	Motivating	Sophisticated
Annoying	Cutting edge	Fragile	Not secure	Stable
Appealing	Dated	Fresh	Not valuable	Sterile
Approachable	Desirable	Friendly	Novel	Stimulating
Attractive	Difficult	Frustrating	Old	Straightforward
Boring	Disconnected	Fun	Optimistic	Stressful
Businesslike	Disruptive	Gets in the way	Ordinary	Time-consuming
Busy	Distracting	Hard to use	Organized	Time saving
Calm	Dull	Helpful	Overbearing	Too technical
Clean	Easy to use	High quality	Overwhelming	Trustworthy
Clear	Effective	Impersonal	Patronizing	Unapproachable
Collaborative	Efficient	Impressive	Personal	Unattractive
Comfortable	Effortless	Incomprehensible	Poor quality	Uncontrollable
Compatible	Empowering	Inconsistent	Powerful	Unconventional
Compelling	Energetic	Ineffective	Predictable	Understandable
Complex	Engaging	Innovative	Professional	Undesirable
Comprehensive	Entertaining	Inspiring	Relevant	Unpredictable
Confident	Enthusiastic	Integrated	Reliable	Unrefined
Confusing	Essential	Intimidating	Responsive	Usable
Connected	Exceptional	Intuitive	Rigid	Useful
Consistent	Exciting	Inviting	Satisfying	Valuable
Controllable	Expected	Irrelevant	Secure	
Convenient	Familiar	Low maintenance	Simplistic	



Which of these is "the boy"?



le chat



l'homme



le croissant



le garçon

CHECK

Appendix D



Translate this sentence



VL
.....

o

omicron

n

alpha

nee

CHECK

Appendix E: Quantitative Results

Participant	Daily Goal	Task 1 Error	Task 1 Time	Task 2 Error	Task 2 Time	Task 3 Error	Task 3 Time	Total Time
User C	15 min a day	3	11:27	11	8:13	0	:45	19:48
User D	15 min a day	0	13:42	7	5:09	0	:36	19:27
User E	10 min a day	5	4:43	26	2:29	6	:32	10:32
User F	10 min a day	4	9:56	20	6:27	0	1:03	16:13
User G	5 min a day	1	6:07	16	3:07	4	:29	9:43

Appendix F: Microsoft Desirability Results

User C	Satisfying	Unpredictable	Overwhelming	Fun	Entertaining	Appealing
User D	Friendly	Motivating	Engaging	Clear	Organized	Easy to use
User E	Straight forward	Motivating	Customizable	Useful	Innovative	Easy to Use
User F	Confusing	Simplistic	Appealing	Clean	Convenient	Time saving
User G	Confusing	Creative	Effective	Fast	Convenient	Easy to Use

Appendix G

Novice User Interview

1. Introduction

- a. Hi, my name is Sam. The goal of this study is to identify any design issues with Duolingo and to help me understand where this tool could be improved. This study should last about 45 to 60 minutes. No breaks are scheduled, but we can stop at anytime if you need one. You are also free to completely stop this interview at any point with no negative repercussions. I will be giving you 3 different tasks with Duolingo to perform on this test device]. I want to emphasize that we are not testing you personally or your own ability to complete these tasks. The goal of this study is to see where this technology could be improved, so please be honest and don't worry about performing everything correctly. Also, I would like your permission to record this session for reference later when writing my report. Do I have your permission to record?

2. Pre-Interview

- a. I'm just going to ask you some background questions before we begin
- b. What is your name?
- c. How old are you?
- d. What is your major?
- e. What are some of your hobbies?
- f. How many languages do you speak? How fluently?
- g. Have you ever wanted to learn another language?
- h. Have you taken a language class?
- i. Have you ever used something like Duolingo before?

3. So now we're going to begin. I'm going to give you three tasks to complete and I'm going to ask you to say aloud your process in completing this task. Please be vocal about anything that crosses your mind during this process, any data is great data.

- a. First, please select a language you want to learn
- b. Now, complete a lesson
- c. (So you've just run out of hearts)(Say you want to practice the things you've just learned) Please go to the practice mode
- d. Finally, you've collected all these (gems)(linglots), please buy an outfit for Duo in the store

4. Thank you so much. We are almost done. I just need you answer some questions and fill out this questionnaire

- a. Of all the things you tried today, what was easy for you? Why?
- b. Was there anything that was especially difficult? Why?
- c. Was there anything confusing or surprising? Why?
- d. Anything else you'd like to share?
- e. Okay, after you fill out this questionnaire, you are done with the interview