

Project Proposal by Nick Jenkins njenkins30

This paper forms a proposal for an educational technology development project. The project will produce a language learning tool for beginner Japanese students. The tool is intended to be a supplement to Duolingo – a popular language learning app. In this paper we will discuss the motivations and context for wanting to develop a tool in this space. We will discuss at a high level the description of the tool to be created, before getting into a detailed schedule of tasks, timings and reasons for task inclusion. Lastly, we will discuss the major project milestones.

Motivations and context

The motivations to want to create a language learning tool are born of both personal desire and to contribute to an area which has not realised it's full potential. As the tool is intended to be a supplement to Duolingo, we will discuss this tool in depth, discussing its strengths and weaknesses. As per previous assignments, we will revisit some of the educational technologies which Duolingo uses and does not use. We will highlight the importance of these technologies to good learning outcomes. We will refer to some educational technology research that shows that there are missed opportunities in the Duolingo platform, which gives us an opportunity to contribute through the development project. Lastly, we will discuss how this separate tool would retain many of the vital strengths of Duolingo.

As discussed, there is a personal desire to create a language learning tool. I have previously used the Duolingo system for learning Portuguese. My initial interest in Duolingo was to learn Japanese, however, it took over five years since I joined the app to be released. The timing of the Japanese language release was fortuitous, as I have recently moved to Japan to live for the next six months. I am eager to learn Japanese, however the course for Japanese in Duolingo is comparatively lacking. It contains less spoken and comprehension exercises and less content overall. This problem provides an intrinsic motivation to want to contribute to the area. Further to this, I am interested in taking the opportunity to learn and deploy web technologies to produce an application.

By way of general overview - Duolingo is an app for language learners which uses many educational technologies. Some of which include gamification, spaced repetition learning, and social aspects. The gamification seeks to turn the coursework into 'level's, and exposure to a language to 'xp' (experience). A spaced repetition model ensures that students repeat areas which are less well understood over time intervals¹. For the social aspect, students can add friends, join clubs, even participate in classrooms. This allows students to track their progress as well as utilise competition to motivate learning.

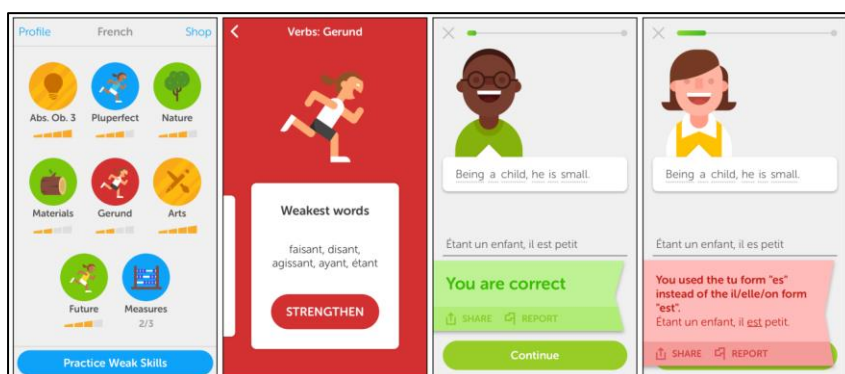


Figure 1 from Duolingo whitepaper. Left to right Skill Tree, Skill screen detail, Correct response, Incorrect response

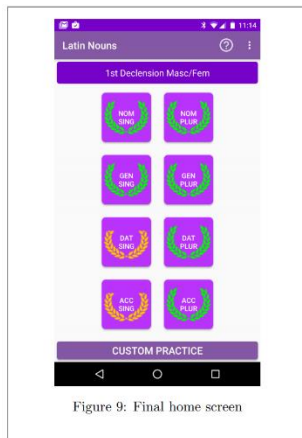


Figure 9: Final home screen

Other students in previous semesters have also studied Duolingo and created similar application extensions. For example, Eric Ianni's Latin Noun learning tool pictured to the left. Taking elements from popular apps including Duolingo and Rosetta Stone, this tool assists Latin learners in memorising noun conjugations using spaced repetition principles. The research into spaced repetition and its effects on recall are well explored in this paper; Ianni's app utilises the phenomenon to increase the effectiveness of memory encoding.

The tool being proposed would also include a SRM. Duolingo developers have published papers regarding their SPMⁱⁱ, and I plan on trying to implement the Leitner System (1972) referenced in their paper.

The origins of Duolingo begin to paint a picture of its weaknesses, and show us an opportunity to build a supplementary tool. Duolingo was mainly founded to help with crowdsourced translation, meaning written and reading comprehension is a primary goal of the appⁱⁱⁱ. Some courses such as Spanish and French also contain a spoken language portion, be it asking students to repeat back phrases, or to speak a phrase in a high-quality voice talent audio file. Many new courses however do not have this functionality. Despite being some of the most popular and keenly anticipated languages^{iv}, Japanese and Korean have a poor amount of verbal content. Another issue is that because of the translation focus, conversational speaking is less of a focus. This means content is geared towards learning language syntax and vocabulary, rather than person to person communication.

Further research into pedagogical theory indicated several key areas required for language learning. While Duolingo handles most of them, there are several that it doesn't, such as;

- Learning how to naturally delay a conversation (Fillers such as Ummms and ahhs)^v
- Asking someone to repeat themselves^{vi}
- Using simplified language and combining language (synonyms) to work around not knowing vocabulary^{vii}
- Conversational settings *such as how to make small talk*^{viii}

In addition to these, it would also help students to prepare for handling money, and learn how to interrogate a speaker for a new phrase, word or explanation. The aim of a new tool would be to directly teach some or all of these areas. The amount and scope of content would be determined following the implementation of the base technology platform.

There are many good parts to the Duolingo platform which would be retained in a new tool. Firstly, it would help to make it obvious that it is intended to be a supplement to Duolingo. It would mean that students felt familiar with the mechanics and could confidently interact with a new tool. Some of the areas that Duolingo does well are the breadth of subjects available, the social aspects, the spaced repetition algorithm, the gamification, the staged progress, and the positive reinforcement.

There are a few technologies I would like to include in a platform to help achieve an improved learning experience;

1. Web hosted with database
2. Use of microphone and speaker
3. Speech to text technologies
4. Text translation technologies

5. Text similarity engines
6. Chatbot technology

These technologies would be used in several ways. For instance, 1-2 can help a learner to listen to Japanese, record their own attempt, and play back their attempt, and repeat until satisfied. Adding in the 3-4th element would allow the computer to attempt to grade/judge a person's pronunciation; exercises where a learner repeats a phrase until it is correctly recorded and translated (or near enough). Adding the 5th and 6th elements allows us to move beyond rote exercises and swap in similar words, or new questions. My background is in text analysis, so I plan to use existing tools to produce and curate a list of similar words in English, along with some basic conversational threads. Adding to the above, we also mentioned the SPM and modularisation, to help learners tackle areas they are prepared to learn.

Example of webpage, game elements and aims. Some thought has been given to branding, colours and font usage.

Otsukare

good work!

Otsukare (good work) is a language learning game designed to help you become more confident speaking in Japanese.

You play the role of Kimio, a hard working character that must repay his families debts! Join Kimio as you search for work, complete tasks and reunite with family members.

This app is designed to be a complement to the popular language learning app Duolingo. We encourage you to practice there and fight for freedom here! This app is designed to be spoken/comprehension based and put you in real world situations helping you to understand:

- Meeting people
- Times and dates
- Directions and navigation
- Paying for items

Our hope is that you become more comfortable speaking Japanese, and help Kimio rescue his family!

Run errands

Save your family

Learn Japanese

Description of the development track tools

As discussed above, the intention is to create a supplementary tool to Duolingo. It would be a web based application where the primary function is to enhance and add to the Japanese Duolingo course experience.

The website would include microphone and audio playback functionality, with quizzes and artificial interactions (chatbot).

The web application, learning algorithms and text analytics elements would be written in python and flask. The front end would include JavaScript elements to capture and present user input. The project will make use of Google services (translation, transcription, text-to-speech) for input processing. The web site would initially be hosted locally, but would be moved to a remote service

such as PythonAnywhere. It would utilise a postgres database for a personalised user based experience, and for storing resources.

The completion of this project relies on some personal leaps in learning about specific technologies. There has been some initial prototyping of transcription, audio recording and playback, and a basic website. The tougher elements would be getting a first ever remote hosted server, with a user login style database, and having good project/script structure to enable both analytical processing alongside web hosting. As a fall back, the project would be hosted locally, perhaps without individual user login ability.

Tasks and Schedules

The following schedule indicates the proposed tasks to take place in certain weeks. Care has been taken to do tasks in an order which best benefits the project, and to spread the workload over the weeks (particularly the 9h vs 7.5 weeks). Gray lines indicate Milestone delivery expected at the end of those weeks. Following the schedule is a list of tasks with additional detail.

Week starting	Hours	Week	Task 1	Task 2	Task 3	Task 4
9/10/2017	7.5	1	Scope technology	Setup database	Basic website (local)	
16/10/2017	9	2	Create Content	Audio recording	Transcribing	
23/10/2017	7.5	3	Create Content	Translation	Text to speech	
30/10/2017	7.5	4	Create Content	Deploy website (web)	Design game elements	Text analysis
6/11/2017	9	5	Create Content	Deploy website (web)	Learning algorithm	
13/11/2017	7.5	6	Create Content	Implement game elements		
20/11/2017	7.5	7	Creating exercises	Chatbot / Scripts		
27/11/2017	9	8	Creating exercises	Chatbot / Scripts	Testing and delivery	

Week 1:

Scope Technology (1.5)

Confirm details of specific technologies to use, such as google service credentials, postgres logins. An early decision means I can get on with creating project

Setup database (3)

Create postgres tables, users, basic logins, data entry method. This will enable multiple users to interact with site including admin, data entry and students.

Basic website (local) (3)

Design the basic site layout for prototyping, styles. Get set up with basic local hosting for developing technology prototypes.

Week 2:

Create Content (12, over 5 weeks)

Input and transcribe Japanese data for page, estimate adding 20 words per hour. To create the basic content for the site, will aim to include around 250 of the most common and useful Japanese words

Audio recording (2)

Create ability to record audio in the browser. This step will be used to collect student input, for use in listening and comparing own voice with an example, and for transcribing.

Transcribing (3)

Transcribe recorded audio into plain text. Will help to display audio as text, so students get a sense of their pronunciation / getting audio transcribing service to understand what is being said

Week 3:

Translation (2)

Translate text from Japanese to English and vice versa. Allows text analysis and judgements of correctness to be made in English.

Text to speech (2)

Create text to speech within browser. Allows content to be provided as audio, to improve student comprehension.

Week 4:

Design game elements (2)

Brainstorm basic game elements to focus the technologies towards a user experience. Designing the game elements shortly after the base technology build, and before creating exercises will help focus the tasks. The game mechanics should be natural, aligned with the material, and engage the student.

Text analysis (2)

Perform text analysis of English/Japanese words to determine interchangeability. A key component of the project will be the ability to swap in similar words and terms into a script/conversation (That is a chair / This is a door), text analysis will support this.

Deploy website (web) (5, over two weeks)

Deploy website onto a hosted platform. This will enable end users to access remotely

Week 5:

Learning algorithm (3)

Implement a spaced repetition algorithm with possible 'RL teacher' modifications. This will ensure the material taught to students is interactive and driven by student's learning progress.

Week 6:

Implement game elements (4)

Implement the specific game elements which were designed previously. As mentioned, game elements will help motivate the students.

Week 7:

Creating exercises (8, over two weeks)

Design and create exercises to be attempted by students. This allows students to piece together the words into more coherent sentences.

Chatbot / Scripts (8, over two weeks)

Design scripts to engage students in conversational learning. This provides students context to learn new material.

Week 8:

Testing and delivery (4)

Ensure product is working and bug free. Ensure a good user experience while on the site.

Intermediate Milestones

To expose classmates to the project and give them the information necessary to give feedback, we propose two intermediate milestones;

1. **Video presentation of the locally hosted prototype.** This would include demonstrations of the transcribing, text-to-speech, audio recording and playback. It would showcase some initial content being created, as well as introduce a basic quiz.
2. **Functional web prototype.** This would include the ability to create a user login, and begin using the tool to learn basic words, with the previously demonstrated speech capabilities. It would also include the spaced repetition algorithm, and modularisation of the content, so users could start using the tool to learn basic Japanese

The step between the functional web prototype and full site would largely include factors of gamification, increased content, and chatbot/script functionality.

i Burr Settles, Brendan Meeder (2016). Journal - <https://s3.amazonaws.com/duolingo-papers/publications/settles.acl16.pdf>

ii As per i.

iii Deb Smit (2014). Article - <https://www.fastcompany.com/3038459/duingos-new-mission-watching-you-learn>

iv Lauren Orsini (2017). Article - <https://www.forbes.com/sites/laurenorsini/2017/05/18/japanese-finally-comes-to-duolingo/#7286f54854cf>

v Dr Stephen Juan (2006). Blog - https://www.theregister.co.uk/2006/05/06/the_odd_body_language_fillers/

vi Abepiusc (2013). Blog - <http://busyteacher.org/14911-how-to-teach-students-ask-for-clarification.html>

vii Stuart Webb (2007). Journal - <http://nflrc.hawaii.edu/rfl/October2007/webb/webb.pdf>

viii Japanese N1-N5 grading (NA). Article - <http://www.jlpt.jp/e/about/levelsummary.html>