© 2015 Thomas Reese

FLIPPING THE LANGUAGE LEARNING CURVE
WORD BY WORD

BY

THOMAS REESE

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Adviser:

Associate Professor Brian Bailey

ABSTRACT

Changing habits to incorporate dedicated learning time into daily life is difficult. This thesis details passive learning through a system that takes advantage of microtasks that occur within the content users are already consuming. I present FlipWord, an application that automatically inserts second language vocabulary in line with a user's native language on every website visited. These words are generally in-context, emphasized through spaced repetition, and reviewable through micro-challenges. User retention over four months has been particularly high, with positive user feedback and reviews.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Second language learning is a difficult but worthwhile struggle around the globe. Learners who succeed in developing second language skills have greater chances to improve social standing [11], are more compassionate of multiple viewpoints [12], and have the opportunity to interact on a more personal level with a heightened number of individuals.

Learning a second language is also perceived as highly difficult. Like anything, it requires persistence, frequent practicing, hours studying, and subsequent revisiting. This difficulty is exacerbated when paired with competing messages on what is the easiest and most efficient learning method or style.

In this thesis, I explore a novel concept of passive microtasking for second language learning. I implemented this concept in FlipWord, a Chrome extension that teaches and tests users' vocabulary in line with content while users browse the Web. The cost to the user is minimal, requiring mere moments of time interacting with FlipWord, without interrupting the main task of Web browsing. Through four months of a beta release, I find that users are continuously interacting with the tool. I also establish that both language novices and veterans are building new knowledge.

1.2 CONTRIBUTIONS

- The invention of a new concept - *passive microtasking* - to facilitate language learning within arbitrary content

- A technical infrastructure that teaches and reinforces learned material by implementing the passive microtasking concept, making learning easier for the user

- Results from a four month field study revealing users are willing to allow content modification in exchange for language learning

Figure 1.1: FlipWord active on a page of content from nytimes.com, with "在" being hovered. Items in blue are informational, items in orange are small tests.

CHAPTER 2

DESIGN PRINCIPLES

Language learning is generally perceived as difficult. Learning any content requires time and persistence. Throughout this thesis, I will explore the following four features which previous work has addressed as both beneficial to learning and implemented in various forms.

The main design principles focused on in this paper are reducing the barrier to entry, microtasking, spaced repetition, and persistence. Many others exist (guidance, collaboration, motivation) [28], but are not core to the initial implementation of FlipWord. They will be considered for future improvements.

2.1 REDUCED BARRIER TO ENTRY

Getting started is the hardest part, when unknown knowledge is most foreign. "Without sufficient motivation, even individuals with the most remarkable abilities cannot accomplish long-term goals...On the other hand, high-motivation can make up for considerable deficiencies both in one's language aptitude and learning conditions" [26].

Nothing can ever replace an internal passion to learn, but tools have been created to try to ease the curve. Teachers have played this role for centuries, introducing content that would otherwise be hidden. What's more, teacher humor is strongly associated with language learning [27]. This can be extended to enjoyment. Duolingo [20] strives to be "fun and effective", delivering gamified content in "bite-sized" lessons while Anki [16] uses spaced repetition to further reduce this barrier.

2.2 MICROTASKING

The ability to complete a hyper-granular task has gained recent traction with systems such as Amazon's Mechanical Turk where a large pool of individuals complete small tasks for payment [22]. The usual use of the term is not clear enough to describe what FlipWord aims to achieve.

FlipWord is passive, in the sense that users are accomplishing a secondary task, without losing focus or context on the main task. Much like unlocking one's phone [2], a microtask takes only a few seconds at most per interaction and then the main task continues as normal. Other terms used with similar meaning include Integrated Microlearning [22] and Wait learning [1].

Unlike unlocking a phone, the language equivalent is still part of the original content, just in a different language. After viewing a translation, the user should then be able to read it in line with the original content as if the word is already part of their own native language, thus effectively skipping the interaction requirement, while still doing the task of building second language skills. This is a major foundation for the core implementation of FlipWord.

2.3 SPACED REPETITION

Spaced repetition is a well-tested method for maintaining gained knowledge over an extended period of time [14]. The most common implementation is SuperMemo [10]. Some form is often used for flashcard systems, particularly common for language learning [15], including systems such as Anki [16] and Memrise [19].

Rather than learning clustered content in distinct time segments as is common in a classroom setting, spaced repetition revisits material before it would

otherwise have been forgotten. In many studies, this has been shown to give users a much greater advantage at remembering material, even months after initially studying. This works for learning, even as humans age. "The reason that children appear to be better learners may have more to do with their environment..." [18].

2.4 PERSISTENCE

"...sustaining motivation remains a major barrier towards progress..." [17]. When using classes, Anki, Memrise, DuoLingo [20], or other methods, it is up to the user to revisit these tools in the future. All occur in isolated environments that require opportunity cost of other habits to be spent instead studying. Once the environment has been left, the user must actively make the decision to re-enter, and do so both regularly and over extended periods of time. This is analogous to application usage on Android devices [4]. The state of any given tool is either in-use or not, with no crossover.

Spaced repetition and microtasking are not sufficient without persistence. Immersion, either physically in a country or learning method, is a commonly discussed concept. Landing in a completely foreign country without any knowledge of the language makes gaining traction with the language. On the contrary, after a little dabbling, much of the learning curve will have been passed.

CHAPTER 3

RELATED WORK

3.1 ACTIVE LEARNING

The standard approach to learning is spending time in a closed environment with the main task of learning and no secondary task. This requires a change in habit and a large time investment.

Anki [16] is a very popular spaced repetition flashcard system that works across mobile and desktop platforms. It uses a spaced repetition algorithm and features easily importable cards with any type of information. It was developed in 2006.

Memrise [19] is a social flashcard website and mobile application that employs gamification and *memes*, images that represent the card content, often with transliterations or mnemonics. It allows anyone to create a custom course and promotes friendly competition within both courses and languages. It is available on web and mobile. It was launched in 2010. Picard [17] is a system very similar to Memrise, with a more creative focus for user memes. It is run on the Facebook platform.

Duolingo [20] is a language learning tool that focuses on small language courses and gamification. Users need to progress through stages, which require completion of a set number of challenges without getting three incorrect. This gives instant gratification and excites the user to continue learning. It uses a welcoming interface and a tree of small challenges to slowly build and refresh knowledge. It includes vocabulary, speaking, and translation training. The tool was created in 2011.

Each of these tools has some form of spaced repetition, which should ensure steady learning. Unfortunately, their major flaw is that which also plagues classroom learning, there is no safety net once the class is over. As long as a user opens a learning tool daily for sufficient time periods, spaced repetition will succeed. Without either being present by default or an intense intrinsic or extrinsic motivation [9], usage will cease. Beyond persistent motivation, studying using an active method requires an investment of time. FlipWord removes the barrier by adding several pause options and a non-isolated environment.

3.2 PASSIVE LEARNING

Several tools have attempted use some combination of in-context and/or passive learning. This occurs within a shared environment, with learning usually as a secondary task. I explore a subset of them here.

WaitChatter [1] is a Chrome extension that makes modifications to the Google chat window within the Gmail interface. It attempts to take advantage of moments of inactivity during chat, appending a small flashcard window for practicing vocabulary. To improve outcomes, it includes context-aware vocabulary terms based on words in the conversation.   Although WaitChatter finds good timing points for inserting learning opportunities, its implementation has several weaknesses. It has a complex install procedure, including jumping between three tabs and reverting the new Hangout chat to the old Google chat system. Over-time, it may prove rather intrusive within chat due to its constant presence. It has experienced a drop-off of at least 70% of users in just over a month.

Mainichi [25] is a new-tab replacement within Chrome, specifically for Japanese and Mandarin. Upon opening a new tab, it displays a single vocab word

and associated image. It has a simple two-tone color scheme for an extremely simple interface. Though Mainichi has a refined user interface, its main shortcoming is due to timing. Time spent in a new tab window pale in comparison to time spent browsing. It is already a secondary task to browsing, which make learning a tertiary task. In addition, the words selected seem to be determined solely on the pictures the authors have drawn rather than value in the language. It also has apparent random selection for each display choice.

Together, these tools provide great value. Learning is a process of discovering new information and reviewing previously discovered information. None of the tools tackle both at once, while also remaining mostly passive. They also appear within niche tasks, rather than during more general and common tasks such as Web browsing. FlipWord builds on their success with greater spread and frequency.

3.3 PASSIVE MICROTASK LEARNING

Passive microtasking is a very new space. The secondary task occurs within the content of the main task and requires a very small time investment for learning to occur.

Two tools pioneered the space FlipWord is entering. Language Immersion for Chrome [7] was the first of its kind. It selects arbitrary chunks of each page and directly Google translates them. There are five selectable levels of page conversion, going from small word clusters to entire paragraphs. Viewing a translation requires clicking the replaced text for in line replacement, while hovering triggers the text to be spoken.

Flewent [3] was released shortly after and builds on Language Immersion as a learning tool. It only translates individual words, randomly selected from visited websites. Those words are saved to a personal dictionary that will begin repeating and focus on that closed dictionary once full. The default dictionary size is 500. Clicking the word shows a popup with lots of extra options, and an additional click within the popup causes the word to be spoken. It is very fast, reliable, and generally well-designed. It also has conversion percentage control and modifiable dictionaries. Flewent reinforced several design decisions once discovered, and has been a great source of motivation.

Both tools created a great base to build on, but fall short of being great learning tools. Neither uses any prior knowledge of the language being studied, or how it relates to the user's native language, other than a direct translation. Both also only go from English to other languages.

Language Immersion's selection of words and partial sentence translations can help a user review language content, but is more confounding than helpful. This causes a large majority of the translations to be the exact quality expected of Google Translate, causing obscuring of the original content. Another large downfall is speed and reliability, usually taking a 5-10 second time frame to convert a page and often break, leaving a large animated obstruction on the page.

Flewent's word choice is one of its biggest weaknesses. Selecting based on page content is clever, building based on what a user would see. However, this leads to approximately 10% broken words, including non-words, acronyms, punctuation, and translation artifacts. It also selects very uncommon words which are rarely used in English, sometimes paired with incorrect translations.

CHAPTER 4

USAGE SCENARIO

Kai is a college student preparing to take her first foreign language class, Mandarin. if things go according to plan, she will be headed to China a year later for an exchange program and wants to increase her Chinese to a sufficient level for decent day to day communication.

Figure 4.1: An example of how FlipWord translates words on a page. Before (top) and after (bottom) hover.

She decides to use FlipWord to make the initial learning easier, and clicks one button on the FlipWord website to instantly download and install the extension. From there, she clicks a button to open the FlipWord popup, selects Chinese as the language to learn, and clicks Continue.
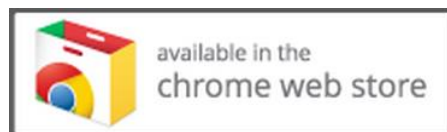

Figure 4.2: One-click install button on website.
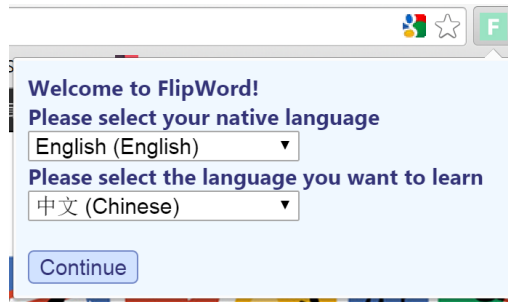
Figure 4.3: Language selection.

From there, Kai uses the tool for two days, learning several new words per day. While browsing for a total of 4 hours, she sees 230 flips out of a total of 1800 flipped, hovering over only 76 of them. She also answers 41 flashcards and types 9 words. The first time she receives a small "+60" notification for a correct answer reinforces that she is learning and builds excitement for learning. After hearing the words spoken repeatedly, she is even able to begin pronouncing many.

On the third day, Kai is hurrying to finish writing a paper and does not want to deal with Chinese while working. She opens the popup, pauses the application, and continues working. Later that night, after a preset delay, FlipWord automatically starts working again. She continues progressing through learning new vocabulary.

The next day, Kai has free time and decides to actively spend time studying flashcards. The FlipWord system knows the words she has worked on, and automatically progresses to teach new unknown words and review those that she hasn't seen in a few days. After spending an hour, Kai learns and reviews 35 new words, correctly reviews 40 words seen the previous 3 days, and misses 7 of the old words. When the 7 missed words reappear, she learns from the mistakes and answers them correctly. All old and new words will reappear accordingly in the coming week, quickly integrating into her known vocabulary.
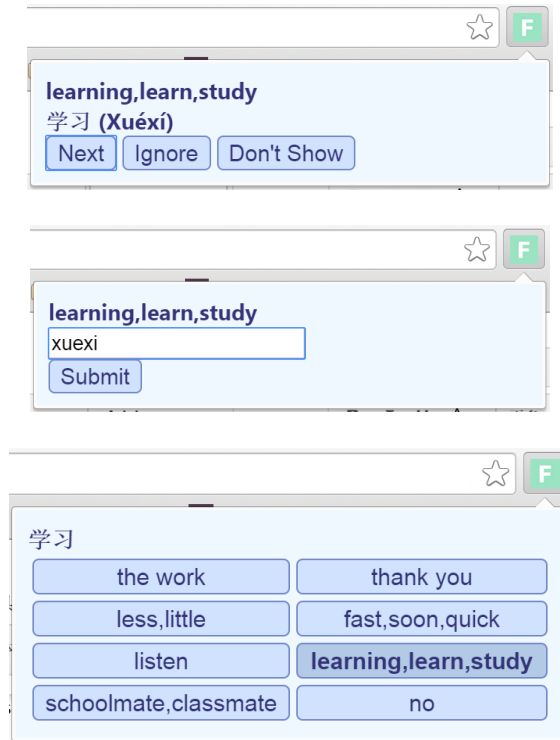
Figure 4.4: Popup flashcard stages

After studying, she is excited about her progress and increases the frequency of flipped words per page from three words to nine. Over the next hour alone, Kai sees three times the words flipped per page compared to the initial rate, nearly the same amount as the first two days combined.
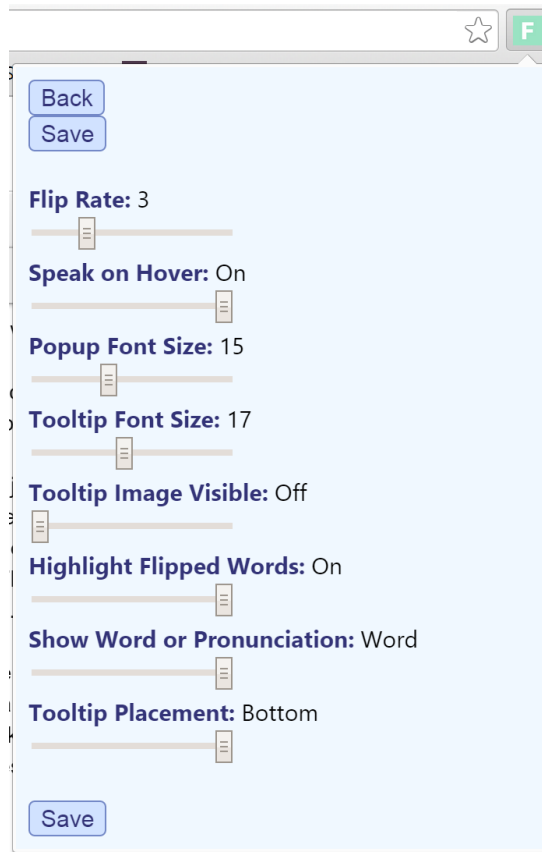
Figure 4.5: FlipWord settings, main option is Flip Rate

Over the course of the three days, Kai spent about one hour and five minutes learning. The additional five hours were already going to be spent on the Web, and she still was able to read very interesting international news articles and a lot of Reddit posts.

If Kai had used only a stand-alone flashcard system, she would have spent an hour learning. The vocabulary learned would also always be out of context and not reviewable if she didn't consistently use the flashcards. The number of possible learning opportunities due to content quantity, total time spent browsing, and continued immersion make her expedited learning experience possible.

CHAPTER 5

DESIGN

5.1 USER INTERFACE

The major design challenge is to spur learning through passive microtasking without impacting the main task. By injecting content directly in line in content being consumed, original content is inherently being removed. With language learning as the secondary task, it is paramount to strike a balance, replacing the maximum amount of original content possible without becoming irritating to the user. FlipWord acts to force immersion and persistence without adding the usual layers of struggle associated with changing one's entire lifestyle.

Replaced elements are, by design, small. They are intended to be encountered independently once arriving at the replaced content. Only the corresponding translation information is displayed, in a manner as to not obscure or make difficult pursuing the following content. While highlighting draws slight attention, it does so minimally to keep from jarring the user experience.

FlipWord has several types of functional in line word replacements: one type contains the corresponding translation, three types test user knowledge. The four current types are: Flip, Flash, Type, and Speak elements. A Flip shows a word's translation in the native language in a tooltip when hovered over, a Flash shows a multiple-choice flashcard, a Type allows the user to type in a translation, and a Speak allows pronunciation practice.

To make the passive microtasking paradigm work, all elements' core feature can be fully utilized without a single mouse click. All interactions cause an instant change, showing different highlighting and/or the addition of a tooltip, to make it clear

that FlipWord is not native content and is active. To preserve size, elements exist with an almost identically sized footprint on the page as the original word.

In order to keep things absolutely passive, only Flip elements are in the desired language. With unknown words, they need to be hovered to read the content. All testing nodes originally show the native language to reduce guessing based on context.

A novel features of this system is that a user will never forget to turn it back on. Every other application [4, 5] requires the user to open the tool. If they ever want the service to pause, the only option is a full off switch. FlipWord currently offers two ways to pause, one for all pages and one for the current page. Thanks to these, it will always automatically resume, as long as the user returns to their web browser. There are also multiple ways to control the number of changes per page, so a reduction can be made without a full service halt.

CHAPTER 6

IMPLEMENTATION

6.1 PLATFORM

FlipWord is currently implemented as a Google Chrome extension. Chrome runs on Windows, Mac, and Linux, as well as Chrome OS, covering the vast majority of desktop operating system users. It runs on every visited webpage during browsing. For speed, all client-side code is written in pure JavaScript without any external packages and can continue to function for the user, even if the server becomes non-functional or the user's device goes offline. There are a few supporting HTML and CSS files.

The server is hosted on an Ubuntu Linux server hosted on DigitalOcean and built using Node.js, Express, and MySQL as the database. Upon initial extension install, the client generates a unique identifier. When the user selects a language, either the pre-computed translation list will be downloaded or will be generated and stored locally indefinitely. Any interactions within the interface are immediately communicated to the server. No other server connections occur.

6.2 ELEMENTS

6.2.1 In-page Elements

The most basic FlipWord element is a Flip (flipped word). Upon reaching a Flip, the user will likely notice the word is not part of their native language and has a slight blue highlighting. To reveal a small tooltip with the replaced word and other possible meanings, the user simply hovers their cursor over the element. When

hovered, the page will speak the word in the non-native language to enforce auditory learning. Moving the cursor away hides the tooltip, allowing the user to continue consuming the original content.

The default number of unique words flipped is three. For very common words (a, the), this results in lots of copies scattered around the page. Less common words or smaller pages might only have three words changed. Sometimes the page simply has no useable content. On average, around five words are turned into flip elements per page.

Figure 6.1: Flip element

In addition to informational Flips, there are several types of testing elements. These appear in the user's native language with an orange highlighting, differentiating them from the desired language. Interspersing tests into the experience reinforces learned material and augments the system's passive understanding of the user knowledge.

One testing element is a Flash (flashcard). Upon hovering, rather than showing the word's equivalent, a number of possible translations are shown to the user. A selection can be made by momentarily hovering or clicking. Any choice will cause the non-native word to be spoken and the tooltip to close. If it is incorrect, the user can hover again to retry. If the selection is correct, the entire element turns green, shows a new tooltip with number of earned points, and changes the Flash into

a Flip. All points are currently fixed at 60, but will be dynamically allocated in future versions.

Figure 6.2: Flash element

Another testing element is a Type (typed response). Once hovered, the Type changes to a lighter orange, shows a darker border, and begins blinking a caret. This brings the Type into focus, which then allows immediate typing of the translation. When typing of the answer is complete, the ENTER key must be pressed. The Type then changes back to a Flip and shows earned points if correct. Acceptable answers are any mixture of the actual word or pronunciation.

ognitive strategies, emerged as positive predictors of
paying attention to word formation, contextual
two test scores. However, visual repetition of new wo

ognitive strategies, emerged as positive predictors of
paying attention to | word | formation, contextual
two test scores. However, visual repetition of new wo

ognitive strategies, emerged as positive predictors of
paying attention to 单词 formation, contextual
two test scores. However, visual repetition of new wo

ognitive strategies, emerged as positive predictors of
paying attention to 单词 formation, contextual
two test scores. However, visual repetition of new wo

d using a wide variety of vocabulary learning strategie
cognitive strategies, e +60 as positive predictors of
paying attention to 字 formation, contextual
two test scores. However, visual repetition of new wo

cognitive strategies, emerged as positive pred
paying attention to 字 formation, contextual
s. However, visual repetition of new words wa

Figure 6.3: Type element

The final type is a Speak (voice recognition). It looks identical to a Flash

before hovered, but with a small microphone icon. When the cursor hovers over the

microphone, a notification appears and the user is able to speak the translation into

their computer. The user is allowed to record several times, until a correct input is

received. It will then convert back to a Flip, just like the Type. A Speak allows a user

to quickly hone their pronunciation, "a difficult and important skill" [6] for learning a

second language.

that their actual performance wasn't
on par with what 🎤 their resume and
credentials suggested.

that their actual performance wasn't
on par with 什么 🎤 their resume and
credentials suggested.

**Figure 6.4: Speak element**

6.2.2 Popup

All functions outside of the main interactions, including pausing and increasing

rate, are available within FlipWord's popup. It is accessed by clicking the 'F' icon in

the browser toolbar.
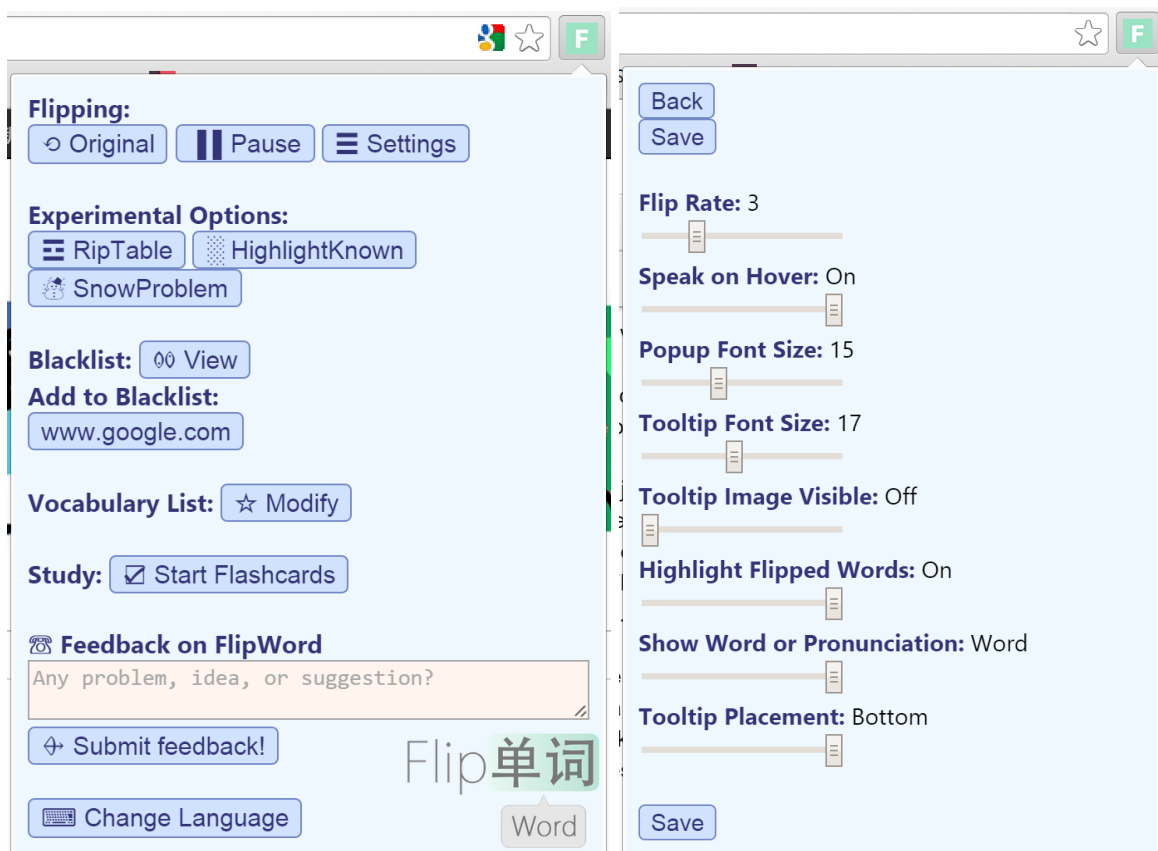
Figure 6.5: Popup and settings window

6.2.2.1 Taking a Break

Occasionally, the user wants to access desired content unobstructed. FlipWord has two methods for making this a reality, neither of which are an off switch.The first and fastest is changing the current page back to its original form by clicking 'Original' in the popup. Any replaced words are reverted to their starting form.

The second is a pause switch. It requires an extra confirmation click, to dissuade users from pausing instead of just reverting the current page. This will turn off any flipping for 30 minutes. This threshold seems to be long enough for most users to complete any task that was causing them an issue without ever becoming a burden.

If they really find that FlipWord is too encroaching, additional options exist to turn down the rate of words flipped. If there are certain pages that FlipWord interferes with or ones where users never want to be bothered, a site blacklist is available. The last resort is to disable the extension completely, which can be accomplished by accessing Chrome's menus.

Figure 6.6: Website blacklist settings

All of these options are in the popup and take a minimum of two clicks. Turning it off completely can be done with a minimum of three clicks, but is less straightforward for an average user, many not even knowing it is an option. Due to the lower barrier, the pause switch has proven to be successful for user longevity.
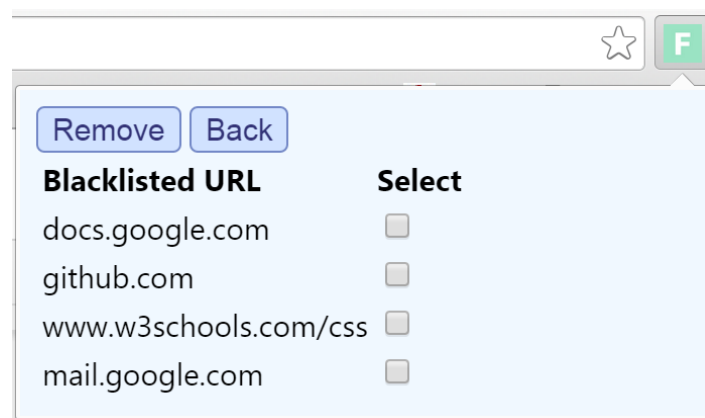
6.2.2.2 Flashcards

When time is available and active learning is desired, FlipWord has flashcards available. Based on all previous user interaction and words learned, it automatically figures out what to review. Flashcard functionality can be seen in Figure 4.4.

6.2.2.3 Gamification

FlipWord's gamification is currently two-fold. The first feature added was regarded as the best part of the learning experience by a subset of users. This was Types. The simple addition of occasional tests made users compete with themselves for learning validation.

This was once again built upon with the addition of Flashes and Speaks. Of all the features added so far, the most applauded from users is Flashes. The simplicity of a single selection to validate and refresh knowledge resonates well with users.

The second current method is allocating points for correct responses. A Flash, Type, and Speak all award points based on answer quality. Correct answers award high point values. Some incorrect answers provide the option to retest (Flash, Speak) a few times with slightly diminished point values, and all types give a somewhat smaller point value for incorrect answers.

Giving instant feedback enhances learning [23], as does gamification [24] (based on feedback, it is positive). Pairing these together encourages a small amount of non-passive interaction, with a much greater benefit due to language recall [13, 21] based or qualitative user feedback.

6.3 VOCABULARY

For an initial implementation, FlipWord builds off word frequency data sets for 21 languages compiled based on parsed scripts from the full catalog of 2012's MovieSubtitles.org. The top 1000 words for each language were kept, with curse words, names, and words from other languages (usually English) manually removed.

These lists were then translated from the language of study back to the user's native language to create a many-to-one word pairing unique to each direction. A server-side script sent each word individually to Google Translate and parsed the JSON response. If it was empty, the word was removed from the list.

For each word, all translations were extracted. Any translations with less than a 1% chance of use were subsequently removed. The resulting lists were saved locally to user machines and made modifiable by users to individually improve any poor translations. All word flips that occur on a site use this reverse translation list, making them a one-to-one pairing, seemingly more reliable than a direct translation in the opposite direction.

CHAPTER 7

FIELD STUDY

7.1 METHODOLOGY

My analysis of FlipWord usage is based on a four month field study from March 15th to July 15th, 2015. Each usage event recorded includes a user identifier, timestamp, page of occurrence, event name, and (if applicable) an associated value. A survey was also conducted at the initial release of the tool. The final piece of information collected is through a user feedback box within the tool. It is shown in Figure 6.5.

The survey was sent to user emails as a link. It was created and hosted using Google Forms. None of the fields were required, other than email. Usage statistics and user feedback were gathered through the tool on a server.

7.2 USER SURVEY

Completion of the first user survey was required to download FlipWord. There are 83 responses. Questions include what languages are desired for study, what features are most exciting, most helpful learning methods for past study, ideas that would improve on those methods, and past tools used. Age was also optionally entered, with 18 as the minimum and 60 as the maximum with an average of 22.8 for all entered values.

For desired languages, a total of 228 were entered for an average of 2.7 per user. Spanish and Mandarin were the top choices with 41 and 37 responses respectively, but Chinese was chosen as the first choice 22 times versus 19 for Spanish. French, Japanese, and German were next with 25, 23, and 23 respectively.

7.3 USAGE STATISTICS

FlipWord has been active consistently over the course of four months for 80 users. Different updates and release dates have occurred across the group. Numbers and graphs are all adjusted accordingly for consistent version and user start dates.

The effectiveness of FlipWord as a potential learning tool can't be understated. When analyzing a user's first month of activity, it becomes clear that users continue to use FlipWord and incorporate it into their daily browsing habits. The number of words they interact with each day plateaus to 5 after 15 days of usage. This is mainly due to less common words. The number of times a particular word gets flipped drops to near zero in only a few days, demonstrating users quickly learn the word's meaning. The flips by the user that occur past the first few days are not from viewing old words, but rather users encountering new ones. It turns out a typical user discovers and learns 46 new words within the first month of using FlipWord.
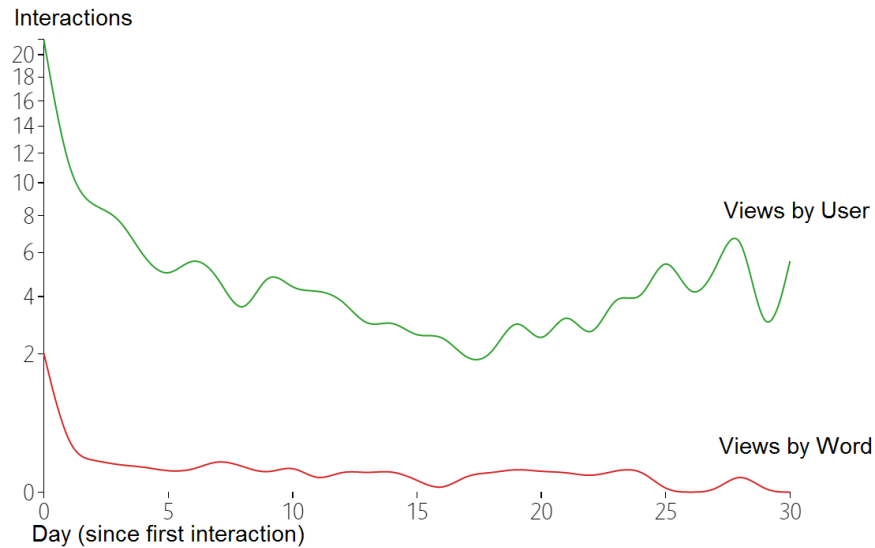
Figure 7.1: First month of average FlipWord usage

After a major update fixing many of the bugs based on feedback and mending a data recording issue, usage remained consistent. Over the course of 11 days, the 33 upgraded users had a total of 40450 words flipped on visited pages. This is an average of 1226 per person for an average of around 111 per day.

For those 40450 words flipped, users encountered and interacted with a total of 1804, or 4.5%, of them. This equates to 55 total per person, 5 per day. Although it cannot be precisely determined, it is estimated that nearly 60% the flipped words encountered were simply read by the user, indicating that the word was learned. This totals to 55 words interacted with per person, and approximately 80 additional words read.

Currently the beta users are located in 8 countries and utilizing 23 language pairs to learn a country's language before they travel to it, help reinforce lessons they are learning in class, and to practice the language of a new country they call home.
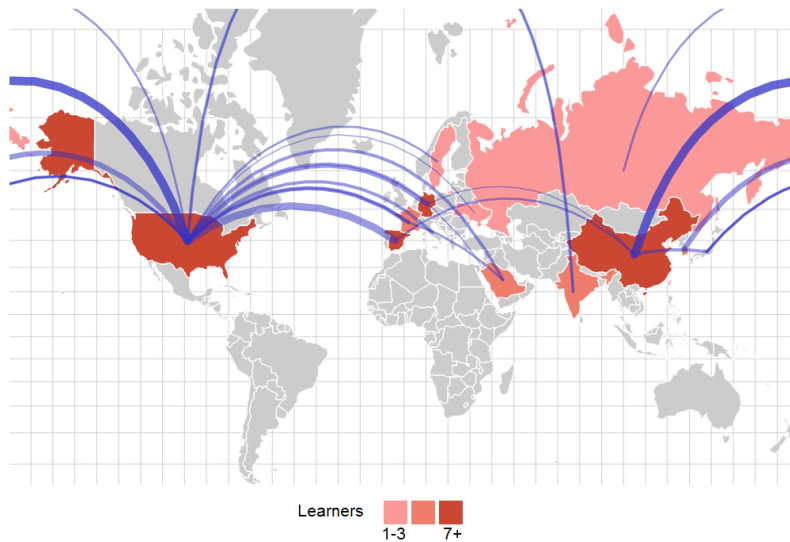
Figure 7.2: Connections between language pairs (country of language base) and highlighting based on total number of users studying that language
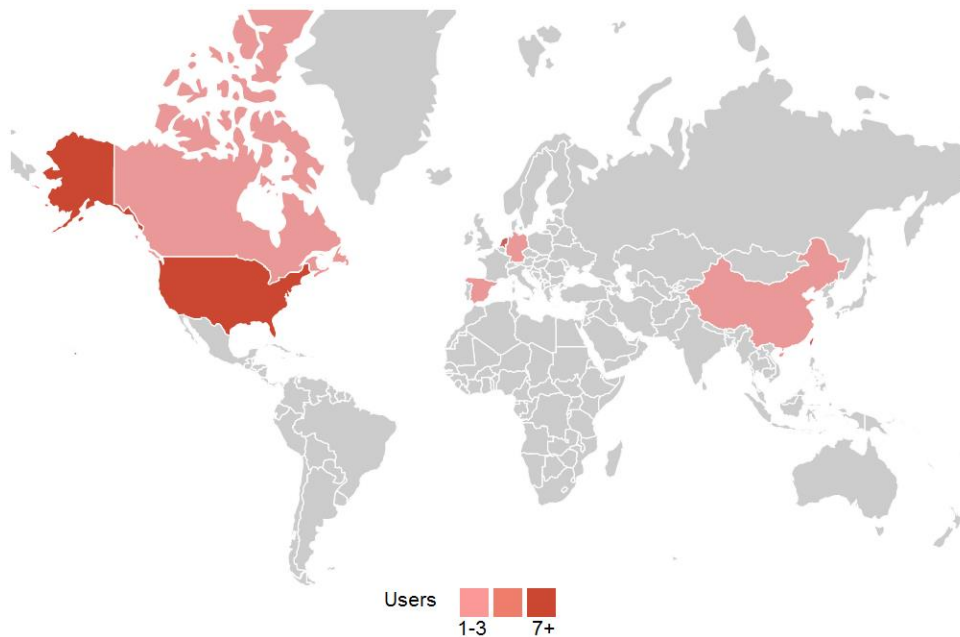
Figure 7.3: Users by physical location

7.4 USER FEEDBACK

For the duration of the four months, a feedback box has been available directly in the popup. This has been users' main method of providing reports of broken features or websites, and their feature ideas and suggestions.

From that box, a total of 68 responses were submitted, totaling 1984 words. Of 30 bug reports, 26 have been fixed. From 25 suggestions, 17 ideas were

implemented. Three responses mention improving the generated data. The remainder include commentary, repeats, and one which is an unchangeable Chrome feature. Several of responses were likely dropped due to a server issue and others are partial due to a limitation introduced in the MySQL database.

Without the active feedback, many bugs would have been left undiscovered. The high desire users have for the tool to be polished clearly shows the vision they have for FlipWord to add value to their daily life.

7.4.1 Selected Feedback

7.4.1.1 Fixed Bugs

- "For businessinsider and forbes the articles end up placing the changed words on their own line"

- "Not positive about this - but it *may* break editing on Wikipedia"

- "When the flipped word is on either extreme side of the page, the popup becomes cut off. You could probably detect mouse position at the point of hover to see if you need to shift the box over. :)"

- "some "phrase words" that are made up of 2 words, like "no one" get translated into "no un" instead of "nadie" for spanish."

- "The word-flipping sometimes slows my browsing down quite a bit, which can be annoying at times"

7.4.1.2 Added Features

- "Can there be an option to switch the language?"

- "Also having a confirmation "feedback was sent" would be nice :)"

- "Would be nice to be able to turn off the pronunciation part, just in case we want to."

- "Can we have a "slider" on how many words/what difficulty of words we want to replace? (Otherwise, it's great!)"

- "Maybe you could allow custom font size for the flipped card? For Chinese characters they can be tricky to read if they're the same size as English words."

7.4.1.3 Troubleshooting and Suggestions

- "I have a hypothesis on the flaky delete. Two tabs are open. Both flip on "foo". I delete "foo" in tab 1, but in tab 2 I interact with the flipped "foo". This causes an update and "foo" is reborn."

- "Quizzes when you have multiple words that flip on the same thing are impossible. :("

CHAPTER 8

DISCUSSION

8.1 USER INTERFACE

The decision was made early in the process that any changes to the page would be minimal, both in speed and page real-estate. Other extensions are usually are quite slow in at least one of them. Early versions had more distracting color schemes and larger elements. Several users asked for the option to remove the highlighting altogether and to reduce the size. These suggestions led to both the implementation of turning highlighting off and reducing the default look.

Initially much of the tooltip design was based on Memrise. Based on its success, images representing the word and access to more word information were put within the tooltip. This led to it growing in size, load time, information density, and even causing an https issue, all for more than what users wanted in a passive microtask. After quite a few feedback responses on lack of image relatability and how to improve the design, tooltips were reduced to the minimum necessary information.

8.2 WORD SELECTION ALGORITHM

Deciding which word to include in each page has been the hardest theory to figure out. Spaced repetition works great when the data set is known, but breaks down when certain words appear rarely. In addition, the long term goal is to implement other learning tools and have the same algorithm work across them, selecting words in all tools based on all previous interactions. This has led to a model which works with passive or active interactions, either independently or mixed.

The selection algorithm has two sets of timing mechanisms which operate independently per word, one for long-term spaced repetition and another for determining when a word has been thoroughly reviewed. If inactive, a word is not eligible for selection to translate or test until a set time in the future. If classified as active, it is eligible until the user has sufficiently reviewed it or believed to had ample opportunities. This is necessary since a user may never interact with a word. Regardless of already learning it or never learning it, the algorithm is not allowed to get stuck on the instance.

All words start as inactive, as users have previously seen none of the words. Words can easily be excluded from learning if the user has prior experience. Upon page load, all active words and inactive words which have exceeded their wait period are aggregated. If the number active is not sufficient (currently 10 needed), words previously unseen are added. Initially the first 10 words become active to fill this need. Inactive is defined as never reached or waiting until a future time to review.

When a word is classified as active, it is assigned a score to reach based on previous interactions. Previously unlearned words require a high level of interaction to learn, while those reviewed several times need just a few moments to refresh user memory. Every interaction moves a word closer to a period of inactivity. Events scored include the number of times words appear translated, are interacted with, and are tested, in order of lowest weight. Until the score is reached, the word is eligible for any type of review.

Once the score is reached, a wait period is determined. The more a word has been tested correctly, the further it is pushed into the future. Incorrectly tested words are set to return fairly quickly. If only interacted with passively, words have a higher frequency of review.

This system is quickly changing based on usage and user feedback. Building on it is one task for future exploration.

8.3 IMPLEMENTATION

As a Chrome extension, all code is run on every page at load time. Without proper precautions, these can easily slow down browsing speeds and impact user perceptions of Chrome. Due to the fairly heavy page analysis and modifications that take place within every site on the Web, performance has been a top priority.

To keep things speedy on even the slowest devices, FlipWord uses no frameworks or libraries, such as jQuery or a tooltip. Many JavaScript functions execute at 100x speed compared to their wrapped jQuery equivalents.

The greatest two technical challenges have been keeping elements consistent across pages and implementing Speak elements, which was done against standard web security measures.

To stay consistent across pages, it has taken a great understanding of the DOM and CSS hierarchy, and persistence in repairing issues. Initial versions often caused shifts in content as elements loaded, which led to further speed issues. Page redraws are the single most expensive computational operation in a webpage.

For Speak elements, a rather complex route to success was taken. Chrome extension code cannot gain direct microphone access. To make this possible, a custom page with HTML5 Voice Recognition is loaded within an iframe, putting it within its own isolated sandbox, but first passing the iframe a reference to the Speak element. Once a word is recognized, the iframe uses a special message passing protocol only available to extensions to a background page, and then to the original page which handles the word accordingly.

8.4 DATA COLLECTION

When releasing a beta version of a fledgling system, there are often issues to fix. In the early releases of FlipWord, one of these issues in particular caused a major hiccup in data gathering.

For an entire month after release, a substantial amount of data failed to reach the server due to a basic web security browser implementation. All data was being sent to the http version of my server, which would be fine if it was my own hosted site. When running on other sites using the https protocol, http content is blocked to protect the integrity of the content.

While it is estimated that 15% of all websites have SSL certificates [8], a much higher percentage of major internet companies have SSL active on their sites, meaning a much greater percentage of usage was failed to be tracked. In order to account for this, the collected data is assumed to represent average usage.

CHAPTER 9

FUTURE WORK AND CONCLUSIONS

9.1 FUTURE WORK

FlipWord started as a specific idea for learning a second language while web browsing. Over time, this vision has grown to encompass other content and methods of learning.

9.1.1 Better Placement Algorithm

The current number of flipped words viewed per day is rather small. This is likely caused by how sparse they are on a given page. Coming iterations will improve on the selection algorithm to replace in major paragraphs and headers early in the document to flip, over those that are in footers or content that users need to scroll down to read.

9.1.2 Mobile Implementation

Operating in Chrome is the flagship platform for validating the model, but moving to mobile of great interest. Desktop usage is declining with mobile phone usage ever increasing. Injecting FlipWord within a mobile browser or other context will bring passive microtasking to a greater breadth of user interactions, and hopefully a greater impact on learning.

9.1.3 Different Domains

The next large expansion will be to move beyond the domain of foreign language learning. Currently several verticals are planned.

9.1.3.1 Native Vocabulary

The same passive microtasking method can be directly applied within a user's native language to improve skills. Using the same language base should result in fewer grammar errors and a less jarring experience when reading a word. This should improve user willingness to increase words flipped per page. For this expansion, no code change is necessary.

An example data set is words on the SAT study list are fairly advanced and uncommon in everyday life. Replacing their more common equivalents with the SAT word in line will build necessary connections in a user's mind. Other data sets could include technical terms for different professions (programming, medical), slang, or vocabulary that has faded from the colloquial language (Shakesperian).

9.1.3.2 Useful Facts

Rather than replacing words within a sentence, another possible option is to inject sentences within paragraphs, leaving the original content otherwise unaltered. Reading small facts repeatedly should be enough to embed it within the user's memory.

An example use case is a student taking a history class. He has to memorize 30 facts for a test the next week. By spending time on Facebook, he discovers that his friend just bought ice cream and that Napoleon died May 5th, 1821. A deadline could be set to make sure they are reviewed amply by the time needed.

9.1.4 More Passive Methods

FlipWord uses a specific type of passive microtask. Utilizing similar tasks in additional places will help fill the maximum gaps possible to emphasize all parts of a language in as many moments as possible.

A tool to enhance grammar and sentence building will utilize additional moments of inactivity or waiting to target non-vocabulary based skills. These include, but are not limited to: opening a new tab, waiting for a download, navigation to sites where little interaction is usually executed, and actions within FlipWord such as while pausing. These may consist of sentence translation based on known vocab, information on grammar rules, or small quizzes on correct sentence structure.

9.1.5 Learning Feedback

To reassure and reinforce learning, additional elements to create a direct feedback loop are in development. Highlighting progress with graphs of interaction and trends, both individual and aggregated will give users a better understanding and positive reminder to continue.

An additional method to strengthen grammar and emphasize progress using natural language based on user knowledge and colloquial language scraped from the Web. It will show sentences based on temporal learning, highlighting sentences understood months ago, a few days earlier, as of the same day, and even what will be known in the near future. This is possible thanks to independent understanding of user interactions with each word.

This natural language feedback can be expounded upon through a 'language diary' highlighting daily accomplishments in a non-repetitive manner. Based on a large template set, weighted scoring of feature importance, active avoidance of repetitive entries, and user-augmented content, this diary offers a quick or detailed view of progress in a user-friendly manner, and the chance to see past milestones.

9.2 CONCLUSIONS

Language learning currently requires a great amount of time and effort. Hours must be spent dedicated to the task of language learning, blocking out any tasks which would otherwise draw time and attention.

This thesis introduces the novel concept of passive microtasking, and a software manifestation in the form of FlipWord. It is a step towards an easier learning future. The hours that go wasted studying material with no progress or eventually retrogresses is not small. The presented implementation incorporates small learning moments into desired content rather than all at once in fixed intervals.

Language word relations, data management, word selection, and language progression are all handled by the system., handling spaced repetition to maximize learned words. The three testing types provide a motivation to remember content and gratification for making progress and change back to the default Flip Element. All inpage elements are activated fully without a single click for maximum passive interaction. The FlipWord implementation succeeds in its goal of triggering passive learning in the smallest microtask possible, remaining as the user's secondary task, reducing barrier to entry. The tiered system for pausing increases the chances for persistence.

My evaluation shows that usage leads to on average learning of 46 words in the first month, and that it is consistent through the following three months. Over 100 words visible to a user per day is similar to the amount of content that would be absorbed in a standard day of class,]with mere seconds of extra time web browsing. The language connections around the world is evidence of the system's flexibility as a general language learning tool. Feedback reveals the features users want and that

they find value in its method. It has become increasingly positive as updates have rolled out.

In the future, other information domains and environments will be explored to evaluate their effectiveness at improving outcomes. With additional usage data and feedback, outcomes can be maximized for user gain and enjoyment. It is my goal to carry this work forward, improving how people learn and opening doors to new personal connections.

REFERENCES

[1] Cai, Carrie J., et al. "Wait-learning: leveraging conversational dead time for second language education." *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2014.

[2] Vaish, Rajan, et al. "Twitch crowdsourcing: crowd contributions in short bursts of time." *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2014.

[3] Groff, Andrew. "Flewent." 4 Mar. 2013. Web. 15 July 2015.

[4] Böhmer, Matthias, et al. "Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage." *Proceedings of the 13th international conference on Human computer interaction with mobile devices and services*. ACM, 2011.

[5] Brown, Barry, Moira McGregor, and Donald McMillan. "100 days of iPhone use: understanding the details of mobile device use." *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. ACM, 2014.

[6] Cenoz, Jasone, and Ma Luisa Garcia Lecumberri. "The acquisition of English pronunciation: Learners' views." *International Journal of Applied Linguistics*9.1 (1999): 3-15.

[7] "Language Immersion for Chrome." *Chrome Web Store*. 2012. Web. 12 July 2015.

[8] "Usage of SSL Certificate Authorities for Websites." *Usage Statistics and Market Share of SSL Certificate Authorities for Websites, July 2015*. Web. 17 July 2015.

[9] Deci, Edward L., and Richard M. Ryan. *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media, 1985.

[10] Wozniak, P. "Super Memory." *Super Memory*. 1987. Web. 3 July 2015.

[11] Kam, Matthew, et al. "Localized iterative design for language learning in underdeveloped regions: the PACE framework." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007.

[12] Guiora, Alexander Z., Robert CL Brannon, and Cecelia Y. Dull. "EMPATHY AND SECOND LANGUAGE LEARNING1." *Language learning* 22.1 (1972): 111-130.

[13] Cepeda, Nicholas J., et al. "Distributed practice in verbal recall tasks: A review and quantitative synthesis." *Psychological bulletin* 132.3 (2006): 354.

[14] Kelley, Paul, and Terry Whatson. "Making Long-Term Memories in Minutes: A Spaced Learning Pattern from Memory Research in Education." *Frontiers in Human Neuroscience* 7 (2013): 589. *PMC*. Web. 22 July 2015.

[15] Phillips, Joseph. "An Investigation into The Effect of Targeted Vocabulary Learning Using a Spaced Repetition Flashcard System on TOEIC Scores."*青山学院女子短期大学紀要* 65 (2011): 55-62.

[16] Elmes, D. "Anki." *Anki*. 1987. Web. 5 July 2015.

[17] Tuite, Kathleen, et al. "Picard: A creative and social online flashcard learning game." *Proceedings of the International Conference on the Foundations of Digital Games*. ACM, 2012.

[18] D Robson. "Old Schooled: You never stop learning like a child." *New Scientist* (2013).

[19] "Memrise." *Memrise*. 2010. Web. 16 July 2015.

[20] "Duolingo." *Duolingo*. Web. 16 July 2015.

[21] Pashler, Harold, et al. "Enhancing learning and retarding forgetting: Choices and consequences." *Psychonomic bulletin & review* 14.2 (2007): 187-193.

[22] Gassler, Gerhard, Theo Hug, and Christian Glahn. "Integrated Micro Learning–An outline of the basic method and first results." *Interactive Computer Aided Learning* 4 (2004).

[23] Dihoff, Roberta E., et al. "Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback." *Psychological Record* 54.2 (2004): 207-232.

[24] Hamari, Juho, Jonna Koivisto, and Harri Sarsa. "Does gamification work?--a literature review of empirical studies on gamification." *System Sciences (HICSS), 2014 47th Hawaii International Conference on*. IEEE, 2014.

[25] *Mainichi*. 2015. Web. 17 July 2015.

[26] Dörnyei, Zoltán. "Motivation in second and foreign language learning."*Language teaching* 31.03 (1998): 117-135.

[27] Wanzer, Melissa Bekelja, and Ann Bainbridge Frymier. "The relationship between student perceptions of instructor humor and students' reports of learning." (1999): 48-62.

[28] Gilson, Zachary, Tristan Wells, and Leah Tatro. "Opportunities for Foreign Language Learning Through New Media."