CS5199 DOER

Student ID: 160000139
Student name: Yuxuan Wang

Description

Project title
An aid to learning and assessing to read foreign languages

Project description
Reading is a fundamental skill to acquire information from all kinds of text resources. For second language (L2) readers and learners, reading is also an inevitable challenge of gaining access to more knowledge and becoming skilled L2 users. Therefore, developing reading ability has become an essential requirement of L2 acquisition.

In the case of linguistics, reading is commonly understood as a multifaceted process to **comprehend** the meaning of the text using readers' morphological, syntactic and semantic knowledge[ref]. Reading is also an implicit process of **learning** and **evaluation**, as readers not only evaluate the texts and their authors by asking themselves *"do I still have an interest in the texts?"* or *"do I agree with the authors?"*, but also evaluate their strengths and weaknesses in their reading abilities by asking themselves *"how well can I understand what the authors say?"* or *"what have I learnt from the texts?"*. Different individuals have diverse educational backgrounds, they also have their reading habits, interests and purposes so that they tend to find and read texts they are interested in.

On the other hand, all readers, even skilled readers, will encounter different levels of reading barriers which distract their attention and reduce their reading efficiency. These reading barriers are much more intractable for L2 readers and learners to address: L2 readers with different native languages (L1) need to understand how their L1s are different from the L2 they are reading. These L1-L2 variations are going to inevitably affect the efficiency and the motivation of L2 reading.

As a result, the development of aids in reading and L2 reading is always a popular topic on language learning and computer science. Nowadays, the prosperity of the Internet and the machine translation technology has significantly changed how people read and develop their reading ability by providing countless linguistic and educational supports in diverse forms. These supports can be classified into three categories concerning the aspects of reading to address the L2 reading barriers and improve reading efficiency:

1. Material support: in the past, people needed to spend large amounts of time in the library or bookstore to search for their interested materials. Today, the Internet and its byproducts, including various search engines, websites and platforms, have provided abundant amounts of multilingual resources that can meet the readability requirements for different readers. In addition to resource support, many language learning software, like Spreeder[ref], also provide dedicated reading ability training.

2. Language assessment support: the majority of reading ability assessments are normally carried out as a part of comprehensive offline language assessments. Recently, some educational organisations have started to provide online computer-based language assessments, such as IELTS[ref] and TOEFL[ref]. Moreover, many language learning software providers, like Duolingo[ref], start to integrate individual language ability assessments into their applications. These assessments are convenient for L2 readers to self-assess and produce accurate self-estimations on their L2 reading ability at different scales.
3. Translation support: advanced machine translation technologies have successfully transformed traditional bulky dictionaries into light-weighted, user-friendly programs. Not only can these ever-evolving programs provide fast and precise multilingual translations on vocabularies, but also sentences and even paragraphs.

Each of these supports meets only a single demand and is lacking integration, L2 readers need to spend additional time opening up multiple individual applications, learning how to use them separately and combining different applications to satisfy their needs. Consequently, hybrid translation applications like Google Translate[ref] and Dictionary Everywhere[ref] are used more frequently today due to their integrated and user-friendly nature, although they purely focus on the translation works and lack assistance on L2 reading ability assessments.

Therefore, an integrated aiding tool for improving L2 reading efficiency and providing L2 reading ability assessments can be developed. This tool is expected to reinforce users' reading ability with regards to the three aspects of L2 reading. In the case of design, it is expected to be implemented as a user-friendly reliable hybrid translation program with additional reading ability evaluation functionalities.

Objectives

This project is an extension on “An aid to learning to read foreign languages”[ref], it aims at developing an integrated language translation and assessment aiding tool targeted at L2 readers to improve their L2 reading ability.

The final software artefact is expected to be a hybrid system composed of three interconnected subsystems with regards to the three aspects of L2 reading:

1. Text extraction and translation (translation system): this subsystem should be able to read the selected text, in the source language, from arbitrary text resources and translate by providing linguistic information in the target language.
2. Note attachment (note system): this subsystem should allow the readers to attach reading notes in parallel with the translation result.
3. Reading ability assessment (assessment system): this subsystem should provide reading ability estimations in accordance with readers’ translation history and recorded notes.

The final software artefact is expected to be implemented as a web browser extension program using HTML, CSS, JavaScript/Node.js and other Internet-related technologies.

Primary objectives

[Translation system] The application shall extract the text selected by the user from arbitrary webpages.

- It shall allow multiple types of selection strategy, e.g. right-drag, copy-and-paste, etc.

[Translation system] The application shall translate input text into corresponding output text.

- It shall display the translated text to the user through small pop-ups or dialogue frames.
- It shall allow users to specify the context of input and output languages, and use appropriate dictionary resources in the translation process.
- *Usability: it should be able to search the text in the dictionary in a short time.*
- *Reliability: it should be able to deliver precise translation results to the users.*

[Note system] The application shall provide an additional input field in parallel with the translated information for note-taking purposes.

- It shall store the <*originText, translateText, noteText*> (note pair) permanently as file or database entries so that the *noteText* can be displayed again when users select *originText* or *translateText* multiple times.
- *Usability: it should allow users to write notes easily and quickly.*

[Assessment system] The application shall be able to evaluate users' reading ability using existing note pairs data and translation histories.

- For a particular user, the reading ability evaluation shall take all his/her note pairs and searching histories, evaluate the relative language level according to linguistic theories and display the evaluation result in terms of textual or numeral reports.
- *Reliability: the evaluation should provide authentic and accurate results to reflect users' real reading ability levels.*

[Software] The application shall support all browsers based on the Chromium browser framework[ref] (e.g. Google Chrome, Microsoft Edge, Opera, etc.)

- The browser extension shall be imported into the browser without reporting errors.
- *Reliability/Security: the application shall not expose any personal information of the readers*

[User evaluation] The application shall be evaluated by L2 readers and proved to be useful in helping improve L2 reading efficiency.

- The participants will be invited to read an L2 article with the aid of this application
- *User evaluation on translation: the participants should be able to fluently read the article by translating their unfamiliar words and sentences.*
- *User evaluation on note-taking: the participants should be able to take note easily*
- *User evaluation on evaluation: the evaluation result should be precise and reflect the real reading ability level.*

Secondary/Tertiary objectives

[Translation system] The application should provide general dictionary interfaces at the backend so that users are able to import other dictionaries.

[Assessment system] The application should provide the evaluation result with respect to some international linguistic standards, such as CEFR[ref].

[Software] The application should support other popular browsers, such as Firefox.

Ethics

This project involves a user evaluation to assess the reliability and performance of software artefact. Participants will be invited to evaluate the artefact voluntarily and anonymously following the school's online evaluation rules[ref].

The ethical issues have been evaluated against the school's preliminary self-assessment form and artefact ethical approval application. The artefact evaluation form will be signed and uploaded through MMS.

Resource

This project requires no additional software or hardware, the artefact development can be completed using the current configuration of school lab machines.