Lab 1 – CESR

Version 3

Blue/ Purple Team

Mohammed Alabdullatif

CS 411W

Professor Price

Professor Brunelle

April 17, 2015

Table of Contents

[**1.** **INTRODUCTION** 2](#_Toc410998409)

[**2.** **CESR PRODUCTION DESCRIPTION**](#_Toc410998410) 2

[2.1 Key Product Features and Capabilities](#_Toc410998411) 3

[2.2 Major Hardware and Software Components](#_Toc410998412) 4

[**3.** **IDENTIFICATION OF CASE STUDY**](#_Toc410998413) 6

[**4.** **CESR PRODUCT PROTOTYPE DESCRIPTION**](#_Toc410998414) 7

[4.1 Prototype Architecture (Hardware/Software) 8](#_Toc410998415)

[4.2 Prototype Features and Capabilities 8](#_Toc410998416)

[4.3 Prototype Development Challenges 8](#_Toc410998417)

[GLOSSARY](#_Toc410998418) 9

[REFERENCES…………………………………………………………………………………………….](#_Toc410998419)10

List of Figures

[Figure 1 - Major Functional Component Diagram](#_Toc411191810) 4

[Figure 2 - Major Hardware Component](#_Toc411191811) 5

List of Tables

[Table 1 - Real World Product vs Prototype](#_Toc411100652) 7

**1- INTRODUCTION**

Mastering a new foreign language takes time and depending on the language can be extremely challenging. The need for this is apparent. Open Doors (2013) reported that 819,644 students came to the United States to pursue university degrees. Data from a 2014 study found that almost 5,000,000 English as second language (ESL) students enrolled in public schools at the elementary, high school and university levels. Such a large number of students would benefit from a software tool aimed at teaching them English. ODU has a department called English Language Center ELC that help foreign students to learn English before starting their studies as academic students. When the student arrived to the U.S.A with student Visa, is allowed to stay for up to two years before beginning his or her academic program. (McKeon, 2014)

Research has found that reading is especially important to academic success and that most ESL students read texts one word at a time. Learning to read in phrase groups and lexical bundles can increase both reading speed and comprehension (Nishida, 2013). The web-based tool CESR is designed to break any text into phrase groups and lexical bundles, and can also highlight parts of speech in individual colors (for example, all nouns may be blue). The web-based tool CESR is designed to break text from any source into phrase groups, and can also highlight parts of speech in individual colors (for example, all nouns may be blue).

**2- CESR PRODUCTION DESCRIPTION**

Web-based tool called CESR is designed to help ESL students improve their reading speed and comprehension. There are two modules used in CESR: a SLASH module designed to break text into phrase groups; a READER module to display these phrase groups dynamically at different speeds and a COLRS module to colorize speech parts such as nouns, verbs, pronouns and prepositions.

The SLASH module will separate sentences into phrase groups which will help improve the reading speed skills of the student. It will also help identify lexical bundles, standard clusters of words used commonly in all writing and speech. By this method, CESR can help the students comprehend complex texts.

The COLRS module will be used to add color on the document’s part of speech to help the students understand the sentence structure and identify key reading strategies such as correct identification of pronoun references. Every part of speech will be colored using JavaScript.

**2.1 Key Product Features and Capabilities**

As a web-based application, CESR can be used in the classroom, in the home or even overseas. CESR will allow students and users in general to increase their reading skills by the two methods: The SLASH-reading algorithm that divides the text into phrase groups and lexical bundles and COLRS, which colorizes individual parts of speech. Furthermore, the READER modules allows dynamic display of the reading with each phrase group and lexical bundle displayed sequentially. Users can control the reading speed through a graphic interface that includes buttons to play, pause and change the speed.

CESR will use a password control to log-in to the web by three types of users: student, instructor and administrator. Students have the ability to view the document, with instructor-identified SLASH and COLRS features. Instructors can control, edit, add and delete documents and users. Moreover, they have the ability to create viewable documents for the students and to keep up with the students’ activity such assignments, including reading speed and the number of words read. Finally, the administrator has all the abilities of student and instructor. Additionally, the administrator can add and remove all type of users.

What makes CESR exceptional is not just the ability to display the document dynamically in phrase groups, but also the ability to give the user the ability to use COLRS to improve their recognition of English grammatical relationships. So far, no website or application uses both methods, breaking texts into phrase units and lexical bundles and also and colorizing part of speech.

**2.2 Major Hardware and Software Components**

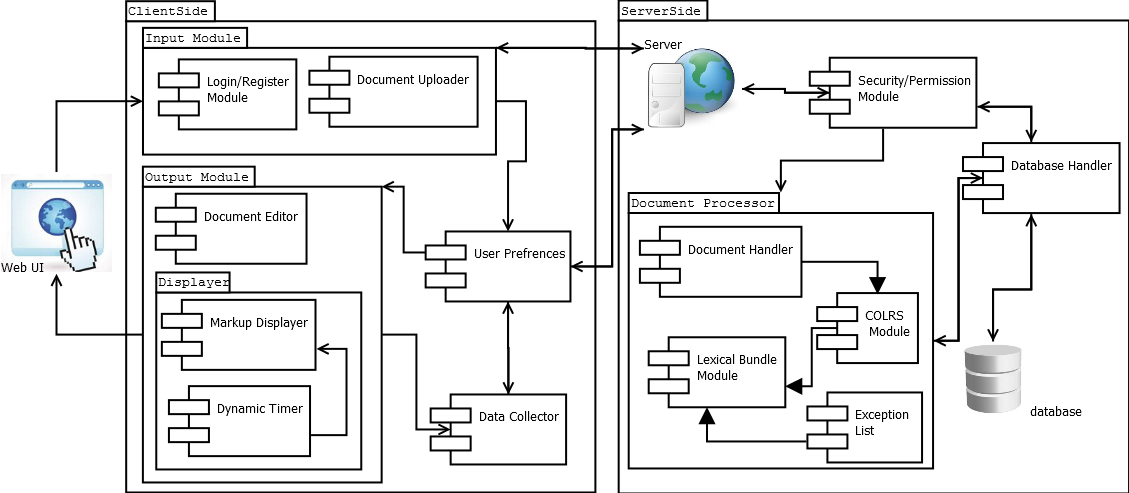
 Figure 1 - Major Functional Component Diagram

Figure 1 shows the major components of the hardware and software of CESR. Because CESR will be a web-based, all a student of instructor needs to use CESR is a computer, a web browser and the Internet. CESR does not require any hardware requirement except a server and a database.

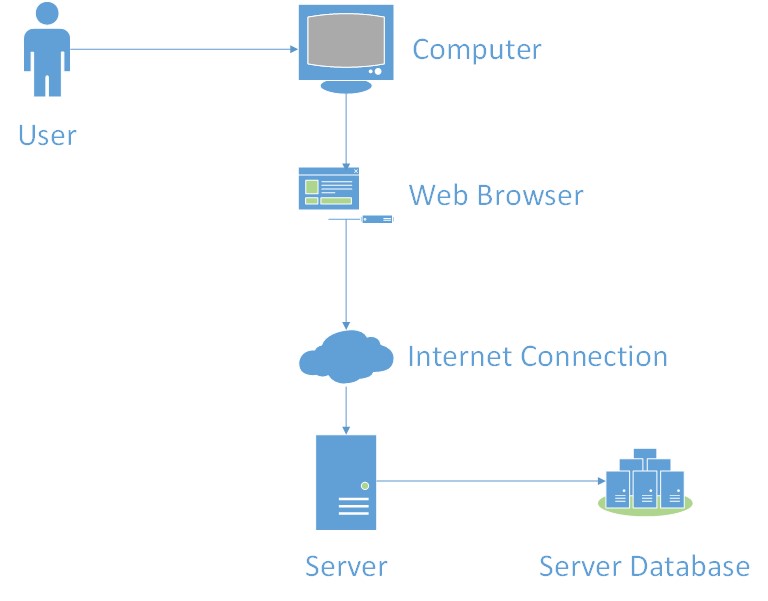


Figure 2 - Major Hardware Component

COLRS module will display each word in a text with different color depending on the part of speech. We used Using Natural Language Processing (NLP) as an open source resource to identify the parts of speech.

The SLASH module will pre-identified parts of speech to break the text into phrase groups and lexical bundles. It will take tokenized streams from the COLRS and will use the parts of speech to identify the phrase groups and the lexical bundles and insert a slash between each one.

The SLASH module receives data from the server then separates the paragraph by phrase groups, placing a slash between each. The COLRS module allows the instructor to chose to color one or more parts of speech. The READER module allows the student or instructor to control the speed of the dynamic display of the phrase groups and lexical bundles.

**3- IDENTIFICATION OF CASE STUDY**

Old Dominion University has more than 25,000 of international students from all over the world pursuing degrees in many programs ranging from finance, engineering, nursing, and computer science. However, for these students to start their academic studies, many l need to prove that their English is good enough for them to succeed academically. In order for them to show that, they need to take tests such as the TOEFL (Test of English for Foreign Learners) the students don't get a high enough scores on the TOEFL then they apply for the English Language Center at ODU.

Greg Raver-Lampman, an instructor at the Old Dominion University English Language Center, teaches English to students who speak English as a second or third language. These students rely on him and other ELC professors to get as much knowledge of the English language before they start their academic classes. These professors do their best to help the students grasp the language for use in future classes. They help them to understand grammar, write paragraphs that later become papers, read articles and books, listen to videos and learn how to listen to discussions, and other techniques to assist in the learning of the language. Often, however, students do not learn as fast as they want to, and tend to be very slow readers. Because of this Professor Raver-Lampman requested help from the Computer Science Senior Design class to design a software solution to help students identify all the different parts and to identify phrase groups and lexical bundles in a target text.

This web application was primarily created to help students at the ELC learn how to increase reading comprehension and speed. This application is superior to existing websites that help other people learn to improve English reading speed and comprehension. This website can be used by not only students, but also professionals that travel to English-speaking countries frequently and want to learn how to read and listen correctly. This is a solution that can be used by anyone that simply wants to learn the English Language and read it efficiently.

**4-** **CESR PRODUCT PROTOTYPE DESCRIPTION**

There are many applications that use some features that CESR has, so CESR has to be unique and close to the real world product. Table 1 will show the disparity between CESR prototype and real world product

|  |  |  |
| --- | --- | --- |
| **Features** | **Real World Product** | **Prototype** |
| Parsing Capabilities | Ability to Parse different kinds of documents | Ability to parse text copy and pasted in text block |
| Text Modification | Ability to modify and store previously parsed documents | Ability to modify and store previously parsed documents |
| Color Capabilities | Ability to color chosen parts of speech using a JSON format and JavaScript functions. | Ability to color chosen parts of speech using a JSON format and JavaScript functions. |
| Slashing Capabilities | Ability to identify Lexical Bundles through the inserting of slashes. | Ability to identify Lexical Bundles through the inserting of slashes. |
| Displaying Lexical Bundles in a single bundle form | Ability to speed up, slow down and pause Lexical Bundles being displayed. | Ability to speed up, slow down and pause Lexical Bundles being displayed. |
| Exception list | Lists of commonly used expressions that would otherwise be incorrectly handled by the SLASH Algorithm. | Lists of commonly used expressions that would otherwise be incorrectly handled by the SLASH Algorithm. |
| Login interface | User Authentication in a stand-alone environment | User Authentication in a stand-alone environment |
| Student Data Reporting | Tracks individual and collective student progress. To include words per minute, total time and total Lexical Bundles. Data to be stored in database. Displayed in graphs and statistics. | Limited basic student metrics will be available such as Lexical Bundles per Minute. |
| Homework Mode | Instructors have the ability to remove coloring of words and have students correctly identify the part of speech. | Not Included. |
| Administrative Privileges | Administrators are able to edit, add, or remove users and saved documents in the system. | Administrators are able to edit, add, or remove users and saved documents in the system. |
| SLASH Document Viewing Mode | Ability to view documents with slashes inserted and SLASH Reader. | Ability to view documents with slashes inserted and SLASH Reader. |

Table - Real World Product disparity Prototype

**4.1 Hardware and Software Prototype Architecture**

Hardware Prototype: the only hardware needed to operate the application is a simple server owned by the user developers. Old Dominion University will provide the developers a Virtual Machine (VM). Software Prototype: use of open source to allow the development into the full product. The user interface will be built in JavaScript. Also, the software will be written on MySQL database and Node.js to be used as the web and application server.

**4.2 Prototype Features and Capabilities**

The CESR prototype has many major features such as, coloring POS in the paragraph that will appear on the browser and the ability to place slashes between phrase groups and lexical bundles. Also, a client can use speed reader to increase reading skills. The client will have the ability to display the document dynamically using the play and pause features to control the speed. Furthermore CESR can be used anywhere all what client need is a computer and internet connection to make it work.

**4.3 Prototype Development Challenges**

CESR will face several challenges. One of them is identifying parts of speech. When the application find an error with identify the part of speech SLASH and COLRS will fail to give the right result. CESR, however, gives the instructor the ability to correct any errors before providing the text to the students. The READER module will also be a challenge for the development because when using SLASH to text into phrase groups and to control the speed at the same time. The developer will have to deal with this challenge before deployment. Because CESR is made for ESL student, the interface must be intuitive and easy to use.

**GLOSSARY**

CESR: Color-Enhanced Slash Reader

Client Side: The user-interface of CESR.

COLRS: Colored Organized Lexical Recognition Software.

ELC: English Learning Center at Old Dominion University.

ESL: English as second language.

JSON: JavaScript Object Notation. A nested data structure commonly used to pass data between a server and a client.

Phrase groups: A discrete cluster of words that conveys meaning within a longer sentence.

Lexical bundle: A group of words that occur repeatedly together in written and spoken English.

NLTK: A suite of libraries and programs for symbolic and statistical natural language

processing (NLP).

Node.js: Open source, cross-platform run-time environment for server-side and networking

applications.

POS: A Part of Speech such as a noun, adjective or verb.

TOEFL: (Test Of English for Foreign Learners) English language proficiency test required by universities for enrollment for internationally based students.

VM: Virtual Machine.

**REFERENCES**

McKeon, D. (n.d.). Research Talking Points on English Language Learners. Retrieved December 11, 2014.

Nishida, H. (2013). Comparison of reading, and grammar as instruction techniques for reading comprehension.

Open Doors 2013 Report. Institute of International Education, Inc11, 11 Nov. 2013. Web. 08 Oct. 2014.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011, January 15). Processing Advantages of Lexical Bundles: Evidence from Self-Paced Reading and Sentence Recall Tasks.

Retrieved December 10, 2014.

ODU International Admission. Enrollment of international students. Retrieved year of 2015

Mikowski, M., & Powell, J. Single Page Applications. Manning Publications 2014.