CS 411 - Lab 1 CLASH Product Description (Blue/Purple Team)

Francia Molina

Old Dominion University

Professor Jannet Brunelle

Professor Hill Price

Table of Contents

1. Introduction 3

(1.1) Key Product Features and Capabilities 4

(1.2) Major Hardware and Software Components 6

(2) Identification of Case Study 7

(3) CLASH Product Prototype Description 8

(3.1) Hardware and Software Prototype Architecture 9

(3.2) Prototype Features and Capabilities 9

(3.3) Prototype Development Challenges 10

Glossary 11

References 12

(1) Introduction

Color Lexical Analysis and Slash Handler (CLASH) is a web interface that includes two modules, ‘COLRS’ and ‘SLASH’. The COLRS’ module colorizes each part of speech (POS) in a text document with a designated color to increase understanding of sentence structure and grammar. The SLASH module parses, breaks down, text into thought groups to help increase reading speed and comprehension. Thought groups are also used inversely with lexical bundles (LB), which are a group of words that occur repeatedly together, presenting one thought. CLASH’s primarily focus is to be used by English as second Language (ESL) students or English Language Learners (ELL).

An ELL is a person who is learning the English language in addition to his or her native language. The percentage of public school students in the United States who were English language learners during school year 2011–12 were 9.1 percent, or an estimated 4.4 million students, compared to school year 2002–03 8.7 percent, or an estimated 4.1 million students (The Condition of Education). In 2011–12, ELL students made up an average of 14.2 percent of total public school enrollment, ranging from 10.9 percent in small cities to 16.7 percent in large cities (The Condition of Education). Old Dominion University (ODU) has an entire department, the English Language Center (ELC), focused on aiding ESL students with their personal, career, and or academic goals involving the English language. ELC helps its ESL students learn the language, but while they work diligently to help, the processes used for teaching reading and grammar is outdated.

The current teaching style followed by professors’ is to write a sentence on the board, then mark each part of speech in regards to grammar. This proves to be time consuming and overall limits the amount of examples that can be provided for the students need. Psychologically, it has been proven that color impacts learning, they relieve eye fatigue, increases information retention, increase productivity and accuracy, and support developmental processes (Engelbrecht, K). CLASH using this principal will implement the COLRS module to use color correlation with POS to increase grammar recognition and retention. Reading assignments are provided to students to increase reading comprehension and reading speed. It is believed that, the more exposure a student has to language through reading, the greater the possibilities that overall language proficiency will increase (Anderson, N.). The issue correlated with this is that often times, in efforts to assist students to increase their reading, teachers overemphasize accuracy. When this occurs, reading fluency is impeded (Anderson, N.). In this day and age, sites like Spreeder have originated in attempts to aid with reading comprehension, but in truth don’t help comprehension due to their focus on speed. It has been proven in sentence recall experiments that LBs were remembered and read more efficiently than word by word reading. The experiments suggest that regular multi-word sequences, LBs, leave memory traces in the brain allowing for better comprehension (Tremblay, Derwing, Libben, Westbury). CLASH is the application that will be aiding teachers’ methods of teaching reading and comprehension by providing a two-in-one product that focuses on colorizing grammar and reading by lexical bundles.

(1.1) Key Product Features and Capabilities

CLASH is a web-based application possessing a Graphic User Interface (GUI), database, slash and coloring algorithms. The GUI will be simple to maneuver around yet have an aesthetically pleasing interface for user (i.e. Student, Teacher, and Administrator). The database (currently 10 gigabyte (GB)) will hold the, back bone of CLASH, text files, parsing and coloring algorithms, program code, etc. The web site will be the face of CLASH that will allow the user to interact with the program. The SLASH module will take a text file and break up the document into lexical bundles with a Slash algorithm inspired by the Korean technique of teaching English. SLASH also has a feature called Slash Player that will be a text stream showing the lexical bundles, of three to five words, one at a time with speed control. The COLRS module will take a text file, previously parsed by SLASH, locate and color POS. It will focus on the 8 traditional parts of speech, excluding interjection, which students learn.

CLASH uses a standard web browser and internet connection. Having logins for all the users and password generator, all users will be required to make an account. Through the user interface, users can copy and paste text, ‘slash’, ‘colorize’, and control the speed at which lexical bundles are displayed in the Slash Player. The CORLS module will display a separate menu with the 8 variances of POS being colorized, as shown in Figure 1.1. The product features three different types of user roles, these roles include Administrator, Instructor, and Student. The Administrator has all the abilities of the Instructor plus the ability to add and or remove Instructors as users. The Instructor will be able to add and or remove Students as users, edit slashed documents, upload documents, and have Student view. The Student will be able to select type of view (SLASH or COLRS), POS colorized, and control reading speed in Slash Player. The application allows for the Instructor to view activity data for the Student users. This activity data will include the student’s current reading speed, and the amount of time spent on the site. CLASH will provide three separate modules that will allow all users to experience a well-rounded environment in which to improve reading comprehension and grammar.



Figure 1.1

(1.2) Major Hardware and Software Components

CLASH will be a web-based application hosted on a server and will accessed through any web browser. Hardware wise there is no special requirements. The user will need a computer with internet or any internet enabled device to access the application. There are three major software components that make up CLASH; the COLRS module, SLASH Module, and Client-side Slash Player. The COLRS Module uses open source Natural Language Processing (NLP) to tokenize and parse the inputted document, it will generalize the POS tags that will later result in the ability to colorize the tags created. The SLASH Module will take the tokenized stream from the COLRS Module and using the slash algorithm break the sentences into 3 to 5 word lexical bundles. This is done accurately by receiving the data from the server and parsing the text according to the POS tag. There will be a sub-component that will create an exception list that will, with time, improve the accuracy of the final output. Figure 1.2 will show the overall hardware and software components being used.

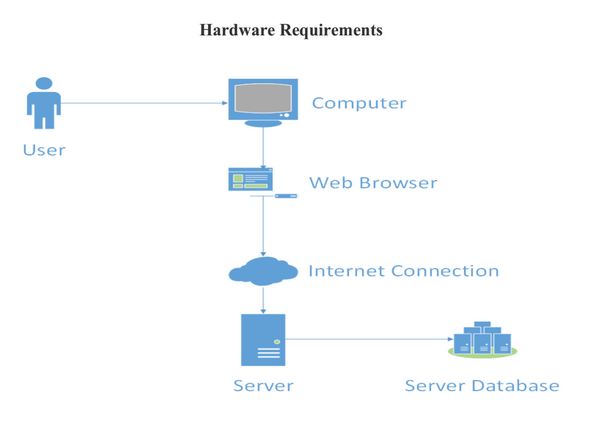


Figure 1.2

(2) Identification of Case Study

ODU is famous for its international student presence. For the 2012-2013 academic school year there were a total of 1,086 international students (Monarch Diversity). A requirement that must be shown while applying for admission is proficiency in English. That is, if the native language is not English, evidence of English proficiency must be demonstrated to enroll in credited courses. If applicants do not have the required test scores, they are allowed admittance through Conditional Admission. The Bridge Program is designed for conditionally admitted students with TOEFL iBT scores from 61 to 79 (PBT 500-550) or IELTS scores from 5.5 to 6.5 (English Proficiency). ELC provides the Bridge Program which runs over two semesters. Once completed, the Undergraduate or Graduate Bridge Program satisfies the university's English proficiency requirement. ELC professors teach the students English grammar, reading skills, and conversing skills. ESL students attempt to grasp as much knowledge of the English language before they start their credited classes. The range of teaching techniques ranges from writing paragraphs that later become papers, reading articles and books, listening to videos and learning how to listen to discussions, and more. Often times, however, students’ experience that they do not learn as quickly as they want to and as a result are slow readers. Greg Raver-Lampman is an instructor at ODUs English Language Center that is looking for a product that will help him and other English teachers’ aid ESL students in their studies. CLASH is that product, combining what Professor Raver-Lampman considers to be the two main issues he sees with ESL students learning English. CLASH is the software solution that will help students identify all the different POS and reading in lexical bundles.

(3) CLASH Product Prototype Description

The CLASH prototype will be a modified version of a Single Page Application (SPA) architecture. A SPA is a web application that fits on a single web page with the goal of providing a similar experience to that of a desktop application that does not reload as the web page changes states. The SPA of CLASH will be built completely in JavaScript (JS), meaning each component and operation will use JavaScript. CLASH will have a traditional html webpage that is called forth by JS, a JavaScript friendly web server, Node,js, and a traditional relational database. A SPA is an ideal platform for CLASH because it offers the look and feel of a desktop application, but the accessibility of a web application. The platform was chosen with users who are not computer savvy in mind. CLASH will not have no software to install by the user. A user management module will be implemented to allow instructors and administrators to add and remove users as needed. Users will simply login to their accounts after being created by Administrator or Instructor. All functions will accessible through simple menus on website. Instructors and Administrators will be able to add student users on a one-by-one basis and they will also be able to add users in bulk with an ODU enrollment file, Administrators will be able to add or remove instructors as well.

(3.1) Hardware and Software Prototype Architecture

 The hardware required for the prototype will essentially be a machine to write the code and server to hold the database. The software required for the prototype will be a collection of open source or liberally licensed software to permit distribution or potential commercialization of the software package. The web and application server will be Node.js. This will allow the HTTP server to establish an incoming network connections from users (the students and instructors), as well as the application handling text processing of documents. At this time this is all that can be provided about the hardware and software architecture.

(3.2) Prototype Features and Capabilities

The CLASH prototype will have three modules as follows: COLRS, SLASH, and Slash Player. The COLRS module will demonstrate the ability to identify parts of speech, distinguishing amongst the various POS with color. The SLASH module will identify lexical bundles throughout a paragraph and display them by the insertion of slashes. The Slash Player will display lexical bundles one at a time, and allow the user the ability to increasing and decreasing the lexical bundle stream. CLASH will be able to take text in two forms, by file upload or by copy and paste. The Administrator will be able to add and remove Instructors and Student roles. Instructor will be able to add and remove Student roles as well as manually edit slashed and colored documents. CLASH will be able to create and update an exception list in the SLASH module. The exception list, is a list that will hold lexical bundles that are common and reoccur multiple times.

(3.3) Prototype Development Challenges

There are multiple challenges that CLASH may face on its developmental journey. One of them is computational time of the exception list. By adding more exceptions to the list, computational time increases due to the increase of size of the list itself. Another challenge is missing attributes; meaning features are not working properly, not working at all, or missing completely. CLASH, being unique and very useful, will have minor hiccups that have to be taken care of when they show up. English is a challenging language and with that comes errors. For example, a possible error that will arise is the incorrect identification of parts of speech and incorrect placement of slashes.

CLASH aims to provide the user the ability to control the speed at which the lexical bundle will show with the “Slash” feature. There is a possibility of having difficulty getting the algorithm that will control the speed to do the proper task. CLASH is meant to be simple for use, meaning having a simple interface. This is to make the product accessible to users with various computer skills levels. There is that possibility that what we believe to be simple is not simple which will in the long run affect the appeal it has for certain users. Furthermore, another challenge that will arise with time is storage capacity. Users will populate our database with their documents and 10GB will not be enough for long term use. Building a working prototype with these features will be tested in a real world situation against a control group (Lab 1).

Glossary

CLASH - Color Lexical Analysis algorithm and Slash Handler

COLRS – Colored Organized Lexical Recognition Software

ELL - English Language Learners

ELC – English Learning Center

ESL – English as second language

GB - Gigabyte

GUI - Graphic User Interface

HTML - HyperText Markup Language

IBT – International Benchmark Test

JS – JavaScript

LB – Lexical Bundle

Lexical Bundle – a group of words that occur repeatedly together within the same register

NLP - Natural Language Processing

Node.js – an open source, cross-platform run-time environment for server-side and networking application.

ODU – Old Dominion University

POS – Parts of Speech

SPA – single page application, is a highly responsive web application that fits on a single page and does not reload as the web page changes states.

TOEFL – Test of English as a Foreign Language

Ubuntu – a Debian-based Linux operating system

References

Anderson, N. (1999, April 1). Improving Reading Speed - Activities for the Classroom. Retrieved February 1, 2015, from http://dosfan.lib.uic.edu/usia/E-USIA/forum/vols/vol37/no2/p2.htm

Engelbrecht, K. (2003). The Impact of Color on Learning (p. 5). Chicago, Illinois: Perkins & Will.

English Proficiency. (2015, February 2). Retrieved February 6, 2015, from https://www.odu.edu/content/odu/admission/proficiency.html

Monarch Diversity. (n.d.). Retrieved February 6, 2015, from https://www.odu.edu/admission/international/global

The Condition of Education 2014. (2014, January 1). Retrieved February 6, 2015, from http://nces.ed.gov/fastfacts/display.asp?id=96

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.