Profile-Based Text Classification for Children with Dyslexia

Chris Litsas, School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

Maria Mastropavlou, Department of Philology, University of Ioannina, Ioannina, Greece

Antonios Symvonis, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Athens, Greece

ABSTRACT

Although extensive research has been conducted in the field of text-readability and user modelling, scholars and researchers have taken into consideration only linguistic complexity in order to classify a text as readable or not. In this paper, the authors move one step forward by considering one more factor, namely intended reader’s skills, and by trying to study text readability from a user-specific perspective. Central to our approach is the notion of the user’s profile which carries information regarding the linguistic difficulties a user with dyslexia may experience. Based on the user’s profile, they develop heuristics for evaluating text’s readability for the specific user. The developed heuristics are incorporated in the text classification services of the iLearnRW project, aiming to facilitate the selection of appropriate/suitable reading resources, written in English or Greek, for children with dyslexia.

Keywords:  Dyslexia, Text Classification, Text Readability, User Modelling

1. INTRODUCTION

A child that learns to read and/or write will practice with several pieces of text. However, not all text is appropriate to be used in the learning process (either for reading or for writing) of a particular child. The level of difficulty (or “degree of appropriateness”) of the text must be carefully considered. For a child without learning difficulties, the degree of appropriateness of a text depends on formal, linguistic factors, the text content as well as the child’s age. For children with learning difficulties, these factors need to be considered with respect to each child’s special educational needs. In that sense, text classification with respect to particular language difficulties encountered in dyslexia would be a very useful process, enabling teachers to select the appropriate content for a particular learner and formulate a more individualized educational plan. The Text Classification approach described in this paper has been developed in the context of iLearnRW (Integrated Intelligent Learning Environment for Reading and Writing), an ongoing European Union funded project.

DOI: 10.4018/ijmstr.2015010102

Copyright © 2015, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
In this paper, we describe the Text Classification Module (TCM) that is incorporated in the iLearnRW software. The design of the TCM aims to provide individualized teaching assistance to children with dyslexia by enabling a teacher or parent to classify texts with respect to the degree of appropriateness for a particular child, based on his/her profile, as well as to search for appropriate content for a particular child. To this end, the notion of text readability was analysed based on linguistic complexity issues which, to a great extent, determine the suitability of a text for a particular learner.

To the best of our knowledge, it is the first time that personalized (based on user’s profile/characteristics) text classification is attempted. Even though in our work we focus on a specific class of users, children with dyslexia, our approach is applicable to the general user.

This paper is organized as follows: We start by describing the basic notions related to linguistic complexity and reading difficulty. We also present a brief review of commonly used readability formulas. Then, we describe the developed user models that capture the reading learning process for the English and Greek languages as well as our user-specific text classification methods (together with an application employing the developed text classification method). Finally, we conclude with the presentation of preliminary evaluation results concerning the Greek Language.

2. LINGUISTIC COMPLEXITY AND READING DIFFICULTY

Text readability is closely related to and even determined by the linguistic complexity of a text in the sense that the readability of a text increases as linguistic complexity decreases and vice versa. Therefore, linguistic complexity is a central notion when dealing with text classification.

Defining linguistic complexity is currently one of the most hotly debated notions in linguistics. In a first quantitative description of linguistic complexity, Blache (2011) identifies the types of constructions that are considered complex and thus difficult to process. She differentiates between local complexity, which refers to structural complexity, difficulty, which involves processing aspects and cognitive load, and global complexity, which refers to the language as a system rather than the complexity of a given realization (see Miestamo (2008) for a similar classification). Of the two levels, local complexity is considered measurable and has drawn considerable attention in the literature. Local complexity therefore includes phonological complexity (e.g. size of phonemic inventory, incidence of marked phonemes, phonotactic restrictions, maximum complexity of consonant clusters), morphological complexity (e.g. extent of allomorphy use and morphophonemic processes), syntactic complexity (e.g. level of clausal embedding and recursion), semantic and lexical complexity (e.g. extensive occurrence of homonymy and polysemy, type/token ratios), pragmatic complexity (e.g. degree of pragmatic inferencing) (see Szmrecsanyi and B. Kortmann (2012) for a review).

3. LINGUISTIC COMPLEXITY AND TEXT COMPLEXITY

The complexity or the degree of challenge of a particular text is the result of combinations and interactions of a variety of factors. These may include linguistic complexity factors, topic familiarity, word difficulty, sentence length, concreteness of ideas and concepts and others. In a description of text complexity, Lipson and Wixson (2003) define a number of factors that affect the readability of a text, which include the number of syllables in the words and the number of words in the sentences, while other linguistic characteristics, such as vocabulary and sentence structure, text organization and the amount of background knowledge that is required of readers are also often taken into account when determining the appropriateness of a text for a particu-
Braided Routing Technique to Balance Traffic Load in Wireless Sensor Networks
www.igi-global.com/article/braided-routing-technique-to-balance-traffic-load-in-wireless-sensor-networks/180663?camid=4v1a