Available online at www.sciencedirect.com
ScienceDirect
Computers & Graphics 31 (2007) 761–777
COMPUTERS & GRAPHICS
www.elsevier.com/locate/cag

Technical Section

A multimedia framework for effective language training [☆]

Markus Gross[a,*,1], Christian Voegeli[b]

[a] ETH Zurich, Switzerland
[b] Dybuster Inc., Switzerland

Received 14 June 2007; received in revised form 21 August 2007; accepted 1 September 2007

Abstract

We present a novel framework for the multimodal display of words using topological, appearance, and auditory representations. The methods are designed for effective language training and serve as a learning aid for individuals with dyslexia. Our topological code decomposes the word into its syllables and displays it graphically as a tree structure. The appearance code assigns color attributes and shape primitives to each letter and takes into account conditional symbol probabilities, code ambiguities, and phonologically confusable letter combinations. An additional auditory code assigns midi events to each symbol and thus generates a melody for each input string. The entire framework is based on information theory and utilizes a Markovian language model derived from linguistic analysis of language corpora for English, French, and German. For effective word repetition a selection controller adapts to the user's state and optimizes the learning process by minimizing error entropy. The performance of the method was evaluated in a large scale experimental study involving 80 dyslexic and non-dyslexic children. The results show significant improvements in writing skills in both groups after small amounts of daily training. Our approach combines findings from 3D computer graphics, visualization, linguistics, perception, psychology, and information theory.

Keywords: Visualization; Information theory; Visual information; Coding; Entropy; Learning; Natural language; Dyslexia

1. Introduction

Dyslexia is traditionally defined as the inability of otherwise intelligent individuals to acquire fluent reading and/or orthographically correct writing skills [1]. The socio-economical implications of dyslexia are significant and often devastating for the individual, who, in many cases, dramatically underperforms in school and profession. Dyslexia occurs predominantly in Western world languages, including English, French, German, or Spanish [1]. It is estimated that about 5–7% of the Western world population suffers from minor or major forms of dyslexia [2].

Dyslexia appears in various types, such as *deep* or *surface*, and *developmental* or *acquired* dyslexia and at different levels of severity and strength. There are multiple causes for dyslexia, which, as of today, are not fully researched yet. Most often, dyslexia develops in early childhood and adolescence [3]. The irregularities in cerebral information processing underlying dyslexia are not fully understood yet and still subject of intensive research in psychology, medicine, neuroscience, linguistics, and other disciplines. A full overview of the exhaustive scientific literature on this subject is beyond the scope of this paper. We will confine ourselves to a summary of the most important findings relevant to our own work.

Researchers have proposed various models for the acquisition of human reading and writing skills. It is widely believed that orthographically correct writing is acquired over three phases: a *visual* phase, a *phonetic* phase, and a final *semantic* phase [2]. A traditional theory for

[☆] Supplemental materials submitted contain two videos demonstrating the system in action and a video report from a German TV station about the system and the user study.

*Corresponding author. Tel.: +41 44 632 7114; fax: +41 44 632 1596.
E-mail addresses: grossm@inf.ethz.ch (M. Gross), cvoegeli@dybuster.com (C. Voegeli).

[1] The author is a professor of Computer Science at ETH Zurich. This work was performed while the co-author worked as a junior researcher at ETH.

0097-8493/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cag.2007.09.001

reading is the *dual route model* distinguishing between a *phonological* and a *lexical* route [4].

More recent theories attribute dyslexia to a neurological disorder with a genetic origin and one school of thought explains dyslexia as a consequence of deficits in the phonological processing of the brain [5]. Scientists [6] also observed correlations between the occurrence of dyslexia and low level transient information processing in the human brain. Another school of thought [7] suggests that dyslexia is caused by specific weaknesses in visual and attentional processes. In particular, there is evidence for deficits in the transient, visual–temporal information processing—as opposed to visual–spatial perception, which is usually well developed in people with dyslexia. Such transient visual activity can be affected by the use of color [8]. A further theory [9] suggests that normal development and disorders of speech perception are both linked to temporospectral auditory processing speed. This key observation has been confirmed by various other authors [10].

Lately neurobiological evidence for reading and writing disorders has been given [11–13]. These researchers found abnormalities in the structure of the temporal cortex of dyslexic children using diffusion tensor imaging. Overall, these recent neurobiological studies seem to confirm the hypothesis that the difficulty in precise auditory timing has a link to language acquisition and comprehension [14]. It has also been hypothesized [15] that musical training may be able to remedy such timing difficulties. Our method compiles all these experimental findings into a novel, multimodal word recoding scheme.

1.1. Therapy and training programs

Numerous therapies of dyslexia have been proposed and applied so far. For instance, a French team showed [16] that a focused and abstract audio–visual training can lead to plastic neural changes in the cortex and thus improve cerebral language processing. Another very successful and scientifically well-founded therapy [9] utilizes the results from above. They developed a series of neuroplasiticity-based training programs that are designed to improve fundamental aspects of oral and written language comprehension and fluency. A further example for dyslexia treatment is LEXY [17], a Dutch therapy. This concept focuses on lexico-phonological deficits and employs the syllabic structure of words as its central element. Besides such scientifically well-founded approaches, there is a wealth of more or less heuristic therapies and learning aids. A comprehensive survey is given in [18,19]. Various commercial multimedia e-learning systems [20] offer computer-based exercises to link words to their semantics and to pictorial information. Strydom and du Plessis [21], for example, present a compendium of cognitive exercises aimed at the development of reading, writing, and other skills partly using color to support learning. Davis and Brown [22] depict words as 3D associations or as scenes sculpted by the user with play dough. While this method associates each word with a spatial representation, it is very cumbersome and of limited success. Overall and while significant progress has been made [9], there is no single commonly agreed-upon therapy for dyslexia to the present day.

1.2. Our approach

The approach presented in this paper is fundamentally different from earlier ones in that it combines concepts from visualization and perception (in parts also used by e.g. [21,22]) with enhanced concepts from 3D graphics, statistical modeling of language, and information theory to design an advanced learning and language acquisition aid for individuals with dyslexia. The heart of our method is an abstract, graphical recoding of the input word. The code consists of a spatio-topological code, an appearance code (color, shape), and an auditory code, as displayed in Fig. 1. This coding transforms the input into a multimodal representation that supports phoneme–grapheme associations. It thus bypasses the distorted cognitive cues of people with dyslexia and builds alternative cerebral

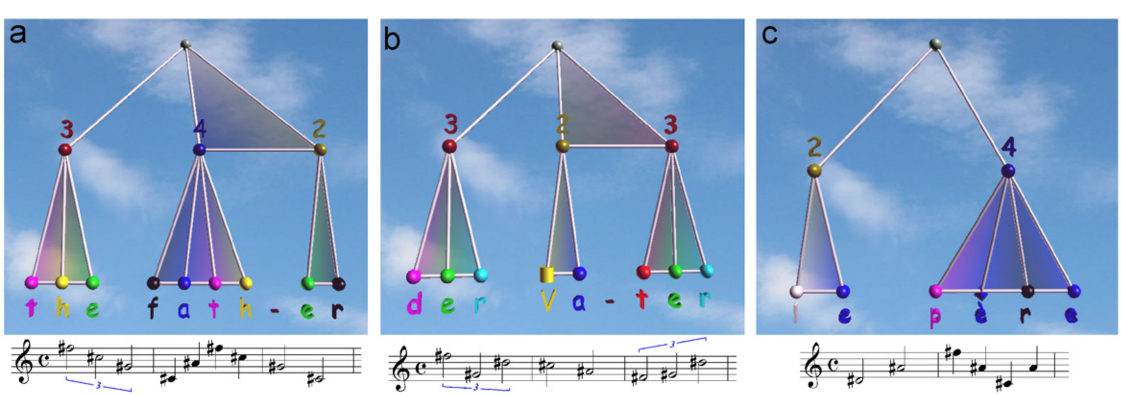

Fig. 1. The string "the father" is displayed using a topological, color, shape, and auditory representation. The recoding retains information in that its overall entropy rate matches the entropy rate of the input symbols. The color code is language specific and the result of an optimization. We depict the codes for English, French, and German.