Ana Sucena, João Falcão Carneiro*, Ana Paula Vale and Fernanda Leopoldina Viana

# The relationship between phonological awareness time and reading abilities: an assessment using self-organizing maps

**Abstract**

**Background:** Classically, the assessment of reading disabilities is based on the accuracy for word and nonword reading, as well as on the accuracy or sensibility measures (such as d′) for phonological awareness tasks. Recent studies indicate that in terms of phonological awareness results, the response time is a more accurate indicator than sensibility measures (such as d′), thus providing an important measure explaining some of the differences between good and poor readers. This article explores the discriminative capability of phonological awareness task time (PATT) in reading disability assessment.
**Methods:** One hundred and eighty-six children were tested using conventional tasks, specifically word reading, nonword reading, and phonological awareness tasks. The word and nonword accuracy and PATT were used to train self-organizing maps (SOM) to classify children into three distinct groups.
**Results:** Phonological awareness response time provides a powerful discriminative measure.
**Conclusions:** Our results indicate that the PATT constitutes a useful selective measure, particularly in the third and fourth grades when classical variables such as word and nonword reading accuracy lose their discriminative capabilities. Also, the use of SOM to classify children's reading abilities can successfully categorize children and capture meaningful measures such as the lexicality effect.

**Keywords:** dyslexia; reading disabilities; self-organizing maps.

***Corresponding author: João Falcão Carneiro**, IDMEC, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, S/N Porto 4200-465, Portugal, E-mail: jpbrfc@fe.up.pt
**Ana Sucena:** Instituto Politécnico do Porto, Porto, Portugal
**Ana Paula Vale:** Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
**Fernanda Leopoldina Viana:** Universidade do Minho, Minho, Portugal

## Introduction

Dyslexia is a learning disability characterized by difficulties in the acquisition of reading and spelling, which may be expressed by failure in accuracy, fluency, or comprehension, despite adequate intelligence [1, 2]. The prevalence of dyslexia across English-speaking children varies between 4% and 8% [3]. The Connecticut Longitudinal Study [4] found a prevalence of 7.8% in second-, 7% in third-, and 5.4% in fourth-grade students. Recent French data [5] indicate that 9% of second graders reveal severe reading difficulties, whereas in Germany [6], dyslexia varies between 6.4% and 8% in second, third, and fourth graders. A recent study run with Portuguese children indicates that dyslexia affects 5.4% of children [7]. The same study indicates that a higher percentage of Portuguese children reveal serious reading difficulties (8.6%), although they did not meet the strict criteria to be diagnosed as dyslexic.

Stanovich [8] applied the Matthew effect to dyslexia, sustaining that just as the rich get richer and the poor get poorer, early good readers eventually turn into fluent readers, whereas early poor readers will tend to lag more and more behind their peers as they progress in school. Children with early reading difficulties will soon cope with difficulties by avoiding reading-related tasks, thus perpetuating the cycle [9]. Early identification of reading difficulties is therefore crucial [10, 11], for which accurate and easy-to-administer screening tests that signal children "at risk of being dyslexic" are needed.

The present study presents an innovative approach by exploring the use of phonological awareness task time (PATT) as a measure capable of successfully discriminating between different reading levels. One hundred and eighty-six children were tested with two classical measures (word and nonword reading) and one innovative measure (time response in a phonological awareness task). These tasks were selected based on research on developmental dyslexia that has shown that dyslexic children perform worse than controls in reading isolated words and nonwords, with a special disadvantage for nonwords, known as lexicality effect [12–15]. Also, dyslexic children reveal a poor phonological awareness [7, 15–17]. These results are coherent with the phonological deficit hypothesis, which is the most consensual hypothesis on the etiology of dyslexia [12, 14, 15]. According to the phonological deficit hypothesis, dyslexia assessment must therefore include isolated word and nonword reading and phonological awareness tasks. Children will be diagnosed as dyslexic if they perform significantly below what should be expected on the basis of their chronological age, IQ, and school grade.

Another contribution of this work relies on the use of self-organizing maps (SOM) [18] to classify the results into three classes: poor, average, and good readers. SOM are feed-forward neural networks that learn to classify its input vectors depending on how they are grouped in the input space. No target data are required, as SOM use a nonsupervised training algorithm. This feature makes their use quite interesting for the application under study because no human intervention is required to classify the results obtained in the different tasks. The use of artificial intelligence techniques in data-mining problems is widespread in very different applications, e.g. speech and image pattern recognition [19], military uses [20], or stock exchange prediction [21]. However, its use for reading or learning abilities assessment has deserved little attention from the scientific community. Some exceptions to this scenario can be found in the studies made by Novák et al. [22], Palacios et al. [23], and Loizou and Laouris [24]. These three studies differ from the one presented here because they use other sources of information as a basis for dyslexia assessment: in Novák et al. [22], eye movements measured using video-oculography technique were used; in Palacios et al. [23], the goal was to diagnose dyslexia in early childhood, so non-writing-based graphical tests were used; finally, in Loizou and Laouris [24], the goal was to identify learning difficulties using a set of tests like the Mental Attributes Profiling System, the Wechsler Intelligence Scale for Children, or rapid naming tests.

The approach followed in this work is different from the above because the input data to the SOM are the results children obtain in a test including the previously described tasks of word and nonword reading and phonological implicit awareness task time. This article is organized as follows. The next sections present an overview of the methods and tests used to assess the reading abilities, a brief introduction on SOM, and the details of the topology and procedures used to train the SOM developed in this work. The results obtained are then shown and analyzed. Finally, the major conclusions drawn from this study are highlighted.

# Methods

## Participants

One hundred and eighty-six Portuguese-speaking children, aged 7.5 to 9.7 years, were tested. Participants were tested in the last trimester of the school year of grades 2, 3, and 4. Two groups of children were assessed: dyslexic (n=72) and non-dyslexic children (n=114). Dyslexic children were preselected from regular schools according to the following criteria: scoring at or below the 5th percentile on a reading level test [25], having no known additional learning or spoken language problems, having an average or above average nonverbal IQ, as measured by the Ravens Coloured Progressive Matrices, being of average socioeconomic background. Non-dyslexic children were selected according to the same criteria, except for the reading percentile, which was fixed to a minimum of 50.

All children were learning to read within a mixed teaching method, which is the most adopted method in Portugal. Informed consent was obtained from parents and school authorities before the start of data collection.

## Materials

### Reading test

Participants were asked to read isolated words (n=132) and derived nonwords (n=108). Words were selected according to their orthographic complexity: one-to-one mapping, rule-based and irregular words. One-to-one mapping words are characterized by a bidirectional grapheme-to-phoneme conversion (GPC); rule-based words are characterized by a one-to-many GPC, for which accurate reading one needs to apply the orthographic rules; irregular words are characterized by a one-to-many GPC, with no underlying orthographic rule. Nonwords were derived from a one-to-one mapping and rule-based words. In the present study, we will refer to the average results for words and nonwords, regardless of the orthographic condition.

### Phonological test

Phonological awareness was tested for the rhyme linguistic unit in bisyllabic words, with CV.CV and CVC.CV syllabic structure, where C represents a consonant and V a vowel. Children were administered a version of the same-different task [26], which consists of judging whether there is a common sound in a pair of words (e.g., <b*ol*so-p*ol*pa> and <xisto-belga> sharing and not sharing the first rhyme, respectively).

## Procedure

Naming tests were administered with the Cognitive Workshop software, developed by the University of Dundee, Scotland, UK, and by the University of Jyvaskyla, Jyvaskyla, Finland which allows accuracy reaction times on-line recording. The items were shown in a 12-inch laptop screen. After a 1000 ms warning signal (*) followed by a 1000 ms delay, the stimulus was presented on the screen for up to 10 s. Participants were required to identify each item as quickly and as accurately as possible. Responses were recorded on-line on a digital sound file, and correct responses were computed. Correct responses and errors were also scored on-line during the experiment and then checked off-line using the digital sound file.

Each child was tested individually. Children were asked to participate in a “word game” and were explained that these tasks did not constitute a school evaluation. Before each task, children were introduced with practice trials. All sessions were run in a quiet room.

The phonological awareness test was orally administered and answered. Children were told there was a clown who liked hearing the same bits of sound. Children were asked to point the “happy clown” if they heard “likewise bits of sounds” and the “sad clown” if they did not.

## Data classification using SOM

Kohonen’s SOM [18] are artificial neural networks with one layer of neurons disposed in a one- or two-dimensional lattice (or map), where each neuron is connected to all the source nodes in the input layer (cf. Figure 1).

These networks are trained so that different regions of the map become active according to the input space division. A trained SOM is therefore capable of recognizing patterns at its inputs. The training algorithm description is beyond the scope of this work and can be found, for instance, in the work of Kohonen [18]. One interesting feature of SOM relies on the fact that they do not require target data, so the classification procedure is unsupervised. When a given data point is input to the SOM, the neuron whose weight vector is the closest to that input pattern is first identified and named the best matching unit. In the present work, the closeness between the inputs and the neurons is measured by the Euclidian distance between the input vector and the weight vectors, which are subsequently trained so they move closer to the input data points. The result of the training phase is a neural network whose neurons are associated with groups or patterns in the input data set. In this study, the SOM used has three neurons corresponding to three reading levels: good, average, and poor reading abilities. All networks were trained for 5000 training epochs because no significant changes were noticed with further training. The inputs used for SOM training were the PATT results and the nonword and word reading accuracy. To obtain a uniform input space range, the phonological task completion times were multiplied by a factor of 20 before being input to the SOM (cf. Figure 2).

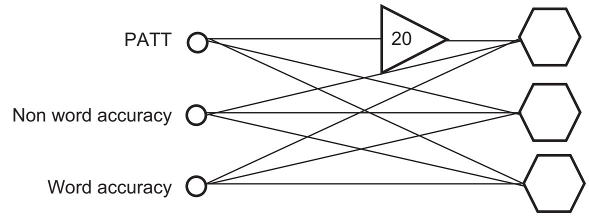


**Figure 2** SOM used in the present study.

## Results

Figure 3 presents the neighborhood distances and number of hits after the SOM training. Regarding the neighborhood distances, lighter colors indicate shorter distances, whereas darker colors indicate the opposite. It is possible to see that classes 2 and 3 are close to each other and away from class 1. This indicates that the input data can be separated into two major classes (class 1 and class 2+3), clearly distinguished from each other. At the rightmost cluster, two subclasses exist, classes 2 and 3.

To provide a more meaningful interpretation of these results, Figure 4 presents the mean of the word and nonword reading accuracy, whereas Figure 5 presents the phonological task completion time.

For both second and third grades, the class corresponding to the worst results (class 1) is clearly detached from the one corresponding to intermediate results (class 2) for both word and nonword accuracy. The quantitative difference between results of class 1 and class 2 for words reaches 44% in the second grade and 31% in the third grade, whereas for nonwords, this difference is about 48% in the second grade and 39% in the third grade. Although not so pronounced, the difference between class 2 and class 3 is clearly visible for both words (12% and 4% for the second and third grades, respectively) and nonwords (15% and 8% for the second and third grades, respectively). Regarding the fourth

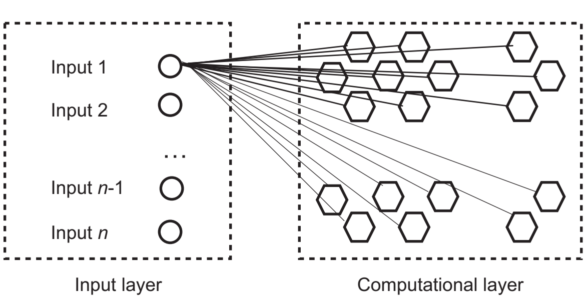


**Figure 1** SOM with two-dimensional hexagonal grid and n inputs (for clarity, only the connections of the first input are shown).

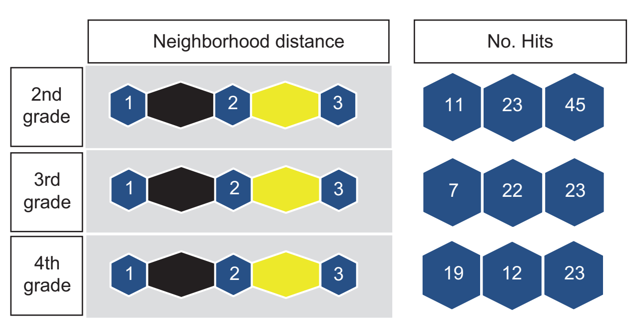


**Figure 3** SOM neighborhood distances and number of hits.

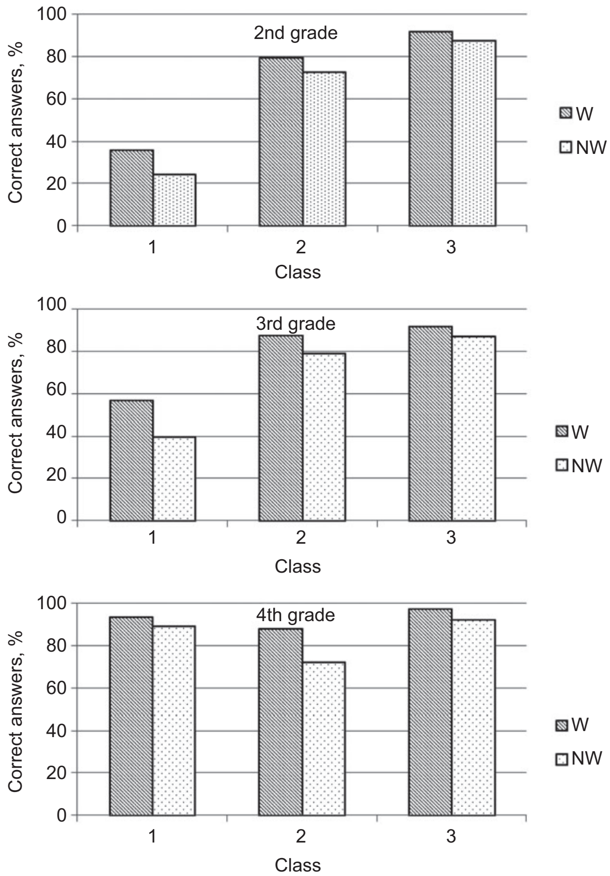


**Figure 4** Word (W) and nonword (NW) percentage correct averages for each class and grade.

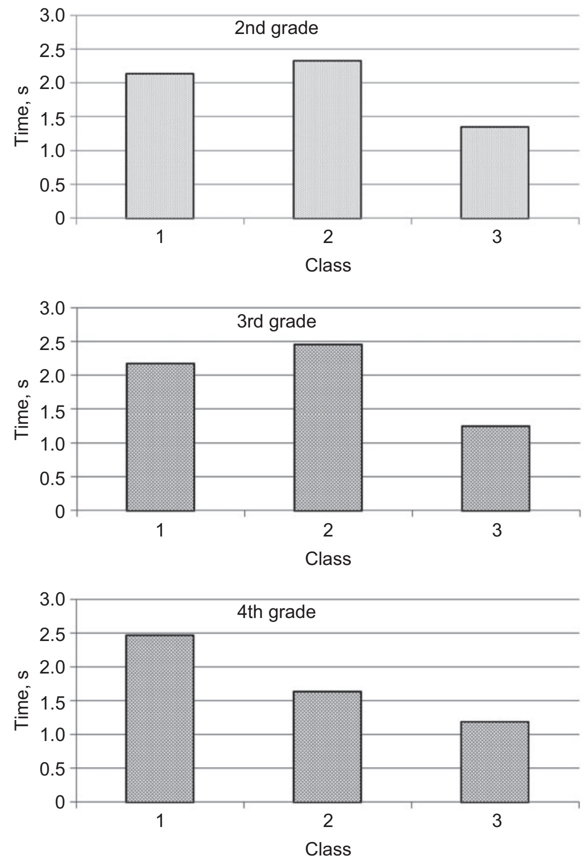


**Figure 5** PATT average for each class and grade.

grade, word and nonword accuracy do not allow a clear distinction among the three classes. In fact, the difference among classes in fourth grade is better understood if the phonological awareness results are inspected (cf. Figure 5): there is a very expressive difference not only between class 1 and class 2 average task times (825 ms) but also between class 2 and class 3 (445 ms). The phonological results have also played an important role on separating class 1 from class 3 in grades 2 and 3 (964 and 1200 ms), although the role of this variable seems less significant in the explanation of differences between class 1 and class 2, as the length of time is roughly the same for both.

Another interesting result in the categorization performed by the SOM can be seen in the average lexicality effect (difference between word and nonword accuracy results) presented in Figure 6. This effect is at least two times quantitatively stronger for class 1 than for classes 2 and 3, both in the second and third grades. The same tendency cannot be observed in the fourth grade, as the lexicality effect loses expression in later years, when the reading experience eventually leads to a better reading accuracy, although the difference remains if the response time is considered [15].

## Discussion

When analyzing the categories identified by the SOM, it is possible to distinguish (i) a clear distance between classes and (ii) the variables that play the most important role to differentiate children.

As described in the Participants section, two groups of children were selected: those with results equivalent or below the 5th percentile and those with results equivalent or above the 50th percentile. In face of this preselection, the SOM was expected to have clearly differentiated at least two groups (poor and good readers). This expectation was confirmed, as shown in Figure 3, so the classification performed by the SOM is in accordance with the reading age test results. Also, the existence of two subclasses with closer, but still differentiable, results replicates the reading age data, in the sense that good and average

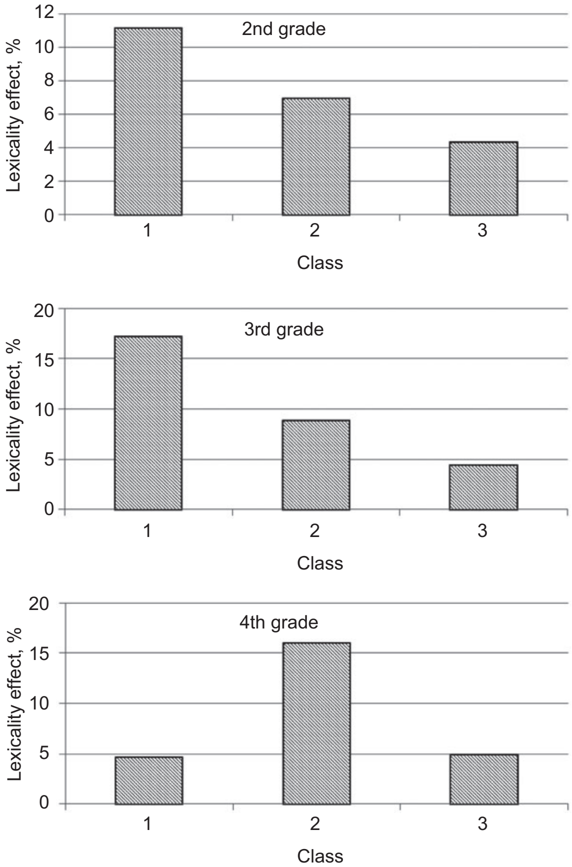


**Figure 6** Lexicality effect average results for each class and grade.

readers constitute two different subclasses, although with results closer to each other than to poor readers.

The SOM took into account both accuracy results (word and nonword reading) and latency results (phonological test) for distinguishing the three classes, although accuracy seems to have played a more important role in the second and third grades, whereas latency seems to have played a more important role in the fourth grade. The weaker importance of accuracy in fourth grade can be understood because even poor readers eventually benefit from experience to acquire a better reading accuracy, thus presenting results closer to normal readers. In future studies, it would be worthwhile to analyze not only the reading accuracy but also the reading latencies because literature suggests that this measure is more sensitive, thus implicating long-lasting results than accuracy [15, 27].

The classification made by the SOM also appears to have taken into account the lexicality effect, which is a remarkable "repetition" of a classic effect differentiating dyslexic children from good readers [14, 15, 28]. This constitutes an interesting result because this effect was not intentionally trained but was inferred by the unsupervised SOM training.

Finally, the phonological results played an important role in all three grades, most particularly in the fourth grade. If only latency results had been inspected in SOM, two distinct categories would have arisen in grades 2 and 3 and a third one in the fourth grade. Again, this is explained by the classic accuracy effect for both words and nonwords (with special disadvantage for nonwords, as expressed by the lexicality effect) during the initial school years, whereas experience eventually leads to closer results in terms of accuracy, although more sensitive measures (such as phonological tests) still highlight the differences between poor and good readers. Indeed, whereas in the fourth grade we cannot differentiate class 1 from class 2 based on accuracy, the difference among all three classes is very clear if phonological results are taken into account.

## Conclusions

This work focused on dyslexia assessment, a learning disability that may play a crucial role in children's progress in school. It was shown that the PATT can constitute a useful selective measure, particularly in the fourth grade, when classical variables such as word and nonword reading accuracy lose their discriminative capabilities. It was also shown that the use of SOM to classify children's reading abilities can successfully categorize children and capture meaningful measures such as the lexicality effect.


### Conflict of interest statement

**Authors' conflict of interest disclosure:** The authors stated that there are no conflicts of interest regarding the publication of this article.
**Research funding:** None declared.
**Employment or leadership:** None declared.
**Honorarium:** None declared.

## References


1. Lyon GR, Shaywitz SE, Shaywitz BA. Defining dyslexia, comorbidity, teachers' knowledge of language and reading. Ann Dyslexia 2003;53:1–14.
2. APA. Diagnostic and statistical manual of mental disorders, 4th ed. (text revision). Washington, DC: APA, 2000.
3. Snowling MJ. State-of-science review: dyslexia. Foresight Mental Capital and Wellbeing Project. The Government Office for Science, 2008. Available at: http://www.bis.gov.uk/assets/foresight/docs/mental-capital/sr-d2_mcw_v2.pdf.
4. Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD. Prevalence of reading disability in boys and girls: results of the connecticut longitudinal study. J Am Med Assoc 1990;264:998–1002.
5. Fluss J, Ziegler J, Ecalle J, Magnan A, Warszawski J, Ducot B, et al. Prévalence des troubles d'apprentissages du langage écrit en début de scolarité: l'impact du milieu socioéconomique dans 3 zones d'éducations distinctes. Arch Pediatr 2008;15:1049–57.
6. Moll KL, Double K. Dissociation between reading and spelling deficits. Sci Stud Read 2009;13:359–82.
7. Vale AP, Sucena A, Viana F. Prevalência Da Dislexia Entre Crianças Do 1° Ciclo Do Ensino Básico Falantes Do Português Europeu. Rev Lusófona Educ 2011;18:45–56 [in Portuguese].
8. Stanovich KE. Matthew effect in reading: some consequences of individual differences in the acquisition of literacy. Read Res Q 1986;21:360–407.
9. Lyytinen H, Erskine J, Kujala J, Ojanen E, Richardson U. In search of a science-based application: a learning tool for reading acquisition. Scand J Psychol 2009;50:668–75.
10. Lyytinen H. State-of-science review: SR-D12. New technologies and interventions for learning difficulties: dyslexia in Finnish as a case study. Foresight Mental Capital and Wellbeing Project: The Government Office for Science, 2008. Available at: http://www.bis.gov.uk/assets/foresight/docs/mental-capital/sr-d12_mcw.pdf.
11. Torgesen JK. Catch them before they fall: identification and assessment to prevent reading failure in young children. Am Educ 1998;22:1–8.
12. Ramus F. Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? Curr Opin Neurobiol 2003;13:212–8.
13. Nag S, Snowling MJ. Cognitive profiles of poor readers of Kannada. Read Writ 2010;24:677–8.
14. Ziegler JC, Goswami UC. Reading acquisition, developmental dyslexia and skilled reading across languages: a psycholinguistic grain size theory. Psychol Bull 2005;131:3–29.
15. Sucena A, Castro SL, Seymour P. Developmental dyslexia in an orthography of intermediate depth: the case of European Portuguese. Read Writ 2009;22:791–810.
16. Cardoso-Martins C, Pennington BF. Qual é a contribuição da nomeação seriada rápida para a habilidade de leitura e escrita? Evidências de crianças e adolescentes com e sem dificuldades de leitura. Psicol Reflex Crít 2001;14:387–97.
17. Landerl K, Wimmer H, Frith U. The impact of orthographic consistency on dyslexia: a German-English comparison. Cognition 1997;63:315–34.
18. Kohonen T. Self-organizing maps. Proc IEEE 1990;78:1464–80.
19. Kohonen T. Engineering applications of the self-organizing map. Proc IEEE 1996;84:1358–84.
20. Ceruti MG, editor. The relationship between artificial intelligence and data mining: application to future military information systems. 2000 IEEE International Conference on Systems, Man, and Cybernetics, 2000.
21. Afolabi MO, Olude O, editors. Predicting stock prices using a hybrid Kohonen self organizing map (SOM). 40th Annual Hawaii International Conference on System Sciences, (HICSS 2007), Hawaii, 2007.
22. Novák D, Kordik P, Macas M, Vyhnálek M, Brzezny R, Lhotská L, editors. School children dyslexia analysis using self organizing maps. Engineering in Medicine and Biology Society, 2004 IEMBS '04 26th Annual International Conference of the IEEE, San Francisco, CA, USA, 2004.
23. Palacios A, Sánchez L, Couso I. Diagnosis of dyslexia with low quality data with genetic fuzzy systems. Int J Approx Reason 2010;51:993–1009.
24. Loizou A, Laouris Y. Developing prognosis tools to identify learning difficulties in children using machine learning technologies. Cogn Comput 2011;3:490–500.
25. Sucena AC, Lui Sao. Aprender a Ler e Avaliar a Leitura. O TIL: Teste de Idade de Leitura, 3rd ed. Coimbra: Almedina, 2010.
26. Treiman R, Zukowski A. Levels of phonological awareness. In: Shankweiler D, editor. Phonological processes in literacy: a tribute to Isabelle Liberman. Hillsdale, NJ: Lawrence Erlbaum Associates, 1991:67–83.
27. Wimmer H, Goswami U. The influence of ortographic consistency on reading development: word recognition in English and German children. Cognition 1994;51:91–103.
28. Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: a dual route cascaded model of visual word recognition and reading aloud. Psychol Rev 2001;108:204–56.