Discriminatory validity of dyslexia screening tasks in French school age children.

Guylaine Le Jan, Régine Le Bouquin Jeannès, Nathalie Costet, Gérard Faucon

Discriminatory validity of dyslexia screening tasks in French school age children

Guylaine Le Jan [1] [2], Régine Le Bouquin Jeannès [1] [2], Nathalie Costet [1] [2], Gérard Faucon [1] [2]

[1] INSERM, U 642, Rennes, F-35000, France
[2] Université de Rennes 1, LTSI, F-35000, France
LTSI, Campus de Beaulieu, Université de Rennes 1, 35042 Rennes Cedex, France
Tel. +33.2.23.23.62.20 - Fax. +33.2.23.23.69.17 – guylaine.lejan@univ-rennes1.fr

Discriminatory validity of dyslexia screening tasks in French school age children

Guylaine Le Jan (1) (2), Régine Le Bouquin Jeannès (1) (2), Nathalie Costet (1) (2), Gérard Faucon (1) (2)

(1) INSERM, U 642, Rennes, F-35000, France
(2) Université de Rennes 1, LTSI, F-35000, France
LTSI, Campus de Beaulieu, Université de Rennes 1, 35042 Rennes Cedex, France
Tel. +33.2.23.23.62.20 - Fax. +33.2.23.23.69.17 – guylaine.lejan@univ-rennes1.fr

***Abstract*—Dyslexia is a specific disorder of language. Researches led on dyslexia origin have conducted to multiple hypotheses and various rehabilitation treatments. In this context, practitioners can be interested in using an automatic tool to help in diagnosing dyslexia. This tool should evaluate children's own deficit and advise adapted rehabilitation. This paper presents the conception of a preliminary test containing the most representative dyslexia evaluation tasks from literature and the first results concerning the discriminatory validity of this preliminary test in French school age children (8-10 years). Moreover a selection of significant tasks to optimize the detection of dyslexia is proposed. These tasks will build up the first step of the automatic tool.**

I. INTRODUCTION

Developmental dyslexia affects in France about 5% of school age children. It is traditionally defined as an enduring and heavy impairment of reading ability in spite of normal intelligence and adequate educational opportunities. Dyslexics have a specific disorder of written language and can have some associated deficits like: attention deficit, visuo-attentional deficit, auditory and memory deficits. Researches led on dyslexia origin have conducted to multiple theories (*i.e.* phonological theory, rapid auditory processing theory, cerebellum theory, *etc*.). These multiple theories created various diagnosis methods and treatments which are sometimes inadequate. In this context, the determination of the more significant tasks and the proposal of an automatic tool to help in diagnosing dyslexia can be of great interest for pratictioners. This tool should evaluate children's own deficit and advise adapted rehabilitation. In order to develop this tool, we must in a first step select dyslexia evaluation tasks and study their capacity to discriminate two populations: children with dyslexia and normal readers. The present article describes a set of selected tasks named "preliminary test" containing the most representative dyslexia evaluation tasks and reports their capacity to detect dyslexia. The first part presents the context of this work. The preliminary test of dyslexia screening is described in a second part. The next sections present the study of its discriminatory validity and propose a reduced preliminary test.

II. CONTEXT

Three forms of developmental dyslexia are generally identified: phonological dyslexia [1], surface dyslexia [3] and mixed dyslexia. Given multiple etiologies, it is particularly difficult to determine its causes. During the 20 last years, many behavioural studies and neurological investigations demonstrated that dyslexics have difficulties in phonological processing in new words reading, phonemic analysis and short-term memory. Such a deficit concerns at least 70% of dyslexic children. Although this phonological theory is defended by a dominant current, four main hypotheses have been suggested: auditory hypothesis, visual hypothesis, magnocellular hypothesis and cerebellar hypothesis. According to auditory theory, dyslexics would have poor performance on a number of auditory tasks - tone discrimination, temporal order judgment [17], repetition tasks [12], backward masking - and would have more difficulties in perceptual discrimination of speech sounds [13]. Visual hypothesis is based on the observation of visual deficits such as bi-ocular fixation instability, a visual stress and visuo-attentional disorder [18]. Magnocellular hypothesis [16] integrates the preceding theories (auditory and visual) in defending a theory that would affect auditory and visual magnocellular pathway. At last, cerebellar hypothesis [10] is based on the observation that some dyslexic children have a motor deficit. They would have a poor capacity in motor coordination and problems of equilibrium, *etc*. Most of these theories are based on a unique modality (auditory, visual or motor). More recent studies show that 40% of bad readers have an auditory deficit, a minority has visual problems and around 30 to 50% of them have a motor deficit [11]. It appears that only phonological deficit constitutes a common factor to developmental dyslexia. The other deficits associated to dyslexia are considered by Ramus [11] as relevant to co-morbidity. The development of an automatic tool to help in diagnosing dyslexia would allow to take in consideration associated deficits in order to facilitate the diagnosis and to guide the remediation. The first step of this elaboration consists in proposing a set of selected tasks for dyslexia and evaluating its discriminatory capacity in school age children.

III. THE PRELIMINARY TEST

A. *Principle*

Twenty French children with developmental dyslexia aged 8-10 years and fifty-six normal school children of the same age were tested in the preliminary test. The poorest reading ability in the normal children group was 18 months below their chronological age and the reading ability of dyslexics was on average 27 months below their chronological age, which defined them as severely impaired in reading the

"Alouette test" defined in § III.B. The preliminary test is a set of selected tasks for dyslexia screening. More precisely, it should recover all suitable deficits implicated in dyslexia and should be able to recognize dyslexic children profile (what are the main deficits? what are the associated deficits? what is the degree of severity?). In order to carry out the test in the best conditions, each child was assessed individually (except for the dictation *cf.* § III.B) during three sessions of 45 minutes. Some experiments were controlled by computer in which a software platform was developed [7]. All sound stimuli were digitally generated and were delivered idiotically via Sennheiser HD 180 earphones.

B. Description of the preliminary test

The preliminary test is composed of 8 categories of tasks:

1) Reading tasks**:**

- The Alouette test [6]: it gives lexical age (*i.e*. reading level). The child reads a text during three minutes. The level of reading is defined by the speed and the accuracy of reading.
- Reading of words and pseudo words: it is carried out on 4 sheets of 20 words which are grouped according to their frequency and regularity on 2 sheets of 20 pseudo-words. Exactitude of pronunciation and reading time are recoded and each sheet is evaluated on 20 points. This test allows to determine the form of dyslexia (phonological, surface or mixed).

2) Memory tasks**:** they are composed of two tasks: number span tasks (forward and backward) and spatial span tasks (forward) via Corsi's blocks. They give a verbal short-term memory span, spatial span and work memory span. The work memory appears to be implicated in reading and notably in its learning. According to some studies, dyslexics would have poor performance in verbal short-term memory [15].

3) Attention tasks**:** it is extracted from the BREV ("Batterie Rapide d'EValuation des fonctions cognitives"): children must cross out as quickly as possible all "3" placed on two tests sheets: one during 20 s and the other during 60 s.

4) Phonological tasks**:**

- Metaphonological tasks: they require phonological awareness. Four different tasks are assessed: phonemic segmentation task (segment the word in phonemes), spoonerism task (switch syllables), initial phonemic omission task (omit the first phoneme of each word presented) and task of rimes judgment (find the word which does not rime with the three others). These tasks return a score on respectively 16, 10, 12, 8 points.
- Phonological automatism task: it regroups two tasks: speed denomination (denominate as quickly as possible a series of letters and a series of colours) and lexical discrimination (recognize if the pronunciation of two words is the same or not).
- Morphology task: children must find a pseudo-affixed word among affixed words (example in French: recoller, *regretter*, repartir, reparler). It gives a score on 6 points. The knowledge of word morphology is considered by some authors as a capacity called up during the reading [3].

5) Motor task**:** it is an extract from NEPSY ("bilan NeuroPSYchologique"): the children must execute manual motor sequences noted on 60 points and an exercise of "tapping", that is an evaluation of digital sleight and motor speed.

6) Visuo-attentional task**:** dyslexics would have difficulties in the treatment of visual information when this information is presented rapidly [18]. A partial report of letters was integrated in the computer: following a central point on computer's screen, a series of 5 letters appears during 250 ms, a dash comes under one of the letters, and then the children must indicate which letter it is.

7) Writing task**:** dictation extracted from the BELEC [9].

8) Auditory tasks**:**

- TMTF (Temporal Modulation Transfer Function) task: this task evaluates the ability to process auditory temporal-envelope cues. According to Lorenzi [8], modulation sensitivity seems poorer for children with dyslexia than for normal children. To address this issue, temporal modulation transfer function is measured. TMTF is the detection threshold of sinusoidal amplitude modulation applied to a white noise carrier, as a function of frequency. For practical purposes, children are asked to listen to two sounds of 500 ms length: a white noise and a modulated white noise at a given modulation depth. For each trial, these two sounds are successively presented in random order to the listener. This test must run less than 45 minutes due to the children fatigability. Consequently, only three modulation frequencies were experimented (4 Hz, 16 Hz and 128 Hz). The test begun by a depth of 0 dB and then changed according to the child response. The threshold detection was obtained using an adaptive two-interval (2I), two-alternative forced-choice (2AFC) [4]. So, three thresholds of modulation for each frequency are noted.
- VOT (Voice Onset Time) tasks: VOT is the time between the release of the consonant and the start of vocal fold vibration (voicing), it is measured in milliseconds (ms). By convention, when voicing starts before the release of the consonant, VOT is negative; when voicing and consonant release happen simultaneously, VOT equals 0 ms; when voicing starts after the release of the consonant, VOT is positive. VOT quantifies the degree of phonetic voicing. The test consists in producing a continuum whose extremities are constituted of two syllables which differ by their VOT and intermediate syllables allow linking the extremities by progressive variation of VOT. A difference of 20 ms between VOT values of two syllables is perceptible only if the syllables belong to distinct phonemic categories. For example, the syllables /ba/ and /pa/ differentiate by respectively negative and positive VOT. The production of several intermediate VOT values generates a continuum of syllables perceived like either /ba/ or /pa/. From a continuum ranging from -40 ms to 40 ms, two exercises are proposed: (*i*) an identification task where the child listens to a syllable. He must indicate if he hears rather /ba/ or /pa/. This test allows evaluating an identification slope that is calculated using a linear regression analysis performed on the data

point 100% /pa/ identification to 100% /ba/ identification, (*ii*) a discrimination task where two syllables are presented. The VOT difference between these two syllables is always 20 ms. In this second exercise, the child must indicate if the syllables are the same or not. Normal subjects present a discrimination peak around a VOT of 0 ms. Such a peak is not recovered for children with dyslexia [14]. Moreover, predicted VOT discrimination values were calculated from VOT identification values. Then, the differences for each pair of syllables between predicted discrimination values and observed discrimination values are summed in order to give another variable.

Finally, all results obtained by children to the preliminary test are stored in a data table where rows correspond to the children and columns to the variables (*e.g.* Alouette test, verbal forward span, *etc.*). So, this data table is composed of many different quantitative variables. First, we want to determine which tests have a real capacity to discriminate dyslexics from normal children and secondly we want to evaluate their performance in the detection of dyslexia.

IV. METHOD

A. *Analysis of discriminatory properties of individual tasks*

A Mann-Whitney test was used to compare the results obtained by normal children and children with dyslexia on each task. This nonparametric test compares the distributions of a variable in two independent samples. A 0.05 significance level was chosen to state whether each task is discriminant (H): H = 0 indicates that the null hypothesis (*i.e* the medians are equal in both groups) cannot be rejected at the 5% level ($p > 0.05$). H = 1 indicates that the medians differ between both groups, then the null hypothesis can be rejected at the 5% level ($p < 0.05$). The tasks identified as non-discriminant were excluded for the next step of the analysis.

B. *Analysis of discriminatory capacity of preliminary test*

A Fisher discriminant analysis including the tasks previously identified as discriminant was then implemented. A parametric analysis was chosen, under the assumptions that each group was normally distributed and the variance/covariance matrix for each group was the same. Furthermore, prior probabilities were assigned to be equal, as the costs of misclassification in both groups. Thus the discriminant function used was linear [4].

A stepwise procedure was then implemented in order to find a parsimonious model and select the best combination of predictors. As our purpose was to predict group membership, the classification accuracy of the resulting function was assessed through the classification matrix which compares classification groups to actual groups. The overall percentage of children correctly classified (hit ratio), the sensitivity (detection rate of dyslexic children), the specificity (detection rate of non-dyslexic children), the false-positive rate (percentage of children classified as dyslexic who were actually not) and the false-negative rate (percentage of children classified as non-dyslexic who were actually dyslexic) were estimated using a bootstrap method.

V. RESULTS

A. *Determination of discriminatory tasks*

The results on Mann-Withney test are given in Table 1.

The scores of visual memory span, attention, lexical discrimination morphology and tapping tasks do not differentiate the two groups ($p > 0.05$). Moreover, contrary to the auditory hypotheses, the TMTF tasks do not seem to have any capacity of discrimination. These results suggest that the modulation thresholds measured in children with dyslexia concern probably a minority of dyslexics. For the reading tasks, only the score for the reading of frequent regular words does not show significant difference between the two groups ($p = 0.166$). But this task is important to determine more precisely the form of dyslexia (phonological, surface or mixed). The other reading tasks have a lower probability ($p < 10^{-4}$). Speed denomination and all metaphonological tasks, except segmentation task, discriminate dyslexics and normal children. We can conclude that a majority of dyslexics have some phonological deficits. The dictation scores and the slope of identification curve in the VOT task appear to be significantly different in both populations.

Variables names	P-value	Hypothesis
Alouette test		
Lexical age (in months)	$< 10^{-3}$	1
Memory tests		
Verbal forward span	< 0.05	1
Verbal backward span	< 0.05	1
Spatial span	NS	0
Attention test		
3 crossing out (60s)	NS	0
3 crossing out (20s)	< 0.05	1
Reading tests		
Frequent regular words	NS	0
Reading time of frequent regular words	$< 10^{-3}$	1
Frequent irregular words	$< 10^{-3}$	1
Reading time of frequent irregular words	$< 10^{-3}$	1
Number of regularization for frequent irregular words	$< 10^{-3}$	1
Few frequent regular words	$< 10^{-3}$	1
Reading time of few frequent regular words	$< 10^{-3}$	1
Few frequent irregular words	$< 10^{-3}$	1
Reading time of few frequent irregular words	$< 10^{-3}$	1
Number of regularization for unfrequent irregular words	$< 10^{-3}$	1
Near phonologically pseudo words	$< 10^{-3}$	1
Reading time of near phonologically pseudo words	$< 10^{-3}$	1
Pseudo words	$< 10^{-3}$	1
Reading time of pseudo words	$< 10^{-3}$	1
Metaphonological tests		
Segmentation	NS	0
Omission	$< 10^{-3}$	1
Judgment of rime	NS	0
Spoonerism	$< 10^{-3}$	1
Phonological automatism tests		
Lexical discrimination	NS	0
Speed denomination letters	< 0.05	1
Speed denomination colours	< 0.05	1
Test of capacity call up during reading		
Morphology	NS	0
Visuo-attentional test		
Partial report of letters (total score)	< 0.05	1
Motor tests		
Manual motor sequences	$< 10^{-3}$	1
Tapping	< 0.10	0
Dictation		
Contextual consistent writing form for unfrequent words	$< 10^{-3}$	1
Contextual consistent writing form for frequent words	$< 10^{-3}$	1
Contextual unconsitent writing form for unfrequent words	NS	0
Minority contextual unconsistent writing form for unfrequent words	$< 10^{-3}$	1
Auditory tests		
TMTF 4	NS	0
TMTF 16	NS	0
TMTF 128	NS	0
Slope of identification VOT curve	< 0.05	1
VOTdiscrimination score	NS	0

Table 1: Mann-Whitney test comparing normal readers to children with dyslexia (NS = Non Significant)

B. Evaluation of the capacity of preliminary test to detect dyslexia

According to the Mann-Whitney test (Table 1), 27 discriminatory variables have been identified. A discriminatory analysis was directly applied on these variables. The classification matrix (Table 2) allows to estimate the quality of the decision rule (Table 3) through different criteria (sensibility, specificity, *etc*.). The global percent of correctly classified is high (89.59%). The discriminatory function identifies correctly the normal readers and the dyslexics: 93.82% of normal readers are classified as non-dyslexics and 77.75% of dyslexics are correctly classified. The false-positive rate (22%) is higher than the false-negative rate (6.18%). So, the high overall hit ratio implies a high false-positive rate too.

		Predicted group			Total
		D	ND	Std dev	
Actual group	D	15.55	4.45	2.1	20
	ND	3.46	52.54	0.72	56

Table 2: Classification matrix (bootstrap estimated frequencies) obtained with 27 variables (D = Dyslexic, ND = Non Dyslexic)

global % of correctly classified (hit ratio)	89.59
Sensibility	77.75
Specificity	93.82
False-positive rate	22.25
False-negative rate	6.18

Table 3: Predictive accuracy of the model including 27 variables

These good classification results were obtained with 27 variables, which is much. We then tried to reduce the number of variables to use, while maintaining or even increasing the predictive efficiency of the discriminatory function.

C. Determination of the best combination of tasks among the preliminary test to increase dyslexia detection

In this part, a stepwise discriminatory analysis was applied to the 27 variables. Three models with respectively three, four and five variables were selected. The variables included in these models are presented in Table 4.

Model with 4 variables
Few frequent regular words Reading time of few frequent irregular words Spoonerism Contextual consistent writing form for frequent words
Model with 5 variables
Few frequent regular words Reading time of few frequent irregular words Spoonerism Contextual consistent writing form for frequent words Slope of identification VOT curve
Model with 6 variables
Few frequent regular words Reading time of few frequent irregular words Spoonerism Contextual consistent writing form for frequent words Slope of identification VOT curve Contextual consistent writing form for unfrequent words

Table 4: Variables included in each model

The three selected models are nested: the second model includes the first model and the last model includes the second model (Table 4). So the choice between the three models can be based on the clinical relevance of the additional tasks that are proposed in the more complex models.

The Figure 1 shows the predictive accuracy of each selected model. Globally the quality of decision rules are better than for the model including 27 variables (*cf.* § 5.B). The best hit ratio and the highest specificity are obtained with the model including 6 variables (94% of individuals correctly classified and 97% of normal readers classified as non-dylexics). The model including 4 variables has the highest detection rate for dyslexia (93% of dyslexics correctly classified). The first model shows the poorest false-positive rate (0.069%) and the third model the poorest false-negative rate (0.03%).

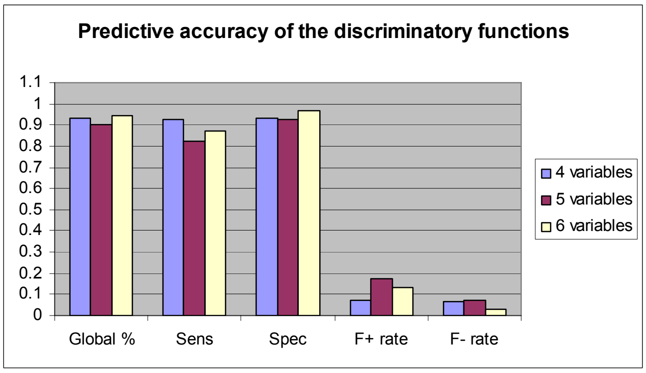

Figure 1: Predictive accuracy (global percent of correctly classified, sensibility, specificity, false-positive rate and false-negative rate) for the stepwise selected model (including four, five and six variables)

In order to favour the dyslexia detection, the reduced preliminary test should have the lowest false-negative rate. So, we propose to select the third model including 6 variables which correspond to 5 tasks: two reading tasks, one spoonerism, one dictation and one VOT identification task.

These 6 variables are deemed pertinent by clinicians. They belong to four different categories of tasks (reading tasks, writing tasks, metaphonological and auditory tasks) and they describe the main observed difficulties in children with dyslexia.

VI. Conclusion

To conclude, this study allowed to evaluate the discriminatory validity of detection tasks proposed in the literature and identified the set of tasks which optimizes the detection of dyslexia. Four tasks seem to have a poor discriminatory capacity: spatial span memory, morphology, lexical discrimination and TMTF tasks. The distributions of spatial span memory and lexical discrimination for dyslexics are the same as those for normal readers. But for the 4Hz TMTF task, 20% of dyslexics have a low threshold detection against 8% of normal readers and for the morphology task, 10% of dyslexics have a score inferior to the worst scores of normal readers. So, they can help to constitute sub-groups.

Moreover, thanks to this study, the preliminary test can be reduced to 5 tasks: 2 reading tasks, the spoonerism task, the dictation and a VOT identification task. The reduced preliminary test constitutes a "pre-test" of an adaptive protocol of dyslexia screening and will be followed by a full test that should help in identifying more precisely sub-groups of dyslexia.

References

[1] M. F. Beauvois and J. Derouesne, "Phonologica alexia: three dissociations", *Journal of Neurology, Neurosurgery and Psychiatry*, 42, 1979, pp.1 115- 1 124.

[2] P. Colé, N. Marec-Breton, C. Royer and J.E. Gombert, "Morphologie des mots et apprentissage de la lecture", *Rééducation orthophonique*, 213, pp. 57-76, 2003.

[3] M. Coltheart, J. Masterson, S. Byng, M. Prior and J. Riddoch, "Surface dyslexia", *Quarterly Journal of Experimental Psychology*, 35, pp. 469-595, 1983.

[4] C.J. Huberty, "Applied Discriminant Analysis", *Wiley-Interscience*, 1994.

[5] W. Jesteadt, "An adaptive procedure for subjective judgments", *Percept. Psychophys,* 28, pp. 85-88, 1980.

[6] P. Lefavrais, "Description, définition et mesure de la dyslexie", Edition E.C.P.A., 1965.

[7] G. Le Jan, N. Troles, R. Le Bouquin Jeannès, G. Faucon, J.E. Gombert, P. Scalart, D. Pichancourt, "Développement d'une plate-forme logicielle en vue de l'élaboration d'un outil d'aide au diagnostic de la dyslexie", *Manifestation des jeunes chercheurs STIC (Majecstic)*, 2006.

[8] C. Lorenzi, A. Dumont and C. Füllgrabe, "Use of temporal cues by children with developmental dyslexia", *Journal of Speech, Language, and Hearing Research*, 43, pp. 1367-1379, 2000.

[9] P. Mousty, J. Leybaert, J. Alegria, A. Content and J. Morais, BELEC : une batterie d'évaluation du langage écrit et de ses troubles. Dans J. Gregoire & B. Piérat (Eds). Evaluer les troubles de la lecture: les nouveaux modèles théoriques et leurs implications diagnostiques, pp. 127-145, Bruxelles : De Boeck Université, 1994.

[10] R.I. Nicolson, A.J. Fawcett, and P. Dean, "Dyslexia, development and the cerebellum", *Trends in Neuroscience*, 24(9), pp. 515-516, 2001.

[11] F. Ramus, S. Rosen, S.C. Dakin, B.L. Day, J.M. Castellote, S. White, & U. Frith, "Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults", *Brain*, 126(4), pp. 841-865, 2003.

[12] V. Rey, S. De Martino, R. Espesser and M. Habib, "Temporal processing and phonological impairment in dyslexia: effect of phoneme lengthening on order judgment of two consonants", *Brain Lan*., 80, pp. 579-591, 2002.

[13] W. Serniclaes, L. Sprenger-Charolles, R. Carré and J.F. Démonet, "Perceptual discrimination of speech sounds in dyslexia", *Journal Speech Language Research*, 44, pp. 384-399, 2001.

[14] W. Serniclaes and L. Sprenger-Charolles, "Categorical perception of speech sounds and dyslexia", *Special issue on language disorders and reading acquisition*, 10, vol. 1, 2003.

[15] L. Sprenger-Charolles, P. Colé, P. Larcet, and W. Serniclaes, "On subtypes of developmental dyslexia: evidence from processing time and accuracy scores", *Canadian Journal of experimental psychology*, 54, pp. 88-104, 2000.

[16] J. Stein, "The magnocellular theory of developmental dyslexia", *Dyslexia*, 7, pp. 12-36, 2001.

[17] P. Tallal, "Auditory temporal perception, phonics, and reading disabilities in children", *Brain Lan*., 9, pp. 182-198, 1980.

[18] S. Valdois, M.L. Bosse, B. Ans, S. Carbonnel, M. Zorman, D. David and J. Pellat, "Phonological and visual processing deficits can dissociate in developmental dyslexia: evidence from two case studies", *Reading and Writing: an Interdisciplinary Journal,* 16, pp. 541-572, 2003.