Screening Dyslexia for English Using HCI Measures and Machine Learning

Luz Rello
HCI Institute
Carnegie Mellon University
Pittsburgh, USA
luzrello@cs.cmu.edu

Enrique Romero
Department of Computer Science
Universitat Politècnica de Catalunya
Barcelona, Spain
eromero@cs.upc.edu

Maria Rauschenberger
Web Science and Social Computing Research Group
Universitat Pompeu Fabra
Barcelona, Spain
maria.rauschenberger@upf.edu

Abdullah Ali
Information School
University of Washington
Washington, USA
xyleques@uw.edu

Kristin Williams
HCI Institute
Carnegie Mellon University
Pittsburgh, USA
krismawil@cs.cmu.edu

Jeffrey P. Bigham
HCI & LTI Institutes
Carnegie Mellon University
Pittsburgh, USA
jbigham@cs.cmu.edu

Nancy Cushen White
Department of Pediatrics
University of California San Francisco
San Francisco, USA
nancycushen.white@ucsf.edu

ABSTRACT

More than 10% of the population has dyslexia, and most are diagnosed only after they fail in school. This work seeks to change this through early detection via machine learning models that predict dyslexia by observing how people interact with a linguistic computer-based game. We designed items of the game taking into account *(i)* the empirical linguistic analysis of the errors that people with dyslexia make, and *(ii)* specific cognitive skills related to dyslexia: *Language Skills, Working Memory, Executive Functions,* and *Perceptual Processes.* . Using measures derived from the game, we conducted an experiment with 267 children and adults in order to train a statistical model that predicts readers with and without dyslexia using measures derived from the game. The model was trained and evaluated in a 10-fold cross experiment, reaching 84.62% accuracy using the most informative features.

CCS CONCEPTS

• **Computers and Society**; • **Social Issues**; • **Assistive Technologies for persons with disabilities**;

KEYWORDS

Dyslexia, screening, early detection, diagnosis, linguistics, serious games, machine learning

DH'18, April 23–26, 2018, Lyon, France

ACM ISBN 978-1-4503-6493-5/18/04…$15.00
https://doi.org/10.1145/3194658.3194675

ACM Reference Format:
Luz Rello, Enrique Romero, Maria Rauschenberger, Abdullah Ali, Kristin Williams, Jeffrey P. Bigham, and Nancy Cushen White. 2018. Screening Dyslexia for English Using HCI Measures and Machine Learning. In *DH'18: 2018 International Digital Health Conference, April 23–26, 2018, Lyon, France.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3194658.3194675

1 INTRODUCTION

More than 10% of the population has dyslexia [11, 26]. The *DSM-V* [1] defines dyslexia as a *specific learning disorder* with a neurological basis. According to the *World Federation of Neurology* it occurs in children who, despite conventional classroom experience, fail to attain the language skills of reading, writing, and spelling commensurate with their intellectual abilities [28]. In summary, dyslexia is frequent, universal and related to school failure. However, it remains under-diagnosed. For instance, in the UK, a country that effectively treats dyslexia as compared with other countries, only 5% of the individuals with dyslexia are diagnosed and given appropriate help, and [7]. It is estimated that over 85% of adult illiterates have dyslexia [7].

Yet, early detection is crucial for addressing dyslexia and effective remediation. Often, students are under-diagnosed because current procedures for diagnosis are expensive [16, 21] and require professional oversight [3, 8]. Our goal is for anyone to know as early as possible if they might have dyslexia in an inexpensive way.

To achieve this goal, we have created a computer game that records a wide variety of web-page interaction measures to screen dyslexia for English. We conducted a user study with 267 participants to collect data to train a machine learning model that is able to correctly determine if a person has (or does not have) dyslexia with 84.62% accuracy.

2 BACKGROUND AND RELATED WORK

The complexity of administering paper-based diagnostic tools, and the time they require, have led educators to turn towards computer based screening methods to derive a quick assessment.

2.1 Commercial Software

Among the available commercial software to detect dyslexia in English there is *Lexercise Screener* [13] and *Nessy* [17]. We could not find studies behind these commercial applications, although they are widely used in practice. To our knowledge, they are not based on a machine learning model predictive of dyslexia, and we have not found publication of their accuracy.

2.2 Computer Based Games

There are a number of computer games designed to screen for dyslexia, but they do not use machine learning models.

Lyytinen *et al.* [15] created the computer game *Literate*, later called *GraphoGame* [14], to identify children at risk in Finland. The game was tested with 12 and 41 children between 6 and 7 years old -with statistically significant differences.

There are three other on-going projects for early risk detection of dyslexia that have not yet reported significant results yet: one approach for Italian tested with 24 pre-schoolers [9], and a language-independent approach *MusVis* evaluated with German, English and Spanish children [19, 24].

2.3 Machine Learning Approaches

Machine learning approaches to predict dyslexia are more recent. In 2015, the first method to screen dyslexia in Spanish was introduced; it used eye-tracking measures from 97 subjects (48 with dyslexia) [21]. Later in 2016, eye-tracking measures were also used to predict dyslexia for Swedish (185 subjects, 97 of then with high-risk of dyslexia)[2]. Both methods used *Support Vector Machines*. Another study detected dyslexia subtypes in the Hebrew language using data derived from existing medical records [12].

The only approach we are aware of to predict risk of dyslexia using features derived from computer-based measures is the game *Dytective* for Spanish. The screener, *Dytective*, was first evaluated with 343 people (95 with diagnosed dyslexia) and attained 83% accuracy in a held-out test set with 100 participants using Support Vector Machines [23]. Later, the model was improved by applying a neural network model (Long Short-Term Memory Networks (LSTMs) [10]) to a larger dataset–4,335 participants (763 with professional dyslexia diagnosis)–attaining 91.97% accuracy [22]. This model was integrated into a free online tool *Dytective* which has been used over 100,000 times.[1] An earlier study piloted *Dytective's* screening measures with 60 English speaking children and found the feature set was promising, but the study did not fully incorporate machine learning methods [25].

We advance these approaches by *(i)* extending these methods to the English language and *(ii)* include a wider number of items targeting cognitive indicators predictive of dyslexia.

[1]https://dytectivetest.org/

3 USER STUDY

We conducted a within-subject study (267 participants) with all participants exposed to the same linguistic items integrated into an online game *Dytective*.

3.1 Procedure and Ethics Statement

Participants completed the experiment remotely, through a computer at home, school, or in a specialized center in the USA (mainly from the states of Pennsylvania, New York and Texas). All participants agreed to participate through an online consent form, and children provided assent along with their parent or legal guardian following protocols approved by our institutional review board (IRB). Parents/legal guardians were specifically warned that they could not help their children complete the study exercises. When schools and specialized centers oversaw participation, parental/legal guardian consent was obtained in advance, and the study was supervised by the school counselor or therapist.

The first part of the study consisted of a questionnaire collecting demographic data. This questionnaire was completed by the participant's supervisor (school counselor or therapist) in cases when the participant was under 18 years of age. Then, following oral instructions, participants were given 20 minutes to complete the test exercises.

3.2 Participants

We recruited 267 participants from one specialized center, three schools, and from individuals with dyslexia who knew about our study through our public call online.

Subjects ranged in age from 7 to 60 years old. We classified these participants into three groups. Of the participants, 52 were diagnosed with dyslexia -Class *D (dyslexia)*- (28 female, 24 male, $M = 11.16, SD = 6.31$) and 206 without a diagnosis of dyslexia served as a control group -Class *N (Not-Dyslexia)*- (94 female, 112 male, $M = 11.89, SD = 5.11$). There were 9 participants at risk of having dyslexia or suspected of having dyslexia -Class *M (Maybe)*- (4 female, 5 male, $M = 17.66, SD = 16.17$).[2]

The first language of all participants was English, although 84 participants spoke another language (mostly Spanish in the Texas area). A total of 224 participants reported having trouble with language classes at school.

3.3 Dependent Measures

Participants' performance was measured using the following *dependent measures* for each of the exercises: (i) Number of *Clicks* per item; (ii) *Hits* (*i.e.*, the number of correct answers); (iii) *Misses* (*i.e.*, the number of incorrect answers); (iv) *Score* (*i.e.*, the sum of correct answers for each stage's problem type); (v) *Accuracy* (*i.e.*, the number of *Hits* divided by the number of *Clicks*; and (vi) *Miss Rate* (*i.e.*, the number of *Misses* divided by the number of *Clicks*).

We later used these performance measures together with the demographic data as features of our prediction model's dataset (see Section 4).

[2]All were either adults or children under observation by professionals, the step before having an official diagnosis.

Language Skills	Working Memory
Alphabetic Awareness	Visual (alphabetical)
Phonological Awareness	Auditory (phonology)
Syllabic Awareness	Sequential (auditory)
Lexical Awareness	Sequential (visual)
Morphological Awareness	**Executive Functions**
Syntactic Awareness	Activation and Attention
Semantic Awareness	Sustained Attention
Orthographic Awareness	Simultaneous Attention
Perceptual Processes	
Visual Discrimination and Categorization	
Auditory Discrimination and Categorization	

Table 1: Indicators used for the design of the test items.

3.4 Materials

We integrated test items into a software game to serve as the primary material of our study.

3.4.1 Design and Implementation. *Dytective* is a cross-platform web-based game built in *HTML5*, *CSS*, *JavaScript* and a *PHP* server and a *MySQL* database. It was designed with a high level of abstraction to make it easily portable for future native implementations.

The interface design of the game implements the guidelines that, according to the latest findings in accessibility research, ensure the best on-screen text readability for this target group. Text is presented in black using a mono-spaced typeface *Courier* and a minimum font size of 14 points [20].

3.4.2 Playing Dytective. At each phase, the player's goal is to accumulate points by solving a linguistic problem type as many times as possible in a 25-second time window. For example, the player hears the target, non-word *crench* and then a board is shown on screen containing the target non-word as well as distractors that are particularly difficult for people with dyslexia to differentiate (See Figure 1 (a)). After each time window, the player continues on to the next item corresponding to a new linguistic problem type.

3.4.3 Content Design. The test items are composed of a set of attention and linguistic exercises addressing three or more of the following indicators belonging to different types of *Language Skills, Working Memory, Executive Function*, and *Perceptual Processes.* These indicators are related to dyslexia [5, 6, 27].

The exercises were designed according to linguistic knowledge and the expertise of dyslexia therapists (specific to the English language). In addition, to assist item selection (exercises) we used the following criteria:

(i) linguistic analyses of 833 confusion sets[3], created from the errors of people with dyslexia writing in English [18]; and
(ii) Performance measures from the linguistic exercises of an online game called *Piruletras.* [4] This game is part of previous work targeting children with dyslexia to improve spelling performance [24]. We selected exercises that were more challenging for the players (those with higher error rates and need for more time to be solved) since those exercises were more likely to manifest dyslexia difficulties.

[3]A confusion set is a small group of words that are likely to be confused with one another–such as *weather* and *whether.*
[4]https://itunes.apple.com/us/app/dyseggxia/id534986729?mt=8

4 DATASET

The dataset is composed of 226 features per participant (i.e., total of 60,342 data points. Each participant from the dataset was marked as *D* if the participant has dyslexia, *N* if not, and *M* (maybe) if the participant suspects that he or she has dyslexia but is not diagnosed. From the dataset we extracted the following features:

1 **Gender** of the participant. A binary feature with two values, *female* and *male.*
2 **Age** of the participant ranging from 7 to 60 years old.
3 **Second language**. A binary feature with two values, *no* and *yes*, when the participant had a second language in case of bilingualism.
4 **Language subject**. This is a binary feature with two values, *no* and *yes*, when the participant declares that she has trouble with language classes at school.

Features from **5** to **226** are **performance measures**; they correspond with the six dependent measures (*Clicks, Hits, Misses, Score, Accuracy,* and *Missrate*) per level played (37 levels).

These features target some of the skills presented in Table 1. Note that all the exercises involve attention, so all these features target the *executive functions* **activation and attention**, and **sustained attention**. In addition, some of them also target **simultaneous attention** when the participant pays attention to a number of sources of incoming information at the same time.

5-28 These features are performance measures related to **alphabetic awareness** and **visual discrimination and categorization**. For these tasks the participant hears the name of a letter, *e.g., d*, and identifies it from among the distractors (orthographic and phonetically similar letters, *e.g. b, q, p*) within a time frame, using a Whac-A-Mole-style game interaction.

29-52 These features relate to **phonological awareness** and **auditory discrimination and categorization**. The participant listens to the sound (phoneme) of a letter and identifies it from among distractors. For example, the participant hears the phoneme /n/ and then a board is shown containing the target <n> as well as distractors. We use distractors that are particularly difficult for people with dyslexia to differentiate (*i.e.*, other phonemes that share phonetic features, such as. nasal and sound consonants.

53-88 These features target **syllabic awareness** and **auditory discrimination and categorization**. The players hear the pronunciation of a syllable (*e.g., /prin/*) and identify its spelling from among orthographic distractors *<pren> <prein>, <prain>, <prean>*, and *<pryn>*.

89-112 These features correspond to a set of exercises where participants identify a word's spelling after hearing its pronunciation *(e.g., /greet/* by discriminating among phonetically and orthographically similar words and/or non-words *(e.g., <create>, <greate>, <great>, <grete>, <greit>, <creet>, <crete>, <creat>*. These features target **lexical awareness**, **auditory working memory**, and **auditory discrimination and categorization**.

113-136 These performance features correspond to exercises targeting **visual discrimination and categorization**, by requiring participants to find as many different letters as possible

Figure 1: Screenshots of the exercises requiring the player to click on the target non-word listed among the distractors; (a) select the different letter; (b) build a correct word by substituting a letter, (c) selecting a letter, (d) or deleting a letter (e).

within a time frame in a visual search task (*e.g. E/F, g/q, c/o, b/d* or *p/q*). See Figure 1, exercise (b).

137-160 These features were extracted from a set of exercises requiring players to listen to a non-word and choose its spelling (*e.g. /lurled/*) from among distractors (*e.g. <rurled>, <larled>, <lurded>, <lurleb>, <lorled>*). These features target **sequential auditory working memory**, and **auditory discrimination and categorization**. See Figure 1, exercise (a).

161-172 These performance features target **lexical**, **phonological**, and **orthographic awareness**; They are derived from exercises requiring participants to supply a missing letter [161-166] or delete an extra letter in a target word [167-172]. See Figure 1, exercises (d) and (e), respectively.

173-178 These performance features target **morphological** and **semantic awareness**. They are collected from exercises requiring participants to find a morphological error in a sentence when there is also a semantic error. For example, in the sentence, *The affect of the wind was to cause the boat's sails to billow.* (The word *affect* should be *effect*).

179-184 These features relate to **syntactic awareness**. Participants find an error in a sentence related to a grammatical or function word that changes, (*e.g., of* instead of *on* in *"Smoking is prohibited of the entire aircraft"*).

185-190 This set of features relates to **phonological**, **lexical**, and **orthographic awareness**. These exercises require to find an error in a sentence and correct it by choosing a letter from a set of distractors. See Figure 1, exercise (c).

191-202 This set of features -**phonological, lexical** and **orthographic awareness** (**Features 191-196**)-require participants to rearrange letters to spell a real word (*e.g., b e c u a s e*) or to rearrange syllables to spell a real word (*e.g., /na/ /na/ /ba*) -**syllabic, lexical,** and **ortho- graphic awareness** (Features **197-202**).

203-208 This set of features, addressing **phonological, lexical** and **orthographic awareness** requires players to separate words to make a meaningful sentence, *e.g.* Change *sheranupthehill* to *e.g. she ran up the hill.*

209-214 This set of features targets **sequential visual working memory** and **visual discrimination and categorization** since they are gathered from exercises where players see a sequence of letters for 3 seconds and then write the sequence discriminating targets from distractors.

215-226 This set of features **(215-220)** targets **lexical** and **orthographic awareness** and requires participants to listen and write a word *(e.g., /make/)* or targets **sequential auditory working memory** and **phonological awareness (221-226)** and requires participants to listen and write a non-word *(e.g., /smay/).*

	Score
Accuracy	84.62%
Precision – Class *D (Dyslexia)*	63.76%
Recall – Class *D (Dyslexia)*	80.24%
Precision – *Class N (Not-Dyslexia)*	93.88%
Recall – Class *N (Not-Dyslexia)*	85.83%

Table 2: Classifier accuracy in the cross validation experiment, using the optimized feature set.

5 RESULTS AND DISCUSSION

To determine whether it is feasible to detect whether a user may have dyslexia, we set up a machine learning experiment. We carried out an experiment with a binary classifier of LIBSVM [4] in the Gaussian Support Vector Machine (SVM) setup. An SVM is a method for supervised machine learning that analyzes data and finds patterns for classification. As other Machine Learning algorithms, given a set of training examples, each marked as belonging to a category, an SVM training algorithm builds a model that assigns new examples into the categories. The particular bias of SVMs is that of constructing a hyperplane (either in the original space or in a transformed one) for the classification output. This hyperplane is constructed by combininig the original input examples with the aim of maximizing the functional margin. Our SVM is trained on the dataset as the one described in Section 4.

We performed a 10-fold cross validation experiment by dividing the data into 10 different roughly equal subsets (10% of the data in each subset). Then, we trained a statistical model on the rest of the data (90%) and tested on the corresponding fold by iterating 10 times; at the end, all data was tested independently. We used 10-fold cross validation because it is normally recommended for smaller datasets when a single train-development test split might not be informative enough.

We randomized the data and used stratified sampling to ensure a similar distribution of data categories in all folds. Participants marked as *M (Maybe)* were assigned to the class *D (Dyslexia)*. Outliers' values in the number of *Clicks* and *Misses* were limited to a maximum fixed value. Subsequently, the data were scaled to zero mean and unit variance.

We analyzed the data for features whose distributions were different between dyslexic and non-dyslexic participants. To that end, a *Kolmogorov-Smirnov* test was performed. The number of *Hits* and *Misses* showed different distributions for a number of exercises.

Table 2 shows the accuracy of the SVM model (Gaussian kernel). This result suggests that the model is able to predict players with dyslexia quite accurately with a final result of 80.24% by using a subset of informative features. Note that the baseline (the percentage of subjects assigned to the class Dyslexia in the data set) is 22.85%.

The most informative features were a set of 10 features composed of *Hits* and *Misses*, *Misses* being the most informative ones at an individual level. These features are performance measures belonging to exercises that target **Alphabetic Awareness**, **Phonological Awareness**, **Visual Discrimination and Categorization** and **Auditory Discrimination and Categorization**. More concretely, these features come from exercises where the participant was required to map (or associate) a letter name or a letter sound with a grapheme (letter or letters). This is consistent with previous literature on dyslexia that focus on the deficit on the phonological component in dyslexia [26, 27].

6 CONCLUSIONS AND FUTURE WORK

We presented a method to screen for risk of dyslexia among English speakers that combines machine learning and web-based interaction data collected from a linguistic game. The method was evaluated with 267 participants and attained 84.62% accuracy on its prediction. These results build on earlier findings from the first version of *Dytective* [23], where only Spanish was considered.

These results should be taken as preliminary, since the model was trained on a small dataset. Further experiments with more participants under other less controlled conditions are needed. Our next step will be to conduct a large-scale study. With positive results, we will integrate the model in a tool to screen risk of dyslexia online. Since estimations of dyslexia are much higher than the actual diagnosed population, we believe this method has potential to make a significant impact.

ACKNOWLEDGMENTS

This paper was developed under a grant from the US Department of Education, NIDRR grant number H133A130057; and a grant from the National Science Foundation (#IIS-1618784).

We thank the *Valley Speech Language and Learning Center in Brownsville* in Texas, and the schools *Winchester Thurston School* (Pittsburgh, PA), and *Ellis School* (Pittsburgh, PA). . We thank the volunteers who participated (Adam Brownold, Elsa Cárdenas-Hagan, Anne Fay, Susan Freudenberg) in supervising the participants. Thanks to Lola Álvarez, Susanne Burger, Debbie Meyer, and Regina Rash for their help with recruiting participants

REFERENCES

[1] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders, (DSM-V).* American Psychiatric Publishing, Arlington, VA.

[2] Mattias Nilsson Benfatto, Gustaf Öqvist Seimyr, Jan Ygge, Tony Pansell, Agneta Rydberg, and Christer Jacobson. 2016. Screening for Dyslexia Using Eye Tracking during Reading. *PloS one* 11, 12 (2016), e0165508.

[3] Elizabeth Carrow-Woolfolk. 1995. *OWLS, Oral and Written Language Scales.* NCS Pearson Incorporated.

[4] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Issue 3. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[5] F. Cuetos and F. Valle. 1988. Modelos de lectura y dislexias (Reading models and dyslexias). *Infancia y Aprendizaje (Infancy and Learning)* 44 (1988), 3–19.

[6] Robert Davies, Javier Rodríguez-Ferreiro, Paz Suárez, and Fernando Cuetos. 2013. Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. *Reading and Writing* 26, 5 (2013), 721–738.

[7] Dyslexia Research Institute. 2015. Dyslexia, Identification. http://www.dyslexia-add.org/issues.html. (January 2015).

[8] Angela Fawcett and Rod Nicolson. 2004. *The Dyslexia Screening Test: Junior (DST-J).* Harcourt Assessment.

[9] Ombretta Gaggi, Giorgia Galiazzo, Claudio Palazzi, Andrea Facoetti, and Sandro Franceschini. 2012. A serious game for predicting the risk of developmental dyslexia in pre-readers children. In *Proc. ICCCN'12.* IEEE, 1–5.

[10] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9, 8 (1997), 1735–1780.

[11] Interagency Commission on Learning Disabilities. 1987. *Learning Disabilities: A Report to the U.S. Congress.* Government Printing Office, Washington DC.

[12] Yair Lakretz and Michal Rosen-zvi. 2015. Probabilistic Graphical Models of Dyslexia. In *Proc. SIGKDD'15.* 1919–1928.

[13] Lexercise. 2016. Dyslexia Test - Online from Lexercise. http://www.lexercise.com/tests/dyslexia-test. (2016). [Online; accessed 18-September-2017].

[14] Heikki Lyytinen, Jane Erskine, Janne Kujala, Emma Ojanen, and Ulla Richardson. 2009. In search of a science-based application: A learning tool for reading acquisition. *Scandinavian journal of psychology* 50, 6 (2009), 668–675.

[15] Heikki Lyytinen, Miia Ronimus, Anne Alanko, Anna-Maija Poikkeus, and Maria Taanila. 2007. Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology* 59, 2 (2007), 109.

[16] Dennis L Molfese. 2000. Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and language* 72, 3 (2000), 238–245.

[17] Nessy. 2011. Dyslexia Screening - Nessy UK. https://www.nessy.com/uk/product/dyslexia-screening/. (2011). [Online; accessed 18-September-2017].

[18] J. Pedler. 2007. *Computer Correction of Real-word Spelling Errors in Dyslexic Text.* Ph.D. Dissertation. Birkbeck College, London University.

[19] Maria Rauschenberger, Luz Rello, Ricardo Baeza-Yates, Emilia Gomez, and Jeffrey P. Bigham. 2017. Towards the Prediction of Dyslexia by a Web-based Game with Musical Elements. In *W4A'17.* 4–7. https://doi.org/10.1145/3058555.3058565

[20] L. Rello and R. Baeza-Yates. 2013. Good Fonts for Dyslexia. In *Proc. ASSETS'13.* ACM Press, Bellevue, Washington, USA.

[21] L. Rello and M. Ballesteros. 2015. Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures. In *Proc. W4A '15.* ACM, Florence, Italy.

[22] L. Rello and M. Ballesteros. 2017. Data Processing System to Detect Neurodevelopmental-Specific Learning Disorders. *United States Patent and Trademark Office* filed in on April 20, 2017, as application 15/493,060 (2017).

[23] L. Rello, M. Ballesteros, A. Ali, M. Serra, D. Alarcón, and J. P. Bigham. 2016. Dytective: Diagnosing Risk of Dyslexia with a Game. In *Proc. Pervasive Health'16.* Cancun, Mexico.

[24] Luz Rello, Enrique Romero, Maria Rauschenberger, Abdullah Ali, Kristin Williams, Jeffrey P. Bigham, and Nancy Cushen White. 2018. Towards language independent detection of dyslexia with a web-based game. In *W4A'18.* Lyon, France. https://doi.org/10.1145/3192714.3192816

[25] Luz Rello, Kristin Williams, Abdullah Ali, Nancy Cushen White, and Jeffrey P. Bigham. 2016. Dytective: Towards Detecting Dyslexia Across LanguagesUsing an Online Game. In *Proc. W4A'16.* ACM Press, Montreal, Canada.

[26] S.E. Shaywitz, M.D. Escobar, B.A. Shaywitz, J.M. Fletcher, and R. Makuch. 1992. Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine* 326, 3 (1992), 145–150.

[27] F. R. Vellutino, J. M. Fletcher, M. J. Snowling, and D. M. Scanlon. 2004. Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry* 45, 1 (2004), 2–40.

[28] World Federation of Neurology (WFN). 1968. Report of research group on dyslexia and world illiteracy. Dallas: WFN. (1968).