A Machine Learning Approach for Dyslexia Screening in a Minoritized Language Context: The Case of Catalan

Enrique Romero
Universitat Politecnica de Catalunya
Barcelona, Spain
eromero@cs.upc.edu

Joaquim Llisterri
Universitat Autonoma de Barcelona
Barcelona, Spain
Joaquim.Llisterri@uab.cat

Miquel Casas[*]
Research Institute of Sant Joan de Deu Hospital
Barcelona, Spain
miquel.casas@sjd.es

Rosa Bosch[†]
Althaia Xarxa Assistencial Universitaria de Manresa
Manresa, Spain
rmbosch@althaia.cat

Marta Ortega Bravo[‡]
Institut Catala de la Salut
Lleida, Spain
mortega.lleida.ics@gencat.cat

Lidia Sanz Borrell
Institut Catala de la Salut
Lleida, Spain
lsanz.lleida.ics@gencat.cat

Ramon Capdevila Bert[§]
University Institute Foundation for Primary Care Research Jordi Gol i Gurina
Lleida, Spain
rcapdevila.lleida.ics@gencat.cat

Mireia Biosca Pamies[§]
University Institute Foundation for Primary Care Research Jordi Gol i Gurina
Lleida, Spain
mbiosca.lleida.ics@gencat.cat

Gemma Terrer Manrique
Institut Catala de la Salut
Lleida, Spain
gterrer.lleida.ics@gencat.cat

Luz Rello
IE University
Madrid, Madrid, Spain
luzrello@gmail.com

Abstract

Learning to read is a fundamental skill for academic success and a key tool for enabling individuals to fully participate in society. However, approximately 10% of children face difficulties in acquiring this skill due to dyslexia, a neurodevelopmental disorder that affects reading and writing acquisition. Developing dyslexia detection methods is particularly challenging in minoritized languages, where the smaller number of speakers makes it difficult to gather the large datasets typically required to train machine learning models. In this work, we present an approach for screening dyslexia in Catalan using a gamified test that combines linguistic exercises with machine learning techniques. To achieve this, we designed the content of a computer game, collected data from 730 children —155 of whom were diagnosed with dyslexia— who played the game, and developed a prediction model using various machine learning classifiers along with targeted feature selection. Our method achieved the highest balanced accuracy when using a Single-Layer Perceptron (SLP) classifier (87.46%) and a linear Support Vector Machine (SVM) classifier (86.67%), both applied to a selected subset of features. These results highlight the potential for cost-effective, online early screening of dyslexia in children who speak minoritized languages, especially in contexts where collecting large datasets is not feasible. The results have been integrated into a freely available online application.[1]

CCS Concepts

• **Human-centered computing → Accessibility technologies**.

Keywords

Dyslexia, Machine Learning, Serious Games, Minoritized Languages

[*]Additional affiliations: Fundació de Investigació Salut I Progrés (FISP), Barcelona, Spain and Department of Psychiatry, Universitat Autònoma de Barcelona, Spain. Part of the project was carried out at the Department of Psychiatry, Vall d'Hebron University Hospital, Barcelona, Spain.

[†]Additional affiliations: SJD MIND Schools Program, Hospital Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain. Biomedical Research Networking Centre for Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain.

[‡]Additional affiliation: University Institute Foundation for Primary Care Research Jordi Gol i Gurina (IDIA-PJGol), Barcelona, Spain.

[§]Additional affiliation: Les Borges Blanques Primary Care Center, Institut Català de la Salut, Lleida, Spain

W4A '26, Dubai, United Arab Emirates

ACM ISBN 979-8-4007-2372-8/26/04
https://doi.org/10.1145/3800424.3800449

[1]Available at www.dytective.org. The overall design of the screener is protected under the United States Patent titled "Data Processing System to Detect Neurodevelopmental-Specific Learning Disorders." The patent, bearing number 11334803, was officially granted on May 17, 2022. The intellectual property rights are assigned to Carnegie Mellon University, located in Pittsburgh, Pennsylvania, USA.

ACM Reference Format:
Enrique Romero, Joaquim Llisterri, Miquel Casas, Rosa Bosch, Marta Ortega Bravo, Lidia Sanz Borrell, Ramon Capdevila Bert, Mireia Biosca Pamies, Gemma Terrer Manrique, and Luz Rello. 2026. A Machine Learning Approach for Dyslexia Screening in a Minoritized Language Context: The Case of Catalan. In *Web4All 2026 23rd International Web for All Conference (W4A '26), April 13–14, 2026, Dubai, United Arab Emirates.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3800424.3800449

1 Introduction

In recent decades, we have witnessed the creation of an increasing number of predictive models for medical screening [52]. All of these methods share a common challenge: they require a substantial amount of data to function effectively. Unfortunately, this limits their applicability in cases where large datasets are not available, such as rare diseases or language disorders in speakers of either minority or minoritized languages.

This work focuses on dyslexia, a universal condition with a neurobiological origin that affects approximately 10% of the population [3]. It is one of the most common learning disorders and, if not detected early and addressed appropriately, can lead to school failure and social exclusion. Dyslexia does not affect general intelligence. However, the U.S. National Center for Learning Disabilities states that learning difficulties are a leading cause of school failure and dropout [11]. For example, the National Center for Education Statistics in the U.S. estimates that between 5 and 8.6 million children in public elementary and secondary schools have dyslexia and are at risk of underperforming academically as a result [55]. A critical step in addressing this issue is the timely detection of dyslexia. Indeed, efforts to detect dyslexia in schools have increased in recent years, as it can only be addressed once it has been identified.

Historically, dyslexia detection methods have been time-consuming and resource-intensive, typically relying on paper-based tests administered under professional supervision [12, 19, 57]. More recently, there has been significant progress in the development of machine learning-based tools aimed at making dyslexia assessment more accessible and affordable [44, 46, 54]. However, these tools often require large datasets for training, which presents a disadvantage for for minority or minoritized groups. Data collection in minority or in minoritized languages is significantly more challenging than in majority languages. Consequently, most machine learning-based dyslexia detection methods have been developed for widely spoken languages with tens or even hundreds of millions of speakers [14, 47, 51].

The main contribution of this work is a method for detecting the risk of dyslexia using machine learning, applied to a small dataset of 750 individuals whose native language is Catalan. To the best of our knowledge, this is the first application of machine learning for screening dyslexia in Catalan.

Our approach demonstrates that, through the careful design of linguistic games based on a specific methodology focused on dyslexia-related error analysis, and combined with targeted feature selection, it is possible to detect the risk of dyslexia even in minoritized language contexts with limited data availability.

The rest of the paper is structured as follows: the next section reviews related work, followed by an explanation of the proposed method (Section 3), the conducted study (Section 4), and the dataset used (Section 5). Finally, we present the results (Section 6), discuss them in Section 7, and conclude in Section 8.

2 Background and Related Work

2.1 Dyslexia Detection

The American Psychiatric Association defines dyslexia as a specific learning disorder that affects approximately 5% to 15% of the global population [4]. According to the International Dyslexia Association, dyslexia is characterized by difficulties with accurate and/or fluent word recognition, as well as poor spelling and decoding abilities. These difficulties typically stem from a deficit in the phonological component of language, which is often independent of other cognitive abilities and not attributable to inadequate classroom instruction [37]. Consequently, a person with dyslexia may experience difficulties with reading and writing regardless of intelligence, native language, socioeconomic status, or education level.

Current methods for screening dyslexia in Catalan require professionals to collect performance measures related to reading and writing through in-person assessments [22, 57]. These evaluations typically include metrics such as reading speed (words per minute), reading and writing errors, word and pseudoword reading, reading fluency, and text comprehension. This traditional approach to dyslexia detection has also historically been used for other languages [12, 19], for which more automated screening methods have gradually been developed.

2.2 Machine Learning Approaches

In recent years, the development of machine learning methods for dyslexia detection has grown significantly, particularly those that leverage various types of data [28]. Initial efforts in this area used eye-tracking data and were applied to languages including [6], Greek [5, 54], Spanish [47], French [16], and Russian [51].

Parallel research has explored the use of brain imaging data to detect dyslexia. For example, functional Magnetic Resonance Imaging (fMRI) data collected during reading tasks have been successfully used for Mandarin Chinese [14], as well as for French, German, and Polish [42]. Similarly, machine learning models trained on electroencephalography (EEG) data—derived from recorded electrical activity in the brain—have been used to detect the risk of dyslexia in Hebrew [21], English [40], Dutch [49], and also in German, French, Hungarian, and Finnish [38].

More recently, researchers have begun to favor data types that are easier and more cost-effective to collect, such as performance metrics derived from computer games. These approaches have shown promising results in languages like English [48], German [43, 44], and Spanish [46], often using games specifically designed to reflect the reading and writing difficulties experienced by children with dyslexia.

Despite dyslexia being the most common language-related learning disorder [3], its detection in minority or in minoritized languages remains a major challenge due to the limited availability of datasets.

2.3 Challenges in Detecting Dyslexia in Catalan

All of these studies share a common characteristic: they have been conducted in populations with a sufficiently large number of speakers or rely on brain imaging data, which does not necessarily require a high number of participants. In contrast, for minoritized languages using data derived from computer games, such studies are scarce due to the challenges involved in collecting datasets large enough to train machine learning models — especially those that include individuals with dyslexia.

This work focuses on dyslexia detection in Catalan, a Romance language fluently spoken by an estimated 9.1 million speakers in Spain [23]. Catalan is not a completely uniform language; due to its historical evolution, broad geographic distribution, and the influence of other languages, it displays significant dialectal diversity, particularly in pronunciation, vocabulary, and some grammatical structures. Therefore, data collection must be conducted from scratch, taking into account these dialectal variations within Spain.

Beyond the limited availability of data, another major challenge lies in detecting dyslexia in languages with transparent orthographies, such as Catalan. In contrast to languages with deep orthographies—like English—where grapheme-to-phoneme correspondence is less consistent and dyslexic difficulties are more apparent, transparent orthographies (e.g., Spanish, Catalan, Finnish) have more regular spelling-to-sound mappings [50]. As a result, individuals with dyslexia in these languages tend to experience milder symptoms, making the condition harder to detect [9, 60]. In fact, dyslexia is often referred to as a "hidden disability" in such contexts, because its manifestations are subtle and often go unnoticed until academic failure occurs—despite the individual having normal intelligence. Unfortunately, by that point, intervention may come too late for optimal impact.

2.4 Originality of our approach

The objective of this paper is to develop a machine learning-based method for detecting dyslexia in Catalan, while addressing three major challenges that go beyond the inherent complexity of the condition itself: (i) the limited availability of data in a minoritized language context, (ii) the linguistic variability of Catalan due to its dialectal diversity, and (iii) the difficulty of detecting dyslexia in a language with transparent orthography.

To the best of our knowledge, this is the first study to apply machine learning methods for screening dyslexia in Catalan. The input data for our approach is derived from interaction metrics collected during an online test consisting of gamified linguistic questions, making the method easy to administer at scale. This work builds upon two previous studies that explored how linguistic games—based on errors commonly made by individuals with dyslexia—can be used to screen for the condition in English [48] and Spanish [46].

3 Method

To create a machine learning model capable of screening the risk of dyslexia and/or literacy difficulties in Catalan, we designed a web-based gamified test and conducted a study with 730 participants. The study followed a within-subject experimental design, allowing us to collect various performance measures that were later used as features in our machine learning model.

3.1 Materials

We integrated the test items into a software game that served as the primary material for our study.

3.1.1 Design and Implementation. The gamified test is a cross-platform, web-based game built using HTML5, CSS, JavaScript, a PHP server, and a MySQL database. It was developed with a high level of abstraction to facilitate future adaptation to native applications. The game's interface follows accessibility design guidelines based on recent research to ensure optimal on-screen text readability for users with dyslexia. Text is presented in black using the monospaced typeface Courier, with a minimum font size of 14 points [7, 45]. The 26 test items and the web application interface are publicly available[2] and included in the Annexes.

3.1.2 Playing the Application. In each phase of the game, the player's objective is to accumulate points by solving as many instances as possible of a particular linguistic problem within a 25-second time window. For example, the player hears the target non-word *silidi* and is shown a board with the target and several distractors—specifically designed to challenge individuals with dyslexia (see Figure 1, left). After each round, the player proceeds to the next item, which corresponds to a different linguistic challenge.

3.1.3 Content Design. The test items consist of a series of exercises designed to assess various indicators associated with dyslexia. These items were jointly created by a phonetician with expert knowledge of Catalan and a specialist in dyslexia and reading disorders. All exercises were reviewed by pediatric professionals who volunteered in the study.

The creation of the exercises followed three main steps or criteria: (i) an empirical analysis of a corpus of errors made by children with literacy difficulties to extract linguistic patterns commonly associated with dyslexia; (ii) the selection of relevant linguistic features and their integration into the design of the exercises; and (iii) the annotation of each exercise with the specific cognitive skill it targets.

First, the mistakes made by children with dyslexia reflect specific patterns associated with the disorder [1, 56]. Thus, the exercises were grounded in an empirical analysis of a Catalan error corpus compiled from texts written by individuals diagnosed with, or suspected of having, dyslexia. We collected a total of 2,386 errors from children in the cities of Lleida and Barcelona, so that the two main geographical varieties of Catalan are represented: Western (Lleida) and Eastern (Barcelona).

Each error was manually annotated with the following linguistic information: correct spelling, wrong spelling, type of error – following Pedler's classification of dyslexic errors [39]– type of grapheme, tonicity, position of the error in the word, syllabic division, number of syllables, lexical category of the word (part of the speech), type of error source (written or spoken mistake), age of the person who made the error (ranging from 7 to 17), and the geographical variant of Catalan (Easter or Western).

[2] https://playdyslexia.org/cat//

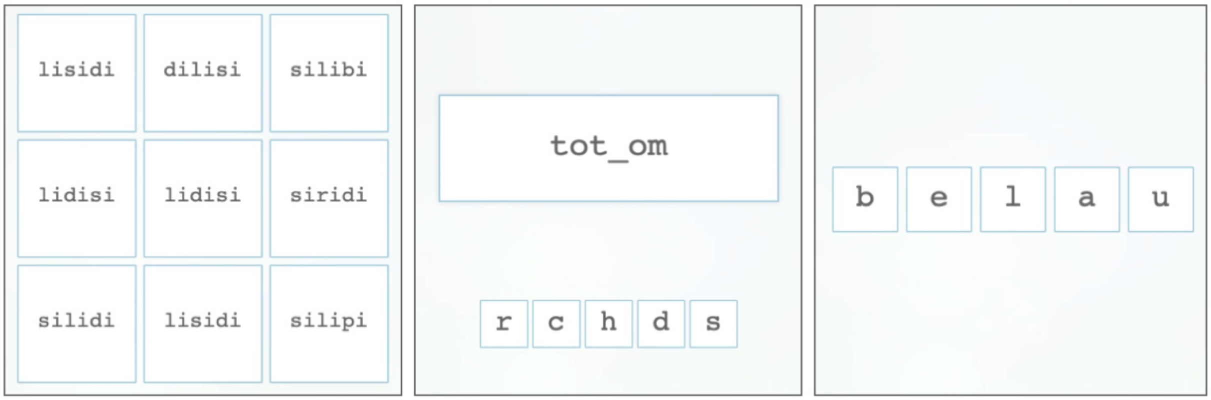

Figure 1: Examples of four test questions: find the non-word *silidi* (left); build a correct word *tothom* ('everybody'), center; and delete a letter to write a correct word *blau* ('blue'), right.

From these annotations, we extracted statistical patterns that guided the design of the test items. For instance, some exercises asked participants to supply a missing letter or remove an extra letter in a target word, based on common insertion or omission errors observed in the corpus. See Figure 1 (center and right) for examples such as "build a correct word" (*tothom*,'everybody' and "delete a letter to write a correct word" (*blau*, 'blue').

Second, the error patterns were integrated into real words. To select the words used in the exercises, we employed NIM *A Web-based Swiss army knife to select stimuli for psycholinguistic studies* developed by a team from the Department of Psychology at Rovira i Virgili University and described in detail by Guasch, Boada, Ferré, and Sánchez-Casas [26]. This tool enabled us to identify words based on criteria such as length, matching substrings, lexical frequency, and part of speech.

Finally, the exercises were categorized based on the different psycholinguistic factors involved in children's cognitive processing [13, 15, 60]. Children with dyslexia exhibit difficulties not only in linguistic areas such as *Linguistic Awareness*, but also in *Working Memory, Executive Functions*, and *Perceptual Processes.* Therefore, each exercise is associated with one or more cognitive indicators, as shown in Table 1, and linked to specific manifestations of dyslexia, also detailed in the same table.

Table 1: Cognitive factors associated with dyslexia used in the creation of test exercises. In parentheses, we include the psycholinguistic studies that show the relationship between those factors and the manifestations of dyslexia.

Linguistic Awareness	
Alphabetic Awareness	[20]
Phonological Awareness	[29]
Syllabic Awareness	[35]
Lexical Awareness	[18]
Morphological Awareness	[25]
Syntactic Awareness	[58]
Semantic Awareness	[27]
Orthographic Awareness	[24]
Working Memory	
Visual (Alphabetical)	[53]
Auditory (Phonology)	[34]
Sequential (Auditory)	[32]
Sequential (Visual)	[41]
Executive Functions	
Activation and Attention	[36]
Sustained Attention	[2]
Simultaneous Attention	[33]
Perceptual Processes	
Visual Discrimination and Categorization	[30]
Auditory Discrimination and Categorization	[31]

4 Experimental Study

4.1 Design

We used a within-subject design, ensuring that all participants contributed equally to all variables. The study involved a total of 730 children, including 155 who were either diagnosed with or suspected of having dyslexia. Each participant played our game for approximately 15 minutes.

4.2 Participants

We conducted experiments with participants aged 7 to 12, along with their supervisors: parents, legal guardians, teachers, or therapists

4.2.1 Recruitment. Participants were recruited in collaboration with hospitals and civil associations from both the Eastern (Lleida) and Western (Barcelona) regions, in order to ensure a representative sample of the two main dialects of Catalan.

In terms of hospital involvement, recruitment was carried out with the support of pediatricians, researchers, and psychiatrists from the following institutions: the Institut Universitari d'Investigació

Table 2: Descriptive data of the participants.

	n	$\overline{age}$	*SD*	f.	m.
Participants without Dyslexia	575	11,34	3,02	289	286
Participants with Dyslexia	155	10,50	2,72	66	89
All	**730**	**11,17**	**3,85**	**355**	**375**

Table 3: Characteristics of the three data sets used in the study according to the age range of the participants (A1, A2 and A3).

Data set	N	Ave. Age	Dyslexia	Female	Male
A1 (Ages 7-8)	115	7.29	27.83% (32)	48.70%	51.30%
A2 (Ages 9-11)	261	10.11	21.07% (55)	50.57%	49.43%
A3 (Ages 12-17)	354	13.55	19.21% (68)	47.18%	52.82%
All	**730**	**11.17**	**26.96%**	**48.17%**	**50.88%**

en Atenció Primària (IDIAP Jordi Gol) *(University Institute for Research in Primary Care, IDIAP Jordi Gol)* at Lleida Hospital, and the Hospital Universitari Vall d'Hebron *(Vall d'Hebron University Hospital)* in Barcelona. The civil associations that collaborated in the study were the Associació Catalana de la Dislèxia *(Catalan Association of Dyslexia)* and the Associació de Dislèxia Lleida *(Lleida Dyslexia Association)*. Participation in the research was completely voluntary.

4.2.2 *Inclusion Criteria.*

The inclusion criteria required a formal dyslexia diagnosis provided by a registered professional. However, participants with a strong suspicion of dyslexia but without a formal diagnosis were also included and grouped with the dyslexic participants for analysis. These cases were explicitly labeled in the dataset to distinguish suspected from confirmed diagnoses.

Participants without dyslexia were recruited through schools and were limited to children and adolescents with no history of language-related difficulties in their academic records. Determining accurate ground truth in dyslexia diagnosis is inherently challenging, as many individuals go undiagnosed, and the true accuracy of professional diagnoses is not always verifiable.

A total of 730 participants were recruited (see Table 2). The group without reading difficulties consisted of 575 individuals (50.26% female, 49.74% male), with ages ranging from 7 to 17 (M = 11.34, SD = 2.72). The group of participants with dyslexia or suspected dyslexia included 155 individuals (42.58% female, 57.42% male), aged 7 to 17 (M = 10.50, SD = 3.85). Of these, 39 participants did not have a confirmed diagnosis of dyslexia, but there was a strong suspicion that they had the condition.

4.2.3 *Bilingualism.*

Determining the native language or mother tongue (L1) of a participant who speaks only one language is straightforward. However, this is not the case for bilingual individuals, who might report speaking two L1s when replying to questionnaires aimed at assessing their linguistic profile. Since both Catalan and Spanish are spoken in Catalonia, in our study we explicitly asked about the languages spoken by the participants and about the language which they consider to be their L1.

All the participants knew Catalan and Spanish, being Catalan the first or dominant language of 626 participants (85,75%) and the second or non-dominant language of the remaining 104 participants (14,25%). In any case, all games and the tool's oral instructions were provided in Catalan.

4.3 Dependent Measures

To quantify task performance, we collected the following dependent measures for each exercise: (i) number of *Clicks*; (ii) number of correct answers *(Hits)*; (iii) number of incorrect answers *(Misses)*; (iv) *Score*, defined as the sum of Hits per set of exercises; (v) *Accuracy*, defined as the number of Hits divided by the number of Clicks; and (vi) *Missrate*, defined as the number of Misses divided by the number of Clicks.

These performance measures were later used, along with demographic data, as features in our machine learning models.

4.4 Compliance and Ethics Statements

The prototypes used in this research comply with the European *General Data Protection Regulation* (GDPR) regarding the processing and protection of personal data [17]. Personal information of the participants' supervisors—such as names or email addresses—is not published and is stored separately from the participants' data, solely for the purpose of communicating results, if consent is given. The name of the child is not collected, and all data are stored on a password-protected web server located in the Netherlands.

4.5 Procedure

After informed consent and parental consent were obtained, participants were given specific instructions to take part in the study. They completed the gamified test for 15 minutes without interruption, as each item in the test had a fixed time limit. All participants used a desktop computer and headphones to complete the task. Participants were supervised while playing the game by an adult supervisor, who was not allowed to assist them in completing the tasks. The supervisors were pediatricians, researchers, and psychiatrists who volunteered for the study. All were trained to administer the study protocol. Finally, the supervisor completed a demographic questionnaire, which included the date of the participant's dyslexia diagnosis (if applicable). All participants and supervisors volunteered to take part in the study.

5 Data sets

Our datasets were derived from the experimental study described in the previous section. In total, we collected 159 features per subject across 730 participants, resulting in 116,070 data points. The data are organized into three separate datasets according to age group.

5.1 Age groups

Each age group completed a customized version of the test based on their language level, resulting in three different datasets (see Table 3). Participants aged 12–17 completed all 26 questions, corresponding to 156 performance features in the dataset. Participants aged 9–11 were not exposed to the items in Exercises 21 and 25, as these were deemed too difficult for their age group. As a result, their dataset contains 144 performance features. Finally, children aged

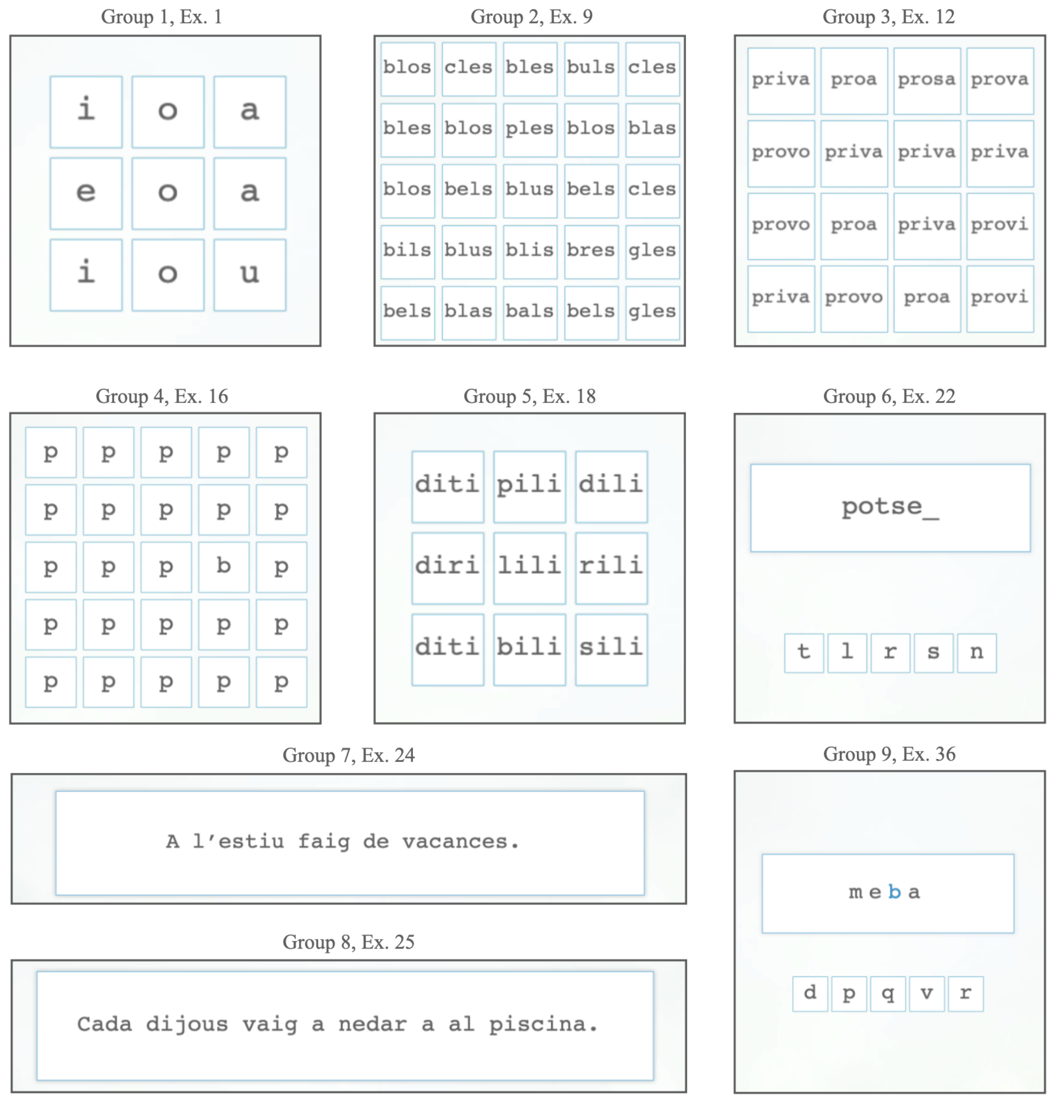

Figure 2: This table presents a representative example of an exercise (Ex.) from each distinct group of items included in the test. The test comprises a total of 26 exercises, organized into groups. In Exercise 1, participants must select the specified letter *e*, and in Exercise 9, the syllable *blis*. In Exercise 12, they are required to identify the word *prova* 'exam'. In Exercise 16, they must identify the letter that differs from the rest. In Exercise 18, the task is to select the non-word *dili*. Exercise 22 requires participants to insert a missing letter to form the correct word *potser* 'maybe'. Additionally, in Exercise 24, they must identify the syntactic error in a list of sentences, while in Exercise 25, they are expected to detect semantic errors. Finally, in Exercise 36, participants must substitute one letter to form a correct word. Each task must be repeated within a fixed time window, during which several words are presented sequentially. In the case presented in the figure, the target word is *meva* 'my'.

6–7 completed only the items appropriate for their developmental level, excluding the more advanced exercises (Exercises 13, 18, 19, 20, 21, 24, 25, and 26). Therefore, this dataset (A1) consists of 108 performance features[3].

[3]The complete dataset used in this study can be downloaded from https://www.cs.upc.edu/~eromero/Downloads/Dyslexia-Catalan-Data.tgz.

5.2 Features

To build our ground truth, we extracted the following features from the dataset, labeled as "Yes" if the participant had a confirmed diagnosis of dyslexia, "No" if not, and "Suspicion" if the participant is suspected of having dyslexia but has not been diagnosed.

Table 4 presents a comprehensive overview of all the features included in our dataset. Features from 1 to 3 correspond to demographic features, while features from 4 to 159 to the performance

Table 4: Description of the demographic and the performance features. For performance-related features, we have included the main cognitive measures that the exercises target.

Demographic features	**Feature**	**Description**
Age	Not used	It ranges from 7 to 17 years old, used to create the different data sets
Gender	Feature: 1 gender	A binary feature with two values, *female* or *male*.
Native language	Feature: 2 mother_language	The first language of the participants, a binary feature with two values: *1* if their first language was Catalan and *0* if they were bilingual, being one the languages Catalan.
Language subject	Feature: 3 language_subject	This is a binary feature with two values: *1* when the participant had failed a language subject at school at least once and *0* when the participant had never fail that subject.
Performance features		
Cognitive abilities	**Feature & Exercise number**	**Description**
Alphabetic Awareness & Sustained Attention	Features: 4-27 corresponding to Exercises 1-4 (Group 1) (i.e., clicks_1, hits_1, misses_1, [...] clicks_2, hits_2, etc.)	These performance features correspond to questions from 1 to 4 (Ex. 1-4). In these tasks the participant hears the name of a letter (*e, p, d,* and *q*) and maps it with the letter among distractors (orthographic and phonetically similar letters) within a time frame, using a Whac-A-Mole game interaction (see Figure 2, Ex. 1). These questions address prerequisites in reading acquisition: *Alphabetic Awareness*, *Phonological Awareness* and *Visual Discrimination and Categorization*.
Syllabic Awareness & Sustained Attention	Features: 28-57 corresponding to Exercises 5-9 (Group 2)	Features targeting *Phonological Awareness*, *Syllabic Awareness* and *Auditory Discrimination and Categorization*. Here the players hear the pronunciation of syllable (*ta*, *des*, *pro*, *cle* or *bles* corresponding to Ex. 5, Ex. 6, Ex. 7, Ex. 8 and Ex. 9, respectively) and map it with its spelling, e.g., *bles* where the distractors are *blas blis blos blus bals bels bils bols buls ples cles gles* and *bres* (see Figure 2, Ex. 9).
Lexical Awareness, Working Memory & Sustained Attention	Features: 58-81 correspond to Exercises 10-13 (Group 3)	These performance features corresponding to a set of exercises (Ex. 10-13) where participants map the pronunciation of a word with its spelling (e.g., *boda*) discriminating among other phonetically and orthographically similar words and/or pseudo-words (e.g., *boba, boca, boga, bola, bota, baba, beba, deba, tuba, buba, suba, loba* or *coba*). These features aim at *Lexical Awareness*, *Auditory Working Memory*, and *Auditory Discrimination and Categorization*, (see Figure 2, Group 3).
Visual Discrimination & Sustained Attention	Features: 82-105 corresponding to Exercises 14-17 (Group 4)	These performance features (Ex. 14-17) target mainly *Visual Discrimination and Categorization*, and *Executive Functions* since the players undertake a visual search task, finding as many as possible different letters within a time frame, e.g., *n/u, e/a, p/b, q/d, u/n, a/e, b/p, d/q, ç/c, l·l/l*, among others. See Figure 2, Ex. 16.
Working Memory, Auditory Discrimination & Sustained Attention	Features: 106-129 corresponding to Exercises 18-21 (Group 5)	Features extracted from a set of exercises (Ex. 18-21) where players listen to a pseudo-word and choose its spelling (e.g., *dili*) among, e.g., *lili, disi, sili, tili, diti, pili, diri, rili, bili* and *sili*, shown in Figure 2, Ex. 18. These features target *Visual Working Memory*, *Sequential Auditory Working Memory*, and *Auditory Discrimination and Categorization*.
Phonological, Orthographic Awareness & Sustained Attention	Features: 130-141 corresponding to Exercises 22-23 (Group 6)	These target mainly *Lexical*, *Phonological*, and *Orthographic Awareness*; extracted from exercises where participants are required to fill in the missing letter in different words that appear within a set time frame, i.e., *potse_* for *potser* 'perhaps' and, i.e., *tot_om* for *tothom* 'everyone' (Ex. 22, features 131-136, or delete the extra letter in the word (Ex. 23, features 137-142), i.e., *balau* for *blau* 'blue', see Figure 1, center and right.
Morphological, Semantic Awareness & Simultaneous Attention	Features: 142-147 corresponding to Exercise 24 (Group 7)	These performance features (Ex. 24) mainly target *Morphological* and *Semantic Awareness*. They are collected from exercises were participants find a morphological error in a sentence -which gives as a result a semantic error-. For instance, in the sentence *A l'estiu *faig de vacances*, 'In the summer I *make on holiday', where the word *faig* 'to make' should be *vaig* 'to go'. See in Figure 2, Ex. 24.
Syntactic Awareness & Simultaneous Attention	Features: 148-153 corresponding to Exercise 25 (Group 8)	Features related (Ex. 25) to *Syntactic Awareness*. Similarly to the previous set of exercises, participants need to find and error in a sentence, being this error in a grammatical or functional word, so the syntactic meaning of the sentence changes. For example, in Figure 2 Ex. 25 *al* 'in' instead of *la* 'the' in *Cada dijous vaig a nedar a *al piscina*, 'Every Thursday I go swimming *to the pool'.
Phonological, Lexical, Orthographic Awareness & Simultaneous Attention	Features: 154-159 corresponding to Exercise 26 (Group 9)	A set of features (Ex. 26, see Figure 2) related to *Phonological*, *Lexical* and, *Orthographic Awareness* since they are extracted from exercises where participant have to find an error in a word, i.e., **meba*, and correct it –*meva* 'my'– choosing a letter from a set of distractors *d p q,* and *v*.

features, derived from their interaction while playing the 26 exercises of the test (6 measures per question presented previously, that is, *Clicks, Hits, Misses, Score, Accuracy,* and *Missrate*). For example, playing Exercise 1 results in six performance features: clicks_1, hits_1, misses_1, score_1, accuracy_1, and missrate_1.

The 26 exercises were grouped according to the cognitive skills they target and the type of user interaction they involve (see nine groups in Figure 2). Specific interaction details for each exercise and item are provided in the description column of Table 4. Additionally, Table 4 indicates the specific dyslexia-related cognitive skills each exercise is designed to assess: *Linguistic Awareness*, *Working Memory*, and *Perceptual Processes*. All exercises also target *Executive Functions*, as they require attentional control. Specifically, all items address *Activation and Attention* and *Sustained Attention*. In addition, some exercises (e.g., Exercises 24–26) also target *Simultaneous Attention*, which involves attending to multiple sources of incoming information simultaneously—such as word recognition, distractor discrimination, and error detection.

6 Results of the Machine Learning Experiments

6.1 Data Preprocessing

As previously mentioned, data contains three different types of features: categorical (Gender, Native language, Language subject), unbounded numeric (Clicks, Hits, Misses) and bounded numeric features in $[0, 1]$ (Accuracy, Missrate).

Different preprocessing procedures were applied to the data. First, we removed 4 examples that contained too large values in too many features. Second, and only for unbounded numeric features, we considered that values above or below a certain threshold gave no additional information. Therefore, values larger/smaller than the mean plus/minus 3 times the standard deviation were set to this value. Third, we imputed missing values of numeric features with the mean. Finally, data were normalized as follows: one-hot encoding was applied to categorical features; unbounded numeric features were linearly scaled to $[-1, 1]$; bounded numeric features were linearly scaled to $[0, 1]$.

6.2 Feature Selection

A Feature Selection (FS) procedure was applied in order to obtain the most informative features for the problem at hand and reduce the final complexity of the models. This FS procedure was performed independently for each age group. Since the initial number of features is too large, a standard wrapper approach would be a very time-consuming procedure. Therefore, FS was performed in two steps.

In the first step, fast FS procedures were applied in order to select small but promising subsets of features with low computational cost. All experiments in this first step were performed with 5 repetitions of a 5-fold cross-validation. Since, as previously mentioned, dyslexia is the least frequent class in the dataset, it is prone to be misrepresented if class imbalance is not addressed. Therefore, different strategies were applied to mitigate this effect and try to accurately represent the dyslexia class. The FS procedures applied were the following:

(1) A filter FS, using as saliency the Mutual Information (MI) of every feature with the class label. Regarding the class imbalance, we artificially balanced every fold by randomly reducing the number of examples of the majority class to be equal to those in the dyslexia class. Different parameters were tested to estimate the respective distributions. The 30 features with the largest mean MI values were selected.
(2) An embedded FS backward selection with Single-layer Perceptrons (SLP), using as saliency of every feature the derivative of the output function with respect to the input (see [59] for details). Different networks were trained varying a number of parameters (optimization algorithm, learning rate, momentum, etc). Two strategies were used to obtain more balanced models:
   (a) Weighting the errors of the examples in the dyslexia class with values larger or equal than those of the majority class. This strategy can be seen as an artificial way to increase the number of examples with dyslexia, leading to a larger recall in this class.
   (b) Weighting the errors of the predictions belonging to the dyslexia class with values larger than or equal to those predictions in the majority class. In this way, the errors of the dyslexia class predictions are considered more important than those of the non-dyslexia class, leading to a larger precision in the former.

   After training, the models maximizing the mean validation F1 in the dyslexia class for each of the weighting strategies were selected. The 25 most salient features for every selected model were extracted and joined in a single subset. Remarkably, most of the features were shared among all selected models.

In the second step, a wrapper backward elimination FS procedure was applied with different classifiers and starting from the subsets obtained in the first step. The classifiers were selected so as to be simple and (to some extent) interpretable:

(1) Single-layer Perceptrons (SLP) with different learning parameters, in the same conditions as reported in the embedded FS procedure
(2) Support Vector Machines (SVM) [10] with a polynomial kernel of degree 3, changing the values of the γ and C parameters
(3) Support Vector Machines (SVM) [10] with a linear kernel, changing the values of the γ and C parameters

All experiments in this second step were performed with 20 repetitions of a 5-4-fold double cross-validation. The saliency of a feature subset was the mean validation F1 in the dyslexia class. Table 5 shows the results of the best models for every age group. As it can be observed, linear models performed slightly better than non-linear ones, both in terms of balanced accuracy and F1 in the dyslexia class, which are considered the most important measures for this problem. Therefore, our final selection consisted of two linear models for every age group: an SLP and an SVM with linear kernel (SVML), which have comparable performance.

Finally, we analyzed the selected features for each age group in the final models, shown in Table 6. We observed that there is a subset of features that were consistently selected by both models:

(1) For group 1: language_subject, clicks_1, clicks_7, misses_8, missrate_8, clicks_9, hits_11, hits_12, misses_14, clicks_16 and misses_22

Table 5: Results of the selected models for every group of age. C1 and C2 stand for class 1 (participants with dyslexia) and class 2 (participants without dyslexia), respectively. The metrics shown are: Balanced Accuracy (BalAcc), Accuracy in class 1 (AccC1), Accuracy in class 2 (AccC2), Total Accuracy (TotAcc), Precision in class 1 (PreC1), Precision in class 2 (PreC1), F1 in class 1 (F1C1) and F1 in class 2 (F1C2).

Group	Model	BalAcc	AccC1	AccC2	TotAcc	PreC1	PreC2	F1C1	F1C2
Group 1	SLP	79.39	70.66	88.12	83.09	70.66	88.12	70.66	88.12
	SVM-Poly3	76.83	65.12	88.54	81.78	69.72	86.24	67.34	87.38
	SVM-Linear	79.48	71.33	87.63	82.94	70.02	88.30	70.67	87.96
Group 2	SLP	87.46	83.89	91.03	89.52	71.39	95.49	77.14	93.20
	SVM-Poly3	86.62	82.91	90.34	88.77	69.62	95.19	75.68	92.70
	SVM-Linear	86.67	81.14	92.21	89.88	73.55	94.82	77.16	93.50
Group 3	SLP	78.99	68.90	89.08	85.21	60.01	92.33	64.15	90.68
	SVM-Poly3	77.62	64.43	90.82	85.75	62.52	91.48	63.46	91.15
	SVM-Linear	79.90	73.12	86.67	84.07	56.61	93.13	63.82	89.79

(2) For group 2: language_subject, clicks_1, hits_2, hits_8, hits_12, misses_20, misses_23 and hits_26
(3) For group 3: hits_1, missrate_9, missrate_12, hits_16, hits_24, misses_24 and hits_25

Remarkably, these features subsets also have similarities between groups 1 and 2 (language_subject, clicks_1 and hits_12), However, there are significant differences between Group 3 and the others, reflecting developmental differences supported by linguistic theory, as explained in the Discussion section.

7 Discussion

7.1 Results Across Groups

This study developed a machine learning-based method to screen dyslexia risk in Catalan-speaking children using a gamified online test. Our models showed strong predictive performance, particularly for Group 2 (ages 9–11), achieving up to 87.46% balanced accuracy with an SLP classifier.

The superior performance observed in the intermediate group (ages 9 to 11) may be attributed to its relatively higher number of participants within a narrower age range. In contrast, the group of 7- to 8-year-olds had fewer participants, many of whom are young and still acquiring foundational reading and writing skills, likely resulting in greater variability. Group 3, although the largest in absolute terms, spans a broader age range (12–17), which may introduce increased heterogeneity. Therefore, the greater homogeneity in Group 2 may explain its stronger performance, while the higher variability in Groups 1 and 3 could have contributed to their comparatively lower results.

These findings are notable given the challenges of limited data and transparent orthography in a minoritized language. Feature selection identified a core set of interaction and demographic features that consistently contributed to model performance across different age groups (Table 6). Note that all exercises were selected at least once indicating that each exercise contributed informative value to the models.

7.2 Features Repeated Across Groups

The analysis revealed that despite working with distinct age cohorts (7–8 years, 9–11 years, and 12–17 years), certain features were recurrently selected as highly informative.

The fact that the language_subject feature appears predominantly in Groups 1 and 2 is consistent with expectations, as this attribute reflects whether a participant has failed a language subject. Particularly at younger ages, a common manifestation of dyslexia is low academic performance, often resulting in the failure of language-related subjects. Therefore, it is logical that for the younger cohorts, this demographic feature would be highly predictive.

Additionally, two other features that consistently appear across all groups are those derived from Exercise 1 and Exercise 12 (clicks_1, hits_12, hits_1, missrate_12). This can be explained as follows: in Exercise 1, participants are required to click on a letter (*e*, see Figure 2, Ex. 1), assessing letter discrimination, a fundamental pre-reading skill and an area commonly impaired in individuals with dyslexia. Given that this is the very first exercise, participants are likely to be more attentive, which could amplify the performance differences between individuals with and without dyslexia. On the other hand, Exercise 12 focuses on lexical discrimination. In this task, participants must click on the correctly spelled word *prova* among distractors (See Figure 2, Ex. 12). This exercise demands extra cognitive effort from participants with dyslexia, as the target word includes a complex syllabic structure, that is, the consonant cluster *pr* at the onset of the syllable *pro*. Consonant clusters are known to be a source of difficulty for individuals with dyslexia across all age groups [8]. Nevertheless, there are more similarities between Groups 1 and 2, indicating that Group 3 differs more from the others, with exercises focused on linguistic skills that develop at later stages.

7.3 Cognitive skills across groups

For the youngest group (7–8 years), the most informative features were linked to alphabetic and syllabic awareness, as well as basic phonological processing and visual discrimination, with early academic struggles (language subject failures) serving as strong predictors (Exercices 1, 7, 8, 9, 11, 12, 14, 16 and 22).

Table 6: Selected features of the final models for every group of age (see the text for details).

	Group 1		Group 2		Group 3	
Feature	SLP	SVML	SLP	SVML	SLP	SVML
language_subject	X	X	X	X		
mother_language					X	
clicks_1	X	X	X	X		
hits_1					X	X
clicks_2			X			
hits_2			X	X		
missrate_2				X		
clicks_4					X	
hits_4						X
missrate_4						X
clicks_7	X	X		X		
hits_7			X			
hits_8			X	X		
misses_8	X	X	X		X	
missrate_8	X	X			X	
clicks_9	X	X				
missrate_9					X	X
clicks_11					X	
hits_11	X	X				
hits_12	X	X	X	X		
missrate_12					X	X
accuracy_14	X					
misses_14	X	X				
missrate_15					X	
clicks_16	X	X				
hits_16					X	X
misses_17	X					
clicks_18						X
clicks_19				X		
misses_20			X	X		
accuracy_22					X	
clicks_22		X				
hits_22		X				X
misses_22	X	X				
misses_23		X	X	X		
hits_24					X	X
misses_24					X	X
hits_25					X	X
clicks_26			X			
hits_26			X	X		

In the intermediate group (9–11 years), similar linguistic features remained important, but working memory, sequential auditory discrimination, and orthographic awareness began to play a greater role, reflecting developmental progress in literacy skills (Exercices 1, 2, 8, 9, 11, 12, 20, 23 and 26). Note that language subject failures is still relevant for this age group.

For the oldest group (12–17 years), the selected features shifted toward higher-order skills. Most informative features came from Exercices 1, 9, 12, 16, 24, and 25. These features are associated with visual discrimination, executive functioning, and more complex morphological, syntactic and semantic awareness tasks. Older students with dyslexia may have developed compensatory strategies for basic phonological tasks, but difficulties persist in tasks requiring sustained attention, fast and accurate visual processing, and complex linguistic manipulation. Interestingly, the 'language subject' feature was less important for this group, possibly because many students with dyslexia have already adapted to school demands, making it harder to identify them based on academic performance alone at this age.

7.4 General Insights

Across all age groups, a pattern emerged: early phonological and visual processing difficulties (alphabetic and lexical awareness) are the strongest indicators in younger children, while more complex language processing features (morphological, semantic, syntactic and orthographic awareness) become more important in older participants (Exercises 23-26). This developmental shift aligns with known models of dyslexia, which highlight how reading difficulties manifest differently over time [60]. Furthermore, exercises targeting simultaneous attention (Exercises 24–26) gained importance in older groups, reinforcing the link between executive functioning and academic reading skills.

Finally, our results demonstrate that even with small datasets, the careful design of test items and thoughtful feature selection can yield robust screening models. This opens promising avenues for affordable and scalable dyslexia screening in educational settings, particularly for languages where large-scale diagnostic data are not readily available.

8 Conclusions and Future Work

In this work, we have presented a novel approach for early screening of dyslexia in Catalan-speaking children by combining linguistic games with machine learning techniques. Our method demonstrates that it is possible to achieve high accuracy (up to 86.67% balanced accuracy) even with a relatively small dataset, addressing key challenges related to minoritized languages and transparent orthographies. By designing a gamified, accessible, and scalable online test, we contribute a low-cost and efficient alternative to traditional dyslexia screening methods that require extensive human resources and time.

The results highlight the relevance of features related to linguistic awareness, working memory, and perceptual processes in predicting dyslexia risk. Moreover, our feature selection analysis revealed consistent indicators across different age groups, suggesting that specific types of linguistic exercises are critical for early identification regardless of developmental stage.

As future work, we plan to extend our methodology to other minoritized languages with similar orthographic transparency challenges, such as Basque (an isolated language) and Galician (a Romance language), which have official status in the territories in which they are spoken in Spain. In addition, our goal is to incorporate longitudinal data to evaluate the predictive power of our models over time. Further research could also explore the use of deep learning models or hybrid systems that combine behavioral and neurocognitive data to improve screening accuracy. Finally, user-centered studies focusing on improving the engagement and adaptivity of the gamified test will be crucial to optimize its practical implementation in educational and clinical settings.

Acknowledgments

We would like to express our sincere gratitude to Dr. Josep Lluís Cruz Cubells, now retired, who was both a physician and the president of the Dyslexia Association of Lleida, and who originally proposed the idea of conducting this research in Catalonia.

We also wish to thank the civil associations — Associació Catalana de Dislèxia, Associació de Dislèxia de Lleida, Asociación de Familias con Dificultades de Aprendizaje (AFDA) en Catalunya — for their invaluable collaboration in helping us collect data from both children with and without dyslexia.

References

[1] Olivia Afonso, Paz Suárez-Coalla, and Fernando Cuetos. 2015. Spelling impairments in Spanish dyslexic adults. *Frontiers in Psychology* 6, 466 (2015).

[2] Tracy Packiam Alloway, Samantha Wootan, and Peter Deane. 2014. Investigating working memory and sustained attention in dyslexic adults. *International Journal of Educational Research* 67 (2014), 11–17.

[3] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5.* American Psychiatric Association, Washington DC.

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

[5] Thomais Asvestopoulou, Victoria Manousaki, Antonis Psistakis, Erjona Nikolli, Vassilios Andreadakis, Ioannis M Aslanides, Yannis Pantazis, Ioannis Smyrnakis, and Maria Papadopouli. 2019. Towards a robust and accurate screening tool for dyslexia with data augmentation using GANs. In *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 775–782.

[6] 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), e0165508 (2016).

[7] British Dyslexia Association. 2012. Dyslexia Style Guide. http://www.bdadyslexia.org.uk/.

[8] Maggie Bruck and Rebecca Treiman. 1990. Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Child Psychology* 50, 1 (1990), 156–178.

[9] Nicola Brunswick. 2010. Unimpaired reading development and dyslexia across different languages. In *Reading and dyslexia in different orthographies*, Sine McDougall and Paul de Mornay Davies (Eds.). Psychology Press, Hove, 131–154.

[10] C. Cortes and V. N. Vapnik. 1995. Support-Vector Networks. *Machine Learning* 20, 3 (1995), 273–297.

[11] Candace Cortiella and Sheldon H. Horowitz. 2014. *The state of learning disabilities: Facts, trends and emerging issues.* New York. 2–45 pages.

[12] Fernando Cuetos, Blanca Rodríguez, Elvira Ruano, and David Arribas. 2014. *PROLEC-R. Batería de evaluación de los procesos lectores, revisada (PROLEC-R: Battery for the Assessment of Reading Processes, Revised).* TEA, Madrid.

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

[14] Zaixu Cui, Zhichao Xia, Mengmeng Su, Hua Shu, and Gaolang Gong. 2016. Disrupted white matter connectivity underlying developmental dyslexia: A machine learning approach. *Human Brain Mapping* 37, 4 (2016), 1443–1458.

[15] 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.

[16] Alae Eddine El Hmimdi, Lindsey M Ward, Themis Palpanas, and Zoï Kapoula. 2021. Predicting dyslexia and reading speed in adolescents from eye movements in reading and non-reading tasks: A machine learning approach. *Brain Sciences* 11, 10 (2021), 1337.

[17] European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679

[18] Kelly Farquharson, Tracy M Centanni, Chelsea E Franzluebbers, and Tiffany P Hogan. 2014. Phonological and lexical influences on phonological awareness in children with specific language impairment and dyslexia. *Frontiers in Psychology* 5 (2014), 838.

[19] Angela J. Fawcett and Rod L. Nicolson. 2011. *Test para la detección de la dislexia en niños (DST-J) (DST-J: Dyslexia Screening Test – Junior).* TEA, Madrid.

[20] Gorka Fraga González, Gojko Žarić, Jurgen Tijms, Milene Bonte, Leo Blomert, and Maurits W van der Molen. 2015. A randomized controlled trial on the beneficial effects of training letter-speech sound integration on reading fluency in children with dyslexia. *PLOS one* 10(12), e0143914 (2015).

[21] Alex Frid and Larry M Manevitz. 2018. Features and machine learning for correlating and classifying between brain areas and dyslexia. *arXiv:1812.10622* (2018).

[22] Amaia García, José Ríos Guillermo, Núria Torrent, Estela López, and Virginia Trémols. 2022. PRODISCAT Pediàtric: un qüestionari útil per detectar la dislèxia en pediatria (PRODISCAT Pediatric: a useful questionnaire for detecting dyslexia in pediatrics). *Pediatria catalana: butlletí de la Societat Catalana de Pediatria* 82, 3 (2022), 99–105.

[23] Generalitat de Catalunya. 2025. Quants parlants té el català dins i fora de Catalunya? Situació del català: preguntes més freqüents. Retrieved July 23 2025 https://llengua.gencat.cat/ca/05_atencio_ciutadania/pmf-preguntes-mes-frequents/?category=pmf-catala.

[24] George K Georgiou, Dalia Martinez, Ana Paula Alves Vieira, and Kan Guo. 2021. Is orthographic knowledge a strength or a weakness in individuals with dyslexia? Evidence from a meta-analysis. *Annals of Dyslexia* 71, 1 (2021), 5–27.

[25] George K Georgiou, Ana Paula Alves Vieira, Kyriakoula M Rothou, John R Kirby, Andrea Antoniuk, Dalia Martinez, and Kan Guo. 2023. A meta-analysis of morphological awareness deficits in developmental dyslexia. *Scientific Studies of Reading* 27, 3 (2023), 253–271.

[26] Marc Guasch, Roger Boada, Pilar Ferré, and Rosa Sánchez-Casas. 2013. NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods* 45, 3 (2013), 765–771.

[27] K Jednoróg, A Marchewka, P Tacikowski, and A Grabowska. 2010. Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia* 48, 9 (2010), 2447–2457.

[28] Shahriar Kaisar. 2020. Developmental dyslexia detection using machine learning techniques: A survey. *ICT Express* 6, 3 (2020), 181–184.

[29] Ioulia Kovelman, Elizabeth S Norton, Joanna A Christodoulou, Nadine Gaab, Daniel A Lieberman, Christina Triantafyllou, Maryanne Wolf, Susan Whitfield-Gabrieli, and John DE Gabrieli. 2012. Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. *Cerebral Cortex* 22, 4 (2012), 754–764.

[30] Arni Kristjansson and Heida Maria Sigurdardottir. 2023. The role of visual factors in dyslexia. *Journal of Cognition* 6, 1 (2023).

[31] Marie Lallier, Sophie Donnadieu, and Sylviane Valdois. 2013. Investigating the role of visual and auditory search in reading and developmental dyslexia. *Frontiers in Human Neuroscience* 7, 587 (2013).

[32] Marie Lallier, Guillaume Thierry, and Marie-Josèphe Tainturier. 2013. On the importance of considering individual profiles when investigating the role of auditory sequential deficits in developmental dyslexia. *Cognition* 126, 1 (2013), 121–127.

[33] Marie Lallier and Sylviane Valdois. 2012. *Dyslexia - A comprehensive and international approach.* IntechOpen, London, Chapter Sequential versus simultaneous processing deficits in developmental dyslexia, 73–108.

[34] Deny Menghini, Alessandra Finzi, Giovanni Augusto Carlesimo, and Stefano Vicari. 2011. Working memory impairment in children with developmental dyslexia: Is it just a phonological deficity? *Developmental neuropsychology* 36, 2 (2011), 199–213.

[35] Carla Míguez-Álvarez, Miguel Cuevas-Alonso, and Ángeles Saavedra. 2022. Relationships between phonological awareness and reading in Spanish: A meta-analysis. *Language Learning* 72, 1 (2022), 113–157.

[36] Elisabeth Moores, Roderick Nicolson, and Angela Fawcett. 2003. Attention deficits in dyslexia: Evidence for an automatisation deficit? *European Journal of Cognitive Psychology* 15, 3 (2003), 321–348.

[37] Orton Dyslexia Society Research Committee. 1994. Definition of Dyslexia.

[38] Shankar Parmar and Chirag Paunwala. 2023. A novel and efficient Wavelet Scattering Transform approach for primitive-stage dyslexia-detection using electroencephalogram signals. *Healthcare Analytics* 3, 100194 (2023).

[39] Jennifer Pedler. 2001. Computer spellcheckers and dyslexics - a performance survey. *British Journal of Educational Technology* 32, 1 (2001), 23–37.

[40] Perera Perera, Harshani Harshani, Mohd Fairuz Shiratuddin, Kok Wai Wong, and Kelly Fullarton. 2018. EEG signal analysis of writing and typing between adults with dyslexia and normal controls. 5, 1 (2018), 62–67.

[41] Beate Peter, Andria Albert, and Shelley Gray. 2021. Spelling errors reveal underlying sequential and spatial processing deficits in adults with dyslexia. *Clinical Linguistics & Phonetics* 35, 4 (2021), 310–339.

[42] Piotr Płoński, Wojciech Gradkowski, Irene Altarelli, Karla Monzalvo, Muna van Ermingen-Marbach, Marion Grande, Stefan Heim, Artur Marchewka, Piotr Bogorodzki, Franck Ramus, et al. 2017. Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia. *Human Brain Mapping* 38, 2 (2017), 900–908.

[43] Maria Rauschenberger, Ricardo Baeza-Yates, and Luz Rello. 2022. A universal screening tool for dyslexia by a web-game and machine learning. *Frontiers in Computer Science* 3, 628634 (2022).

[44] Maria Rauschenberger, Ricardo Baeza-Yates, and Luz Rello. 2025. Screening Dyslexia Using Visual Auditory Computer Games and Machine Learning. *IEEE Access* 13 (2025), 29541–29553.

[45] Luz Rello and Ricardo Baeza-Yates. 2017. How to present more readable text for people with dyslexia. *Universal Access in the Information Society* 16 (2017), 29–49.
[46] Luz Rello, Ricardo Baeza-Yates, Abdullah Ali, Jeffrey P Bigham, and Miquel Serra. 2020. Predicting risk of dyslexia with an online gamified test. *PLOS one* 15(12), e0241687 (2020).
[47] Luz Rello and Miguel Ballesteros. 2015. Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures. In *Proc. W4A '15*. Florence, Italy, 121–128.
[48] 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 *Proceedings of the 2018 International Conference on Digital Health*. 80–84.
[49] Z Rezvani, M Zare, G Žarić, M Bonte, J Tijms, MW Van der Molen, and G Fraga González. 2019. Machine learning classification of dyslexic children based on EEG local network features. *BioRxiv* 569996 (2019).
[50] P. H. K. Seymour, M. Aro, and J. M. Erskine. 2003. Foundation literacy acquisition in European orthographies. *British Journal of Psychology* 94, 2 (2003), 143–174.
[51] Soroosh Shalileh, Dmitry Ignatov, Anastasiya Lopukhina, and Olga Dragoy. 2023. Identifying dyslexia in school pupils from eye movement and demographic data using artificial intelligence. *PLOS one* 18(11), e0292047 (2023).
[52] Mohammad Shehab, Laith Abualigah, Qusai Shambour, Muhannad A Abu-Hashem, Mohd Khaled Yousef Shambour, Ahmed Izzat Alsalibi, and Amir H Gandomi. 2022. Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine* 145, 105458 (2022).
[53] James H Smith-Spark and John E Fisk. 2007. Working memory functioning in developmental dyslexia. *Memory* 15, 1 (2007), 34–56.
[54] Ioannis Smyrnakis, Vassilios Andreadakis, Vassilios Selimis, Michail Kalaitzakis, Theodora Bachourou, Georgios Kaloutsakis, George D Kymionis, Stelios Smirnakis, and Ioannis M Aslanides. 2017. RADAR: A novel fast-screening method for reading difficulties with special focus on dyslexia. *PLOS one* 12(8), e0182597 (2017).
[55] Thomas D Snyder, Cristobal de Brey, and Sally A Dillow. 2018. *Digest of Education Statistics 2016, NCES 2017-094*. ERIC, Washington, DC.
[56] Paz Suárez-Coalla, Sara Ramos, Marta Álvarez-Cañizo, and Fernando Cuetos. 2014. Orthographic learning in dyslexic Spanish children. *Annals of dyslexia* 64, 2 (2014), 166–181.
[57] J. Toro and M. Cervera. 1984. *TALE: Test de Análisis de Lectoescritura (TALE: Literacy Analysis Test)*. Visor, Madrid.
[58] Marie Van Reybroeck. 2020. Grammatical spelling and written syntactic awareness in children with and without dyslexia. *Frontiers in Psychology* 11 (2020), 1524.
[59] A. Vellido, E. Romero, M. Julià-Sapé, C. Majós, À. Moreno-Torres, J. Pujol, and C. Arús. 2012. Robust Discrimination of Glioblastomas from Metastatic Brain Tumors on the Basis of Single-voxel Proton MRS. *NMR in Biomedicine* 25, 6 (2012), 819–828.
[60] 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.

A Materials: Items used in the gamified test

Ex.1 Click on the letter *e* (Distractors: *a i o u*; Grid size = 3).
Ex.2 Click on the letter *p* (Distractors: *a b c ç d e f g h i j k l l·l m n o q r s t u w x y z*; Grid size = 4).
Ex.3 Click on the letter *d* (Distractors: *b p q f h k l t* ; Grid size = 5).
Ex.4 Click on the letter *q* (Distractors: *b d p g j y*; Grid size = 6).
Ex.5 Click on the syllable *ta* (Distractors: *te ti to tu at et it ot ut ba ca da ma pa ra sa ac am an ar as at*; Grid size = 5).
Ex.6 Click on the syllable *des* (Distractors: *das dis dos dus bes ces mes pes res ses tes dec dem den der des det sec sem sen ser ses set*; Grid size = 5).
Ex.7 Click on the syllable *pro* (Distractors: *par per pir por pur bro tro dro cro gro por bor tor dor cor gor*; Grid size = 5).
Ex.8 Click on the syllable *cle* (Distractors: *cla cli clo clu cal cel cil col cul ple ble cle gle pel bel tel del cel gel cre cer rec*; Grid size = 5).
Ex.9 Click on the syllable *bles* (Distractors: *blas blis blos blus bals bels bils bols buls ples cles gles bres*; Grid size = 6).
Ex.10 Click on the word *goma* (Distractors: *gota gola gosa poma coma doma*; Grid size = 3).
Ex.11 Click on the word *ballar* (Distractors: *callar tallar fallar ballat billar badar banyar basar*; Grid size = 4).
Ex.12 Click on the word *prova* (Distractors: *provi provo priva prosa proa*; Grid size = 4).
Ex.13 Click on the word *llibreta* (Distractors: *llebreta llibrera llibrets*; Grid size = 4).
Ex.14 Click on the different letter. The target and distractors are taken from the following vowel pairs: *e/a, a/i, a/o, u/a, i/e, o/e, u/e, o/i, u/i, u/o, a/e, a/i, a/o, a/u, e/i, e/o, e/u, i/o, i/u* and *o/u*. Grid size = 5.
Ex.15 Click on the different letter. The target and distractors come from the following pairs of consonants: *c/ç, F/E, l/l·l, q/g, o/c, f/t, ç/c, E/F, l·l/l, g/q, c/o* and *t/f*. Grid size = 5.
Ex.16 Click on the different letter. The target and distractors come from the following mirror letter pairs: *n/u, e/a, p/b, q/d, u/n, a/e, b/p* and *d/q*. Grid size = 5.
Ex.17 Click on the different letter. The target and distractors are taken from the following rotating letter pairs: *d/b, p/d, q/b, q/p, b/d, b/q, d/p* and *p/q*. Grid size = 5.
Ex.18 Click on the non-word *dili* (Distractors: *lidi rili tili lili sili didi diri diti disi pili bili*; Grid size = 3).
Ex.19 Click on the non-word *diriti* (Distractors: *ditiri riditi ritidi tiridi tidiri piriti biriti diridi diribi*; Grid size = 3).
Ex.20 Click on the non-word *silidi* (Distractors: *sidili lisidi lidisi disili dilisi silipi silibi siridi siliri*; Grid size = 3).
Ex.21 Click on the non-word *tirisili* (Distractors: *tisirili tilisiri tisirili ritisili risiliti rilitisi ritilisi sitirili siliriti siriliti siritili litirisi lisitiri liritisi lisiriti dirisili pirisili birisili tilisili*; Grid size = 3).
Ex.22 Insert a letter to form a correct word. The displayed words are:
- *ana_* (target: *r*; distractors: *n d b p*; correct word: *anar* 'to go'),
- *ta_bé* (target: *m*; distractors: *l n s r*; correct word: *també* 'also'),
- *ha_ia* (target: *v*; distractors: *f t m b*; correct word: *havia* 'had'),
- *ve_mell* (target: *r*; distractors: *m s l rr*; correct word: *vermell* 'red'),
- *tot_om* (target: *h*; distractors: *s r d c*; correct word: *tothom* 'everyone'),
- *par_* (target: *t*; distractors: *s ç d b*; correct word: *part* 'part'),
- *_a* (target: *h*; distractors: *d g r x*; correct word: *ha* 'has'),
- *g_an* (target: *r*; distractors: *i l u rr*; correct word: *gran* 'great'),
- *fe_* (target: *r*; distractors: *ç d z n*; correct word: *fer* 'to make'),
- *_as* (target: *h*; distractors: *d s qu i*; correct word: *has* 'you have'),
- *potse_* (target: *r*; distractors: *s l n t*; correct word: *potser* 'perhaps'),
- *aju_tament* (target: *n*; distractors: *m r l h*; correct word: *ajuntament* 'city council'), and

- *carre_* (target: *r*; distractors: *s m l d*; correct word: *carrer* 'street').

Ex.23 Delete a letter to form a correct word. The presented errors are displayed:
- *heines* (correct word *eines*, 'tools'),
- *balau* (correct word *blau*, 'blue'),
- *exsamen* (correct word *examen*, 'exam'),
- *lliegir* (correct word *llegir*, 'to read'),
- *hinvita* (correct word *invita*, 'invites'),
- *gràrcies* (correct word *gràcies*, 'thank you'),
- *vaixsell* (correct word *vaixell*, 'ship'),
- *germàr* (correct word *germà*, 'brother'),
- *fanstasma* (correct word *fantasma*, 'ghost'),
- *pereguntar* (correct word *preguntar*, 'to ask'),
- *himperi* (correct word *imperi*, 'empire'),
- *guiix* (correct word *guix*, 'chalk'),
- *hoblidar* (correct word *oblidar*, 'to forget'),
- *diistributiva* (correct word *distributiva*, 'distributive'), and
- *hintel·ligència* (correct word *intel·ligència*, 'intelligence').

Ex.24 Find the error in the sentence. The following set of sentences with semantic errors is displayed:
- *A l'estiu *faig de vacances.* ('In the summer I *make on holiday'), word *faig* ('to make') should be *vaig* ('to go').
- *Això ho faré *nones si tinc temps.* ('I will do this *nones if I have time.'), word *nones* ('nones/pray') should be *només* ('only').
- *Quan estic malalt no tinc ganes de *menja.* ('When I'm sick, I don't feel like *eat'), word *menja* ('eat') should be *menjar* ('eating').
- *Si estic *dormin, no m'agrada que em despertin.* ('If I'm *sleep, I don't like to be woken up'), word *dormin* ('sleep') should be *dormint* ('sleeping').
- *Vaig perdre la clau i no vaig poder *entra a casa.* ('I lost the key and couldn't *enters the house'), word *entra* ('enters') should be *entrar* ('enter').
- *Quan *vall en cotxe sempre em marejo moltíssim.* ('When I *valley by car, I always get very dizzy'), word *vall* ('valley') should be *vaig* ('go').
- *Tinc ganes de menjar *pasti de xocolata per berenar.* ('I feel like eating chocate *pasti for snack'), word *pasti* ('kneaded') should be *pastís* ('cake').
- *He de llegir un llibre i escriure'n el resum en una *fixa.* ('I have to read a book and write the summary on a *fixed'), word *fixa* ('fixed') should be *fitxa* ('form/sheet').
- *Com que no *savis les respostes, no has aprovat l'examen.* ('Since you didn't *sages the answers, you didn't pass the exam'), word *savis* ('wise man/sages') should be *sabies* ('you know').
- *Haig d'omplir una *figa amb el meu nom i la meva adreça.* ('I have to fill out a *fig with my name and address'), word *figa* ('fig') should be *fitxa* ('form/sheet').

Ex.25 Find the error in the sentence. The following set of sentences with syntactic errors is displayed:
- *Demà sortiré amb *les meus amics.* ('Tomorrow I will go out with *the my friends'), word *les* ('the') should be *els* (form that goes 'with my').
- *Cada dijous vaig a nedar a *al piscina.* ('Every Thursday I go swimming at *at pool'), word *al* ('at') should be *la* ('the').
- *Si plou, agafa *le paraigües o et mullaràs.* ('If it rains, take *the umbrella or you'll get wet'), the error **le* should be *el* ('the').
- **Durat l'última setmana no s'ha trobat bé.* ('*Durat the last week he/she hasn't felt well'), the error **durat* should be *durant* ('during').
- *Visc *la edifici més alt del meu carrer.* ('I live *the tallest building on my street'), word *la* ('the') (feminine) should be *l'* ('the') (masculine).
- *Les llibretes que em vas demanar, no *els tinc.* ('The notebooks you asked me for, I don't have *the'), word *els* ('the') (masculine) should be *les* ('them') (feminine).
- *Aquest llibre *le trobaràs a la biblioteca.* ('This book you will find *the in the library'), the error **le* should be *el* ('it').
- *Per entrar *la metro s'ha de comprar un bitllet.* ('To enter *the metro you need to buy a ticket'), word *la* ('the') should be *al* ('to').
- *M'agraden *les macarrons amb salsa de tomàquet.* ('I like *the macarroni with tomato sauce'), word *les* ('the') (feminine) should be *els* ('the') (masculine).
- *El meu pare treballava moltes hores i no *le veia gaire.* ('My father worked many hours and I didn't see *him much'), the error **le* should be *el* ('him').

Ex.26 Replace a letter to form the correct word. Participants must choose a letter from a given set and substitute it to create the correct word. The following are the words with errors to be corrected by substitution:
- *meba* (target: *v*; distractors: *d p q*; correct word: *meva* 'mine').
- *sentral* (target: *c*; distractors: *z ss ç*; correct word: *central* 'central').
- *abui* (target: *v*; distractors: *d q c*; correct word: *avui* 'today').
- *pun* (target: *nt*; distractors: *nd ns mt*; correct word: *punt* 'was').
- *ere* (target: *a*; distractors: *u s ç*; correct word: *era* 'point').
- *tera* (target: *rr*; distractors: *ss ll l·l*; correct word: *terra* 'earth').
- *masa* (target: *ss*; distractors: *ç rr z*; correct word: *massa* 'too much').
- *coxes* (target: *tx*; distractors: *g tg tj*; correct word: *cotxes* 'cars').
- *trevall* (target: *b*; distractors: *f p d*; correct word: *treball* 'job').
- *qualsebol* (target: *v*; distractors: *d p f*; correct word: *qualsevol* 'anyone').