Towards Language Independent Detection of Dyslexia with a Web-based Game

Maria Rauschenberger
WSSC
Universitat Pompeu Fabra
maria.rauschenberger@upf.edu

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

Ricardo Baeza-Yates
WSSC
Universitat Pompeu Fabra
rbaeza@acm.org

Jeffrey P. Bigham
HCI and LT Institutes
Carnegie Mellon University
jbigham@cs.cmu.edu

ABSTRACT

Detecting dyslexia is important because early intervention is key to avoid the negative effects of dyslexia such as school failure. Most of the current approaches to detect dyslexia require expensive personnel (*i.e.* psychologists) or special hardware (*i.e.* eye trackers or MRI machines). Also, most of the methods can only be used when children are learning how to read but not before, necessarily delaying needed early intervention. In this work, we present a study with 178 participants speaking different languages (Spanish, German, English, and Catalan) with and without dyslexia using a web-based game built with musical and visual elements that are language independent. The study reveals eighth game measures with significant differences for Spanish children with and without dyslexia, which could be used in future work as a basis for language independent detection. A web-based application like this could have a major impact on children all over the world by easily screening them and suggest the help they need.

CCS Concepts

•**Human-centered computing** → **Empirical studies in accessibility; Accessibility design and evaluation methods;** •**Software and its engineering** → *Interactive games;*

Keywords

Dyslexia; Detection; Pre-Readers; Serious Games; Web-based Assessment; Universal Screening; Language-Independent;Visual; Musical; Information Processing; Gamification

W4A '18, April 23–25, 2018, Lyon, France

ISBN 978-1-4503-5651-0/18/04…$15.00

DOI: https://doi.org/10.1145/3192714.3192816

1. INTRODUCTION

The American Psychiatric Organization defines dyslexia as a *specific learning disorder* which affects around 5% to 15% of the world population [1]. Those affected by dyslexia usually have difficulties in reading and writing, independently from the mother tongue. Dyslexia does not affect the general intelligence of a person. Hence, people with dyslexia understand the meaning of the words but do not always know how to spell or read a word. Often this results in bad grades at school and frustration for students and parents over many years (40% to 60% of children with dyslexia have psychological difficulties [32]). These are common indicators for detecting a person with dyslexia. Other indicators for detecting dyslexia relate to linguistic skills, *e.g.*, *prosodic* or *phonological awareness* [9], differences in reading and spelling error rates of people with and without dyslexia [3, 31], and differences in game measures derived from phonological awareness, letter recognition or word recognition [28]. These and other language related indicators have been used in various software for screening, prediction or detection of dyslexia (we use this three terms indistinguishably). Examples of software for detecting dyslexia in English are **Lexercise Screener** [15], **Nessy** [19] and **Dytective** for both English [29] and Spanish [28].

All these reading and spelling applications are language dependent. This means on one hand that the content of the application needs to be adapted for every new language which is time and resource consuming. On the other hand, only people who already have language acquisition can be tested.

Children with dyslexia can learn the spelling of words or decode words for reading. But they need more time, as well as special and intense treatment. For example, two years instead of one for learning how to spell phonetically accurate words [32]. Hence, to give children with dyslexia more time to practice, avoid frustration and the possibility to succeed, early detection is needed.

Detecting dyslexia in children before they learn to read and write is difficult because the indicators above all use manifestations of reading and writing. This means that children can be detected only after they begin to learn to read and write. This puts students with dyslexia behind. Therefore, new ways of detecting the risk of having dyslexia are needed for pre-readers. Prior studies show approaches to predict future language acquisition of pre-readers *e.g.*, from

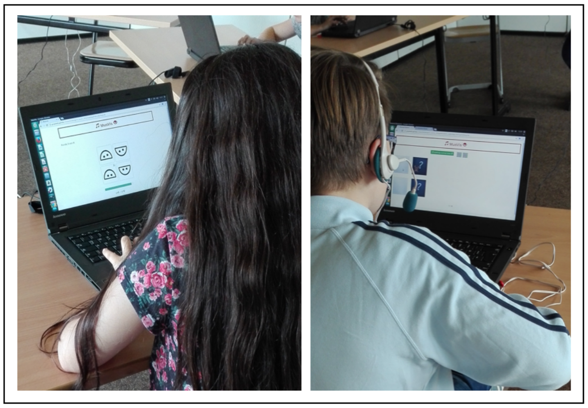

Figure 1: Participants playing the visual part (left) and the musical part (right) of the Game *MusVis*. Photos included with the adults' permission.

newborn with brain recordings [16], from infants with *rapid auditory cues* [2], and from kindergarten children with the perception of *visual-spatial attention* [5].

Previous research has related speech perception difficulties to auditory processing, phonological awareness and literacy skills [30, 34]. Phonological deficits of dyslexia have been linked to basic auditory processing [10]. The auditory perception of children with dyslexia has been proven to be related to the sound structure [12] as well as to the auditory working memory [17]. None of these require reading ability, and may be useful in detecting dyslexia.

Related research suggests that reading impairments are due to the visual-spatial attention and poor coding instead of phonological difficulties [36]. Apart from that, visual discrimination and search efficiency are used as predictors for future reading acquisitions [5]. This prior work motivated us to design our game content with musical and visual elements to create a language independent environment to analyze the differences in the game measures between children with and without dyslexia. For the content design of the musical and visual elements prior knowledge of language acquisition, phonological awareness, letter naming or letter recognition is not needed.

To create the musical elements, we used acoustic parameters in the musical part of our game *MusVis*. To create the visual elements, we designed different visual representations similar to visual features of annotated error words from people with dyslexia [23, 27] and designed the game as a simple search task which does not require language acquisition. Additionally, the participants need to store chunks of information in their short-term memory for both parts of the game.

Next, we present the first results of the game measures collected from children with and without dyslexia while playing the game *MusVis* (see Figure 1) as well as the game content. With a pre-study like this where participants already diagnosed with dyslexia are participating, we reduce the time to find the indicators and increase the chances of making a promising approach before smaller children participate in a long-term study. In this study, we do find game measurements which we can use as indicators to distinguish readers with and without dyslexia after playing our game.

2. RELATED WORK

Dyslexia is a specific reading disorder, which is probably caused by the *phonological skills deficiencies associated with phonological coding deficits* [35]. A person with dyslexia has visual and auditory difficulties that cause problems in reading and writing. It does not affect how intelligent a person with dyslexia is [1]. We focus in related work on digital approaches (games) to predict, screen or detect pre-readers or content unrelated to the knowledge of phonological awareness or letter naming as well as on the perception of sound and visual cues.

Audenaeren *et al.* present a tool called *DIESEL-X* [7] which intends to predict the possibility of a child having dyslexia. The tool includes three mini-games to measure dyslexia related indicators (*e.g.*, 'letter knowledge, FM detection, end-phoneme recognition' [7]). An example task is that the child is asked to find a certain letter in the game. The focus of the development was the gameplay and motivation for a player. To the best of our knowledge, the validation of the prediction model is not published yet.

Another screening computer-based game is *AGTB 5–12* for children at the age of five to twelve [33]. The game has twelve tasks in total and every task takes seven minutes (total game duration is 87 minutes). The tasks focus on the phonological working memory processing, the central working memory, and the visual-spatial working memory.

In a similar vein, the *Bielefelder Screening* provides nine tasks for children at the last year of kindergarten [33]. In 20 to 25 minutes the children do different types of tasks on phonological perception, phonological working memory processing, long-term memory, and visual attention. The results are the categorization of risk groups for dyslexia. The published accuracy of the prediction process has not been found.

Gaggi *et al.* [6] publish their preliminary results with a sample size of 24 participants (last year of kindergarten) using *six different visual and/or auditory games*. The game performance was compared to detect a child but the games were originally developed to train different skills of a child with dyslexia. There is no information about the total duration of the game.

2.1 Sound Perception and Dyslexia

Different theories and empirical results motivated us to use modifications of *acoustical parameters* as game content. For example, the *rapid auditory processing deficit hypothesis* assumes individuals with dyslexia have problems processing short auditory cues. Another theory, claims the dynamic change of the acoustical parameters cause the difficulties [10].

Since the *phonological grammar* of music [22] is similar to the prosodic structure of language, music *i.e.,* a combination of acoustical parameters, can be used to imitate these features. Studies showed a significant difference in the perception of readers with dyslexia on the syllable stress compared to the control group at the age of 9 [8].

For example, the rise time of a sound could imitate stress levels on syllables. Additionally, findings suggest a relation between rise time perception and the *prosodic* and *phonological development* [12].

Even newborns respond automatically to the complex task of perceiving music [37] and show differences in the perception of *sensitivity to native versus non-native rhythmic stress* [9] by the age of 5 months or to the phonemic length by 6 months [16].

Because of the similarities of music and language different acoustic parameters of sound have been explored and proven significance in the perception of children at the age of 8 to 13 years with and without dyslexia [12] *e.g., rise time*, *short duration (100ms)*, *intensity*, and *rhythm*. Also, the perception of pitch and its patterns relates to reading skills which are one difficulty people with dyslexia have [30, 37].Furthermore, a lab study showed different behavior of infants ($m = 7.5$ months) using complex sound frequencies to predict child's language skills at the age of three [2].

A recent study found evidence that dyslexia-associated genes are related to the encoding of sounds in the auditory brainstem [18].However, there are musicians with dyslexia which scored better on auditory perception tests than the general population [17]. At the same time, these participants score worse on tests of auditory working memory, *i.e.,* the ability to keep a sound in mind for seconds. This observation is in line with the results on perceptions for short duration sounds [12] and the findings on the *prosodic similarity effects* of participants with dyslexia [9]. One connection between the difficulties in the perception of language and music seems to be the short-term memory and the recall of information chunks [9]. Since people with dyslexia have short-term memory difficulties [14, 21] questions like "Which sound did you hear first" or "Which sound is pitched higher?" would determine the groups [12].

Huss *et al.* [12] already showed that significant performance differences can be found using musical metrical structure between people with and without dyslexia between the age of 8 and 13 years old in a controlled setting.

2.2 Visual Perception and Dyslexia

Apart from the auditory perception results already mentioned, previous research suggests that the cause of reading impairments could be partly due to the visuo-spatial (also called visual-spatial [5]) attention and poor visual coding instead of the auditory difficulties [36]. These would mean that the difficulties people with dyslexia have in reading and writing are due to a poor decoding of visual cues, *e.g.,* letter recognition, especially, for error cases where a person has a good phonological awareness but difficulties in reading non-words.

Further, findings also provide evidence that the cause of dyslexia might be due to *a more basic cross-modal letter-to-speech sound integration deficit* and the *pre-reading visual parietal-attention* [5]. They are able to predict reading acquisition in preschoolers with the visual-spatial attention. An example of a visual-spatial attention task is a search task (searching for symbols) which shows significant differences in the error rate for poor readers in first grade [5].

The analysis of error words from children with dyslexia shows that the wrong and correct letters in errors words are visually similar as well as through different languages, *e.g.,* English, Spanish [27] or German [23]. The annotated error and correct letters show similarities in different visual features called *mirror letter* (*e.g.,* $< n > < u >$) or *fuzzy letter* (*e.g.,* $< s >$ and $< z >$). The letters have also similarities in the *vertical* (*e.g.,* $< m >$) and *horizontal* symmetries (*e.g.,* $< e >$) through the visual features [27].

However, to predict pre-readers which have no knowledge of language acquisition, phonological awareness or letter recognition is still a challenge. Our game *MusVis* aims to distinguish between readers with and without dyslexia with the derived measurements of our game. This is the first step towards a prediction of dyslexia with measurements derived from a game for readers and pre-readers.

The related work focused on using one evidence, *e.g.,* local visual search on one game. We are combining findings from previous literature, which are known to cause troubles for children with dyslexia as explained before, to create a game environment to find solid differences for predicting dyslexia in the future. At the same time, the game should be fun and not too difficult. We expect people with dyslexia to make more mistakes and take more time than the control group.

This is a language independent approach which has the potential of detecting pre-readers by just translating the instructions of the game because it is independent from knowledge of, *e.g.,* word recognition, letter recognition or phonological awareness.

3. GAME DESIGN

The game *MusVis* aims to measure differences in how children with and without dyslexia react on musical and visual cues. For that reason, we designed a musical (see Figure 2) and a visual part (see Figure 4) of the game *MusVis* with features extracted from the literature. Because of the different perceptions involved (sound and visual), the game design of each part is different but both games trigger the short-term memory.

As is well known, children have more difficulties to pay attention over a longer period of time. Therefore, the two games have four stages each and eight rounds that need less than 10 minutes to play. We used game mechanics, *e.g., rewards (points, score), feedback (instant feedback, progress bar, visible status) or challenges (time limit)* and game components, *e.g., story* for the game design. The content design, user-interface, interaction and implementation for the musical and visual part of the game are described in the following subsections.

3.1 Content Design for Musical Elements

The musical part is adapted from the already existing visual game *Memory.*[1] This game was chosen because it is a well-known child game and could be easily transformed to musical elements. The musical elements were already evaluated with a five user study to discover usability problems that could influence the prediction approach [25].

The participant is asked to find all similar musical elements instead of the same pictures and no time pressure is given. To avoid a random match in the first two clicks, the participants always listen to the first two different sounds of the round. The last two cards and clicks are always correct because these are the remaining cards of the memory game. The musical element is played once when clicked.

This part has four stages which are counter-balanced with *Latin Squares* [4]. Each stage is assigned to one acoustic parameter of sound, *i.e., frequency, length, rise time, rhythm* and three musical elements are created, for example, with different rhythms. Each stage has two rounds with, first two, and then three, musical elements that must be matched. The sounds arrangement for every round are in random order.

[1]An example of a visual memory game can be found on https://goo.gl/vhWmYs.

	Features					General		
	Complex vs. simplex	**Pitch**	**Sound duration**	**Rise time**	**Rhythm**	**Shot-term memory**	**PSE***	**CAPS****
Literature								
Overy [21]			x		x	x		
Huss et al.[12]	x	x		x	x			
Goswami et al.[9]	x			x		x	x	x
Yuskaitis et al. [37]		x						x
Johnson [14]						x		
Stage								
frequency	x	x	x			x	x	x
length			x			x	x	x
rise time	x		x	x		x	x	x
rhythm	x		x		x	x	x	x

Table 1: Mapping of the evidence from literature to distinguish a person of dyslexia, the features and general assumptions and the stages of the musical part of the game *MusVis*. *phonological similarity effect; **correlation acoustic parameters speech

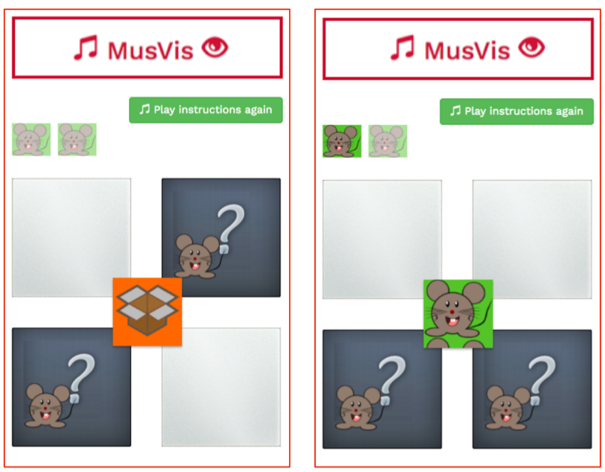

Figure 2: Example of the musical part from the game *MusVis* for the first two clicks on two sound cards (left) and then a pair of equal sounds is found (right). The participant is asked to find two equal musical elements by clicking on sound cards in a row.

The musical elements are generated with a simple sinus tone using the free software *Audacity*. The exact parameters of each musical element are described in the preliminary usability study [25] and the Musical Elements are available at *GitHub* [24]. The musical elements generated for the game *MusVis* are designed with the knowledge of previous literature. We present the mapping of the literature in Table 1 which provide evidence to distinguish a person of dyslexia to our designed stages for the musical part of the game *MusVis* and the following is a short summary.

Same for all musical elements: Each acoustic stage has three musical elements (we use MP3 for sound files). Only one acoustic parameter is changing within a stage.

Stage *frequency:* The frequencies used are in the auditory perception range of a person starting from 440 Hz. We combine the simple tone with a relatively short duration of 0.350s. Each musical element of this stage differs by 0.25 of a semitone (440 Hz to 452.8929 Hz to 446.3998 Hz). We use for the first round of two sound pairs the 440 Hz and 446 Hz musical elements.

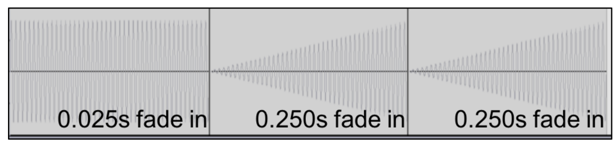

Figure 3: Waveform for the order of intervals for one musical element of the stage *Rise Time*. The example starts with a 0.025s fade-in interval and then a 0.250s interval followed by a 0.250s fade-in interval.

Stage *length:* Each musical element of this stage has a different duration (0.350s, 0.437s, 0.525s), *i.e.*, tone length. The differences between the length of each musical element follow the suggested short duration (100msec) from Huss *et al.* [12]. We use for the first round of two sound pairs the 0.350s and 0.525s musical elements.

Stage *rise time*: At this stage, each musical element is designed with either a short fade in of 0.025s or a fade in of 0.250s or a fade out of 0.250s. We use for the first round of two sound pairs the 0.025s and 0.250s fade ins.

Stage *rhythm:* At this stage, the musical elements are designed with two intervals of rise time equal to 0.250s fade in and one interval equal to 0.025s fade in, in a different order. The order of fade in for each musical element is changed according to the limit of possibilities (see example in Figure 3). We always use for the first round the two sound pairs with the order of rise time interval *0.025s, 0.250s, 0.250s* and musical element with the rise time order reversed.

To keep the game duration short, we only include very promising and easy to deploy acoustic parameters. Parameters like *intensity*, even though, showed significant differences in controlled environment studies could not easily be controlled in our online remote study, as different personal computers and headphones might produce different volume levels due to hardware or software diversity.

3.2 Content Design for Visual Elements

The visual part of the game is similar to the interaction of Whac-A-Mole. We adapted the interaction design and content for this purpose, as shown in Figure 4. For the visual game, we design elements that have the potential of making

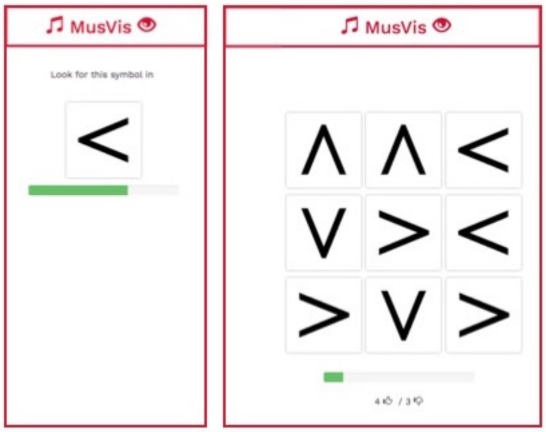

Figure 4: Example of the visual part of the game *MusVis* with the priming of the target element *symbol* (left) and then the nine-squared design including the distractors for each *symbol* (right).

more elements with similar features and represent horizontal and vertical symmetries which are known to be difficult for a person with dyslexia [23, 27, 36].

At the beginning, participants see the target visual element (see Figure 4, left) for three seconds. They are asked to remember this visual element. After that, the participants are presented with a setting where the target visual element and distractors are displayed (see Figure 4). Within 15 seconds the participants try to click as often on the target visual element as possible. The arrangement of target and distractor elements is randomly changed after every click.

The visual part has 4 stages which are counter-balanced with *Latin Squares* [4]. Each stage is assigned to one visual type (***symbol, z, rectangle, face***) and four visual elements for each stage are presented. One visual element is the target which the participants need to find and click (see Figure 5, top). The other three visual elements are *distractors* for the participants. Each stage has two rounds (in total the number of rounds is 8) with first a 4-squared and then a 9-squared design (see Figure 4, right). The target and all three distractors are displayed in the 4-squared design. In the 9-squared design, the target is displayed twice as well as the distractor two and three. Only distractor one is displayed three times. The stages from Figure 5 are summarized next.

Stage *symbol:* This stage uses two lines connected in an angle of less than 30° as the target visual element and creates a vertical symmetry. The distractor one is mirrored while the distractor two and three are rotated by 90° and -90°.

Stage *z:* The target visual element for this stage is created with two lines parallel to each other connected with a diagonal line. The diagonal line is drawn from the top right line end to the down left line end. This creates vertical and horizontal symmetry of the visual element. This representation looks very similar to the letter *z* but we do not use the phonological awareness of the letter, *i.e.* the participants do not need to know that this is also an existing letter of the Latin alphabet. The distractor one is mirrored while the distractor two and three are rotated by 90° and -90°.

Stage *rectangle:* This stage is the shape of a square and a right-angled triangle. These shapes have by design vertical and horizontal symmetries which we use to create a complex

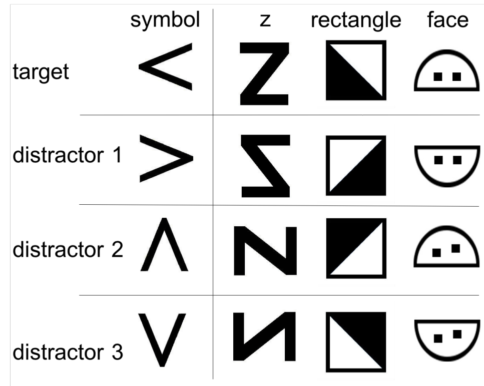

Figure 5: Overview of the designed visual elements. The figure shows the target element (top) and distractor elements (below) for the four different stages (*z, symbol, rectangle, face*) of the visual part of the game *MusVis*.

target. The outline shape is the square and two triangles are placed within the square. The 90° corner of one triangle is placed in the up-right corner of the square and the other triangle in the below-left corner of the square. This creates a visual element with different ways to perceive similarities within the element. The distractors are rotated by 90°, 180° and 270°.

Stage *face:* The target visual element has three visual cues combined (two symmetric dots placed horizontally and an outline around them). The outline is first a straight horizontal line under the two dots and connects the ending with a bow around the two dots. The whole target element is symmetrical on the vertical line. The target is rotated 180° for the first and third distractor. Additionally, the two dots are slightly staggered up and down for the second and third distractor.

3.3 User-Interface and Implementation

To support the readability for parents and supervisors we used a large font size (minimum 18 points) [26]. The interactive elements (cards to be clicked within the game) are large enough to be clicked easily. The presentation of interactive elements (sound cards/squares) are the same within each game and do not differ in color or shape to avoid differences in the perception.

Both games are implemented as a web-application using JavaScript, jQuery, CSS, HTML5 and a backend with a PHP server and a MySQL database. One reason for this is access simplicity for remote online-studies. Another reason is the advantage of adapting the application for different devices in future research studies.

4. EXPERIMENTAL STUDY

We conducted an independent within-subject design study with 178 participants and included only participants which are either diagnosed with dyslexia ($n = 67$) or without dyslexia ($n = 111$). Every participant played all the game rounds in their mother language with the same game content. Only the study and game instructions (audio and text) are translated into the mother tongue. We recruited Spanish participants diagnosed with dyslexia mostly over public social media calls from the non-profit organization *ChangeDyslexia* (https://changedyslexia.org/). German participants diagnosed with dyslexia have been mostly recruited over social media calls in support groups. The control group was recruited with the collaboration of two Spanish schools and two German schools.

4.1 Procedure

The communication with the participants was mostly via email. The web-application was played at home or in the school with one researcher (authors of the paper) present or always available through digital communication.

First, the parents or supervisor filled out the demographic questionnaire for gaining background information, *e.g.,* age of the participant, dyslexia diagnosis (yes/no/maybe) and the mother tongue.

This was followed by explaining instructions for the user study to the parents or supervisor, *e.g.,* turn up the volume, use headphones, play without interruptions or explain and help your child only with the instructions of the games.

Then a short video story for the musical part was played. After that every participant played first the musical and then the visual part of *MusVis* (see Figure 1). Finally, the parents answered two feedback questions and left their contact details.

Each input method (*computer* vs. *tablet*) needs to be analyzed separately. We decided to use a laptop or desktop computer for two reasons: (1) From prior game evaluation [6, 28] we know that readers are able to interact with the device and (2) these devices are still more available than tablets [13].

4.2 Participants

The analysis data includes only the data from participants who played all 16 rounds of the game *MusVis* with a computer. Dropouts happened mostly because participants used a different browser (e.g., *Internet Explorer* instead of *Google Chrome*) or a different device (tablet instead of a computer).

To have a more accurate analysis we made sure we know the status of a participant (diagnosed or not) and excluded participants which reported they might have dyslexia. Only participants that showed no indication of dyslexia or with an official diagnosis, by a medical doctor or equivalent, were included. Thirteen participants were suspected of dyslexia and therefore, taken out of the analysis.

We report separately the results for the Spanish participants ($n = 108$), German participants ($n = 57$) and an analysis with all languages for the language independent variables where we added English ($n = 6$), and Catalan ($n = 7$). For the dependent variables (DV) which show indications of the same tendency of results and are therefore considered as *language independent* .

For the analysis with all languages ($n = 178$), we considered for the dyslexia group, 67 participants which were diagnosed of dyslexia (33 female, 34 male). Their ages ranged from 7 to 12 years ($m = 9.8$, $sd = 1.4$). For the control group, we considered 111 participants (67 female, 44 male). Their ages ranged from 7 to 12 years ($m = 10.5$, $sd = 1.5$).

For the Spanish participants, we considered 41 participants diagnosed of dyslexia (23 female, 18 male). Their ages ranged from 7 to 12 years ($m = 9.5$, $sd = 1.1$). For the control group, we took into account 67 participants (42 female, 25 male). Their ages ranged from 7 to 12 years ($m = 10.0$, $sd = 1.2$).

For the German participants, we considered 17 participants diagnosed with dyslexia (5 female, 12 male). Their ages ranged from 7 to 12 years ($m = 10.7$, $sd = 1.4$). For the control group, we had 40 participants (21 female, 19 male). Their ages ranged from 7 to 12 years ($m = 11.4$, $sd = 1.4$).

4.3 Dependent Measures

The *dependent variables* we collected from the user interaction with the web-based game *MusVis* were for both parts of the game *time intervals of clicks and total number of clicks*.

For the musical part, we additionally collected the *duration of each round and average click time* (we calculated the average click time by dividing the duration with the total number of clicks). Because of the gameplay, the first three cards need less rethinking to find where the sound is the same. Therefore, we consider for the musical part the first 4 clicks as one interval, *4th click interval*, to measure the duration. Every click interval after that can be used as well and we choose the *6th click interval*, which even exists when the participant finishes the game in the shortest click sequence possible.

For the visual part we collected *time to the first click, number of hits or correct answers, number of misses or non-correct answers, efficiency* (we calculated the efficiency by dividing the time of the last click by hits) and *accuracy* (we calculated the accuracy as hits divided by the total number of clicks).

5. RESULTS

In order to find out whether we have new indicators to predict people with dyslexia after playing *MusVis*, we analyzed the dependent variables for our independent within-subject study for the three groups: *Spanish, German, all languages*. We applied first the *Shapiro-Wilk test*. All variables ($n = 54$) were not normally distributed and we applied, therefore, the independent *Wilcoxon Test*. All analyses were conducted with a *Bonferroni correction* to avoid type I errors (2.4e-3). We present the results for Spanish in Table 2, for German in Table 3, and for all languages in Table 4.

The DVs are categorized for Spanish and German according to the tendency that participants with dyslexia compared to the control group had within each language (see Table 5). An example of a language independent variable is the DV *hits* because the dyslexia group has in German and Spanish significantly less correct clicks (Spanish 5.7; German 5.6) than the control group (Spanish 6.6; German 6.3). The DV *duration* is an example of the opposite trend because the dyslexia group for Spanish takes significantly more time while the German participants with dyslexia take less time compared to their language control group. Only if the tendency was similar, the DV were included in the overview of all languages (Table 4). We consider the variables in Table 5 as a first step to provide evidence towards a language independent detection.

We use the *effect size* to estimate the likely size of the effect in the population. The effect size (r) for the *Wilcoxon Test* [4] is calculated as

$$r = \frac{z}{\sqrt{N}}$$

where z is the z-score and N is the number of observations.

We use the effect size in Tables 2 and 3, only for the significant results. First, we report the results for the musical part and then for the visual part of the game.

Total number of clicks (music) is not language independent. The tendency of results is opposite between partici-

Dependent variables Spanish	**Control mean**	**sd**	**Dyslexia mean**	**sd**	**Wilcoxon W**	**p-value**	**z**	**effect size**
Musical								
Total clicks	11.0	5.5	11.3	6.0	86231	0.63	-0.48	0.05
4th click interval	1.6s	0.7s	2.0s	1.3s	63658	**7e-12**	-6.80	0.66
6th click interval	1.5s	0.8s	1.7s	1.2s	76762	**2e-3**	-3.13	0.30
Duration	27.5s	17.1s	34.3s	27.0s	72316	**1e-5**	-4.38	0.42
Average click time	2.5s	0.8s	3.0s	1.2s	59028	**2e-16**	-8.11	0.78
Visual								
Total clicks	8.0	3.3	6.7	2.7	110000	**3e-10**	6.25	0.60
Time to first click	2.3s	1.4s	2.7s	1.8s	75566	**5e-4**	-3.47	0.33
Hits	6.6	2.9	5.7	3.0	105670	**5e-7**	5.02	0.48
Misses	1.3	3.1	1.0	1.8	86340	0.62	-0.50	0.05
Accuracy	0.60	0.50	0.57	0.50	90432	0.43	0.83	0.08
Efficiency	2.8s	2.6s	3.1s	2.8s	73301	**4e-5**	-4.10	0.39

Table 2: Overview of all reported dependent variables for the musical and visual part of the game *MusVis* for Spanish only ($n = 108$).

Dependent variables German	**Control mean**	**sd**	**Dyslexia mean**	**sd**	**Wilcoxon W**	**p-value**	**z**	**effect size**
Musical								
Total clicks	10.8	5.4	10.6	4.4	20880	0.49	-0.70	0.09
4th click interval	1.8s	0.8s	2.0s	1.2s	19218	0.05	-2.00	0.26
6th click interval	1.7s	0.7s	1.6s	0.7s	21580	0.89	-0.14	0.02
Duration	28.5s	16.9s	27.9s	13.0s	20542	0.34	-0.95	0.13
Average click time	2.6s	0.8s	2.6s	0.5	19.708	0.11	-1.59	0.21
Visual								
Total clicks	7.2	3.1	6.8	2.7	23887	0.10	1.67	0.22
Time to first click	2.4s	1.5s	2.5s	1.1s	19314	0.06	-1.90	0.25
Hits	6.3	2.8	5.6	2.6	24675	0.02	2.28	0.30
Misses	0.9	2.1	1.2	2.3	20718	0.36	-0.92	0.12
Accuracy	0.60	0.49	0.59	0.49	22084	0.83	0.30	0.04
Efficiency	2.8s	2.3s	3.2s	2.9s	19357	0.06	-1.87	0.25

Table 3: Overview of all results reported dependent variables for the musical and visual part of the game *MusVis* for German only ($n = 57$).

pants with dyslexia (Spanish $m = 11.3$ & German $m = 10.6$) compared to participants without dyslexia (Spanish $m = 11.0$ & German $m = 10.8$). This means that German participants with dyslexia click less compared to the German control group and Spanish participants with dyslexia click more compared to the Spanish control group. The *total number of clicks* did not reveal significant differences on *total clicks* for Spanish ($W = 86231$, $p = 0.63$, $r = 0.05$) or German ($W = 20880$, $p = 0.49$, $r = 0.09$). The effect size for Spanish and German is nearly zero, so is considered as it has no effect [4].

Click time interval (music) is not language independent over all click intervals and we, therefore, do not report any click intervals for all languages. Hence, participants with dyslexia (Spanish 4th click interval $m = 2.0$s & German 4th click interval $m = 2.0$s & Spanish 6th click interval $m = 1.7$s) take more time before they make the next click than the control group (Spanish 4th click interval $m = 1.6$s & German 4th click interval $m = 1.8$s & 6th click interval $m = 1.5$s). But German participants with dyslexia (6th click interval $m = 1.6$s) take less time before they make the next click than the German control group ($m = 1.7$s). The *4th time interval* ($W = 63658$, $p =$ 7e-12, $r = 0.66$) as well as the *6th click interval* ($W = 76762$, $p = 2e-3$, $r = 0.30$) is significant for Spanish but not for German ($W = 21580$, $p = 0.89$, $r = 0.02$). The effect size for *4th click time interval* Spanish is considered as large where the effect size for *6th click time interval* for Spanish and the *4th click time interval* for German is considered as medium [4]. We report only the fourth and sixth click interval for the musical game since the first three intervals do not show, as expected due to the game design, any significant differences between the groups.

Duration (music) is not language independent. Hence, Spanish participants with dyslexia ($m = 34.3$s) take more time to find all pairs and finish the round than the Spanish control group ($m = 27.5$s). But German participants with dyslexia ($m = 27.9$s) take less time before they find all pairs than the German control group ($m = 28.5$s). The *duration* is significant for Spanish ($W = 72316$, $p = 1e-5$, $r = 0.42$) but not for German ($W = 20542$, $p = 0.34$, $r = 0.13$). The effect size for Spanish is considered as medium and for German as small [4].

Average click time (music) is not language independent. Because Spanish participants with dyslexia ($m = 3.0$s) take in average more time for a click than the Spanish control group ($m = 2.5$s). But German participants with and without dyslexia take in average the same time ($m = 2.6$s). Spanish participants with dyslexia significantly spend more time for each click ($W = 59028$, $p =$ 2e-16), $r = 0.78$) where

Dependent variables Visual	Control mean	sd	Dyslexia mean	sd	Wilcoxon W	p-value
Total clicks	7.6	3.2	6.8	2.7	276120	**3e-7**
Time to first click	2.4s	1.5s	2.6s	1.6s	210850	**3e-4**
Hits	6.5	2.9	5.8	2.9	272180	**4e-6**
Accuracy	0.60	0.49	0.58	0.49	240780	0.66
Efficiency	2.8s	2.5s	3.1s	2.7s	209740	**2e-4**

Table 4: Overview of all language independent results for the visual part of the game ($n = 178$).

we cannot measure a difference for German ($W = 19708$, $p = 0.11$), $r = 0.21$). The effect size for Spanish is considered as large and for German as small [4].

Total number of clicks (visual) is language independent. Participants with dyslexia ($m = 6.7$) significantly clicked less times than participants without dyslexia ($m = 8.0$) for Spanish ($W = 110000$, $p =$ 3e-10, $r = 0.60$). The effect size for Spanish is considered as large [4]. The German participants have the same trend for the control group ($m = 7.2$) compared with the group of participants with dyslexia ($m = 6.8$, $W = 23887$, $p = 0.10$, $r = 0.22$). The effect size for German is considered as small [4]. Because the trend is the same, we provide the analysis for the *total number of clicks* which confirm the significant difference ($W = 276120$, $p = 3e - 7$).

Time to the first click (visual) is language independent. This means that participants with dyslexia (Spanish $m = 2.6$s) and German ($m = 2.5$s) take more time before they make the first click than the control group (Spanish $m = 2.3$s & German $m = 2.4$s). The *time to the first click* is significant for Spanish ($W = 89450$, $p =$ 1e-3, $r = 0.30$) but not for German ($W = 19314$, $p = 0.06$, $r = 0.25$). The effect size for Spanish and German is considered as medium [4]. Because the trend is the same for both languages even though it is only significant for Spanish, we provide the analysis for the *time to the first click* with a significant difference for all languages ($W = 210850$, $p = 3e - 4$).

Hits is language independent. Hence, participants with dyslexia (Spanish $m = 5.7$s) & German ($m = 5.6$s) have less hits than the control group (Spanish $m = 6.6$s & German $m = 6.3$s). The *hits* is significant for Spanish ($W = 105670$, $p =$ 5e-7, $r = 0.48$) and for German ($W = 24675$, $p = 0.02$, $r = 0.30$). The effect size for Spanish and for German is considered medium [4]. Because the trend is the same for both languages, we provide the analysis for the *hits* with a significant difference for all languages, $W = 272180$, $p =$ 4e-6.

Misses is not language independent. Hence, Spanish participants with dyslexia ($m = 1.0$) make less mistakes than the Spanish control group ($m = 1.3$). But German participants with dyslexia ($m = 1.2$) make more mistakes than the German control group ($m = 0.9$). *Misses* has no significant difference for Spanish ($W = 86340$, $p = 0.62$, $r = 0.05$) or German ($W = 20718$, $p = 0.36$, $r = 0.12$). The effect size is considered for both languages small [4].

Accuracy is language independent. There were no differences for participants with dyslexia (Spanish $m = 0.57$ & German $m = 0.59$) and the control group (Spanish $m = 0.60$ & German $m = 0.60$) in *accuracy*. The *accuracy* is not significant different for Spanish ($W = 90432$, $p = 0.43$, $r = 0.08$) or German ($W = 22084$, $p = 0.83$, $r = 0.04$). Because the trend is the same for both languages, we provide the analysis for the *accuracy* with no significant difference for all languages ($W = 240780$, $p = 0.66$).

Efficiency is language independent. Hence, participants with dyslexia (Spanish $m = 3.1$s & German $m = 3.2$s) take more time for a hit than the control group (Spanish $m = 2.8$s & German $m = 2.8$s). The *efficiency* is significant for Spanish ($W = 73301$, $p = 4e - 5$, $r = 0.39$) but not for German ($W = 19357$, $p = 0.06$, $r = 0.25$). The effect size for both languages is considered as medium [4]. Because the trend is the same for both languages, we provide the analysis for the *efficiency* with a significant difference for all languages ($W = 209740$, $p = 2e - 4$).

Children and parents provided positive ($n = 44$) and negative ($n = 7$) feedback about the gameplay or content. Translated positive example quote from a boy (8 years) who participated in a school: *This was so cool! It was the best day at school ever*; from the web feedback input field of a girl (12 years): *it was fun and not boring!*; or a boy (10 years): *I love this game.* The positive feedback was provided by all age groups. Translated negative example quotes from a girl (12 years): *not exciting more boring* or a boy (12 years): *game started to fast.*

6. DISCUSSION

The measurement data taken from the game *MusVis* show that Spanish participants with dyslexia behave differently than their control group. Differences can be reported for the musical game for: *4th click interval, 6th click interval, duration, and average click time.* For the visual part the following measurements can be reported as indicators: *total clicks, time to the first click, hits, and efficiency.* Besides, similar tendencies can be reported for the variables of the visual part: *total clicks, time to the first click, hits, accuracy, and efficiency* (see Table 5). We can show with our results over all languages that the effect for each measurement is confirmed even if we cannot draw strong conclusions about our sample size on the comparison of German vs. Spanish speaking participants. Spanish has 8 significant indicators and we expected to reproduce the same amount of significant indicators with more German participants.

In general, all participants found the game easy to understand and only children at the age of 12 complained about missing challenges. The amount of positive feedback and engagement of all age groups let us conclude that the **game mechanics and components** applied are also positive to perceive the *MusVis* as a game and not as a test.

Dyslexia is known to be present **across different languages and cultures** [1]. The assumption that the tendency for the indicators are similar over all languages cannot be proven for all indicators in our study, e.g. German participants with dyslexia start faster to click (music) than the Spanish participants compared to their language control group. We can exclude external factors like different applications or different study set up as a possible influence on opposite tendency. According to the results we may have

Musical	Language Independent
Total clicks	No
4th click interval	No
6th click interval	No
Duration	No
Average click time	No
Visual	**Language Independent**
Total clicks	**Yes**
Time to first click	**Yes**
Hits	**Yes**
Misses	No
Accuracy	**Yes**
Efficiency	**Yes**

Table 5: Overview of all dependent variables showing the language independent results between the German and Spanish groups.

to assume that not all indicators for dyslexia are language independent and have cultural dependencies. To confirm this assumption we will need an equal larger number of participants for both language groups (Spanish and German).

The variables *time to first click (visual & music)* and *total number of clicks (visual & music)* provide apart from the cultural society also dependencies of the **game content and game design**. Otherwise, we could not explain the trend difference between the musical and visual part for *total number of clicks*, *i.e.*, total clicks for visual is significantly different than for music. Additionally, the analyses of the musical part of the game present two limitations: (1) participants could select a correct pair by chance, and (2) participants could click through the game board without listening to the sounds.

Children with dyslexia are detected by their slower reading or **error rate** as explained in the introduction [3, 31]. Therefore, we designed our game with content that is known to be difficult to differentiate for children with dyslexia to measure the errors and the duration. Nevertheless, from previous literature we knew that children with dyslexia do not make more mistakes in games than the control group [28]. We can confirm that *misses* did not reveal significant differences for German or Spanish either. It might be possible that we cannot compare errors in reading and writing with errors in this type of games. Then, we cannot explain (yet) why the Spanish control group made more mistakes than the Spanish group with dyslexia.

Spanish children without dyslexia take significantly more time to find all pairs and finish the musical game. Children without dyslexia take more time before they *click the first time* (visual) for all languages. The reason for that might be, the time they need to **process the given information** in the auditory processing [34] or recall the information from the short-term memory [9] for auditory and visual. However, participants with dyslexia from the German group are nearly as fast as the control group in finding all pairs (music) which might be due to the **cultural differences.**

The musical and visual elements are designed on purpose to be more difficult to process for people with dyslexia than without. Therefore, children with dyslexia are expected to need more time (duration) which might be due to a **less distinctive encoding of prosody** [9] and is line with indicator of slower reading. Considering that children with dyslexia need more time to process information we observe this behavior as well for our indicators. For example, participants with dyslexia from the Spanish group take more time on the *4th click interval* and also on the *average click time* compared to the control group. Both results are significant and with a large effect size of 0.7 and 0.8, we can estimate what effect would be also in the whole population [4].

A person with dyslexia has difficulties to read and write independently of the mother tongue which also appears when learning a second language [11, 20]. The analysis of errors from children with dyslexia show similar error categories for Spanish, English [27], and German [23] which show similarities of the perception between the languages. Our results suggest that we can measure a significant difference on four indicators for the visual game with the same tendency between Spanish, German, English, and Catalan. These means that a person with dyslexia might perceive our visual game content similar independent of the mother tongue. Further research needs to be done to confirm the results but this first pilot study shows strong evidence that it will be possible to measure dyslexia using our content, approach and game design with the same language independent content for different languages.

7. CONCLUSIONS AND FUTURE WORK

We presented a game with musical and visual elements that collects measures that show differences between people with and without dyslexia for different languages. With our user study, we found eighth significant indicators for Spanish and four significant indicators which are language independent to distinguish a person with and without dyslexia. This is preliminary evidence for a prediction approach for dyslexia that would work across different languages.

The next step is to test how language dependent this approach for different language is and compare tablets (touch) and desktop computer (mouse) input. We already started to pilot studies in Catalan and English to be able to compare the different languages and study whether the game content is truly language independent. For future work, we are planning to conduct a large-scale study to be able to apply a machine learning model to predict dyslexia based on the eight significant indicators for Spanish.

Additionally, we will carry out a longitudinal study with pre-readers to be able to predict children before they acquire reading and writing skills. This would provide children with dyslexia, more time to practice and compensate their difficulties before starting school. Long term we plan to offer a game, *MusVis*, with the musical and visual indicators that could be language independent, applicable for pre-readers and easily accessed from any given location with a tablet, laptop or desktop computer. This will leverage the opportunities of children with dyslexia being able to predict their difficulties when they are pre-readers, in time to effectively intervene.

8. ACKNOWLEDGMENTS

This paper and content were partially supported by the ICT PhD program of Pompeu Fabra through a travel grant and funded by the *fem:talent Scholarship* from the *Applied University of Emden/Leer* as well as by the *Deutschen Lesepreis 2017* from the *Stiftung Lesen* and the *Commerzbank-Stiftung*. We deeply thank M. Jesús Blanque and R. Noé López, *School Colegio Hijas de San José*, Zaragoza; S. Tena, A. Carrasco y E. Méndez, innovation team of *Colegio Leonardo da Vinci*, Madrid; *ChangeDyslexia*, Barcelona; in Spain, and F. Hansch, *Grundschule Ofenerdiek*, Oldenburg;

L. Klaus, *Peter-Ustinov-Schule*, Eckernförde; H. Marquardt, *Gorch-Fock-Schule*, Eckernförde; A. Tammena and W. Volkmann, *Johannes-Althusius-Gymnasium*, Emden; M. Batke and J. Thomaschewski, *Hochschule Emden/Leer*, Emden; and P. Stümpel, *AncoraMentis*, Rheine, in Germany. We also thank all parents and children for playing *MusVis*. Finally, thanks to H. Witzel for his advice during the development of the visual part and to M. Blanca, and M. Herrera on the translation for the Spanish version.

9. REFERENCES

[1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, May 2013.

[2] Benasich et al. Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural Brain Research*, 136(1):31–49, 2002.

[3] Coleman et al. A comparison of spelling performance across young adults with and without dyslexia. *Assessment for Effective Intervention*, 34(2):94–105, 2008.

[4] Field et al. *How to design and report experiments*. SAGE Publications, 2003.

[5] Franceschini et al. A causal link between visual spatial attention and reading acquisition. *Current Biology*, 22(9):814–819, may 2012.

[6] Gaggi et al. Serious games for early identification of developmental dyslexia. *Comput. Entertain. Computers in Entertainment*, 15(4), 2017.

[7] Geurts et al. DIESEL-X: A game-based tool for early risk detection of dyslexia in preschoolers. In *Describing and Studying Domain-Specific Serious Games*, pages 93–114. Springer International Publishing, 2015.

[8] Goswami et al. Impaired perception of syllable stress in children with dyslexia: A longitudinal study. *Journal of Memory and Language*, 69(1):1–17, 2013.

[9] Goswami et al. Prosodic similarity effects in short-term memory in developmental dyslexia. *Dyslexia*, 2016.

[10] Hämäläinen et al. Basic auditory processing deficits in dyslexia. *Journal of Learning Disabilities*, 46(5):413–427, 2013.

[11] Helland et al. Dyslexia in English as a second language. *Dyslexia*, 11(1):41–60, feb 2005.

[12] Huss et al. Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, 47(6):674–689, jun 2011.

[13] IDC Worldwide. Shipment forecast of tablets, laptops and desktop PCs worldwide from 2010 to 2019 (in million units). *Statista*, 2020:2020, 2016.

[14] D. J. Johnson. Persistent auditory disorders in young dyslexic adults. *Bulletin of the Orton Society*, 30(1):268–276, jan 1980.

[15] Lexercise. Dyslexia test - Online from Lexercise. http://www.lexercise.com/tests/dyslexia-test, 2016. [Online; accessed 18-September-2017].

[16] Lyytinen et al. Dyslexia - Early identification and prevention. *Current Developmental Disorders Reports*, 2(4):330–338, dec 2015.

[17] Männel et al. Phonological abilities in literacy-impaired children: Brain potentials reveal deficient phoneme discrimination, but intact prosodic processing. *Developmental Cognitive Neuroscience*, 23:14–25, 2016.

[18] Neef et al. Dyslexia risk gene relates to representation of sound in the auditory brainstem. *Developmental Cognitive Neuroscience*, 24(February):63–71, 2017.

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

[20] J. Nijakowska. *Dyslexia in the foreign language classroom*. Multilingual Matters, 2010.

[21] K. Overy. Dyslexia, temporal processing and music: The potential of music as an early learning aid for dyslexic children. *Psychology of Music*, 28(2):218–229, oct 2000.

[22] R. F. Port. Meter and speech. *Journal of Phonetics*, 31:599–611, 2003.

[23] Rauschenberger et al. A language resource of German errors written by children with dyslexia. In *LREC 2016*, Paris, France, may 2016.

[24] Rauschenberger et al. Supplement: DysMusic musical llements: Towards the prediction of dyslexia by a web-based game with musical elements, 2017. https://doi.org/10.5281/zenodo.809783.

[25] Rauschenberger et al. Towards the prediction of dyslexia by a web-based game with musical elements. In *W4A'17*, pages 4–7, 2017.

[26] Rello et al. Good fonts for dyslexia. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, page 14, 2013.

[27] Rello et al. A resource of errors written in Spanish by people with dyslexia and its linguistic, phonetic and visual analysis. *Language Resources and Evaluation*, pages 1–30, 2016.

[28] Rello et al. Dytective: Diagnosing risk of dyslexia with a game. In *Proc. Pervasive Health'16*, Cancun, Mexico, 2016.

[29] Rello et al. Screening dyslexia for english using hci measures and machine learning. In *Proc. Digital Health'18*, Lyon, France, 2018.

[30] Rolka et al. A systematic review of music and dyslexia. *Arts in Psychotherapy*, 46:24–32, 2015.

[31] Schulte-Körne et al. Familial aggregation of spelling disability. *Journal of Child Psychology and Psychiatry*, 37(7):817–822, oct 1996.

[32] G. Schulte-Körne. Diagnostik und therapie der lese-Rechtscheib-Störung: The prevention, diagnosis, and treatment of dyslexia. *Deutsches Ärzteblatt international*, 107(41), 2010.

[33] Steinbrink et al. *Lese-Rechtschreibstörung*. Springer Berlin Heidelberg, 2014.

[34] P. Tallal. Improving language and literacy is a matter of time. *Nature reviews. Neuroscience*, 5(9):721–728, 2004.

[35] Vellutino et al. Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1):2–40, jan 2004.

[36] Vidyasagar et al. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences*, 14(2):57–63, 2010.

[37] Yuskaitis et al. Neural mechanisms underlying musical pitch perception and clinical applications including developmental dyslexia. *Current neurology and neuroscience reports*, 15(8):51, 2015.