

Speech Technology to Support Phonics Learning for Kindergarten Children at Risk of Dyslexia

Stine Fuglsang Engmose
University College Absalon
Trekroner Forskerpark 4
DK-4000 Roskilde, Denmark
stfe@pha.dk

Peter Juel Henriksen
Danish Language Council
Adelgade 119B
DK-5400 Bogense, Denmark
pjh@dsn.dk

Abstract

We present the AiRO learning environment for kindergarten children at risk of developing dyslexia. The AiRO frontend, easy to use for pupils down to 5 years old, introduces each spelling task with pictural and auditive cues. AiRO responds to spelling attempts with phonetic renderings (synthetic voice). Below, we introduce the didactic and technical principles behind AiRO before presenting our first experiment with 49 kindergarten pupils. Our subjects were pre- and post-tested on reading and spelling. After four weeks of AiRO-based training the experimental group significantly out-performed the control group, suggesting that a new CALL-based pedagogical approach to prevent dyslexia for some children may be within reach.

1 Background

An early, but influential study¹ found that 12% of adult Danes had reading difficulties inhibiting their professional life. Dyslexia is a well-described cause of reading difficulties but until recently, dyslexia was studied only superficially

in the Danish education system, leaving teachers little prepared to engage proactively (Pihl and Jensen, 2017). It is problematic if difficulties in reading are not met with appropriate support because adults with poor reading and writing skills are strongly overrepresented among those who have low-paid jobs and short educations (Rosdahl et al., 2013). Among dyslectic 25/26-year-olds, only 69% completed secondary school, compared to 81% among peers (Egmont, 2018). However, early intervention can lessen the problem significantly. Vellutino and Scanlon (2002) report that special training programs for pupils from the age of 7 years reduced the proportion of bad readers from 9% to 1.5%. Effective intervention should be based on intensive, sustained, and individually tailored courses focused on the relations between letters and sounds (Elbro and Petersen, 2004; Elbro, 2021). A solid grip of phonics is a necessary precondition to solid reading and spelling skills (Ehri, 2005; National Reading Panel, 2000; Share, 1995). Early intervention, more than anything else, holds a strong potential for societal and personal gains with dyslexia (Gellert et al., 2018). "We believe that CALL might hold a potential as a supplement to teacher's instruction in a didactic programme of early intervention. As will be clear in the following, our approach concerns a specific CALL setup with a pronounced focus on the writing situation. More specifically, we have developed a didactic tool for use in classrooms, exploiting a very close stimulus-response cycle from student

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Licence details: <http://creativecommons.org/licenses/by/4.0/>.

¹ Elbro et al (1995). Similar figures have been reported from other Western countries.

production ("spelling") to system response ("correction" or "confirmation") with a level of granularity down to the individual letter/phone combination. To our knowledge, no other interactive training tool on the market for children at risk of dyslexia (such as Gissel & Andersen, 2021, Messer & Nash, 2018, and Solheim et al., 2018) use the same level of granularity."

2 Introduction to AiRO

The project AiRO², that we present results from in this paper, seeks to meet some of these societal and personal challenges. We expect that kindergarten children at risk of dyslexia can benefit from an early intervention characterized by a learning environment with positive interaction and corrective feedback. More specifically, a child with poor command of phonics will benefit from a quick and simple response (affirming or correcting) to their spelling attempt. A dedicated teacher can of course provide ideal feedback, but teachers' attention is limited in a classroom with more than 20 kindergarten children. AiRO is developed as an interactive learning tool to supplement ordinary teacher lead instruction.

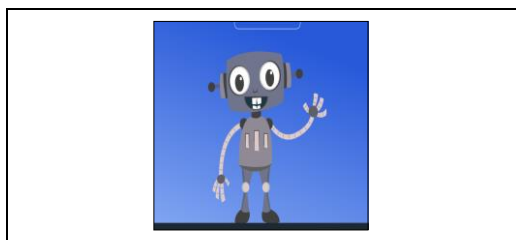


Figure 1. AiRObot greets the kindergarten children at the AiRO frontpage

2.1 AiRObot - your classroom assistant

Seen from the kindergartener's point of view, AiRO is a friendly robot (see the AiRObot in figure 1) presenting manageable spelling tasks, beginning from simple one-letter words and continuing slowly but steadily (depending on the pupil's profile and performance) with ever more demanding words.

AiRO is intended for use in classrooms or small groups. Individual pupils or a small group can use AiRO while the rest of the class are following the regular education. When using AiRO in school, headphones are mandatory; the application is however also available to the pupils at home.

In the following sections, we present AiRO's underlying didactic, linguistic, and computational principles. We also report on our recent experiments with pupils in the Danish pre-primary school (49 subjects). Finally we discuss some future perspectives.

3 Linguistic principles and technical design

To develop spelling and reading skills children must among others acquire and be able to use phonics rules. This is the objective of the CALL-based pedagogical approach for children at risk of dyslexia, AiRO.

Looking at the research of phonics instruction as an early intervention, Danish professor in reading sums up generations of research (Elbro, 2021) in the following headings. For phonics instruction to be helpful for children at risk of dyslexia it should be characterized by being:

- Systematic, e.g. introducing letter-sound-connections that are stable and frequent before connections that are less stable or rare
- Direct, e.g. instruction where words are chosen, in such a manner that the letter-sound-connections introduced can be practiced

² AiRO \approx CALL-based pedagogical approach for children at risk of dyslexia (In Danish Adaptiv it-baseret støtte til børn i Risiko for Ordblindhed)

- Applied, using phonics for reading and spelling words with support and feedback
- Intensive and extensive, small groups of 3-4 students or 1 on 1, daily 30 min. of practice, lots of time spend on the students practicing
- Motivating, making the progress of the student visible to the student and providing lots of task variation to deal with the students slow progress
- At the students instructional level, and progressing slowly

The CALL-based pedagogical approach is designed to create a learning situation with the above characteristics.

In AiRO the user are presented to 3 new and 3 earlier practiced target words at each level. At the initial level, target words are short (1-2 letters) with V, CV and VC structure (e.g. "å" *stream*, "is" *ice cream*) and straightforward pronunciation (see how target words are presented in figure 2). Only letters E, I, L, S, Å are used, and only the most basic letter-to-sound rules are in play. In general, rules trained at one level carry over to the next so that easier rules are practiced before more difficult ones. A total of 20 letter-to-sound rules are covered. The entire course comprises 16 levels, first focusing on the vowels and fricatives, then gradually introducing the plosives. The purpose is to create a learning situation that systematically and directly introduces the user to phonics applied in spelling with abundant opportunity for the user to practice at the appropriate level of instruction and progression.



Figure 2. How target words are presented in AiRO

The target words are accompanied with a picture, and the pronunciation of the specific word. To ensure that the child practices the intended word and also, has the possibility to access the pronunciation an unlimited number of times, a play button is provided.

The user responds by spelling the target word as best they can, letter by letter. For each keystroke, AiRO responds with an auditive rendering of the word-so-far (pronounced by a synthetic voice). Each letter entered by the user is immediately analyzed for correctness, response time, and other metrics. A sound file (synthetic speech) is generated in response, returned to the frontend and played without delay. In order to stimulate the learning process, the system responses must of course support the correct use of letter-sound-correspondances and discourage wrong ones. Later in the development of spelling it must support correct spellings and discourage spelling errors, in other words, be effective cues of promotion and inhibition and thus provide a relevant feedback that supports and encourages the user to apply their knowledge of letter-sound-connections when spelling. A speech generation algorithm was therefore designed with a close look to orthographic, phonetic and didactic theory. The algorithm, called Aspera³ (Articulated Spelling Response Algorithm), is presented in some detail below.

With the word completed, an encouraging greeting is given, and a new word presented. The process is spiced up with a little game logic (points and praise). The purpose is to visualize the progress of the student.

3.1 A challenging phonetics

Among the European languages, Danish is often considered to be the most vowel-rich. Approximately 39 phonetic symbols are needed

³The name Aspera is inspired by the proverb *per Aspera ad Astra*, "through hardships to the stars"

to represent the distinctive vowel sounds (compared to ≈ 18 for Swedish and ≈ 20 for Norwegian). This unusual diversity has to do with two historical developments, (i) early influence from Low German replaced the Scandinavian rolled [r] by the German velar, thereby introducing several new phonetic vowels, (ii) the tonal system (still preserved in Swedish and Norwegian) was replaced in Danish by the 'stød'-feature, also adding to the inventory of vowels (Jespersen, 1897-99, 478; Brink and Lund, 1975, I §§8-26, II §36). Even with the extra alphabetic letters Æ Ø Å, Danish orthography still has only 9 vowel letters for 39 vowel sounds. Not surprisingly, the Danish graphemes are heavily overloaded with phonetic renderings. Some examples are given in table 1.

For these reasons, among others, Danish letter-to-sound rules are unusually hard to master (for humans and NLP-applications alike). This is not good news for children at risk of developing dyslexia who often have difficulties with the so called 'phonological attention'. AiRO's didactic design pays special attention, therefore, to the vowel-related intricacies.

"rejsefeber" [rAJs0fe:!bC]
E → [A][0][e:][C]
"trestjernet" [trzsdjaR!n0D]
E → [z][a][0]
"tempererede" [tEmp0rz:!CD0]
E → [E][0][z:][C][0]

Table 1. Frequent phonetic renderings of letter E.⁴

⁴Word translations: *three starred*; *travel fever*; *tempered*.
Phonetic forms are shown in brackets. [:] is prolongation, [!] is stød (cf. the full SAMPA table at www.dsn.dk). SAMPA is IPA compatible but more keyboard friendly.

3.2 The well-formed syllable - and beyond

The Danish syllabic structure is governed by principles of phonology restricting the scope and location of the individual language sounds, very similar to the other Germanic languages (e.g. English; cf. Grønnum, 1998, chap.13). These are typical examples:

- The nasal [N] occurs only post-vocally, as in "ping" [peN] *ping*; "vinge" [veN0] *wing*; "ting" [teN!] *thing*
- [h] occurs only syllable-initially, as in "hø" [hø:] *hay*; "påhit" [pÅhid] *whim*
- Plosives [p][t][k] weaken to [b][d][g] in all positions except syllable-initially: "tip" [tib] *hint*; "skat" [sgad] *treasure*; "stærk" [sdaRg] *strong*

Certain sound combinations never occur in Danish syllables, and this fact makes them particularly suitable in the inhibitory function mentioned above. For instance, if the pupil targets the word "gnaven" (*grumpy*) by producing the letters 'N' - 'G' - 'A', the system can respond by uttering the 'impossible' syllable [Na], signalling the anomaly long before the word is completed. The 'unnatural' sound thus becomes an effective stimulus utilising the language knowledge that the child already possesses. In order to fully exploit the didactic potential of 'forbidden sounds', our speech synthesizer must of course be phonetically complete, in the sense of being able to pronounce any phone combination accurately, including those never occurring in Danish words. We call this capability **hyper-articulation**. At this time, there is no hyper-articulating speech synthesis for Danish on the market, so the AiRO project has had to develop its own voice, HYPERDAN, based on the principle of diphone resynthesis (a technology particularly suited to hyper-articulation; Henrichsen 2004).

3.3 Progressive response

Each spelling session begins with AiRO selecting a fresh target word *T* with the phonetic form *P*

(say "sofa" pronounced [so:fa]). T is presented to the pupil (with picture and sound). The pupil begins spelling (by typing 'S'), and AiRO responds with the corresponding sound ([s]).

Input	Auditive response
"S"	[s]
"O"	[so:]
"F"	[so:f]
"A"	[so:fa]

Table 2: Illustration of progressive response

In flawless sessions (such as in table 2) the spoken feedback progresses continuously, in the sense that each speech production repeats and extends the preceding one until P is met. The feedback thus provides continuous confirmation that the speller remains on the right track. This didactic approach we term **progressive response**.⁵

How are the proper input-response patterns to be computed in order to support progressive response? In the simplest case where T and P are of identical length (i.e. consists of the same number of symbols), each letter maps to a single phone (as in "s-o-f-a"). For $|T| < |P|$ (T shorter than P) some of the letters extend the spoken response by more than a single phone (e.g. "t-a-x-i" [t-A-gs-i] *taxi*). However, for $|T| > |P|$ the mapping is less straight-forward (e.g. "ch-au-ff-ø-r" [S-o-f-ø-R!] *driver*) as some of the letters do not correspond to phonetic increments in any simple way, putting the progressive response at risk. Our solution is to allow the inclusion of sub-phones in Aspera's output. Aspera may thus choose to reconstrue the phonetic form of a target word (say "hvidt" [vid] *white*) as a string of sub-phones ($[v_1-v_2-i-d_1-d_2]$) ensuring that T

⁵ Observe that the intermediate phonetic feedback (such as [so:f] in the example above) may not correspond to any known word. Even when the given (intermediate) input accidentally matches an existing Danish word T_x (e.g. 'SO' [so:!] *sow*), the phonetic feedback will not in general match T_x 's pronunciation (compare [so:] and [so:!]).

and P can still be aligned, maintaining the progressive response.

Consequently, the synthetic voice must be able to accurately pronounce sub-phones (e.g. the first and second half of phone [v] represented by $[v_1-v_2]$). The AiRO synthesis was developed with special attention to this aspect of hyper-articulation.

3.4 Polarised feedback

What happens, or should happen, when the child makes a spelling error? Consider a target word T consisting of letters $t_1-t_2-t_3-..-t_n$ and an intermediate input sequence P deviating from T , e.g. $P = t_1-t_2-p-$ (where $p \neq t_3$). The spoken feedback for P must then be clearly distinct from the feedback for $t_1-t_2-t_3-$ to provide an inhibiting effect. Here, for once, the complex Danish word-to-sound rules come in handy. Due to linguistic factors hinted at above, almost every string of letters has more than one phonologically acceptable pronunciation (if any at all).⁶ A nonsense word "hog" could thus be faithfully pronounced in Danish as [hCg], [håg], [håW], [ho:], [hOW] etc. Aspera exploits this ambiguity by always maximizing the phonetic distance between responses for correct and incorrect input (of course within the limits of phonological well-formedness). We term this principle **polarized feedback**. The phonetic distance is calculated based on the acoustic features of the individual phones. We will not pursue the details here; a journal article presenting the Aspera algorithm in formal detail is in preparation.

In case the input does not map to any phonologically acceptable pronunciation at all (say, having no vowels), Aspera's strategy is trivial: the input string then maps to the signature pronunciation of each letter (e.g. [e] for letter E; [gs] for letter X). This will necessarily produce an odd-sounding response – an inhibiting cue by nature.

⁶ This fact is a real challenge when developing Danish artificial voices, as experienced in trains, cars, call centers, home assistants, etc. where delusive pronunciations are commonplace.

4 Kindergarteners testing AiRO

AiRO was tested for the first time by kindergarteners in the Danish primary school during November 2021. Fifty kindergarteners were selected from 9 kindergarten classes. Kindergarten pupils are between 5 and 6 years old. In Danish kindergarten classrooms children are taught linguistic awareness, phonics, and reading and spelling of simple words (Juil and Elbro, 2005).

4.1 Design

We designed this testing as an effect study with an experimental group ($n=26$) and a business as usual control group ($n=24$), following Bryman (2016).

From each kindergarten classroom we selected 4-6 subjects based on their (low) scores in the national screening test (Sprogvrdering: BUVUM, 2019). Parental consent was acquired for each participating subject. The reading professional at the schools helped us evenly distribute subjects with mild and severe spelling difficulties in the two conditions of the study.

Before and after the intervention the 49 subjects' spelling and reading skills were evaluated with customized versions of screening tests developed in Engmose (2019). These test focuses on phonics applied in spelling and reading. Each subject's attention to language sounds and knowledge of letters was also assessed with standardized tests from Language Assessment 3-6 (BVUM, 2019).

4.2 Description of the intervention

Before the intervention the participating teachers and reading professionals were given a two-hour introductory course. They were introduced to the design of the study, the purpose of the intervention, and how they should instruct and assist the pupils during the intervention.

Only subjects in the experimental group had access to AiRO, while the control group received

ordinary instruction. The experimental group worked with AiRO during four weeks, four days a week, 10-15 minutes each time.

The intervention in the experimental group began with an individual introduction to AiRO and a guided practice of the first two levels. This was done by the teachers. The kindergarteners worked unattended⁷ for the remaining levels (3-16). The participating subjects could ask questions to the teacher at all times. Due to too much noise in some of the kindergarten classrooms some teachers ended up separating the children working with AiRO from the remaining classroom e.g. in a nearby smaller room.

4.3 Descriptive statistics

For both spelling and reading we compared the control and the experimental group at pre- and posttest. Table 3 and 4 show descriptive statistics for both groups (experimental and control) at pre and posttest. For each measure the number of items (#items) and minimal and maximal score values (min-max) of the scale are listed. The descriptive statistics are the number of participants (N), mean performance (M), standard deviation (SD) and range of performance (Range). Notice, that scores are calculated as how far they are from correct, meaning that lower scores are better.

Measure (#items;min-max)	M (SD)	Range	N
AiRO group			
Spelling (10;0-28)	18 (9)	41-3	23
Reading (12;0-72)	53 (9)	64-31	26
Control group			
Spelling (10;0-28)	16 (7)	29-5	20
Reading (12;0-72)	45 (18)	72-4	22

⁷ Most of the pupils found it difficult to log on to their personalized AiRO-homepage and needed help for this step throughout.

Table 3: Descriptive statistics from pretest

Measure (#items,min-max)	M (SD)	Range	N
AiRO group			
Spelling (10;0-28)	11(7)	25-1	21
Reading (12;0-72)	25 (14)	43-1	15
Control group			
Spelling (10;0-28)	12 (9)	36-0	16
Reading (12;0-72)	40 (20)	68-6	10

Table 4: Descriptive statistics from posttest

Notice in table 3 and 4 that not all 49 subjects were actually fully tested. This was due to corona-related challenges. These missing data affects the generalizability of our analysis as reported in section 4.4.

4.4 Results

For both spelling and reading we compared the control and the experimental group at the beginning and at the end of the experiment. We used paired t-test (two-tailed). In the experimental group these analyses showed significantly strengthened spelling, $t(20) = 5.127$, $p < .001$, $d = 1.12$, and reading, $t(14) = 7.566$, $p < .001$, $d = 1.95$. For the control group reading was also significantly strengthened, $t(9) = 4.312$, $p = .002$, $d = 1.36$, but spelling was not, $t(14) = 1.977$, $p = .068$, $d = 0.51$.

We used the two-way mixed ANOVA to determine whether there is an interaction effect between time of testing (pre- and posttest) and group (experimental and control). For reading we found a significant interaction effect between the two groups and time, $F(1, 23) = 8.552$, $p = .008$, partial $\eta^2 = .271$. This interaction was due to more progress in the experimental group than in the control group. For reading, the experimental group thus significantly out-performed the control group which received ordinary class teaching during the intervention period. For spelling the pattern was similar, but there was not a significant interaction effect between the

two groups and time, $F(1, 34) = 0.980$, $p = .329$, partial $\eta^2 = .028$.

5 Conclusion

As mentioned before most Danish teachers have received very little formal education about dyslexia in young children. This is one of the barriers to providing the needed support for students at risk of dyslexia or students with dyslexia in primary school. In Denmark, every second adult dyslectic report that they have never received individual offers from the education system, such as one-on-one teaching, special courses (in or outside class) or indeed personalized help of any sort (Mejding et al., 2017; Egmont 2018).

The CALL-based pedagogical approach in AiRO is a starting point for exploring new ways to support the early and later stages of reading and spelling acquisition for struggling readers.

Given the promising results from our first small experiment with kindergarten children at risk of dyslexia, we feel encouraged to develop AiRO further. We are currently making preparations for a new and updated AiRO-tool (AiRO2), capable of screening its users while servicing them, providing the teacher with status reports on the performance of the class as a whole and of the individual pupils.

Acknowledgments

We would like to express our special gratitude to all the participating teachers, reading professionals, and pupils from Sydfalster Skole and Susåskolen, Taastrup Realskole, Holmegaard-skolen, Arenaskolen, and Fladsåskolen. The AiRO project was funded by the Danish Ministry of Research (Innovationsfonden).

References

Brink, L. and Lund, J. (1975) *Dansk Rigsmål*. Part I and II. København: Gyldendals Forlag.

- Bryman, A. (2016). *Social research methods*. Oxford University Press.
- BUVM (2021). <https://www.uvm.dk/statistik/grundskolen/elever/skolestart> [visited 05.01.22]
- BUVM (2019). https://emu.dk/sites/default/files/2020-09/GSK_FællesMål_Børnehaveklassen.pdf [visited 05.01.22]
- BVUM (2019). Vejledning til Sprogvurdering 3-6 <https://www.sprogvurdering.dk/> [visited 05.01.22]
- Elbro, C., Møller, S., and Nielsen, E.M. (1995). Functional reading difficulties in Denmark. A study of adult reading of common texts. *Reading and Writing*, 7, 257-276.
- Egmont Fonden (2018). *Survey og registeranalyse - børn og unge med ordblindhed rapport*. København: Egmont Fonden.
- Ehri, L.C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, 9(2): 188-167.
- Elbro, C. and Petersen, D. K. (2004). Long-term effects of phoneme awareness and letter name training. An intervention study with children at risk of dyslexia. *Journal of Educational Psychology*, 96(4), 660-670.
- Egmont Fonden (2018). *Survey og registeranalyse - børn og unge med ordblindhed rapport*. København: Egmont Fonden.
- Engmose, S.F. (2019). *IT-støttet børnestavning: studier af børnestavnings rolle i den tidlige skriftsproglige udvikling*. Ph.D. thesis Copenhagen: Københavns Universitet, Det Humanistiske Fakultet.
- Gellert, A.S., Poulsen, M. and Elbro, C. (2018). Ordblindhed. *Samfundsøkonomen*, (1), 22-24.
- Gissel, S.T., and Andersen, S.C. (2021). A cluster-randomized trial measuring the effects of a digital learning tool supporting decoding and reading for meaning in grade 2. *J Comput Assist Learn*, 37, 287– 304.
- Grønnum, N. (1998) *Fonetik og Fonologi*. København: Akademisk Forlag.
- Henrichsen, P.J. (2004). The Twisted Tongue: Tools for Teaching Danish Pronunciation Using a Synthetic Voice. *Copenhagen studies in language*, 30, 95-111
- Jespersen, O. (1897-99) *Fonetik*. København: Det Schuboeske Forlag.
- Juul, H. and Elbro, C. (2005). *Sproglige færdigheder ved starten af børnehaveklassen. Rapport om en undersøgelse gennemført for Undervisningsministeriet efteråret 2004*. København: Center for Læseforskning.
- Mejdning, J., Neubert, K. and Larsen, R. (2017). *En international undersøgelse om læsekompetence i 3. og 4. klasse*. Rapport. PIRLS 2016.
- Messer, D., and Nash, G. (2018). An evaluation of the effectiveness of a computer-assisted reading intervention. *Journal of Research in Reading*, 41(1), 140-158.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington DC: The National Institute of Child Health and Human Development.
- Pihl, M.D. and Jensen, T. (2017). *Færre læreruddannede i folkeskolen*. Danmarks Lærerforenings hjemmeside: <https://www.dlf.org/nyheder/2017/marts/ae-raadet-hver-10-laereransatte-har-ikke-en-uddannelse> [visit 05.01.22]
- Rosdahl, A., Fridberg, T., Jakobsen, V., and Jørgensen, M. (2013). *Færdigheder i læsning, regning og problemløsning med it i Danmark*. København: SFI.
- Share, D.L. (1995). Phonological Recoding and Self-teaching: sine qua non of Reading Acquisition. *Cognition*, 55(2): 151-218.
- Solheim, O. J., Frijters, J. C., Lundetræ, K., and Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography: A group randomised controlled trial. *Learning and Instruction*, 58, 65-79.
- Vellutino, F.R. and Scanlon, D.M. (2002). The Interactive Strategies approach to reading intervention. *Contemporary Educational Psychology*, 27(4), 573-635.