Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), pages 4631–4636
Marseille, 20-25 June 2022

AiRO - an Interactive Learning Tool for Children at Risk of Dyslexia

Peter Juel Henrichsen, Stine Fuglsang Engmose
Danish Language Counsil, University College Absalon
Adelgade 119B, DK-5400 Bogense, Denmark; Trekroner Forskerpark 4, DK-4000 Roskilde, Denmark
pjh@dsn.dk, stine@pha.dk

Abstract

This paper presents the AiRO learning tool, which is designed for use in classrooms and homes by children at risk of developing dyslexia. The tool is based on the client-server architecture with a graphical and auditive front end (providing the interaction with the learner) and all NLP-related components located at the back end (analysing the pupil's input, deciding on the system's response, preparing speech synthesis and other feedback, logging the pupil's performance etc). AiRO software consists of independent modules for easy maintenance, e.g., upgrading the didactics or preparing AiROs for other languages. This paper also reports on our first tests 'in vivo' (November 2021) with 49 pupils (aged 6). The subjects completed 16 AiRO sessions over a four-week period. The subjects were pre- and post-tested on spelling and reading. The experimental group significantly out-performed the control group, suggesting that a new IT-supported teaching strategy may be within reach. A collection of AiRO resources (language materials, software, synthetic voice) are available as open source. At LREC, we shall present a demo of the AiRO learning tool.

Keywords: dyslexia, spelling acquisition, Danish, CALL, speech synthesis

1. Introduction

According to official statistics, analphabetism is now almost eliminated in the industrialized parts of the world. The authors' own country, for example, boasts a literacy rate of 99.0 percent.[1] But does this mean that all citizens feel comfortable reading and writing? Far from it. An early, but very influential study[2] found that 12 percent of adult Danes had reading difficulties hampering their professional life. Yet the problem of "functional illiteracy" has only rarely been addressed politically in Denmark, in part due to a lack of theoretical knowledge of *dyslexia* among school teachers. Until recently, dyslexia was studied only superficially in the education system, leaving teachers little prepared to engage proactively (Pihl and Jensen, 2017). As a result, too few reports from the schools on this pervasive societal problem reach the politicians.

Adults with low literacy rates are strongly overrepresented among those who have short educations, low-paid jobs or are unemployed (Rosdahl et al., 2013). Among dyslectic 25/26-year-olds, only 69% completed secondary school as opposed to 81% among peers (Egmont, 2018). Early intervention can, however, lessen the problem significantly. Vellutino and Scanlon (2002) report that special training programs for pupils from the age of 7 years can reduce 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, 2021; Elbro and Petersen, 2004). Mastering the letter-to-sound rules is the first step on the road to solid reading and spelling skills (Ehri, 2005; National Reading Panel, 2000; Share, 1995). Above all else, early intervention holds a potential for significant personal and societal gains (Gellert et al., 2018).

The ongoing project AiRO, presented in this paper, attemps to meet some of these challenges.[3] Our tenet is that young children at risk of dyslexia can benefit from a learning environment saturated with interaction and feedback that systematically and directly teaches the children the Danish spelling systems. More specifically, a child with poor command of the letter-to-sound rules will benefit from a learning situation with quick turnaround from letter production to corrective feedback. A dedicated schoolteacher can of course provide the ideal sparring, but teachers' attention is a scarce resource. Therefore, we have developed an interactive learning tool as a supplement to mainstream teaching.

In the following, we present the AiRO tool and discuss some of its underlying didactive, linguistic, and computational principles. Thereafter, we will report on our recent experiments with pupils in the Danish pre-primary school (49 subjects). Finally we discuss some future perpectives.

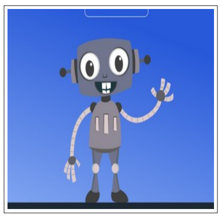

Figure 1. The AiRO frontpage (the guy is called AiRObot)

2. AiRO as a teaching assistant

Seen from the user's point of view, AiRO is a friendly agent (see fig. 1) presenting spelling tasks in a progressive fashion, beginning from trivial one-letter words and continuing (depending on the pupil's profile and performance) with ever more challenging words.

Each target word is presented to the user with a pictogram or a photograph, accompanied by its pronunciation (in synthetic speech). The user responds by spelling the target word as best she can, letter by letter. For each new keystroke, AiRO responds with an auditive rendering of

1 https://worldpopulationreview.com/country-rankings/literacy-rate-by-country (visited Jan 3 2022)

2 Elbroe et al (1995). Similar figures have been reported from other Western countries.

3 AiRO ≈ *IT-based support for children at risk of dyslexia*

the word-so-far (pronounced by a synthetic voice). As soon as the word is completed, an encouraging greeting is given, and a new word presented. The process is spiced up with a little game logic (points and praise); but most users actually find the interaction amusing in itself (see section 4 with data from our recent experimental study).

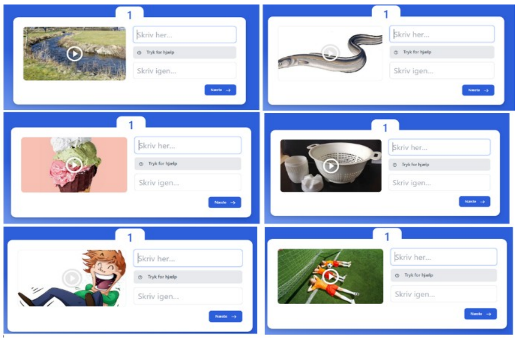

Figure 2. Presentations of target words

At the initial level, the target words are short (1-2 letters) with V, CV and VC structure (e.g. "å" *stream*, "is" *ice cream*) and straightforward pronunciation (fig. 2). Only the letters E, I, L, S, Å are used, and the letter-to-sound rules involved are basic. In general, rules trained in one level carry over to the next so that easier sounds and sound rules are trained before more difficult ones. A total of 20 letter-to-sound rules are covered. The entire course has 16 levels, focusing on the vowels and fricatives first, then gradually introducing the plosives.

3. AiRO's technical design

Each letter entered by the pupil is sent to the back-end application for analysis (correctness, response time, and other metrics). Based on the analysis, a sound file (synthetic speech) is generated and returned to the front end for immediate playback. In order to stimulate the learning process, the auditory responses must of course support the correct spelling and discourage spelling errors, i.e. be effective cues of promotion and inhibition. The speech generation algorithm, called Aspera,[4] was therefore designed with a close look to orthographic, phonetic and didactic theory. Aspera (Articulated Spelling Response Algorithm) was developed specifically for the AiRO project and is presented in some detail below (the busy reader could skip to section 3.6 for a summary).

3.1 The challenge: Danish sounds and letters

Among all the European languages, Danish is arguably the most vowel-rich. Approximately 38 phonetic symbols are needed to represent the distinctive vowel sounds (compared to ≈20 for Norwegian, ≈18 for Swedish, and ≈14 for English). This diversity has to do with two historical factors, (*i*) early influence from Low German swopping the Scandinavian rolled [r] for the German velar, thereby introducing several new phonetic vowels, (*ii*) the gradual transformation of the toneme-1 syllables (preserved in Swedish and Norwegian) into stød-wovels, adding to the vowel inventory (Jespersen, 1897-99, 478; Brink and Lund, 1975, I §§8-26, II §36). Even though these language changes have been accommodated by introducing extra alphabetic symbols (Æ Ø Å), Danish orthography still has only 9 vowel letters for 38 sounds. Not surprisingly, the graphemes are heavily overloaded with phonetic renderings. Fig. 3 is an example.

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

Figure 3. Frequent phonetic renderings of letter E.[5]

For these reasons in particular, and for other reasons as well, Danish letter-to-sound rules are unusually hard to master (for humans and NLP-applications alike). This is bad news for children at risk of developing dyslexia, often facing problems with socalled 'phonological attention'. AiRO's didactic design pays special attention, therefore, to the vowel-related intricacies.

3.2 Well-formed syllables - and beyond

The Danish syllable is organized by phonological principles restricting the scope and location of the phones (i.e. the individual language sounds). Most rules of Danish phonology are similar to other Germanic languages (e.g. English; cf. Grønnum, 1998, chap.13). These are examples:

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

The fact that certain sound combinations never occur in Danish syllables makes them particularly suitable in the inhibitory function mentioned above. For example, if the child targets the word "gnaven" (*grumpy*) but begins with N-G-A-, the system responds by uttering the 'impossible' syllable [Na], and the mistake becomes clear long before the word is completed. The unnatural sound thus acts as 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 be phonetically complete, in the sense that it is able to pronounce any phone combination accurately, including ones never occurring in Danish words. We call this capability **hyper-articulation**.

3.3 Hyper-articulation

At this time, there is no hyper-articulating speech synthesis for Danish on the market; so the AiRO project had to develop its own. Our voice, called HyperDan, is

4 inspired by the proverb *per Aspera ad Astra*, "through hardships to the stars"

5 Word 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).

built on the principle of diphone resynthesis (CMU Festival, www.festvox.org). This technology, though a bit out of fashion, allows an outstanding control over the linguistic parameters (Henrichsen 2004) and was therefore the natural choice for HyperDan. We created a library of 4761 diphones: the fully instantiated matrix of 69 phones (including the pseudo-phone *pause*). This unusually rich distribution with four recorded variants for each vowel phone (±prolongation and ±stress) was needed to meet the requirements of hyper-articulation: phonetic completeness[6], prosodic control, ultra-low speech rates (as an option). Not surprisingly, the vowel system is thus heavily over-represented in the HyperDan diphone library in comparison with other diphone-based voices for Danish (typically using 800-1500 diphones only).

3.4 Progressive response

As mentioned in section 2, a spelling session begins with AiRO selecting a target word *T* with the phonetic form *P* (say "sofa" pronounced [so:fa]). *T* is presented to the pupil (as picture and sound). The pupil begins spelling by typing 'S', and AiRO responds with the corresponding sound [s]. From there the session continues as follows.

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

In flawless spelling sessions (such as this one) the spoken feedback progresses continuously, in the sense that each speech production extends the preceding one until *P* is met, at which point AiRO concludes the session with an acknowledging comment ("Good job!" or similar). The spoken feedback thus provides continuous confirmation that the speller is still on the right track. This didactive approach we term **progressive response**.

How are the proper input/response patterns to be computed in order to support progressive response? In the simplest case where *T* and *P* have the same length (i.e. 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* and *P* are still alignable (maintaining the progressive response). Consequently, the synthesiser must be able to accurately pronounce even sub-phones (e.g. the first and second half of phone [v] represented by [v_1-v_2]). The AiRO synthetic voice was developed with special attention to this aspect of hyper-articulation.

3.5 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 Þ deviating from *T*, e.g. Þ = t_1-t_2-þ- (where þ ≠ t_3). The spoken feedback for Þ 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).[7] A nonsense word "hog" could thus be faithfully pronounced [hCg], [hååg], [hååW], [hååW!], [håå:!W], [ho:!], [hOW] etc. Aspera exploits this ambiguity by always maximizing the phonetic distance between responses for correct and incorrect input (within the limits of phonological well-formedness). We term this principle **polarized feedback**. The phonetic distance is calculated based of the acoustic features of the individual phones. We will not pursue the details here; an extended paper is in preparation presenting the Aspera algorithm in technical details for computational linguistst.

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.

3.6 AiRO design principles - a summary

We are now in a position to state the main principles underlying the AiRO tool's technical design.

Seen from the child's point of view, working with AiRO must be simple, amusing and instructive at the same time. Our users are very young and not trained (or inclined) to follow formal instructions. This means that the AiRO-application must be largely self-explanatory.

From a programmer's perspective, these requirements apply.

- **Orthographic constraints**: None. Aspera must convert any input (string of letters) to a phonetic form, phonologically well-formed when possible.
- **Phonetic constraints:** None. The AiRO speech synthesis must be able to articulate any phonetic form (even if phonologically ill-formed).
- **Didactic constraints:** The auditory response must provide relevant cues (promoting and inhibiting), guiding the spelling process closely.

The phonetic productions of Aspera (driving the synthetic voice) must comply with the formal principles of hyper-articulation, progressive response, and polarised feedback.

6 The stød-vowels (SAMPA symbol [!]) are emulated with a sharp F0-drop. Consider the sequence of pitch values 110-110-115-120-104-120 in the scm-table for 'ting' [teN!]:

```
(voice_hyperdan) (set! utt1 (Utterance Segments (
 (pau 0 (0 110)) (t 0.0725 (0 110)) (e 0.12 (0 115))
 (N 0.065 (0 120)) (N 0.025 (0 104)) (pau 0 (0 120))
))) (utt.synth utt1) (utt.save.wave utt1 "ting.wav")
```

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

- **Hyper-articulation**
  - The synthetic voice can accurately pronounce any phone string (even if 'anomalous')
  - The synthetic voice can pronounce any subpart of a phone (allowing sub-phones in *P*-forms)
- **Progressive response**
  - Flawless spelling is confirmed by spoken responses continuously extending towards *P*
- **Polarized feedback**
  - Spelling errors trigger spoken responses maximally distinct from the correct lettering

4. AiRO's first days at school

During November 2021, AiRO was tested for the first time by pupils in the Danish primary school. Forty-nine pupils (and their participating teachers) were selected from nine kindergarten classes, in Denmark children 6 years of age. Kindergarten pupils are taught linguistic awareness as well as reading and spelling of simple words (Juul and Elbro, 2005).

4.1 Experimental setup

Our experiment was designed as an effect study with an experimental group (*n*=26) and a control group (*n*=23), following Bryman (2016). We selected 4-6 subjects from each class based on their (low) scores in the national screening test (Sprogvurdering: BUVM, 2019). The participating teachers helped us evenly distribute subjects with mild and severe spelling difficulties in the experimental and control groups. Parental consent was acquired for each participating subject.

The 49 subjects' spelling and reading skills were evaluated before and after the survey with customized versions of screening tests developed in Engmose (2019). Each subject's attention to language sounds and knowledge of letters was also assessed with standardized tests from Language Assessment 3-6 (BVUM, 2019). The subjects worked with AiRO during four weeks, four days a week, 10-15 minutes each time.

All subjects were tested before and after the intervention period. Only subjects in the experimental group had access to AiRO, while the control group received ordinary instruction.

4.2 The intervention period

The teachers were given a two-hour introductory course explaining the purpose of the intervention, the experimental setup, and their own commission (including how to instruct and assist the pupils during the experiment).

The pupils were then introduced to AiRO individually by a teacher and guided through the first two levels. For the rest of the levels (3-16) the pupils worked unattended.[8] The pupils could ask questions to the teacher at all times.

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

4.3 Experimental data

Table 1 and 2 show descriptive statistics for both groups (experiment

control) at pretest and posttest. For each measure the number of items (#items) is shown, minimal and maximal score values (min-max), number of participants (N), mean performance (M), and range of performance (Range). Scores are calculated as how far they are from correct, meaning that lower scores are better.

Notice in table 1 and 2 that not all 49 subjects were actually fully tested (viz. pretested for spelling and reading, posttested for spelling and reading). This was due to corona-related challenges.

Measure (min-max, #items)	N	M	Range
Experimental group			
Spelling (0-28, 10)	23	17.4	11.6-23.2
Reading (0-72, 12)	26	52.3	42.0-60.0
Control group			
Spelling (0-28, 10)	21	17.8	16.8-19.0
Reading (0-72, 12)	22	51.0	45.0-56,0

Table 1. Data from pretest

Measure (min-max, #items)	N	M	Range
Experimental group			
Spelling (0-28, 10)	23	8.3	4.2-11.6
Reading (0-72, 12)	17	35.7	27.0-43.0
Control group			
Spelling (0-28, 10)	21	11.1	5.2-17.2
Reading (0-72, 12)	14	40.0	32.0-49.0

Table 2. Data from posttest

4.4 Results and observations

Preliminary analyses of our test data (summarized in table 1 and 2) indicate that the experimental group's spelling and reading competence was significantly strengthened[9] during the test period, especially concerning the attention to the basic letter-to-sound rules (the ability to aplit words into their component sounds). The experimental group also significantly out-performed the control group which received ordinary class teaching during the intervention period. Last but not least the pupils found it amusing to work with AiRO, as reported by the participating teachers, who also reported that the AiRO activites could be accommodated without disturbing the rest of the class.

9 For spelling: $t(23)=13.0$, $p<.001$ (two-tailed); for reading: $t(17)=68.1$, $p<.001$ (two-tailed).

5. AiRO resources to share

A number of databases and program modules have been developed in the AiRO-project, including

- MISPEL, a dictionary of Danish misspelled words produced by dyslectic children and adults
- DYSPRO, a catalogue of typical learning profiles observed in Danes with writing difficulties (expressed as computer-readable rules)
- Resource files for HyperDan (synthetic voice for Danish meeting the requirements of hyper-articulation)

These resources will be made publicly available (open-source) by September 2022. Contact the authors for more info.

6. Concluding remarks

As mentioned before, dyslexia and dyslectic pupils tend to be ignored both in school and in the public debate. Most teachers have received very little formal education about dyslexia in young children, and they do not know how to respond to their own observations of spelling difficulties, especially in the very first classes where didactic intervention is most effective. Consequently, the government is not being sufficiently informed about the consequences of not taking action. And for that reason, most teachers receive very little education about dyslexia in young children. This deadlock is not at all new, and no remedy is in sight.

The AiRO project group wants to make a contribution by exploring new teaching strategies in the early stages of reading acquisition. Given the positive response from the schools and the promising results of our experiments with very young schoolchildren, 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. Some early signs of dyslexia are not difficult to detect automatically (can be done with machine learning or even simple rule-based methods).

In Denmark, every second adult dyslectic has 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). It's about time that we relieve the burden of dyslexia for everyone – because we can.

Acknowledgements

The AiRO project was made possible by a grant from the Danish Ministry of Research (Innovationsfonden), for which we are grateful. We would like to express our special gratitude to all the participating teachers and pupils from Arenaskolen, Fladsåskolen, Holmegaard-skolen, Taastrup Realskole, Sydfalster Skole and Susåskolen.

Bibliographical 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. thesisCopenhagen: Københavns Universitet, Det Humanistiske Fakultet.

Gellert, A.S., Poulsen, M. and Elbro, C. (2018). Ordblindhed. *Samfundsøkonomen*, (1), 22-24.

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 Schubotheske 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.

Mejding, J., Neubert, K. and Larsen, R. (2017). *En international undersøgelse om læsekompetence i 3. og 4. klasse*. Rapport. PIRLS 2016.

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.

Vellutino, F.R. and Scanlon, D.M. (2002). The Interactive Strategies approach to reading intervention. *Contemporary Educational Psychology*, 27(4), 573-635.