British Journal of Education, Society & Behavioural Science
4(11): 1504-1524, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

A Randomized Controlled Trial of a Large Scale Systematic and Explicit Language and Literacy Intervention in Danish Daycares: The SPELL Study

Dorthe Bleses[1*], Anders Højen[1], Line Dybdal[2], Philip Dale[3], Laura Justice[4], Shayne Piasta[4], Justin Markussen-Brown[1], Marit Clausen[1], E. F. Haghish[1,5], Burcak Aktürk Ari[1], MetteKjær Andersen[1] and Werner Vach[5]

[1]Center for Child Language, University of Southern Denmark, Odense, Denmark.
[2]Rambøll Management Consulting, Aarhus, Denmark.
[3]Department of Speech and Hearing Sciences, University of New Mexico, Albuquerque, USA.
[4]Crane Center for Early Childhood Research and Policy, The Ohio State University, Columbus, USA.
[5]Clincial Epidemiology, Center of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Germany.

Authors' contributions

Authors may use the following wordings for this section: This work was carried out in collaboration between all authors. Authors DB, AH, WV LD, PD, LJ and SP developed the study idea, designed the study and developed the protocol. Author WV designed the statistical analysis plan, Authors DB and WV drafted the manuscript. All authors have been involved in the development of measurement instruments and data collection, have revised the manuscript critically and have given their final approval of the version to be published.

Study Protocols

Received 29th May 2014
Accepted 25th June 2014
Published 14th July 2014

ABSTRACT

Background: Children's early language and literacy skills are key predictors of later educational outcomes. Children from low socioeconomic status and immigrant

**Corresponding author: Email: bleses@sdu.dk;*

backgrounds have elevated risk for adverse educational outcomes. Research suggests that systematic and explicit curriculum-based language and literacy interventions may have positive impacts on children's outcomes. However, the majority of this research has consisted of efficacy trials conducted in the U.S. on a relatively small number of children from low-income homes and focussed on efficacy instead of effectiveness. As a large-scale effectiveness trial, the Structured Preschool Efforts for Language and Literacy *(SPELL)* study will test the generality of the value of systematic and explicit instruction by evaluating the effects of an intervention in Denmark conducted under "real-life" circumstances including all children in the participating daycares addressing questions both relating to the efficacy and effectiveness of the intervention.

Methods: The *SPELL* study is a cluster-randomized trial of a language and literacy intervention which uses storybook reading with systematic and explicit instruction to promote language and literacy skills in Danish children. Based on sample size calculations, we intend to include 128 daycare centers, each with on average three classrooms and 20 children/two daycare educators in each classroom, summing to 7.680 children and 768 educators. Daycare educators will implement a 20-week intervention which provides an explicit scope and sequence of language and literacy instruction over 40 lessons. The potential added value of coupling *SPELL* with an expanded professional development program and with a supplemental program implemented by parents will be tested. The primary outcome is the change of language scores from pre to post intervention in all participating children. A range of secondary outcomes and moderator analyses will be conducted.

Discussion: It is uncertain how findings from studies of systematic and explicit interventions in the US will translate into non-US societies, and how effective such interventions are at a large scale. This study will be the first large scale study to address both questions. The findings will indicate the generality of the effects of systematic and explicit instruction in a non-US context and identify implementation factors that affect the efficacy of language and literacy interventions when carried out at scale.

Keywords: Language and literacy intervention; daycare and home setting; randomized controlled trial; large-scale effectiveness study.

Registration: (ISRCTN04394487)

1. BACKGROUND

A recent systematic review has shown that children's early language and literacy skills are key predictors of later educational outcomes, particularly in reading achievement [1]. Children who arrive at formal schooling with well-developed language and literacy skills have better academic trajectories than children with lags in these areas [2,3]. Notably, some children exhibit an elevated risk for adverse educational outcomes due to the presence of family related factors that affect early language and literacy skills, such as having parents with low-income or low educational background or having parents with an immigrant background. Child-related factors such as language impairment or having a developmental disability may also increase the risk for language delays. Bleses and colleagues have shown that children facing social disadvantage in Denmark perform poorly on a battery of language and literacy tests compared to other children, with half of bilingual immigrant children in kindergarten performing more than one standard deviation below the mean of monolingual children [4]. By the end of primary school about 40% of bilingual immigrant children in

Denmark have no functional reading skills as indicated by PISA results [5]. A 30 year follow-up study in Denmark of children with significant language delay at school entry demonstrated disadvantaged trajectories in later life outcomes with reported literacy difficulties, unemployment and low socioeconomic status at rates significantly higher than in the general population [6]. The language and literacy development in young children can therefore be regarded as a central public health issue that warrants preventive initiatives, and there is a pressing need to improve early childhood practice with the overall aim of promoting language and literacy development in Danish children.

Based on research that has demonstrated the critical importance of early life experience for establishing the brain architecture necessary for future language and literacy (and other) development, and that such abilities are responsive to environmental enrichment [7], early childhood education research has investigated how children's acquisition of language and literacy skills can best be enhanced through early interventions [8]. A new meta-analysis of the effects of 84 daycare programs, nearly all of them focusing on children from low-income families, estimated an average effect size of .35 standard deviation, with substantial variation between studies [9].(Note that we use the term "daycare" to refer to any early childhood program serving children ages 3-5. Similarly, we use the term "educator" to refer to any adult serving children in early childhood programs.) A critical element in effective interventions in daycares appears to be the provision of stimulating and supportive interactions between daycare educators and children in which daycare educators elicit verbal responses and actions from children, and foster engagement in learning based on an effective use of curricula that support systematic and explicit instruction [10,11]. Systematic instruction utilizes learning sequences that are guided by a rigorous understanding of how children learn; explicit instruction makes it clear to children what they have to attend to within an activity by orienting them towards the goals of the activity, thus helping them to engage in the learning process. Similar to the meta-analysis referenced above another recent evaluation of specific daycare curricula has documented substantial variations in effects across studies indicating that developmentally-focused research-based curricula (as opposed to commercial curricula) have higher positive impact on children's language and literacy development (effect sizes ranged from –.38 to .68 across all curricula) [12]. Evidence is mixed as to whether children from low-income backgrounds benefit the most from early language and literacy interventions compared to children from more advantaged backgrounds or vice versa [1,10]. Similar results have emerged from studies of language and literacy interventions in home settings. A meta-analysis of family literacy programs found that shared book reading combined with additional literacy activities showed a mean effect size of .21 on both language and literacy outcomes in children, with some variation across studies (effect sizes ranged from .14 to .28). Parallel language and literacy activities in both daycares and homes may augment the effects of early interventions [13]. However, additional support such as modelling of positive interactions and feedback may be needed for parents to implement literacy interventions with success [14].

Inspired by Bronfenbrenner's bio-ecological model of development, in which the child is considered at the centre of reciprocal social and cultural influences, research has identified a number of environmental factors that impact the effectiveness of language and literacy interventions. Structural quality in daycares, such as the number of children in a classroom, the ratio of daycare educators and children, staff qualifications and the quality of the psychical learning environment (book materials, literacy area, writing materials, etc.) all appear to create the conditions in which high quality adult-child interactions can take place (process quality), but such structural factors are not sufficient to ensure the effectiveness of interventions [10,15-17]. Children's interactions with educators, peers and tasks may also

affect possible outcomes; in particular, children's engagement with tasks is predictive of language and literacy outcomes [18]. Children's exposure to peers' language and literacy competences is associated with outcomes, i.e. children with less-developed skills in language and literacy may benefit little from enrolment in classrooms in which their classmates also exhibit less-developed skills [19].Finally, associations between the quality of the home literacy environment and child outcomes indicate that parents with higher socioeconomic status are more apt to provide an enriched environment with stimulating learning materials and language and literacy practices that promote stronger language and literacy skills in their children [20-23].

Systematic and explicit interventions delivered with high intervention fidelity appear to be of particular benefit for children with initially low language and literacy skills, whereas children with better initial skills benefit more from child-managed activities [24]. However, an important challenge emerges from findings indicating that, whereas educators show high fidelity in terms of using specified materials, following lesson plans, etc., the quality of their interactions are generally lower which thereby reduce the intended benefits [25]. Randomized controlled trials studies indicate that practice-focused professional development aiming at enhancing the quality of interactions through the provision of specific knowledge of effective practice, practice in detecting effective adult-child interactions through video analyses, and knowledge of children's language and literacy skills is successful in changing the behaviour of daycare educators [26]. Relatively few intervention studies use intervention fidelity data effectively to investigate the extent to which fidelity is associated with outcomes [27,28], and the extent to which low implementation fidelity is responsible for differences in effect sizes across intervention studies is not clear.

In summary, the body of research suggests that systematic and explicit curriculum-based language and literacy interventions have positive impact on children's outcomes. However, the majority of this research has been conducted in the US and it is unclear how such findings will translate into societies like Denmark for a number of reasons. First, in Denmark 97% of children between age three and five are in public daycare settings, but the Danish daycare system has traditionally emphasized a broader-based learning context to a large extent based on child-initiated activities and with little systematic and explicit instruction and little focus on literacy skills. Secondly, the professional development of daycare educators has been largely theory-oriented, with little focus on language and literacy practice. Third, the role of the home learning environment may have a reduced role in Denmark when compared with the U.S., given that 85% of Danish children already attend daycare at one year of age, whereas the attendance rate is much lower among American children. In general, socioeconomic status (SES) variation in the Danish population as a whole is less extreme than in the US. Additionally, the majority of rigorously conducted studies of developmentally focused curricula have had extensive involvement of researchers in the intervention, a relatively small numbers of children from mainly low-income homes and a focus on demonstrations of efficacy. The effectiveness of curricula implemented under "real-life" circumstances and at a large scale is not clear, given the well-known difficulties in taking interventions to scale [10,29].

In this study we will therefore test the generality of systematic and explicit instruction by evaluating the effects of an intervention called SPELL (Structured Preschool Efforts in Language and Literacy) in Denmark which has a pedagogical and educational context different from the American context. It is a large-scale effectiveness study conducted under "real-life" circumstances including all children in the participating daycares independent of parental background. The *SPELL* intervention is adopted from the developmentally focused

research-based curriculum Read It Again (RIA). RIA is a 30-week, 60 lesson curriculum-based intervention that targets a systematic and explicit progression of skills in four language and literacy domains (phonological awareness, vocabulary knowledge, print knowledge and narrative ability) based on existing scientific knowledge. These learning domains represent high-priority instructional targets as they are consistently linked to later developments in word recognition and reading comprehension [1]. The choice of *RIA* as basis for developing the *SPELL* intervention is motivated by the fact that prior feasibility studies of RIA have shown medium to large-sized effects for language and literacy skills in children [30]. Additionally, *RIA* builds upon a large literature demonstrating the value of shared storybook reading in promoting language and literacy learning, a practice that to some extent is already used in Denmark and that can be readily embedded within most general curricula. Finally, *RIA* is designed to be used for children at-risk through an explicit focus on critical skills that relate to immediate and future academic success and was specifically designed to be used "at scale" in the sense that it is a low-cost, easily accessible program requiring little professional development support, and can be readily implemented in Danish daycares.

2. TRIAL OBJECTIVES

Our primary aim is to use a causally interpretable design (randomized controlled trial) to evaluate the impact of the *SPELL* intervention on children's language and literacy development, particularly in children at-risk due to low socioeconomic or bilingual immigrant backgrounds, and on daycare educators' process quality. The theory of change in *SPELL* is that children's outcomes will be improved by providing systematic and explicit instruction addressing high-priority language and literacy domains in a way that is developmentally targeted to the individual child and ensure repeated exposure. Additional, it is anticipated that repeated self-reflections by the educator on the actual implementation of the critical elements of the intervention, as well as educator's repeated monitoring of the child's progress will support implementation fidelity and the targeting of the intervention to the needs of the individual children and therefore children's outcomes. We hypothesize to achieve similar effects with *SPELL* as have been reported for *RIA*. Our study is designed as an effectiveness study, focusing on a real-life setting, recognizing that lack of fidelity in implementation of the intervention may affect the effect sizes negatively. We will however, address the effect of the interventions in a Danish context when delivered with high efficacy as well.

A second aim is to examine the potential added value of coupling *SPELL* with an expanded professional development program for daycare educators. The theory of change is that children's outcomes will benefit from helping daycare educators identify children's varying language skills, and by providing daycare educators with knowledge about and practice in identifying and using individually targeted instructional interactions. We hypothesize that those children, whose daycare educators receive extra professional development which emphasizes the use of strategies to differentiate instruction across the full range of child abilities, will make significantly increased language and literacy gains compared to *SPELL Basic*.

A third aim is to examine whether coupling *SPELL* with home-based companion *SPELL* activities implemented by parents will add value. The theory of change is that the higher dosage of the intervention through aligned exposure in both daycare and home settings will improve children's outcome. We hypothesize that children whose caregivers implement *SPELL* home companion activities, featuring two weekly adult-child storybook reading

lessons with targeted language and literacy activities will make significantly increased language and literacy gains compared to *SPELL Basic*.
Throughout the project we will also test the hypotheses that a range of factors at daycare centre-, daycare educator-, parent- and child-level will moderate the effectiveness of the intervention. One goal of these analyses will be to determine which children benefit most from the intervention and under which circumstances. We address how contextual factors in the daycare centers affect implementation fidelity. Based on the literature, we will investigate how specific features of classroom quality (structural and process) and the composition of classrooms in terms of child background and language status is associated with the delivery of the intervention and children's gains in specific domains. Similarly, we will investigate whether the intervention has differential effects on child outcomes depending on children's parental backgrounds and home learning environment. Finally, we will map the association among various implementation fidelity measures and the outcomes achieved.

3. METHODS AND ANALYSES

3.1 Background and Rationale

The *SPELL* study is a cluster-randomized trial of a language and literacy intervention which uses storybook reading with systematic and explicit instruction to promote language and literacy skills in Danish children. We aim to assess the effects of the *SPELL* intervention as implemented in daycares when implemented alone, and also to test added benefits when combined with a) additional professional development of daycare educators focussed on differentiating implementation for children, and b) parent-implemented *SPELL* activities aligned with the daycare educator-implemented intervention. Usual practice in daycares is used as comparator. Based on sample size calculations we will recruit 128 daycare centers with a target sample size of 7.680 children from several municipalities in Denmark. Major parts of the design and methods used in the trial are similar to another randomized trial (LEAP) that is currently being implemented.

3.2 Study Population and Eligibility Criteria

All children of age three and above (in practice up to age six) in the recruited daycare centers will be eligible for participation. Centres to be invited will be purposely sampled to over-represent daycares with high concentrations of children at-risk for language and literacy delays based on social disadvantage and immigrant status. The only exclusion criterion pertains to children with developmental delays that may be integrated in daycares.

3.3 Experimental Intervention

To fit into the Danish educational context, some modifications and further development of *RIA* have been made, mainly that a) *SPELL* is reduced in length as a 30 week program is not judged feasible by practitioners, and b) *SPELL* is implemented in small group settings instead of large group settings. The basic adaptation of *RIA* to *SPELL* has been performed in consultation with an international research group (which includes the original developer of *RIA*) and details of the adaptation (choice of books, instructions for use) were further developed in a collaborative process with daycare educators in order to integrate practice-oriented feedback from the end-users. *SPELL* and the experimental intervention's four arms are described below.

3.3.1 SPELL Basic

Daycare educators will implement a 20-week intervention which provides an explicit scope and sequence of language and literacy instruction over 40 lessons implemented at an intensity of two per week. Each lesson is implemented in a small-group format and lasts approximately 30 min. The same set of 10 high quality story books will be used repeatedly, i.e. each book will be read four times as the vehicle for instruction. Educators will access the lessons using iPads with an app featuring lesson plans and a logging tool (implementation notes). Each lesson involves two components: 1) *SPELL lesson plan*, which soft-scripts a sequence of step-by-step instruction featuring a *before-, during-, and after-reading* activity as well as suggested language that daycare educators can use to support children's learning during each activity, and 2) Learners' Ladder, which presents specific differentiation strategies to differentiate instruction within the lesson on the basis of individual children's response to the lesson. Four approaches (three of them representing further development of *RIA*) are used to promote daycare educators' implementation fidelity. First, each daycare educator participates in a two-day introduction course prior to implementation. The introduction course will be based on a practice-focused approach in which daycare educators receive step-by-step guidance in how to implement *SPELL* lessons and activities. Second, in addition to a paper version daycare educators receive access to the *SPELL* manual on an iPad, which provides all 40 sequence lesson plans, and will receive the set of 10 books used with *SPELL*. Third, daycare educators maintain implementation notes on the iPad for 1) documentation of use of critical elements in the intervention, individual children's overall engagement, the experiences of individual children, and record their usage of specific differentiation strategies towards individual children and 2) for tracking individual children's progress towards each *SPELL* objective three times during the 20-week-study period. Both elements also form the basis for the daycare educator's self-reflections on the implementation of lessons. Fourth, each lesson plan, including a learning ladder with examples of differentiation strategies targeted to the learning objectives of each lesson, will be developed in two formats; a full version for preparation and a short version only containing keywords and brief introduction to the activity which can be used during each lesson.

3.3.2 SPELL Basic + Expanded Professional Development (SPELL Basic + EPD)

Daycare educators in this arm will implement all *SPELL Basic* components as described in the preceding section, but will also receive extra professional development to better support differentiation of instruction for at-risk children. Specifically, these daycare educators will receive: a) an additional, intensive two-day training (in weeks three and six of the intervention) to increase their substantive knowledge regarding the development and instructional needs of children at-risk for language and literacy problems, bilingualism and cultural sensitivity, and the differentiation strategies included in *SPELL*. The educators will receive a language profile based on the pre-test data of each child in their groups (each educator will have two *SPELL*-groups), specific training in recognizing and practicing differentiation strategies by watching others and own video recordings of *SPELL* lessons; and a weekly refresher session in which the two educators working in each classroom conduct small assignments (e.g. watching videos with the aim of recognizing and discussing the implementation of differentiation strategies or the engagement of individual children in the intervention) to assist implementation fidelity.

3.3.3 SPELL Basic + Learning activities in the home (SPELL Basic + Home)

Daycare educators in this arm will implement all *SPELL* components as described for the *SPELL Basic* arm. In addition, parents are asked to implement *SPELL* activities two times a week at home to complement the *SPELL* intervention that children receive in their classrooms. Parents in this arm will receive companion materials (the same 10 books, 20 instructional lesson plans for each book reading that address the same learning domains and learning targets as in *SPELL Basic*, and a reading calendar with stickers to mark each reading session). The home activities are completely aligned in time with the SPELL lessons in the daycare centers in terms of books and the learning objectives to be addressed. Parents are provided with a small manual and will have access to videos that illustrate how parent-led storybook reading at home using the *SPELL* lesson plans can support children's language and literacy development. All materials will be translated into the most commonly spoken foreign languages in the included municipalities. Implementation will be tracked via weekly structured diaries completed by the parents.

In these three active intervention arms, all manuals and test materials will be easily accessible through an internet-based platform (Rambøll Results, see below) in order to limit practical obstacles to implementation.

3.3.4 Business as usual(Control)

Daycare educators in the Control arm will not receive the *SPELL* intervention materials or *SPELL* training. Rather, they will continue to provide the types of instruction typical in Danish childcare centres. Daycare educators will receive one day of training on social inclusion relevant to their daily educational practice.

3.4 Study Conduct

The study uses a multilevel cluster-randomized trial utilizing a four-arm pretest-posttest design to estimate the impact of SPELL on child outcomes and daycare educator outcomes. In two sequential waves, daycare centers will be randomly assigned to one of four arms: Control, SPELL Basic, SPELL Basic + EPD, or SPELL Basic + Home. Pre-test and post-test data will be collected for children and daycare educators, and children will be assessed in a one and two-year follow-up.

The study started in summer 2012 with selecting the participating daycare centers for the first wave. Daycare centers are informed in November 2012 about the arm they should implement. Baseline questionnaire data at daycare centre-, daycare educator-,parent- and child level will be collected in November 2012. Video recordings of the daycare centre activities will also be collected at this time. In January 2013 all daycare educators assigned to one of the SPELL arms receive two days of initial training in the implementation of the intervention and receive their iPads. All daycare educators will perform pre-test language and literacy assessment of all participating children at this time. The intervention starts in the fourth week of January, 2013. Video recordings of lessons will be performed three times. A post-test language assessment, post-questionnaire data collection and video recordings will be made in May and June 2013. The second wave of the study will start in the autumn 2013 following the same plan and ends in May 2014. Further assessments of all children are planned in a one-year and two-year follow-up.

To support study implementation a monitoring and evaluation system, Rambøll Results, has been developed especially for delivering real time data on performance at different levels of aggregation to use in managing implementation and performance. This will allow for continuous feedback to project leaders in the municipalities in order to support implementation of the intervention in the daycare centers.

3.5 Outcomes and Data Collection

A variety of measurement instruments will be used for obtaining information on outcomes, moderators and fidelity (see overview of measurement instruments in Table 1 and variables to be extracted for the analysis in Table 2).There are five different sources of these measures: a) language assessments using structured tests performed by the educators; b) questionnaires to be completed by parents, educators and daycare centre leaders; c) observational ratings of activities in classrooms, *SPELL* lessons and the structural environment; d) implementation notes and progress monitoring checklists completed by daycare educators and parents as part of the intervention; e) background information made available from Statistics Denmark.

3.5.1 Language and literacy assessment

The language status of the children is assessed by an instrument which is a minor adaptation of Language assessment of children: 3-6 [31]. The instrument consists of seven language and literacy subscales: vocabulary, language comprehension, sound discrimination, rhyme, word-, syllable- and phoneme deletion, letter identification and communicative strategies. Sound discrimination is only assessed in three-year-old-children; rhyme, deletion and letter identification only in children of age four and above. The three subscales sound discrimination, rhyme, and deletion will be summarized into one measure of phonological awareness. Finally we summarize all seven subscales into two measures: A proximal measure, including phonological awareness and letter identification, and a distal measure including vocabulary, language comprehension and communicative strategies. A total combined score based on all subscales will be computed. The language and literacy assessment is a revised version of the language and literacy assessment battery that is administered by educators in a vast majority of Danish municipalities as part of a national screening program.

3.5.2 Questionnaires

Several questionnaires will be used to collect background information. Daycare educators are asked to fill out the well-known SDQ [32]. All daycare centre leaders and educators will also be asked about various aspects of the implementation context, including structural, cultural and organizational quality. In addition, daycare educators are asked questions related to the implementation context and program feasibility using two researcher-developed instruments developed by ourselves with inspiration from the implementation literature [33-36]. All parents are asked to complete a home learning environment questionnaire (*HLEQ*). The *HLEQ* is experimenter-developed but inspired by a range of home learning environment instruments which have been used in other research projects examining the home literacy environment [37-43]. Data from all these questionnaires will be aggregated with the intention of quantifying the dimensions given in Table 2. Internal consistency will be addressed using psychometric analyses such as factor analysis.

Table 1. Project measure table, Description of major measurement instruments

Measure	Construct	Psychometric quality	Administration
Revised LA: 3-6. Revised Language Assessment Test: 3-6 [31]	Language and literacy development	Internal consistency coefficient (Cronbach's alpha) for subscales across studies= .64 - .89; Correlations between subscales=.25-.70. Revised as part of project. Internal consistency coefficient (Cronbach's alpha) and concurrent validity will be established	A test battery administered by daycare educators
SDQ. Strengths and difficulties questionnaire [32]	Psycho-social development	Internal consistency coefficient (Cronbach's alpha) for subscales across studies= .73; Cross-informant correlation=mean: 0.34, Retest stability after 4 to 6 months=.62)	Questionnaire completed by educators
HLEQ. Children's learning in the home	Quality of home learning environment	Developed as part of project. Internal consistency coefficient (Cronbach's alpha) for subscales= .65 - .81	Questionnaire completed by parents
CLASS. Classroom Assessment Scoring system. PreK [44]	Process quality of language and literacy instruction	Inter-rater reliability (within 1)=.78-.93; Stability coefficient (test/ retest)=.84 - .91; Internal consistency coefficient (Cronbach's alpha) for subscales across studies=.76 to .94; Concurrent validity with ECERS across subscales=45- .63	Observational instrument administered at daycare educator level
CLOP[45]	Quality of physical classroom literacy quality	Internal consistency coefficient (Cronbach's alpha) for print environment composite score= .61; Intra-class coefficient for reliability= 0.34, Retest stability fall-spring test months=.42). Concurrent validity with SABR	Observational instrument administered at daycare centre level
Context of implementation questionnaire	Structural, cultural, organizational and practice quality of implementation	Developed as part of project. Internal consistency coefficient (Cronbach's alpha) for subscales= .70 - .90	Questionnaire completed by daycare administrators and daycare educators
Social validity (only experimental arms)	Educators' attitudes, feelings, opinions about the current conduct of SPELL	Developed as part of project. Internal consistency coefficient (Cronbach's alpha) for subscales= .61 - .92	Questionnaire completed by daycare educators

Implementation notes (only experimental arms)	Intervention frequency and quality, interactions with children, use of differentiation strategies	N/A	Process measures collected as part of the intervention
Child Progress Monitoring reports (only experimental arms)	Acquisition of instructional targets for each of the four language domains	N/A	Observational assessment performed as part of the intervention
Fidelity checklist (only experimental arms)	Fidelity and quality of instruction (dosage, adherence and use of differentiation strategies)	N/A	Observational instrument administered at daycare educator level to code videos of SPELL lessons
Reading checklist (only SPELL Basic +HOME)	Parental reading frequency	N/A	Checklist competed by parents

** ECERS, The Early Childhood Environment Rating Scale; SABR, Systematic assessment of book reading*

3.5.3 Observational ratings

As part of their participation in the study, daycare educators will be asked to video record a shared book reading session, a language instruction session and a mealtime pre- and post intervention. These recordings will be observed and scored for various aspects of process quality using *CLASS* [44].We will use a modified version of the observational instrument *CLOP* [45] to rate the structural quality of the learning environment of the daycare centers. In addition, daycare educators assigned to one of the SPELL arms will be asked to video record themselves implementing *SPELL* at the beginning, middle and end of the intervention. These video recordings will be observed and scored for different aspects of implementation fidelity that represent the critical components of the intervention [27,28].All scoring will be performed by student assistants who are trained by the research team until they can reproduce gold standard scores of the researchers with at least 80% accuracy.

3.5.4 Implementation notes

Daycare educators complete implementation notes addressing questions targeted to the critical components of the intervention after each lesson on iPads which enable them to complete fidelity measures in real-time in their classrooms, enhancing the scope and accuracy of fidelity data. The implementation notes also inform evaluation of child exposure and engagement in the intervention. Three times during the intervention (after lesson 6, 20 and 36) they also assess children's progress on the specific learning objectives of each of

the four learning domains addressed in *SPELL*. Parental fidelity in the home arm is measured through a weekly structured diary.

3.5.5 Background information from statistics denmark

The Danish Central Personal Number System allows us to obtain for each child and parent information on ethnicity, family structure, education, income, use of social welfare etc. To quantify the SES of the children, a probabilistic SES score developed for the VIDA-project[46] will be used. This score is based on five variables from which the probability of occurrence of a "child case" (i.e. children for whom professionals have significant concerns about their development and start examining the child's needs systematically) can be determined from a statistical model with considerable accuracy. The five variables are "Parents living at the same address", "Family income below 10th percentile", "Family depending on social welfare", "One parent has more than a basic education" and "Parents are first or second generation immigrants". The latter variable will be also used to describe the ethnic background of the children.

3.5.6 In depth study

To obtain a measure of the effect of *SPELL* on children's task orientation, and to investigate if *SPELL* has any spill-over effects on individual children's educator and peer interactions, a small subset of children in wave two will be observed pre- and post intervention in their daycares with the *in CLASS* instrument[18] and also while they participate in sample *SPELL* lessons. This subset of children will also be tested individually by student assistants with a range of assessment tools (EVT2 (expressive vocabulary)[47], PPVT4 (receptive vocabulary) [48], Action Picture Test (narrative competencies and linguistic complexity)[49], and Bus Story Test(ability to give coherent descriptions of continuous series of events)[50]to obtain a more in-depth measure of possible intervention effects.

3.6 Primary and Secondary Outcomes

The primary outcome of the study is the change on the seven language assessment scores from pre- to post intervention. Secondary and other outcomes are mentioned in Table 2.

Table 2. Overview of variables for outcome, moderator and fidelity analyses

Variable	Source	Time of measure
Educator outcome		
Instructional support (concept development, quality of feedback, language modelling)	*CLASS*	Baseline and post intervention
Parent outcome		
Quality of Parent-Child Interaction, Frequency of Literacy Teaching, Parent Beliefs, Frequency of Reading, Child Interest, Literacy Support in the Home Environment, Frequency of Technology Use Frequency of Television Use	*HLEQ*	Pre and post intervention (*SPELL Basic + Home*)
Child outcomes		
Language and literacy development. Proximal subscales: phonological awareness (sound discrimination, rhyme, deletion) and letter	*LA: 3-6*	Pre and post intervention

identification. Distal subscales: vocabulary, comprehension and communicative strategies		
Psycho-social development: Emotional difficulties, behavioural problems, hyperactivity/attention problems, difficulties with peers, social strengths	*SDQ*	Pre and post intervention
Daycare moderators		
Daycare size, child-adult ratio, ratio of educated staff, sick leaves, educator turnover rates, organization (planning) and cultural quality (leadership and cooperation, joint attitudes	*CIQ*	Baseline
Physical literacy environment	*CLOP*	Baseline
Educator moderators		
Age, educational background, participation in professional development, work experience, support from daily leader, development of own practice, attitudes and use of new educational methods, occurrence and frequency of educational activities	*CIQ*	Baseline
Emotional support, classroom organization, instructional support	*CLASS*	Baseline
Attitudes towards SPELL, judgment of effects and usefulness of SPELL, association of SPELL with current educational practices	*SFQ*	Baseline
Parent moderators		
SES composite (parents living together, income, education, social welfare, current employment, immigrant status), ethnic background	**Statistics Denmark**	**2011-2013** **Baseline**
Linguistic background, Quality of Parent-Child Interaction, Frequency of Literacy Teaching, Parent Beliefs, Frequency of Reading, Child Interest, Literacy Support in the Home Environment, Frequency of Technology Use	***HLEQ***	
Frequency of Television Use		
Child moderators		
Age (at baseline), gender	Statistics Denmark	2013
Bilingual status, type of bilingualism	*HLEQ*	Baseline
Peer effect moderators (at child group, classroom and daycare level)		
Peer SES, peer pre-test language status (median and fraction below 10%)	Statistics Denmark *LA: 3-6*	Baseline Pre intervention
Educator fidelity		
Dosage and adherence to intervention, quality of differentiation	Implementation notes, FC, CMC	During intervention
Parent fidelity		
Frequency of reading and activities, child's engagement	Reading diary	During intervention (*SPELL Basic +HOME*)

Child fidelity		
Days absent, engagement, dosage of differentiation strategies	Implementation notes	During intervention

HLEQ, Home learning environment questionnaire; CIQ, Context of implementation questionnaire; SFQ, Social feasibility questionnaire; FC, fidelity checklist; CMC, Child Monitoring Checklist

3.7 Selection of Daycare Centers and Randomization

All municipalities that met the following criteria are offered participation: 1) willingness to participate with a large number of daycare centers, 2) willingness to randomly assign daycare centers to one of the four arms, 3) commitment to support the local implementation, and 4) commitment to participate in further development of the manuals and materials. From these municipalities we will select a sample representing the regional variation in Denmark with respect to SES composition of the parent population and daycare organisation. From each municipality we randomly select daycares from two strata based on the median probabilistic SES score, oversampling daycare centers with a high median SES score (i.e. with many children with a low SES background). The child population as of August 1st, 2012 will be used for these computations. Municipalities are allowed to exclude a few daycare centers from the random selection, if a successful implementation of any intervention in that daycare centers is regarded as unlikely. Educators in each classroom will divide children in small groups of five to six children based on their familiarity with the children and age of the children. Educators will then randomly be assigned to child groups.

We will use a stratified conditional randomization to assign the daycare centers to intervention arms. First, we will rank the daycare centers according to the median SES score and build groups of eight daycare centers with similar SES scores. Within each group we will randomly assign the daycare centers to the eight combinations of arm and wave. Randomization will be repeated until a sufficient balance of the number of daycare centers per arm and per wave within each of the participating municipalities is achieved.

3.8 Statistical Analysis

3.8.1 Handling of drop outs

Children who start the intervention, but fail to obtain a pre-test language and literacy assessment are regarded as initial drop outs. Children who start the intervention, are language and literacy assessed at pre-test, but are not assessed at post-test, are considered as final drop outs. The frequency of initial and final drop outs will be analysed overall, as well as with respect to any association with baseline characteristics of children, educators and daycare center, focusing on detection of differences in drop out mechanisms across the intervention arms. In the case of drop-out rates above 10% and evidence for differences in the drop out mechanisms, all subsequent analyses will be combined with multiple imputation. Otherwise, a complete case analysis will be performed.

3.8.2 Handling of missing information on moderators or fidelity variables

Non response in these variables will be analysed accordingly with respect to their associations with the characteristics mentioned above, including measurements during intervention where appropriate. In the case of substantial missing rates, multiple imputation will be used to ensure full use of the available information.

3.8.3 Assessment of intervention effects (main analysis)

Differences between the intervention arms with respect to language scores will be assessed using a hierarchical linear model using change scores as outcomes; child group, educator, class room, and daycare centeras random effects, and municipalities, intervention arms and the pre-test language score as covariates. A global *p*-value referring to the null hypothesis of no intervention effect and estimates of all pairwise differences with 95% confidence intervals and corresponding *p*-values based on the Kramer-Tukey method will be reported. Effect sizes will also be calculated and reported. In the primary analysis, the proximal measure will be analysed prior to the distal measure, and significance of the latter is only assumed if it is also reached for the proximal measure. This way *p*-values need not to be further adjusted. This principle is applied separately for the global comparison as well as the pairwise comparisons. In secondary analyses, all single language subscales, the phonological awareness measure, the total combined score and the further outcome measures will be analysed accordingly. No further adjustments for multiplicity will be performed in the secondary analyses.

3.8.4 Subgroup analysis

The above mentioned analyses will be repeated in two subgroups: a) All children with a non-Danish ethnic background; b) The 10% children with lowest SES according to the VIDA-score.

3.8.5 Analysis of moderators

All single moderators will be analysed for their predictive value for the intervention effect by adding them as a single covariate together with interaction(s) with the intervention(s) to the model mentioned above. In a first analysis, only the predictive value for the intervention difference between the control group and all SPELL intervention arms will be analysed. In a second analysis, all pairwise differences will be considered. Besides the p-value of the interaction, estimated intervention effects with confidence intervals at the lower and upper 10% percentile of the distribution of the moderator and estimates of the change point from a negative to a positive effect will be reported. If several moderators can be identified as predictive, multivariate models will be considered to identify factors acting independently.

3.8.6 Analysis of fidelity measures

In the three SPELL arms, a variety of fidelity measures will be available (cf. Table 2). For any of these measures showing sufficient variation, we will analyse the association with the language change scores by considering a hierarchical linear model with child group, educator, classroom and daycare centre as random effects, and municipalities, the three intervention arms, and the fidelity measure as covariates. In addition, adjusted analyses will be performed adding the baseline measurement and principal component based summary scores of children, parent, educator, and daycare characteristics as covariates. The estimated association of each measure with confidence intervals and p-values will be reported. The analyses will be repeated separately for the SPELL Basic + PD and SPELL Basic + Home arm allowing us to incorporate the arm-specific fidelity measures. Models with several fidelity variables as covariates will be used to identify fidelity measures acting independently. In addition, the effect attributable to each fidelity measure on the intervention differences between the three SPELL arms will be estimated.

3.8.7 Efficacy analysis

To assess the potential effect of the interventions under ideal circumstances, the main analysis will be repeated by including only educators from the three SPELL arms who on average are above the median fidelity (i.e. educators, who actually performed the intervention with a high degree of consistency and accuracy). Here as a summary measure we will use the best predictor of changes in the proximal language score based on all feasibility variables.

3.8.8 Sample size and power calculations

We intend to include 128 daycare centers with an average of three class rooms and 20 children/two educators in each class room, summing to 7680 children. This reflects the typical structure of a Danish daycare center with respect to the age span three to six. Assuming eight municipalities, balanced randomization within each municipality, an intra-class correlation for daycare center and educator of 0.05 and of class room and child group of 0.01 and an effect size of 0.2 for all three active interventions compared to control, we have a power of 95% to demonstrate a significant difference at the 5% level between all four intervention arms using the above mentioned model without pre-test score values as covariates. Adding the latter will further improve the power. We will also achieve a power of 92% to demonstrate a significant interaction with a binary moderator with a prevalence of 0.33 independent of institution, classroom, educator and child group and reducing the effect to 0 in 1/3 of the children using the approach outlined above.

4. ETHICS

In all aspects of the proposed project, the safety and wellbeing of participants - including children, parents, educators and supporting organizations – will be given the highest priority. The study is registered with the Danish Data Protection Agency. In conducting the project, we will pay considerable attention to a wide range of ethical issues as in the implementation of any study. Due to the registration with the Danish Data Protection Agency, the project is categorised as public research. For public research projects of significant societal importance it is not required to ask for consent from each parent in Denmark. Second, in our measurements of children, families, and professionals, testing and intervention procedures we will use established techniques which are known to be safe for general use with a particular group or individual. This is particularly important in this project since children from a range of different backgrounds will participate. Third, participants' right to privacy, confidentiality and anonymity will be strictly observed. The protocols will adhere to relevant provisions outlined in The Act on Processing of Personal Data. No identifying information will be provided in any publication or educational material. Fourth, parents, professionals and associated daycare centers will always be provided with knowledge of main results to honour their contribution to this work.

5. DISCUSSION

Children's early language and literacy skills are central for later educational outcomes. Language and literacy skills particularly predict later reading achievement and children with lags in these areas have difficulties catching up. Meta-analyses suggest a positive impact on children's language and literacy development from interventions in daycares which provide stimulating and supportive interactions between daycare educators and children based on

an effective use of curricula that support systematic and explicit instruction. Yet, most such studies have been conducted in the US and under circumstances that allowed the researchers a relatively high level of control over the implementation. Therefore, we lack important knowledge of how systematic and explicit interventions will translate into non-US societies and how effective such interventions are at a large scale with less control over the implementation.

In this study we evaluate the effects of an intervention called *SPELL* in a large-scale effectiveness study conducted under "real-life" circumstances including all children in the participating daycares independent of parental background. The research will show if systematic and explicit intervention have positive effects for various subgroups of children in a country with a different pedagogical and educations settings than the US (efficacy). Additionally, the project will indicate the effect of a systematic and explicit curriculum-based intervention when carried out at scale and identify implementation factors that contribute to a successful implementation in daycares.

The findings from this study which documents a rigorous evaluation of a systematic and explicit instruction intervention addressing efficacy as well as effectiveness questions, will contribute important new information to the field of early childhood intervention with a focus on language and literacy development.

ETHICAL APPROVAL

Educational research in Denmark cannot be approved by a Regional Ethics Committeeas the ethics committee only review biomedical research. Instead, all research studies must be approved by the Danish Data Protection Agency which safeguards the protection of individuals with regard to the processing of personal data and on the free movement of such data. The trial protocol was approved by the register in 2012 (Journal number: 2012-41-0030).

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. National Early Literacy Panel. Developing early literacy. A Scientific Synthesis of Early Literacy Development and Implications for Intervention. National Institute for Literacy; 2008.
2. Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, et al. School readiness and later achievement. Developmental psychology. 2007;43(6):1428.
3. Heckman JJ. Skill formation and the economics of investing in disadvantaged children. Science. 2006;312(5782):1900-2.
4. Bleses D, Højen A, Jørgensen RN, Jensen KØ, Vach W. Sprogvurdering af 3-årige (09) - karakteristika og risikofaktorer. Working papers, Center for Child Language, e-prints. 2010;10.
5. Egelund N, Nielsen CP, Rangvid BS. PISA Etnisk 2009: etniske og danske unges resultater i PISA 2009: AKF; 2011.

6. Elbro C, Dalby M, Maarbjerg S. Language learning impairments: a 30 year follow up of language impaired children with and without psychiatric, neurological and cognitive difficulties. International Journal of Language & Communication Disorders. 2011;46(6):437-48.
7. Shonkoff JP. Building a new biodevelopmental framework to guide the future of early childhood policy. Child development. 2010;81(1):357-67.
8. Bronfenbrenner U, Morris PA. The bioecological model of human development: Wiley Online Library; 2006.
9. Duncan GJ, Magnuson K. Investing in Preschool Programs. The Journal of Economic Perspectives. 2013;27(2):109-32.
10. Yoshikawa H, Weiland C, Brooks-Gunn J, Burchinal MR, Espinosa LM, Gormley WT, et al. Investing in Our Future: The Evidence Base on Preschool Education; 2013.
11. DeBaryshe BD, Gorecki DM. An experimental validation of a preschool emergent literacy curriculum. Early Education and Development. 2007;18(1):93-110.
12. PreschoolCurriculumEvaluationResearchConsortium. Effects of Preschool Curriculum Programs on School Readiness. Report from the Preschool Curriculum Evaluation Research Initiative. National Center for Education Research, Institute of Education Sciences. U.S. Department of Education; 2008.
13. van Steensel R, McElvany N, Kurvers J, Herppich S. How effective are family literacy programs? Results of a meta-analysis. Review of Educational Research. 2011;81(1):69-96.
14. Skibbe LE, Justice LM, Bowles RP. Implementation processes associated with a home-based phonological awareness intervention for children with specific language impairment. International Journal of Speech-Language Pathology. 2011;13(2):110-24.
15. Mashburn AJ, Pianta RC, Hamre BK, Downer JT, Barbarin OA, Bryant D, et al. Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development. 2008;79(3):732-49.
16. Burchinal M, Vandergrift N, Pianta R, Mashburn A. Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. Early Childhood Research Quarterly. 2010;25(2):166-76.
17. Guo Y, Justice LM, Kaderavek JN, McGinty A. The literacy environment of preschool classrooms: contributions to children's emergent literacy growth. Journal of Research in Reading. 2012;35(3):308-27.
18. Downer JT, Booren LM, Lima OK, Luckner AE, Pianta RC. The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. Early childhood research quarterly. 2010;25(1):1-16.
19. Justice LM, Petscher Y, Schatschneider C, Mashburn A. Peer Effects in Preschool Classrooms: Is Children's Language Growth Associated With Their Classmates' Skills? Child development. 2011;82(6):1768-77.
20. Sawyer BE, Justice LM, Guo Y, Logan JA, Petrill SA, Glenn-Applegate K, et al. Relations among home literacy environment, child characteristics and print knowledge for preschool children with language impairment. Journal of Research in Reading; 2013.
21. Dobbs-Oates J, Pentimonti JM, Justice LM, Kaderavek JN. Parent and child attitudinal factors in a model of children's print-concept knowledge. Journal of Research in Reading; 2012.
22. Manolitsis G, Georgiou GK, Tziraki N. Examining the effects of home literacy and numeracy environment on early reading and math acquisition. Early Childhood Research Quarterly; 2013.

23. Miser TM, Hupp JM. The influence of socioeconomic status, home environment, and childcare on child language abilities. Current Psychology. 2012;31(2):144-59.
24. Connor CMD, Morrison FJ, Slominski L. Preschool instruction and children's emergent literacy growth. Journal of Educational Psychology. 2006;98(4):665.
25. Justice LM, Mashburn AJ, Hamre BK, Pianta RC. Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. Early Childhood Research Quarterly. 2008;23(1):51-68.
26. Hamre BK, Pianta RC, Burchinal M, Field S, LoCasale-Crouch J, Downer JT, et al. A Course on Effective Teacher-Child Interactions Effects on Teacher Beliefs, Knowledge, and Observed Practice. American Educational Research Journal. 2012;49(1):88-123.
27. Hamre BK, Justice LM, Pianta RC, Kilday C, Sweeney B, Downer JT, et al. Implementation fidelity of MyTeachingPartner literacy and language activities: Association with preschoolers' language and literacy growth. Early Childhood Research Quarterly. 2010;25(3):329-47.
28. Darrow CL. The Effectiveness and Precision of Intervention Fidelity Measures in Preschool Intervention Research. Early Education & Development. 2013;24(8):1137-60.
29. Lonigan CJ, Farver JM, Phillips BM, Clancy-Menchetti J. Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. Reading and Writing. 2011;24(3):305-37.
30. Justice LM, McGinty AS, Cabell SQ, Kilday CR, Knighton K, Huffman G. Language and literacy curriculum supplement for preschoolers who are academically at risk: A feasibility study. Language, Speech, and Hearing Services in Schools. 2010;41(2):161.
30. Bleses D, Vach W, Jørgensen RN, Worm T. The internal validity and acceptability of the Danish SI-3: a language screening instrument for 3-year-olds. Journal of Speech, Language, and Hearing Research. 2010;53:490-507.
32. Obel C, Dalsgaard S, Stax H, Bilenberg N. Spørgeskemaom barnets styrker og vanskeligheder (SDQ-Dan). Ugeskrift for læger. 2003;165(5):462-73.
31. Fixsen DL, Naoom SF, Blase KA, Friedman RM. Implementation research: A synthesis of the literature; 2005.
32. Rønnov LM, B. Implementeringsforskning om forebyggelse – en baggrundsrapport. Sundhedsstyrelsen; 2010.
33. Lehman WE, Greener JM, Simpson DD. Assessing organizational readiness for change. Journal of substance abuse treatment. 2002;22(4):197-209.
34. Witt JC, Martens B. Assessing the acceptability of behavioral interventions used in classrooms. Psychology in the Schools; 1983.
35. Marjanovič Umek L, Podlesek A, Fekonja U. Assessing the home literacy environment. European Journal of Psychological Assessment. 2005;21(4):271-81.
36. Melhuish EC, Phan MB, Sylva K, Sammons P, Siraj-Blatchford I, Taggart B. Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. Journal of Social Issues. 2008;64(1):95-114.
37. Weigel DJ, Martin SS, Bennett KK. Ecological influences of the home and the child-care center on preschool-age children's literacy development. Reading Research Quarterly. 2005;40(2):204-33.
38. DeBDeBaryshe BD, BinderJC. Development of an instrument for measuring parental beliefs about reading aloud to young children. Perceptual and Motor Skills. 1994;78(3c):1303-11.

39. Nebrig MR. Parent and teacher perceptions of home activities to encourage emergent literacy: ProQuest; 2008.
40. Matheny AP, Wachs TD, Ludwig JL, Phillips K. Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. Journal of Applied Developmental Psychology. 1995;16(3):429-44.
41. Pianta R, La Paro K, Hamre B. CLASS: Classroom Assessment Scoring System Manual Preschool (Pre-K) Version. Charlottesville, VA: Center for Advanced Study of Teaching and Learning. www. virginia. edu/vprgs/CASTL; 2006.
42. Dynia J. The Literacy Environment of Early Childhood Special Education Classrooms. Predictors of Print Knowledge: The Ohio State University; 2012.
43. Griffin EA, Morrison FJ. The unique contribution of home literacy environment to differences in early literacy skills. Early Child Development and Care. 1997;127(1):233-43.
44. Jensen B, Holm A, Wang C, Kousholt D, Ravn I, Larsen MS, et al. Vidensbaseret indsats overfor udsatte børn i dagtilbud-modelprogram: Statusrapport 1. Design og metode. DPU, Aarhus Universitet; 2011.
45. Williams KT. EVT-2: Expressive Vocabulary Test: Pearson Assessments; 2007.
46. Dunn LM, Dunn DM. Peabody picture vocabulary test, (PPVT-4). Minneapolis, MN: Pearson Assessments; 2007.
47. Renfrew CE. Action picture test: Winslow; 1997.
48. Renfrew CE. Bus Story Test: A test of narrative speech: Winslow; 1997.

APPENDIX

DETAILS OF PREPARATION OF LANGUAGE ASSESSMENT DATA

For the vocabulary scale a delayed entry at item 21 for children above the age of 48;0 was recommended. Consequently, for all children not tested for the first 20 items these items are counted as correct responses. For both the vocabulary and the segmentation scale a discontinuing rule is applied. Whenever the discontinuing rule is applied, all remaining items are counted as incorrect responses. For the rhyming and comprehension scales a stopping rule is applied, too. Here the remaining items are counted as correct responses with a weight of 0.25, corresponding to a response rate at chance level.

For each of the seven scales a summary score is intended to be computed whenever - after the steps above - information on more than 50% of the items are available. Missing values at the item level are substituted by predictions from a two-parameter Rasch model fitted by maximum likelihood. At pre-test, the Rasch model is based on all measurements, at post-test the model is fitted separately for each arm. If an imputation was made, the summary scores were rounded to the next integer. For children who fail to pass the two test items in the rhyming and segmentation part, no score is computed.

The summary score of each language scale is transformed to a percentage score by estimating for each possible value of the score the probability to reach maximally this value. These probabilities are derived for each month of age and each gender group separately from a logistic regression model with the binary outcome to be equal or below the value and age as the only covariate, fitted to all children of the same sex with an age difference of maximally 4 months. Children from families with a non-Danish speaking environment are not included when fitting this model.

Combined scores will be based on averaging the available pseudo z-transformed percentile scores of the scales involved, which are then again transformed into percentage scores. In computing the pseudo-z-scores, all percentages are truncated at 0.99 and 0.01, and z-score above 0 are reduced by 50%. So the pseudo-z-scores range from -2.32 to 1.16.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=596&id=21&aid=5316