Comprehending L2 Comprehension: A Study of Arabic-Swedish Bilingual Preschoolers' Performance on a Swedish Proficiency Test

Annika Andersson, Hanna Lindfors & Kristina Hansson

Abstract

About a fifth of all children in Sweden learn the societal language Swedish outside of the home, i.e., they have Swedish as a second language (L2). Many of these children have lower socioeconomic status (SES) backgrounds, which predicts lower language proficiency. The aim of the present study is twofold: to contribute to a greater understanding of L2-Swedish proficiency in preschoolers with lower SES backgrounds, and to find out how proficiency tests should be adapted for bilingual children such that the tests are valid, i.e., unbiased to the language status (L1 or L2). We investigate test performance on a Swedish receptive language proficiency test (the *Comprehension scale* of The New Reynell Developmental Language Scales, NRDLS) which has a monolingual norming sample. The participants are 51 bilingual children (3-5-years of age) with Arabic as their L1, and who attend preschools in Swedish neighborhoods with lower SES. Results indicate that in contrast to the norming sample, bilingual children's raw scores for subsections of the test are not progressively more difficult. Thus, we need to be aware that bilingual children's high proficiency in a particular aspect of the language does not necessarily imply that they are proficient in aspects that would be considered easier from a monolingual perspective. In addition, there are indications that unfamiliarity with L2 lexical items, that are typically acquired early in L1, causes bilingual children to fail on tasks aimed at assessing syntactic skills, even though they appear to understand the syntactic pattern. We conclude with suggestions for special considerations and adaptations to assess individual L2-comprehension in preschoolers more accurately, such that practitioners in turn can support the children's language development.

Introduction

According to the Swedish statistic central bureau approximately 25% of preschoolers in Sweden are bilingual, many of which have lower socio-economic backgrounds (SES) (SCB 2021). The combination of lower SES and having the societal language as a second language (L2) is a predictor for lower language proficiency which often leads to lower academic achievement (Carlisle et al. 1999; Gottardo 2002; Gottardo et al. 2008; Grönqvist & Niknami 2017; Hammer et al. 2007; Hoff 2013; Miller et al.

https://doi.org/10.15626/hn.20224802

2006; Páez & Rinaldi 2006; Proctor et al. 2005). Important for the current study, preschoolers' language proficiency predicts their subsequent grades in school. This includes math, reading, and to some extent also social skills (Murphy et al. 2016; Pace et al. 2019). Yet, we know little of the development of L2-Swedish proficiency in preschool children with lower SES backgrounds and how to best measure their L2-proficiency.

Therefore, in this paper, we investigate how preschoolers with Arabic as their first language (L1) and Swedish as their L2[1] perform on a Swedish language proficiency test. Our aim is twofold; to contribute to a greater understanding of L2-Swedish proficiency in bilingual preschoolers with lower SES backgrounds, and to find out if and what adaptations of the test procedure are needed in order for test results to better capture children's L2-proficiency. This is important since an understanding of children's various language proficiencies and needs forms a basis for planning appropriate education. To reach this aim, we test the bilingual preschoolers with a receptive language test; the *Comprehension Scale* of the Swedish version of the New Reynell Developmental Language Scales (NRLDS; Edwards et al. 2017). Our specific research questions are:

1. a. How do bilingual children with lower SES backgrounds perform when assessed with NRDLS in comparison to the monolingual Swedish norming sample?
   b. How is this performance associated with chronological age?

2. a. Is the intended incremental degree of difficulty of the sections in the test reflected in the bilingual children's performance?
   b. If not, which sections or items deviate?

While bilingualism *per se* is not problematic for language learning or academic success, for some families, bilingualism is associated with lower SES as measured by the proxy maternal or parental educational level (Bornstein et al. 2003; Grönqvist & Niknami 2017) and type of neighborhood (OECD 1995). These measures of SES are strong predictors for academic success. Indeed, the relationship between lower language proficiency and lower SES has been reported repeatedly in studies of both monolingual children (Basit et al. 2015; Hart & Risley 1995; Hoff 2003a, 2003b; Law et al. 2017; Law et al. 2011; Pan et al. 2005), and bilingual children (Andersson et al. 2019; McClintock & Baron 1979; Schwartz & Stiefel 2006; Umbel et al. 1992).

[1] Throughout this paper we will refer to the native language or mother tongue, as the first language (L1, Arabic for our participants) and the language acquired outside of the home as the second language (L2, Swedish for our participants) even though children in the study differ in age of first exposure to the L2.

Importantly, it is not SES in itself, but rather the quality and quantity of exposure to the language that is related to SES and that in turn affects language proficiency in both monolingual children (Bornstein et al. 2003) and bilingual children (Cobo-Lewis et al. 2002; Duursma et al. 2007; Genesee et al. 2005; Hammer et al. 2007; Oller et al. 2007; Umbel & Oller 1994; Umbel et al. 1992). Thus, for bilingual children less and lower quality exposure to the language explain a lower L2 proficiency than is required to succeed in a school system that is based on the L2 (Corson 1996).

When assessing bilingual children's language proficiency all languages should be considered (Nayeb et al. 2021; Sanchez et al. 2013). However, this is not always feasible (Ebbels et al. 2019; Law et al. 2017) which results in that children typically are tested in their L2 only. Not only is it problematic that only one language is tested, but there is, in addition, a concern of the testing procedure itself affecting the assessment. For instance, the very experience of being tested might be unfamiliar to bilingual children (Fuste-Herrmann et al. 2006; Garcia 1991). Also, the particular tasks and stimulus material that are used are expected to, due to e.g., cultural diversity or a non-overlapping proficiency in both languages, affect children's results. The non-overlapping proficiency across languages have for instance been reflected in u-shaped difficulty levels for Spanish-English bilingual children rather than the expected linear increased difficulty level that is reported for monolingual children (Hickey 1972; Restrepo & Silverman 2001; Umbel et al. 1992). This u-shaped difficulty level has been argued to be due to the early acquired words and constructions being learnt in the home in L1 but not L2, such that the initial sections in a test are particularly difficult in L2. This initial difficulty is followed by words and constructions used in preschools and that therefore are potentially easier for the child. Even if this is known, manuals for proficiency tests do not suggest different order of sections when testing bilingual children.

Overall, the sequence of learning a language as an L2 is similar to the sequence of learning the language as an L1: 1) receptive vocabulary prior to expressive vocabulary, 2) higher frequency words prior to lower frequency words, 3) general terms prior to specific terms, 4) positive polarity prior to negative polarity, 5) morphologically related (e.g., *bike, biking*) prior to lexically distinct (e.g., *car*, *driving*), 6) words describing experiences prior to words describing beliefs (Fuste-Herrmann et al. 2006; Lindholm et al. 1979; Peña et al. 2003), and 7) similar order of acquisition of grammatical structures (Salameh et al. 2004). In addition, acquisition of grammar and syntax is related to the vocabulary within the same language, such that not until a certain size of the vocabulary is reached can the child uncover syntactic patterns and relate words to each other (Bates & Goodman 1997; Conboy & Thal 2006; Marchman et al. 2004). However, the specific input and learning context for an infant or toddler learning L1 differ from the input and context of a toddler learning an L2, especially outside of the home

(Bialystok et al. 2010). Furthermore, multilingual children have to distribute their vocabulary over two or more languages. Clearly, there are several obvious problems when relating bilingual children's test scores to monolingual test norms.

Even if there is an urge to understand the development of L2-Swedish proficiency in preschoolers with lower SES backgrounds to enable support of their language development there is an apparent lack of studies focusing on this group. Crucially, as it is not feasible to develop specific L2-proficiency measures for children of different L1s, we need to know how to best measure their L2-Swedish proficiency with the tests that are available.

Present study

Therefore, this study aims to contribute to a greater understanding of preschool children's L2 proficiency and to facilitate professional assessments of children's L2 comprehension with available tests. In this explorative and descriptive study, we analyze how 51 Arabic-Swedish bilingual 3-5-year-old children perform on a test of their L2—the Swedish version of NRDLS (Edwards et al. 2017). This test assesses children's comprehension and production of single words, combining words, building sentences, verb morphology, pronouns, complex sentences, and inferencing. It thus has a focus on grammar and combining words, rather than on vocabulary, therefore using a basic vocabulary (words small children usually know in their L1). It is used to diagnose developmental language disorder in monolingual children and to point to areas of the child's language skills that need more detailed investigation. It is intended for children in the age range from 2 years of age to 6 years and 11 months and is normed on monolingual children with typical language development. In the present study, we restrict the testing to the Comprehension scale to measure children's receptive language skills as this develops prior to productive skills.

We include a background questionnaire targeting SES background, age of acquisition (AoA), and language exposure and usage to explore the relationships between these factors in an Arabic-Swedish population. We expect to replicate previous studies showing these measures to correlate with language proficiency (e.g., Duursma et al. 2007). Further, we expect SES, especially maternal educational level (Bornstein et al. 2003), to be related with language experience (as measured as exposure and usage), and L2 proficiency (Andersson et al. 2019).

We expect bilingual preschoolers with lower SES backgrounds to show lower proficiency in comparison to the monolingual Swedish NRDLS norms. This expectation is supported by our previous findings, of 6-8-year-old bilingual children with lower SES backgrounds who performed significantly below norms. This was true for the core language score in the Swedish version of the Clinical Evaluation of Language Fundamentals

(CELF, Fourth Edition; Semel et al. 2013) for Arabic-Swedish bilingual children (Andersson et al. 2019) and for the English version for Spanish-English bilingual children (Andersson 2012; Wiig et al. 2004). We are also expecting divergence from the intended incremental degree of difficulty for each subsequent section of the test replicating previous findings (e.g., Restrepo & Silverman 2001). However, in addition, we expect to replicate an increased proficiency with age, as was reported in a recent publication regarding Arabic-Swedish bilingual 4-8-year-old children (Bohnacker et al. 2021).

More importantly, when comparing the raw scores across sections we hope to find patterns indicating specific challenges and, similarly, which sections are uncomplicated for this group of bilingual children. These results should assist us when suggesting how to modify test administration for more valid results that in turn can be of help to practitioners when creating strategies on how to support children's L2 development.

Method

Participants

This study is part of a project including an intervention study focusing on the development of early math that will be presented elsewhere. Here we focus on the L2-proficiency testing that took place pre-intervention. The intervention project was approved by the Regional Ethical Review Board in Lund (Dnr: 2018/957). For that project, we recruited preschools that were interested in being part of the math intervention and which had 3- to 5-year-old bilingual children with L1-Arabic and L2-Swedish. Children were recruited from the participating preschools, which served mainly or only immigrant children in neighborhoods with lower SES. Although many children were born in Sweden they tended not to have been exposed to much Swedish until they started preschool. All Arabic-Swedish bilingual preschoolers were invited to take part in the study.

Caregivers signed informed consent forms prior to children being invited to take part in the study. Five children were excluded due to not submitting birthdates (3) or not being able to carry out the test at all (2). The final sample (*whole group*) consisted of 51 children (Table 1). Of the 51 children 29 completed the entire test (i.e., from subtest A through subtest H, Table 2). This group is called the *full-test subgroup* (Table 1). The results from the full-test subgroup were crucial for answering the second research question regarding the presence of an incremental degree of difficulty of the sections in NRLDS. Therefore, there are separate descriptions of this subgroup and their results.

Table 1. Demographics for the entire group and for the full-test subgroup

Group	*n* (F)	Age
Whole group (*SD*) range	51 (23)	4;9 (0;10) 2;11-6;1
Full-test subgroup (*SD*) range	29 (16)	4;11 (0;10) 3;2-6;1

Note. *n*, number of children in the group; F, number of females within brackets; Age given in years;months, for age means are given followed by standard deviations, *SD*, in brackets and range below.

Table 2. Target structures with examples from the English version of NRDLS

Section	Task	Examples	Items
A	Selecting objects	*Var är bordet?* 'Where's the table?'	10
B	Relating two objects	*Göm skeden i lådan* 'Hide the spoon in the box'	10
C	Verbs	*Gör så att apan sitter* 'Make Monkey sit' *Peka på apan flyger* 'Show me monkey flying'	10
D	Sentence building	*Gör så att kaninen går* 'Make Rabbit walk' *Peka på Kaninen äter ett äpple* 'Show me Rabbit eating an apple'	10
E	Verb morphology	*Peka på tjejen som dricker* 'Show me to the girl who drinks' *Peka på killen som sprang* 'Show me the boy who ran'	6
F	Pronouns	*Kramar mormor sig själv?* 'Is the grandmother hugging herself?' *Målar mamman av henne?* 'Is the mother painting her?'	6
G	Complex sentences	*Tjejen som har en hatt springer* 'The girl who is wearing a hat is running' *Elefanten blir buren av killen* 'The elephant is carried by the boy'	10
H	Inferencing	*Vem känner sig väldigt glad?* 'Who is feeling very happy'	10

Note. Items indicates the number of items (and thus also maximum available points, i.e., total 72) for each target structure. For complex sentences children were asked to "show me" followed by the sentence in the examples

Procedures

When arriving at the preschool the first author (referred here to as the *test administrator*), who has a background as a preschool teacher, spent time playing with children focusing on the children whose parents had signed consent forms. During circle time she presented herself and the reason for

her being there. This included a presentation of three soft animals that were used in the math intervention, which is not discussed further here. The test administrator ended the presentation by showing a bag which contained the material for the NRLDS testing and told children that she would like some of them (naming the children with parental consent) to join her in a separate room where they would look at the content of the bag together. Most children were excited and curious and wanted to be the first to play in the separate room with her, while a few were shyer. Therefore, in some cases, a preschool teacher would come along and was then seated next to but just behind the child to not inadvertently help the child during the test. In all cases, the test administrator made sure that the child was aware that s/he could leave the room to go and play with the other children whenever s/he wanted to and was also sensitive to any signs of the child being ready to leave. However, the children, even those who were shy and quieter, seemed interested in the tasks such that the sessions were only aborted by the administrator as according to the instructions for administering the test. The warm-up section for NRDLS (see below) was implemented with all children since they as a group were expected to have few prior experiences of testing (refer to e.g., Fuste-Herrmann et al. 2006).

Teachers were asked to hand out a questionnaire in Swedish with Arabic translations in envelopes to caregivers. This questionnaire asked for parental educational level, parents' arrival in Sweden, the child's and parents' AoA (as measured as first time of exposure to L2), and the child's exposure and use of L1 and L2 during weekdays and weekends. Unfortunately, some teachers were reluctant to hand out the questionnaires (see Appendix A for an English translation) that targeted private information. For those who received the questionnaire, an Arabic speaking liaison met parents at the preschools and helped filling in the form to support parents who were not literate or had challenges when reading and filling in the questionnaire, an approach that have proven successful in prior studies (e.g., Andersson 2012). The liaison was involved in the math intervention and spent time with the children at the preschools and was well known by the parents. However, even with the support from this liaison our approach failed in that only parents to 18 children (8 females) received and chose to fill in the information. Although it is not clear how representative the results from this self-selected group are, we have chosen to analyze the data from these questionnaires as results could potentially open up new questions or concerns. The results will be discussed as preliminary, and readers will be reminded of the limitations of the results for making inferences to a larger sample.

NRDLS Test

The NRDLS (Edwards et al. 2011) is designed to be attractive to small children, with tasks making use of both objects and pictures. The

Comprehension scale consists of 72 test items for which one point is scored for each successful answer (maximum of 72 points), in 8 sections (A-H) with an incremental degree of difficulty (Table 2). All but two sections consist of 10 test items. The focus is on grammatical skills, understanding forms of words (morphology) as well as relations between words in word combinations (syntax) and ranges from comprehension of single words (identification of named objects), through comprehension of prepositional phrases, simple sentences, verb morphology and pronouns to comprehension of complex sentences (relative clauses and passives) and inferencing.

Prior to testing, there is an optional warm up section where the child is asked to point to his or her body parts, such as feet, tummy, and nose. The following sections (A-H) and tasks (numbered 1 or 2 in sections with 10 points, except for H) have an incremental degree of difficulty (refer to Table 2). In section A, the child's task is to identify an object from an array of five items that differ for section A1 and A2. In the following sections (B-H), the tasks are to manipulate objects or figures according to instructions or to point to a picture matching an expression. More specifically, section B, *Relating two objects* consists mainly of prepositional phrases (B1), but also a few coordinated words (B2, e.g., *Ge mig äpplet och sängen*, 'Give me the apple and the bed'). In section D, *Sentence building*, items consist of SV, SVO, and SVOAdv-sentences (subject, verb, object, and adverbial, e.g., *Kaninen kittlar nallen med en sopborste,* 'the rabbit tickles the teddy bear with a broom'). *Verb morphology* items in section E target comprehension of present vs. past tense, while *Pronouns* in section F tests comprehension of reflexive vs. non-reflexive pronouns. Finally, *Complex sentences* in section G tests comprehension of relative clauses and passives, while section H test inferences (*Vem känner sig väldigt glad?* 'Who is very happy?'). The sections A-G have trial items to make sure the child understands what s/he is expected to do prior to moving on to the test items. The NRLDS instructions for the test administrator are to follow the instructions and protocol strictly to ensure the test is carried out in the same manner as for the norming population. The test manual also has detailed scoring instructions. The *Comprehension scale* can take up to 20 minutes to administer. The recommendation is to always carry out the three first sections (A-C) and then to try the first part of section D, irrespective of whether the child responds correctly or not. After that, the recommendation is to interrupt when the child has completely failed two sub-sections in a row.

Some adaptations to the test were made, taking cultural diversity into account (Letts & Sinka 2011). These were restricted to items in section D. *Krama* 'hug' was exchanged for *klappa* 'pat' (*Gör så att apan klappar nallen* 'Make the monkey pat the teddy') and *pussa* 'kiss' for *putta* 'push' (*Gör så att kaninen puttar nallen* 'Make the rabbit push the teddy') as suggested by the Arabic speaking liaison.

Analyses

For NRDLS, the children's responses were scored according to the manual. For answering the first research question regarding how bilingual children with lower SES backgrounds perform on this test, we computed raw scores rather than standard scores since the latter are based on the monolingual norming sample. That is, it is normed on age rather than length of exposure which would be interchangeable for monolingual children, however not for bilingual children. Means and standard deviations of total raw scores for each of the sections were computed for each of the 51 participating children. Of these children, testing for 22 stopped before the last section (H), following the rules described above. Thus, it was possible to carry out the whole test with 29 of the children, which constituted the full-test subgroup. We report the results for this subgroup separately as their results are relevant for the second research question regarding the incremental difficulty levels across sections.

Further, the Pearson's correlation coefficient was calculated to explore the correlation between total raw score and age, which constituted the second portion of the first research question. According to Cohen's convention (Cohen 1988) correlations below .30 were considered weak, while .30-.49 were considered moderate and correlations of .50 and above strong. In addition, means and standard deviations for answers to the questionnaire were calculated as well as Pearson's correlation coefficients for relationships between L2-proficiency (total raw score), L2-experience (an average of usage and exposure combined), the SES variables maternal and paternal educational level, and AoA. Importantly, however, these results are explorative in nature and will only be discussed as pilot data, which potentially open up new questions or concerns for L2 acquisition for Arabic-Swedish bilingual preschoolers with lower SES backgrounds. These results will be discussed as future directions in the discussion section.

Results

In this section, we will first present the results pertaining to the first research question regarding how the group of bilingual preschoolers performed on the proficiency test and how these results are related with chronological age. This is followed by the results from the full-test subgroup, which is crucial for answering the second research question focusing on the intended incremental degree of difficulty and deviations from the intended linear pattern. In the last portion of the result-section we present the preliminary results from the questionnaire exploring the relationships between L2-experience, AoA, SES, and L2-proficiency.

The bilingual children's total scores and section scores are presented as average raw scores in Table 3 (including number of children carrying out each section). Averages were calculated for the children that took part in the particular section. As a comparison to the whole group's range of total

scores the monolingual norming population with a similar age span (2;11-6;1) showed a narrower range not including any raw scores below 38 (range, 39-67; Edwards et al. 2017). Importantly, in the present study also the full-test subgroup performed considerably lower than the norming population even if the range was narrower than for the whole group (Table 4). Yet, for all 51 children there was a moderate positive correlation between chronological age and raw scores ($r = 30$, $p < .05$; Figure 1). In figure 1, we show the large spread of scores in relation to the regression line illustrating the large variance within the group. The positive relationship was stronger ($r = .56$, $p < .005$) for the 29 children in the full-test group in comparison with the whole group. Thus overall, bilingual children tended to score higher on the test with age.

Table 3 Average scores for each section for the whole group

	Total	**A (10)**	**B (10)**	**C (10)**	**D (10)**	**E (6)**	**F (6)**	**G (10)**	**H (10)**
M	35.9	7.8	4.9	8.4	7.0	3.2	3.1	5.3	5.3
SD	20.0	2.4	3.5	2.1	2.4	1.4	1.3	2.2	2.4
Range	1-68	1-10	0-10	3-10	1-10	0-6	0-6	2-10	0-9
n	51	51	49	43	42	39	32	30	29

Note. Total indicates total raw scores[2] in A-H. A-H indicates sections in the test, refer to Table 2, with maximum score possible within brackets. *M*, average scores; *SD*, standard deviation; Range, minimum to maximum; *n* = number of children taking part in this section.

Table 4 Average scores for each section for the full-test subgroup

	Total	**A (10)**	**B (10)**	**C (10)**	**D (10)**	**E (6)**	**F (6)**	**G (10)**	**H (10)**
M	50.5	9.0	7.0	9.3	8.0	3.5	3.3	5.2	5.3
SD	10.1	1.4	2.8	1.2	1.7	0.9	1.1	2.2	2.4
Range	26-68	5-10	0-10	6-10	4-10	2-6	2-6	2-10	0-9

Note. Total indicates total raw scores in A-H. A-H indicates sections in the test, refer to Table 2, with maximum possible within brackets. *M*, average scores given in columns; *SD*, standard deviation; Range, minimum to maximum.

[2] While the current study had no expectation of gender differences previous studies have reported on these in both L1 and L2 (e.g., Zambrana et al. 2012). Therefore, we included an independent t-test ($t(49) = 1.42$, $p = .162$) confirming the numerical gender difference in total scores to be non-significant (girls: M = 40.2, SD = 19.0, and boys: M = 32.3, SD = 20.5).

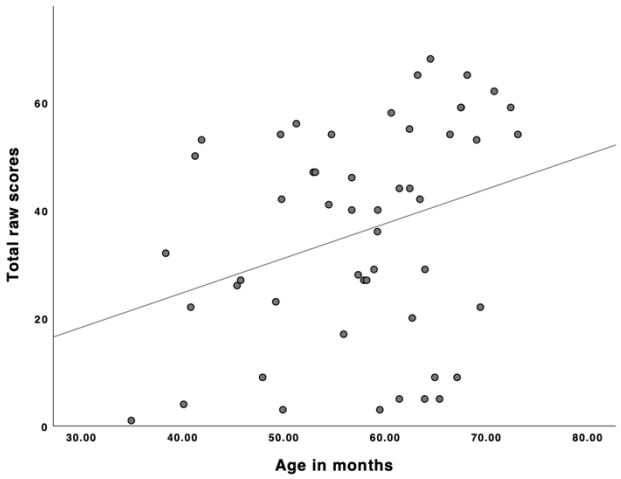

Figure 1. Correlation between age and total scores for the whole group. Age is measured in months.

The degree of difficulty across sections for a monolingual sample is increasing incrementally. We should therefore expect mean scores to be lower for each section (from A through H). The decreasing number of participants in the whole group who concluded a section suggests an incremental difficulty level (Table 3). However, from the full-test subgroup, we learn that the progression through the test on average was not incremental (Figure 2). Note that in Table 4 actual raw scores are given while the box plots in Figure 2 show percentage correct which accounts for differences in number of possible total scores across sections (6 or 10). Rather than an incremental difficulty level, the mean score for section B (*Relating two objects*) was strikingly low compared to the following sections, and the mean score on section C (*verbs*) was unexpectedly the highest. Further, there were no difference in scores between the four last sections E through H (refer to Figure 2). However, the larger variance in the last two sections (G and more so for H) in comparison to that of E and F could be seen as an indication of higher difficulty for these sections, at least for some children.

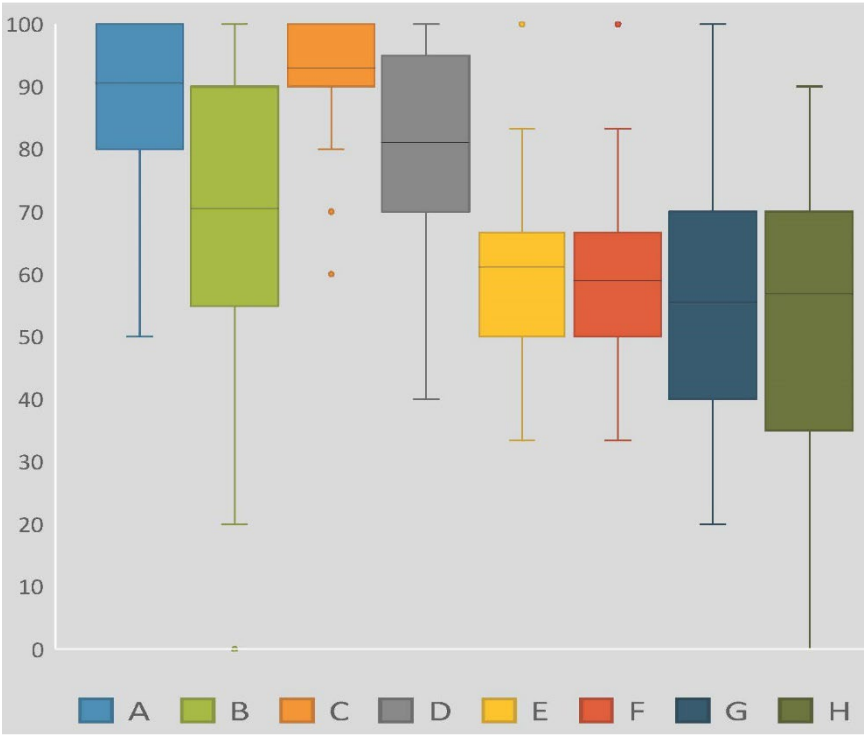

Figure 2. Percentage raw scores of possible scores for each subsection for the full-test subgroup. The boxes indicate quartile range two and three while whiskers indicate the first and the fourth quartile range. The bars indicate means (as this represents percentage correct these are not identical to the values in Table 4). Outliers are indicated by circles above or below whiskers.

At the item level, the items with the highest proportion of correct responses were *bollen* 'ball', *apan* 'monkey', and *pennan* 'pen' in section A2 (*Var är…* 'show me the…'), *ankan* 'duck', *bordet* 'table', and *strumpan* 'sock' in section A1 (*Var är…* 'show me the…'), acting out the verbs *sitter* 'sit', *klappar händerna* 'clap hands', and *går* 'walk' in section C1, and *Gör så att apan hoppar* 'make Monkey jump' in section D1. The test items with the lowest number of correct responses (0-2 children) were items in section B2, E, G, and H. Section B2 considered prepositions where the child was instructed to put a teddy bear in relation to a truck. The bilingual preschoolers in the full-test group had particular difficulties with four of the five prepositions: *framför* 'in front of', *bredvid* 'next to', *bakom* 'behind', and *under* (but not *på* 'on'). In section E: *Verb morphology,* the child first looked at one image of an action, for instance of two girls brushing their hair followed by two images; one with one of the girls brushing her hair and one with the other girl who had a brush in her hand but was not brushing her hair. The child was asked to point to the image that showed the girl that

borstade 'brushed' her hair (note the past tense of *borsta* 'brush') and was expected to point to the girl that did not brush her hair any longer. The section included two training items where the test administrator and the child could talk about the images and what they represented after the child had chosen an image. On the test, children would typically answer by pointing to the image representing present tense that is, where the action was being performed rather than the image suggesting that the action had taken place (past tense). This resulted in children to typically succeed on the three items representing present tense and fail on the three items representing past tense. Further, children had difficulties with the items in section G (*Complex sentences*) that included passive voice, and the specific item *Vems dotter har födelsedagskalas* 'Whose daughter has a birthday party' in section H (*Inferencing*).

As questionnaires were received and filled in only by a few parents (n = 18) we compared children of this partly self-selected group with children whose parents did not receive or fill in the questionnaire (n = 33). Children of these two groups did not differ on either total scores ($t(49) = .735$, p = .466; children with questionnaires, M = 38.7, SD = 28.1; children without questionnaires, M = 34.3, SD = 21.1), or age ($t(49) = -.324$, p = .747; children with questionnaires, M = 4:9, SD = 0:9; children without questionnaires, M = 4;8, SD = 0;9). Below follow the results from this preliminary exploration of SES, AoA, and L2-experience, and these variables' relationship with L2-proficiency.

Each parent was asked to check the box that represented their highest degree of educational level, as according to Hollingshead's educational levels (1975; Table 5, see Appendix A for an English translation of the questionnaire). While parents in this group had on average a high school degree the variances in maternal and paternal educational level was high. Educational levels ranged from less than seven years in school to having a university degree which would indicate a higher SES background even if living in a lower SES neighborhood. Parents reported children to have been exposed to Swedish for an average just shy of two years, although here too the variance was high (Table 5). This large variation was similar to what we experienced when talking to parents that did not hand in a questionnaire and what we learnt from teachers when discussing children's exposure overall. That is, while many children were born in Sweden, many had not been exposed to Swedish prior to preschool. In addition, some children had arrived as immigrants more recently. These results and anecdotes suggest shorter length of exposure than expected for their age.

Table 5. Explorative data on L2 experience and SES background

	Maternal Ed	Paternal Ed	AoA	L2-exposure weekday	L2-usage weekday	L2-exposure weekend	L2-usage weekend
Mean	4.1	4.1	1;11	50%	52%	28%	37%
(*SD*)	(2.1)	(2.0)	(1;6)	(21%)	(25%)	(23%)	(28%)
Range	1-7	1-7	0-5	10-80%	20-95%	0-90%	0-95%

Note. Ed, educational level, from Hollingshead (1975), 1 below seven years of schooling, 2 nine years of schooling, 3 at least one year in high school, 4 high school graduate, 5 at least one year in college or university, 6 BA or BS, and 7 University graduate, MS, MA, PhD. AoA, age of acquisition given in years;months.

Moving on to language use, even fewer data points were collected. Notably, only 15 parents gave information regarding the exposure (amount heard) over weekdays and weekends, while 16 parents gave information of the usage (amount spoken). Parents filled in portions (in percentages) of L1 and L2 exposure and usage for each of the following, the child is exposed to L1 and L2 during weekdays, the child is using L1 and L2 over weekdays, and the same for weekends for each language. By including questions on both L1 and L2 we could ensure that the sum would not be higher than 100%, which could have been an indication of parents not understanding the question. For this very small group, there were significant differences indicating, as expected, that children's L2 experiences were larger during weekdays compared to weekends (exposure: $t(14) = 3.42$, $p = .004$; usage: $t(15) = 2.76$, $p = .015$; refer to Table 5 for means).

Results from the explorative bivariate correlations between the two SES variables (maternal and paternal educational level), AoA, L2-experience (exposure and usage averaged together over weekdays and weekends separately) and L2 proficiency (total raw scores) follows below (refer to Table 6).

While parental educational levels were strongly correlated ($r = .76$, $p < .001$) only paternal educational level had a significant and strong relationship with L2 experience during weekdays and a moderate relationship with AoA that was approaching significance (Table 6). These relationships suggested earlier acquisition and more L2-experience over weekdays with higher paternal education. The measure of first exposure to L2 (AoA) was strongly associated with more L2 experience over weekends suggesting longer exposure was related to more experience with L2 outside of preschool. Further, a strong relationship between L2 experience over weekdays and proficiency and a moderate relationship that was just approaching significance for the same relationship over weekends suggested proficiency increased with experience and vice versa for the group of children whose parents received and handed in the questionnaire.

Table 6. Explorative bivariate correlations between background and proficiency measures

	AoA	L2-experience weekdays	Total raw scores
Paternal Ed	-.45$^{\#}$	.52*	
L2-experience weekdays			.52*
L2-experience weekends	-.68**		.42$^{\#}$

Note. Ed, educational level from Hollingshead (1975), L2 represents average of L2 spoken and heard (i.e., use and exposure) over weekdays and weekends. $^{\#}p < .10$, $^{*}p < .05$, $^{**}p < .01$

Discussion

This study was aimed at furthering our understanding of L2-Swedish proficiency in preschoolers with lower SES backgrounds and explore results that could indicate how proficiency tests should be adapted to be more valid for bilingual children. Specifically, we wanted to answer the question of how 3-5-year-old bilingual children with Arabic as their L1, performed when their L2-Swedish receptive proficiency was tested with the Swedish version of the NRDLS. As a summary of the results and as expected, the participating children's average total scores were lower than test norms for monolingual Swedish-speaking children. While chronological age correlated with total score, this relationship was only moderate, and the variation was larger than reported in the norming Swedish population of the test (Edwards et al. 2017: 24 figure 4.2). The progression through the test did not follow the intended incremental degree of difficulty. That is, some early sections were more difficult than later sections, and a few specific lexical items were particularly difficult.

Below we will further our discussions on how these children as a group performed on the test, after which we will discuss the failure to establish an incremental difficulty level across sections. This is followed by a discussion on specific items in the test that showed an unexpected difficulty level. We conclude that special considerations and adaptations are needed to accurately assess preschoolers' individual L2-comprehension and give some suggestions. We did not collect enough background questionnaires for being able to draw any inferences from this data. Yet, we will prior to our conclusion discuss the results from the self-selected group that received and handed in the questionnaire as these results can be informative for future studies.

On a group level, the bilingual preschoolers' total scores were lower than expected for their age range. This was due to a wider range of total scores in this group compared to the norming group, such that some children scored much lower than expected. Importantly, however, the total scores also included bilingual children who reached higher scores than expected of a monolingual child of the same age. A previous study (Håkansson et al. 2003) found that Arabic-Swedish children varied with respect to whether their performance was stronger in Arabic or Swedish suggesting that the reason for some children in the current study scoring very low would be that

they were dominant in their L1. Even though it is likely the case for the children with very short exposure to Swedish, teachers often suggested that children low in L2 proficiency tended to be low also in L1 proficiency and vice versa. In an evaluation of language screening tests performed at child health centers (including only one or both languages) Nayeb et al. (2021) found that three times as many multilingual children than monolingual children in the age range 2;5 – 2;9 performed at a level indicating developmental language disorder. Furthermore, a study focusing on comprehension and production of nouns and verbs by 4-8-year-old Arabic-Swedish speaking children found similar proficiencies in both languages (Bohnacker et al. 2021). This indicates, thus, that there could be several reasons for lower proficiency for the group, such as low levels of L2 experience, being L1 dominant, but also having overall low language proficiencies. Importantly, however it is worth noticing that in the group of Arabic-Swedish preschoolers from lower SES neighbourhoods some children performed at or above the level indicated by the monolingual norming sample.

Notably, we observed that many children mixed up the cuddly toy animals, confusing for example *rabbit* and *teddy,* in the section testing comprehension of intransitive verbs and sentences consisting of two and up to four elements (section D). This suggests that children who may master basic syntax (i.e., how to build simple sentences), but do not understand a specific vocabulary item fail, and their scores on the test will thus not represent their competence. Accordingly, we believe the proficiency test to be biased against bilingual children. Indeed, another example was that very few children responded correctly to the very first test item (A1) *Var är koppen* '*Show me the cup*'. It could be due to this being the first item, so that despite the warm-up items (pointing to body parts) the children were not prepared for the test procedure. However, we suspect that it was because *cup* specifically is a word that is not so often used in a preschool context (Bialystok et al. 2010). This suspicion was strengthened when a large majority of the children responded correctly to the items that followed (*table* and *duck*).

The bilingual preschoolers showed specific difficulties also on test items in section B2 focusing on prepositions that are notoriously challenging to acquire in L2 (e.g., Bowerman 1996; Bratož 2014) presumably due to the different classifications of space in different languages. In this section, the child is asked to place a teddy in relation to a truck, which makes perspective a confounding factor. Moreover, the size of the objects is odd – the teddy is larger than the truck, which makes *Put Teddy under the truck* strange (the arrangement is rather *Truck on Teddy*). According to the Swedish NRDLS manual (Edwards et al. 2016) it seems to have been the case that the youngest children in the norming population had difficulties particularly with this section. However, interestingly, where the whole

group and the full-test subgroup differed with respect to order of difficulty was in section B, which was comparatively less difficult for the full-test subgroup. It is possible that mastering prepositions is an indication of overall higher proficiency in L2 Swedish as acquiring prepositions require a certain vocabulary size (e.g., Conboy & Thal 2006).

The language proficiency test was originally developed to show an incremental degree of difficulty over each section, yet some bilingual children had apparent difficulties even in the first section. According to the NRDLS manual, after section C the testing should be terminated if the child is not able to answer in a section. By visually inspecting the number of children who took part in a specific section we can be led to believe that there was an incremental degree of difficulty as fewer and fewer children were included. However, to more properly investigate the existence of an incremental difficulty across sections a better approach would be to investigate the scores for children who took the full test.

From the full-test subgroup we learned that, in addition to specific items in the first section (A) and prepositions in section B, the test of tense was especially difficult. Children did well on the three present tense items (with all being correct on *dricker* 'drinks' and only one of the children choosing the incorrect option for *jonglerar* 'juggles' and another child choosing the incorrect option for *puttar* 'pushes'). However, on the past tense items they performed far below chance (10-24% with 50% representing chance level). It is possible that as in everyday speech regular verb-suffixes are not properly pronounced such that e.g., *borstade* 'brushed' is pronounced *borsta'* which lends a fully pronounced *borstade* as a low frequent word for the children. The other two verbs in past tense were irregular verbs, *sprang* 'ran' and *drog* 'pulled'. Of the three past tense forms, children performed best on *sprang* 'ran'. This is presumably a frequent verb at the preschool in both present and past form. Further, the high scores on present tense verbs tested in section C suggest frequent exposure to verbs labelling activities often performed in the preschools. Importantly however, the lack of a progression in degree of difficulty in the test for the present population of bilingual children is noteworthy.

Regarding the research questions, we have thus established that bilingual children perform lower on the test than expected for their age even though age is positively related with their performance. As expected, we did not find an incremental difficulty level across sections. However, both the lower scores for section B and unexpected higher scores for section C could be explained by the vocabulary that is used at preschools.

According to the total scores and scores in each separate section, we would like to suggest the following modifications to the test when testing bilingual children, especially Arabic-Swedish preschoolers. First, we would suggest initiating the test session with verbs (section C). The test items are familiar to the child who can commence the session with succeeding, which

is likely to positively affect their self-confidence and thus also their results on subsequent sections. Indeed, several children did not respond at all to some items, which could have been an indication of not feeling confident to make a guess after several early failures. Section C should be followed by the second subsection of A (identical questions and procedures but pointing to another set of five objects in reference to A1), which seemed more familiar to children than the items in A1. Also, the cup could be moved to the last position within section A1, such that the expected unfamiliarity of this item would not have such a profound effect.

Further, results from this study also suggest a change of order such that relating two objects (section B) is administered after sentence building (section D) when testing bilingual children. This order would adhere to this groups' pattern of increasingly difficult sections, as shown by number of children being able to engage in each section and especially, the results from the full-test subgroup. Thus, such a modification of section order would allow for more valid results in relation to children's proficiencies.

As mentioned in the introduction, it is important never to take for granted that a bilingual child knows all words that are part of an early L1 vocabulary (Bialystok et al. 2010). Thus, even if sections are ordered according to our suggestions, it would be important to initiate later sections even if the child would fail on earlier sections. We would also suggest, prior to initiating the testing session, as a warm-up, to let children play with the cuddly toy animals to ensure they know their labels. This approach would ensure that children do not fail on items focusing on grammatical structures due to not having full mastery of specific lexical items. Indeed, modifications when administering tests to bilingual children include ensuring that they have experience with the content and tasks assessed in the test. These include rewording or expanding instructions, allowing additional time to respond, providing credit for responses that use a different dialect than the standard, the use of additional demonstration items, and asking children to explain incorrect answers (Saenz & Huer 2003). We are confident that these changes to the administration of the test will make the results more valid. This will in turn ensure better knowledge of what the challenges are for the particular child. Thus, the test can be informative to the test administrator who can develop an intervention.

We would now like to turn to the results from the questionnaire. Importantly, the group that received and chose to answer the questionnaire was small. Even if this group did not differ from those who did not hand in a questionnaire either on age or on total scores it is not possible to know if these results are representative for the whole group. Also, we cannot be certain that we can make any inferences to other bilingual preschoolers from these results. Yet, it is likely that these preliminary results will provoke new questions and are therefore included here.

As alluded to in the introduction, SES has repeatedly been shown to be related to children's language development (e.g., Hoff 2013; Schwab & Lew-Williams 2016). Traditionally, the proxy maternal educational level has explained variances in children's language proficiency (Bornstein et al. 2003; Letts et al. 2013; Magnuson et al. 2009; Zambrana et al. 2012). For example, Letts et al. (2013) found that children of mothers with more years of education performed better on the NRDLS than those with mothers with minimum years of education. However, the same pattern was not replicated in the current study as neither maternal nor paternal educational level correlated with the child's performance on NRDLS. Thus, similar to a previous study of older Arabic-Swedish speaking children in Sweden that did not find a relationship between family SES and vocabulary scores (Bohnacker et al. 2021). Importantly though, L2 experience (usage and exposure averaged) during weekends correlated both with NRDLS score and with paternal educational level, suggesting an indirect association between paternal educational level and the child's proficiency in Swedish. It is possible that paternal educational level propels the family's integration into the Swedish society through which the child receives more exposure in form of, for example, social interactions and media to L2. Indeed, this replicates previous research that did not find amount of exposure to L2 but rather that the proportion of L2 input provided by native speakers explained variance in children's L2 proficiency (Place & Hoff 2011). Indeed, the results from the correlation analyses for the self-selected group that handed in questionnaires suggest that there is a complex interaction between different background factors, such as parental SES, AoA, and L2 experience that together influence the child's L2 proficiency. In particular, higher paternal educational level, earlier AoA, and higher exposure to L2 during weekends seem to be favorable factors that are associated with each other and directly or indirectly with the child's proficiency in the L2 and should therefore be explored further in larger populations.

Conclusions

There is great variation in the performance of children in the age range 3-5 years with L1-Arabic, who attend preschools in Swedish neighborhoods with lower SES, when their L2-Swedish comprehension is assessed. These bilingual preschoolers perform better with age although the range is larger than expected from a monolingual sample. Specific challenges were prepositions, reflexive pronouns, and past tense forms of verbs in addition to specific lexical items. Although the results from a language proficiency test administered as according to instructions are informative, we conclude that special considerations are needed to accurately understand preschoolers' individual L2-comprehension. For instance, professionals should reorder the sections and adapt their qualitative L2 assessments so that children do not

fail on tasks aimed at assessing syntactic comprehension due to a possible unfamiliarity with certain L2 lexical items.

Acknowledgements

We thank preschools and preschoolers who were interested to be part of our study. We gratefully acknowledge funding from Marcus and Amalia Wallenberg Foundation [MAW 2016.0069].

References

Andersson, Annika (2012), *Second language acquisition in 6- to 8-year-old native Spanish-speaking children: ERP studies of phonological awareness, semantics, and syntax*. PhD Monograph, Psychology Department, University of Oregon.

Andersson, Ketty, Hansson, Kristina, Rosqvist, Ida, Åhlander, Viveka L., Sahlén, Birgitta & Sandgren, Olof (2019), "The contribution of bilingualism, parental education, and school characteristics to performance on the Clinical Evaluation of Language Fundamentals: Swedish", *Frontiers in Psychology* 10(1586). https://doi.org/10.3389/fpsyg.2019.01586

Basit, Tehmina N., Hughes, Amanda, Iqbal, Zafar & Cooper, Janet (2015), "The influence of socio-economic status and ethnicity on speech and language development", *International Journal of Early Years Education* 23(1): 115–133. https://doi.org/10.1080/09669760.2014.973838

Bates, Elizabeth & Goodman, Judith C. (1997), "On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia and real-time processing", *Language & Cognitive Processes* 12(5/6): 507–584.

Bialystok, Ellen, Luk, Gigi, Peets, Kathleen F. & Yang, Sujin (2010), "Receptive vocabulary differences in monolingual and bilingual children", *Bilingualism* 13(4): 525–531. https://doi.org/10.1017/S1366728909990423

Bohnacker, Ute, Haddad, Rima & Öberg, Linnéa (2021), "Arabic-Swedish-speaking children living in Sweden: Vocabulary skills in relation to age, SES and language exposure", *Journal of Home Language Research* 4(1): 1–18. https://doi.org/10.16993/jhlr.37

Bornstein, Marc H., Hahn, Chun-Shin, Suwalsky, Joan T. D. & Haynes, O. Maurice (2003), "Socioeconomic status, parenting, and child development: The Hollingshead four-factor index of social status and the socioeconomic index of occupations", in Marc H. Bornstein & Robert H. Bradley (Eds), *Socioeconomic status, parenting, and child development*. Mahwah, US: Lawrence Erlbaum Associates Publishers, pp. 29–82.

Bowerman, Melissa (1996), "Learning how to structure space for language: A crosslinguistic perspective", in Paul Bloom, Mary A. Peterson, Lynn Nadel & Merrill F. Garrett (Eds), *Language and space. Language, speech, and communication*. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics, pp. 385–436.

Bratož, Silva (2014), "Teaching English locative prepositions: a cognitive perspective", *Linguistica* 54(1): 325–337. https://doi.org/10.4312/linguistica.54.1.325-337

Carlisle, Joanne F., Beeman, Margaret, Davis, Lyle Hull & Spharim, Galila (1999), "Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual", *Applied Psycholinguistics* 20(4): 459–478.

Cobo-Lewis, Alan B., Pearson, Barbara Z., Eilers, Rebecca E. & Umbel, Vivian M. (2002), "Effects of bilingualism and bilingual education on oral and written English skills: A multifactor study of standardized test outcomes", in D. Kimbrough Oller & Rebecca E. Eilers (Eds), *Language and literacy in bilingual children* (Vol. 2). Tonawanda, N.Y.: Multilingual Matters Ltd., pp. 64–97.

Cohen, Jacob (1988), "The effect size index: d". In *Statistical power analysis for the behavioral sciences* (Second ed., pp. 20–26). Hillsdale: Lawrence Erlbaum Associates.

Conboy, Barbara T. & Thal, Donna J. (2006), "Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers", *Child Development* 77(3): 712–735.

Corson, Peter (1995), *Using English words*. Dordrecht: Kluwer Academic Publishers.

Duursma, Elisabeth, Romero-Contreras, Silvia, Szuber, Anna, Proctor, Patrick, Snow, Catherine, August, Diane & Calderon, Margerita (2007), "The role of home literacy and language environment on bilinguals' English and Spanish vocabulary development", *Applied Psycholinguistics* 28(1): 171–190.

Ebbels, Susan H., McCartney, Elspeth, Slonims, Vicky, Dockrell, Julie E. & Norbury, Courteney F. (2019), "Evidence-based pathways to intervention for children with language disorders", *International Journal of Language & Communication Disorders* 54(1): 3–19. https://doi.org/10.1111/1460-6984.12387

Edwards, Susan, Letts, Carolyn & Sinka, Indra (2011), *The New Reynell Developmental Language Scales.* [Language test]. London: GL Assessment.

Edwards, Susan, Letts, Carolyn & Sinka, Indra (2017), *The New Reynell Developmental Language Scales.* [Language test]. Swedish version by Lundeborg Hammarström, Inger, Kjellmer, Lisen & Hansson, Kristina. Stockholm: Hogrefe.

Fuste-Herrmann, Belinda, Silliman, Elaine R., Bahr, Ruth H., Fasnacht, Kyna S. & Federico, Jeanne E. (2006), "Mental state verb production in the oral narratives of English- and Spanish-speaking preadolescents: An exploratory study of lexical diversity and depth", *Learning Disabilities Research & Practice* 21(1): 44–60.

Garcia, Georgia E. (1991), "Factors influencing the English reading test performance of Spanish-speaking Hispanic children", *Reading Research Quarterly* 26(4): 371–392.

Genesee, Fred, Lindholm-Leary, Kathryn, Saunders, William & Christian, Donna (2005), "English language learners in U.S. schools: An overview of research findings", *Journal of Education for Students Placed at Risk* 10(4): 363–385.

Gottardo, Alexandra (2002), "The relationship between language and reading skills in bilingual Spanish-English speakers", *Topics in Language Disorders* 22(5): 46–70.

Gottardo, Alexandra, Collins, Penny, Baciu, Iuliana & Gebotys, Robert (2008), "Predictors of grade 2 word reading and vocabulary learning from grade 1 variables in Spanish-speaking children: Similarities and differences", *Learning Disabilities Research & Practice* 23(1): 11–24.

Grönqvist, Hans & Niknami, Susan (2017), "The school achievements of refugee children: lessons from Sweden", *Nordic Economic Policy Review* 520: 159–183.

Hammer, Carol Scheffner, Lawrence, Frank R. & Miccio, Adele W. (2007), "Bilingual children's language abilities and early reading outcomes in Head Start and kindergarten", *Language, Speech, and Hearing Services in Schools* 38(3): 237–248.

Hart, Betty & Risley, Todd R. (1995), *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H Brookes Publishing.

Hickey, Tom (1972), "Bilingualism and the measurement of intelligence and verbal learning ability", *Exceptional Children* 39(1): 24–28.

Hoff, Erica (2003a), "Causes and consequences of SES-related differences in parent-to-child speech", in Marc H. Bornstein & Robert H. Bradley (Eds), *Socioeconomic status, parenting, and child development.* Mahwah, US: Lawrence Erlbaum Associates Publishers, pp. 147–160.

Hoff, Erica (2003b), "The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech", *Child Development* 74(5): 1368–1378.

Hoff, Erica (2013), "Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps", *Developmental Psychology* 49(1): 4–14. https://doi.org/10.1037/a0027238

Hollingshead, August B. (1975), *Four factor index of social status* [Working Paper published by the author].

Håkansson, Gisela, Salameh, Eva-Kristina & Nettelbladt, Ulrika (2003), "Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment", *Linguistics* 41: 255–288.

Law, James, Charlton, Jeanna, Dockrell, Julie, Gascoigne, Marie, McKean, Christina & Theakston, Anna (2017), *Early Language Development: Needs, provision, and intervention for preschool children from socio-economically disadvantage backgrounds.* Newcastle University. https://eprints.ncl.ac.uk/file_store/production/242270/99F15C18-813F-4447-BE90-0BBDAA211E0C.pdf

Law, James, McBean, Kirsty & Rush, Robert (2011), "Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage", *International Journal of Language & Communication Disorders* 46(6): 657–664. https://doi.org/10.1111/j.1460-6984.2011.00036.x

Letts, Carolyn, Edwards, Susan, Sinka, Indra, Schaefer, Bianca & Gibbons, Wendy (2013), "Socio-economic status and language acquisition: children's performance on the new Reynell Developmental Language Scales", *International Journal of Language & Communication Disorders* 48(2): 131–143. https://doi.org/10.1111/1460-6984.12004

Letts, Carolyn & Sinka, Indra (2011), *The New Reynell Devemopmental Language Scales. Multilingual toolkit*. London: GL Assessment.

Lindholm, Kathryn J., Padilla, Amado M. & Romero, Arturo (1979), "Comprehension of relational concepts: Use of bilingual children to separate cognitive and linguistic factors", *Hispanic Journal of Behavioral Sciences* 1(4): 327–343.

Magnuson, Katherine A., Sexton, Holly R., Davis-Kean, Pamela E. & Huston, Aletha C. (2009), "Increases in maternal education and young children's language skills", *Merrill-Palmer Quarterly* 55(3): 319–350.

Marchman, Virginia A., Martínez-Sussmann, Carmen & Dale, Philip S. (2004), "The language-specific nature of grammatical development: Evidence from bilingual language learners", *Developmental Science* 7(2): 212–224.

McClintock, Evie & Baron, James (1979), "Early intervention and bilingual language comprehension", *Hispanic Journal of Behavioral Sciences* 1(3): 229–245.

Miller, Jon F., Heilmann, John, Nockerts, Ann, Iglesias, Aquiles, Fabiano, Leah & Francis, David J. (2006), "Oral language and reading in bilingual children", *Learning Disabilities Research & Practice* 21(1): 30–43.

Murphy, Kimberley A., Language and Reading Research Consortium (LARRC) & Farquharson, Kelly (2016), "Investigating profiles of lexical quality in preschool and their contribution to first grade reading", *Reading and Writing* 29(9): 1745–1770. https://doi.org/10.1007/s11145-016-9651-y

Nayeb, Laleh, Lagerberg, Dagmar, Sarkadi, Anna, Salameh, Eva-Kristina & Eriksson, Mårten (2021), "Identifying language disorder in bilingual children aged 2.5 years requires screening in both languages", *Acta Paediatrica* 110(1): 265–272. https://doi.org/10.1111/apa.15343

Oller, D. Kimbrough, Pearson, Barbara Z. & Cobo-Lewis, Alan B. (2007), "Profile effects in early bilingual language and literacy", *Applied Psycholinguistics* 28(2): 191–230.

Organisation of Economic Co-operation and Development (1995), *Our Children at Risk*. OECD/CERI.

Pace, Amy, Alper, Rebecca, Burchinal, Margaret R., Golinkoff, Roberta Michnick & Hirsh-Pasek, Kathy (2019), "Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school", *Early Childhood Research Quarterly* 46: 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001

Páez, Mariela & Rinaldi, Claudia (2006), "Predicting English word reading skills for Spanish-speaking students in first grade", *Topics in Language Disorders* 26(4): 338–350.

Pan, Barbara Alexander, Rowe, Meredith L., Singer, Jidotj D. & Snow, Catherine E. (2005), "Maternal correlates of growth in toddler vocabulary production in low-income families", *Child Development* 76(4): 763–782.

Peña, Elizabeth, Bedore, Lisa M. & Rappazzo, Christina (2003), "Comparison of Spanish, English, and bilingual children's performance across semantic tasks", *Language, Speech, and Hearing Services in Schools* 34(1): 5–16.

Place, Silvia & Hoff, Erica (2011), "Properties of dual language exposure that influence 2-year-olds' bilingual proficiency", *Child Development* 82(6): 1834–1849. https://doi.org/10.1111/j.1467-8624.2011.01660.x

Proctor, C. Patrich, Carlo, Maria, August, Diane & Snow, Catherine (2005), "Native Spanish-speaking children reading in English: Toward a model of comprehension", *Journal of Educational Psychology* 97(2): 246–256.

Restrepo, Maria Adelaida & Silverman, Stacy W. (2001), "Validity of the Spanish Preschool Language Scale-3 for use with bilingual children", *American Journal of Speech-Language Pathology* 10(4): 382–393.

Saenz, Terry Irvine & Huer, Mary Blake (2003), "Testing strategies involving least biased language assessment of bilingual children", *Communication Disorders Quarterly* 24(4): 184–193.

Salameh, Eva-Kristina, Håkansson, Gisela & Nettelbladt, Ulrika (2004), "Developmental perspectives on bilingual Swedish-Arabic children with and without language impairment: a longitudinal study", *International Journal of Language and Communication Disorders* 39: 65–91.

Sanchez, Serafin V., Rodriguez, Billie Jo, Soto-Huerta, Mary Esther, Villarreal, Felicia Castro, Guerra, Norma Susan & Flores, Belinda Bustos (2013), "A case for multidimensional bilingual assessment", *Language Assessment Quarterly* 10(2): 160–177. https://doi.org/10.1080/15434303.2013.769544

SCB (2021), https://Statistikdatabasen.scb.se, Demografi: Antal personer med utländsk eller svensk bakgrund (fin indelning) efter region, ålder och kön (år 2002–2020). Valda variabler: Region—riket, utländsk/svensk bakgrund—alla fyra tillgängliga variabler, ålder—1–6 i 1-årsklasser, kön—inte valt, och år—2020.

Schwab, Jessica F. & Lew-Williams, Casey (2016), "Language learning, socioeconomic status, and child-directed speech", *Wiley interdisciplinary reviews - Cognitive science* 7(4): 264–275. https://doi.org/10.1002/wcs.1393

Schwartz, Amy Ellen & Stiefel, Leanna (2006), "Is there a nativity gap? New evidence on the academic performance of immigrant students", *Education Finance and Policy* 1(1): 17–49. https://doi.org/10.1162/edfp.2006.1.1.17

Semel, Eleanor, Wiig, Elisabeth & Secord, Wayne A. (2013), *Clinical evaluation of language fundamentals, fourth edition (CELF-4), Swedish version*. Stockholm: Pearson Assessment.

Umbel, Vivian M. & Oller, D. Kimbrough (1994), "Developmental changes in receptive vocabulary in hispanic bilingual school children", *Language Learning* 44(2): 221–242.

Umbel, Vivivan M., Pearson, Barbara Z., Fernandez, Maria C. & Oller, D. Kimbrough (1992), "Measuring bilingual children's receptive vocabularies", *Child Development* 63(4): 1012–1020.

Wiig, Elisabeth H., Secord, Wayne A. & Semel, Eleanor (2004), *The Clinical Evaluation of Language Fundamentals* (Preschool 2nd edition ed.). San Antonio, TX. The Psychological Corperation: Harcourt Assessment, Inc.

Zambrana, Imac Maria, Ystrom, Eivind & Pons, Fransisco (2012), "Impact of gender, maternal education, and birth order on the development of language comprehension: A longitudinal study from 18 to 36 months of age", *Journal of Developmental & Behavioral Pediatrics* 33(2): 146–155. https://doi.org/10.1097/DBP.0b013e31823d4f83

Appendix A

Appendix A

Date: _____ Subject id: _____

Questions based on Hollingshead, and language background forms from www.sprakenshus.se

Answer in as much detail as possible and ask for explanations as soon as you think they would help you to move on.

Background

Child's date of birth (YY MM DD): _____

Gender of the child (tick a box):

Girl ☐ Boy ☐ Other ☐ Do not want to specify ☐

---

1. Your relationship with the child:

_____

2. Level of education:

Mother	Father	Other caregiver (if applicable, specify): _____	
☐	☐	☐	Less than 7th grade
☐	☐	☐	Junior high school (9th grade)
☐	☐	☐	Partial high school (10th or 11th grade)
☐	☐	☐	High school graduate
☐	☐	☐	Partial college (at least one year) or specialized training
☐	☐	☐	Standard college or university graduation
☐	☐	☐	Graduate professional training (MA, MS, eller PhD)

Appendix A

Date: _____ Subject id: _____

Language history:

3. a) What country are you from? _____

b) When did you arrive in Sweden? _____

4. a) Which country is the other parent from? _____

b) When did the other parent arrive in Sweden? _____

5. Which language(s) is/are the child's mother tongue:

---

7. At what age did the child start learning Swedish: _____

8. On a typical **weekday** (Monday-Friday), of the languages that the child **hears,** how much is mother tongue, Swedish or other languages? (indicate in % so that the total is 100%)

_____	_____	_____
Mother tongue	Swedish	Other language (please specify)

9. On a typical **weekday** (Monday-Friday), of the languages **spoken** by the child, how much is mother tongue, Swedish, or another language? (indicate in % so that the total is 100%)

_____	_____	_____
Mother tongue	Swedish	Other language (please specify)

Appendix A

Date: _____ Subject id: _____

10. On a typical **weekend** (Saturday-Sunday), of the languages that the child **hears,** how much is mother tongue, Swedish or other languages? (indicate in % so that the total is 100%)

_____	_____	_____
Mother tongue	Swedish	Other language (please specify)

11. On a typical **weekend** (Saturday-Sunday), of the languages **spoken** by the child, how much is mother tongue, Swedish or other languages? (indicate in % so that the total is 100%)

_____	_____	_____
Mother tongue	Swedish	Other language (please specify)

12. How many hours does the child spend at the preschool during a typical week: _____

13. Does the child receive any mother tongue education through the preschool?

☐ Yes ☐ No

13b). If Yes:
How many hours a week? _____

Thanks for all your help!