



Arabic-Swedish-Speaking Children Living in Sweden: Vocabulary Skills in Relation to Age, SES and Language Exposure

UTE BOHNACKER 

RIMA HADDAD 

LINNÉA ÖBERG 

**Author affiliations can be found in the back matter of this article*

RESEARCH



ABSTRACT

This paper investigates the receptive and expressive vocabulary skills of 100 Arabic-Swedish-speaking children ages 4;0–7;11 growing up in Sweden. We explore how vocabulary in this under-researched population is affected by age, socio-economic status (SES), age of onset, daily exposure and home language use in the family (parents, siblings, extended family and friends) and via mother tongue instruction. Comprehension and production of nouns and verbs were assessed with the Arabic and Swedish versions of the Cross-linguistic Lexical Tasks (CLTs; Haman et al., 2015). Background information was collected via a parental questionnaire. In our cross-sectional study, comprehension was better in the minority home language (Arabic) than in the majority language (Swedish) for the youngest (4-year-old children), but this difference levelled out at ages 5, 6 and 7. There was a clear and positive effect of age on receptive and expressive vocabulary scores in both languages. For neither language was there any effect of SES (parental education). Age of onset and daily exposure had a measurable effect on Swedish vocabulary scores, whilst for Arabic, daily exposure and input in the home played an important role: Children whose parents mostly spoke Arabic to them had significantly higher Arabic vocabulary scores than other children. The complex interplay of environmental and individual-level factors on vocabulary skills is also illustrated by four case studies. These results from a Swedish context complement vocabulary studies of other language combinations and reveal the importance of input for the development of vocabulary in bilingual children.

CORRESPONDING AUTHOR:

Ute Bohnacker

Uppsala University, SE

ute.bohnacker@lingfil.uu.se

KEYWORDS:

Arabic; Cross-linguistic Lexical Task; exposure; home language use; Swedish; vocabulary

TO CITE THIS ARTICLE:

Bohnacker, U., Haddad, R., & Öberg, L. (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): 4, pp. 1–18. DOI: <https://doi.org/10.16993/jhlr.37>

In Sweden, more and more children are growing up with several languages; 29% speak a home language other than Swedish, the majority language of society (National Authority for Education, 2021). These children's bilingualism is not well researched, and their minority home language development has hardly been documented at all. (*Home languages* are defined as languages that are not the majority language of society and are spoken by the child's family and/or close community.) Of the minority home languages spoken by Swedish preschool and school-age children, Arabic is by far the most prevalent: according to the National Authority for Education (2017), 25% of the children in *förskoleklass* (a preparatory year in between nursery school and primary school), and 23% of the children in primary and lower secondary school (grades 1–9) are Arabic-speaking.

That so many children are growing up with Arabic and Swedish is a result of Sweden's history of migration. Migrants from Arabic-speaking countries have mostly come to Sweden as refugees or asylum seekers, fleeing wars, conflict, persecution and hardship, or as spouses or close relatives of refugees, so-called 'family unification' (Swedish Migration Agency, 2017). Large-scale immigration from Arabic-speaking regions did not start until the 1980s and can be summed up in three waves: The first wave, in the 1980s, came from Lebanon in the wake of the civil war; the second wave, mainly from Iraq, came in the 1990s and 2000s, following the 1990/91 Gulf war and the 2003 Iraq war; and the third, most recent and largest wave came from Syria after civil war broke out in 2011, culminating in 2015 (Bohnacker, 2017; Statistics Sweden, 2017). As a result, Arabic-Swedish-speaking bilingual children include first-generation children born abroad (having mostly arrived – directly or indirectly – from Syria as refugees with their families, or from Iraq via family unification), as well as second-generation 'heritage-language' Arabic-speaking children born in Sweden (whose parents migrated mainly from Iraq and Lebanon, often having been refugees themselves or relatives of a refugee).

After Swedish, Arabic is now known as the language with the second-highest numbers of native speakers in the country (National Authority for Education, 2021; Parkvall, 2016). Yet very little is known about the home language and majority language skills and development of these Arabic-Swedish-speaking children (pace Holmström, 2015; Salameh, 2003, 2011). Studies of larger groups have been lacking, and the influence of age and factors such as SES and exposure has not been investigated. The present paper explores the lexical skills of Arabic-Swedish-speaking children for a relatively large group ($N = 100$, ages 4;0–7;11), in both the minority home language Arabic and the majority language Swedish, as modulated by age and a number of environmental factors.

A large body of research has shown that vocabulary is a domain that is strongly affected by language input; and as children receive more input in one language, their vocabularies grow, and scores on vocabulary assessment tasks are expected to increase. Outside of Sweden, bilingual vocabulary development has regularly been found to be influenced by age, SES and patterns of language input/exposure; however, these factors do not necessarily affect the home language and the majority language in the same way. For instance, many studies document robust vocabulary gains with age in preschoolers and school-age pupils for the majority language, whilst vocabulary growth with age in the minority home language is often weaker, despite substantial input (e.g., Gagarina et al., 2017; Ganuza & Hedman, 2019; Gathercole et al., 2013; Hoff et al., 2014; Lindgren, 2018). SES, often operationalised via parental education, has also been found to affect vocabulary development differentially, where high SES is positively related to gains in the majority language but not in the minority language (Armon-Lotem et al., 2011; Cobo-Lewis et al., 2002a, 2002b; Leseman, 2000; Meir & Armon-Lotem, 2017; Prevoo et al., 2014). How strongly age and environmental factors affect bilingual vocabulary may also vary depending on how vocabulary knowledge is measured, which age range is studied, and whether receptive or expressive vocabulary is assessed (Buac et al., 2014; Gathercole et al., 2013; Thordardottir, 2011). Moreover, bilingual vocabulary development is influenced by whether the language pair is closely related or not, with many or few cognates, since less exposure is necessary to learn new words that sound (nearly) alike and mean the same (Lindgren & Bohnacker, 2020; Sheng et al., 2016). Arabic and Swedish are typologically distant languages with very few cognates, and the vocabulary tasks used in the present study contain no cognates. Arabic/Swedish children thus do not get any 'lexical help' from their home language Arabic when learning words in the majority language Swedish, or vice versa.

The Swedish national context differs from many other countries in terms of social circumstances, institutionalised childcare and minority home language support. Nursery school ('preschool') starts at a very young age (age 1 or 2), and most children, including migrant children, attend nursery school for a major part of the day (6–8 hours), as it is offered to everyone irrespective of income or status. For children who have Arabic as their home language, just as for other bilingual groups, early and extensive nursery attendance may positively impact Swedish proficiency, and with time, some children might develop a preference for Swedish (over Arabic). Swedish is the principal national language and has majority sociocultural status. At the same time, official language policy encourages multilingualism and the upkeep of home languages other than Swedish (Education Act, 2010; Languages Act, 2009), including home language education, which is called *modersmålsstöd* 'mother tongue support' in nursery school and *modersmålsundervisning* 'mother tongue instruction' (MTI) in *förskoleklass* and school. These are extra classes of typically 40–60 min/week devoted to developing oral proficiency (and later, literacy) in the minority home language, typically organised by the municipality. In addition, Arabic lessons for children are organised via private initiatives by associations and congregations. Arabic is one of the languages with the highest MTI attendance (43% in *förskoleklass*, and 66%, in school, National Authority for Education, 2017), and is widely used as a family home language. Moreover, a number of nurseries nowadays have a substantial intake of Arabic-speaking children, and also hire Arabic-speaking staff (including substitute staff). This means that Arabic-speaking children may have the opportunity to communicate, at least partly, in their home language with staff and other children at their Swedish daycare institution, outside the home. These factors are likely to support the upkeep of the minority language Arabic.

In Sweden, the lexical skills of bilingual children have only recently begun to be investigated in relation to environmental factors. Bohnacker et al. (2016) studied effects of age and language input on expressive vocabulary in the home languages (German, Turkish) of 38 German-Swedish-speaking and 40 Turkish-Swedish-speaking preschoolers (ages 4;0–6;11). In their cross-sectional sample, there were no age-related vocabulary gains in the minority language, neither for German nor Turkish. However, there were clear input effects: Children whose parents spoke only or mostly the minority language to the child and to each other scored significantly higher on expressive vocabulary than those children whose parents did not. For children who received less minority language input at home, having friends who spoke that language measurably boosted their minority language vocabulary scores.

Öztekin (2019) investigated effects of age, SES and language input on the receptive and expressive vocabulary skills of 102 Turkish-Swedish-speaking children ages 4;0–8;1; this study included the 40 Turkish-Swedish bilinguals from Bohnacker et al. (2016). In Öztekin's cross-sectional sample, receptive and productive vocabulary scores increased with age in both languages, but these gains were greater in the majority language Swedish than in the home language Turkish (in line with international findings). SES (parental education) did not affect the scores on any of the vocabulary measures. There were clear input effects though. Children whose parents spoke only or mostly Turkish to the child and to each other had significantly higher Turkish vocabulary scores than children whose parents did not. Children who received 80% or more Swedish input during the day had significantly higher Swedish vocabulary scores than other children. A longitudinal follow-up confirmed Öztekin's cross-sectional findings: When a subsample of 10 Turkish-Swedish-speaking children was retested two years later, every child improved their receptive and productive vocabulary scores in Swedish from age 4 to age 6, but not in Turkish. For Turkish, some children showed an increase, others stagnated or decreased in vocabulary scores, which could in part be related to changes in family language activities (Öztekin, 2019, Ch. 6). For the minority language then, it was not primarily age, but rather language input quantity and quality that made a difference for continued vocabulary development.

Despite the fact that more than one fourth of Sweden's children are growing up as bilinguals, there are very few studies that investigate children's home-language vocabulary as well as Swedish, and, as summarised above, even fewer studies do so in relation to exposure and/or other environmental factors. The current study aims to address this knowledge gap by exploring the lexical skills of 100 Arabic-Swedish-speaking children in both languages. The children's receptive and expressive vocabularies in Arabic and Swedish are assessed with comparable vocabulary tasks. This paper investigates how lexical performance is related to

chronological age, socio-economic status (SES, as measured by parental education), age of onset, daily exposure, and patterns of home language use. Four case studies are discussed as well. The study thus contributes to the growing international knowledge base on vocabulary development in bilingual children, but from a Swedish perspective.

METHOD

PARTICIPANTS

The participants were 100 Arabic-Swedish-speaking children aged 4;0–7;11 (see [Table 1](#)), growing up in Eastern Central Sweden. They were recruited by contacting a large number of (pre)schools, as well as associations and congregations that offered activities for Arabic-speaking children, and via personal connections of our Arabic-speaking research team. Only children who were able to speak both Arabic and Swedish were included in the study. The children attended 53 different (pre)schools in the Greater Stockholm and Uppsala regions.

	4-YEAR-OLDS	5-YEAR-OLDS	6-YEAR-OLDS	7-YEAR-OLDS	TOTAL
N	22	25	29	24	100
Girls/boys	12/10	10/15	12/17	16/8	50/50
Mean age	4;5	5;6	6;6	7;7	6;1
Age range	4;0–4;11	5;0–5;11	6;0–6;11	7;1–7;11	4;0–7;11

Table 1 Participants, sex and age (years; months).

Parents of children willing to participate gave informed written consent and filled in a background questionnaire. The participant information presented below is based on the data from the questionnaires. At the time of testing, children did not have any known hearing, language or neuropsychiatric disorder according to parental report.

Roughly half of the children (54%) were born in Sweden, and the others (46%) had migrated with their family from an Arabic-speaking country (or in one case, from a third country). All children had Arabic as their home language, mostly a Middle Eastern (Levant) variety, as is typical for the Arabic-speaking population in Sweden (Bohnacker, 2017). According to parental report, the majority of the children spoke Syrian (42%), Palestinian (27%) or Iraqi (17%) varieties; 9% spoke Lebanese and 4% Egyptian Arabic. The children were often exposed to more than one variety of Arabic, beyond the dominant variety spoken at home. Most children were bilingual, but a handful of children were also exposed to a third language, which was either Kurdish (Sorani) or English.

In nearly all families (95%), both parents were native (L1) speakers of Arabic; in a few cases, information on this was available only for one parent, and in three cases definitive information was missing. In one family, both parents stated that their L1 was not Arabic (but presumably Neo-Aramaic). Virtually all parents were first-generation immigrants, with residence lengths varying from 10 months to 31 years. A few parents had come to Sweden as children, but most had immigrated as adults. Only one parent had been born in Sweden.

With these family constellations, it will come as no surprise that 98% of the children (98/100) had received regular input in Arabic from birth. One child was reported to have started to hear Arabic shortly after age 1, and for one child, such information was missing. For Arabic then, there was hardly any variation in age of onset. By contrast, age of onset varied considerably for Swedish. 30% of the children were exposed to Swedish before age 2;0 (only 6% before age 1;0). For 42%, regular exposure to Swedish started between 2;0 and 3;11, for 25% between 4;0 and 5;11, and for 2% not until after age 6;0. (For one child, this information was missing.) Most children were thus sequential bilinguals, acquiring Swedish as a second (L2) language. Twenty children had had less than two years (24 months) of exposure to Swedish at the time of testing. Via (pre)school, these children were immersed in the Swedish language and could complete all tasks in Swedish. We decided not to exclude children with short residence lengths or late exposure to Swedish a priori, as we wanted to explore the relationship between length of exposure and vocabulary knowledge statistically. As long as the children could complete the tasks in both languages, they were included in the study.

All children attended Swedish institutional childcare, mostly 25–40 hours a week. The 4- and 5-year-olds, as well as three 6-year-olds, attended *förskola* (nursery/preschool). All other 6-year-olds attended Swedish-medium *förskoleklass* (a preparatory year for primary school), and the 7-year-olds were in first grade of primary school. Generally, schooling was in Swedish, but 15 children had attended or were attending a bilingual Arabic-Swedish nursery and 2 children a bilingual English-Swedish nursery, according to parental report.

Many participants lived in linguistically and culturally diverse, socio-economically disadvantaged urban neighbourhoods. The individual children came from a wide variety of socio-economic backgrounds, both concerning parental occupations and education, where all levels from less than six years of primary education to doctorate degrees were represented (i.e., levels 0–8 on the 9-level ISCED 2011 classification (UNESCO Institute for Statistics, 2012). This variation can be considered typical of Arabic-speaking families in Sweden.

MATERIALS

Cross-linguistic Lexical Tasks (CLTs)

The Cross-linguistic Lexical Task (CLT; Haman et al., 2015) is a picture-based vocabulary task consisting of four parts: noun comprehension, verb comprehension, noun production, and verb production. Each part has 30 test items and 2 practice items, with a maximum score of 60 points for comprehension and production each. The comprehension parts are picture-identification tasks; the child hears the target word embedded in a prompt question (e.g., *Who is pouring?* for verb comprehension) and has to select the corresponding picture from an array of four coloured pictures. The production parts are picture-naming tasks, where the child is shown one coloured picture at a time and has to answer the experimenter's prompt question (e.g., *What is that?* for noun production) with a word. CLTs are currently available in 29 languages (www.psychologia.pl/clts/) and have been specifically developed for assessing vocabulary in both languages of bilingual children.¹ The present study used the Swedish CLT version (Ringblom et al., 2014) and an adapted Arabic CLT version (Haddad, 2017), which are suitable for the age range in question and the only existing vocabulary tasks that are comparable for Swedish and Arabic. An earlier Lebanese Arabic CLT version developed in Lebanon (Khoury Aouad Saliby et al., 2017) could not straightforwardly be used with our participants, as only few spoke Lebanese Arabic. The existing Lebanese version was therefore adapted to the Arabic dialects most relevant in the Swedish context. For the CLT production tasks, the Lebanese target words needed to be complemented by other dialect synonyms, particularly Syrian, Palestinian and Iraqi, as well as Modern Standard Arabic (MSA). For the CLT comprehension tasks, new prompt questions had to be constructed for all test items in the relevant dialects, so as not to disadvantage children by asking them about a word in a variety they do not understand. This process of adaptation involved extensive consultation of Arabic dialect dictionaries and Semitics experts, checks with parents of participating children, and systematic consultations of native-speaker informants (ages 30–45 years) from nine different locations in the Levant and Iraq, speaking different varieties, as well as MSA. The informants received lists of the target concepts in Swedish, English or Lebanese Arabic, and were asked how they would express the concept in their variety of Arabic. By comparing and classifying the responses, up to 5 lexical forms could be identified for each target concept that were likely to be known to a young speaker of a certain variety. These lexical forms were then incorporated into the prompts for CLT comprehension and added as synonyms to the target answers for CLT production. The adaptations were piloted. Thus, four different adaptations were developed for Syrian, Palestinian, Lebanese and Iraqi Arabic (Haddad, 2017, in progress).

Parental questionnaire

The parental questionnaire used in the present study was specifically developed in 2014–2016 for a large-scale childhood multilingualism research project, BiLI-TAS (Bohnacker 2013–2019). Questionnaire prototypes were constructed in several languages, including Swedish and Arabic. They were discussed with native-speaker linguists, minority community members and speech-language pathologists, piloted, and finalised after some changes. The questionnaire contains 36 questions about family background and attitudes, variety/varieties of Arabic, the

child's language development, input quality and quantity in both languages. Parents filled in the questionnaire in the language of their choice (Arabic or Swedish). Here, we analyse the answers to the following questionnaire items, and relate them to the child's vocabulary scores.

- For how long has the child been exposed to the majority language Swedish?
- How much daily exposure does the child currently receive (i) in the home language Arabic, (ii) in Swedish?
- What is the level of education of each parent?
- Which language(s) does each parent speak to the child, and how much?
- Does the child hear the home language from (i) siblings, (ii) friends/playmates, (iii) extended family and friends of the family?
- Does the child receive home language instruction?

Procedure

The study has been planned and carried out in accordance with Swedish legislation on research ethics and data protection and adheres to the university ethical code of conduct (Codex) that came into place half-way through the BiLI-TAS research project.

Each child was assessed with the CLT as part of a test battery that also included narrative tasks and non-word repetition tasks in each language. Children were seen on two separate occasions, one in Arabic, and one in Swedish; the order of the language was counterbalanced. Sessions lasted 30–45 minutes (including roughly 15 minutes for the CLT) and were held in a convenient location for the family – in a quiet room at (pre)school, in the home, or at a community centre, with a median interval of 7 days between sessions. Tasks were administered by trained native speakers (the second and third authors, and two Arabic-speaking research assistants). The experimenter spoke to the child only in the language of testing. This was done to be able to assess the knowledge of Arabic and Swedish as separate, individual languages. Care was taken to match the variety of Arabic spoken by the experimenter and the CLT vocabulary items as closely as possible to the variety spoken by the child. All sessions were audio- and video-recorded, so that child responses could be checked afterwards.

The CLT was administered via coloured picture booklets, following the standard procedure described in Haman et al. (2015, p. 221). The experimenter gave neutral feedback (*aha*, *mhm*, *okay*) regardless of whether the child's response was correct or not. Responses were written down on paper forms. The child was always praised at the end, irrespective of the actual outcome, and rewarded with stickers.

DATA TREATMENT

Scoring

All CLT child responses were transcribed and scored. The maximum score for each subtask was 60 points. Every child completed all four subtasks. The total number of responses was 24,000 (= 100 children × 2 languages × 120 test items (i.e., 60 for comprehension + 60 for production)). The scoring was done by the authors (native speakers of Swedish and Arabic).

One point was awarded for each correct response in the language of testing. For the comprehension tasks, only target picture identification was scored as correct. For the production tasks, a point was awarded if the child produced the target word, for example, Arabic *dabdab*, *ṣaḥaḥ* or *ḥaba* ('crawl') on the Arabic CLT, or Swedish *krypa* ('crawl') on the Swedish CLT, in response to a picture of a baby crawling. Moreover, the following responses were also scored as correct: (i) adult-like synonyms, (ii) words that were more specific than the target word and corresponded to the picture (e.g., Swe. *meta* 'to angle' instead of the target *fiska* 'to fish'), and (iii) word forms that were pronounced slightly off-target but were still recognisable as the target lemma. All other types of responses were scored as incorrect. Thus, words not in the target language, words that corresponded to the picture but were less specific than the target word (e.g., Swe. *städa* 'clean' instead of *sopa* 'sweep'), paraphrases and circumlocutions, forms belonging to a different word class, and forms that phonologically and/or morphologically strongly deviated from the target word,

were scored zero. The scoring of items was carefully checked for consistency. Unclear items were discussed by the authors and Arabic- and Swedish-speaking team members until consensus was reached. Whenever necessary, the audio and video recordings were consulted.

This resulted in individual CLT scores for each child and language.

Statistical analysis

CLT vocabulary scores are reported separately for comprehension and production, and for the home language (Arabic) and the majority language (Swedish). First, we investigated effects of age as a continuous variable (in months) on the vocabulary scores, using four linear regression models. Then we explored effects of the following environmental factors on the vocabulary scores: (i) age of onset, (ii) current daily exposure, (iii) SES, (iv) hearing language from parents, (v) hearing language from siblings, (vi) hearing home language from friends/playmates, (vii) hearing home language from extended family and friends of the family, (viii) home language instruction.

Age of onset (AoO) for Swedish is the reported age from which the child regularly received input in Swedish (this often coincided with preschool entry). AoO information was missing for one child. For statistical purposes, AoO was transformed into length of exposure to Swedish, by subtracting AoO (in months) from the child's chronological age (in months). Then, Pearson correlations were run. Length of exposure for Arabic could not be tested statistically separately from chronological age, as it was uniform (98/100 children were exposed to Arabic from birth).

The child's current *daily exposure* to each language was estimated on a seven-point scale ranging from almost only Arabic (Arabic 95%/Swedish 5%) to almost only Swedish (95% Swedish/5% Arabic). Parents could also note a different distribution. For one child, information on daily exposure was missing. We used a three-way split and categorised children into 'mostly exposed to Arabic' (60%–95% Arabic, $N = 31$), 'even exposure' (50%/50%, $N = 38$), and 'mostly exposed to Swedish' (60%–95% Swedish, $N = 30$). To explore the effect of daily exposure on vocabulary scores, one-way ANOVAs were run.

For all other, binary factors, independent samples t-tests (Welch) were carried out on the mean vocabulary scores.

SES was operationalised as parental education. The questionnaire queried the highest level of education of each parent. Free-form answers were coded according to the 9-level ISCED 2011 classification (UNESCO Institute for Statistics, 2012). Education level was then averaged across both parents of a child. When information was only available for one parent (e.g., single-parent households), that parent's education level was used. Children where information on education was missing for both parents ($N = 6$) were excluded from the SES analysis. Children were classified into two groups: low-SES (ISCED 0–3, i.e., non-completed primary school up to completed secondary education, $N = 31$) vs. high-SES (ISCED 4–8, i.e. tertiary education from college up to completed doctorate, $N = 63$).

As for *hearing the home language from parents*, the questionnaire asked for estimates of how much each parent spoke to the child in each language. This was done on separate scales for each parent, ranging from '(almost) only Arabic' to '(almost) only Swedish'. According to these answers, children were classified into one of two groups: the 'only/mainly Arabic' group ($N = 79$), if *both* parents indicated that they '(almost) only' or 'mainly' spoke Arabic to the child, and the 'other' group ($N = 20$). (For one child, such information was missing. For single parents and for families who provided this information for one parent only, a single value was used.) 'Other' meant that one or both parent(s) spoke a substantial proportion of a language other than Arabic (e.g., Swedish) to the child.

The factors *hearing the home language from siblings*, *from friends/playmates*, *from extended family and friends of the family* were each queried by a yes-or-no question (e.g., 'Does your child hear Arabic from siblings?'). Whether the child attended Arabic home language lessons (either in the form of municipal classes or via private initiatives), was also queried by a yes-or-no question. Based on the parents' answers, the child was assigned to a 'yes' or 'no' group for each of these factors.

RESULTS

OVERALL RESULTS IN THE TWO LANGUAGES

Table 2 shows the results for all four CLT subtasks for the entire group (ages 4;0–7;11).

	ARABIC COMPREHENSION	ARABIC PRODUCTION	SWEDISH COMPREHENSION	SWEDISH PRODUCTION
Mean (SD)	47.5 (7.5)	32.7 (12.3)	45.6 (10.8)	30.8 (11.7)
Range	25–59	1–53	18–60	10–53

Table 2 CLT scores for all participants (N = 100). For each subtask, max = 60.

Unsurprisingly, the vocabulary scores were significantly higher for vocabulary comprehension than for production in both Arabic ($t(99) = 18.51$, $p < .001$, $d = 1.85$ (large effect size)) and Swedish ($t(99) = 30.56$, $p < .001$, $d = 3.06$ (large effect size)). Within each modality, there was no significant difference between the languages (comprehension: $t(99) = 1.71$, $p = .090$; production $t(99) = 1.04$, $p = .301$).

AGE

The overall scores in the previous section hide a development with age for all four vocabulary measures, as illustrated by the breakdown by age group in **Tables 3** and **4**.

	4-YEAR-OLDS (N = 22)	5-YEAR-OLDS (N = 25)	6-YEAR-OLDS (N = 29)	7-YEAR-OLDS (N = 24)
Arabic Comprehension				
Mean (SD)	41.5 (7.6)	46.7 (6.5)	48.5 (7.5)	52.4 (3.6)
Range	25–52	27–56	31–58	45–59
Arabic Production				
Mean (SD)	25.5 (12.2)	32.6 (11.7)	34.5 (13.2)	37.1 (9.2)
Range	1–42	11–48	10–53	16–51

Table 3 Arabic CLT scores by age group. Max = 60.

	4-YEAR-OLDS (N = 22)	5-YEAR-OLDS (N = 25)	6-YEAR-OLDS (N = 29)	7-YEAR-OLDS (N = 24)
Swedish Comprehension				
Mean (SD)	36.0 (8.3)	44.9 (9.0)	46.9 (10.2)	53.3 (8.5)
Range	18–52	29–59	27–60	27–60
Swedish Production				
Mean (SD)	22.3 (7.2)	29.1 (10.0)	31.6 (11.5)	39.5 (11.2)
Range	10–41	15–48	11–48	12–53

Table 4 Swedish CLT scores by age group. Max = 60.

At age 4, vocabulary comprehension in the home language Arabic was significantly better than in the majority language Swedish ($t(21) = 2.33$, $p = .030$), though effect size was small ($d = 0.50$). No differences obtained at other ages. (In the interest of space, these statistics are not reported here.)

Performance on all four vocabulary measures increased with age (Arabic comprehension ($F(1,98) = 32.07$, $R^2 = 0.25$, $p < .001$), Arabic production ($F(1,98) = 11.6$, $R^2 = 0.11$, $p < .001$), Swedish comprehension ($F(1,98) = 34.79$, $R^2 = 0.26$, $p < .001$) and Swedish production ($F(1,98) = 26.4$, $R^2 = 0.21$, $p < .001$)). **Figures 1** and **2** illustrate this age development in the form of regression lines drawn on scatterplots of child scores plotted against their age in months. Comprehension scores in Arabic are comparatively high and close together; whilst production scores show a large spread. For both Arabic and Swedish comprehension, some children score at or near ceiling, but for vocabulary production, none do. The steeper slope of the regression lines and the size of the respective R^2 -values indicate that vocabulary scores increase more strongly with

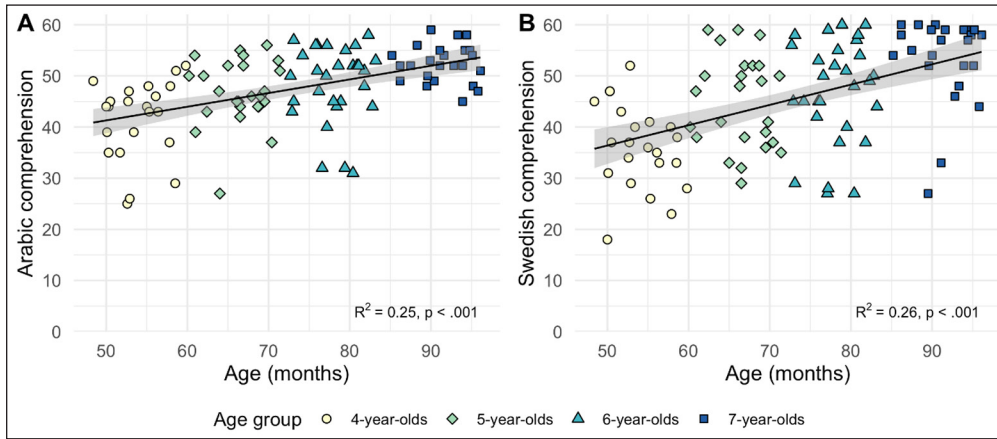


Figure 1 The relationship between age (in months) and vocabulary comprehension scores for Arabic **(A)** and Swedish **(B)**. The band along the regression line indicates a 95% confidence interval.

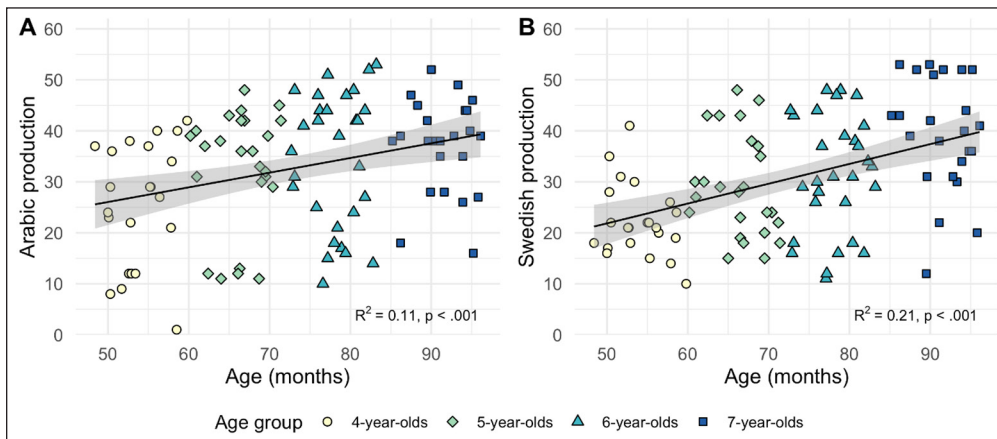


Figure 2 The relationship between age (in months) and vocabulary production scores in Arabic **(A)** and Swedish **(B)**. The band along the regression line indicates a 95% confidence interval.

age in the majority language Swedish than in the home language Arabic, particularly so for vocabulary production. Despite these age trends, individual scores are very scattered. In the following sections, we explore environmental factors that may explain some of this variability in the data.

AGE OF ONSET FOR SWEDISH

Age of onset, here recoded as a continuous variable of length of exposure to Swedish, was positively correlated both with Swedish vocabulary comprehension and production scores. Both correlations (Pearson) were significant and of medium strength (comprehension: $N = 98$, $r = 0.63$, $p < .001$; production: $N = 98$, $r = 0.61$, $p < .001$). Since the children vary greatly regarding age of onset (see Participants section), this result was not surprising. Age and LoE to Swedish showed some degree of co-variation (as expected), but it was not very large ($N = 97$, $r = 0.44$, $p < .001$). Thus, they both add unique explanatory value.

CURRENT DAILY EXPOSURE

One-way ANOVAs showed a significant effect of the proportion of language exposure a child receives throughout the day: Children who received the majority (60%–95%) of their daily exposure in one language had significantly higher vocabulary production and comprehension scores in that language (Arabic comprehension: $F(2,96) = 3.82$, $p = .025$, $\eta^2 = 0.07$; Arabic production: $F(2,96) = 14.36$, $p < .001$, $\eta^2 = 0.23$; Swedish comprehension: $F(2,96) = 9.95$, $\eta^2 = 0.17$, $p < .001$; Swedish production: $F(2,96) = 14.99$, $p < .001$, $\eta^2 = 0.24$).² The boxplots in **Figure 3** visualise this relationship for vocabulary comprehension (3A) and vocabulary production (3B).³ In both languages, the effects of exposure are stronger for production than for comprehension, as suggested by the size of eta-squared.

In order to explore the data further, we plotted the individual children's Arabic vocabulary production scores by age, estimated daily exposure, and dialectal variety of Arabic. In the scatterplot in **Figure 4**, the shape of the dots indicates the dialectal variety of the child (Syrian, Palestinian, Iraqi, Lebanese, Egyptian), and the colour of the dots indicates exposure. Children of all ages are spread across the three exposure groups (mostly Arabic,

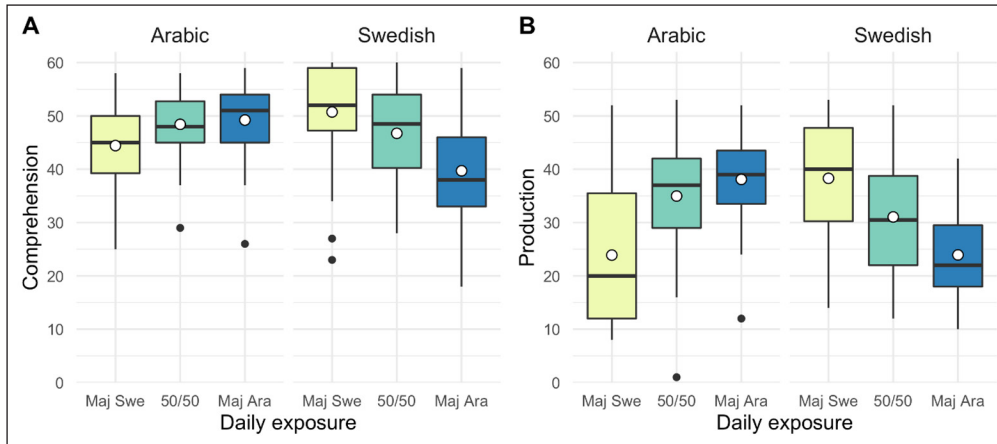


Figure 3 Vocabulary comprehension scores (left) and production scores (right) per exposure group: ‘Maj(ority) Swedish’ (60%–95% daily exposure to Swedish) vs. ‘50/50’ (even, 50% Swedish/50% Arabic) vs. ‘Maj(ority) Arabic’ (60%–95% daily exposure to Arabic). Max score = 60. White dots indicate means.

50/50, or mostly Swedish). High and low scores are spread across all dialects, and the three exposure groups also spread out across the dialects, which suggests that the task does not disadvantage any particular dialect group. However, 19 children of different ages with very low Arabic production scores cluster at the bottom of the graph. For nearly all of them, daily exposure was ‘mostly Swedish’, which suggests a relation between daily exposure and CLT scores.

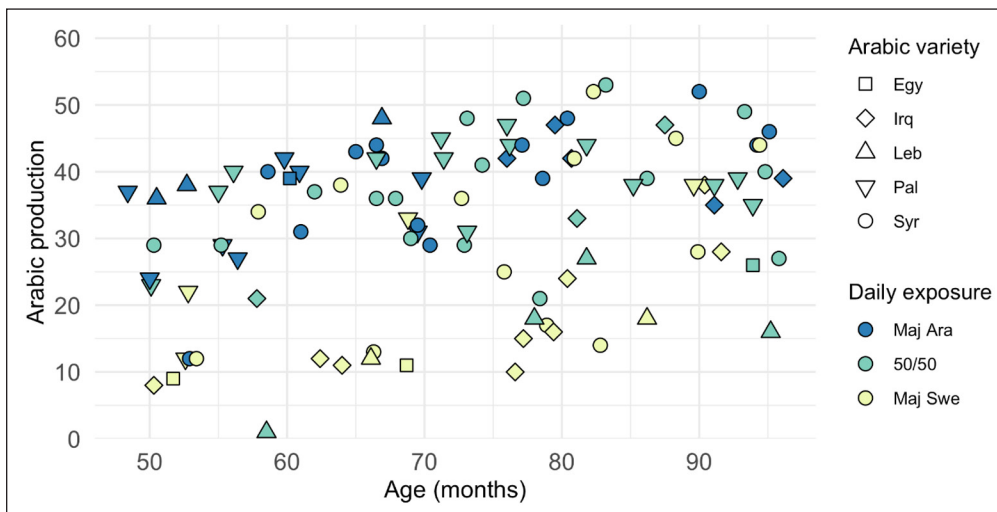


Figure 4 Arabic vocabulary production scores as a function of age (in months), in relation to proportion of current daily exposure and dialectal variety.

SOCIO-ECONOMIC STATUS (SES)

As described in the Method section, SES was operationalised via parental education, with a split between no tertiary education (‘low-SES’, ISCED-levels 0–3, $N = 31$) and tertiary education (‘high-SES’, ISCED-levels 4–8, $N = 63$). Welch’s t -tests showed no effect of SES for any of the four vocabulary measures (Arabic comprehension $t(47.06) = -0.21$, $p = .832$; Arabic production $t(62.64) = 1.81$, $p = .074$; Swedish comprehension $t(63.09) = -1.72$, $p = .090$; Swedish production $t(64.52) = -1.31$, $p = .196$).⁴

HEARING THE HOME LANGUAGE FROM PARENTS

It was very common that both parents reported that they spoke only/mostly Arabic to the child ($N = 79$). There were clear effects of the parents’ language use, as illustrated in the boxplots in **Figure 5**. Children whose parents spoke to them only/mostly in Arabic scored significantly higher in both Arabic comprehension ($t(23.65) = 3.42$, $p = .002$, $d = 0.95$ (large effect size)) and production ($t(30.05) = 4.12$, $p < .001$, $d = 1.02$ (large effect size)) than the other children ($N = 20$). Children whose parents spoke only/mostly in Arabic to them did not score lower in Swedish comprehension than other children ($t(29.51) = -1.16$, $p = .256$). They did score lower in Swedish vocabulary production though the difference was barely significant ($t(28.34) = -2.08$, $p = .047$, $d = -0.52$ (moderate effect size)).

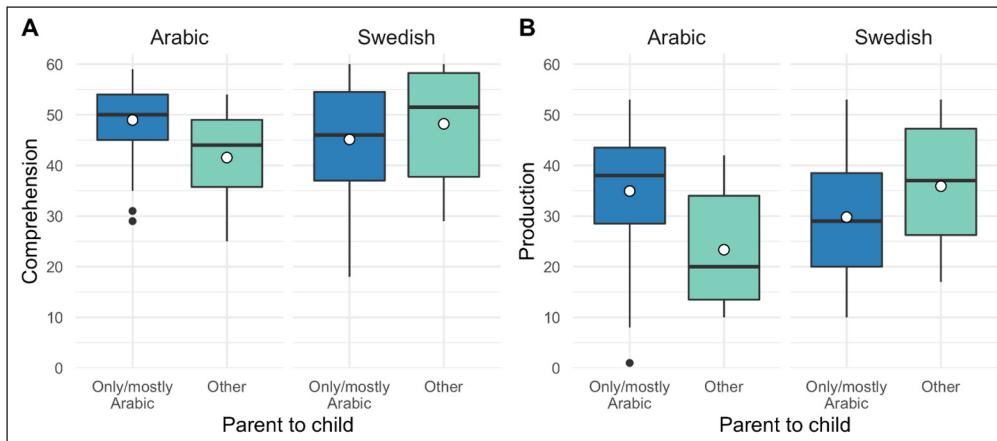


Figure 5 Vocabulary comprehension scores (left) and production scores (right) by parent-to-child home language input: ‘Only/mostly Arabic’ vs. ‘Other’. Max score = 60.

HEARING THE HOME LANGUAGE FROM SIBLINGS

Seven children did not have siblings and were excluded from the sibling analyses. Most children ($N = 69$) were reported to hear at least some Arabic from their sibling(s). Such children had significantly higher Arabic vocabulary production scores than the other children ($N = 26$) ($t(45.93) = -2.25$, $p = .029$, $d = -0.51$ (moderate effect size)). There was no difference between the groups for Arabic vocabulary comprehension ($t(55.55) = -1.41$, $p = .165$) nor for Swedish comprehension ($t(52.26) = -1.65$, $p = .105$). Children who heard Arabic from their sibling(s) did however have somewhat lower Swedish production scores than those with no Arabic from their sibling(s) ($t(52.58) = 2.33$, $p = .0244$, $d = 0.51$ (moderate effect size)).

HEARING THE HOME LANGUAGE FROM FRIENDS AND FROM EXTENDED FAMILY

We also investigated potential input effects from friends or playmates who speak (at least some) Arabic to the child. According to parental report, a third of the children ($N = 30$) had Arabic-speaking friends; 69 children did not. When comparing the two groups, having Arabic-speaking friends had a very small, barely significant, positive effect on Arabic production ($t(74.01) = -2.05$, $p = .044$, $d = -0.42$). For Arabic comprehension scores, the very small difference just failed to reach significance ($t(79.75) = -1.95$, $p = .055$, $d = -0.39$).

The children who heard Arabic from extended family and friends of the family ($N = 74$) were compared to those who did not receive any such input ($N = 25$). There were no significant differences between the groups (Arabic comprehension ($t(41.98) = -1.04$, $p = .302$); Arabic production ($t(45.45) = 0.45$, $p = .651$)).

MOTHER-TONGUE INSTRUCTION (MTI)

Two thirds of the children ($N = 65$) were attending Arabic home language classes. 34 children did not attend such classes. Arabic lessons were mostly organised by the municipality (in the form of ‘mother tongue support’ for pre-schoolers or ‘mother tongue instruction’ for *förskoleklass* and school-age children), or in some cases privately by associations or congregations. Attendees scored significantly higher than non-attendees in Arabic vocabulary comprehension ($t(50.54) = -3.84$, $p < .001$, $d = -0.86$ (large effect size)) and production ($t(53.16) = -3.62$, $p < .001$, $d = -0.80$ (moderate effect size)). However, attendance did not distribute evenly across age: Whilst the majority of 4- and 5-year-olds did not attend Arabic classes, nearly all 6-year-olds (79%) and 7-year-olds (92%) did attend. Since older children tend to have higher Arabic vocabulary scores (see Age section), the difference between MTI-attendees and non-attendees may not be a real ‘MTI effect’ but largely an effect of age.⁵

DISCUSSION

For this sample of 100 Arabic-Swedish-speaking preschoolers and first-graders growing up in Sweden, there was a clear and positive effect of age on receptive and expressive vocabulary in both Arabic and Swedish, as measured via CLT scores. As the CLT is a relatively new tool, it is worth stating that there were no floor effects or any pronounced ceiling effects (though a handful of 5-, 6-, and 7-year-olds scored near ceiling on the comprehension tasks).

The increase in vocabulary production scores with age was stronger for the majority language Swedish, in line with similar trends observed in other settings (e.g., for English/Spanish in the USA: Cobo-Lewis et al., 2002a, 2002b; Hoff et al., 2014; for English/Welsh in Wales: Gathercole et al., 2013; for Hebrew/Russian in Israel: Armon-Lotem et al., 2011; for Swedish/Somali in Sweden: Ganuza & Hedman, 2019; for Swedish/German in Sweden: Lindgren & Bohnacker, 2020). However, the present results do not quite match those of an earlier study of 40 Turkish-Swedish-speaking and 38 German-Swedish-speaking 4-to-6-year-old preschoolers growing up in Sweden (Bohnacker et al. 2016). For these children's home-language expressive vocabulary, no effect of age could be discerned (no results were reported for the majority language). Possibly, there may have been a lack of power in the study by Bohnacker et al. (2016) (group size and age range were smaller than in the present study). Our results do align with a study of 102 Turkish-Swedish-speaking children who were identical in age range with our participants and were assessed with equivalent CLT tasks (Öztekin, 2019). Even though overall vocabulary gains with age were stronger in Swedish, Öztekin's youngest participants showed better comprehension in their minority home language (Turkish) than in the majority language Swedish, just like our participants. At older ages, such differences between the languages evened out. However, Öztekin's youngest participants (age 4 and 5) had far better vocabulary production in the home language (Turkish) than in Swedish, whilst our youngest group had equally low vocabulary production in Arabic and Swedish. For Öztekin's oldest children (age 6 and 7), vocabulary comprehension and production scores did not differ between the languages, just as they did not differ in the present study.

Even though our data evinced clear age effects for both receptive and expressive vocabulary, scores varied substantially across children. For the oldest children ranges were still very large (Tables 3 and 4), and especially for expressive vocabulary, variation did not appear to decrease with age (Fig. 2). This suggests that bilingual vocabulary development is strongly affected by variables other than chronological age, and we therefore proceeded to investigate a number of these.

Age of onset to (L2) Swedish correlated significantly with Swedish vocabulary; the earlier the children were exposed, the higher were their comprehension and production scores. Since length of exposure is often related to cumulative amount of input in a language, this finding is hardly surprising and also in agreement with other reports in the literature (e.g., Armon-Lotem et al, 2011; Gagarina et al., 2017). It is documented here for the first time for bilingual children growing up in a Swedish context.

The proportion of *current daily exposure* also had a measurable effect on vocabulary scores: Children who received their daily input mainly in one language (60%–95%) had significantly higher vocabulary production and comprehension scores in that language; this held for both the majority and the minority language. These results rhyme well with international findings (e.g., Buac et al., 2014; Klassert & Gagarina, 2010; Thordardottir, 2017) and also with Öztekin's (2019) findings of daily exposure effects for 102 Turkish-Swedish-speaking children the same age as the participants of the present study.⁶

Another way to explore input was to look at patterns of language use in the home. Here, *hearing the minority language from parents* had a measurable positive effect: Children had significantly higher vocabulary comprehension and production scores in Arabic when both parents mostly spoke in Arabic to the child. This is in agreement with international findings (e.g., Buac et al., 2014; Cobo-Lewis et al., 2002b; Hoff et al., 2014; Klassert & Gagarina, 2010; Thordardottir, 2011), but shown here for the first time for bilingual Arabic-speaking children. In addition to parents, we also investigated the effect of *other home-language input providers* such as siblings, friends/ playmates and extended family and friends of the family on the children's vocabulary scores, but except in the case of siblings (cf. Bridges & Hoff, 2014), their contributions did not reach significance, at least not on our coarse-grained measure ('Does the child hear any Arabic from X?' – Yes/No).

Participating in mother tongue instruction (MTI) did have an effect on Arabic vocabulary scores; however, since MTI attendance was strongly related to age, its attendance may not add unique explanatory value. The time allocated to MTI classes is limited (typically 1 hour/week) and may constitute a negligible amount of input for 4-to-7-year-olds.⁷ MTI and other sources of home-language input should be explored in more detail in the future.

Concerning SES, for neither Arabic nor Swedish was there any effect of family SES on vocabulary scores. This result echoes findings for same-age Turkish-Swedish bilingual children growing up in Sweden (Bohnacker et al., 2016; Öztekin, 2019), but it is at odds with a large body of international research, according to which SES influences vocabulary skills, at least for the majority language (Buac et al., 2014; Calvo & Bialystok, 2014; Cobo-Lewis et al. 2002a, 2002b; Leseman, 2000; Meir & Armon-Lotem, 2017; Pearson, 2007; Prevoo et al., 2014). There could be several reasons behind our divergent finding. One may be the way SES is operationalised, as we used the average of both parents' education levels as a proxy for SES, whilst some other studies have used maternal education, parental occupation, family income, residential area, or a combination of them. It is also possible that SES interacts differently with other kinds of vocabulary measures (and the aforementioned studies used different tests). Yet another explanation could be the Swedish setting, where children may have more equal opportunities in Sweden than in some of the settings other studies were carried out in. It is generally assumed that children from higher SES families more easily get access to the type of language learning experiences that stimulate vocabulary growth. However, daycare and preschool in Sweden is comprehensive and accessible from a very early age, regardless of parental education level and social and economic situation. As a consequence, some differences between children from different SES backgrounds may be levelled out. Language development may be influenced more strongly by the quantity and quality of input in daycare/school and patterns of language use and language-fostering activities in the home, than by a simple measure of SES.

Alongside variation dependent on age (4;0–7;11) and age of onset, we have seen that vocabulary skills vary substantially across children of similar ages. To account for some of this variation, we have explored several environmental factors for the children as a group. However, our sample consists of 100 individual children, and each child's vocabulary skills are a result of many factors, individual-level and environmental ones. We will try to showcase this complex interplay with case studies of four individual children.

Table 5 contrasts a 4-year-old girl with very high Arabic scores (BiAra4-24) with a 7-year-old girl with very low Arabic production scores (BiAra7-12). For both children, parents were born in an Arabic-speaking country and have Arabic as their L1, and family SES is low, but this is where the similarities end. BiAra4-24's family came to Sweden 3 years ago, her parents speak Arabic to her, and she hears Swedish at preschool (from age 3). Noteworthy are the family's frequent language- and literacy-fostering activities at home, plus an exceptional number of hours of MTI, all of which are likely to boost the Arabic vocabulary skills of this top-performing young child.

By contrast, BiAra7-12's parents have lived in Sweden for decades, the child was born in Sweden, entered preschool early and she hears mostly Swedish during the day. Her parents speak both Arabic and Swedish to her. BiAra7-12 has become Swedish-dominant (with very high Swedish CLT scores) and may be on her way towards receptive bilingualism, as her Arabic vocabulary production scores are very low.

Table 5 Very high Arabic scores at age 4 vs. very low Arabic scores at age 7.

BIARA4-24, AGE 4;10	BIARA7-12, AGE 7;1
- Arabic CLT comprehension score: 51/60 (age group mean: M = 41.6) - Arabic CLT production score: 40/60 (age group mean: M = 25.6)	- Arabic CLT comprehension score: 49/60 (age group mean: M = 52.4) - Arabic CLT production score: 18/60 (age group mean: M = 37.2)
- Born in Arabic-speaking country - Daily exposure: 60% Arabic/40% Swedish - Both parents: L1 Arabic - Parents' residence in Sweden: 3 + 3 years - Parents speak mostly Arabic and a little Swedish to the child - 2 older siblings speak Arabic to the child - Joint book reading in Arabic 1–2 times/week and storytelling in Arabic almost every day, none in Swedish - Arabic MTI (4 hours/week, private lessons) - Low SES (ISCED-level 3 + 2, car mechanic + housewife) - Late AoO for Swedish ("after 3;0", preschool entry age unknown) - Above-mean Swedish CLT scores for age 4 (comprehension: 38/60, production: 24/60)	- Born in Sweden - Daily exposure: 95% Swedish/5% Arabic - Both parents: L1 Arabic - Parents' residence in Sweden: 14 + 30 years - Parents speak 50% Arabic/50% Swedish to the child - 2 older siblings speak Arabic and Swedish to the child - No information on book reading and storytelling - All media consumption is in Swedish - Arabic MTI (1 hr/week, municipal) - Low-mid SES (ISCED-level 3 + 4, nursery staff + cook) - Early AoO for Swedish ("1;3", at preschool entry) - Very high Swedish CLT scores (comprehension: 60/60, production: 53/60)

Table 6 contrasts two 7-year-olds (BiAra7-15, BiAra7-17) with seemingly similar backgrounds but widely divergent vocabulary results. Both were born in an Arabic-speaking country and came to Sweden only three years ago, their parents have Arabic as their L1 and they speak (only/mostly) Arabic to the child. Older siblings also speak Arabic to the child. Average parental education (SES) is low. Both children attend MTI. For both, age of onset (AoO) to Swedish (via preschool) was late, around age 5. Nevertheless, BiAra7-15 has very high Swedish and Arabic CLT scores; comprehension is near ceiling in both languages, and production is also very high, despite late AoO. By contrast, BiAra7-17's CLT scores are low in both Swedish and Arabic.

Upon closer inspection, some environmental and individual-level differences emerge that may have influenced the children's vocabulary development. The parents of (high-performing) BiAra7-15 carry out book-reading and storytelling activities with their child nearly every day, unlike BiAra7-17's parents. BiAra7-15 attended preschool at twice the rate than BiAra7-17 (30 vs. 15 hours/week). BiAra7-17 had to change (pre)school several times due to relocation, and also stayed home for some time while waiting for a place at school, which may explain why BiAra7-17's Swedish vocabulary is less developed. Moreover, high-performing BiAra7-15 is characterised by her parents as an "early talker", with "faster language development than her peers". By contrast, BiAra7-17 is described by the parents as being fascinated by English, spending much of his spare time watching English video clips online and playing video games. As a consequence, less time is spent interacting with others in Arabic and Swedish. This may have contributed to the child's low performance on Arabic and Swedish vocabulary, along with the aforesaid schooling interruptions.

Table 6 Two 7-year-olds with similar backgrounds but very different vocabulary scores.

BIARA7-15, AGE 7;10	BIARA7-17, AGE 7;11
- Swedish CLT comprehension score: 57/60 (age group mean: M = 53.3) - Swedish CLT production score: 44/60 (age group mean: M = 29.5)	- Swedish CLT comprehension score: 44/60 (age group mean: M = 53.3) - Swedish CLT production score: 20/60 (age group mean: M = 29.5)
- Born in Arabic-speaking country - Daily exposure: 40% Arabic/60% Swedish - Both parents: L1 Arabic - Parents' residence in Sweden (3 + 3 years) - Both parents speak only Arabic to the child - Older siblings speak Arabic to the child - Joint book reading and storytelling in Swedish almost every day, and 1–2 times/week in Arabic - Arabic MTI (1.5 hrs/week, municipal) - Low SES (ISCED-levels 2 + 3; goldsmith + sales person) - Late AoO for Swedish ("after age 5;0", preschool entry at 4 but 3 months absence, preschool attendance 30h/week) - Very high Arabic CLT scores (comprehension: 58/60, production: 44/60)	- Born in Arabic-speaking country - Daily exposure: 40% Arabic/40% Swedish/20% English (via video games) - Both parents: L1 Arabic - Parents' residence in Sweden (3 + 3 years) - Parents speak only/mostly Arabic to the child, one parent speaks a little English - Older siblings speak Arabic to the child - Joint book reading in Swedish 1–2 times/week, none in Arabic, storytelling in Arabic 1–2 times week, none in Swedish - Arabic MTI (1.5 hrs/week, municipal) - Mixed SES (ISCED-level 6 + 2; engineer + unemployed) - Late AoO for Swedish ("after age 5;0"; preschool entry at 4;9, but attended only 15h/week) - Low Arabic CLT scores for age 7 (comprehension: 47/60, production: 27/60)

All in all, these case studies reflect the individual-level diversity in vocabulary development in both languages. Children may have very different proficiency levels in the two languages due to differences in their exposure history and current language use. However, children with seemingly similar backgrounds may also have different proficiency levels in each language due to differences in current language use.

CONCLUSION

Lexical skills are a cornerstone of language proficiency. Rich and diversified vocabularies help children to use language in different contexts and support their literacy development and academic achievements. As age increases, so do lexical skills, but these gains are modulated by other factors such as the quantity and quality of input, which vary for children growing up with two languages. As a result, some bilingual children will develop rich receptive and expressive vocabularies in both languages, whilst others have unevenly distributed lexical skills, and yet others may end up with only limited receptive skills in one of their languages.

This paper has presented a first, cross-sectional snapshot of this process for 100 4-to-7-year-old Arabic-Swedish-speaking children growing up in Sweden. Through the lens of the CLT vocabulary tasks, we have explored receptive and expressive vocabularies in both languages, in relation to age and a number of environmental factors. The future will tell how these children's vocabularies develop beyond age 7, but between age 4 and 7, vocabulary gains in both languages were found to be clearly and positively related to chronological age and daily exposure. For Swedish, length of exposure was important as well. Arabic vocabulary knowledge was measurably boosted if the parents predominantly spoke Arabic to the child. Thus, for the development and upkeep of minority-language lexical skills, extensive input from parents seems to be crucial. Notably, no stagnation in the minority language was evident. We suggest that this is linked to the relatively extensive use of Arabic in the children's families, and possibly also to the fact that Arabic-Swedish-speaking children in Sweden have opportunities to be exposed to and use their minority language outside the home setting.

The present study opens up many avenues for further research. The interaction of the aforesaid factors should be investigated, along with potential effects of sibling interaction, mother-tongue instruction and other input sources. It would also be interesting to study the children's vocabulary skills in relation to other aspects of language proficiency.

Finally, we would like to emphasise that our results at group level do not necessarily hold for every individual child. Although there were clear effects of age and length of exposure for the group as a whole, large vocabulary gains in a new language (here, Swedish) can also be made in a relatively short period of time, if the input conditions are favourable, as illustrated by the individual case studies. Conversely, children with longer lengths of exposure do not always do better. Although there were clear positive effects of minority-language daily exposure and parent-to-child input for the group as a whole, the use of a certain language in the home does not by itself guarantee that the child will acquire satisfactory vocabulary skills in that language.

NOTES

- 1 The CLT was developed during COST Action IS0804 (2009–2013) as part of a battery of testing materials for bilingual children. Test items are basic, concrete, every-day words with high imageability. For each CLT language version, 120 test items are selected from a common base of 300 concepts for objects and actions and paired with pictures from an accompanying picture database. Items are selected to represent a mix of different levels of difficulty and estimated age-of-acquisition (see Łuniewska et al., 2016) in the specific language. The CLTs are thus not translations of each other. For a detailed description of the CLT and of how it was constructed, see Haman et al. (2015). The validity of the CLT has been shown in a large-scale study with 17 languages (Haman et al., 2017).
- 2 The difference between the exposure groups 'mostly Swedish' and 'mostly Arabic' was significant for all four vocabulary measures. The difference between 'mostly Swedish' and 'even', and between 'even' and 'mostly Arabic' was significant for some vocabulary measures, but is not reported in detail here due to space restrictions.
- 3 In the third author's PhD thesis (Öberg, 2020), reported daily exposure to Arabic and to Swedish is treated as two separate variables ranging from 5 to 95. The level of daily exposure correlates positively with both the comprehension and production scores in the respective language, confirming the results presented here.
- 4 We also split parental education into three categories (low ($N = 31$), mid ($N = 34$), high ($N = 29$)) and ran one-way ANOVAs, but again there was no significant effect of SES for any vocabulary measure (Arabic comprehension $F(2,91) = 0.46$, $p = .634$; Arabic production $F(2,91) = 2.30$, $p = .106$; Swedish comprehension $F(2,91) = 1.48$, $p = .234$; Swedish production $F(2,91) = 0.84$, $p = .435$). Other splits, or treating parental education as a continuous variable, did not yield any significant effects either. For space reasons, these statistics are not reported here.
- 5 This question is explored further in Bohnacker et al. (2021).
- 6 Note that Öztekin put the split at 80%, not at 60%; thus for each language, she compared children with 80% or more of their daily input in one language to all other children (Öztekin, 2019: 112–113). Since very few of our participants received 80% or more of their daily input in one language, we could not test these extremes statistically.
- 7 At higher ages, MTI attendance does have measurable effects on home language vocabulary, as lessons add up year after year and also involve literacy activities (Ganuza & Hedman 2019).

ACKNOWLEDGEMENTS

This research was supported by the Swedish Research Council (VR), Grant 2013-1309, and the Bank of Sweden Tercentenary Foundation (RJ), Grant P19-0644:1, to Ute Bohnacker.

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Ute Bohnacker  orcid.org/0000-0002-6715-7470

Uppsala University, SE

Rima Haddad  orcid.org/0000-0002-3183-2622

Uppsala University, SE

Linnéa Öberg  orcid.org/0000-0001-9253-4234

Uppsala University, SE

REFERENCES

- Armon-Lotem, S., Walters, J., & Gagarina, N.** (2011). The impact of internal and external factors on linguistic performance in the home language and in L2 among Russian-Hebrew and Russian-German preschool children. *Linguistic Approaches to Bilingualism*, 1(3), 291–317. DOI: <https://doi.org/10.1075/lab.1.3.04arm>
- Bohnacker, U.** (2013–2019). Language impairment or typical language development? Developing methods for linguistic assessment of bilingual children in Sweden. *Research project funded by the Swedish Research Council (VR 421-2013-1309), commonly referred to as BiLI-TAS (Bilingualism, Language Impairment, Turkish, Arabic, Swedish).*
- Bohnacker, U.** (2017). Sveriges arabisktalande befolkning. *Internal report*. Department of Linguistics and Philology, Uppsala University (29 October 2017).
- Bohnacker, U., Haddad, R., Lindgren, J., Öberg, L., & Öztekin, B.** (2021). Ordförrådsutveckling hos arabisk-svensktalande och turkisk-svensktalande barn i förskoleåldern och vid skolstart. *Språk och stil NF*, 31(1), 75–107. DOI: <https://doi.org/10.33063/diva-434152>
- Bohnacker, U., Lindgren, J., & Öztekin, B.** (2016). Turkish- and German-speaking bilingual 4-to-6-year-olds living in Sweden: Effects of age, SES and home language input on vocabulary production. *Journal of Home Language Research*, 1, 17–41. DOI: <https://doi.org/10.16993/jh1r.26>
- Bridges, K., & Hoff, E.** (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, 35(2), 225–241. DOI: <https://doi.org/10.1017/S0142716412000379>
- Buac, M., Gross, M., & Kaushanskaya, M.** (2014). The role of primary caregiver vocabulary knowledge in the development of bilingual children's vocabulary skills. *Journal of Speech, Language and Hearing Research*, 57(5), 1804–1816. DOI: https://doi.org/10.1044/2014_JSLHR-L-13-0055
- Calvo, A., & Bialystok, E.** (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130(3), 278–288. DOI: <https://doi.org/10.1016/j.cognition.2013.11.015>
- Cobo-Lewis, A. B., Pearson, B. Z., Eilers, R. E., & Umbel, V. C.** (2002a). Effects of bilingualism and bilingual education on oral and written English skills: A multi-factor study of standardized test outcomes. In D. K. Oller, & R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 64–97). Multilingual Matters. DOI: <https://doi.org/10.21832/9781853595721-005>
- Cobo-Lewis, A. B., Pearson, B. Z., Eilers, R. E., & Umbel, V. C.** (2002b). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multi-factor study of standardized test outcomes. In D. K. Oller, & R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 98–117). Multilingual Matters. DOI: <https://doi.org/10.21832/9781853595721-006>
- Education Act (Skollagen).** SFS 2010: 800.
- Gagarina, N., Posse, D., Düsterhöft, S., Golcher, F., & Topaj, N.** (2017). Bilingual lexicon development in German in preschool children with the home languages Russian and Turkish. In H. Peukert & I. Gogolin (Eds.), *Dynamics of linguistic diversity* (pp. 125–142). John Benjamins. DOI: <https://doi.org/10.1075/hsl.6.07gag>
- Ganuza, N., & Hedman, C.** (2019). The impact of mother tongue instruction on the development of biliteracy: Evidence from Somali-Swedish bilinguals. *Applied Linguistics*, 40(1), 108–131. DOI: <https://doi.org/10.1093/applin/amx010>
- Gathercole, V. C. M., Thomas, E. M., Roberts, E. J., Hughes, C. O., & Hughes, E. K.** (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. C. M. Gathercole (Ed.), *Issues in the assessment of bilinguals* (pp. 20–55). Multilingual Matters. DOI: <https://doi.org/10.21832/9781783090105-004>
- Haddad, R.** (2017). Cross-Linguistic Lexical Tasks: Selected Arabic dialects. *Adapted from the Lebanese version (CLT-ARA)*. (Unpublished material).

- Haman, E., Łuniewska, M., Hansen, P., Simonsen, H., Gram, Chiat, S., et al. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-Linguistic Lexical Tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics*, 31(11–12), 818–843. DOI: <https://doi.org/10.1080/02699206.2017.1308553>
- Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem, N. Meir & J. de Jong (Eds.), *Assessing multilingual children: Disentangling bilingualism from Language Impairment* (pp. 196–240). Multilingual Matters. DOI: <https://doi.org/10.21832/9781783093137-010>
- Hoff, E., Rumiche, R., Burridge, A., Ribot, K. M., & Welsh, S. N. (2014). Expressive vocabulary development in children from bilingual and monolingual homes: A longitudinal study from two to four years. *Early Childhood Research Quarterly*, 29, 433–444. DOI: <https://doi.org/10.1017/S0305000910000759>
- Holmström, K. (2015). *Lexikal organisation hos en- och flerspråkiga skolbarn med språkstörning*. Unpublished PhD thesis. Lund University.
- Khoury Aouad Saliby, C., Kouba Hreich, E., & Messarra, C. (2017). Cross-Linguistic Lexical Tasks: Lebanese Version (CLT-ARA). Unpublished material.
- Klassert, A., & Gagarina, N. (2010). Der Einfluss des elterlichen Inputs auf die Sprachentwicklung bilingualer Kinder: Evidenz aus russischsprachigen Migrantenfamilien in Berlin. *Diskurs Kindheits- und Jugendforschung*, 4, 413–425.
- Languages Act (Språklagen). SFS 2009: 600.
- Leseman, P. (2000). Bilingual vocabulary development of Turkish preschoolers in the Netherlands. *Journal of Multilingual and Multicultural Development*, 21(2), 93–112. DOI: <https://doi.org/10.1080/01434630008666396>
- Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4–6*. Studia Linguistica Upsaliensia 19. Uppsala University.
- Lindgren, J., & Bohnacker, U. (2020). Vocabulary development in closely-related languages: Age, word type and cognate facilitation effects in bilingual Swedish-German preschool children. *Linguistic Approaches to Bilingualism*, 10(5), 587–622. DOI: <https://doi.org/10.1075/lab.18041.lin>
- Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., et al. (2016). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, 48, 1154–1177. DOI: <https://doi.org/10.3758/s13428-015-0636-6>
- Meir, N., & Armon-Lotem, S. (2017). Independent and combined effects of socioeconomic status (SES) and bilingualism on children's vocabulary and verbal short-term memory. *Frontiers in Psychology*, 8, 1442. DOI: <https://doi.org/10.3389/fpsyg.2017.01442>
- National Authority for Education (Skolverket). (2017). <https://www.skolverket.se/statistik-och-utvardering/statistik-i-tabeller/grundskola/skolor-och-elever> (as per school year 2016/2017). Retrieved 24 July 2017.
- National Authority for Education (Skolverket). (2021). <https://www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokC&verkform=Grundskolan&omrade=Skolor%20och%20elever&lasar=2020/21&run=1; www.skolverket.se/skolutveckling/statistik/sok-statistik-om-forskola-skola-och-vuxenutbildning?sok=SokC&verkform=F%C3%B6rskoleklass&omrade=Elever&lasar=2020/2F21>. Retrieved 14 April 2021.
- Öberg, L. (2020). *Words and non-words: Vocabulary and phonological working memory in Arabic-Swedish-speaking 4–7-year-olds with and without a diagnosis of Developmental Language Disorder*. Studia Linguistica Upsaliensia 27. Uppsala University.
- Öztekin, B. (2019). *Typical and atypical language development in Turkish-Swedish bilingual children aged 4–7*. Studia Linguistica Upsaliensia 25. Uppsala University.
- Parkvall, M. (2016). *Sveriges språk i siffror: Vilka språk talas och av hur många?* Morfem & Språkrådet.
- Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*, 28(3), 399–410. DOI: <https://doi.org/10.1017/S014271640707021X>
- Prevoo, M. J. L., Malda, M., Mesman, J., Emmen, R. A. G., Yeniad, N., Van Ijzendoorn, M. H., & Linting, M. (2014). Predicting ethnic minority children's vocabulary from socioeconomic status, maternal language and home reading input: different pathways for host and ethnic language. *Journal of Child Language*, 41(5), 963–984. DOI: <https://doi.org/10.1017/S0305000913000299>
- Ringblom, N., Håkansson, G., & Lindgren, J. (2014). Cross-Linguistic Lexical Tasks: Swedish version (CLT-SWE). Unpublished material.
- Salameh, E.-K. (2003). *Language impairment in Swedish bilingual children: Epidemiological and linguistic studies*. Lund University.
- Salameh, E.-K. (2011). Grammatisk och fonologisk utveckling på svenska och arabiska vid tvåspråkig undervisning. In L. Bergman, I. Ericsson, N. Hartsmar, L. Lang, B. Liljefors Persson & C. Ljungberg (Eds.), *Educare 3. Tema: Tvåspråkig undervisning på svenska och arabiska i mångkulturella storstadsskolor*. Malmö högskola.

- Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A.** (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, 141, 229–238. DOI: <https://doi.org/10.1016/j.jecp.2015.09.007>
- Statistics Sweden (SCB).** (2017). <https://www.statistikdatabasen.scb.se> (census data as per 31 December 2016). Retrieved 4 May 2017.
- Swedish Migration Agency (Migrationsverket).** (2017). <https://www.migrationsverket.se/Om-Migrationsverket/Statistik/Oversikter-och-statistik-fran-tidigare-ar.html>. Retrieved 4 May 2017.
- Thordardottir, E.** (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 15(4), 426–445. DOI: <https://doi.org/10.1177/1367006911403202>
- Thordardottir, E.** (2017). Are background variables good predictors of need for L2 assistance in school? Effects of age, L1, amount, and timing of exposure on Icelandic language and nonword repetition scores. *International Journal of Bilingual Education and Bilingualism*. DOI: <https://doi.org/10.1080/13670050.2017.1358695>
- UNESCO Institute for Statistics.** (2012). *International Standard Classification of Education: ISCED 2011*. <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>.

TO CITE THIS ARTICLE:

Bohnacker, U., Haddad, R., & Öberg, L. (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): 4, pp. 1–18. DOI: <https://doi.org/10.16993/jhlr.37>

Submitted: 15 June 2020

Accepted: 21 March 2021

Published: 17 June 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Home Language Research is a peer-reviewed open access journal published by Stockholm University Press.

