




ORIGINAL ARTICLE

Interdependence between L1 and L2: The case of Syrian children with refugee backgrounds in Canada and the Netherlands

Elma Blom^{1,2*} , Adriana Soto-Corominas³ , Zahraa Attar¹, Evangelia Daskalaki⁴ and Johanne Paradis⁴ 

¹Utrecht University, Department of Education and Pedagogy, Utrecht, Netherlands, ²The Arctic University of Norway UiT, Department of Language and Culture, Tromsø, Norway, ³Universitat Internacional de Catalunya, Department of Applied Linguistics, Barcelona, Spain and ⁴University of Alberta, Department of Linguistics, Edmonton, Canada

*Corresponding author. Email: W.B.T.Blom@uu.nl

(Received 8 July 2020; revised 5 March 2021; accepted 22 April 2021; first published online 15 July 2021)

Abstract

Children who are refugees become bilingual in circumstances that are often challenging and that can vary across national contexts. We investigated the second language (L2) syntactic skills of Syrian children aged 6–12 living in Canada ($n = 56$) and the Netherlands ($n = 47$). Our goal was to establish the impact of the first language (L1 = Syrian Arabic) skills on L2 (English, Dutch) outcomes and whether L1–L2 interdependence is influenced by the length of L2 exposure. To measure L1 and L2 syntactic skills, cross-linguistic Litmus Sentence Repetition Tasks (Litmus-SRTs) were used. Results showed evidence of L1–L2 interdependence, but interdependence may only surface after sufficient L2 exposure. Maternal education level and refugee camp experiences differed between the two samples. Both variables impacted L2 outcomes in the Canadian but not in the Dutch sample, demonstrating the importance to examine refugee children's bilingual language development in different national contexts.

Keywords: refugee children; sentence repetition; transfer; Interdependence Hypothesis; L2 exposure

Worldwide there are over 20 million refugees, many of whom are children (UNHCR, 2019). Due to their experiences as a refugee, these children are at an increased risk for low abilities in their first language (L1), while also learning a new and second language (L2) (Kaplan et al., 2016). Reaching a high level in both languages is important for academic (Dickinson & Tabors, 2001), socioemotional (Tseng & Fuligni, 2000), economical (Agirdag, 2014), and cognitive reasons (Bialystok et al., 2009). However, in immigration contexts, successful bilingual development is not guaranteed. The L1 may develop more slowly and could stagnate because the L2 is omnipresent, valued, and supported (Blom, 2019; Haman et al.,

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

2017; Hoff, 2018). L2 development could be negatively affected by a later age of onset (Bleakley & Chin, 2010; *cf.* Pfenninger & Singleton, 2018), and limited input (Chondrogianni & Marinis, 2011; Hammer *et al.*, 2012; Paradis, 2011; Place & Hoff, 2016; Sorenson Duncan & Paradis, 2020). Development of the L1 and L2 of children with refugee backgrounds are particular concerns because of adverse learning conditions, for example, displacement, exposure to trauma, disruption of family structures, frequent transitions, limited social support, interrupted schooling, and socioeconomic disadvantages (Graham *et al.*, 2016; Kaplan *et al.*, 2016; Sirin & Rogers-Sirin, 2015).

Adopting an ecological approach to refugee experiences (Lustig, 2010), we investigated language outcomes of Syrian refugee children across two national contexts, Canada and the Netherlands. For several years in a row, Syria has been the main country of origin for refugees in Europe and in Canada. Millions of Syrian refugees crossed the Mediterranean Sea to Europe, where most went to Germany (532,100), followed by Sweden (109,000), Austria (49,000), and the Netherlands (32,100) (UNHCR, 2019). Tens of thousands have crossed the Atlantic Ocean to North America. Canada has admitted over 70,000 refugees from Syria (Government of Canada, 2020). The specific goals of this study were to determine: (1) whether L1 Syrian Arabic abilities predicted emergent L2 English and L2 Dutch abilities, and the role of length of L2 exposure in this relation and (2) if bilingual development patterns for these Syrian refugees were similar in Canada and the Netherlands.

Refugee experiences and language development

It is often assumed that children learn language almost automatically. However, language development, including simultaneous (Hoff, 2018) and sequential bilingual development (Pfenninger & Singleton, 2019), is impacted by the environmental conditions under which children learn a language, resulting in individual differences (Armon-Lotem & Meir, 2019; Chondrogianni & Marinis, 2011; Paradis, 2011). Previous studies on child bilingual development have investigated contextual factors in children with migration backgrounds (e.g., Spanish L1 speakers in the USA, Russian L1 speakers in Israel). While these populations share many characteristics with refugee children (e.g., learning and maintaining a minority L1 and majority L2, limited L1 resources, being educated in the L2, biculturalism), different pre-migration experiences (e.g., interrupted schooling and traumatic experiences), and often *post*-migration experiences (e.g., resettlement and integration challenges) may place refugee children in a more vulnerable position for learning in general, and language learning specifically.

For example, among young refugees, the prevalence of posttraumatic stress disorder (PTSD) is high, ranging between 19 and 54%, whereas in non-refugees, the estimates are between 2 and 9% (Bronstein & Montgomery, 2011). PTSD is linked to lower cognitive functioning and may affect learning negatively (Yasik *et al.*, 2007). Parents who are refugees may struggle with mental disorders and settlement issues. Research has demonstrated that children with depressed parents have lower language outcomes (Ahun *et al.*, 2017; Paulson *et al.*, 2009). Less sensitive and

responsive parenting behaviors of depressed parents, as well as less positive engagement behaviors and less reading to their children may contribute to this effect (Paulson et al., 2009; Sohr-Preston & Scaramella, 2006). Moreover, in the case of refugee children, fewer opportunities for L1 interactions may arise because family structures are often disrupted, potentially affecting L1 acquisition. In addition, L2 acquisition could be hampered because L2 learning opportunities are confined to school hours with few one-on-one interactions. Moreover, many refugee children experience interrupted schooling before and after arriving in the host country (Browder, 2018; Herzog-Punzenberger et al., 2017), negatively affecting the acquisition of L1 and L2 academic registers (Brown et al., 2006). These examples suggest that findings about bilingual language development from previous child L2 research do not necessarily generalize to refugee children, and that targeted research on the language development of refugee children is needed.

Cross-country research could add further insights, as illustrated by the two national contexts in which the current study is situated. In Canada, more than half of the Syrian refugees are government-assisted refugee families (Korntheuer et al., 2017) who are selected based on their vulnerability and whose initial resettlement is entirely supported by the government. In contrast, in the Netherlands, the majority of Syrian refugees enter the country using private funds. In general, the socioeconomic status (SES) of Syrian refugees will thus be lower in Canada than in the Netherlands. Lower SES (often indexed by maternal education) is associated with lower L2 outcomes of both migrants (Golberg et al., 2008; Prevoo et al., 2013; Sorenson Duncan & Paradis, 2020) and refugee children (Paradis et al., 2020). In addition, Syrian refugee children in Canada are more likely than those in the Netherlands to have experienced adverse conditions in refugee camps, for example, faced violence, discrimination, poverty, and malnutrition (Harrell-Bond, 2000; Krause & Gato, 2019), which could all have an impact on cognitive and socioemotional well-being, which, in turn, could affect children's uptake of linguistic input (Soto-Corominas et al., 2020).

The current study investigated Syrian refugee children across two countries to obtain insight into developmental risk factors for low language outcomes, focussing primarily on the role of the L1. Migrant children are at risk of a lower rate of L1 development, in particular, if L2 exposure starts early (Bedore et al., 2016). In the case of Syrian refugee children, low socioemotional well-being (Soto-Corominas et al., 2020) has been found to be related to low L1 abilities. A well-developed L1 may, however, equip refugee children with an important experience that fosters L2 development, an issue to which we turn below.

L1–L2 interdependence

A prominent question in the field of child bilingualism, which has not yet been addressed for refugee children, is whether L2 learning depends on the L1, and if refugee children transfer abilities. Transfer of abilities is central in the well-known *Interdependence Hypothesis* (Cummins, 1979), which holds that access to abilities acquired in the L1 facilitates learning in the L2, and vice versa. The Interdependence Hypothesis posits that “on a cognitive level, the languages are not separate but connect with each other by means of a common underlying

proficiency” (Cummins, 2017, p. 106). This type of transfer does not refer to the transfer of linguistic structures, which can result in either positive effects if structural properties of the L1 and L2 overlap or negative effects if structural properties of the L1 and L2 differ (Castilla *et al.*, 2009; Larsen-Freeman & Long, 1991). Rather, it refers to abilities that facilitate learning and that are to some degree independent of language-specific structural linguistic properties. Cummins mentions, for example, abstract thinking, problem-solving, verbal reasoning, metalinguistic awareness, metacognitive and metalinguistic learning strategies, pragmatic aspects of language use, and conceptual knowledge as examples of transferable underlying proficiencies. Furthermore, this notion of transfer refers to processes by which both languages have access to the same store of knowledge, rather than to processes by which knowledge “moves” from one language to the other (Riches & Genesee, 2006; Rolstad & MacSwan, 2014).

In line with the assumption that the languages connect on a cognitive level, various studies demonstrate that lower level nonlinguistic literacy skills (e.g., phonological awareness, decoding abilities) may be more likely to be transferred than oral language skills (Lechner & Siemund, 2014; Melby-Lervåg & Lervåg, 2011; Proctor *et al.*, 2017; Riches & Genesee, 2006; Rolstad & MacSwan, 2014; Verhoeven, 1994). For example, Verhoeven (1994), who investigated Turkish L1/Dutch L2 children reports limited predictive effects for L1 vocabulary and morphology measures, in contrast with L1 phonological awareness and pragmatic knowledge, which showed respectively moderate and strong relationships with L2 phonological awareness and pragmatic knowledge.

Other studies do find support for L1–L2 relationships in oral language. The meta-analysis of Melby-Lervåg and Lervåg (2011) shows a significant, albeit weak, correlation between L1 and L2 oral language based on measures of vocabulary. In addition, Goodrich *et al.* (2016) found that bilingual Spanish–English children did readily acquire translational equivalents of words known in one language but not the other, suggesting that they transfer specific concepts. Further support for transfer in the domain of vocabulary comes from a study in which a vocabulary intervention in children’s home language/L1 (English) was associated with vocabulary gains in the school language/L2 (Hebrew) (Armon-Lotem *et al.*, 2021). Studies that examined correlations between L1 and L2 grammatical outcomes report more mixed findings (Méndez *et al.*, 2019). For example, Gottardo (2002) did not find significant cross-language relationships for Spanish–English bilinguals, whereas Castilla and colleagues (2009) found that L1 Spanish grammatical abilities predicted L2 English grammatical abilities.

As the two languages of a bilingual may share a conceptual store (Kroll & Stewart, 1994; Kroll *et al.*, 2010), transfer effects in the domain of vocabulary could indeed be expected, as vocabulary items consist of mappings between phonological forms and conceptual representations. In the domain of grammar, transfer effects may be related to metalinguistic awareness, specifically morphological and syntactic awareness. Morphological awareness refers to the ability to combine and decompose words into morphemes, while syntactic awareness refers to knowledge of the rule system that governs how words are combined into larger units. Several studies show that bilingual children transfer morphological awareness (for an overview, see Chen & Schwartz, 2018), and suggest that joint experiences in two languages are

associated with heightened morphological and syntactic awareness (Bialystok et al., 2014). Rauch et al. (2012) argue that metalinguistic awareness is closely related to the concept of language aptitude, which is the potential that an individual has to learn languages. It would thus be expected that children who are proficient in the L1 and have developed metalinguistic awareness in this language, will also be relatively proficient in the L2 because they transfer their metalinguistic awareness ability.

In sum, findings on L1–L2 interdependence in oral language are discrepant, warranting further research. Moreover, it is unknown to what extent the specific bilingual group of refugee children transfers abilities that they have developed in the L1 to promote learning the L2. It is not uncommon that policymakers and educators view the L1 as an obstacle for the successful L2 development of migrant children (Cummins, 2000; Van Der Wildt et al., 2017), instead of a resource. L1 skills are, however, important for socioemotional development, harmonious family relations, and for the overall well-being of migrant children (DeCapua & Wintergerst, 2009; Han, 2010; Liu et al., 2009; Tannenbaum & Berkovich, 2005). If positive relationships exist between the L1 and L2 of refugee children, this would provide additional grounds for promoting L1 maintenance.

Length of L2 exposure as a moderator

Interdependence between the L1 and L2 may be moderated by and depend on contextual factors (Verhoeven, 1994). For example, L2 exposure may influence L1–L2 relationships because interdependence might be delayed and not present at the outset of L2 development. The reason for this discontinuity in L1–L2 interdependence could be that a certain degree of L2 experience is required before children can map words in the new language to known concepts, or recruit and apply their metalinguistic abilities to foster L2 learning.

The moderating role of L2 exposure fits in with the *Linguistic Threshold Hypothesis* (Cummins, 1979), which holds that the learner needs a sufficient level of bilingualism to benefit from their bilingualism by experiencing, among other things, improved linguistic skills. Empirical evidence for linguistic thresholds is limited, because few studies have researched this topic (Prevo et al., 2015; Rolstad & MacSwan, 2014). Feinauer et al. (2017) found that among young Spanish–English bilinguals the rate of transfer depends on oral proficiency in the L2. Prevo et al. (2015) observed that children’s relative use of Turkish L1 and Dutch L2 affected the relation between Turkish and Dutch vocabulary growth. Specifically, in their study, the relationship between growth in the two vocabularies was stronger for children who used relatively more Turkish, suggesting that sufficiently developed L1 skills enable children to use the L1 as a resource. To our knowledge, no previous research has investigated the length of L2 exposure as a moderator of positive relationships between refugee children’s L1 and L2 abilities, or, more in general, in the abilities in the L1 and L2 of child second language learners. We do so in the present study by employing a sentence repetition task to measure L1 and L2 outcomes.

Present study

Few studies have investigated the language development of refugee children, even though the bilingual development in this group of children raises concerns (Kaplan *et al.*, 2016). Accordingly, in this study, we employed different adaptations of the Litmus Sentence Repetition Tasks (Litmus-SRTs) in order to tap into L1 and L2 syntactic knowledge of children who are Syrian refugees (see for a similar approach Hamann *et al.* (2020)). We did so with two main objectives in mind: (1) to investigate risk factors for refugee children's L2 abilities, with a particular focus on the role of L1 abilities and the interaction between L1 abilities and length of L2 exposure and (2) to compare patterns across two different national contexts, Canada and the Netherlands.

In terms of the first research goal, we asked: *To what extent are L2 outcomes of Syrian refugee children living in Canada and the Netherlands measured with the Litmus-SRT predicted by L1 outcomes, length of L2 exposure, and the interaction between L1 outcomes and length of L2 exposure?* Correlations between the L1 and L2, and whether L1 outcomes predict L2 outcomes provide information on L1–L2 interdependence and transfer of linguistic abilities. From prior research, we expected the length of L2 exposure to affect L2 outcomes (Blom *et al.*, 2012; Chondrogianni & Marinis, 2011; Paradis *et al.*, 2017; Roesch & Chondrogianni, 2016; Sorenson Duncan & Paradis, 2020). Crucially, and in line with the Threshold Hypothesis, we predicted that L2 exposure would moderate L1–L2 interdependence.

In order to obtain a cross-national perspective, we conducted the same research across two national contexts: Canada and the Netherlands. Collecting data in two independent samples is important for replication purposes and determining the robustness of results. It furthermore allows us to detect contextual differences that could be related to refugee policies that vary from country to country. With regards to the second research objective, we asked: *To what extent are patterns found for Syrian refugee children in Canada and the Netherlands comparable, and could different patterns be related to maternal education and refugee camp experiences?* Maternal education and refugee camp experiences were singled out as potentially relevant variables as Syrian refugees in Canada and the Netherlands are likely to differ on these variables, as discussed above. As argued in the introduction, these variables could affect L2 development, potentially modulating L1–L2 relations differently in the two national contexts. Tentatively, and based on findings suggesting that a higher level and more use of the L1 and L2 promote transfer (Feinauer *et al.*, 2017; Prevoo *et al.*, 2015), we predicted that interdependence between the L1 and L2 could be weaker in Canada than in the Netherlands, because increased refugee camp experiences and lower SES for the Syrian refugees in Canada may imply lower levels and less use of the L1 and L2.

Method

Participants

Data for the present study were collected within two partly parallel projects in Canada and the Netherlands. The Canadian project is a longitudinal study among children who migrated to Canada that started in 2018 and continues until 2021. In total, 133 children participated in the first data collection of the Canadian study

(year 2018), and 122 participated in the second one (2019) due to participant attrition (i.e., participants becoming untraceable or moving away). The data from either rounds 1 or 2 of 56 children from the Canadian sample were selected for the purpose of having two samples across the two nations that were comparable in age and L2 exposure (age at testing and arrival, and length of residence and L2 schooling). There were 30 boys and 26 girls. The average age of the children was 8 years and 8 months ($SD = 1.52$, range = 6.08–12.33). The Dutch project was of a smaller scale and was not longitudinal. Data were collected between October 2018 and January 2019 from 52 Syrian children. To achieve the optimal matching, data from 5 children with 50 months of residence in the Netherlands or more were excluded from the present study. The final Dutch sample included 16 boys and 31 girls. The average age of the children was 8 years and 7 months ($SD = 1.87$, range = 5.67–12.17). Further details on the trimming and matching processes for the two samples are given in the *Matching* section below. Information on the characteristics of the matched samples is described in the Results section.

Materials

Litmus Sentence Repetition Task (Litmus-SRT)

Litmus-SRTs were used to test L1 and L2 skills. SRTs tap into lexical knowledge (Klem et al., 2015) and verbal short-term memory (Alloway & Gathercole, 2005), but they are mostly a measure of syntactic competence (Frizelle & Fletcher, 2014; Polišenská et al., 2015). The Litmus-SRT was conceptualized and developed within the COST Action IS0804 (*Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment*) in order to create comparable SRTs for a wide range of different languages to be able to assess bilingual children's language abilities in both languages (Marinis & Armon-Lotem, 2015). An increasing number of studies have used this instrument for research purposes.¹ The Syrian Arabic Litmus-SRT consists of 31 sentences: 6 simple sentences in the active voice (with modals, negation, and adjectives), 3 coordinated sentences, 3 simple sentences in the agentless passive voice, 3 object topicalizations, 6 questions (2 who, 2 what, 2 which), 4 subordinate clauses, and 6 relative clauses. The English Litmus-SRT consists of 31 sentences as well: 6 simple sentences in the active voice (with modals and auxiliaries), 3 coordinated sentences, 6 simple sentences in the passive voice (3 short and 3 long passives, i.e. without and with a *by*-phrase), 6 questions (2 who, 2 what, 2 which), 4 subordinate clauses, and 6 relative clauses. The Dutch Litmus-SRT consists of 30 sentences: 9 simple sentences in the active voice (with modals, negation, and combinations of auxiliaries and modal verbs), 6 simple sentences in the passive voice (3 short and 3 long passives), 6 questions (2 who, 2 what, 2 which), 3 subordinate clauses, and 6 relative clauses. All stimuli are included in Appendix 1. The Litmus-SRTs were designed to be of comparable difficulty in the three languages so that correlations across languages would be meaningful. To this end, stimuli included structures that were of a similar syntactic complexity across languages. Whether transfer as defined by Cummins (1997) is dependent on typological overlap is not well researched. However, major differences between the two national settings are unlikely as L1 Syrian Arabic is an Afro-Asiatic Semitic

language, whereas L2 English and L2 Dutch are both West Germanic languages and typologically closely related.

Alberta Language Environment Questionnaire-4 (ALEQ-4, Paradis et al., 2020)

The ALEQ-4 was completed by the parents and was used to collect information on the participants, their language use, and pre-migration experiences (e.g., time spent in refugee camps, schooling in Arabic). In terms of language use, parents were asked to indicate how much Arabic versus English/Dutch children used with older and younger siblings and with parents and other relatives. This information was collected using 1-5 descriptors (1 = Mainly or only Arabic, 2 = Usually Arabic/L2 sometimes, 3 = Arabic and L2, 4 = Usually L2/Arabic sometimes, 5 = Mainly or only L2). Language input to the child and language output from the child were assessed separately. Parents were asked about the frequency with which their children engaged in language-rich activities in English/Dutch and Arabic in a given week using a 1-5 scale (1 = 0-1 hr, 2 = 1-5 hr, 3 = 5-10 hr, 4 = 10-20 hr, 5 = 20+ hr). Activities included listening/speaking activities (television, YouTube, WhatsApp, music), reading/writing activities (books, websites, messaging), playing with friends, and extracurricular activities (homework clubs, sports, religious activities). Individual rating scale scores were obtained and composite scores, estimating the richness of the English/Dutch and Arabic environments, were calculated by adding the rating scale numbers and dividing by the total number of scales to generate a proportion score. Parents were also asked questions about their own education background (including English/Dutch training), the number of children in their family, their self-rated fluency in English/Dutch (using 1-5 descriptors; 1 = Not fluent, 2 = Limited fluency, 3 = Somewhat fluent, 4 = Quite fluent, 5 = Very fluent) and use of English/Dutch outside the home (also using 1-5 descriptors; 1 = 0-1 hr, 2 = 1-5 hr, 3 = 5-10 hr, 4 = 10-20 hr, 5 = 20+ hr). The complete ALEQ-4 is available as online Supplementary Materials in Paradis et al. (2020).

Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997)

This screening questionnaire is composed of 25 items and can be used to assess emotional and behavioral difficulties in children aged 4-16. The SDQ produces five subscales: hyperactivity, conduct, emotional, peer relationship problems, and pro-social behavior. Scores can be considered independently, amalgamated into externalizing (hyperactivity and conduct) and internalizing (emotional and peer problems) scores, or all combined into one total difficulties score (Goodman et al., 2010). With validated translations into over 30 languages, including Arabic, the SDQ is one of the most widely used questionnaires on child mental health worldwide (Goodman & Scott, 1999). However, prior studies have argued that the cross-cultural validity of certain items within this test may not be completely adequate (Kersten et al., 2016; Thabet et al., 2000).

For this study, parents completed the SDQ for each participant. Both in Canada and the Netherlands, the validated Arabic translation of the SDQ was used.

Procedures

Recruitment

In Canada, participants were recruited in three English-majority cities, Edmonton, Toronto, and Waterloo, after ethics approval. In the three cities, members of the Arabic community who were involved in the project as research assistants disseminated the call for participants within the community. In Toronto, principals in public schools serving refugees were also asked to disseminate the study to students' parents. Finally, some participants were recruited via word-of-mouth. Participants who were interested and fulfilled the requirements (i.e., were of Syrian origin *and* arrived in Canada as refugees) were recruited. In the Netherlands, participants were recruited in seven cities and towns, Maarssen, Breukelen, Houten, Overvecht, Woerden, Nieuwegein and Loenen aan de Vecht, all located in the Dutch province of Utrecht. A Syrian research assistant recruited the participants through her network, among which were directors of Arabic weekend schools, and snowball sampling. In addition, social media was used. A text with relevant information for possibly interested participants was posted on Facebook in groups of the Syrian community in the Netherlands. The same text was shared via Whatsapp within Syrian networks. Parents who were interested contacted the research assistant. Their informed consent was obtained prior to data collection.

Testing

In Canada, children were either tested in their homes or in a quiet room at a school. The Litmus-SRT was part of a larger test battery (see Paradis et al., 2020). Language order was counterbalanced across participants in the full sample, so that half of the participants completed the English tasks first and half the Arabic tasks. Task order was randomized across participants. Parents completed the ALEQ-4 as an interview delivered in Arabic by a native speaker of the language. Literate parents completed the Arabic version of the SDQ about their child on their own, filling in a paper copy. Parents with low literacy skills completed the SDQ as an interview. Whether parents had sufficient skills to fill the SDQ in on their own was determined by parents themselves. In the Netherlands, most participants were tested in a quiet room at home. This was not always their own home: in each city or town, a Syrian parent offered her home for testing. A schedule was created to test the children individually and avoid overlap between participants. Part of the participants was tested in a quiet room at a neighborhood center where the Arabic school is held. Language order was counterbalanced across participants, so that half of the participants started with the Dutch SRT, and half the Arabic SRT. Arabic ALEQ-4 copies and validated Arabic translations of the SDQ were sent by post to the participants. Most parents preferred to fill in the paper copy of the questionnaire while having the research assistant over the phone.

Scoring

SRTs can be scored in different ways depending on the purpose. In both samples, SRTs were initially scored in two manners, *verbatim* and *structural accuracy*. The results of the current study are presented for verbatim scoring. In this type of

scoring, a participant's response is scored as 1 if the target sentence is repeated entirely verbatim and as 0 if the participant makes one or more changes to it. Disfluencies and phonological errors are not considered as changes. The structural accuracy scoring considered the nature of the repetition of the target syntactic structure only, disregarding errors that did not affect the target structure. In the literature, comparisons of scoring schemes, including verbatim and structural accuracy scoring, have yielded different results (Abed Ibrahim & Fekete, 2019). For this reason, we also provide information on structural accuracy scoring in Appendix 2. Since the structural accuracy scoring correlated strongly with the verbatim scoring and analyses using both as outcome variables yielded largely the same results, we did not retain the structural accuracy scoring in the main text for reasons of focus.

In Canada, seven research assistants who were native speakers of English and four research assistants who were native speakers of Arabic transcribed and scored the Litmus-SRTs for the two data collection periods. Twenty-five percent of the recordings in each language in each round were transcribed and scored by a second research assistant. An α coefficient of inter-rater reliability was obtained for each language for each round, and they all were between .91 and .94. In the Netherlands, three research assistants who were native speakers of Dutch and two assistants who were native or near-native speakers of Syrian Arabic transcribed and scored Litmus-SRT; 25% of the Litmus-SRT was independently scored a second time. The α coefficients of inter-rater reliability were 1.00 for Dutch and .99 for Arabic, indicating almost perfect agreement (Cohen, 1960).

Matching

Matching of the two samples was done by manually trimming the Canadian sample considering participants at both data collection periods. Initial preliminary analyses of the Canadian sample ($n = 133$) indicated that participants were significantly older and had a significantly shorter residence in the host country than those in the Dutch sample. These two variables were the only two considered during the trimming: older participants and participants with shorter residence were trimmed until samples were matched according to inferential tests. Of the final 56 participants from the Canadian sample, the data for 35 corresponds to the first time of data collection, and the data for 21 corresponds to the second time, with no participants being included at both time points. As mentioned above, 5 participants of the original Dutch sample ($n = 52$) were also eliminated because they exceeded the Canadian range in length of residence significantly.

Data analysis

Descriptive and inferential statistics for this study were carried out using R (version 3.6.3; R Core Team, 2020). Specific information about the analytical approach is included below according to the research objective is addressed.

Description of the matched samples

In order to compare the two samples and to explore the relations between different variables within each sample, we ran parametric two-tailed independent samples t

tests and Pearson correlations using the base package. Cohen's *d* effect sizes were calculated using the package *effsize* (version 0.7.8; Torchiano, 2020). In describing the L1–L2 performance of the matched samples on the Litmus-SRT, we employed two-tailed paired samples *t* tests.

Association between L2 outcomes and L1 level/L2 exposure

We address our first research objective in two ways. First, we employed Pearson correlations to investigate the relationship in performance in the two languages and their association to L2 exposure. Second, for each sample, we ran separate generalized mixed-effects regression models with a binomial distribution using the *lme4* package (version 1.1-21; Bates et al., 2015). For each sample, the model included L2 verbatim scores as the dependent variable, which could take the value 1 (verbatim repetition) or 0 (non-verbatim repetition). The random-effect structure included one random intercept for Participant and one for Item, without any random slopes. The initial model contained three predictors (i.e., fixed effects): L1 verbatim score, length of L2 exposure, and the interaction between the two. We then followed backwards selection on the fixed effects, which were centered around 0 and standardized to improve the interpretability of the intercept and prevent any issues related to differences in scale. At each step, the predictor with the highest *p*-value was eliminated and the reduced model was compared to the previous model using a log likelihood ratio test with the *anova* function, and visual inspection of the AIC and BIC values. The reduced model was retained when it did not entail a significant loss of model fit. The optimal models for the two samples are discussed in detail in the Results section. The C-index of concordance for these models was computed using the *somers2* function of the *Hmisc* package (version 4.3-1; Harrell, 2020), following Baayen (2008).

Effect of maternal education and refugee camp experiences on L2 outcomes

The Canadian and Dutch samples are likely to differ on SES/maternal education and refugee experiences. To address our second research goal of exploring the effect of such differences, we performed follow-up analyses. The restricted sample size limited the number of predictors we could include in the models without overfitting them. Thus, we fitted follow-up models where only one predictor was added to the optimal models for associations between L1 level/length of L2 exposure and L2 outcomes. To evaluate model fit, the above procedures were employed.

Results

Description of matched samples

Demographic information about the two samples is summarized in Table 1 (see Appendix 3 for a correlation matrix of these variables). The two groups did not differ in age (testing, arrival), length of residency in Canada/Netherlands, and schooling in the L2 (English/Dutch). In addition, there were no differences in Arabic schooling, richness of Arabic in the home, and in maternal and paternal interaction and fluency in the L2 (English/Dutch). In the Dutch sample, there was only one child with refugee camp experience. In the Canadian sample, 24 children had refugee camp experiences of

Table 1. Characteristics of the Canadian and Dutch samples

	Canada (n = 56)		Netherlands (n = 47)	
	Mean (SD)	Min–Max	Mean (SD)	Min–Max
Age (months)	105.62 (18.27)	73–148	105.23 (22.49)	68–146
Age of arrival (months)	74.48 (18.78)	45–121	72.57 (23.64)	28–117
Length of residency (months)	31.14 (6.23)	15–47	32.66 (10.23)	14–48
L2 schooling (months)	23.96 (6.32)	4–40	23.83 (8.14)	7–39
L1 schooling (months)	7.65 (9.84)	0–36	12.87 (17.68)	0–60
Time in refugee camp (months)	8.86 (16.05)	0–48	0.04 (0.29)	0–2
Number of children in family	4.30 (1.50)	1–7	3.23 (0.91)	2–5
Language use parent interaction ^{ab}	1.10 (0.23)	1–2	1.26 (0.38)	1–2.25
Language use sibling interaction ^{ac}	1.58 (0.92)	1–5	2.34 (0.97)	1–4
L2 richness	0.45 (0.14)	.20–.90	0.58 (0.19)	.25–.85
L1 richness	0.40 (0.10)	.20–.64	0.37 (0.10)	.20–.56
Maternal interaction in L2 ^d	2.14 (1.18)	1–5	2.30 (1.20)	1–5
Maternal fluency in L2 ^e	2.73 (1.00)	1–5	2.64 (0.74)	1–4
Maternal education in years	10.39 (3.70)	4–17	13.38 (3.37)	6–17
Paternal interaction in L2 ^d	2.80 (1.47)	1–5	2.94 (1.34)	1–5
Paternal fluency in L2 ^e	2.89 (1.00)	1–5	3.04 (1.01)	1–4
Paternal education in years	9.84 (4.04)	0–17.50	13.32 (4.17)	68–146

^aInformation was obtained for input to the child and output from the child and was subsequently averaged into *Language use* (Scale: 1 = Mainly or only Arabic, 2 = Usually Arabic/L2 sometimes, 3 = Arabic and L2, 4 = Usually L2/Arabic sometimes, 5 = Mainly or only L2). ^bScale is an average of language use with mother and father. ^cScale is an average of older and younger siblings. ^dCalculated from scale: 1 = 0–1 hr, 2 = 1–5 hr, 3 = 5–10 hr, 4 = 10–20 hr, 5 = 20+ hr. ^eCalculated from scale: 1 = Not fluent, 2 = Limited fluency, 3 = Somewhat fluent, 4 = Quite fluent, 5 = Very fluent.

Table 2. Descriptive information about sentence repetition outcomes in the Canadian (L1 Arabic, L2 English) and Dutch (L1 Arabic, L2 Dutch) studies, % correct (SD), number correct (SD), and minimum–maximum score^a

	L1 verbatim accuracy	L2 verbatim accuracy
Canada (<i>n</i> = 56)	66.13% (21.86) 20.5 (6.78) 5-30	35.43% (26.05) 10.98 (8.08) 0-28
The Netherlands (<i>n</i> = 47)	61.98% (25.41) 19.21 (7.88) 1-29	36.45% (29.12) 10.94 (8.74) 0-28

^aThe maximum score that could be obtained is 31, 31, and 30 for the Arabic, English, and Dutch Litmus-SRT, respectively.

between a month and 2 years. The families in the Canadian sample were larger ($t(92.711) = 4.4417, p < .001$) and the parents were, on average, less educated than the parents in the Dutch sample (mothers: $t(100.26) = -4.2903, p < .001$; fathers: $t(96.856) = -4.2798, p < .001$). In addition, in the Dutch sample the families used the L2 slightly more in interaction between parents and participants ($t(73.575) = 2.4854, p = .02$) and between siblings and participants ($t(95.731) = 4.0808, p < .001$) and L2 richness was higher compared to the Canadian sample ($t(82.322) = 3.9313, p < .001$).

Age and exposure can be related to L2 children, complicating the interpretation of these variables (Unsworth & Blom, 2010). In the Canadian sample, length of L2 schooling showed a positive correlation with age at testing, $r(54) = .34, p = .01$ indicating that older children tended to have longer L2 schooling (see Appendix 3). The strength of the correlation is weak, showing that there is relatively little overlap between age and length of L2 schooling. Length of L2 schooling did not correlate with age of arrival, $r(54) = .07, p = .60$. In the Dutch sample, the correlation between L2 schooling and age at testing, $r(45) = .25, p = .09$, and between L2 schooling and age at arrival, $r(45) = -.16, p = .27$, did not reach significance. In the present study, we use the length of L2 schooling, and not length of residency in the host country, as a proxy for L2 exposure since onset of and sustained exposure to the L2 are more likely associated with schooling in these newcomer families.

In addition to the demographic and linguistic information presented in Table 1, information was collected about the children's socioemotional well-being, and the extent to which children showed internalizing and externalizing problem behaviors, combined into a total difficulties score where higher numbers indicate more socioemotional difficulties. In the Canadian sample, the $M(SD)$ total difficulties score was 9.60 (4.00), and in the Dutch sample, it was 8.64 (5.10). This difference was not statistically significant ($t(87.165) = 1.0123, p = .31$). It should be noted that nine participants from the Canadian sample are not included in the calculation of the total difficulties score due to missing data.

In both samples, participants obtained higher scores in the L1 compared to the L2 (see Table 2). Specifically, in the Canadian sample, a two-tailed paired samples t test showed a significant difference, $t(55) = 7.8018, p < .001$ (Cohen's $d = 1.27$; large effect size). Similarly, a two-tailed paired samples t test showed a significant difference in the Dutch sample, $t(46) = 7.1748, p < .001$ (Cohen's $d = 0.93$; large effect size). These

Table 3. Pearson correlations between L1 and L2 based on Litmus-SRT scores and length of L2 schooling for the Canadian (upper right corner) and Dutch studies (lower left corner)

	L1 verbatim	L2 verbatim	L2 schooling
L1 verbatim		.25.	-.10
L2 verbatim	.61***		.30*
L2 schooling	.16	.64***	

Note: . = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Table 4. Output of optimal model for Canadian sample L2 verbatim scores. Fixed effects are scaled and centered

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-1.167	0.364	-3.201	< .001***
L1 verbatim scores	0.675	0.254	2.662	.008**
Length of L2 schooling	0.726	0.254	2.858	.004**

results suggest that most children are (still) dominant in L1 Arabic. This language dominance pattern is somewhat stronger in the Canadian compared to the Dutch sample, as shown by the larger effect sizes in the former sample.

Effect of L1 level and L2 schooling on L2 outcomes

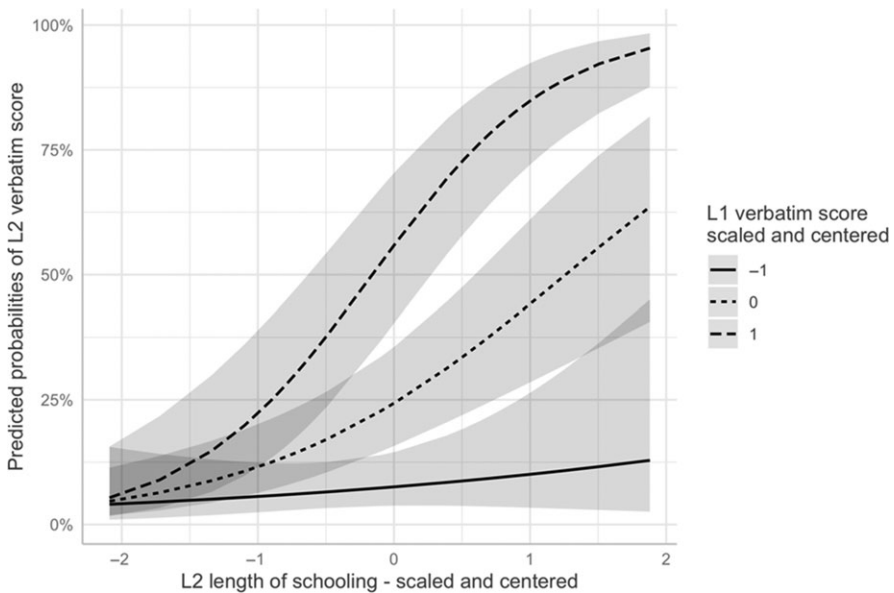
As a first approach of investigating the relationship between L1 and L2 performance and L2 schooling, we ran Pearson's correlations. The results in Table 3 demonstrate that L1 and L2 outcomes are positively correlated and correlations reach significance, though in the Canadian study this correlation shows only a trend (see also Table A4 in Appendix 2). L2 outcomes are positively and significantly correlated with the length of L2 schooling in both samples. Overall, very similar patterns arise in both samples, but the strength of the relations is considerably stronger in the Dutch compared to the Canadian sample.

To fully address our first objective, we ran a mixed-effects logistic regression for each sample (see Data analysis section). For the Canadian sample, the optimal model is shown in Table 4. For questions regarding the interpretation of logistic regression model results, which are in log-odds, we direct readers to Winter (2020). In this sample, L1 verbatim scores and length of L2 schooling were both positive predictors, and the model with the interaction between these two factors was not a better fit than the model without this interaction, $\chi^2(1, 6) = 1.8731$, $p = .17$. The optimal model had a C-index of concordance of .91, indicating excellent discrimination (Hosmer et al., 2013).

Table 5 shows the optimal model for the Dutch sample. In this sample, the interaction between L1 scores and length of L2 schooling was significant and we unpack this interaction below. This model also had outstanding discrimination ($C = .91$).

Table 5. Output of optimal model for Dutch sample L2 verbatim scores. Fixed effects are scaled and centered

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-1.135	0.277	-4.101	<.001***
L1 verbatim scores	1.370	0.211	6.485	<.001***
Length of L2 schooling	0.900	0.201	4.481	<.001***
L1 verbatim scores: Length of L2 schooling	0.585	0.236	2.475	.013*

**Figure 1.** Marginal effects of interaction between L1 verbatim scores and L2 length of schooling for the Dutch sample. This plot was created using the package *ggeffects* (version 0.14.2; Lüdtke, 2018).

In order to understand the relationship between L1 scores and L2 schooling in the Dutch model, we plotted the marginal effects of these interaction terms (Figure 1) and, using visual binning with one cut-point, compared subsamples based on shorter and longer length of L2 schooling (Table 6). The interaction depicted in Figure 1 can be broadly interpreted in the following way: in the Dutch sample, the effect of the length of L2 schooling is modulated by L1 performance, with L2 schooling having a greater effect on participants with better L1 performance. That is, participants with short L2 schooling, regardless of L1 performance, are unlikely to repeat the L2 sentence verbatim (i.e., L1 performance has little effect on participants with short L2 exposure). On the other hand, for participants with longer L2 schooling, those with better L1 performance are more likely to repeat the L2 sentence verbatim.

Table 6. L1–L2 correlations in short and long L2 exposure subsamples in the Canadian and Dutch samples

	L2 schooling mean (minimum–maximum)	Pearson correlation L1 and L2 outcomes for verbatim
Canada short ($n = 31$)	20.13 (4–24)	.06
Canada long ($n = 25$)	28.70 (25–40)	.48*
Netherlands short ($n = 26$)	17.73 (7–23)	.54**
Netherlands long ($n = 21$)	31.38 (23.5–39)	.79***

Table 7. Model for Canadian sample L2 verbatim scores including maternal education. Fixed effects are scaled and centered

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	–1.154	0.348	–3.317	< .001***
L1 verbatim scores	0.554	0.230	2.410	.016*
Length of L2 schooling	0.963	0.241	4.001	< .001***
Maternal education	0.827	0.239	3.456	< .001***

Likewise, Table 6 shows that in the Dutch sample, the correlation between L1 and L2 is considerably stronger for the long L2 exposure subsample compared to the short L2 exposure subsample. What Table 6 also demonstrates is that although the interaction effect between L1 level and L2 schooling did not reach significance for the Canadian sample, a similar pattern can be observed.

Effect of maternal education and refugee camp experiences on L2 outcomes

The model with maternal education for the Canadian sample, shown in Table 7, had a significantly better fit than the model without this variable (Table 4); $\chi^2(1, 6) = 10.858$, $p < .001$. According to the results, there was an association between more educated mothers and stronger L2 performance. Adding maternal education to the optimal model for the Dutch sample shown in Table 5 did not result in a better fit; $\chi^2(1, 7) = 0.4464$, $p = .50$.

Adding refugee camp experiences to the model without it (Table 4) resulted in a significantly better fit for the Canadian sample, $\chi^2(1, 6) = 8.4763$, $p = .004$. As shown in the model output in Table 8, the time spent in a refugee camp was a negative predictor, indicating that there was an association between longer time spent in refugee camps and lower L2 verbatim scores. We did not attempt to replicate this model in the Dutch sample since only one participant had spent time at a refugee camp.

Table 8. Model for Canadian sample L2 verbatim scores including time spent in a refugee camp. Fixed effects are scaled and centered

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-1.168	0.352	-3.321	< .001***
L1 verbatim scores	0.584	0.236	2.480	.013*
Length of L2 schooling	0.829	0.238	3.482	< .001***
Time in refugee camp	-0.728	0.242	-3.015	.003**

Discussion

In this study, we investigated the syntactic abilities of 6- to 12-year-old Syrian refugee children in Canada and the Netherlands using a multilingual assessment instrument, the Litmus-SRT. Our first goal was to determine whether or not children's L1 Syrian Arabic abilities are related to L2 English and Dutch abilities, and to establish if such L1–L2 interdependence is moderated by the length of L2 exposure. Positive relationships between the L1 and L2 indicate that the L1 could be a resource for L2 learning, instead of an obstacle. Our second goal was to determine if patterns are the same in samples from Canada and the Netherlands and whether different patterns could be due to cross-national differences.

Notable differences between the two cross-national samples, anticipated beforehand based on refugee policies, were a lower maternal education and more refugee camp experiences in the Canadian compared to the Dutch sample. Other cross-national differences included family size, with Canadian families being significantly larger, and language environments, with Dutch participants using a larger proportion of L2 in the home than Canadian participants and having comparatively richer L2 environments. The two samples were comparable in the age of arrival and testing, in participant mental health difficulties, and in length of residence and L2 schooling.

Participants in both samples scored better in L1 Syrian Arabic than in L2 English and Dutch in the Litmus-SRT, demonstrating that participants are still L1-dominant. L1 dominance patterns were slightly more pronounced in the Canadian sample. We hypothesize that this could be related to the lower maternal education in this sample. Different mechanisms could drive the relation between L1 versus L2 development and maternal education. Mothers who are more educated tend to use the L2 more frequently with their children, while less educated mothers use the L1 relatively more often (Prevo et al., 2013). This link between maternal education and L1 versus L2 use at home is reflected in our study: as mentioned above, Dutch participants were exposed to the L2 slightly more in the home and had higher L2 richness than Canadian participants. More L2 use at home could hypothetically help L2 development; this hypothesis is, however, not widely supported for school-aged children (Chondrogianni & Marinis, 2011; Kaltsa et al., 2019; Paradis et al., 2017), and it is more likely that a higher degree of L2 use negatively affects children's L1 development and/or that more highly educated mothers have more means to create opportunities for L2-rich activities for their children. In addition, the

possibility exists that there is somewhat more L2 exposure in the Dutch sample (at the expense of L1 exposure) because the Dutch Kindergartens are full day, whereas this is not the case for all of the Canadian kindergartens in Alberta/Edmonton, resulting in somewhat less hours per week of L2 exposure for 8.93% of the Canadian sample. We hope to investigate the mechanisms underlying the different cross-national patterns of dominance in future research.

L1–L2 interdependence and L2 exposure

A common approach to identify L1–L2 interdependence is to establish the amount of shared variance and investigate cross-language correlations. Previous research on cross-language correlations in the domain of grammar (morphology, syntax) reports discrepant results, raising the question of whether there is L1–L2 interdependence and whether for L1 and L2 learning a bilingual child has access to the same store of knowledge, as has been posited by the Interdependence Hypothesis. Our results demonstrate that L1 Syrian Arabic predicts children's L2 English and Dutch outcomes. These findings tie in with those of Castilla and colleagues (2009) who found that Spanish L1 morphosyntax abilities predicted English L2 morphosyntax, but differ from those of Gottardo (2002) and Verhoeven (1994) who found limited evidence for L1–L2 relations in the domain of grammar. Typological proximity (Larsen-Freeman & Long, 1991) does not explain the different findings, as both Castilla *et al.* (2009) and Gottardo (2002) investigated Spanish L1/English L2 children and our study showed interdependence between L1 and L2 abilities, despite typologically distant L1 and L2s.

In order to understand the differential findings on the interdependence of grammar skills, timing and exposure may be relevant as this transfer of skills may be dependent on children's L1 and L2 proficiency levels (Cummins, 1979; Feinauer *et al.*, 2017; Prevoo *et al.*, 2015). Sufficient L1 knowledge to positively boost the L2 may be more likely at older ages (Castilla *et al.*, 2009), and in children who use the L1 frequently (Prevoo *et al.*, 2015). Although age does not explain why Castilla *et al.* (2009), who tested younger children than Gottardo (2002), found more robust evidence of interdependence, a relatively well-developed and frequently used L1 may have supported this cross-linguistic association in the children we investigated. Both the Canadian and the Dutch samples showed that children with longer L2 exposure performed better in the L2, replicating previous research (Blom *et al.*, 2012; Chondrogianni & Marinis, 2011; Paradis *et al.*, 2017; Roesch & Chondrogianni, 2016; Sorenson Duncan & Paradis, 2018). Length of L2 exposure modulated the effect of L1 on L2 abilities in the Dutch sample in line with the Threshold Hypothesis (Cummins, 1979), confirming that children need to have some L2 experience to benefit from their L1 skills. Prevoo and colleagues (2015) also found support for moderated L1–L2 relationships. Their results highlight the importance of frequent L1 use to facilitate the cross-linguistic transfer of skills in bilingual Turkish–Dutch children, in contrast with our results, which demonstrate the importance of L2 experience for bilingual Syrian Arabic–Dutch children. Whereas, the children in our study have recently arrived in the Netherlands, are L1-dominant, and use the L1 at home most of the time, the children studied by Prevoo *et al.* (2015) are second or third-generation migrants in the Netherlands who are

exposed to the L2 from birth and are raised in homes where the L2 is frequently used. Thus, for children who have strong L1 skills and are relatively new to the L2, interdependence relies more heavily on sufficient L2 experience, while for children whose L1 skills are more under pressure, sufficient L1 experience is more important. In the Canadian sample the interaction effect between L1 ability and length of L2 exposure did not reach significance or improved the model fit. However, the L1–L2 correlations in subsamples that differed in length of L2 exposure suggested that also in the Canadian sample L1–L2 correlations were stronger for children with longer L2 exposure, in line with the Dutch data.

In sum, findings from both samples reveal significant L1–L2 interdependence, indicating that 6- to 12-year-old refugee children's L1 abilities form the basis of common underlying proficiencies that promote L2 learning even though the L1 and L2 are typologically distant. The positive L1–L2 relationships underscore the conclusion that L1 knowledge does not hinder L2 development, but might, in fact, be a useful resource for children.

Exploration of modulating factors: maternal education and refugee camp experiences

Due to different refugee policies, the Syrian refugee populations in Canada and the Netherlands are different. We singled out two variables that presented cross-national differences, maternal education, and refugee camp experiences, as factors that could weaken the associations between the L1 and the L2 in the Canadian sample specifically. While L1–L2 patterns were very similar across the Canadian and Dutch samples, supporting their robustness, they were more pronounced and stronger in the Dutch sample, in line with our predictions.

Interestingly, the two national contexts differed with respect to the role of maternal education. Both maternal education and time spent in a refugee camp predicted L2 outcomes in the Canadian sample, in addition to L1 outcomes and length of L2 schooling. Maternal education did not contribute to explaining variation in L2 Dutch outcomes, despite a higher mean than in the Canadian sample but similar spread. In line with various other studies, the Canadian sample showed that more educated mothers had children who scored better in the L2 (Golberg et al., 2008; Sorenson Duncan & Paradis, 2020). These mothers may have more resources and cultural capital to stimulate their children's development in the L2 through, for example, literacy-related activities (Prevoe et al., 2013). However, this does not seem to be the case since, as shown in Appendix 3, the correlation between L2 richness and maternal education is not significant in the Canadian sample ($r(54) = -0.02, p = .87$). It is also unlikely that more educated mothers support their children's L2 skills via oral interactions in the Canadian sample since Syrian Arabic was reported to be used most of the time in the home, as shown in Table 1. Overall low L2 use, including L2 use of highly educated mothers, could explain why maternal education did not predict L2 outcomes in the Dutch study, but it does not explain why maternal education did predict L2 outcomes in the Canadian study.

An alternative, and perhaps most likely explanation of why maternal education predicts L2 outcomes in Canada but not in the Netherlands is that English, unlike Dutch, is present in Syria. Also, English is taught as a subject at Syrian schools so more

educated mothers are more likely to help their children in the L2 from the very beginning. Thus, the onset of using English is different from the onset of using Dutch, which may have helped Syrian highly educated mothers in Canada to support their children with school tasks in the L2, stimulating their children's L2 development, even though they do not tend to use the L2 in conversational interactions with their children. Support for this hypothesis is provided by the correlations and is shown in Appendix 3. The correlation between maternal education and maternal L2 fluency is significant in the Canadian sample ($r(54) = .49, p < .001$) but not in the Dutch sample ($r(45) = .17, p = .25$). On the other hand, the correlation between maternal L2 fluency and length of residency in the host country is not significant in the Canadian sample ($r(54) = .03, p = .81$) but it is in the Dutch sample ($r(45) = .60, p < .001$). This means that Syrian mothers in the Netherlands are more fluent in Dutch due to longer residency in the host country, whereas Syrian mothers in Canada are more fluent in English due to higher education in the home country. This finding supports our hypothesis that Syrian mothers in Canada were better equipped than those in the Netherlands to support their children's L2 development from the beginning of their residency in the host country, explaining the diverging model fits. This is a possible hypothesis at this time that remains for future research to be tested.

Refugee camp experiences were not investigated in the Dutch study, as hardly any children had such experiences. Time spent in refugee camps was negatively related to L2 outcomes in the Canadian study. Living in a state of limbo, with no access to education and faced with a lack of freedom, children's development may slow down in refugee camps. In addition, they can face adverse conditions that permanently limit their ability to learn, such as malnutrition (Harrell-Bond, 2000), affecting not only the uptake of linguistic input, which, in our study could be reflected in weaker effects of length of L2 exposure, but also the strength of positive L1–L2 relations. The influence of refugee camp experiences on L2 outcomes is important for clinical practice in terms of exclusionary factors for attributing low language abilities to a congenital language disorder.

Overall, the cross-national differences observed across the two samples in this study demonstrate the need to consider factors at the macrosystem level, which may indirectly affect refugee children's bilingual development even after resettlement (Lustig, 2010).

Limitations and future research

To compare the Canadian and Dutch samples, we matched them on age and exposure. Matching increases the comparability of the samples, but it also reduced the sample size, which limited the number of predictors that we could simultaneously evaluate in the statistical models. The second limitation concerns the sampling. The samples were not a priori recruited to be representative of the refugee populations in both countries, and although they reflected what would be expected based on national refugee policies, they may not be fully representative. Third, the correlational data in our study do not provide insight into specific transfer mechanisms. In relation to this, it should be noted that the L1–L2 correlations do not necessarily stem from L1 to L2 dependencies (Castilla *et al.*, 2009). L1 development shows stable and meaningful individual differences, which are linked to endogenous factors (Kidd & Donnelly, 2020). Some children are thus better language learners than

others in both the L1 and the L2. Language learning ability, and specific mechanisms that underlie language learning such as verbal short-term memory, statistical learning, and analytical reasoning, would thus be a relevant third variable. Future research could investigate the role of third variables by researching L1–L2 relations factoring out the variance shared through general language learning ability.

Conclusions

This study on 6-to-12-year-old bilingual L1 Syrian Arabic children with refugee backgrounds found evidence of L1–L2 interdependence in terms of syntactic skills. Interdependence may however not surface in the earlier phases of L2 development when children have had only limited L2 exposure. Interdependence between the two languages provides an argument in favor of L1 maintenance in order to support L2 development. The Canadian and Dutch samples showed overlap in many demographic and linguistic variables, but there were important differences related to maternal education level and refugee camp experiences, and unique patterns of interdependence. Maternal education level impacted L2 outcomes in the Canadian but not in the Dutch sample. Refugee camp experiences predicted L2 outcomes in the Canadian sample, while in the Dutch sample, hardly any child had refugee camp experiences. It is important to examine bilingual language development in refugee children in different contexts, as they may fare differently.

Acknowledgments. The authors would like to thank all the participants and their families for taking the time to participate in the study, as well as all the research assistants who collected and processed the data. The Canadian study was made possible by the funding from the Social Sciences and Humanities Research Council of Canada through a Partnership Grant (Paradis) and an Insight Developmental Grant (Paradis). The Dutch study was funded through an Asperia Grant of the Dutch Research Council (NWO), which was given to Elma Blom.

Note

1. Albanian, Lebanese Arabic, Palestinian Arabic, Saudi Arabic, Syrian Arabic, Catalan, Gaelic, German, English, Farsi, French, Standard Modern Greek, Cypriot Greek, Hebrew, Irish, Italian, Croatian, Lithuanian, Dutch, Norwegian, Malay, Maltese, (European) Portuguese, Russian, Serbian, Spanish, and Turkish (<https://www.litmus-srep.info/>).

References

- Abed Ibrahim, L., & Fekete, I. (2019). What machine learning can tell us about the role of language dominance in the diagnostic accuracy of German LITMUS non-word and sentence repetition tasks. *Frontiers in Psychology*, 9, 2757, doi: 10.3389/fpsyg.2018.02757
- Ahun, M.N., Geoffroy, M.C., Herba, C.M., Brendgen, M., Séguin, J.R., Sutter-Dallay, A.L., Boivin, M., Tremblay, R.E., Côté, S.M. (2017). Timing and chronicity of maternal depression symptoms and children's verbal abilities. *The Journal of Pediatrics*, 190, 251–257. <https://doi.org/10.1016/j.jpeds.2017.07.007>
- Armon-Lotem, S., Restrepo, M. A., Lipner, M., Ahituv-Shlomo, P., & Altman, C. (2021). Vocabulary gains in bilingual narrative intervention. *Language, Speech, and Hearing Services in Schools*, 52(1), 436–448. https://doi.org/10.1044/2020_LSHSS-19-00117
- Armon-Lotem, S., & Meir, N. (2019). The nature of exposure and input in early bilingualism. In A. De Houwer & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 193–212). Cambridge, UK: Cambridge University Press.
- Agirdag, O. (2014). The long-term effects of bilingualism on children of immigration: student bilingualism and future earnings. *International Journal of Bilingual Education and Bilingualism*, 17(4), 449–464, doi: 10.1080/13670050.2013.816264

- Alloway, T. P., & Gathercole, S. E.** (2005). The role of sentence recall in reading and language skills of children with learning difficulties. *Learning and Individual Differences*, *15*(4), 271–282.
- Armon-Lotem, S., & Meir, N.** (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew. *International Journal of Language & Communication Disorders*, *51*, 715–731, doi: [10.1111/1460-6984.12242](https://doi.org/10.1111/1460-6984.12242)
- Attar, Z., Blom, E. & Le Pichon-Vorstman, E.** (2020). Towards more multilingual practices in the mathematics assessment of young refugee students: effects of testing language and validity of parental assessment. *International Journal of Bilingual Education and Bilingualism*, doi: [10.1080/13670050.2020.1779648](https://doi.org/10.1080/13670050.2020.1779648)
- Baayen, R. H.** (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge, UK: Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Walker, S.** (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7.
- Bedore, L. M., Peña, E. D., Griffin, Z. M., & Hixon, J. G.** (2016). Effects of Age of English Exposure, Current Input/Output, and grade on bilingual language performance. *Journal of Child Language*, *43*(3), 687–706. <https://doi.org/10.1017/S0305000915000811>
- Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H.** (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89–129, doi: [10.1177/1529100610387084](https://doi.org/10.1177/1529100610387084)
- Bialystok, E., Peets, K. F., & Moreno, S.** (2014). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied Psycholinguistics*, *35*(1), 177–191, doi: [10.1017/S0142716412000288](https://doi.org/10.1017/S0142716412000288)
- Bleakley, H., & Chin, A.** (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American Economic Journal: Applied Economics*, *2*(1), 165–92, doi: [10.1257/app.2.1.165](https://doi.org/10.1257/app.2.1.165)
- Blom, E.** (2019). Domain-general cognitive ability predicts bilingual children’s receptive vocabulary in the majority language. *Language Learning*, *69*, 292–322, doi: [10.1111/lang.12333](https://doi.org/10.1111/lang.12333)
- Blom, E., Paradis, J. & Sorenson Duncan, T.** (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular –s in child L2 English. *Language Learning*, *62*, 965–994, doi: [10.1111/j.1467-9922.2012.00715.x](https://doi.org/10.1111/j.1467-9922.2012.00715.x)
- Bronstein, I., & Montgomery, P.** (2011). Psychological distress in refugee children: A systematic review. *Clinical Child and Family Psychology Review*, *14*, 44–56, doi: [10.1007/s10567-010-0081-0](https://doi.org/10.1007/s10567-010-0081-0)
- Browder, C.** (2018). Recently resettled refugee students learning English in US high schools: The impact of students’ educational backgrounds. In Shapiro S., Farrelly R. & Curry M. J. (Eds.), *Educating Refugee-background Students: Critical Issues and Dynamic Contexts* (pp. 17–32). Bristol, UK: Blue Ridge Summit, Multilingual Matters.
- Brown, J., Miller, J., & Mitchell, J.** (2006). Interrupted schooling and the acquisition of literacy: Experiences of Sudanese refugees in Victorian secondary schools. *Australian Journal of Language and Literacy*, *29*(2), 150–162.
- Castilla, A. P., Restrepo, M., & Perez-Leroux, A. T.** (2009). Individual differences and language interdependence: A study of sequential bilingual development in Spanish-English preschool children. *International Journal of Bilingual Education and Bilingualism*, *12*(5), 565–580, doi: [10.1080/13670050802357795](https://doi.org/10.1080/13670050802357795)
- Chen, X., and Schwartz, M.** (2018). Morphological awareness and literacy in second language learners: a cross-language perspective. *Reading and Writing*, *31*, 1685–1694, doi: [10.1007/s11145-018-9866-1](https://doi.org/10.1007/s11145-018-9866-1)
- Chondrogianni, V. & Marinis, T.** (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, *1*(3), 318–345, doi: [10.1075/lab.1.3.05cho](https://doi.org/10.1075/lab.1.3.05cho)
- Chondrogianni, V.** (2008). Comparing child and adult L2 acquisition of the Greek DP: Effects of age and construction. In Haznedar, B. & Gavrusseva, E. (Eds.). *Current Trends in Child Second Language Acquisition: A Generative Perspective* (pp. 97–142). Amsterdam: John Benjamins.
- Cohen, J.** (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46, doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)
- Conti-Ramsden, G., Botting, N. & Faragher, B.** (2001). Psycholinguistic markers for Specific Language Impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*, 741–748, doi: [10.1111/1469-7610.00770](https://doi.org/10.1111/1469-7610.00770)

- Cummins, J.** (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222–251, doi: [10.3102/00346543049002222](https://doi.org/10.3102/00346543049002222)
- Cummins, J.** (2000). *Language, Power, and Pedagogy: Bilingual Children in the Crossfire*. Clevedon, UK: Multilingual Matters.
- Cummins, J.** (2017). Teaching for transfer in multilingual school contexts. In García O., Lin A., & May, S. (Eds.) *Bilingual and Multilingual Education. Encyclopedia of Language and Education* (3rd ed.) (pp. 103–115). New York, NY: Springer International Publishing.
- DeCapua, A. & Wintergerst, A. C.** (2009) Second-generation language maintenance and identity: A case study. *Bilingual Research Journal*, *32*(1), 5–24, doi: [10.1080/15235880902965672](https://doi.org/10.1080/15235880902965672)
- Dickinson, D. K., & Tabors, P. O.** (Eds.). (2001). *Beginning Literacy with Language: Young Children Learning at Home and School*. Baltimore, US: Paul H Brookes Publishing.
- Feinauer, E., Hall-Kenyon, K., & Everson, K.** (2017). Rethinking the Linguistic Threshold Hypothesis: Modeling the linguistic threshold among young Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *20*(5), 886–902. doi: [10.1017/S1366728916000626](https://doi.org/10.1017/S1366728916000626)
- Frizelle, P. & Fletcher, P.** (2014), Relative clause constructions in children with specific language impairment. *International Journal of Language & Communication Disorders*, *49*, 255–264. doi: [10.1111/1460-6984.12070](https://doi.org/10.1111/1460-6984.12070)
- Golberg, H., Paradis, J., & Crago, M.** (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, *29*(1), 41–65, doi: [10.1017/S014271640808003X](https://doi.org/10.1017/S014271640808003X)
- Goodman, R.** (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.
- Goodman, A., Lamping, D. L., & Ploubidis, G. B.** (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*(8), 1179–1191.
- Goodman, R., & Scott, S.** (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful?. *Journal of Abnormal Child Psychology*, *27*(1), 17–24.
- Goodrich, J., Lonigan, C., Kleuver, C., & Farver, J.** (2016). Development and transfer of vocabulary knowledge in Spanish-speaking language minority preschool children. *Journal of Child Language*, *43*(5), 969–992. doi: [10.1017/S030500091500032X](https://doi.org/10.1017/S030500091500032X)
- Gottardo, A.** (2002). The relationship between language and reading skills in bilingual Spanish-English speakers. *Topics in Language Disorders*, *22*(5), 46–70. doi: [10.1097/00011363-200211000-00008](https://doi.org/10.1097/00011363-200211000-00008)
- Government of Canada.** (2020). Canada - Admissions of Resettled Refugees by Country of Citizenship and Immigration Category [Data set]. http://www.cic.gc.ca/opendata-donneesouvertes/data/IRCC_Resettled_0012_E.xls
- Graham, H.R., Minhas, R.S. & Paxton, G.** (2016). Learning problems in children of refugee background: A systematic review. *Pediatrics*, *137*(6), e20153994, doi: [10.1542/peds.2015-3994](https://doi.org/10.1542/peds.2015-3994)
- Haman, E., Wodniecka, Z., Marecka, M., Szweczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miękisz, A., Kacprzak, A., Banasik, N., & Foryś-Nogala, M.** (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology*, *8*, 1444, doi: [10.3389/fpsyg.2017.01444](https://doi.org/10.3389/fpsyg.2017.01444)
- Hamann, C., Chilla, S., Ibrahim, L. A., & Fekete, I.** (2020). Language assessment tools for Arabic-speaking heritage and refugee children in Germany. *Applied Psycholinguistics*, *41*(6), 1375–1414.
- Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., & Goldstein, B.** (2012). Predicting Spanish-English bilingual children's language abilities. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1251–1264, doi: [10.1044/1092-4388\(2012\)11-0016](https://doi.org/10.1044/1092-4388(2012)11-0016)
- Han, W. J.** (2010). Bilingualism and socioemotional well-being. *Children and Youth Services Review*, *32*(5), 720–731, doi: [10.1016/j.childyouth.2010.01.009](https://doi.org/10.1016/j.childyouth.2010.01.009)
- Harrell Jr., F. E.** (2020). Hmisc: Harrell Miscellaneous. R package version 4.3-1. <https://CRAN.R-project.org/package=Hmisc>
- Harrell-Bond, B.** (2000). Are refugee camps good for children? In *New Issues in Refugee Research: Working Paper* (Vol. 29). UN. High Commissioner for Refugees (UNHCR).

- Herzog-Punzenberger, B., Le Pichon-Vorstman, E. & Siarova, H.** (2017). *Multilingual Education in the Light of Diversity: Lessons Learned, NESET II report*, Luxembourg: Publications Office of the European Union, doi: [10.2766/71255](https://doi.org/10.2766/71255).
- Hoff, E.** (2018). Bilingual development in children of immigrant families. *Child Development Perspectives*, 12(2), 80–86. doi: [10.1111/cdep.12262](https://doi.org/10.1111/cdep.12262)
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X.** (2013). *Applied Logistic Regression* (Vol. 398). Hoboken, NJ: John Wiley & Sons.
- Kaltsa, M., Tsimpli, I. M., & Argyri, F.** (2019). The development of gender assignment and agreement in English-Greek and German-Greek bilingual children. *Linguistic Approaches to Bilingualism*, 9(2), 253–288, doi: [10.1075/lab.16033.kal](https://doi.org/10.1075/lab.16033.kal)
- Kaplan, I., Stolk, Y., Valibhoy, M., Tucker, A., & Baker, J.** (2016). Cognitive assessment of refugee children: Effects of trauma and new language acquisition. *Transcultural Psychiatry*, 53(1), 81–109, doi: [10.1177/1363461515612933](https://doi.org/10.1177/1363461515612933)
- Kersten, P., Czuba, K., McPherson, K., Dudley, M., Elder, H., Tauroa, R., & Vandal, A.** (2016). A systematic review of evidence for the psychometric properties of the Strengths and Difficulties Questionnaire. *International Journal of Behavioral Development*, 40(1), 64–75.
- Kidd, E. & Donnelly, S.** (2020). Individual differences in first language acquisition. *Annual Review of Linguistics*, 6(1), 319–340, doi: [10.1146/annurev-linguistics-011619-030326](https://doi.org/10.1146/annurev-linguistics-011619-030326)
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C.** (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, 18(1), 146–154.
- Korntheuer, A., Pritchard, P. & Maehler, D.** (2017). Structural context of refugee integration in Canada and Germany, *GESIS-Schriftenreihe*, 15.
- Krause, U., & Gato, J.** (2019). Escaping humanitarian aid in camps? Rethinking the links between refugees' encampment, urban self-settlement, coping and peace. *Die Friedens-Warte*, 92(1-2), 76–97.
- Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W.** (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, 13(3), 373–381, doi: [10.1017/S136672891000009X](https://doi.org/10.1017/S136672891000009X)
- Kroll, J. F., & Stewart, E.** (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174, doi: [10.1006/jmla.1994.1008](https://doi.org/10.1006/jmla.1994.1008)
- Larsen-Freeman, D. & Long, M. H.** (1991). *An Introduction to Second Language Research*. London, UK: Longman.
- Lechner, S., & Siemund, P.** (2014). Double threshold in bi- and multilingual contexts: preconditions for higher academic attainment in English as an additional language. *Frontiers in Psychology*, 5, 546, doi: [10.3389/fpsyg.2014.00546](https://doi.org/10.3389/fpsyg.2014.00546)
- Liu, L. L., Benner, A. D., Lau, A. S., & Kim, S. Y.** (2009). Mother-adolescent language proficiency and adolescent academic and emotional adjustment among Chinese American families. *Journal of Youth and Adolescence*, 38, 572–586, doi: [10.1007/s10964-008-9358-8](https://doi.org/10.1007/s10964-008-9358-8)
- Lüdtke D** (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software*, 3*(26), 772. doi: [10.21105/joss.00772](https://doi.org/10.21105/joss.00772)
- Lustig, S. L.** (2010). An ecological framework for the refugee experience: What is the impact on child development? In Evans, G.W & Wachs, T.D. (Eds), *Chaos and Its Influence on Children's Development: An Ecological Perspective* (pp. 239–251). Washington, DC: American Psychological Association.
- Marinis T., & Armon-Lotem S.** (2015). Sentence repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing Multilingual Children: Disentangling Bilingualism From Language Impairment* (pp. 95–124). Bristol, UK: Multilingual Matters.
- Melby-Lervåg, M. & Lervåg, A.** (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: a meta-analysis of the correlational evidence. *Journal of Research in Reading*, 34, 114–135. doi: [10.1111/j.1467-9817.2010.01477.x](https://doi.org/10.1111/j.1467-9817.2010.01477.x)
- Méndez, L. I., Hammer, C. S., Lopez, L. M., & Blair, C.** (2019). Examining language and early numeracy skills in young Latino dual language learners. *Early Childhood Research Quarterly*, 46, 252–261.
- Paradis, J., Rusk, B., Sorenson Duncan, T., & Govindarajan, K.** (2017). Children's second language acquisition of English complex syntax: the role of age, input, and cognitive factors. *Annual Review of Applied Linguistics*, 37, 148–167, doi: [10.1017/S0267190517000022](https://doi.org/10.1017/S0267190517000022)

- Paradis, J., Soto-Corominas, A., Chen, X., & Gottardo, A. (2020). How language environment, age, and cognitive capacity support the bilingual development of Syrian refugee children recently arrived in Canada. *Applied Psycholinguistics*, *41*(6), 1255–1281. doi: [10.1017/S014271642000017X](https://doi.org/10.1017/S014271642000017X)
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, *1*(3), 213–237, doi: [10.1075/lab.1.3.01par](https://doi.org/10.1075/lab.1.3.01par)
- Paulson, J.F., Keefe, H.A., & Leiferman, J.A. (2009). Early parental depression and child language development. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *50*, 254–262. <https://doi.org/10.1111/j.1469-7610.2008.01973.x>
- Pfenninger, S.E. & Singleton, D. (2019). Starting age overshadowed: The primacy of differential environmental and family support effects on second language attainment in an instructional context. *Language Learning*, *69*, 207–234. doi: [10.1111/lang.12318](https://doi.org/10.1111/lang.12318)
- Place, S., & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition*, *19*(5), 1023–1041. doi: [10.1017/S1366728915000322](https://doi.org/10.1017/S1366728915000322)
- Polišenská, K., Chiat, S. & Roy, P. (2015). Sentence repetition: What does the task measure?. *International Journal of Language & Communication Disorders*, *50*, 106–118. doi: [10.1111/1460-6984.12126](https://doi.org/10.1111/1460-6984.12126)
- Prevo, M., Malda, M., Mesman, J., Emmen, R., Yeniad, N., van Ijzendoorn, M., & Linting, M. (2013). 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: [10.1017/S0305000913000299](https://doi.org/10.1017/S0305000913000299)
- Prevo, M., Malda, Emmen, R., Yeniad, N., & Mesman, J. (2015). A context-dependent view on the linguistic Interdependence Hypothesis: Language use and SES as potential moderators. *Language Learning*, *65*(2), 449–469, doi: [10.1111/lang.12099](https://doi.org/10.1111/lang.12099)
- Proctor, C.P., Harring, J.R. & Silverman, R.D. (2017). Linguistic interdependence between Spanish language and English language and reading: A longitudinal exploration from second through fifth grade. *Bilingual Research Journal*, *40*(4), 372–391, doi: [10.1080/15235882.2017.1383949](https://doi.org/10.1080/15235882.2017.1383949)
- Rauch, D. P., Naumann, J., & Jude, N. (2012). Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish–German bilinguals. *International Journal of Bilingualism*, *16*(4), 402–418, doi: [10.1177/1367006911425819](https://doi.org/10.1177/1367006911425819)
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Riches, C., & Genesee, F. (2006). Literacy: crosslinguistic and crossmodal Issues. In F. Genesee, K. Lindholm-Leary, W. M. Saunders, & D. Christian (Eds.), *Educating English Language Learners: A Synthesis of Research Evidence* (pp. 64–108). Cambridge, UK: Cambridge University Press. doi: [10.1017/CBO9780511499913.004](https://doi.org/10.1017/CBO9780511499913.004)
- Roesch, A., & Chondrogianni, V. (2016). “Which mouse kissed the frog?” Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of wh-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, *43*(3), 635–661, doi: [10.1017/S0305000916000015](https://doi.org/10.1017/S0305000916000015)
- Rolstad, K., & MacSwan, J. (2014). The facilitation effect and language thresholds. *Frontiers in Psychology*, *5*, 1197, doi: [10.3389/fpsyg.2014.01197](https://doi.org/10.3389/fpsyg.2014.01197)
- Sirin, S. R., & Rogers-Sirin, L. (2015). *The educational and mental health needs of Syrian refugee children* (p. 13). Washington, DC: Migration Policy Institute.
- Sohr-Preston, S.L. & Scaramella, L. V. (2006). Implications of timing of maternal depressive symptoms for early cognitive and language development. *Clinical Child and Family Psychology Review*, *9*, 65–83. <https://doi.org/10.1007/s10567-006-0004-2>
- Sorenson Duncan, T., & Paradis, J. (2020). How does maternal education influence the linguistic environment supporting bilingual language development in child second language learners of English? *International Journal of Bilingualism*, *24*(1), 46–61, doi: [10.1177/1367006918768366](https://doi.org/10.1177/1367006918768366)
- Soto-Corominas, A., Paradis, J., Al Janaideh, R., Vitoroulis, I., Chen, X., Georgiades, K., Jenkins, J. & Gottardo, A. (2020). Socioemotional wellbeing influences bilingual and biliteracy development: evidence from Syrian refugee children. In M. M. Brown & A. Kohu (Eds.), *BUCLD 44 Proceedings*, 620–633. Somerville, MA: Cascadilla Press.

- Tannenbaum, M., & Berkovich, M.** (2005). Family relations and language maintenance: Implications for language educational policies. *Language Policy*, *4*, 287–309, doi: [10.1007/s10993-005-7557-7](https://doi.org/10.1007/s10993-005-7557-7)
- Thabet, A. A., Stretch, D., & Vostanis, P.** (2000). Child mental health problems in Arab children: application of the strengths and difficulties questionnaire. *International Journal of Social Psychiatry*, *46*(4), 266–280.
- Torchiano, M.** (2020). effsize: Efficient Effect Size Computation, doi: [10.5281/zenodo.1480624](https://doi.org/10.5281/zenodo.1480624)
- Tseng, V., & Fuligni, A. J.** (2000). Parent-adolescent language use and relationships among immigrant families with east Asian, Filipino and Latin American backgrounds. *Journal of Marriage and the Family*, *62*(2), 465–476, doi: [10.1111/j.1741-3737.2000.00465.x](https://doi.org/10.1111/j.1741-3737.2000.00465.x)
- United Nations High Commissioner for Refugees (UNHCR).** (2019). Global Trends: Forced Displacement in 2018. Retrieved May 29th, 2020, from <https://www.unhcr.org/globaltrends2018/>
- Unsworth, S., & Blom, E.** (2010). Comparing L1 children, L2 children and L2 adults. In E. Blom & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 201–222). Amsterdam/Philadelphia: John Benjamins.
- Van Der Wildt, A., Van Avermaet, P. & Van Houtte, M.** (2017). Opening up towards children's languages: Enhancing teachers' tolerant practices towards multilingualism. *School Effectiveness and School Improvement*, *28*(1), 136–152, doi: [10.1080/09243453.2016.1252406](https://doi.org/10.1080/09243453.2016.1252406)
- Verhoeven, L.T.** (1994). Transfer in bilingual development: The Linguistic Interdependence Hypothesis revisited. *Language Learning*, *44*, 381–415. doi: [10.1111/j.1467-1770.1994.tb01112.x](https://doi.org/10.1111/j.1467-1770.1994.tb01112.x)
- Winter, B.** (2020). *Statistics for Linguists: An Introduction Using R*. London/New York: Routledge.
- Yasik, A., Saigh, P.A., Oberfield, R.A., & Halamandaris, P.V.** (2007). Posttraumatic stress disorder: Memory and learning performance in children and adolescents. *Biological Psychiatry*, *61*(3), 382–388, doi: [10.1016/j.biopsych.2006.06.005](https://doi.org/10.1016/j.biopsych.2006.06.005)

Appendix 1. Stimuli used in Litmus-SRTs

Table A1. Stimuli used in English SRT

Structure	Items
Declarative ($k=6$)	She can bring the glass to the table.
	They have been riding the bicycle around the backyard.
	They are eating the bananas in the park.
	The kitten could have bounced the ball down the stairs.
	The boy must sweep the floor in the kitchen.
	The teacher has been looking at us all day.
Short passive ($k=3$)	She was stopped at the big red lights.
	The children were taken to the office.
	He was pushed hard against the ground.
Long passive ($k=3$)	The cow was kicked in the leg by the donkey.
	She was seen by the doctor in the morning.
	The mother was followed by the girl.

(Continued)

Table A1. (Continued)

Structure	Items
Question ($k=6$)	What did the mother cook in the evening?
	Who have they seen near the front door?
	Which picture did he paint at home yesterday?
	What did the father buy last month?
	Who did the girl meet in the library yesterday?
	Which drink did the neighbour spill in the house?
Coordinated ($k=3$)	The mother is shopping and the child is studying at home.
	The dog barks outside and the child cries inside.
	Our neighbor cleans the car and his son plays basketball.
Subordinate ($k=4$)	If the weather is warm, we can go to the park.
	Before the girl eats dinner, she will play with the computer.
	The children will get a present if they clean the house.
	The child ate breakfast after he washed his face.
Relative ($k=6$)	The boy that the neighbour helped has lost his way.
	They should wash the baby that the mother is feeding.
	The horse that the farmer pushed kicked him in the back.
	The mother made the meal that the children are eating.
	The children enjoyed the candy that they tasted.
	The team that my brother cheered for won the race.

Table A2. Stimuli used in Syrian Arabic SRT

Structure	Items
Declarative (<i>k=6</i>)	<i>ʕæm yæklu lmo:z bəlʕne:ne</i> PROG 3-eat-PL the-bananas in-the-park 'They are eating the bananas in the park.'
	<i>rʕi:ʔi byəʔder yərfæʕ hækersi lkbi:r</i> friend-my 3SG.M-can 3SG.M-lift this-the-chair the-big 'My friend can lift this big chair.'
	<i>læzem tʔətʕ lkæ:se ʕætʕʕawle</i> must 3SG.F-put the-cup on-the-table 'She must put the glass on the table.'
	<i>mæ bteħsen lʔænse tji:b lktæ:b ʕæsʕsʕaf</i> NEG 3SG.F-can the-teacher 3SG.F-bring the-book to-the-class 'The teacher cannot bring the book to the class.'
	<i>mæ jereb lwælæd hæli:b səxan əssəbəħ</i> NEG drank the-child milk hot the-morning 'The child did not drink hot milk in the morning.'
	<i>mæ ʔære:t ʔesʕa helwe bəlmædræse</i> NEG read-1SG story beautiful in-the-school 'I did not read a beautiful story at school.'
Short passive (<i>k=3</i>)	<i>ətæxædu lwlæ:d ʕælmæktæb</i> PASS-taken-3PL the-children to-the-office 'The children were taken to the office.'
	<i>ndæfæf bʔuwe ʕælʔardʕ</i> PASS-pushed in-hard to-the-ground 'He was pushed hard against the ground.'
	<i>txæ:lafet ʕælijæ:ra lhæmra</i> PASS-fined-3SG.F on-the-light the-red 'She got a fine at the red light.'
Topicalization (<i>k=3</i>)	<i>lʔm leħʔa esʕsʕabi ʕæjʕa:reʕ</i> the-mother followed-her the-boy to-the-street 'As for the mother, the boy followed her to the street.'
	<i>lbaʔara lhma:r dʕarabæ barra</i> the-cow the-donkey hit-her outside 'As for the cow, the donkey hit her outside.'
	<i>lʔæb eddoktɔ:r fæħasʕ əssəbəħ</i> the-father the-doctor examined-him the-morning 'As for the father, the doctor examined him in the morning.'
Question (<i>k=6</i>)	<i>mi:n əlli ʕæfəɔ lbənt</i> who that saw-3SG.F-him the-girl <i>bəlmæktæbe mbæ:reħ</i> in-the-library yesterday? 'Who did the girl see [him] in the library yesterday?'

(Continued)

Table A2. (Continued)

Structure	Items
	<p><i>mi:n fæ:fu jænɒ lbæ:b lʔæmæ:mi:</i> who have.seen-3PL near the-door the-front 'Who have they seen near the front door?'</p> <p><i>ju: ftæra lʔæb eʃfaher lmædʰi</i> what bought the-father the-month the-last 'What did the father buy last month?'</p> <p><i>ju: tʰabxet lʔm lmæsæ</i> what cooked-3SG.F the-mother the-evening 'What did the mother cook in the evening?'</p> <p><i>ʔyyæ fara:b kæb lja:r bəlbe:t</i> which drink spilled the-neighbor in-the-house 'Which drink did the neighbor spill in the house?'</p> <p><i>ʔyyæ sʰu:ra rasæmæ bəlbe:t mbæreh</i> which picture drew-her in-the-house yesterday 'Which picture did he draw [it] at home yesterday?'</p>
Coordinated (k=3)	<p><i>lkælb sʔæm yʃawwi barra w</i> the-do PROG 3SG.M-bark-SG outside and <i>lwæləd sʔæm yibki juwwa</i> the-child PROG 3SG.M-cry-SG inside 'The dog is barking outside and the child is crying inside.'</p> <p><i>lʔm sʔæm tʔtsæwwa? w</i> the-mother PROG 3SG.F-shop and <i>lwæləd sʔæm yidros bəlbe:t</i> the-boy PROG 3SG.M-study in-the-home 'The mother is shopping and the boy is studying at home.'</p> <p><i>ja:rnæ sʔæm yinædʰef essæyyæra w</i> neighbor-our PROG 3SG.M-clean the-car and <i>əbn-ɔ sʔæm yə-lʃab bətʰtʰ a:be</i> son-his PROG 3SG.M-play with-the-ball 'Our neighbor is cleaning the car and his son is playing with the ball.'</p>
Subordinate (k=4)	<p><i>fetʰer lwæləd bæʃdmæ: yassæl wəʃɔ</i> had.breakfast the-child after washed face-his 'The child had breakfast after he washed his face.'</p> <p><i>ʔabəlmæ: tʔtʃəʃæ lbənt rəh</i> before 3SG .F-has.dinner the-girl FUT <i>təlʃab sʔælkɔmbiuter</i> 3SG.F-play on-the-computer 'Before the girl has dinner, she will play with the computer.'</p> <p><i>rəh ya:xdu lwlə:d hdi:ə</i> FUT 3-get-PL the-children gift <i>izæ naɔdʰdʰafu lbe:t</i> if clean-3PL the-house 'The children will get a present if they clean the house.'</p>

(Continued)

Table A2. (Continued)

Structure	Items
	<p><i>izæ kæ:n ejɜau dæ:fi</i> if was the-weather warm <i>fi:næ nət^llaɟ ʒæljnene</i> can-1PL 1PL-go.out to-the-park ‘If the weather is warm, we can go to the park.’</p>
Relative (<i>k=6</i>)	<p><i>ənbasat^lu lwlæ:d bəljɔkɔlata</i> enjoyed-3PL the-children with-the-chocolate <i>əlli ʔækæluwa</i> that ate-3PL-her ‘The children enjoyed the chocolate that they ate [it].’</p> <p><i>lʔm sæwet et^labxa yælli</i> the-mother made-3SG.F the-meal that <i>ʒæm yæ:kluwa lwlæ:d</i> PROG 3-eat-PL-her the-children ‘The mother made the meal that the children are eating [it].’</p> <p><i>læ:zem yħæməmu ez^lɣi:r əlli</i> should 3-wash-PL the-baby that <i>ʔmmɔ ʒæm ətt^laʒmi:</i> mother-his PROG 3SG.F-feed-him ‘They should wash the baby that his mother is feeding [him].’</p> <p><i>lħs^la:n əlli lfællæħ d^larab</i> the-horse that the-farmer hit <i>rafæso ʒæd^lahrɔ</i> kicked-him on-back-his ‘The horse that the farmer hit kicked him on his back.’</p> <p><i>es^ls^labi əlli sæ:ʒædɔ lja:r</i> the-boy that help-him the-neighbor <i>d^layyæʒ t^lari: ʔɔ</i> lost way-his ‘The boy that the neighbor helped [him] has lost his way.’</p> <p><i>fæ:z lfari:ʔ əlli ʒæjæʒɔ</i> won the-team that encouraged-him <i>æxi bæssiba:ʔ</i> brother-my in-the-race ‘The team that my brother cheered for [it] won the race.’</p>

Table A3. Stimuli used in Dutch SRT

Structure	Items
Declarative ($k=9$)	<i>Zij kan het meisje zien door het raam</i> she can the girl see through the window 'She can see the girl through the window.'
	<i>Het meisje moet het vliegtuig op het bord tekenen</i> the girl must the plane on the board draw 'The girl has to draw the plane on the board.'
	<i>De jongen moet de vloer in de keuken vegen</i> the boy must the floor in the kitchen wipe 'The boy must wipe the floor in the kitchen.'
	<i>De man had niet met de auto naar huis mogen rijden</i> the man had not with the car to house may drive 'The man should not have driven home by car.'
	<i>De agent had ons een bekeuring kunnen geven</i> the police officer had us a ticket can give 'The officer could have given us a ticket'
	<i>De muis ziet de kaas in de schaal niet</i> the mouse sees the cheese in the bowl not 'The mouse does not see the cheese in the bowl'
	<i>Hij had naar de boot moeten zwemmen</i> he had to the boat must swim 'He had to go to the boat and swim'
	<i>Zij heeft de schapen op de boerderij niet kunnen voeren</i> she has the sheep on the farm not can feed 'She was unable to feed the sheep on the farm'
	<i>De kat had de bal van de trap af kunnen gooien</i> the cat had the ball off the stairs off can throw 'The cat could have thrown the ball down the stairs'

(Continued)

Table A3. (Continued)

Structure	Items
Question ($k=6$)	<p><i>Wie heeft de aap natgespetterd bij de zee?</i> who has the monkey wet splattered at the sea 'Who splattered the monkey at the sea?'</p>
	<p><i>Wat heeft de prinses vorige maand gekocht?</i> what has the princess last month buy 'what did the princes buy last month?'</p>
	<p><i>Wie hebben zij bij de trap gezien?</i> who have they at the stairs see 'Who did they see near the stairs?'</p>
	<p><i>Wat hebben zij gisteren in de sneeuw gevonden?</i> what have they yesterday in the snow find 'What did they find in the snow yesterday?'</p>
	<p><i>Welke foto heeft hij gisteren gemaakt?</i> which photo has he yesterday made 'What photo did he take yesterday?'</p>
	<p><i>Welke fles liet de jongen op de grond vallen?</i> which bottle let the boy on the floor fall 'What bottle did the boy drop on the floor?'</p>
Short passive ($k=3$)	<p><i>Bij de grote rode lampen werd ze gestopt</i> at the big red lamp was she stopped 'At the big red lamp, she was stopped.'</p>
	<p><i>Ze werd hard tegen de grond geduwd</i> she was hard against the floor pushed 'She was pushed hard against the floor.'</p>
	<p><i>De jongen in de snoepwinkel werd geholpen</i> the boy in the candy shop was helped 'The boy in the candy shop was helped.'</p>
Long passive ($k=3$)	<p><i>Ze werd 's ochtends door de dokter bezocht</i> she was in-the-morning by the doctor visit 'She was visited by the doctor in the morning.'</p>
	<p><i>De jongen werd bang gemaakt door de clown</i> the boy was scared made by the clown 'The boy was scared by the clown.'</p>
	<p><i>De lunch werd door de timmerman gegeten</i> the lunch was by the carpenter eaten 'Lunch was eaten by the carpenter.'</p>

(Continued)

Table A3. (Continued)

Structure	Items
Subordinate ($k=3$)	<p><i>De mensen krijgen een cadeau als ze het</i> the people get a present if they the <i>huis schoonmaken</i> house clean-make ‘People get a present if they clean the house.’</p>
	<p><i>Als de kinderen zich gedragen gaan</i> if the children themselves behave go <i>we de tuin in</i> we the garden in ‘If the children behave, we will go into the garden.’</p>
	<p><i>Hij had zijn vriendin niet meegenomen als</i> he had his girlfriend not with-taken if <i>ze boos was</i> she angry was ‘He wouldn’t have brought his girlfriend if she was angry.’</p>
Relative ($k=6$)	<p><i>De kinderen proefden de snoepjes die</i> the children tasted the candies that <i>ze kregen</i> they got ‘The children tasted the candies they got.’</p>
	<p><i>De bij die de man inslikte had hem</i> the bee that the man swallowed had him <i>pijn gedaan</i> pain done ‘The bee that the man swallowed had hurt him.’</p>
	<p><i>De aap aaide de koe die de worm</i> the monkey stroke the cow that the worm <i>bang had gemaakt</i> scared had made ‘The monkey stroked the cow that had scared the worm.’</p>
	<p><i>De jongen die de postbode hielp was</i> the boy that the postman helped was <i>verdwaald</i> lost ‘The boy helping the postman was lost.’</p>
	<p><i>Hij moet de knuffel wassen waarmee het kind</i> he must the stuffy wash where-with the child <i>slaapt</i> sleeps ‘He should wash the stuffed animal that the child sleeps with.’</p>
	<p><i>De pony die de boer bekeek duwde hem opzij</i> the pony that the farmer looked push him aside ‘The pony that the farmer looked at pushed him aside.’</p>

Appendix 2. Structural accuracy scoring

For the structural accuracy scoring, a participant's response was scored as 1 if the child repeated the target syntactic structure, regardless of other changes to the sentence. Otherwise, the response was coded as 0. The criteria for a sentence to be considered structurally accurate depended on the target syntactic structure (Marinis & Armon-Lotem, 2015). For example, for a sentence with a subordinate clause to be considered preserved, participants' responses had to contain two clauses, one main and one subordinated, and a subordinator.

The structural accuracy scoring was carried out simultaneously with the verbatim scoring. Therefore, the same team of scorers as described in the main text completed both types of scoring. For the Canadian sample, full reliability of structural accuracy scoring was ensured by having every sentence scored independently twice by a team of trained scorers and all disagreements were settled by group discussion. In the Netherlands, twenty-five percent of the Litmus SRT was independently scored a second time. The α coefficients of inter-rater reliability for structural accuracy for this site were .93 for Dutch and .88 for Arabic.

As mentioned in the text, verbatim and structural accuracy scores were strongly correlated, as shown in Table A4 (a more complete version of Table 3).

Structural accuracy results for the L1 and L2 for both samples appear in Table A5. As is the case for verbatim scores, significant differences were found when comparing L1 and L2 scores with two-tailed paired-samples *t*-tests. For the Canadian sample: $t(55) = 6.1802$, $p < .001$ (Cohen's $d = 0.94$; large effect size). For the Dutch sample: $t(46) = 3.6391$, $p < .001$ (Cohen's $d = 0.47$; small effect size).

Table A4. Pearson correlations between L1 and L2 based on Litmus-SRT scores and length of L2 schooling for the Canadian (upper right corner) and Dutch samples (lower left corner)

	L1 Verbatim	L2 Verbatim	L1 Structural accuracy	L2 Structural accuracy	L2 schooling
L1 Verbatim		.25.	.92***	.30*	-.10
L2 Verbatim	.61***		.27*	.89***	.30*
L1 Structural accuracy	.93***	.53***		.35**	-.11
L2 Structural accuracy	.64***	.87***	.60***		.30*
L2 schooling	.16	.64***	.07	.54***	

^a . = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Table A5. Descriptive information about sentence repetition outcomes in the Canadian (L1 Arabic, L2 English) and Dutch (L1 Arabic, L2 Dutch) samples, % correct (SD), number correct (SD), and minimum-maximum score.^a

	L1 Structural accuracy	L2 Structural accuracy
Canada ($n = 56$)	83.93% (12.58) 26.02 (3.90) 13-31	64.11% (25.24) 19.88 (7.83) 3-31
Netherlands ($n = 47$)	81.95% (15.07) 25.40 (4.67) 13-31	70.71% (26.54) 21.21 (7.96) 5-30

^aThe maximum score that could be obtained is 31, 31, and 30 for the Arabic, English and Dutch Litmus-SRT, respectively.

Appendix 3

Table A6. Correlation matrix of variables in the Canadian sample (upper right corner) and Dutch sample (lower left corner)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Age		.944***	.087	.342**	.516***	.001	.053	-.138	.047	.157	.145	.237.	.173	-.175	.135	.131	-.034
2. Age of arrival	.903***		-.247.	.071	.607***	-.008	-.044	-.269*	-.040	.012	.180	.187	.157	-.079	.081	.103	.061
3. Length of residency	.113	-.325*		.789***	-.318*	.027	.289*	.405**	.259.	.424**	-.118	.130	.033	-.275*	.152	.072	-.282*
4. L2 schooling	.252.	-.163	.931***		-.061	.168	.293*	.200	.217	.452***	.047	.245	-.091	-.316*	.313*	.224.	-.184
5. L1 schooling	.685***	.622***	.071	.172		-.111	-.185	-.009	.199	.121	.291*	.102	.318	.263.	.064	.155	.391**
6. Time in refugee camp	NA	NA	NA	NA	NA		.189	-.077	-.137	-.143	.108	-.076	-.352**	-.212	-.097	-.260	-.337*
7. Number of children in family	-.006	-.172	.385**	.352*	.209			-.061	.042	-.012	-.194	.026	-.466***	-.465***	-.352**	-.401**	-.465***
8. Language use parent interaction	.053	.126	-.174	-.205	-.006	NA	-.035		.547***	.456***	-.031	.031	.331*	.140	.044	.105	-.070
9. Language use sibling interaction	-.075	-.073	.002	.019	-.123	NA	-.251.	.336*		.592***	-.228.	-.174	.181	-.098	.096	.152	.084
10. L2 richness	.118	-.089	.465**	.441**	.022	NA	.193	-.084	.217		-.019	.020	.316*	-.023	.207	.178	.088

(Continued)

Table A6. (Continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
11. L1 richness	-.265.	-.166	-.199	-.153	-.208	NA	.203	-.321*	-.401**	-.144		.150	.278*	.401**	.200	.268*	.288*
12. Maternal interaction in L2	.080	.028	.112	.101	-.026	NA	-.224	.177	.397**	.562***	-.375**		.310*	.062	.414**	.197	-.056
13. Maternal fluency in L2	.088	-.175	.599***	.609***	.103	NA	.193	-.052	.191	.582***	-.340*	.546***		.487***	.149	.243.	.300*
14. Maternal education in years	.012	-.066	.178	.170	.232	NA	.125	-.505***	-.429**	.082	.179	.020	.171		.168	.379**	.500***
15. Paternal interaction in L2	.179	.106	.152	.162	.250.	NA	-.041	.067	-.025	-.351*	-.439**	-.077	.134	.127			.353**
16. Paternal fluency in L2	.093	.006	.194	.204	.237	NA	.036	-.132	-.228	-.204	-.265.	-.241	-.009	.262.	.757***		.605***
17. Paternal education in years	-.025	-.116	.212	.230	.140	NA	.180	-.647***	-.444**	.117	.220	-.111	.102	.749***	.172	.510***	

*Correlation not attempted for the Dutch sample since only one participant reported time spent in a refugee camp.