Journal of Child Language (2022), **49**, 741–773
doi:10.1017/S0305000921000246

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Sources of variation at the onset of bilingualism: The differential effect of input factors, AOA, and cognitive skills on HL Arabic and L2 English syntax

Adriana SOTO-COROMINAS[1], Evangelia DASKALAKI[2], Johanne PARADIS[2], Magdalena WINTERS-DIFANI[2], and Redab AL JANAIDEH[3]

[1]Universitat Internacional de Catalunya, Barcelona, Spain, [2]University of Alberta, Canada, and [3]Ontario Institute for Studies in Education, University of Toronto, Canada
Address for correspondence: Adriana Soto-Corominas, Department of Applied Linguistics, Universitat Internacional de Catalunya, Barcelona, 08017, Spain. E-mail: asotoc@uic.cat

(Received 14 October 2019; revised 14 May 2020; accepted 1 March 2021;
first published online 26 May 2021)

Abstract

Despite growing research on individual differences in child bilinguals, few studies have focused on the development of syntax, included both languages, and studied newly arrived school-age migrant children. Accordingly, this study investigated the syntactic development of heritage language (HL) Syrian Arabic and L2 English by Syrian refugee children ($N = 119$) recently arrived in Canada using a sentence repetition task. Regression analyses showed that a partially overlapping set of child-level (input and cognitive skills) and language-level (syntactic structure) factors accounted for performance in each language. HL performance was particularly sensitive to language, cognitive, and input variables indexing cumulative HL exposure. L2 performance, however, was sensitive to cognitive and environmental variables indexing current and cumulative L2 use. Finally, despite stronger performance in Arabic than in English, results revealed interdependence between the two languages, indicating that participants with stronger syntactic abilities in their HL tended to have stronger syntactic abilities in their emerging L2.

Introduction

A growing body of research has investigated individual differences and the factors that promote or hinder bilingual children's development in their heritage language (HL) and second language (L2). The vast majority of this research, however, has focused on the acquisition of vocabulary and morphosyntax (see review in Chondrogianni, 2018 and Unsworth, 2016a). On the other hand, the research on individual differences in the acquisition of syntax is comparatively limited, as this linguistic domain has traditionally been considered relatively robust to variable individual linguistic experience (e.g., Hopp, Steinlen, Schelletter & Piske, 2019; Unsworth, 2016b).

Another limitation of the extant body of research on early bilingual development is that it has mostly included children from immigrant or second-generation families who have spent most or all of their childhood in the host country (Paradis, 2019). The present study investigates the syntactic development of bilingual Arabic–English Syrian refugee children who emigrated to Canada between the ages of 4–10, as a cohort, between 2015 and 2018. This population presents relevant similarities and differences when contrasted with the samples frequently reported in child bilingual language acquisition research. In terms of similarities, refugee children, as many other migrant children, acquire their HL at home and their L2 through exposure in the L2 majority community and at school. In contrast, the first-generation refugee children in this study have had considerable exposure to the HL (Syrian Arabic) in a quantitatively and qualitatively rich Arabic-speaking environment prior to immigration. As such, this study describes the early stages of bilingual acquisition of HL Syrian Arabic and L2 English by Syrian refugee children, before any dominant language shift has taken effect (Montrul, 2016; Murphy, 2018). In doing so, this is one of the few studies to investigate the two languages of bilingual school-age children at the onset of their bilingualism, allowing for a direct comparison of the factors that influence the development in the two systems.

Refugee children present differences from other migrant children, at least in the Canadian context, regarding factors outside of language structure that may influence their acquisition patterns: children may have interrupted schooling prior to migration, refugee families may be under-resourced, and parents may be less educated and underemployed, which means that fathers, as well as mothers, are likely to spend time in the home (Hadfield, Ostrowski & Ungar, 2017; Kanu, 2008). While research on individual differences in L2 acquisition in children is growing, no systematic research exists examining how language environment factors shape L2 and HL acquisition at the early stages post-migration in this population of bilingual children.

In light of the gaps in research on bilingual syntactic development and first-generation refugee children, the objectives of this study were threefold: (1) to compare the syntactic abilities in the HL Syrian Arabic and L2 English of refugee children in the early stages of bilingual language acquisition, (2) to investigate whether the more established HL syntactic abilities may support the developing L2 abilities, and (3) to investigate what child-level factors (language environment or age/cognition-related) and language-level factors (structure-related) predict individual variation in their syntactic abilities in the two languages.

In order to provide a background for this study, we first begin by describing extant research on the child-level and language-level factors that may influence syntactic development in both the HL and the L2. We then describe the theory of Interdependence (Cummins, 1979, 2000), which provides a framework for investigating the relationship between the HL and L2 abilities in sequential bilingual children. Before delving into the present study, we provide a description of the relevant morphosyntactic properties of Arabic and English.

Factors influencing HL and L2 development of syntax

Research on bilingual development has demonstrated that both child- and language-level factors may influence the development of vocabulary and morphology in bilingual children (for an overview, see Armon-Lotem & Meir, 2019; Unsworth, 2016a; for the HL, see Unsworth, 2017; for L2, see Chondrogianni, 2018;

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Paradis, 2019). Comparatively fewer studies have explored whether these factors are also relevant for the SYNTACTIC abilities of bilingual children. In what follows, we provide an overview of the literature focusing on the bilingual (HL/L2) syntactic development of children in immigrant contexts.

Child-level factors: language environment

Bilingual children's language environment may vary considerably in quantitative and qualitative terms. The main child-level factors that have been investigated in terms of their relation to syntactic development are home language use, length of exposure, language-environment richness, and socioeconomic status.

In terms of the relative HL/L2 use at home, a positive relationship has been found between more HL use at home and stronger syntactic abilities in the HL (Albirini, 2014; Daskalaki, Chondrogianni, Blom, Argryri & Paradis, 2019; R. Jia & Paradis, 2020). For example, Albirini's (2014) retrospective study of heritage speakers of Arabic in the United States showed that stronger Arabic skills (in terms of fluency, grammaticality, and complex syntax) were predicted by the amount of Arabic use over time. The effect of L2 language use on L2 syntactic development is far less clear, with some studies suggesting a negative association (Sorenson Duncan & Paradis, 2020) or no association at all (Paradis, Rusk, Sorenson Duncan & Govindarajan, 2017). For instance, Kaltsa, Prentza and Tsimpli (2020), using a Greek sentence repetition task (SRT), found that even though the use of Greek at home during preschool years correlated with better overall accuracy in the SRT productions by sequential bilinguals (HL Albanian; mean age of 8.6), current amount of Greek was not correlated with their L2 Greek accuracy. The differential effect of HL vs. L2 use at home could be due to diverse reasons. First, the HL may be more sensitive than the L2 to fluctuations in the home environment. Due to its status as a minority language, the presence of the HL in the community tends to be limited; therefore, children's acquisition of the HL would rely more heavily on its use in the home compared to the L2, which is the majority language of the broader community and the school system. It is also possible that the effect of L2 use in the home is modulated by the L2 proficiency of parents, who are often L2 learners themselves (Sorenson Duncan & Paradis, 2020; Paradis, 2019). Depending on their degree of L2 proficiency, they may be more or less likely to produce diverse syntactic structures.

Variation in the length of exposure (LOE) to the two languages may also influence the syntactic development of bilingual children. The effect of LOE on HL syntactic development has been explored mostly indirectly through its association with age-related variables that we discuss later on. By contrast, the effect of L2 LOE on L2 syntactic development has played a prominent role in the L2 literature. For example, a positive effect of longer LOE to the L2 has been reported for a range of syntactic phenomena and languages (e.g., Chondrogianni & Marinis, 2011; Paradis et al., 2017; Roesch & Chondrogianni, 2016; Sorenson Duncan & Paradis, 2020). Of particular relevance to the present study are the results from the SRT literature, which are conflicting. Armon-Lotem, Walters and Gagarina (2011) and Chiat, Armon-Lotem, Marinis, Polišenská, Roy, Seeff-Gabriel and Gathercole (2013), who investigated different groups of bilingual children, found a positive effect of longer LOE for some groups and a null effect for others within the same study. Similarly, Meir, Walters and Armon-Lotem (2016) found no significant correlations between LOE to Hebrew and bilingual children's (L1 Russian-L2 Hebrew) performance on a Hebrew SRT task.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

In addition to quantity (i.e., relative HL/L2 use at home and overall LOE), bilingual children's language experience may vary in terms of its HL/L2 richness. The richness of the linguistic environment refers to the frequency with which children engage in HL/L2 language-rich activities, such as reading books, socializing with peers, and participating in community events or activities (Paradis, 2019). Very few studies have explored the association between HL richness and HL syntax, and those that have done so have focused on the role of HL formal instruction and/or literacy. For example, Bayram, Rothman, Iverson, Kupisch, Miller, Puig-Mayenco and Westergaard (2017) found that being literate in Turkish predicted stronger abilities with Turkish syntactic structures (passives) in Turkish-German bilingual youth in Germany. On the other hand, Flores and Barbosa (2014) found no association between the hours of formal instruction in Portuguese and clitic placement among heritage child learners of Portuguese in Germany. Research on the effect of L2 richness on L2 syntactic abilities is also scarce, but results are consistently positive. Paradis et al.'s (2017) study on the production of complex syntax by English child L2 learners (cL2ers) showed that richness of the L2 environment was a positive predictor of greater use of L2 complex syntax in spontaneous conversation and elicited narration. Similarly, Kaltsa et al. (2020) found that children who engaged in more literacy activities in L2 Greek had higher accuracy on a Greek SRT.

Finally, the few studies that have examined the effect of socioeconomic status (SES) on syntax report a positive association between parental education/occupation and children's syntactic abilities in their two languages. For instance, Armon-Lotem et al.'s (2011) study on Russian-German bilingual children found that maternal education was predictive of children's performance in the HL SRT, whereas parental occupation was predictive of children's performance in the L2 SRT. To explain this correlation, the authors invoked research suggesting that higher SES might index parents' proficiency in the L2 and, more generally, L2 input quality/richness. In line with this conclusion, Sorenson Duncan and Paradis (2020) found that maternal education levels had an indirect effect on children's L2 syntactic abilities: maternal education predicted L2 proficiency and this, in turn, was a proximal predictor of children's syntactic abilities.

Child-level factors: age of onset and cognitive capacities

In addition to the child-external/environmental factors discussed so far, child-internal factors such as age of onset of acquisition (AOA) of the L2 and cognitive skills have also been found to influence bilingual development.

AOA bears implications for the development of the two languages of the bilingual. On the one hand, children with an older AOA of the L2 have had a longer period of being monolingual in their HL and, consequently, are more likely to develop and retain strong abilities in their HL (see Montrul, 2008, 2016). In this sense, AOA is not only an index of cognitive abilities, but also an index of cumulative exposure to the HL. Specifically for the children in this study, a later AOA indicates a longer period of time living in a quantitatively and qualitatively rich Arabic-speaking environment prior to immigration. For the L2, a later AOA indexes the cognitive and linguistic maturity at the onset of L2 acquisition, and it may also index the setting where the onset of acquisition took place (e.g., at home, in preschool, school). For the HL, Albirini's (2018) study of Arabic–English bilingual children in the US found that an older AOA correlated positively with children's performance in HL Arabic in three tasks targeting various syntactic and morphosyntactic phenomena. Similarly, Meir,

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Walters and Armon-Lotem (2017) found that Russian-Hebrew bilingual children with an older AOA (after age 2) showed higher accuracy with respect to Subject-Verb agreement in their HL Russian. In the case of cL2 syntax, on the other hand, results have generally either shown a negative association between older AOA and L2 syntactic abilities (Meir et al., 2017; Roesch & Chondrogianni, 2016) or no association at all (Kaltsa et al., 2020). Chiat et al. (2013), in particular, reported significant negative effects of older AOA in their SRT studies for some groups of bilingual children but no AOA effect for others. Importantly, all of these studies relied on group comparisons between simultaneous and sequential bilingual children to determine AOA effects.

Finally, cognitive abilities, a factor that as mentioned above often correlates with AOA, have also been shown to predict bilingual children's syntactic abilities. Paradis et al. (2017) found that immigrant children's verbal working memory and analytical reasoning scores were predictors of the number of complex English sentences produced. Whether these abilities are also associated with children's HL syntactic abilities remains unexplored.

Language-level factors

Whereas child-level factors determine individual differences between children, language-level factors determine differential acquisition rates for specific structures (Paradis, 2019). Typically, structures that are syntactically complex and/or structures that stabilize late in monolingual acquisition are good candidates for protracted acquisition in bilingual contexts (Gathercole, 2007; Tsimpli, 2014).

Effects of type of structure have been reported in a number of SRT studies. Meir et al. (2016) found that L1 Russian-L2 Hebrew bilinguals were outperformed by their monolingual Russian-speaking peers on the Russian SRT primarily on structures involving case. On the Hebrew SRT, on the other hand, they performed similarly to their Hebrew monolingual peers in all structures, except object relatives, where they scored significantly lower. Similarly, Kaltsa et al. (2020), using a Greek SRT, found that both monolingual and bilingual groups of children scored significantly lower in syntactic structures involving clitics, whereas Chiat et al. (2013), using an English SRT, showed that monolingual and bilingual children performed better on short and simple sentences than on long and complex sentences.

Overall, the results of these studies suggest that syntactically complex sentences (long distance dependencies/subordination) that rely on morphological cues, such as case morphology and/or object clitic-NP agreement, might pose a burden on (bilingual) learners.

Interim summary: child-level and language-level factors

Taken together, the studies reviewed in this section show that the acquisition of syntax by bilingual children, like the acquisition of vocabulary and morphology, may be affected by different child-level and language-level factors, though the effect of these factors might differ for the HL and the L2. This observation has to be treated with caution as most of the existing studies focus on either the HL or the L2 and examine the effect of only a few factors in isolation.

Accordingly, the present study builds on this line of research by making three novel contributions: first, we test the effect of a larger set of child-level and language-level factors on children's syntactic development. This allows us to investigate the effect of a certain factor while controlling for the effect of other variables. Second, we investigate an

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

understudied population, Syrian refugee children in Canada, who might differ from the general body of bilingual children in important aspects, including their pre-migration experiences with their HL, as well as the quantity and richness of their current language environment. Finally, we test children's syntactic development in both their HL and L2. In this way, we explore whether the same cluster of factors are relevant for HL and L2 syntactic development. Furthermore, since both languages are tested, we have the opportunity to investigate whether the more stabilized HL grammar may support the developing L2 one. We turn our attention to this potential relationship next.

Interdependence

The relationship and influence between the two developing systems of bilinguals have been studied from different perspectives, one of them being The Developmental Interdependence Hypothesis (Cummins, 1979, 2000). This hypothesis posits that certain aspects of L1/HL skills can support the development of the L2, independently of typological considerations. There is a robust body of research that shows that interdependence applies to literacy skills and to metalinguistic skills that support reading, such as phonological awareness (e.g., Hammer, Lawrence & Miccio, 2007; Tabors, Paez & López, 2003; Verhoeven, 1994; see other references in Cummins, 1991). Whether a similar relationship exists for oral language skills is less clear. Conflicting findings in terms of interdependence of vocabulary skills across the two languages have emerged, with some studies reporting evidence of an interlinguistic relationship between the vocabulary abilities of the L1/HL and the L2 and some reporting no evidence for an association (Tabors et al., 2003; Uchikoshi, 2006; Verhoeven, 1994; see review in Méndez, Hammer, Lopez & Blair, 2019). In terms of syntactic and morphosyntactic skills, evidence is similarly mixed. Gottardo (2002), who used a cloze task to assess Spanish-English bilingual six-year-olds' syntactic and morphosyntactic abilities, found no support for interdependence between the two languages. On the other hand, Castilla, Restrepo and Perez-Leroux (2009) found strong evidence for interdependence between the L1 Spanish and L2 English of four-year-old bilinguals in terms of grammatical development. These authors used a comprehensive test battery together with a measure of mean length of utterance in the two languages. Finally, limited evidence of interdependence in terms of morphosyntactic skills, measured with an SRT, was found by Verhoeven's (1994) longitudinal study. This author found a significant contribution of L1 Turkish morphosyntactic skills to the L2 Dutch skills of six-year-old bilinguals only at the beginning of Grade 1, but not one year later.

In summary, the interdependence hypothesis has received strong support in terms of literacy and metalinguistic skills related to literacy. On the other hand, far less conclusive evidence exists for interdependence of oral language skills, including syntax. Accordingly, a further objective of the present study is to explore whether interdependence between the HL (Syrian Arabic) and the L2 (English) is observed in terms of syntactic development.

Before moving on to the present study, we briefly review the relevant syntactic properties of Syrian Arabic in relation to English.

Morphosyntactic properties of Syrian Arabic and English

Arabic belongs to the Semitic branch of the Afro-Asiatic family of languages and comprises a number of varieties. In addition to the Modern Standard variety

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

(MSA), which is the official variety of Arab governments, schools, and print publications, there are numerous regional/spoken varieties that are used in everyday conversations and informal settings, including TV shows and sports (Albirini, 2016). These spoken varieties lack a standardized written form, they are acquired naturalistically, and their mutual intelligibility depends on their geographical distance (Saiegh-Haddad & Henkin-Roitfarb, 2014). The Syrian variety, in particular, belongs to the Levantine geographical/linguistic group (Aoun, Benmamoun & Choueiri, 2010).

In this section we introduce the basic morphosyntactic properties of the Syrian Arabic structures that are of interest to the present study and compare them to their English counterparts. Our description is based on two main sources (Brustad, 2000; Cowell, 1964), as well as on examples and judgements provided by two native speakers and one highly proficient speaker of the Syrian variety.

Simple declarative sentences

Simple declarative sentences in Syrian Arabic typically follow an SVO (1a) or a VSO (1b) word order (Brustad, 2000: 361). Verbs inflect for person, number, and gender (Cowell, 1964: 421), while subjects, being inflected on the verb, are often omitted (2) (Brustad, 2000: 317; Cowell, 1964: 418). In this respect, Arabic differs from English, a language with a rigid SVO word order that typically requires subjects to be overtly realized.

(1) a. *Neʕmæt* *dæff-et* *l-bæːb*[1]
neemat pushed-3SG.F the-door
b. *dæff-et* *Neʕmæt* *l-bæːb*
pushed-3SG.F Neemat the-door
'Neemat pushed the door.'

(2) *dæff-et* *l-bæːb*
pushed-3SG.F the-door
'(She) pushed the door.'

Topicalization

In addition to SVO and VSO, other word orders can be used in Syrian Arabic to change the information packaging of the sentence. Object topicalization, in particular, involves a sentence initial object that is co-indexed with a clitic in the post-verbal position, as illustrated in (3) (Brustad, 2000: 348–351).

(3) *lœilœ* *ʕeʃʔ- a* *qais*
Layla adored-her Qays
'(As for) Layla, Qays adored her.'

In this example, the object (*lœilœ*) is sentence initial and co-indexed with the object clitic *-a* 'her'. Importantly for the comparison at hand, even though English allows for the fronting of objects, it does not have object clitics.

[1]The abbreviations used in the Arabic examples are the following: F = feminine, FUT = future, M = masculine, PASS = passive marker, PL = plural, SG = singular.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Passive

Whereas topicalization changes the typical word order of the sentence by pre-posing the object, passivization changes the typical mapping between semantic roles and grammatical relations. As an example, we may consider the active construction in (4) and its passive counterpart in (5).

(4) *neʃmæt* *dœff-et* *l-bœːb*
neemat pushed-3SG.F the-door
'Neemat pushed the door.'

(5) *n-dœfœf* *l-bœːb*
PASS–pushed the-door
'The door was pushed.'

Note that, differently from English passive constructions, Syrian Arabic passive constructions do not license an agent phrase and are often ambiguous between what Cowell (1964: 239) calls "a true passive" interpretation (where an external Agent is implied even though it is not morphologically realized) and a "mediopassive" interpretation (where no external Agent is implied).

Information questions

The last monoclausal structure that we consider is the information question. In Syrian Arabic, information questions contain sentence-initial interrogative words, on par with their English counterparts. This is illustrated with examples (6a-c):

(6) a. *miːn* *ʃəf-t* Ø *bə-l-matʕam*
who saw-2SG.M in-the-restaurant
'Who did you see in the restaurant?'
b. *ʔœyyœ* *mumœssel* *ʃəf-t-ɔ* *bə-l-matʕam*
which actor saw-2SG.M-him in-the-restaurant
'Which actor did you see in the restaurant?'
c. *miːn* *əlli* *ʃəf-t-ɔ* *bə-l-matʕam*
who that saw-2SG.M-him in-the-restaurant
'Who did you see in the restaurant?'

Even though all the examples above are object questions (in that the interrogative pronoun is understood as the object of the verb) they differ in the strategy they use to encode the dependency between the sentence-initial interrogative word and the canonical object position (for MSA, see Aoun et al., 2010, Chapter 6). Example (6a) illustrates the gap strategy, where the object position is empty (Ø). Example (6b) illustrates the resumptive strategy, where the object position is occupied by a clitic (*-ɔ* 'him') that is co-indexed with the fronted interrogative. Finally, in (6c), the object clitic (*-ɔ* 'him') is embedded within a relative clause introduced by the relative complementizer *əlli* 'that'. Of the three strategies only the gap strategy is available in English.

Coordination and subordination

In addition to the monoclausal sentences discussed above, we also consider biclausal sentences created either through coordination or through subordination. The two

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

strategies are illustrated in (7) and (8), respectively, with Arabic. As evident by the translation, they are structurally similar in English.

(7) *sœ:miœ* *ʕəml-et* *l-ʕɒʃœ* *w* *tˁəlʕ-et*
Samia prepared-3SG.F the-dinner and left-3SG.F
'Samia prepared dinner and left.'

(8) *rɒħ* *y-lʕɒb* *l-wlœd* *bə-l-tˁa:be* *izœ* *sœ:ʕad* *ʔmm-ɔ:*
FUT 3SG.M-play the-child with-the-ball if help mother-his
'The child will play with the ball if he helps his mother.'

Relative clauses

Finally, a special subtype of subordination involves relative clauses (RCs). Similar to English, RCs in Arabic are head initial, which means that they follow the head/antecedent that they modify. Furthermore, they are introduced by the complementizer *əlli* and, in the case of object relativization, they preferably use the resumptive strategy (Brustad, 2000: 89–111). As an example, we may consider the object RC in (9).

(9) *bœ-ʕref* *l-mumœssile* *əlli* *rɒħ* *y-ʃu:f-œ* *ʔœħmœd*
1SG-know the-actress that FUT 3SG.M-see-her Ahmed
'I know the actress that Ahmed will see' (adapted from Aoun et al., 2010)

Note that in (9) the RC is introduced by the complementizer *əlli* 'that', whereas the head/antecedent *l-mumœssile* 'the actress' is co-referential with the object resumptive clitic *-œ* 'her' within the RC. Overall, the availability of the resumptive strategy and postverbal subjects, together with the obligatory realization of the relative complementizer, differentiate Arabic from English object relatives.

Present study

The main goal of this study was to assess the HL Syrian Arabic and L2 English syntactic skills of Syrian refugee children ($N = 119$) ages 6–13 who recently immigrated to Canada, focusing on the factors that predict abilities in each language and on potential relationships of interdependence between the two languages. In order to do so, we investigated participant performance on an SRT by measuring OVERALL ACCURACY and SYNTACTIC ACCURACY in responses. The former measured the proportion of correct/incorrect morphemes and therefore factored in lexical as well as morphosyntactic errors. Syntactic accuracy, on the other hand, measured whether participants repeated the target syntactic structure in their productions accurately. We asked the following research questions:

1) Do Syrian refugee children perform better in their HL Arabic than L2 English with respect to overall accuracy and syntactic accuracy on an SRT? Are the same trends observed for overall accuracy and syntactic accuracy? Is interdependence observed between languages?

We predicted that at this early stage in their bilingual development (i.e., after an average of two years of onset of bilingualism) children would perform better in their HL Arabic than in their L2 English with regard to both overall accuracy and syntactic accuracy.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

We also predicted that there would be less individual variation in the HL than in the L2 since these participants had several years of HL exposure obtained in a quantitatively and qualitatively rich Arabic-speaking environment that would have allowed for their grammar to stabilize prior to immigration. Furthermore, if interdependence holds for oral language skills as it does for literacy skills, there should be a positive association between children's performance in the Arabic and English SRT.

2) What is the contribution of the following child- and language-level factors in predicting overall accuracy and syntactic accuracy in HL Syrian Arabic and L2 English: AOA, cognitive skills, richness of the environment, relative HL/L2 use, LOE, parental education, and target syntactic structure? Do the same factors predict performance in both languages and measures?

Given the special profile of our population (short residency in Canada and varied length of residency in Syria before migration), we predicted a differential effect of child-level variables on HL Syrian Arabic and L2 English syntax. Specifically, we predicted that, at this early stage of L2 exposure, measures indexing CURRENT amount/richness of the language environment might be more predictive for the L2 English than for the HL Arabic. This is because, differently from the studies reviewed in our introduction, our study focuses on newcomers with longer exposure to their HL in the home country and shorter exposure to the L2 in the host country. With respect to language-level variables, and in line with the literature discussed in our introduction, we predicted that target syntactic structure would be a significant predictor of performance in both languages. That is, we expected participants to be less accurate with structures involving long-distance dependencies; especially the ones that rely on morphology. Finally, in terms of the influence of factors across the two measures, we hypothesized that if syntax, like vocabulary and morphology, is sensitive to individual differences, then children's performance in syntactic accuracy should be affected by the same range of factors as their performance in overall accuracy.

Method

Participants

This study presents data from Wave 1 of an on-going longitudinal study. It includes 119 participants (57 females), all of them Syrian refugee children. These participants had a mean age of 9.37 years ($SD = 1.97$, range = 6–13). At the time of testing, they had resided in three English-majority cities in Canada for a mean of 24.46 months ($SD = 9.09$; range = 10–37). All participants and their parents were native speakers of Syrian Arabic. None of the children had been exposed to English before their arrival to Canada. The 119 participants came from 67 families, indicating that there were siblings in the sample. Most families did not emigrate into Canada directly from Syria. Ninety-one participants spent some months in Arabic-speaking countries (Jordan, Lebanon, Egypt, and Qatar) and 15 participants emigrated into Canada from Turkey. More details on participant characteristics appear in the first section of the *Results* and in Table 3.

Procedures

The measures in the present study are part of a larger battery of linguistic and non-linguistic tests; only measures whose results are reported in this study are discussed here.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Participants were met once or twice (for about 90 minutes total) for each language. Testing took place either in children's homes or at school. Parent questionnaires were administered at home or at the children's school. Only one language was used and tested in a given session. Language order was counterbalanced: the first session for each participant was in the opposite language from the participant tested before. The order of the tasks was fully randomized across participants. The research assistants who carried out the testing for Arabic and English were native speakers of the language.

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

To elicit information on the child-level variables that might have affected participants' performance in their two languages, we used the ALEQ-4. This questionnaire was delivered to parents in an interview format in Arabic by a native speaker of Arabic and gathered demographic information on both the family and the child (age, length of residency in Canada, English AOA, length of schooling in English and Arabic, and parental education), together with information on the amount of language use at home as well as the overall richness of the language environment.

More precisely, to collect information on the relative use of Arabic/English in the home, parents described, using a 1–5 scale, how much Arabic and English was used in the household by each relative (1 = Mainly or only Arabic, 2 = Usually Arabic/English sometimes, 3 = Arabic and English, 4 = Usually English/Arabic sometimes, 5 = Mainly or only English). While this information was collected for each member of the family in terms of output given to and received from the child, we calculated composite scores of relative Arabic/English use across parents, on the one hand, and siblings, on the other.

To estimate the frequency of Arabic and English oral language and literacy activities (language environment richness), parents were asked to indicate, for each language, how many hours per week participants spent 1) reading and writing (including books, messaging, and homework), 2) speaking and listening (such as TV, music, and Skype), 3) taking part in extra-curricular activities (e.g., sports, clubs, and religious services), and 4) playing with friends. The measure of Arabic richness also included the number of hours per week of HL classes. Parents estimated the frequency of these activities using a 1–5 scale: (1 = 0–1 hours, 2 = 1–5 hours, 3 = 5–10 hours, 4 = 10–20 hours, 5 = 20+ hours). From these scales, a richness proportion score (0–1) was calculated for each language, where 1 would indicate high frequency of language-rich activities. Participant characteristics are summarised in Table 3 below.

Matrix Analogies Test (MAT; Naglieri, 1985)

To evaluate participants' cognitive abilities, we used the MAT, a test that measures non-verbal analytical skills. Participants were administered two subtests of the MAT: reasoning by analogy and spatial visualization. Both subtests asked participants to select the picture that best completed a matrix. Since only two subtests were used, a standard score could not be computed. Instead, a compound score of the two subtests was calculated, which ranged from 0 to 32. Participants' scores appear in Table 3. Instructions for this task were minimal and were given in Arabic.

Sentence Repetition Task (SRT)

An SRT was used to provide information on participants' abilities in both languages. In SRTs, participants are asked to listen to a sentence and to repeat it as close as possible to the original one. An SRT was chosen for this study because it conferred two specific

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 1. Structures used in the English SRT

Name	Syntactic structure
Declarative ($k = 6$)	Simple, active declarative sentences with auxiliaries and/or modals
Short passive ($k = 3$)	Short passives (no agent *by*-phrase)
Long passive ($k = 3$)	Long (with agent *by*-phrase) passive
Question ($k = 6$)	Object questions
Coordinated ($k = 3$)	Coordinated sentences
Subordinate ($k = 4$)	Conditional/temporal subordinate clauses
Relative ($k = 6$)	Object relative clauses (Object-Object and Subject-Object)

advantages: it allowed us to include a wide range of syntactic structures that varied in complexity, and it provided a source for comparison across other studies that have also employed an SRT to assess language skills. Despite early concerns that SRTs may mostly tap onto memory skills, recent empirical research has provided proof of construct validity for SRTs: these tasks effectively measure language ability at a lexical, grammatical, and speech production level (Klem, Melby-Lervåg, Hagtvet, Lyster, Gustafsson & Hulme, 2015). More specifically, morphosyntactic skills have been found to be central in SRT performance. Polišenská, Chiat and Roy, P. (2015), who delivered an SRT to 4–5-year-old children, manipulated a variety of linguistic factors in their stimuli: prosody, semantic plausibility, lexicality, and syntactic grammaticality. Their results showed that grammaticality was the most significant factor in determining children's ability to repeat a target sentence, effectively demonstrating that children must access their morphosyntactic skills in order to recall and reconstruct a target sentence. For this reason, recent research has used SRTs to tap onto and assess morphosyntactic and syntactic skills in a variety of languages (Armon-Lotem et al., 2011; Kaltsa et al., 2020; Meir et al., 2017; see recent review of other studies in Komeili, Marinis, Tavakoli & Kazemi, 2020).

The SRT used in this study was comprised of 32 stimuli (one practice item and 31 scored items). The stimuli for the SRT were based on Marinis and Armon-Lotem's (2015) LITMUS SRT for English. The English version consisted of the seven structures shown in Table 1. All sentence stimuli included basic pragmatic and lexical content that would be familiar to most children, regardless of cultural affiliation or SES. One example for each structure is shown in (10–16) and all the stimuli appear in Appendix 1. The English SRT stimuli were recorded by a native speaker of Canadian English.

(10) She can bring the glass to the table. (Declarative)
(11) She was stopp-ed at the big red light-s. (Short passive: no agentive phrase)
(12) She was se-en by the doctor in the morning. (Long passive: with agentive phrase)
(13) Who have they se-en near the front door? (Question)
(14) Our neighbor clean-s the car and his son play-s basketball. (Coordinated)
(15) If the weather is warm, we can go to the park. (Subordinate)
(16) The children enjoy-ed the candy that they tast-ed. (Relative)

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 2. Structures used in the Arabic SRT

Name	Structure
Declarative ($k = 6$)	Simple active declarative sentence with modals and/or particles
Short passive ($k = 3$)	Short passive (no agent *by*-phrase)
Topicalization ($k = 3$)	Topicalization
Question ($k = 6$)	Object questions
Coordinated ($k = 3$)	Coordinated
Subordinate ($k = 4$)	Conditional/temporal subordinate clauses
Relative ($k = 6$)	Object relative clauses (Object-Object and Subject-Object)

The Syrian Arabic SRT was developed using the English SRT as a model to enable close comparisons between children's performance on both tasks. Syrian Arabic was chosen over MSA for this task since it is the variety children use at home with their parents and siblings, and because many of the children had interrupted or no experience with schooling in Arabic, which is the foundation for learning MSA. Stimuli were first translated in the spoken variety by a native speaker of the closely-related variety of Jordanian Arabic. Subsequently, two native speakers of Syrian Arabic acted as consultants and changes were made to some items for the Syrian variety specifically. A few stimuli items received variable judgments by the native speakers, reflecting the dialectal variation within Syria. Stimuli used for the Arabic SRT were the ones agreed upon by either all or at least two of the Arabic native speakers. The structures for the Arabic SRT appear in Table 2, example stimuli appear in (17–23), and all the stimuli appear in Appendix 2. Note that short passives (i.e., passives without an agentive phrase) in Arabic are often ambiguous between a passive and a mediopassive interpretation (Cowell, 1964). Note also that long passives (i.e., passives with an agentive phrase) do not exist and were thereby replaced with object topicalizations. Even though long passives and object topicalizations are not structurally equivalent, they are comparable in complexity and pragmatic function in the sense that they both reverse the canonical order of Agent and Patient (on the functional equivalence of long passives and object topicalizations, see Cowell 1964, and El-Yasin, 1996). The English SRT and the Arabic SRT had a similar average number of morphemes per sentence (10.03 and 10.06, respectively). The Arabic SRT stimuli were recorded by a native speaker of Syrian Arabic, who is a recently arrived refugee.

(17) *lœzem* *t-ħɒtˁ* *l-kœ:se* *ʕœ-tˁ-tˁawle* (Declarative)
must 3SG.F-put the-cup on-the-table
'She must put the glass on the table.'

(18) *n-dœfœʃ* *b-ʔuwe* *ʕœ-l-ʔardˁ* (Short Passive: no agent)
PASS-pushed in-hard to-the-ground
'He was pushed hard against the ground.'

(19) *l-ʔm* *leħʔ-a* *esˁ-sˁabi* *ʕœ-ʃ-ʃaːreʕ* (Topicalization)
the-mother followed-her the-boy to-the-street
'As for the mother, the boy followed her to the street.' (lit. 'The mother, the boy followed her to the street.)

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

(20) *miːn ʃœːf-u jœnb l-bœːb l-ʔœmœːmiː* (Question)
who have.seen-3PL near the-door the-front
'Who have they seen near the front door?'

(21) *l-ʔm ʕœm tət-sœwwaʔ* (Coordinated)
the-mother PROG 3.SG.F-shop
w l-wœlœd ʕœm yi-drɔs bə-l-beːt
and the-boy PROG 3.SG.M-study in-the-home
'The mother is shopping and the boy is studying at home.'

(22) *rɒħ y-aːxd-u l-wlœːd hdiːə* (Subordinate)
FUT 3-get-PL the-children present
izœ nadˤdˤaf-u l-beːt
if clean-3PL the-house
'The children will get a present if they clean the house.'

(23) *ənbasatˤ-u l-wlœːd bə-ʃ-ʃɔkɔlata əlli ʔœkœl-u-wa* (Relative)
enjoyed-3PL the-children with-the-chocolate.F that ate-3PL-her
'The children enjoyed the chocolate that they ate [it].'

Administration of the SRT

Participants were presented with the 32 pre-recorded stimuli of the SRT, one at a time, using a laptop (PowerPoint) while wearing noise-cancelling headphones. There were two breaks during the task. Participants' repetitions were recorded for later transcription, scoring and analysis. The first sentence, which was not scored, was a practice item. Participants were allowed to listen to the practice item more than once to ensure they understood the task mechanics but were not allowed to listen to the other stimuli more than once.

Scoring of overall accuracy

Accuracy on an SRT can be scored in different ways. We employed two types of scoring for this study: overall accuracy and syntactic accuracy. OVERALL ACCURACY was a measure of how closely the stimuli sentences were repeated. It was captured as a percentage score indicating the percentage of correct morphemes in participants' productions. Substitutions, omissions, additions, and movements of content and functional morphemes were considered errors, even if they did not render the production ungrammatical, and all types of errors carried the same weight. On the other hand, mispronunciations or retracings were not counted as errors. Below is one example of scoring for an English and for an Arabic stimulus:

(24) The teacher has been look-ing at us all day. (10 morphemes)
(25) The teacher has look-*ed* at us all day. (Production example; 8 correct morphemes + 2 errors [omission of *been* and replacement of *-ing* by *-ed*]; Score: 80%)
(26) *miːn ʃœːf-u jœnb l-bœːb l-ʔœmœːmiː* (Target sentence; 8 morphemes)
'Who have they seen near the front door?'
(27) *miːn ʃœːf ø jœnb l-bœːb l-ʔœmœːmiːʔ* (Production example; 7 correct morphemes + 1 error [omission of *-u*]; Score: 87.5%).

Three native Arabic speakers transcribed and scored the recordings for Arabic and four native English speakers did so for English. Inter-rater reliability was assessed on

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

the number of errors for 25% of the recordings. Results of the two raters were compared to obtain the Krippendorff's α coefficient of reliability using the `irr` package on R (Gamer, Lemon, Fellows & Singh, 2019). Krippendorff's α is more appropriate to assess the rate of agreement on interval data than other measures, such as percent agreement (Hayes & Krippendorff, 2007). The Krippendorff's α coefficient for English and Arabic overall accuracy was .94 and .93, respectively. Data with $\alpha \geq .80$ is considered reliable (Krippendorff, 2004, p. 241).

Scoring of syntactic accuracy

The second type of scoring assessed whether participants had repeated the target syntactic structure accurately or not regardless of other errors in their production. This type of scoring narrowed down the aspects of performance to be assessed by disregarding lexical and morphological errors that did not compromise the target syntactic structure. Hence it was aimed at being a more direct proxy of syntactic abilities than overall accuracy. We chose sentence-recall ability as the measure of language because there is a strong link between sentence repetition and syntactic competence (Frizelle & Fletcher, 2014; Gallimore & Tharp, 1981; Geers & Moog, 1978; Kidd et al., 2007; Polišenská et al., 2015).

Productions that did not contain at least one inflected verb were considered unscoreable for the measure of syntactic accuracy. Unscoreable sentences made up for 1.8% and 6.8% of all productions in Arabic and English, respectively. The remaining (i.e., scoreable) sentences were analyzed for syntactic accuracy. The requirements for each structure to be considered preserved in participants' repetitions depended on the target structure. For example, for subordinate structures to be considered preserved, participants' productions had to contain two clauses (one embedded) and a subordinator. Examples (28–29) illustrate two repetitions for the sentence "*The children will get a present if they clean the house*" (subordinate). Example (28) was scored as syntactically accurate because the requirements for the subordinate structure were met despite five errors of omission or replacement of different content and function words. On the other hand, (29) was considered to not have been repeated accurately because even though there were two clauses with the target content and function words, the subordinator "if" had been replaced by the coordinator "and".

(28) Boys get present if they clean house.
(29) The children will get a present *and* they clean the house.

Syntactic accuracy coding was conducted after the initial coding for overall accuracy for the dataset was completed. In order to ensure close control and full reliability, every sentence was scored independently twice and all disagreements were settled by group discussion.

Results

Descriptive statistics

Participant characteristics appear in Table 3. Differently from most studies on bilingual children, age at testing is not an index of LOE to English in this sample. As these participants form a cohort which arrived in Canada at approximately the same time (and have therefore spent a similar amount of time in Canada), age at testing and

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 3. Participant characteristics

	Mean	SD
Age at testing (years)	9.37	1.97
Length of residency in Canada (months)	24.46	9.09
Age of onset of English acquisition (years)[a]	7.64	1.98
Length of English schooling (months)	18.92	5.57
Length of Arabic schooling (months)	13.25	14.09
Maternal education (years)	9.69	4
Paternal education (years)	9.84	3.96
Relative Arabic/English use with parents[b,d]	1.09	.25
Relative Arabic/English use with siblings[c,d]	1.59	.96
English environment richness[e]	.43	.10
Arabic environment richness[e]	.41	.13
Non-verbal analytical (MAT) score[f]	8.65	5.67

Note. [a]Measured as onset of English schooling. [b]Relative scores averaged across both parents for input to and output from the child. [c]Relative scores averaged across all siblings for input to and output from the child. [d]Measured on scale 1 = Mainly or only Arabic, 2 = Usually Arabic/English sometimes, 3 = Arabic and English, 4 = Usually English/Arabic sometimes, 5 = Mainly or only English. [e]Proportion score of frequency of reading/writing, speaking/listening, extra-curricular, and playing with friends in a week. Arabic richness included heritage language classes. [f]Out of 32.

English AOA are strongly correlated ($r = .96$, $p < .001$). Given this almost perfect correlation, we only included English AOA in the statistical modelling below.

For these participants, LOE to English can be indexed by length of English schooling or length of residency in Canada. Again, these two variables are strongly correlated ($r = .85$, $p < .001$). For all statistical modelling below, we used length of English schooling as a measure of exposure to the L2, since families do not tend to use English in the home (see Table 3), and therefore schooling in English provides a more reliable measure of exposure to the language.

Participants' schooling in Arabic, which was just over one full academic year ($M = 13.25$, $SD = 14.09$), was on average shorter than schooling in English ($M = 18.92$, $SD = 5.57$). This illustrates the fact that most participants had interrupted schooling prior to immigration. In fact, of the 119 participants, only 71 had some schooling in Arabic (i.e., in MSA); ranging from 1 month of schooling to 60. Even though 15 participants spent time in Turkey prior to migration to Canada, only one child had some schooling in Turkish. Arabic schooling was moderately correlated with age ($r = .58$, $p < .001$) and English AOA ($r = .64$, $p < .001$) indicating that older participants and participants that were older at the onset of bilingualism had had longer schooling pre-migration.

Parental education was around 10 years, indicating that the majority of the parents in the sample had primary or secondary schooling only. This signals that the sample was skewed toward low SES. Maternal and paternal education were moderately correlated in the sample ($r = .59$, $p < .001$). The measures of Arabic/English use with parents and siblings indicated that our participants lived in Arabic-dominant households (as scores close to "1" indicate Arabic only was spoken). In fact, due to the limited variation in language use with parents, this variable could not be included in any statistical models. In terms of language environment richness,

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

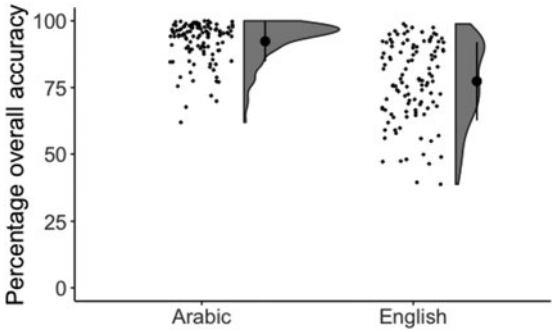

Figure 1. Participant overall accuracy reported in mean percentage of correct morphemes that were repeated accurately. Points show individual participants, which appear jittered. Width of split violin shows density. The point over the density plot signals the sample mean and the bar, one standard deviation below and above the mean.

participants showed a similar but low frequency of activities in English and in Arabic (scores of .60 or higher out of 1.0 would be mid to high frequency; Paradis, 2011; Paradis, Soto-Corominas, Chen & Gottardo, 2020). MAT scores were moderately correlated with age ($r = .44$, $p < .001$), which was expected since scores were not age-referenced. By including MAT scores in the model, we were able to specify the variation due to these cognitive skills separately from general cognitive maturity due to age/AOA.

Performance on the SRT

To evaluate children's syntactic abilities in Arabic and English (RQ 1), we produced two scores: overall accuracy and syntactic accuracy.

Overall accuracy

Participant overall accuracy is displayed in Figure 1. Scores are percentages, indicating the mean percentage of morphemes that were repeated accurately (i.e., a 100% would indicate that the participant repeated all morphemes correctly for all sentences).

In the Arabic task, participants' accuracy was 92.36% ($SD = 7.57\%$, range = 61.91–100%) and in the English task it was 77.34% ($SD = 14.73\%$, range = 38.90–98.89%). A paired Wilcoxon signed-rank test with continuity correction (suitable for non-normally distributed samples) found that participants were significantly more accurate in Arabic than in English ($p < .001$; Cohen's $d = 1.28$, large effect size).

Syntactic accuracy

Figure 2 displays participants' syntactic accuracy in percentages. The mean for Arabic was 85.58% ($SD = 12.33\%$, range = 32.26–100%) and for English it was 57.62% ($SD = 27.69\%$, range = 3.23–100%). According to a Wilcoxon signed-rank test, participants were significantly more accurate in Arabic than in English ($p < .001$; Cohen's $d = 1.12$, large effect size).

A breakdown of participant scores divided by syntactic structure is provided in Figure 3 for Arabic and Figure 4 for English. The statistical analysis on structural performance is reported in the section below.

Interdependence

As part of the first research question, we asked whether interdependence would be observed between the two languages. In order to answer this, we ran partial

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

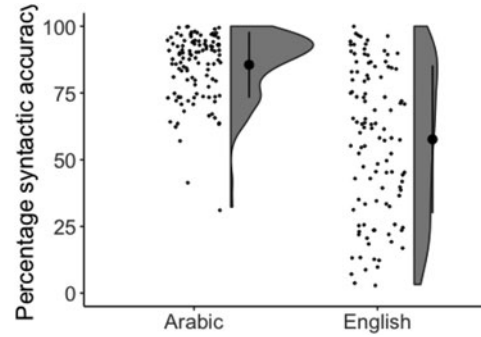

Figure 2. Participant syntactic accuracy reported as percentage of structures that were repeated accurately. Points show individual participants, which appear jittered. Split violin shows density. The point over the density plot signals the sample mean and the bar, one standard deviation below and above the mean.

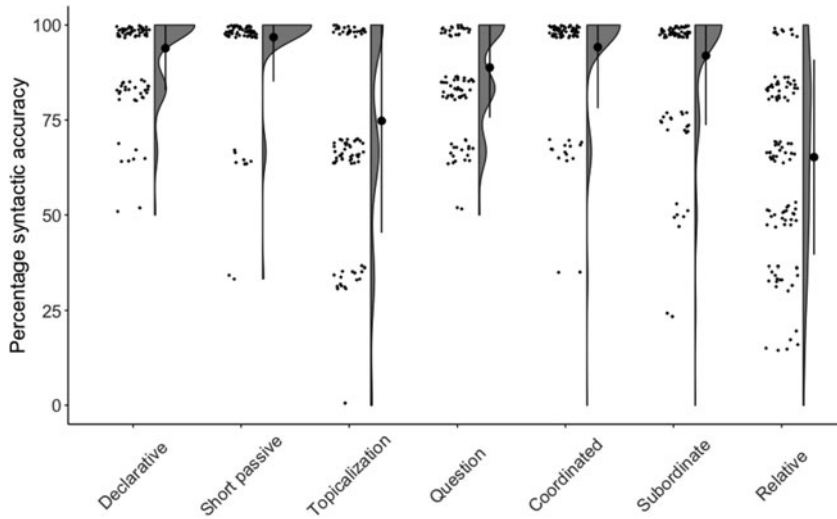

Figure 3. Participant syntactic accuracy in Arabic reported by structure. Points show individual participants, which appear jittered to avoid overlap due to discreteness. Split violin shows density. The point over the density plot signals the sample mean and the bar, one standard deviation below and above the mean (capped at 100%).

correlations, keeping length of English schooling (i.e., LOE to the L2) constant, between performance in Arabic and English in terms of overall accuracy and syntactic accuracy. Both correlations were positive, moderate, and significant (overall accuracy: $r = .55$, $p < .001$; syntactic accuracy: $r = .47$, $p < .001$), indicating that participants who performed better in one language tended to perform better in the other language with regard to both outcome measures. This was also the case when correlations were run separately for each structure, with coefficients ranging between $r = .22$-$.33$ (all $p < .001$).

Factors predicting individual variation

Analyses to determine which predictors influenced overall performance and syntactic accuracy in both languages (RQ 2) were conducted using logistic linear mixed-effects regression with the package lme4 in R (Bates, Maechler, Bolker & Walker, 2015). For

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

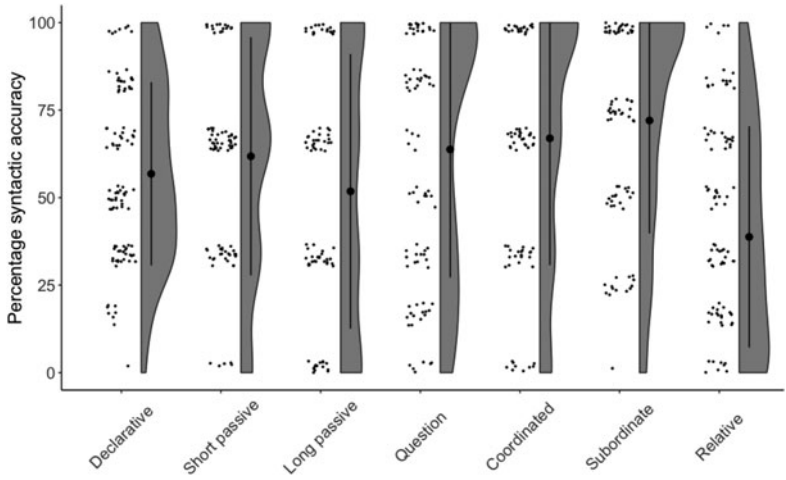

Figure 4. Participant syntactic accuracy in English reported by structure. Points show individual participants, which appear jittered to avoid overlap due to discreteness. Split violin shows density. The point over the density plot signals the sample mean and the bar, one standard deviation below and above the mean (capped at 100%).

all models, random effects (intercepts) included item and participant. Participant was nested within family in order to control for any variability arising from the fact that some participants were siblings. Information about the dependent variable and the fixed effects (predictors) for each model appears below. In all cases, an initial model was fit that included all predictors and backwards selection, using log-likelihood ratio tests and inspection of AIC values, was followed to obtain the optimal model.

Predictors of overall accuracy

The dependent variable for the model analyzing overall accuracy was a proportion score of correct/incorrect morphemes for each sentence. The initial model for overall accuracy in Arabic included the following fixed effects: English AOA, length of Arabic schooling, non-verbal analytical skills, Arabic richness, maternal and paternal years of education (entered separately), relative English/Arabic use with siblings, syntactic structure, and sentence length. SYNTACTIC STRUCTURE was a categorical variable with seven levels (declarative, passive, topicalization, question, coordinate, relative, and subordinate) and six levels in English (declarative, passive, question, coordinate, relative, and subordinate). SENTENCE LENGTH was only included in order to control for the fact that longer sentences contained more morphemes and, hence, more potential for mistakes. All fixed effects were numerical and were scaled and centered around 0. The optimal model for Arabic overall accuracy is shown in Table 4. English AOA and non-verbal analytical skills were positive predictors, indicating that participants who started acquiring English later in life and participants with higher analytical skills were more accurate on the Arabic SRT. There was a significant main effect of syntactic structure, which appears unpacked in the analysis for syntactic accuracy.

The initial model for overall accuracy in the English SRT included the following fixed effects: English AOA, length of English schooling, non-verbal analytical skills,

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 4. Coefficient table for fixed effects in optimal model predicting overall accuracy in Arabic SRT

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.686	0.286	12.873	<.001***
English AOA	0.507	0.107	4.736	<.001***
Non-verbal analytical skills	0.243	0.107	2.275	.023*
Structure Question	−0.479	0.378	−1.265	.206
Structure Passive	0.328	0.481	0.682	.495
Structure Topicalized	−0.922	0.463	−1.993	.046*
Structure Coordinated	−0.210	0.659	−0.319	.750
Structure Subordinate	−0.287	0.432	−0.663	.508
Structure Relative	−1.211	0.415	−2.920	.004**
Sentence length	−0.206	0.218	−0.947	.343

Note. Estimates are in log-odds. All numerical predictors have been scaled and centered around 0. Reference level for structure is declarative (with auxiliary or modal).

Table 5. Coefficient table for fixed effects in optimal model predicting overall accuracy in English SRT

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.597	0.128	12.464	<.001***
English AOA	0.221	0.079	2.794	.005**
Non-verbal analytical skills	0.234	0.076	3.060	.002**
Length of English schooling	0.239	0.076	3.122	.002**
English richness	0.273	0.077	3.536	<.001***
Maternal years of education	0.475	0.075	6.310	<.001***
English/Arabic use with siblings	0.276	0.082	3.364	.001**
Sentence length	−0.410	0.107	−3.828	<.001***

Note. Estimates are in log-odds. All numerical predictors have been scaled and centered around 0.

English richness, maternal and paternal years of education, relative English/Arabic use with siblings, sentence structure, and sentence length. All fixed effects were numerical and were scaled and centered around 0.

The optimal model for English is shown in Table 5. While English AOA (i.e., onset of schooling), length of English schooling, non-verbal analytical skills, English richness, maternal years of education, and English use with siblings were positive predictors (i.e., the higher, the more accurate the participant was on the task), sentence length was a negative predictor. That is, participants were significantly less accurate with increasing sentence length.

Predictors of syntactic accuracy

The analysis of syntactic accuracy only considered errors that affected the syntactic structure and disregarded other (lexical, morphological) errors. In order to ascertain

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 6. Coefficient table for fixed effects in optimal model predicting structural accuracy in Arabic SRT

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.961	0.506	7.826	<.001***
Length of Arabic schooling	0.482	0.134	3.599	<.001***
Structure Question	−0.682	0.669	−1.018	.309
Structure Passive	0.667	0.858	0.777	.437
Structure Topicalized	−2.210	0.792	−2.792	.005**
Structure Coordinated	0.254	0.839	0.302	.762
Structure Subordinate	−0.653	0.740	−0.881	.378
Structure Relative	−2.864	0.656	−4.363	<.001***

Note. All fixed effects had been scaled and centered around 0.

whether participants' syntactic performance was affected by the same range of factors as their overall accuracy, a second logistic regression model was fit for each language. The dependent variable in these models was syntactic structure preserved (1) or not (0) for each individual sentence, not the percentage of sentences whose structure were repeated accurately (displayed in Figure 2). The full models for syntactic accuracy included the same child-level factors from the models for overall accuracy, together with syntactic structure. As explained above (see Procedures), stimuli that did not contain an inflected verb were considered unscoreable and thus were not included in the models below.

The optimal model for Arabic syntactic accuracy appears in Table 6. This model had a .90 *C*-index of concordance, indicating excellent goodness-of-fit (Levshina, 2015). Only syntactic structure and length of Arabic schooling contributed significantly to this model. That is, participants with longer Arabic schooling tended to repeat the target syntactic structure accurately more often. In order to assess accuracy with individual syntactic structures, we ran post-hoc contrasts with a Tukey adjustment with the package `emmeans` in R (Lenth, 2019). The contrasts indicated that relative clause structures had been repeated significantly less accurately than declaratives ($p < .001$), questions ($p = .01$), short passives ($p < .001$), coordinated structures ($p = .003$), and subordinated structures ($p = .03$). In turn, topicalized structures were also repeated significantly less accurately than short passives ($p = .03$). The pairwise contrast between topicalized and declarative sentences almost reached significance ($p = .08$).

The optimal model for English syntactic accuracy did not contain syntactic structure. That is, the model without structure was as predictive as the model with structure ($\chi^2(5,15) = 9.5853$, $p = .088$; AIC with structure = 2768.8; AIC without structure = 2768.4). The optimal model is shown in Table 7. This model had a C-index of concordance of .90, indicating outstanding model fit. Non-verbal analytical scores, length of English schooling, English richness, maternal and paternal years of education, and English use with siblings emerged as positive predictors.

Discussion

The present study sought to describe the syntactic development in the HL Syrian Arabic and L2 English at the early stages of bilingualism of Syrian refugee children, a

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Table 7. Coefficient table for fixed effects in optimal model predicting structural accuracy in English SRT.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.718	0.292	2.460	.014*
Non-verbal analytical scores	0.477	0.147	3.238	.001**
Length of English schooling	0.536	0.148	3.634	<.001***
English richness	0.413	0.162	2.550	.011*
Maternal years of education	0.585	0.183	3.198	.001**
Paternal years of education	0.366	0.185	1.981	.048*
English/Arabic use with siblings	0.505	0.162	3.123	.002**

Note. All fixed effects had been scaled and centered around 0.

population whose language development had not yet been addressed specifically. This study had three main goals: (1) to determine whether the syntactic skills are stronger for the HL than for the L2 after two years of bilingualism, (2) to determine whether there is a cross-linguistic relationship (i.e., interdependence) between the syntactic skills in the two languages, and (3) to explore the child-level and language-level factors that predict performance in each language.

Overall accuracy and syntactic accuracy within and across the two languages

The first research question asked whether participants would perform better in HL Arabic than in L2 English and whether interdependence relationships would be observed. Consistent with our predictions, participants' overall accuracy and syntactic accuracy was significantly higher in HL Arabic than in L2 English (Figures 1–2). As indicated by the means and SDs, participants performed mostly at ceiling in both Arabic measures and showed less individual variation than in English. These findings contrast with a study of sequential Russian-Hebrew bilingual children who performed better in L2 Hebrew (Meir et al., 2016). The discrepancy between the results of Meir et al. (2016) and the results of the present study were to be expected. In contrast to the preschoolers studied in Meir et al. (2016), the Syrian refugee children in this study were older and were born and raised in an Arabic-speaking country. Furthermore, they had a later L2 AOA (91.68 months vs. 34.60 months) and a shorter LOE to the L2 (19 months vs. 39 months). The performance gap between the HL Arabic and the L2 English of our participant sample is likely to diminish and eventually change polarity in the subsequent years, accompanied with increased variability in HL abilities, in line with what has been reported for other HL populations born and raised in the host country (Montrul, 2016). However, our findings indicate that attrition of HL abilities is not apparent in the first two years of residency in the host country.

Turning to the possibility of interdependence between HL and L2 syntactic abilities, our results indicated that there was indeed a moderate, positive correlation between performance in HL Arabic and L2 English. These results are in line with Castilla et al.'s (2009) study and are consistent with the hypothesis that the HL-L2 interdependence is not confined to literacy but it can also extend to oral language skills, including syntax. Importantly, this study was not designed to test whether this

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

interdependence stems from a language-general underlying ability that promotes performance in both languages (i.e., a "common underlying proficiency" in Cummins' [2000] words) or from the linguistic skills in the HL, which directly support the development of linguistic skills in the L2 (Castilla et al., 2009). In either case, what is clear is that, at these early stages of bilingualism, strong HL skills provide a strong foundation for the development of the L2.

Sources of individual differences

Our second research question sought to determine what child- and language-level factors would predict syntactic performance at this early stage of bilingualism and whether both languages and scores (overall accuracy and syntactic accuracy) would be predicted by a similar group of factors.

HL Arabic vs. L2 English

Our initial hypotheses were partially borne out: fewer child-level factors influenced performance on the Arabic than on the English SRT. However, contrary to our initial predictions, syntactic complexity only impacted performance in Arabic.

Starting with child-level factors, it is important to reiterate that English AOA and chronological age were correlated almost perfectly in this sample, as they would be in any sample that immigrated as a cohort. For this reason, the modelling only included AOA and the effect of this variable could be potentially confounded with the effects of age. Specifically, we found that an older English AOA predicted better performance for overall accuracy in both languages. We propose that this effect is due to different reasons. For the HL, AOA is not only a measure of cognitive maturity and skills; it is also a measure of cumulative Arabic input and quality of this input: children who started acquiring English later in life were immersed for longer in a rich Arabic-speaking environment where they received input from a variety of speakers and sources. These results are consistent with the previous literature: a later L2 onset allows bilingual children to develop and retain strong abilities in their HL (Albirini, 2018; Montrul, 2008, 2016). In the case of the L2, the positive effect of older AOA may be related to greater cognitive maturity associated with older age (Chondrogianni, 2018). As mentioned in the introduction, this is one of the few studies to investigate AOA as a continuous variable with regards to syntactic performance in the L2 of sequential bilinguals, instead of comparing sequential bilinguals' performance to that of simultaneous bilinguals (cf. Chiat et al., 2013; Kaltsa et al., 2020). These results point to an advantage of an older AOA in sequential bilingual children at least at the early stages of bilingualism. It remains to be seen whether this advantage remains with longer exposure to the L2.

Length of Arabic schooling was a significant predictor of Arabic syntactic accuracy. As explained in the Results, length of Arabic schooling was moderately correlated with age and English AOA. The fact that it was Arabic schooling, and not English AOA, which remained in the optimal model for syntactic accuracy indicates that it was not just the quantity of cumulative input pre-migration that influenced syntactic performance, but also the high quality of this exposure – namely, exposure at school. Since exposure at school is to MSA rather than to Syrian Arabic, this finding also suggests some interdependence of syntactic skills between different varieties of a language; in other words, that rich input in MSA can support syntactic development in Syrian Arabic. This possibility merits further research.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Length of schooling in English was used as a proxy for length of English exposure in the present study. In previous research, LOE has been found to be a positive predictor of syntactic development of the L2 in diverse languages for different syntactic phenomena, though the SRT literature has provided conflicting findings (Armon-Lotem et al., 2011; Chiat et al., 2013). Our study found unequivocal evidence that LOE is a positive predictor of L2 performance both in terms of overall accuracy and syntactic accuracy.

HL/L2 relative input in the home was only measured through sibling interaction, since at this early stage of bilingualism, parents used Arabic most of the time and the limited range did not allow for the inclusion of this variable in the modelling. As reviewed in the introduction, a positive effect of current input in the home has been reported consistently for the production/comprehension of complex syntax for the HL (Daskalaki et al., 2019; R. Jia & Paradis, 2020) but not for the L2 (Kaltsa et al., 2020; Paradis et al., 2017). Our study found the opposite results: for Syrian refugee children, using more English (and less Arabic) with siblings did not affect the performance in HL Arabic but it was a positive predictor for overall accuracy and syntactic accuracy in L2 English. This discrepancy between the results of the present study and those of previous HL/L2 studies is explained when the nature of the participant samples and methodologies is considered. For the HL, Daskalaki and colleagues (2019) and R. Jia and Paradis (2020) included children who had, in their majority, been born and raised in the host country. On the other hand, the sample of the present study was born and raised in Syria and had short residency in the host country at the time of testing. As suggested above, at this early stage of bilingualism, factors that account for cumulative input (i.e., L2 AOA and Arabic schooling) influence HL performance more strongly than fluctuations of the current input in the HL. It is likely, however, that compounded long-term effects of current lack of input quantity/quality become apparent as participants become older (Albirini, 2014). That is, the effects of (lack of) input may not have an immediate effect on the HL but may accumulate in later years. The discrepancy between our findings regarding the positive effect of English use with siblings for the L2 and the null results in Kaltsa et al. (2020) and Paradis et al. (2017) may be due to methodological differences: these two studies provided measures of current language use which considered parents and siblings in combination (and in the case of Kaltsa's study, together with other measures of language preference). Our study suggests that considering input providers independently (parents vs. siblings) may provide important insights on the relationship between syntactic skills and home language use. In this population, whose parents may be underemployed and less likely to acquire L2 proficiency in the work setting, siblings may be key in children's bilingual development by virtue of being exposed to high quality English input at school. Alternatively, since directionality cannot be assumed from the modelling employed in this study, it is possible that children with higher L2 abilities use this language more often with their siblings.

This study has contributed to the limited body of research that has addressed the effects of language environment richness on the development of HL and L2 syntax. For the L2, richness of the English environment was a positive predictor both in terms of overall accuracy and syntactic accuracy, in line with Kaltsa et al. (2020) and Paradis et al. (2017). For the HL, Arabic richness was not a significant predictor, suggesting again that at this stage of bilingualism, it is the cumulative linguistic experiences that matter for syntactic development.

Together with Armon-Lotem et al. (2011), this was one of the few studies to investigate the effect of SES, using maternal and paternal education as a proxy, on

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

participants' syntactic performance on the HL and L2. The results of our study partially replicate those of Armon-Lotem and colleagues. First, in line with the previous study, higher maternal education predicted better performance on the L2 for the current participant sample, both for overall accuracy and syntactic accuracy. In addition, paternal education positively predicted syntactic accuracy in the L2. As mentioned in the introduction, refugee fathers may be underemployed and, as a result, may spend more time in the home, and thus influence the linguistic development of the children. It is possible that more educated parents are more proficient in the L2, as previous studies have suggested (Armon-Lotem et al., 2011; Sorenson Duncan & Paradis, 2020). However, the relationship between SES and L2 performance in our sample is difficult to pinpoint since parents mostly use Arabic when speaking to their children. Unlike Armon-Lotem et al. (2011), we did not find an effect of SES on the HL. This contrast may be due to differences in the participant samples. Children in Armon-Lotem et al. (2011) were younger ($M = 5.48$) than children in this study ($M = 9.37$) and had a younger AOA of the L2 ($M = 2.36$) than the children in this study ($M = 7.64$). We speculate that the effect of SES on the HL may be stronger during HL development, as syntax is acquired. In this sense, our participants have had more time for the HL to stabilize than those in Armon-Lotem et al. (2011). It is possible, nevertheless, that SES plays a more important role again after prolonged exposure to the L2, when the HL faces the possibility of attrition.

This was the first study to investigate the effects of non-verbal cognitive skills on the development of HL syntax. We found that this child-internal factor affected HL overall accuracy but not syntactic accuracy, a finding to which we return below. In line with Paradis et al. (2017), we found that cognitive skills positively predicted L2 syntactic skills, both for overall accuracy and syntactic accuracy.

With respect to language-level variables, regression analyses showed that syntactic structure was a significant predictor for overall accuracy and syntactic accuracy in Arabic but not in English. Its effect for Arabic was further unpacked with post-hoc contrasts, which found that object relatives, together with topicalizations, had been preserved significantly less often than other structures. Relatives and topicalizations pose a high level of syntactic complexity as they integrate a long-distance dependency between a fronted object and an object clitic that need to agree in morphological features. This is a dependency that has been shown to be challenging for bilingual heritage speakers and monolingual speakers of Arabic (Albirini, 2018). In the case of English, on the other hand, there was no effect of structure. Given the overall lower performance in the English SRT, the null effect of syntactic structure could be due to lack of sensitivity to L2 syntactic complexity at this early stage of acquisition. That is, participants' lack of access to the abstract syntactic representations in their L2 may have triggered a reliance, rather, on surface-level information (i.e., vocabulary). This requires further study since null results, such as this one, can only be interpreted cautiously.

Overall accuracy vs. syntactic accuracy

We finally consider the contrast between the two outcome scores in terms of the factors that predicted them. Overall, we found that syntactic performance is sensitive to individual and environmental factors, in line with what has been found for vocabulary and morphology. Both overall accuracy and syntactic accuracy were indeed predicted by a similar cluster of child-external and child-internal variables.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

We address here the two discrepancies between the two outcome scores that require future attention. First, non-verbal analytical skills predicted overall accuracy but not syntactic accuracy in the HL, while they predicted both measures in the L2. We hypothesize that this difference is due to the fact that the target syntactic structures have mostly stabilized in participants' HL grammar, given their age and late L2 AOA. In the case of the developing L2, on the other hand, the effect of these cognitive abilities may be due to two reasons. (1) Cognitive abilities might be predictive of children's ability to recall the target structure. That is, participants may rely on cognitive skills to recall the syntactic structure in the system that is developing (L2) but not in the system that has stabilized (HL). (2) Cognitive skills may be predictive of children's ability to acquire the L2 syntax. In this sense, non-verbal analytical skills may aid bilinguals not in RECALLING the target structure directly but in ACQUIRING the L2 syntax. Different theories of language acquisition make different predictions regarding the involvement of domain-general cognitive skills in the acquisition of syntax and the design of this study does not allow us to probe further into this question.

The second discrepancy in the contrast between the two outcome scores occurred in English. English AOA was predictive of overall accuracy but not of syntactic accuracy in the L2. This may be due to a methodological artefact: given the moderate correlation between AOA and non-verbal analytical skills ($r = .44$, $p < .001$), the latter factor could have suppressed the effect of English AOA in the optimal model for syntactic accuracy.

Implications

Overall, this study complements the existing research on individual differences in three main respects. First, it demonstrates that HL and L2 syntax is affected by a number of child-level (input and cognitive skills) and language-level (syntactic structure) factors. Second, our results show that the relative effect of each of these factors may differ depending on the language (HL vs. L2) and the bilingual's stage in development. In this regard, we found that at the initial stages of bilingual development, HL syntax appears more sensitive to cumulative measures of input (AOA, HL schooling) than to measures of current input (relative input, richness), whereas the L2 is influenced by both current and cumulative measures. The finding that the HL is more sensitive to cumulative linguistic experiences at the onset of bilingualism should not be taken as a reason to encourage language shift in the home. The effects of limited current HL input will most likely add up after prolonged exposure to the L2, resulting in attrition of the previously stabilized HL system. Therefore, the use of the HL in the home is key in allowing these developing bilinguals to receive sustained input in the minority language. Finally, we found evidence of interdependence between the HL and L2 syntactic systems of the bilingual: better performance in the HL was related to better performance in the L2. While the origin of this interdependence is difficult to establish, the L2 English syntax of these Syrian refugee children may take years to reach the proficiency of their HL, and therefore, providing opportunities for their HL to develop may be beneficial for their overall linguistic development in the two languages. That is, HL proficiency, far from undermining L2 proficiency, should be considered a useful tool for L2 development. This may be especially important for younger children, who, due to historical and social circumstances, have had interrupted schooling, shorter time in the home country, and a younger onset of bilingualism.

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Acknowledgements. We would like to first thank the families for taking the time and effort to participate in this research. We also acknowledge the important contribution of the student assistants in all three cities (Edmonton, Waterloo and Toronto) who collected the data. We also thank Aisha Barise, Zahraa Attar and Lina Abed Ibrahim for their contributions in the IPA transliteration and glossing of the SRT Items as well as for providing Syrian Arabic translations and judgements for our syntactic section. We would also like to thank Aseel Dalal for her role as a primary consultant for the Syrian community. This research was funded by the Social Sciences and Humanities Research Council of Canada (Partnership Grant and Insight Developmental Grant; Chen and Paradis), for which we are grateful.

References

Albirini, A. (2014). Toward understanding the variability in the language proficiencies of Arabic heritage speakers. *International Journal of Bilingualism*, *18*(6). 730–765.

Albirini, A. (2016). *Modern Arabic Sociolinguistics. Diglossia, variation, codeswitching, attitudes and identity*. New York: Routledge.

Albirini, A. (2018). The Role of Age of Exposure to English in the Development of Arabic as a Heritage Language in the United States. *Language Acquisition*, *25* (2). 178–196.

Aoun, E. J., Benmamoun, E., & Choueiri, L. (2010). *The Syntax of Arabic*. Cambridge, UK: Cambridge University Press.

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. 193–211). Cambridge University Press.

Armon-Lotem, S., Walters, J., & Gagarina, N. (2011). The impact of internal and external factors on linguistic performance in the home language and in L2 among Russian-Hebrew and Russian-German preschool children. *Linguistic Approaches to Bilingualism*, *1*(3), 291–317.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7.

Bayram, F., Rothman, J., Iverson, M., Kupisch, T., Miller, D., Puig-Mayenco, E., & Westergaard, M. (2017). Differences in use without deficiencies in competence: passives in the Turkish and German of Turkish heritage speakers in Germany. *International Journal of Bilingual Education and Bilingualism*. 1–21.

Brustad, E. K. (2000). *The syntax of spoken Arabic. A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, DC: Georgetown University Press.

Castilla, A. P., Restrepo, M. A., & 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.

Chiat, S., Armon-Lotem, S., Marinis, T., Polišenská, K., Roy, P., Seeff-Gabriel, B., & Gathercole, V. C. M. (2013). Assessment of language abilities in sequential bilingual children: the potential of sentence imitation tasks. *Issues in the assessment of bilinguals*, 56–89. doi:10.21832/9781783090105-005

Chondrogianni, V. (2018). Child L2 acquisition. In: D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.) *Bilingual Cognition and Language. The state of the science across its subfields* (pp. 103–126). Amsterdam: John Benjamins.

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.

Cowell, W. M. (1964). *A Reference Grammar of Syrian Arabic*. Washington, DC: Georgetown University Press.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222–251.

Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge: Cambridge University Press.

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon; Multilingual Matters.

Daskalaki, E., Chondrogianni, V., Blom, E., Argryri, F., & Paradis, J. (2019). Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research 35*(3), 421–455.

El-Yasin, M. (1996). The Passive Voice: A Problem for the English-Arabic Translator. *Babel*, *42*(1), 18–26.

Flores, C., & Barbosa, P. (2014). When reduced input leads to delayed acquisition: a study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism*, *18*(3), 304–325.

Frizelle, P., & Fletcher, P. (2014). Relative clause constructions in children with specific language impairment. *International Journal of Language & Communication Disorders*, *49*(2), 255–264. https://doi.org/10.1111/1460-6984.12070

Gallimore, R., & Tharp, R. G. (1981). The interpretation of elicited sentence imitation in a standardized context. *Language Learning*, *31*(2), 369–392. https://doi.org/10.1111/j.1467-1770.1981.tb01390.x

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement*. R package version 0.84.1. https://CRAN.R-project.org/package=irr

Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morphosyntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 224–247.

Geers, A. E., & Moog, J. S. (1978). Syntactic maturity of spontaneous speech and elicited imitations of hearing-impaired children. *Journal of Speech and Hearing Disorders*, *43*(3), 380–391. https://doi.org/10.1044/jshd.4303.380

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

Hadfield, K., Ostrowski, A., & Ungar, M. (2017). What can we expect of the mental health and well-being of Syrian refugee children and adolescents in Canada?. *Canadian Psychology/Psychologie canadienne*, *58*(2), 194–201.

Hammer, C. S., Lawrence, F. R., & Miccio, A. W. (2007). Bilingual children's language abilities and early reading outcomes in Head Start and kindergarten. *Language, Speech, and Hearing Services in Schools*, *38*, 237–248.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, *1*(1), 77–89.

Hopp, H., Steinlen, A., Schelletter, C., & Piske, T. (2019). Syntactic development in early foreign language learning: Effects of L1 transfer, input, and individual factors. *Applied Psycholinguistics*, *40*(5), 1241–1267. doi:10.1017/S0142716419000249

Jia, R., & Paradis, J. (2020). The acquisition of Mandarin relative clauses by Mandarin heritage children. *Linguistic Approaches to Bilingualism*, *10*(2), 153–183.

Kaltsa, M., Prentza, A., & Tsimpli, I. M. (2020). Input and literacy effects in simultaneous and sequential bilinguals: The performance of Albanian–Greek-speaking children in sentence repetition. *International Journal of Bilingualism*, *24*(2), 159–183.

Kanu, Y. (2008). Educational needs and barriers for African refugee students in Manitoba. *Canadian Journal of Education*, *31*, 915–940.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and cognitive processes*, *22*(6), 860–897. https://doi.org/10.1080/01690960601155284

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.

Komeili, M., Marinis, T., Tavakoli, P., & Kazemi, Y. (2020). Sentence Repetition in Farsi-English Bilingual Children. *Journal of the European Second Language Association*, *4*(1), 1–12. DOI: https://doi.org/10.22599/jesla.55

Krippendorff, K. (2004). *Content Analysis, an Introduction to Its Methodology*. Thousand Oaks: Sage.

Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.2. https://CRAN.R-project.org/package=emmeans

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. In S. Armon-Lotem, J. de Jomg, & N. Meir (Eds.). *Methods for Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*. Multilingual Matters.

Meir, N., Walters, J., & Armon-Lotem, S. (2016). Disentangling SLI and bilingualism using sentence repetition tasks: The impact of L1 and L2 properties. *International Journal of Bilingualism*, *20*(4), 421–452.

Meir, N., Walters, J., & Armon-Lotem, S. (2017). Bi-directional cross-linguistic influence in bilingual Russian-Hebrew children. *Linguistic Approaches to Bilingualism*, *7*(5), 514–553. http://doi.org/10.1075/lab.15007.mei

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

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor* (Vol. 39). John Benjamins Publishing.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge, United Kingdom: Cambridge University Press.

Murphy, V. A. (2018). Literacy development in linguistically diverse pupils. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual Cognition and Language: The state of the science across its subfields* (Vol. 54), (pp. 315–323). Amsterdam: John Benjamins.

Naglieri, J. A. (1985). Matrix Analogies Test: Short Form (MAT-SF). Psychological Corporation.

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.

Paradis, J. (2019). English Second Language Acquisition from Early Childhood to Adulthood: The Role of Age, First Language, Cognitive and Input Factors. *BUCLD 43 Proceedings*.

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*, 1–20.

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.

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language & Communication Disorders*, *50*(1), 106–118.

Roesch, A. D., & 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.

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In *Handbook of Arabic Literacy* (pp. 3–28). Springer, Dordrecht.

Sorenson Duncan, T., & Paradis, J. (2020). How does maternal education influence the linguistic environment supporting bilingual language development in child L2 learners of English? *International Journal of Bilingualism*, *24*(1), 46–61.

Tabors, P. O., Paez, M., & López, L. M. (2003). Dual language abilities of bilingual four-year olds: Initial findings from the early childhood study of language and literacy development of Spanish-speaking children. *NABE Journal of Research and Practice*, *1*(1), 70–91.

Tsimpli, I. M. (2014). Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, *4*(3), 283–313.

Uchikoshi, Y. (2006). English vocabulary development in bilingual kindergarteners: What are the best predictors?. *Bilingualism: Language and Cognition*, *9*(1), 33–49.

Unsworth, S. (2016a). Quantity and quality of language input in bilingual language development. In: E. Nicoladis & S. Montanari (Eds.) *Lifespan perspectives on bilingualism: factors moderating language proficiency* (pp. 136–196). Washington, DC: American Psychological Association.

Unsworth, S. (2016b). Early child L2 acquisition: Age or input effects? Neither, or both? *Journal of Child Language*, *43*(3), 608–634

Unsworth, S. (2017). Quantifying language experience in HL development. In M. Schmid, & B. Köpke (Eds.), *The Oxford handbook of first language attrition*. Oxford: Oxford University Press.

Verhoeven, L. T. (1994). Transfer in bilingual development: The linguistic interdependence hypothesis revisited. *Language Learning*, *44*(3), 381–415.

Appendix A. Arabic SRT

Structure	Example
Declarative ($k = 6$)	*ʕœm yœ-kl-u l-mɔːz bə-l-jneːne* PROG 3-eat-PL the-bananas in-the-park 'They are eating the bananas in the park.' *rfiːʔ-i byə-ʔder yə-rfœʕ hœ-l-kersi l-kbiːr* friend-my 3SG.M-can 3SG.M-lift this-the-chair the-big 'My friend can lift this big chair.' *lœzem t-ħɒtˁ l-kœːse ʕœ-tˁ-tˁawle* must 3SG.F-put the-cup on-the-table 'She must put the glass on the table.' *mœ bt-eħsen l-ʔœnse t-jiːb l-ktœːb ʕœ-sˁ-sˁaf* NEG 3SG.F-can the-teacher 3SG.F-bring the-book to-the-class 'The teacher cannot bring the book to the class.' *mœ ʃereb l-wœlœd ħœliːb səxən əs-sɒbəħ* NEG drank the-child milk hot the-morning 'The child did not drink hot milk in the morning.' *mœ ʔœreː-t ʔesˁa ħelwe bə-l-mœdrœse* NEG read-1SG story beautiful in-the-school 'I did not read a beautiful story at school.'
Short passive ($k = 3$)	*ət-œxœd–u l-wlœːd ʕœ-l-mœktœb* PASS-taken-3PL the-children to-the-office 'The children were taken to the office.' *n-dœfœʃ b-ʔuwe ʕœ-l-ʔardˁ* PASS-pushed in-hard to-the-ground 'He was pushed hard against the ground.' *t-xœːlaf-et ʕœ-l-iʃœːra l-ħœmra* PASS-fined-3SG.F on-the-light the-red 'She got a fine at the red light.'
Topicalizations ($k = 3$)	*l-ʔm leħʔ-a esˁ-sˁabi ʕœ-ʃ-ʃaːreʕ* the-mother followed-her the-boy to-the-street 'As for the mother, the boy followed her to the street.' *l-baʔara, l-ħmaːr dˁarab-œ barra* the-cow the-donkey hit-her outside 'As for the cow, the donkey hit her outside.' *l-ʔœb ed-doktɔːr fœħasˁ-ɔ əs-sɒbəħ* the-father the-doctor examined-him the-morning 'As for the father, the doctor examined him in the morning.'
Question ($k = 6$)	*miːn əlli ʃœf-ət- ɔ l-bənt* who that saw-3SG.F-him the-girl *bə-l-mœktœbe mbœːreħ* in-the-library yesterday? 'Who did the girl see [him] in the library yesterday?'

(*Continued*)

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

(*Continued.*)

Structure	Example
	miːn ʃœːf-u jœnb l-bœːb l-ʔœmœːmiː who have.seen-3PL near the-door the-front 'Who have they seen near the front door?'
	ʃuː ʃtœra l-ʔœb eʃ-ʃaher l-mədˤi what bought the-father the-month the-last 'What did the father buy last month?'
	ʃuː tˤabx-et l-ʔm l-mœsœ what cooked-3SG.F the-mother the-evening 'What did the mother cook in the evening?'
	ʔyyœ ʃaraːb kœb l-jaːr bə-l-beːt which drink spilled the-neighbor in-the-house 'Which drink did the neighbor spill in the house?'
	ʔyyœ sˤuːra rasœm-œ bə-l-beːt mbœreħ which picture drew-her in-the-house yesterday 'Which picture did he draw [it] at home yesterday?'
Coordinated ($k = 3$)	*l-kœlb ʕœm y-ʕaww-i barra w* the-dog PROG 3SG.M-bark-SG outside and *l-wœlœd ʕœm y-ibk-i juwwa* the-child PROG 3SG.M-cry-SG inside 'The dog is barking outside and the child is crying inside.' *l-ʔm ʕœm tət-sœwwaʔ w* the-mother PROG 3SG.F-shop and *l-wœlœd ʕœm yi-drɔs bə-l-beːt* the-boy PROG 3SG.M-study in-the-home 'The mother is shopping and the boy is studying at home.' *jaːr-nœ ʕœm y-inœdˤef es-sœyyœra w* neighbor-our PROG 3SG.M-clean the-car and *əbn-ɔ ʕœm yə-lʕab bə-tˤ-tˤ aːbe* son-his PROG 3SG.M-play with-the-ball 'Our neighbor is cleaning the car and his son is playing with the ball.'
Subordinate ($k = 4$)	*fetˤer l-wœlœd bœʕd-mœː ɣassœl wəʃʃ- ɔ* had.breakfast the-child after washed face-his 'The child had breakfast after he washed his face.' *ʔabəl-mœː tə-tʕaʃʃœ l-bənt rɒħ* before 3SG .F-has.dinner the-girl FUT *tə-lʕab ʕœ-l-kɔmbiuter* 3SG.F-play on-the-computer 'Before the girl has dinner, she will play with the computer.' *rɒħ y-aːxd-u l-wlœːd hdiːə* FUT 3-get-PL the-children gift *izœ nadˤdˤaf-u l-beːt* if clean-3PL the-house 'The children will get a present if they clean the house.'

(*Continued*)

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

(*Continued.*)

Structure	Example
	izœ kœ:n eʒ-ʒɑu d œ:fi if was the-weather warm *fi:-nœ nə-tˤlaʕ ʕœ-l-jnene* can-1PL 1PL-go.out to-the-park 'If the weather is warm, we can go to the park.'
Relative ($k = 6$)	*ənbasatˤ-u l-wlœ:d bə-ʃ-ʃɔkɔlata* enjoyed-3PL the-children with-the-chocolate *əlli ʔœkœl-u-wa* that ate-3PL-her 'The children enjoyed the chocolate that they ate [it].' *l-ʔm sœw-et etˤ-tˤabxa yœlli* the-mother made-3SG.F the-meal that *ʕœm y-œ:kl-u-wa l-wlœ:d* PROG 3-eat-PL-her the-children 'The mother made the meal that the children are eating [it].' *lœ:zem y-ħœməm-u ez-ˤɣi:r əlli* should 3-wash-PL the-baby that *ʔmm-ɔ ʕœm ət-tˤaʃm-i:* mother-his PROG 3SG.F-feed-him 'They should wash the baby that his mother is feeding [him].' *l-ħsˤa:n əlli l-fœllœħ dˤarab* the-horse that the-farmer hit *rafœs- ɔ ʕœ-dˤahr- ɔ* kicked–him on-back-his 'The horse that the farmer hit kicked him on his back.' *esˤ-sˤabi əlli sœ:ʕœd- ɔ l-ja:r* the-boy that help-him the-neighbor *dˤayyœʕ tˤari: ʔ -ɔ* lost way-his 'The boy that the neighbor helped [him] has lost his way.' *fœ:z l-fari:ʔ əlli ʃœjœʕ- ɔ* won the-team that encouraged-him *œx-i bə-s-siba:ʔ* brother-my in-the-race 'The team that my brother cheered for [it] won the race.'

https://doi.org/10.1017/S0305000921000246 Published online by Cambridge University Press

Appendix B. English SRT

Structure	Example
Declarative ($k = 6$)	She can bring the glass to the table. They have been rid-ing the bicycle around the backyard. They are eat-ing the banana-s in the park. The kitten could have bounc-ed the ball down the stair-s. The boy must sweep the floor in the kitchen. The teacher has been look-ing at us all day.
Short passive ($k = 3$)	She was stopp-ed at the big red light-s. The children were tak-en to the office. He was push-ed hard against the ground.
Long passive ($k = 3$)	The cow was kick-ed in the leg by the donkey. She was se-en by the doctor in the morning. The mother was follow-ed by the girl.
Question ($k = 6$)	What did the mother cook in the evening? Who have they se-en 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 shopp-ing and the child is study-ing at home. The dog bark-s outside and the child crie-s inside. Our neighbor clean-s the car and his son play-s basketball.
Subordinate ($k = 4$)	If the weather is warm, we can go to the park. Before the girl eat-s dinner, she will play with the computer. The children will get a present if they clean the house. The child ate breakfast after he wash-ed his face.
Relative ($k = 6$)	The boy that the neighbour help-ed has lost his way. They should wash the baby that the mother is feed-ing. The horse that the farmer push-ed kick-ed him in the back. The mother made the meal that the children are eat-ing. The children enjoy-ed the candy that they tast-ed. The team that my brother cheer-ed for won the race.

Cite this article: Soto-Corominas A, Daskalaki E, Paradis J, Winters-Difani M, Janaideh RA (2022). Sources of variation at the onset of bilingualism: The differential effect of input factors, AOA, and cognitive skills on HL Arabic and L2 English syntax. *Journal of Child Language* **49**, 741–773. https://doi.org/10.1017/S0305000921000246