Annual Review of Applied Linguistics, page 1of 20.

doi: 10.1017/S0267190517000022

Children's Second Language Acquisition of English Complex Syntax: The Role of Age, Input, and Cognitive Factors

Johanne Paradis
University of Alberta,
Johanne.Paradis@ualberta.ca

Brian Rusk
University of Alberta

Tamara Sorenson Duncan
University of Alberta

Krithika Govindarajan
University of Alberta

ABSTRACT

The goal of this study was to determine (a) the similarities and dissimilarities between child L2 and L1 acquisition of complex sentences and (b) the individual difference factors predicting L2 children's acquisition of complex sentences. We analyzed language samples from 187 English L2 children with diverse L1s (Age_{mean} = 5;10 [years;months]; English $exposure_{mean}$ = 17 months). Children used various types of complex sentences at all levels of L2 exposure, including sentences with relative clauses, which are late-acquired by L1 learners. Mixed logistic regression modeling revealed that longer exposure to English in school, richer English environments outside school, larger L2 vocabulary, superior verbal memory and visual analytic reasoning contributed to greater use of complex sentences. L1 typology did not impact complex sentence use in the L2. Overall, L2 children used more complex sentences within a few months of English L2 exposure than what is reported for L1 children aged 2;0–4;0, revealing an advantage for an older age of acquisition. The predictive role of input and cognitive factors, as well as vocabulary, in children's use of complex sentences is more consistent with constructivist than generativist accounts of L2 syntactic acquisition.

INTRODUCTION

Research on complex sentences in English first language (L1) acquisition shows that input factors and cognitive or conceptual maturity play a role in development (Diessel, 2004; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Vasilyeva, Waterfall, & Huttenlocher, 2008). This is significant because, while vocabulary acquisition is expected to be largely input driven and sensitive to cognitive factors across theoretical frameworks, generative and constructivist theories differ regarding the expected role of input and cognitive factors in the acquisition of syntax (Vasilyeva et al., 2008). Regarding child second language (cL2) acquisition, morphological acquisition has received more attention than the development of syntax. Furthermore, existing studies on cL2 complex syntax examine comprehension rather than production (e.g., Unsworth, 2016). Consequently, the developmental patterns and mechanisms of syntactic production in cL2 learners of English are largely unknown. The goal of this study was to examine complex sentence production in cL2 learners (cL2ers) of English using parallel methods to existing L1 acquisition studies.

The outcomes of this study have both theoretical and applied relevance. First, comparisons between L1 and cL2 acquisition of complex syntax would contribute to debates about the role of age in cL2 acquisition. On one hand, older age of acquisition is often construed as a limiting factor in ultimate attainment for L2 morphosyntax (Abrahamsson & Hyltenstam, 2009; Paradis, Tulpar, & Arppe, 2016; Unsworth, 2016). On the other hand, an older age of acquisition onset means children are more cognitively and linguistically mature when they are learning a target construction, and research shows that older age of acquisition within early childhood can result in a faster rate of L2 morphosyntactic development (Blom & Bosma, 2016; Paradis, 2011; Paradis & Jia, 2016; Paradis, Rice, Crago, & Marquis, 2008). Second, research on cL2 acquisition of complex syntax would contribute to theoretical accounts regarding the mechanisms underlying cL2 acquisition. Unsworth (2016) argued that different learning mechanisms, including the role of input factors, underlie Dutch L2 children's acquisition of morphology versus syntax-semantics, taking a generative perspective. In contrast, a constructivist perspective would predict parallel learning mechanisms for L2 vocabulary, morphology, and syntax (Paradis, 2011; Paradis & Jia, 2016). Finally, research on cL2 acquisition of complex syntax would be relevant to educational concerns in countries like Canada, the United States, and Britain, where there are a sizable number of English L2 learners from migrant families in schools (Murphy & Evangelou, 2016). English cL2ers face the challenge of learning the language of instruction at the same time as they are learning academic content. English cL2ers' abilities with complex syntax could be especially relevant to their academic performance, given that complex sentences are more frequent in written than oral language (Dabrowska, & Street, 2006; O'Donnell, 1974; Purcell-Gates, 2001).

Complex Sentences in English

In the most essential sense, complex sentences are sentences containing more than one clause. There are four common types of clauses in English complex sentences: coordinated, complement, adverbial, and relative. Coordinated sentences contain two independent clauses, for example, *she went to a movie, and she ate popcorn*. By contrast, other complex sentences involve a matrix or main clause and at least one dependent clause. In the following examples, the matrix clauses are underlined. To illustrate these components, consider a sentence such as, *<u>he said</u> that he didn't*

TABLE 1. *Examples of Complex Sentence Types in English*

Example	Complex sentence type	
[[She went to a movie] and [she ate popcorn]]$_{\mathbf{M}}$	Coordinated	
[He said [that he didn't go to a movie last night]$_{CC}$]$_{\mathbf{M}}$	Complement	Finite
[He wanted [to go to a movie last night]$_{CC}$]$_{\mathbf{M}}$		Nonfinite
[She went to a café [after she saw the movie last night]$_{AC}$]$_{\mathbf{M}}$	Adverbial	Finite
[She went to a café [after seeing a movie last night]$_{AC}$]$_{\mathbf{M}}$		Nonfinite
[He saw a man [who was dressed as Spiderman]$_{RC}$ at the Spiderman movie]$_{\mathbf{M}}$	Relative	Finite
[He saw a man [wearing a Spiderman costume]$_{RC}$ at the Spiderman movie]$_{\mathbf{M}}$		Nonfinite

Note. Brackets mark clauses. Boldface brackets with subscript M = matrix clause. Subscripts CC = complement clause, AC = adverbial clause, RC = relative clause

go to a movie last night. In this instance, "*he said*" is the matrix clause, where the argument structure of the verb "*said*" subcategorizes for a clausal complement, and thus, "*he didn't go to a movie last night*" is the dependent complement clause. In contrast, the other types of dependent clauses provide extra information, that is, modification. There are two main types of modifying dependent clauses: adverbials and relatives. Adverbials, often introduced by an adverb, modify the matrix clause, for example, *she went to a café after she saw the movie last night*. In contrast, relative clauses modify a noun or noun phrase in a sentence, for example, *he saw a man who was dressed as Spiderman at the Spiderman movie* . In this sentence, the dependent clause modifies the noun phrase "*a man*," and not the full matrix clause, as is the case for adverbials.

In addition to this variation in function, dependent clauses can vary in form. For example, dependent clauses can be finite or nonfinite, with nonfinite clauses consisting of infinitival or gerund constructions. Compare *she went to a café after she saw the movie last night* with *she went to a café after seeing a movie last night*. Furthermore, some connectives, like relative pronouns *who* and *that*, can be omitted grammatically in certain constructions. Compare *he said that he didn't go to a movie last night* with *he said ∅ he didn't go to a movie last night.* However, the infinitival *to* cannot be omitted, for example, **he wanted ∅ go to a movie last night*. A summary of these complex sentence types is given in Table 1. (For more details on English complex syntax, see Diessel, 2004; Huddleston & Pullum, 2002; Quirk, Greenbaum, Leech, & Svartvik, 1985.)

Complex Sentences in English L1 Acquisition

Diessel (2004), Vasilyeva et al. (2008), and Huttenlocher et al. (2010) carried out longitudinal, corpus-based studies with transcribed language production samples from English-learning children that also included the speech of their caregivers. Complex sentences emerged in children's speech around 2;2 (years;months) on average (Diessel, 2004; Huttenlocher et al., 2010; Vasilyeva et al., 2008), and on an individual basis, children started to use complex sentences about one year after they used simple sentences or word combinations in spontaneous speech (Vasilyeva et al., 2008). The proportion of all sentences that were complex was small in children's speech across the preschool years, as Diessel (2004) and Vasilyeva et al. (2008) found that complex sentences were less than 10% of all sentences from ages 2;0–4;0, with use of complex sentences increasing from a minimum of 1% at emergence to an average of 10% to 14% of all sentences by age 4;0. Furthermore, Diessel argued that some early emerging complex sentences are likely just prefabricated formulas since children's cognitive abilities and schematization in syntactic representation are underdeveloped early on. Different types of complex sentences emerged at different ages. The first to emerge were complement clauses, followed by coordinated clauses, adverbial clauses, and then relative clauses (Diessel, 2004; Vasilyeva et al., 2008). Importantly, this order of emergence also corresponds to the relative frequency of complex sentence types in children's speech, that is, at all ages, the frequency of complement clauses is higher than relative clauses (Vasilyeva et al., 2008). This relative frequency also corresponds to the distribution in caregiver speech (Diessel, 2004). Thus, one question researchers have asked is whether the relative frequency of use and the order of emergence of complex sentences in L1 learners is driven by the complexity of the syntactic constructions or by input frequency. For example, relative clauses consist of a clause embedded within a noun phrase within a larger matrix sentence. Complement clauses are embedded within the matrix sentence, following the main verb as part of its argument structure. From the perspective of both the construction schema and the semantic processing requirements, sentences with relatives are seen to be more complex than sentences with complement clauses. Complexity of construction as an explanatory factor assumes that cognitive maturity, codeveloping with language in L1 learners, could limit early and frequent use of complex sentences in general and more difficult complex sentences in particular, that is, relative clauses. This could mean that cL2ers might be less constrained and would use more complex sentences earlier on, given that they are, by definition, older (and thus more cognitively mature) when they are exposed to English input. Furthermore, if distributional frequency in the ambient language were determining which complex sentence types children used more often, then the relative proportion of complex sentence types used might be similar between L1 and cL2.

Aside from the distribution of complex sentence types in the ambient language, other input factors have been investigated with respect to individual differences in the L1 acquisition of complex sentences. For example, family socioeconomic status (SES) influences the number of complex sentences and the diversity of syntactic types in speech to children, and in turn, children from higher SES families use more and more diverse forms of complex sentences (Huttenlocher et al., 2010; Vasilyeva et al., 2008). Furthermore, the frequency of matrix clause frames and clause types used in caregiver speech predicted the use of those frames and types by individual children (Diessel, 2004; Huttenlocher et al., 2010). Therefore, since variation in the input determines individual differences in the rate of L1 acquisition of complex syntax in production, input factors could also be expected to drive individual differences in cL2 acquisition of complex syntax.

Factors Influencing cL2 Acquisition of Morphosyntax

Much recent research has focused on cL2 acquisition of verbal and nominal morphology. Findings show that cL2 morphological acquisition is protracted, meaning it can take years for children to approach the accuracy levels of their monolingual peers even when L2 acquisition onset is relatively young (Jia & Fuse, 2007; Paradis et al., 2016; Thomas, Williams, Jones, Davies, & Binks, 2014; Unsworth, 2016; Unsworth et al., 2014). Variations in the quantity and quality of L2 input and output that children experience influence their rate of L2 morphological acquisition (Armon-Lotem, Walters, & Gagarina, 2011; Bohman, Bedore, Peña, Mendez-Perez, & Gillam, 2010; Jia & Fuse, 2007; Paradis, 2011; Thomas et al., 2014; Unsworth et al., 2014). Child-internal linguistic factors, such as L1 typology and size of L2 vocabulary, also influence L2 morphological acquisition in children (Blom & Paradis, 2013; Blom, Paradis, & Sorenson Duncan, 2012; Paradis, 2011; Unsworth et al., 2014). Furthermore, nonlinguistic cognitive factors like age of acquisition (AOA [older is advantageous]), verbal memory and visual analytic reasoning impact cL2 morphological acquisition (Blom & Bosma, 2016; Paradis, 2011; Paradis et al., 2016).

Comparatively less research has examined individual difference factors in cL2 acquisition of complex syntax, and no study to date has examined complex sentences in production. Nevertheless, some studies reveal that cL2ers' comprehension of *wh-* questions, comprehension of passives, and grammaticality judgments of *that*-trace effects, can be influenced by length of L2 exposure, amount of L2 spoken at home, AOA (older is advantageous), SES, and maternal L2 fluency (Chondrogianni & Marinis, 2011; Gathercole, 2007; Roesch & Chondrogianni, 2016; Rothman et al., 2016). By contrast, other studies have found asymmetries between cL2 morphology and syntax with respect to cognitive and input factors. In a study comparing different subdomains of morphosyntax, Unsworth (2016) found that Dutch cL2 verbal morphology and verb placement in production were predicted by the amount of L2 exposure, but children's comprehension of sentences with object scrambling (syntax-semantics interface) was not. Similarly, Armon-Lotem et al. (2011) found that cL2 Hebrew verb morphology was affected by AOA and input factors, but complex syntax was not. Overall, the evidence for input and cognitive factors impacting cL2 acquisition is more robust for morphology than syntax.

Whether input factors differentially impact separate linguistic subdomains raises theoretical implications. We begin our discussion of these implications by outlining some key differences between generative and constructivist theories, drawing on Ambridge and Lieven (2011), Behrens (2008), and Eisenbeiss (2009). A generative perspective assumes that acquisition unfolds through an interaction between exposure to input and innate, domain-specific mechanisms: the principles and parameters of Universal Grammar. Linguistic representation is modularized in the generative approach, as well as the rest of cognition; therefore, computational components of language-like syntax are separate from the lexicon and, in turn, from other cognitive systems. A constructivist (or usage-based, emergentist) approach is nonnativist and thus assumes that acquisition is driven by interactions between input properties, communicative exchanges, and domain-general mechanisms like cognitive abilities (with a particular emphasis on the role of input properties). A constructivist approach assumes a nonmodular linguistic representation, where monomorphemic words, multimorphemic words, and syntactic constructions are stored in an interactive network, which interfaces with other cognitive systems. Critically, examining predictors of individual differences as an approach to uncovering mechanisms of language acquisition is common in research within the constructivist framework (Ambridge & Lieven, 2011), and this is the approach taken in this study.

Within a generative approach, some kind of "syntactic exceptionalism," with respect to the role of input in acquisition, is possible. For example, Unsworth (2016) argued that the syntactic interpretation rules underlying Dutch object scrambling are part of invariant Universal Grammar principles, therefore, would not be expected to be sensitive to amount of L2 exposure in acquisition. In contrast, Unsworth (2016) found a role for amount of L2 input in predicting verb morphology and verb placement (V2), the latter argued to be a case of input being needed to trigger the V2 parameter for Dutch. Furthermore, modular representation means that no tight relationship is expected between the acquisition of syntax and the size of the lexicon on a generative approach (for open-class words in particular), nor would a systematic role for nonlinguistic cognitive factors in syntactic acquisition be predicted. On a constructivist approach, "syntactic exceptionalism" would be unexpected because parallel domain-general mechanisms should underlie acquisition across different linguistic subdomains, and thus interrelationships between the lexicon, morphology, and syntax would be predicted in acquisition. For example, Paradis (2011) and Paradis and Jia (2016) found that both input and cognitive factors impacted cL2ers' rate of acquisition of English morphology, morphosyntax, and vocabulary. Blom et al. (2012) and Blom and Paradis (2013) found that a larger cL2 lexicon predicted greater accuracy with cL2 morphology.

The Present Study

This study is based on a corpus of language samples from English cL2ers. Language samples include spontaneous speech in play sessions and storytelling from a wordless picture book. We asked the following research questions:

1. Do cL2ers follow the same patterns of complex sentence acquisition as L1 children? We hypothesized that because cL2ers are more cognitively and linguistically mature than their younger monolingual peers, they might have earlier emergence and greater use of complex sentences, even with a few months of L2 exposure. We also hypothesized that cL2ers might produce all types of complex sentences early on, instead of the sequence of emergence shown in L1 acquisition, also due to

greater maturity. Testing these hypotheses will reveal whether older AOA confers advantages on young child learners, as found in other studies of cL2 acquisition.

2. Is cL2 acquisition of complex sentences influenced by input and cognitive factors? Is there an association between L2 vocabulary and syntax? We hypothesized that input, cognitive factors, and L2 vocabulary would predict individual differences in cL2ers' production of complex sentences, which is consistent with prior research on cL2 morphology. Analyses aimed at this hypothesis will yield evidence relevant to distinguishing between generative and constructivist approaches.

METHOD

Participants

The participants in this study were 187 English cL2 learners from migrant (immigrant and refugee) families residing in English-majority Canadian cities (Edmonton and Toronto). Parents were all foreign born and were L2 speakers of English. Children were both Canadian and foreign born, but as an inclusion criterion, children had to be exposed exclusively or primarily to their L1 before beginning to learn English at preschool or school. Children's mean age was 5;10, their mean L2 AOA was 4;4, and their mean length of exposure to English was 17 months (see Table 2 for full descriptives).

The children had diverse L1 backgrounds, which is consistent with the multilingual composition of Canadian cities (Statistics Canada, 2011): Arabic, Cantonese, Hindi, Mandarin, Punjabi, Spanish, and Urdu. Probing for L1 transfer in children's acquisition of complex syntax was not a goal of this study; however, we recognize the potential of L1 transfer in complex syntax (e.g., relative clauses: Kidd, Chen, & Chiu, 2015). In order to reduce confounds, our sample is structured such that there is no difference between the mean lengths of L2 exposure or mean ages among the L1 groups. This was achieved by starting with a sample of 220 and trimming it to 187, so that each L1 group would have equivalent mean exposure and age as determined by nonsignificant one-way ANOVAs. We also entered L1 as a predictor in our regression models to specify L1-based variance. We revisit the issue of L1 transfer in the "Discussion" section of this article.

This sample of children overlaps with—but is not identical to—samples in other studies (e.g., Blom & Paradis, 2013; Paradis, 2011). Importantly, no existing study has examined complex syntax or included the language production data from these children, as were analyzed in this study.

Procedures

Language Samples. Children provided two language samples and were tested either at home or at school by a graduate or senior undergraduate student in the Department of Linguistics. One was a 20-minute, recorded, spontaneous conversation sample, where the children played with the student assistant. The second was a recording of the child producing a narrative based on a wordless picture series from the Edmonton Narrative Norms Instrument (Schneider, Dubé & Hayward, 2005). Children were shown the pictures first, and then the book was held so the child could see it and the assistant could not. Children were then asked to tell the story to the assistant, and their story generation was recorded. Both the spontaneous sample and the narrative sample were transcribed using the CHAT system (MacWhinney, 2000; www.childes.psy.cmu.edu). Transcription interrater reliability was achieved as follows: A different student assistant retranscribed 10% of partial transcripts from a diverse sample of participants in terms of ages, L1s, and lengths of exposure to English ($n = 45$). Original and new transcriptions were then checked for word-for-word accuracy, yielding reliability scores of 88%–97%. Because it is possible that the more formal narrative task might yield more complex sentences than the spontaneous speech task (Elin Thordardottir, 2008), task type was entered as a predictor in regression models.

Language Sample Coding. Sentences were extracted from the language sample transcripts and entered into a dataframe. Sentences were defined as an utterance with at least one main verb or a be-copula. Children had a mean of 156 utterances ($SD = 71$) to contribute to our corpus. Next, sentences were coded for whether they were simple or complex. Complex sentences were those that had two or more clauses, signaled by at least two main verbs, or one main verb and one be-copula. Complex sentences were coded for type: coordinated clause (COR), complement clause (CC), adverbial clause (AC), relative clause (RC), and ambiguous (AMB). Dependent clauses could be finite or nonfinite; we did not differentiate between them in this study. Ambiguous sentences were those containing two main verbs, indicating that they were attempts at complex sentences, but omission of connectives, odd word choice, or other errors made them difficult to classify. For example, *he don't like to the sand is here to build the castle* (Spanish L1, 4 months English, age = 5;3) could be an attempt at a CC or perhaps an RC. Finally, morphological and word choice errors were ignored in this classification system, but complex sentences were coded for syntactic errors, that is, word order errors and ungrammatical substitution or omission of connectives (relative pronouns like *that* or *who* or infinitival *to*). Intercoder reliability was measured using a similar system as described above for transcription: 10% of the sample was recoded anew, and the results were compared with the original. This process revealed 98% agreement for simple versus complex, and for complex type, agreement ranged from 86% to 99%. Disagreements were settled by consensus.

Parent Questionnaire. Parents were given a questionnaire in the form of an interview, with interpreter if necessary, at home by a student assistant (Alberta Language Environment Questionnaire; ALEQ Paradis, 2011; https://www.ualberta.ca/linguistics/cheslcentre/questionnaires). Questions asked yielded numeric variables for the following input factors: length of L2 exposure (months in preschool or school), L2 use to child (from family members to child; 5-point rating scales for family members contributed to proportion score range of 0–1.0), child use of L2 (from child to family members; 5-point rating scales for family members contributed to proportion score range of 0–1.0), richness of L2 environment (outside school; points assigned to frequency of language-rich

TABLE 2. *Characteristics of English L2 Children*

Characteristic	Mean	SD	Description
Age at testing	5;10	0;07	Years; months
AOA	4;4	0;10	Years; months
Length of L2 exposure	17	9	Months in preschool/school
L2 use to child	.30	.21	Proportion range = 0–1.0
Child use of L2	.40	.26	Proportion range = 0–1.0
Richness of L2 environment	.59	.16	Proportion range = 0–1.0
Maternal education	14	3	Years
Maternal L2 fluency	2.2	1.1	5-point scale, range= 0–4
Verbal memory	7.4	1.9	Standard mean = 10, 1 SD = 3
Analytic reasoning	106	12	Standard mean = 100, 1 SD = 15
L2 vocabulary	86	16	Standard mean = 100, 1 SD = 15

activities like reading, playing with friends in the L2, etc., yielding a score range of 0–1.0), maternal education (years), and maternal L2 fluency (5-point rating scale). Questions also yielded the variables age at testing and AOA for each child. Descriptions for these variables are in Table 2. For more details about ALEQ administration and scoring, see Paradis (2011) and Paradis and Jia (2016).

Cognitive Tasks. Children were given a verbal, short-term memory task, nonword repetition, from the Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999). Children were asked to repeat words recorded on CD that increased in syllable length or difficulty. Their repetitions were recorded, then transcribed and scored according to the tester's manual, yielding standard scores. To measure visual analytic reasoning, children were given the Columbia Mental Maturity Scales (Burgemeister, Hollander Blum & Lorge, 1972), where they were presented with a series of shapes and had to detect the shape that did not match the others in the sequence. Responses were scored according to the tester's manual, yielding standard scores (Table 2). These tasks were used in Paradis (2011) to examine the role of cognitive factors (components of language aptitude) on cL2 morphological acquisition.

Vocabulary Task. Children were administered the Peabody Picture Vocabulary Test to estimate receptive vocabulary size (PPVT-III, Dunn & Dunn, 1997). Responses were scored according to the tester's manual, yielding standard scores (Table 2).

RESULTS

Characteristics and Distribution of Complex Sentences

The language samples comprised a corpus of 29,239 sentences, where 18% of all sentences were complex and 82% were simple. On an individual basis, children produced a mean of 17% (*SD* = 8%) complex sentences. Children's sentences

TABLE 3. *Examples of Complex Sentences Produced by English L2 Children*

Complex Type	Example
Coordinated	*take this knife and kill um Snow White.* PU, 6;10, exp2eng = 26
	she is not little but she is only big like this. AR, 5;0, exp2eng = 5
Complement	*she tried to fix it.* UR, 6;11, exp2eng = 26
	I'm sorry that I pull your balloon. UR, 7;0, exp2eng = 8
Adverbial	*then they were sad because they didn't have balloons.* SP, 5;11, exp2eng = 5
	when you do bad choices you get on the wall. SP, 5;0, exp2eng = 5
Relative	*sometimes I'd show my ship that I got from Santa Claus.* MA, 6;8, exp2eng = 26
	the library um there was a persons that read stories to us. SP, 5;1, exp2eng = 5
Ambiguous	*he pull the doctor take to him.* PU, 6;10, exp2eng = 26
	and then he go home eat. MA, 5;3, exp2eng = 8

Note. L1 abbreviations: AR = Arabic, CA = Cantonese, MA = Mandarin, PU = Punjabi, SP = Spanish, UR = Urdu. exp2eng = months of exposure to English in preschool/school

contained few of the syntactic errors we coded for. Just 4% of all complex sentences included errors with a connective, for example, *and sometime I go to the park play soccer* (age = 6;4, 16 months English, L1 = Arabic), and 1.6% contained odd word order, for example, *I just like (to) eat a grass those animals* [target: *I just like to eat those animals on the grass*] (age = 5;2, 27 months English, L1 = Mandarin).

Children produced all complex sentence types coded for at all levels of exposure to English. Example sentences are presented in Table 3, where, for each complex sentence type, an example from a child with relatively lower English exposure and an example from a child with relatively higher English exposure are given.

In order to understand if there was an order of emergence, as in L1 acquisition (e.g., complement >> coordinated >> adverbial >> relative), the proportion of complex sentence types produced at different intervals of English L2 exposure (low = < 12 months; medium = 13–24 months; high = > 25 months) are plotted in Figure 1. Visual inspection of Figure 1 indicates that the proportional distribution of complex sentence types follows this order, but that this order does not change substantially as a function of more exposure to English. Pearson correlations between children's mean proportion of complex sentence type and length of L2 exposure were significant for relative clauses ($r(181) = .23$, $p = 0.0022$); visible in Figure 1 as a slight increase from mean 0.051 proportion of complex sentences as RCs at low exposure to mean 0.091 proportion of complex sentences as RCs at high exposure. A significant correlation was also found for ambiguous sentences ($r(181) = -.22$, $p = 0.0033$), which decreased from low to high exposure: 0.047 to 0.015. No other significant correlations emerged. Therefore, proportional distribution of complex sentence types produced was fairly stable across length of exposure to English, and thus, this analysis revealed no order of emergence.

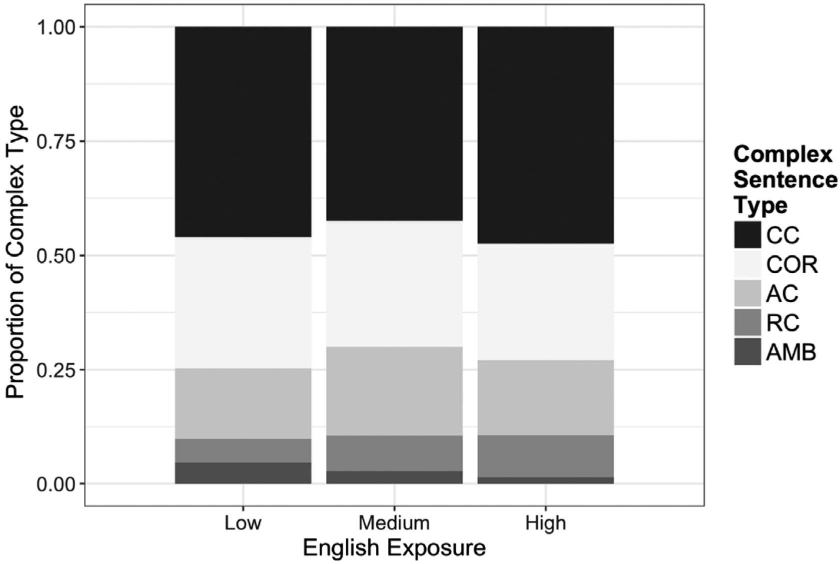

FIGURE 1. Proportion of complex sentence types produced by English L2 children at low, medium, and high exposure to English. CC = complement clause, COR = coordinated clause, AC = adverbial clause, RC = relative clause. AMB = ambiguous. Low exposure = < 12 months, medium exposure = 13–24 months, high exposure = > 25 months.

Factors Predicting Use of Complex Sentences

In order to understand which combination of factors predicted children's use of complex sentences, we conducted mixed logistic regression modeling, using glmer from the lme4 package in R Core (Bates, Maechler, & Bolker, 2015; R Core Team, 2015). The dependent variable was whether a sentence was complex (yes/no), and child and matrix verb lemma (item) were random factors. Fixed factors included the following: task (spontaneous or narrative), age at testing, AOA, length of L2 exposure, L2 use to child, child use of L2, richness of L2 environment, maternal education, maternal L2 fluency, verbal memory, analytic reasoning, L2 vocabulary, and L1. Because the fixed effects in these models varied in their ranges, all continuous fixed effects were scaled before modeling, that is, transformed to have more homogeneous ranges. Correlations between predictors were also inspected for issues of collinearity; there was one issue. Vocabulary and exposure were correlated with each other ($r(184) = 0.54$, $p = 1.11\text{e-}15$). As a consequence, a new residualized variable for vocabulary was created by predicting the variation in vocabulary using exposure to English (for details on this method, see Blom et al., 2012). The residualized and original variable were highly correlated ($r(174) = 0.87$, $p = 2.2\text{e-}16$), suggesting that the new variable accurately represents vocabulary scores while also being sufficiently independent from exposure. There were no other issues of collinearity.

TABLE 4. *Fixed Effects of Optimal Model Predicting English L2 Children's Use of Complex Sentences*

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	− 2.24732	0.09299	− 24.178	< 2e-16***
Task	0.21759	0.04024	5.408	6.38e-08***
Length of L2 exposure	0.22492	0.04193	5.364	8.14e-08***
Richness of L2 environment	0.08229	0.04130	1.992	0.04633*
Verbal memory	0.11279	0.04121	2.737	0.00619**
Analytic reasoning	0.07588	0.04102	1.849	0.06440†
L2 Vocabulary (Residual)	0.19636	0.04399	4.463	8.06e-06***

Note. For task, the storytelling task was the reference level. † = $p < 0.08$, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Nested model comparisons were conducted to determine the most parsimonious model—a model with the smallest number of fixed effects that is still significantly better than a reduced model ascertained by comparing deviances. This analysis was conducted on 28,715 sentences from 174 children—because some of the 187 children had missing data for one or more of the fixed effect factors, due to reasons such as a child refusing to complete a task or failure of the recording equipment.

Fixed effects that were not significant in any model were age at testing, AOA, L2 use to child, child use of L2, maternal education, maternal L2 fluency, and L1. Details of the fixed effects of the optimal model are in Table 4. This model provides a significantly better fit than a reduced model ($\chi^2 = 18.715$, $p = 1.518\text{e-}05$). Note that while the analytic reasoning fixed effect is only a trend in the optimal model, the model is a better fit with this variable; moreover, analytic reasoning scores do positively predict complex sentence use in a reduced model without vocabulary scores ($\beta = 0.12731$, std. error = 0.04213, $z = 3.022$, $p = 0.002510$). Each of the fixed effects are plotted in relation to proportion of complex sentence use for individual children in Figures 2 to 6. Full interpretation of how these fixed effects shape children's use of complex sentence is in the "Discussion" section.

DISCUSSION

The goal of this study was to examine English cL2 acquisition of complex sentences in production, which has not been studied previously. A corpus of language samples from 187 English L2 children was coded and analyzed to address two research questions. One research question asked if the patterns found for English L1 acquisition would be similar or dissimilar in cL2ers, and this was aimed at understanding the role of older age in syntactic acquisition. Another research question asked whether input and cognitive factors and L2 vocabulary would predict individual difference in cL2ers' use of complex sentences; this question was aimed at determining if children's acquisition patterns were consistent with a constructivist approach.

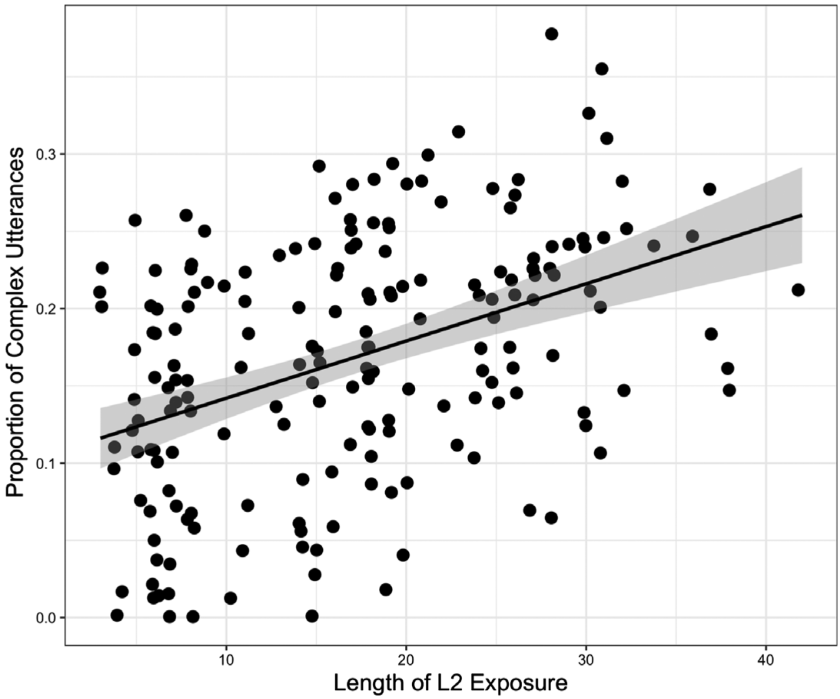

FIGURE 2. Scatterplot of proportion of complex sentence use by each child as a function of length of exposure to English in months.

cL2 Compared to L1 Acquisition of Complex Syntax: The Role of Age

Our analyses show some parallels in complex sentence acquisition between these young cL2ers and what has been reported for L1 learners. L1 children's use of complex sentences increases with age, and similarly, that of cL2ers increases with exposure to English (Table 4). Because age and exposure to the ambient language are the same in L1 acquisition, we can conclude that more experience with the language is associated with greater use of complex sentences in both groups. The distribution of different complex sentence types in L1 learners' speech is not equal, and this is also the case for cL2ers. Both groups of children show the same distribution pattern: complement >> coordinated >> adverbial >> relative. This pattern mirrors the distribution in caregiver speech to L1 learners (Diessel, 2004; Huttenlocher et al., 2010; Vasilyeva et al., 2008) and possibly in the language more generally, although future corpus-based research would be needed to ascertain this.

Our analyses, however, also revealed some divergence from the patterns noted for monolingual children. First, cL2ers used a higher proportion of complex sentences in their language production: 18% for cL2 learners versus 10% for L1

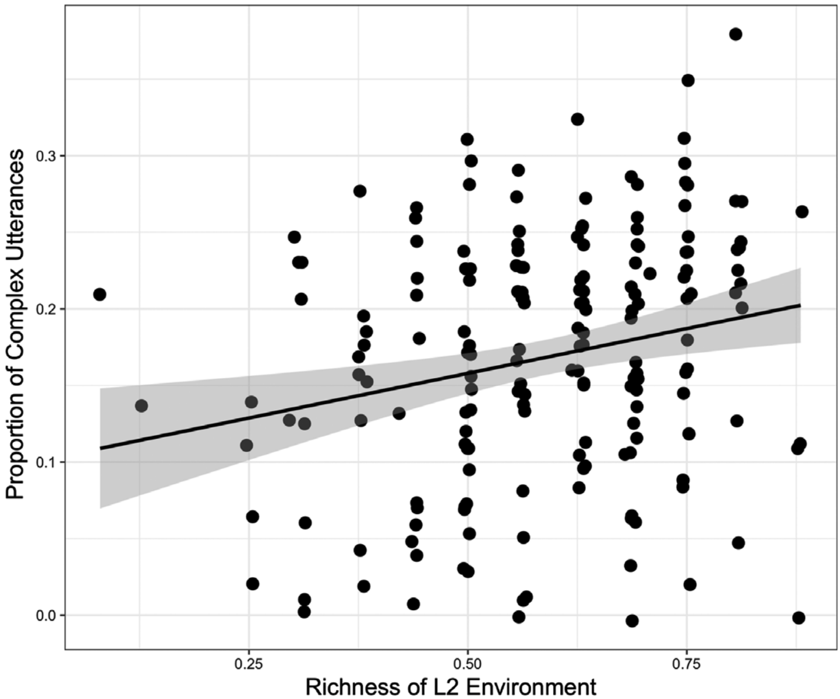

FIGURE 3. Scatterplot of proportion of complex sentence use by each child as a function of richness of the L2 environment.

learners aged 2;0–4;0. Even cL2ers with a few months of exposure to English produced complex sentences. Second, cL2ers did not demonstrate any order of emergence in that all complex sentence types were produced by the low English exposure group (Figure 1), and proportional distributions did not change much with increasing L2 exposure. These differences suggest an advantage for an older AOA in the acquisition of complex syntax (cf. Rothman et al., 2016), where greater cognitive and linguistic maturity permits cL2ers to use complex syntax earlier in their English development. However, it should be noted that neither age at testing nor AOA were significant predictors of complex syntax use *within* the cL2 sample. It is possible that the ranges of AOAs and ages within the sample were too narrow to have had an impact. It could also mean that the age advantage for cL2ers lies in the contrast between toddlers and young school age children. In the case of the latter, we cannot rule out combined effects of school-language experience and older AOA: a topic to be tackled in future research. Finally, the role of SES/maternal education was prominent as an input quality factor in the L1 research but not in this study, and we found instead that richness of the L2 environment—a composite of activities and interactions with native-speaker oral and written input—emerged as the input quality predictor of complex sentence use (Table 4). One reason for this difference could be that maternal education in the

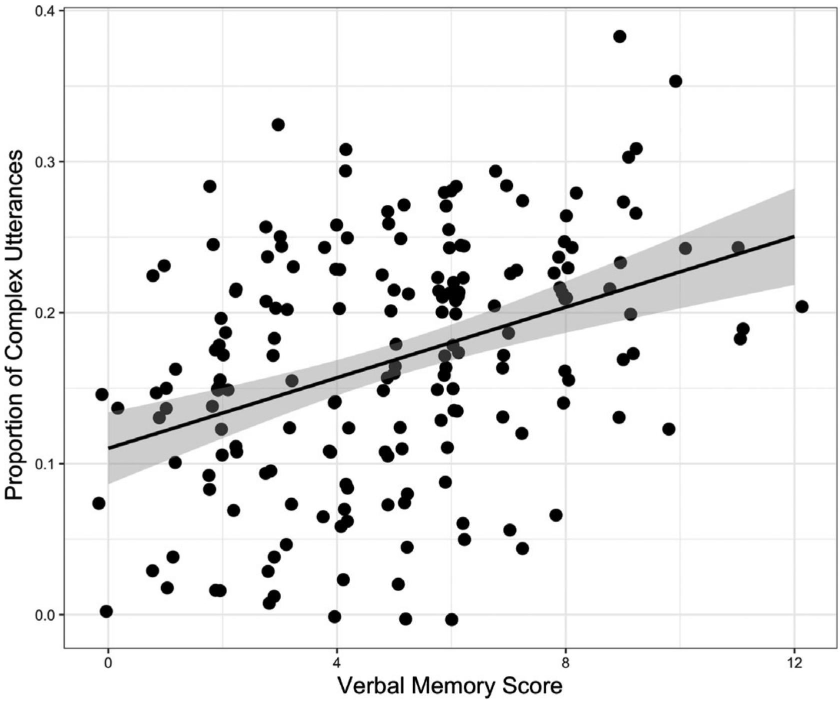

FIGURE 4. Scatterplot of proportion of complex sentence use by each child as a function of verbal memory scores.

sample in this study is skewed high (mean of 14 years in Table 2 implies most had postsecondary education), and thus, there might not have been sufficient variation in maternal education levels to detect a strong impact on children's L2 development (cf. Paradis, 2011). Furthermore, mother's level of fluency in the L2 was not high (2.2 on 0–4 scale; Table 2), and so if mothers were using the L2, they might not have used as many complex sentences compared to high-SES monolingual mothers.

Factors Predicting cL2 Complex Syntax and Constructivist Accounts

The optimal model (Table 4) indicates that children's production of complex sentences was higher on the spontaneous than the narrative task, contra expectations (Elin Thordardottir, 2008). Since narrative tasks have a closer relationship with written language, one might expect children to have produced more complex syntax using this genre. However, we used a story generation rather than a story retell narrative task, and it is possible that if children had been asked to retell a story that was told or read to them, more complex sentences would have emerged. Another reason could be that fewer complex sentences were used on the narrative task

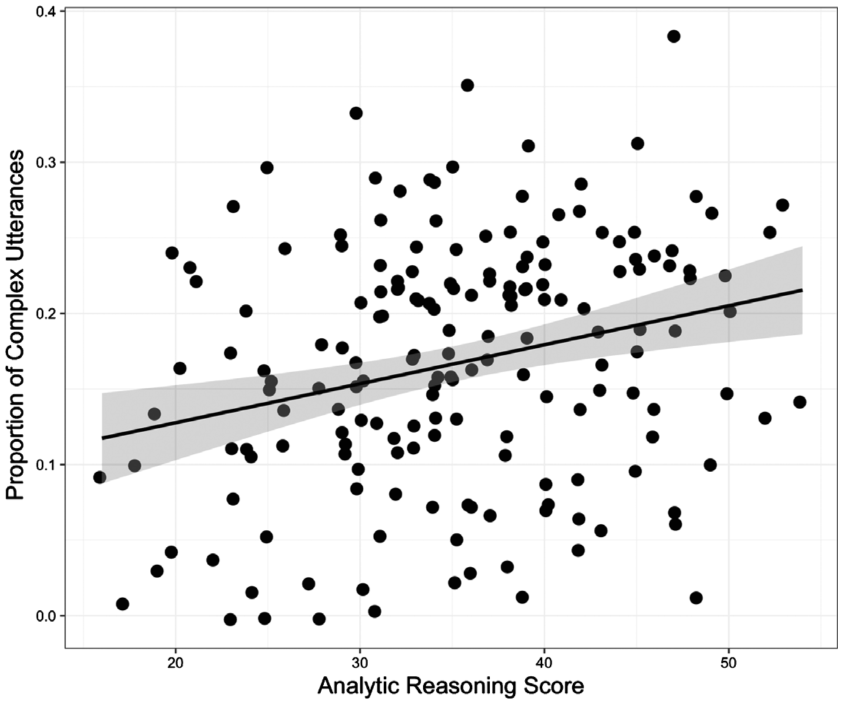

FIGURE 5. Scatterplot of proportion of complex sentence use by each child as a function of analytic reasoning scores.

because it yielded a shorter language sample than the 20-minute play session. For the narrative task, children produced a mean of 75 sentences, and for the spontaneous language sample, they produced a mean of 87 sentences; the latter was significantly longer ($t(186) = 3.13$, $p = 0.002$). The model results also show that children's production of complex sentences increased along with greater exposure to English in school, richer English L2 environments, superior verbal memory, superior analytic reasoning, and larger L2 vocabularies (figures 2–6). Therefore, L2 input quantity and quality factors and cognitive and cross-domain linguistic factors are all associated with the acquisition of complex syntax. Notably, speaking and hearing more English at home did not influence children's use of complex sentences. This finding is consistent with other research with different samples of cL2ers (Paradis & Kirova, 2014) raising the issue of the quality of L2 input at home from nonfluent speakers (Hoff, Welsh, Place, & Ribot, 2014). L1 background did not influence the proportion of complex sentences children produced, and thus, the diverse L1s in our sample were not a source of confounding variation regarding our key outcome variable. Nevertheless, future studies using more fine-grained methods could reveal a role of L1 transfer the cL2 acquisition of complex syntax in production.

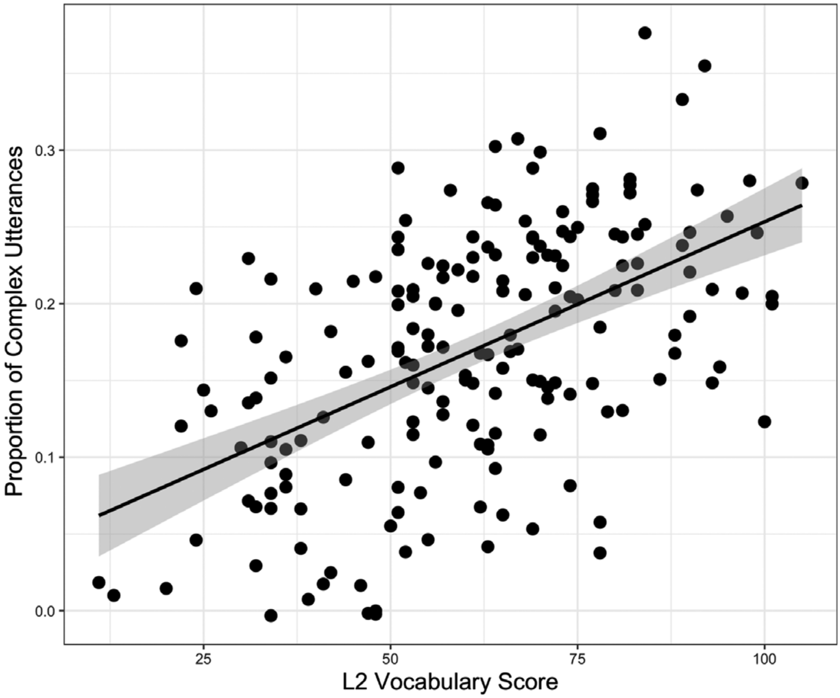

FIGURE 6. Scatterplot of proportion of complex sentence use by each child as a function of L2 receptive vocabulary scores.

Overall, these results support constructivist approaches because they indicate that input and cognitive factors are mechanisms in cL2 syntactic acquisition and that associations between the lexicon and syntax are apparent in this process. Furthermore, prior research has found similar results for cL2 morphology (Blom et al., 2012; Paradis, 2011). It could be argued that the findings for the role of input are not necessarily inconsistent with generativist approaches. Unsworth (2016) found that amount of L2 input did not impact cL2ers' comprehension of a Dutch structure at the syntax-semantics interface, but found amount of L2 input did influence the acquisition of another morphosyntactic structure. Perhaps differentiated mechanisms of learning are apparent in subtle aspects of comprehension, especially in the case of structures for which acquisition based on the input alone is not obvious (see Unsworth 2016). Furthermore, Unsworth adopted a generative model where parameter triggering is sensitive to amount of input and argued that this explains why cL2 learners of Dutch vary in morphosyntax involving setting the V2 parameter. We acknowledge that these arguments have merit. Nevertheless, these arguments pertain mainly to the role of input quantity in cL2 acquisition; the roles of input quality, cognitive factors, and syntax-lexicon connections make the findings of this study incompatible with generative approaches overall.

CONCLUSION

This study has shown that complex syntax is a relative strength in cL2 acquisition of English. Older AOA appears to confer advantages on these young cL2ers because they use all types of complex sentences earlier than L1 children, in terms of length of exposure to the language. They also have low rates of syntactic errors—errors in word order and with connectives—in their complex sentences. At the same age and level of English exposure, cL2ers' accuracy with English verb morphology is low (Blom et al., 2012; Paradis, 2011), indicating that morphology is a relative weakness. The relative strength that English cL2 children display with complex syntax could serve them well in the context of learning the language of instruction at the same time as academic content in the early education years. Input factors (length of English exposure in school, richness of the English environment outside school) and cognitive factors (verbal memory and analytic reasoning), as well as English vocabulary size, predicted individual differences, suggesting that these are mechanisms driving cL2 acquisition of complex syntax forward. Finally, sensitivity to input variation, interactions between cognitive abilities and linguistic development, and associations across linguistic subdomains in development are consistent with a constructivist model of acquisition. One worthwhile direction for future research would be to directly examine the L2 input from native and nonnative speakers to cL2 learners, since the present study measured input factors indirectly via parent report.

Acknowledgments

We would like to thank our research assistants, Stephanie Thomlinson, Khadija Farooq, Nicole Salvador, and Julie Sawchuk, for their excellent work in coding the data. Research for this article was made possible by funding from the Social Sciences and Humanities Research Council of Canada (standard research grant # 410-2010-0123 to Johanne Paradis).

REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*(2), 249–306.

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge, UK: 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. (2015). lme4 package for linear mixed-effects models using "eigen" and S4. *Journal of Statistical Software, 67*(1), 1–48.

Behrens, H. (2008). *Corpora in language acquisition research: History, methods, perspectives*. Amsterdam, The Netherlands: John Benjamins.

Blom, W. B. T., & Bosma, E. (2016). The sooner the better? An investigation into the role of age of onset and its relation with transfer and exposure in bilingual Frisian-Dutch children. *Journal of Child Language, 43*, 581–607.

Blom, E., & Paradis, J. (2013). Past tense production by English second language learners with and without impairment. *Journal of Speech, Language, and Hearing Research, 56*, 1–14.

Blom, E., Paradis, J., & Sorenson Duncan, T. (2012). Effects of input properties, vocabulary size and L1 on the development of third person singular –s in child L2 English. *Language Learning, 62*(3), 965–994.

Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A., & Gillam, R. B. (2010). What you hear and what you say: Language performance in early sequential Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism, 13*, 325–344.

Burgemeister, B., Hollander Blum, L., & Lorge, I. (1972). *Columbia Mental Maturity Scale*. New York, NY: Psychological Corporation.

Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, morphology and complex syntax in successive bilingual children. *Linguistic Approaches to Bilingualism, 1*, 223–248.

Dabrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*, 604–615.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge, U.K.: Cambridge University Press.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Elin Thordardottir. (2008). Language-specific effects of task demands on the manifestation of specific language impairment: A comparison of English and Icelandic. *Journal of Speech, Language, and Hearing Research*, *51*, 922–937.

Eisenbeiss, S. (2009). Generative approaches to language learning. *Linguistics, 47*(2), 273–310.

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

Hoff, E., Welsh, S., Place, S., & Ribot, K. (2014). Properties of dual language input that shape bilingual development and properties of environments that shape dual language input. In T. Grüter & J. Paradis, *Input and experience in bilingual development* (pp. 119–140). Amsterdam, The Netherlands: John Benjamins.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge, UK: Cambridge University Press.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*(4), 343–365.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research, 50*(5), 1280–1299.

Kidd, E., Chen, A., Chiu, J. (2015). Cross-linguistic influence in simultaneous Cantonese–English bilingual children's comprehension of relative clauses. *Bilingualism: Language and Cognition*, *18*, 438–452.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Murphy, V., & Evangelou, M. (2016). *Early childhood education in English for speakers of other languages*. London, UK: British Council. Retrieved from www.britishcouncil.org.

O'Donnell, R. (1974). Syntactic differences between speech and writing. *American Speech, 49*, 102–110.

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., & Jia, R. (2016). Bilingual children's long-term outcomes in English as a second language and sources of individual differences in their rate of English development. *Developmental Science, 20*(1), e12433. doi: 10.1111/desc.12433

Paradis, J., & Kirova, A. (2014). English second-language learners in preschool: Profile effects in their English abilities and the role of home language environment. *International Journal of Behavioral Development, 38*(4), 342–349.

Paradis, J., Rice, M., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child L2 from L1 and SLI. *Applied Psycholinguistics, 29*, 1–34.

Paradis, J., Tulpar, Y., & Arppe, A. (2016). Chinese L1 children's English L2 verb morphology over time: Individual variation in long-term outcomes. *Journal of Child Language, 43*, 553–580.

Purcell-Gates, V. (2001). Emergent literacy is emerging knowledge of written, not oral, language. *New Directions for Child and Adolescent Development, 92*, 7–22.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Essex, UK: Longman.

R Core Team. (2015). *The R project for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.r-project.org/

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.

Rothman, J., Long, D., Iverson, M., Judy, T., Lingwall, A., & Chakravarty, T. (2016). Older age of onset in child L2 acquisition can be facilitative: Evidence from the acquisition of English passives by Spanish natives. *Journal of Child Language*, *43*, 662–686.

Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Retrieved from University of Alberta Faculty of Rehabilitation Medicine website: www.rehabresearch.ualberta.ca/enni

Statistics Canada. (2011). Linguistic characteristics of Canadians (Catalog No. 98-314-X2011001). Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf

Thomas, E. M., Williams, N., Jones, L. A., Davies, S., & Binks, H. (2014). Acquiring complex structures under minority language conditions: Bilingual acquisition of plural morphology in Welsh. *Bilingualism: Language and Cognition, 17*(3), 478–494.

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

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., & Tsimpli, I. (2014). On the role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics, 35*, 765–805.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11*(1), 84–97.

Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.